STUDIES ON NUCLEOSIDE ANALOGS

I. TOTAL SYNTHESIS AND PROPERTIES OF N(1)-(β-D-ERYTHROFURANOSYL)-URACIL AND ITS 2',3'-CYCLIC PHOSPHATE, A SUBSTRATE ANALOG FOR RIBONUCLEASE

II. IMPROVEMENTS IN THE DETERMINATION OF CONFORMATIONS OF NUCLEOSIDES BY NMR SPECTROSCOPY

A Dissertation

Presented to

the Faculty of the Department of Chemistry

College of Arts and Sciences

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Robert Freeman Coombes

May 1972

ACKNOWLEDGEMENTS

I gratefully thank Dr. James R. Cox, Jr., for the original suggestion and continued interest in the research problem presented in this thesis.

I sincerely thank my wife, Wanda, who typed the rough draft of this dissertation and assisted in the chromatography and leukemia testing.

I thank the members of my committee, Dr. Cox, Dr. S. Welch, Dr. M. R. Willcott, III, Dr. A. P. Kimball, and Dr. J. L. Bear for advice in the research and for their criticisms of the rough draft.

I thank Dr. R. Matcha and Dr. T. L. Lemke and all the other professors and students who advised or assisted me in various parts of the research program.

STUDIES ON NUCLEOSIDE ANALOGS

I. TOTAL SYNTHESIS AND PROPERTIES OF N(1)-(β-D-ERYTHROFURANOSYL)-URACIL AND ITS 2',3'-CYCLIC PHOSPATE, A SUBSTRATE ANALOG FOR RIBONUCLEASE

II. IMPROVEMENTS IN THE DETERMINATION OF CONFORMATIONS OF NUCLEOSIDES BY NMR SPECTROSCOPY

An Abstract of

A Dissertation

Presented to

the Faculty of the Department of Chemistry

College of Arts and Sciences

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Robert Freeman Coombes

May 1972

ABSTRACT

The first synthesis of an N(1)-(β-D-erythrofuranosyl)-pyrimidine nucleoside was carried out by two methods which unambiguously established the structure of the desired nucleoside analog and a number of side products. D-Erythrose triacetate (V) was synthesized according to the literature method from glucose in four steps. The nmr spectrum of V was completely interpretable by first order rules, and the interpretation was confirmed with the use of a paramagnetic shift reagent, $Eu(DPM)_3$. Compound V is predominantly the β-anomer. Reaction with hydrogen chloride in dichloromethane afforded a good yield of the unstable acylglycosyl chloride, VI, as an oil containing only the β-anomer. The mercuri and mercuric cyanide reactions failed to produce an isolable yield of a pyrimidine nucleoside from VI because of instability of the halogenose. The Hilbert-Johnson reaction of VI with 2,4-bis(trimethylsilyloxy)pyrimidine afforded predominantly the α-anomer(VII) of the N(1)-glycoside, and the desired β-nucleoside analog was obtained in only fair yield. Reaction of the triacetate V with 2,4-bis(trimethylsilyloxy)pyrimidine in the presence of stannic chloride afforded a satisfactory yield of the desired N(1)-(2',3'-di-O-acetyl-β-D-erythrofuranosyl)-uracil(VIII), along with two O-nucleosides, and a bis(erythrofuranosyl)-uracil. Compound VIII was hydrolyzed

to N(1)-(β-D-erythrofuranosyl)-uracil(X), which was found to be inactive against leukemia L1210 in mice. Reaction of X with trimethylphosphite in the presence of trifluoroacetic acid afforded the 2'(3')-monophosphite derivatives, which were oxidized with hexachloroacetone to give N(1)-(β-D-erythrofuranosyl)-uracil-2',3'-cyclic monophosphate, XV. Compound XV will be studied as a substrate analog for ribonuclease.

A computer program was developed for accurately calculating the dihedral angles for twist, intermediate, and envelope conformations of any degree of puckering for both cyclopentane and tetrahydrofuran rings. The most popular literature tables for determining conformations of nucleosides by nmr spectroscopy were shown to contain numerous large inaccuracies of up to 15°. Dihedral angles of hydrogen substituents of fourteen nucleosides were calculated from literature X-ray data, and it was determined that distortions of up to 26° (compared to ring dihedral angles) commonly occur. In light of all of the data described above and of our own and literature correlations of the Karplus equation to molecules of fixed conformation, a more reliable method of determining conformations of nucleosides in solution was developed. The conformations of VIII, X, and other molecules containing five-member rings are discussed.

TABLE OF CONTENTS

TABLE OF CONTENTS CONTINUED

TABLE OF CONTENTS CONTINUED

LIST OF TABLES

LIST OF TABLES CONTINUED

LIST OF FIGURES

LIST OF FIGURES CONTINUED

CHAPTER I

INTRODUCTION

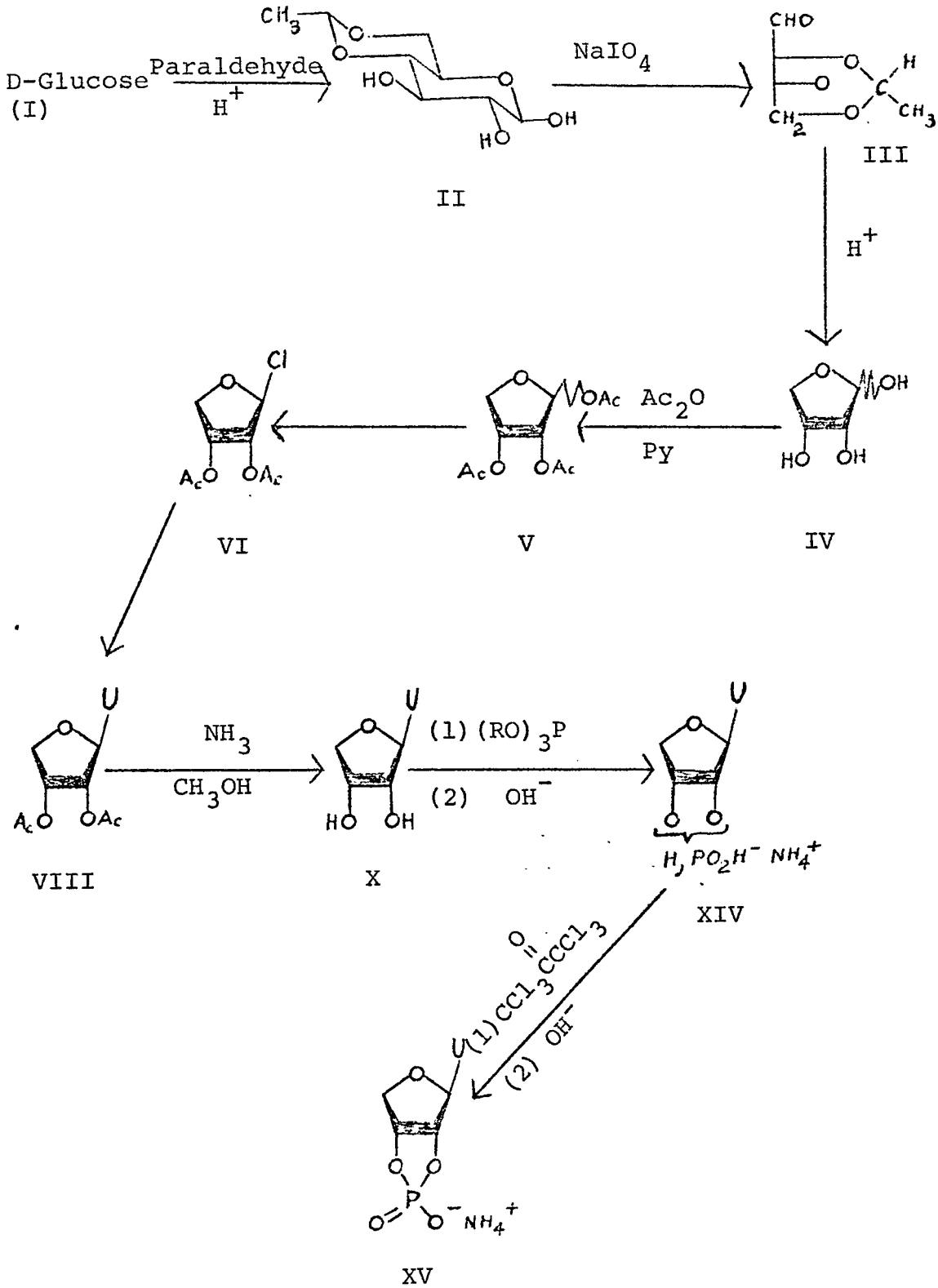
DESCRIPTION OF THE PROJECT

The objective of this research project was to synthesize 1-(β-D-erythrofuranosyl)-uracil-2',3'-cyclic monophosphate (XV) in order to study its ribonuclease-catalyzed hydrolysis reaction by stopped flow temperature jump techniques. It was also of interest to test the nucleoside analog, N(1)-(β-D-erythrosyl)-uracil, for anticancer activity and to study its conformation. The original plan was to follow the synthetic outline below (eq. 1). Unexpected difficulty was encountered in the preparation of the nucleoside from the polyacylglycoside, because of unusual instability of the polyacylglycosyl halide. A number of methods of nucleoside synthesis were applied until a method was found which afforded a good yield of the desired product. These studies enabled us to gather enough product data to furnish proof in favor of one recently proposed mechanism of the Hilbert-Johnson reaction over another.

Pyrimidine Nucleoside Synthesis

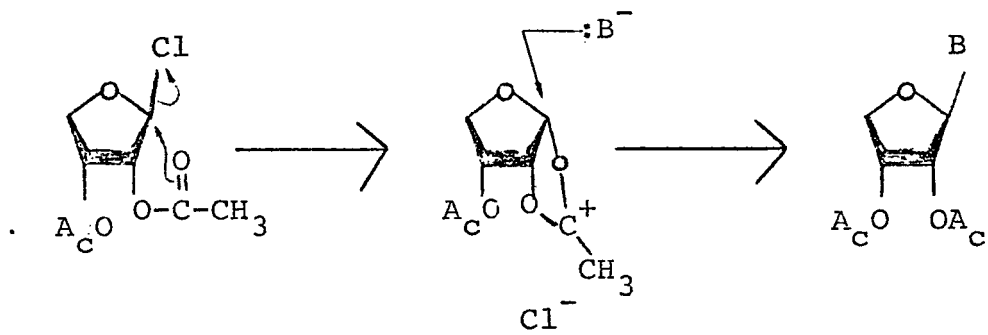
Nucleosides are usually synthesized from a derivative of the corresponding sugar, usually a polyacylglycosyl halide.[1,2,3,4] In general, an O-protected glycosyl halide is assumed to react by an S_N1 type of mechanism in which the halide dissociates before reaction with the heterocycle. The

(Eq.1)


D-Glucose
(I)
Paraldehyde
H+
NaIO4
II
III
H+
IV
Ac2O
Py
V
VI
NH3
CH3OH
VIII
(1) (RO)3P
(2) OH−
X
XIV
(1) CCl3CCCl3
(2) OH−
XV

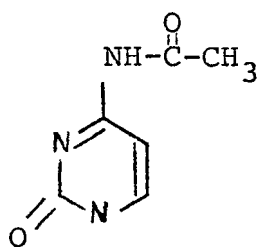
advantage of acetyl or benzoyl protecting groups at the O(2) position is that in many reactions neighboring group participation by this group occurs, forming a dioxolenium cation before reaction with the heterocycle. If all of the glycoside forms a dioxolenium cation which then reacts with the heterocycle, the reaction product will be stereospecifically trans-1,2 in spite of the S_N1 character of the rate-controlling step because displacement of the O(1) of the dioxolenium cation is an S_N2 process (Eq. 2).

(Eq.2)

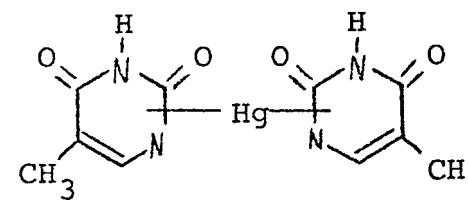


Steric hindrance by the O(2) protecting group of the halogenose is an important factor in determining preferential stereochemistry of these reactions when a dioxolenium cation does not form, and this factor also favors the formation of a trans-1,2 product. Another stereochemically important factor is the configuration of the halide. The departing halide tends to hinder attack by the heterocycle, and this factor favors partial inversion of configuration at C(1).

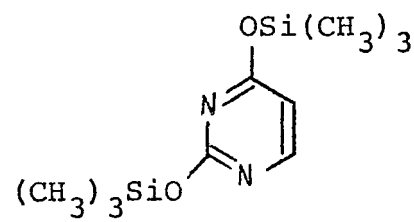
Three types of pyrimidine compounds are commonly used to produce nucleosides: unprotected or exocyclic N-protected pyrimidine bases of nucleosides or nucleoside analogs (e.g., N-acetyl cytosine), their mercury salts (e.g., dithiminyl mercury), or their O-alkyl or trimethylsilyl derivatives (e.g., 2,4-bis(trimethylsilyloxy)pyrimidine). These three types of bases are associated with three different reactions: the mercuric cyanide reaction, the mercuri reaction, and the Hilbert-Johnson reaction, respectively.



N(4)-acetyl cytosine



Dithiminyl mercury (exact structure uncertain)



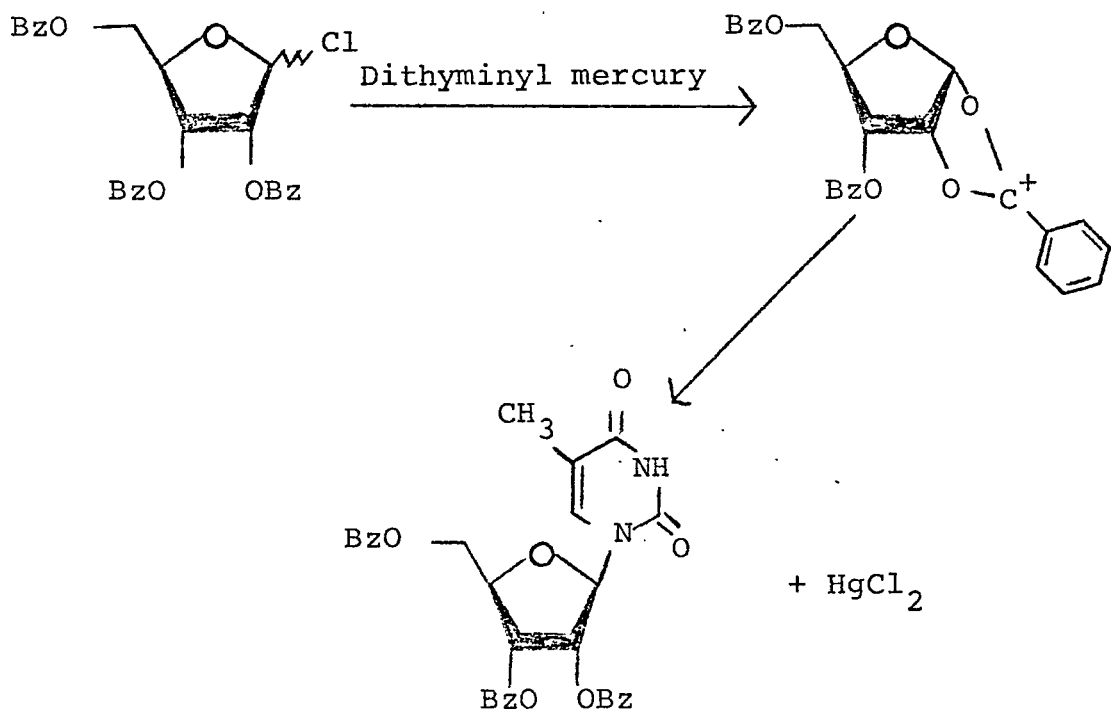
2,4-bis(trimethylsilyloxy)-pyrimidine

The Mercuri Reaction

In the mercuri reaction[1,2] the mercury salts of nucleoside bases react with O-protected glycosyl halides in

refluxing toluene or xylene to give nucleoside products (Eq. 3). When 2-acyloxyglycosyl halides of either anomeric configuration are treated with a mercury salt of a pyrimidine nucleoside base, a trans-1,2-N(1)-glycoside is usually the only isolable product. The stereospecificity of the reaction is evidence that the reacting species is the dioxolenium cation, and one function of the mercuric group is presumably to catalyze removal of the halide to facilitate complete ionization so that the dioxolenium ion forms.

(Eq.3)



T.L.V. Ulbricht's group has isolated side products in the mercuri reaction in which pyrimidine base hetero atoms

other than N(1) were glycosylated, especially oxygen atoms.[6] When silver salts of pyrimidine bases were used, O-nucleosides were commonly isolable. Their studies indicate that N(1)-pyrimidine nucleosides are formed by two mechanisms in the mercuri reaction. One mechanism involves direct N-glycosidation, while the other mechanism proceeds through an initially formed O-nucleoside followed by O→N rearrangements catalyzed by mercuric halide. Unstable O-glycosides, such as deoxyribosides, cleave and presumably decompose rather than undergoing the O→N rearrangement. Since the site or sites of coordination of the mercuric ion to the heterocyclic moieties of mercury salts of pyrimidines are unknown, it is difficult to explain why O-nucleosides are initially formed in some mercuri reactions but not in others. One hypothesis made by Ulbright, based on the observation that dithiminyl mercury gives an O-deoxyriboside while monothiminyl mercury gives an N(1)-deoxyriboside, is that the hetero atom coordinates to the mercuric ion in the position of glycosidation.[6] This mechanism has been proven in the chloromercuripurine series, where the points of coordination have been determined.[7] However, O-glucosides were later isolated in a reaction of monothiminyl mercury.[8] Perhaps monothiminyl mercury contains both N- and O- coordinated mercuric ions.

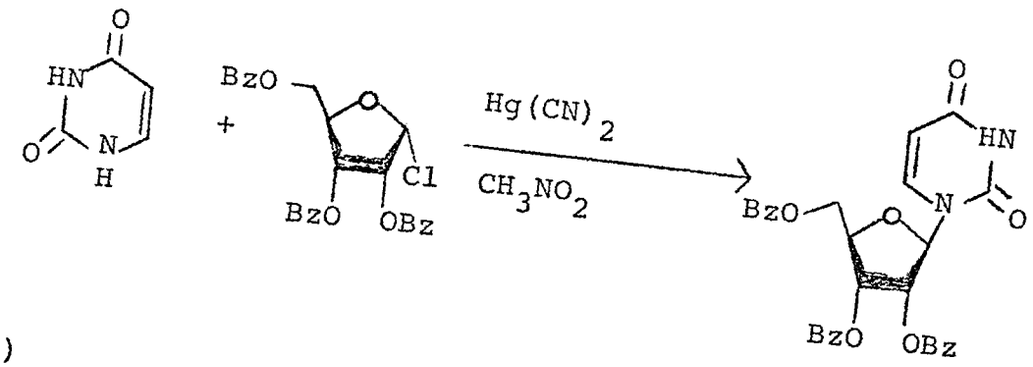
In a modification of the mercuri reaction, N(6)-benzoyladenosine mercuri chloride, $TiCl_4$, and polyacylglycoside are

refluxed together in a solvent such as 1,2-dichloroethane.[9,10] The polyacylglycosyl halide is generated in situ and probably reacts with the mercury salt to produce initially the trans-1,2-nucleoside, which subsequently is equilibrated by $TiCl_4$ to form a mixture of α- and β-anomers.

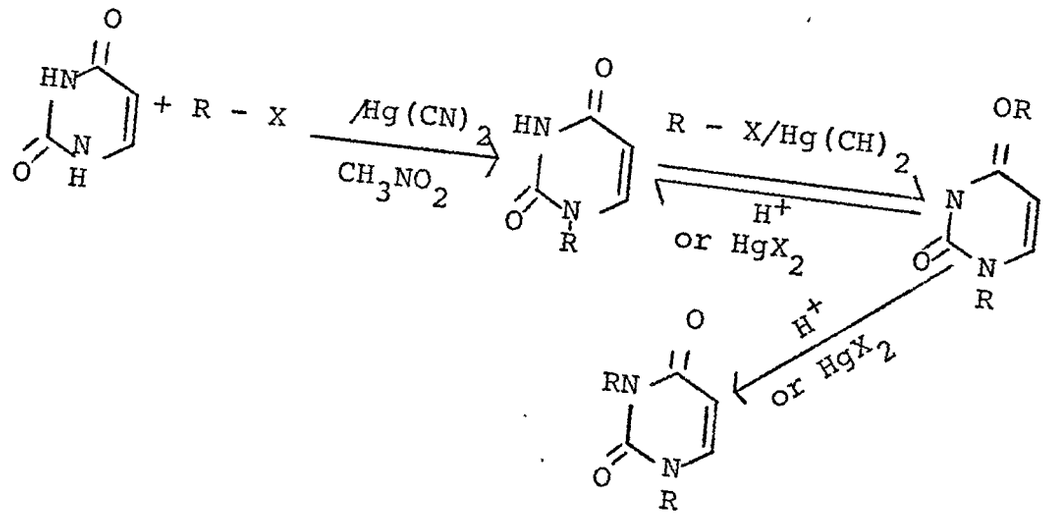
The Mercuric Cyanide Reaction

The pyrimidine bases also may be glycosylated by poly-O-acylglycosyl halides using mercuric cyanide catalyst and acetonitrile or nitromethane solvent (Eq. 4).[11,12] Only trans-1,2-nucleosides are produced, and both N(1)O(4)- and N(1)N(3)-bis-glycosides have been encountered when uracil was used. At present, the simplest mechanism which accounts for all observed experimental results is initial N(1)-glycosidation, followed sometimes by O(4)-glycosidation to produce a bis-glycoside. The O-glycoside may either undergo O→N rearrangement or cleavage, and both rearrangement and cleavage may be catalyzed by either mercuric halide or hydrogen halide (Eq. 5). The low boiling polar solvents, together with molecular sieves to remove hydrogen halide, favor mercuric halide catalyzed cleavage of O-nucleosides, thereby producing only N(1)-glycoside in best yields.[12]

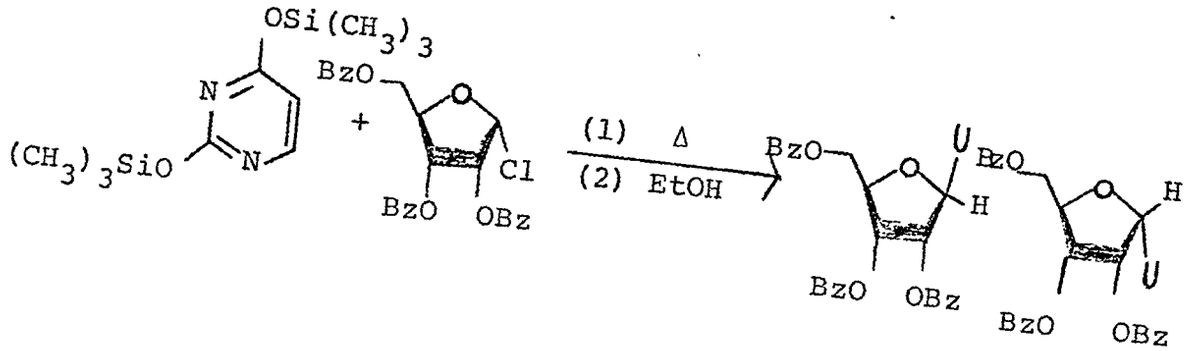
(Eq.4)



(Eq.5)



(Eq.6)



The Hilbert-Johnson Synthesis

In the Hilbert-Johnson synthesis, which has recently been reviewed,[3] halogenoses react with alkoxy or trimethylsilyloxy-pyrimidines with or without solvent (Eq. 6). Both α- and β-anomers are often formed from 2-acyloxyhalogenoses, implicating a different intermediate or mechanism from the dioxolenium cation of the heavy metal salt reactions. Reactions of alkoxypyrimidines in more polar solvents such as acetonitrile form more stereoselectively trans-1,2 products than reactions in nonpolar solvents or without solvents, suggesting that at least two mechanistic pathways are operating, and that in more polar solvents the dioxolenium ion mechanism is more important due to stabilization by greater solvation energy. Alkoxypyrimidines also sometimes give O-nucleoside products, but no O-nucleosides have been reported from trimethylsilyloxypyrimidines. Trimethylsilyloxypyrimidines are less reactive, and are used in the Hilbert-Johnson reaction without sovent at temperatures of 150°-200°.[13]

The studies of W. G. Overend[14] on solvolytic reactions of polyacylglycosyl halides suggest that the Hilbert-Johnson reaction with the weakly nucleophilic heterocycles used proceeds by an S_N1 mechanism in nonpolar or weakly polar media. T. L.V. Ulbright[15] has postulated, nevertheless, that an S_N2 displacement competes with a dioxolenium cation intermediate in the Hilbert-Johnson reaction of trans-1,2-acylglycosyl

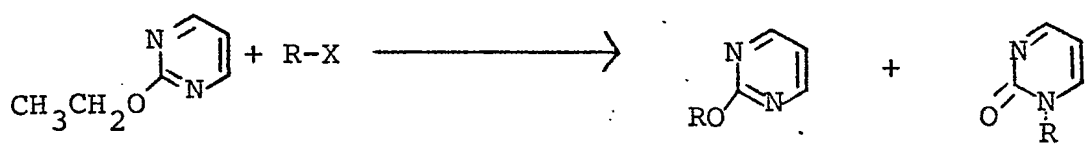
halides. Sorm[16,17] has postulated that the alternate mechanism in the Hilbert-Johnson reaction proceeds through a closely associated ion pair intermediate. The ion pair could react to give glycosides of both anomeric configurations. In more polar solvents, complete dissociation resulting in a dioxolenium ion intermediate is favored by greater solvation of the ions. Mercuric bromide increases the rate of the Hilbert-Johnson reaction but lowers the yield. Sorm attributed both the increased rate and lowered yield to faster dissociation of the halide catalyzed by mercuric bromide to form the less stable dioxolenium ion of the glycoside. He did not attempt to reconcile this explanation with the high yield of the stereospecifically trans-1,2 mercury salt reactions.

If O-nucleosides are formed, a switch from alkyloxy pyrimidine to trimethylsilyloxy pyrimidine may eliminate this undesirable side product; for example, 2-ethoxypyrimidine gives O-nucleoside side products in a Hilbert-Johnson reaction with 2,3,4,6-tetraacetyl-β-D-glucopyranosyl bromide with no solvent (Eq. 7),[18] but 2-trimethylsilyloxypyrimidine does not (Eq. 8).[19,20] When cis-1,2-acylglycosyl halides are used in the Hilbert-Johnson reaction, the trans-1,2-nucleoside can be expected as the predominant product and is often the only isolable product. When trans-1,2-acylglycosyl halides are used, either cis- or trans-1,2 nucleosides may be the predominant products. In the reaction of 2,3,4,6-tetraacetyl-

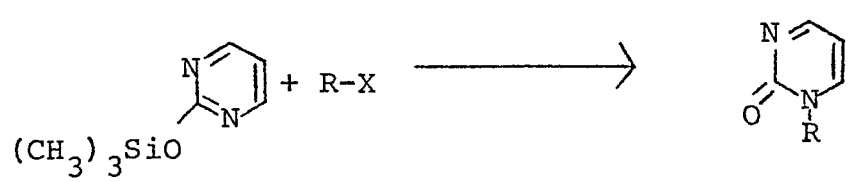
α-D-glucopyranosyl bromide with 2,4-diethoxy-5-methylpyrimidine with no solvent, only the β-nucleoside was isolated;[21] but when triacetylribofuranosyl bromide (presumably β) was used, only the α-nucleoside was isolated.[22] In nonpolar solvent, both α- and β-nucleosides were obtained from the triacetylribofuranosyl bromide.[22] When non-participating protecting groups are used, as in tribenzyl-α-D-arbinofuranosyl chloride,[23] the predominant product usually has the anomeric configuration opposite that of the starting halide.

A recently discovered reaction utilized trimethylsilyloxy pyrimidine in presence of stannic chloride catalyst to produce nucleosides from polyacyl sugars (Eq. 9)[24,25] Only trans-1,2 glycosides have been produced from 2-acyloxyacyl-glycosides in this reaction.

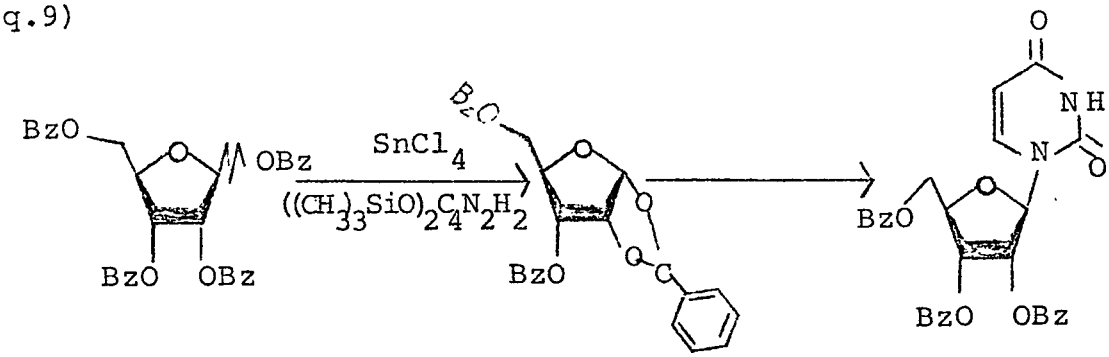
(Eq.7)



(Eq.8)



(Eq.9)



Conformations of Nucleosides

The conformational aspects of primary interest in these compounds are the conformation of the 5-member ring and the preferred rotational angle or range of angles about the bond between the sugar (ribose) and the base. Figure 1 is a Haworth projection which shows uridine in the anti-conformation. Figure 2 is uridine pictured in one possible envelope conformation of the ribose portion, C(3')-endo.

The knowledge of the shape and physical properties of the various nucleotides in solution is essential for the understanding of the structures of DNA and RNA. In addition, nucleotide and nucleoside interactions with enzymes are of current research interest, as are the antibacterial and anti-cancer properties of many of these compounds and analogs of them. Molecular conformation may be related to biological activity.

It has been widely considered that 5-membered rings which include only carbon and oxygen atoms may occupy either

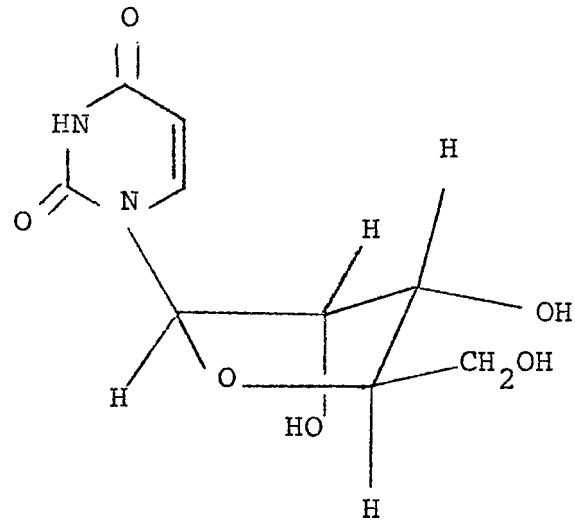


Figure 1. Uridine in Anti Conformation

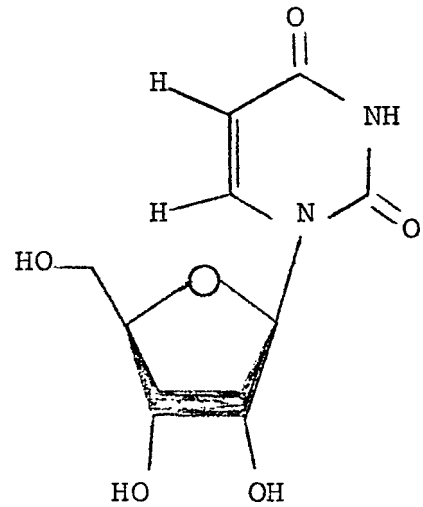


Figure 2. Uridine in C(3')-endo Conformation

"envelope" (also called C_s) conformations (Figure 3) or "twist" conformations (also called C_2) (Figure 4).[26] In Figure 3 one atom is out of plane, while in Figure 4 two atoms are equidistant out of plane.[26] Any such puckered conformation has increased "Baeyer" (angle) strain over that of the planar conformation, but has decreased "Pitzer" (eclipsing) strain. In planar cyclopentane all the hydrogens are eclipsed. There are 10 C_s conformations resulting from there being 5 ring atoms which may be out of plane, and two sides that may be out of plane. There are also 10 C_2 conformations, making a "total" of 20. The energy advantage of puckered forms of cyclopentane over planar is small, and there is very little energy difference if any between envelope or half-chair or any of the intermediate conformations. The molecule is thus in a rapid state of conformational flux called "pseudorotation." The bulky substituents of nucleosides and nucleotides therefore are the determining factor of exactly what conformation or conformations exist in these molecules.[27] X-ray diffraction shows that the conformations of crystalline nucleosides and nucleotides are usually somewhere between C_s and C_2, but more like C_s.[28] It should be noted that an additional parameter is needed to define these conformations, that being the distance in angstroms that the puckered atoms are out of plane. Adenosine Monophosphate is

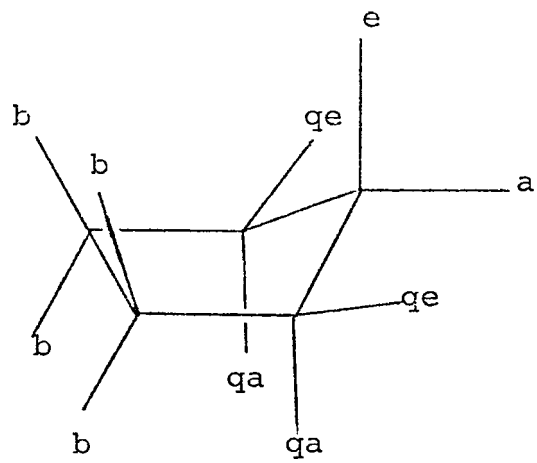


Figure 3. Envelope Conformation

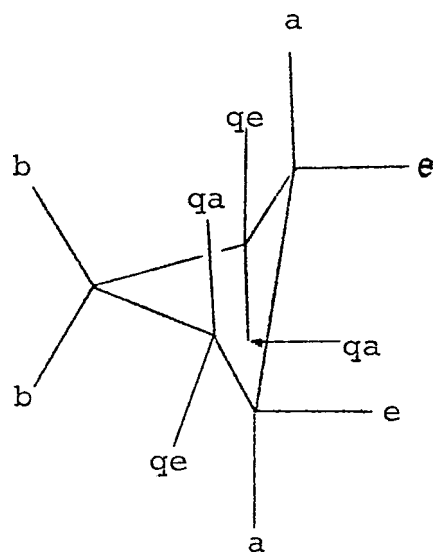


Figure 4. Twist Conformation

thus called C(3')-endo in solid phase,[29] but might also be called C(3')-endo .537Å, C(2')-exo .120Å.

While conformations of these compounds may be determined quite accurately in crystal lattice using X-ray diffraction, forces governing conformation in solution may be different. X-ray data are valuable as a guide, however. This method shows that the ribose β-nucleosides and nucleotides are either C(2')-endo or C(3')-endo. Deoxyribose counterparts are either C(2')-endo, C(3')-endo, or C(3')-exo in the crystal.[28]

The method used to estimate the probable conformation or the nature of the conformational equilibrium in the furanose ring in solution is nmr spectroscopy. M. Karplus provided a theoretical equation for determination of dihedral angles of protons on adjacent carbons.[30] In Figure 5 is shown a dihedral angle using a Newman projection. The Karplus equations used by Jardetsky for nucleosides and nucleotides[31,32,33,34] are

$$0° \leq \Phi \leq 90° \qquad J = 8.5 \cos^2 \Phi - 0.28$$

$$90° \leq \Phi \leq 180° \qquad J = 9.5 \cos^2 \Phi - 0.28$$

Figure 5 shows the general form of the Karplus curve. J represents the coupling constant between the protons. Although Karplus warned against quantitative application of his equation, carbohydrate chemists continue to believe in

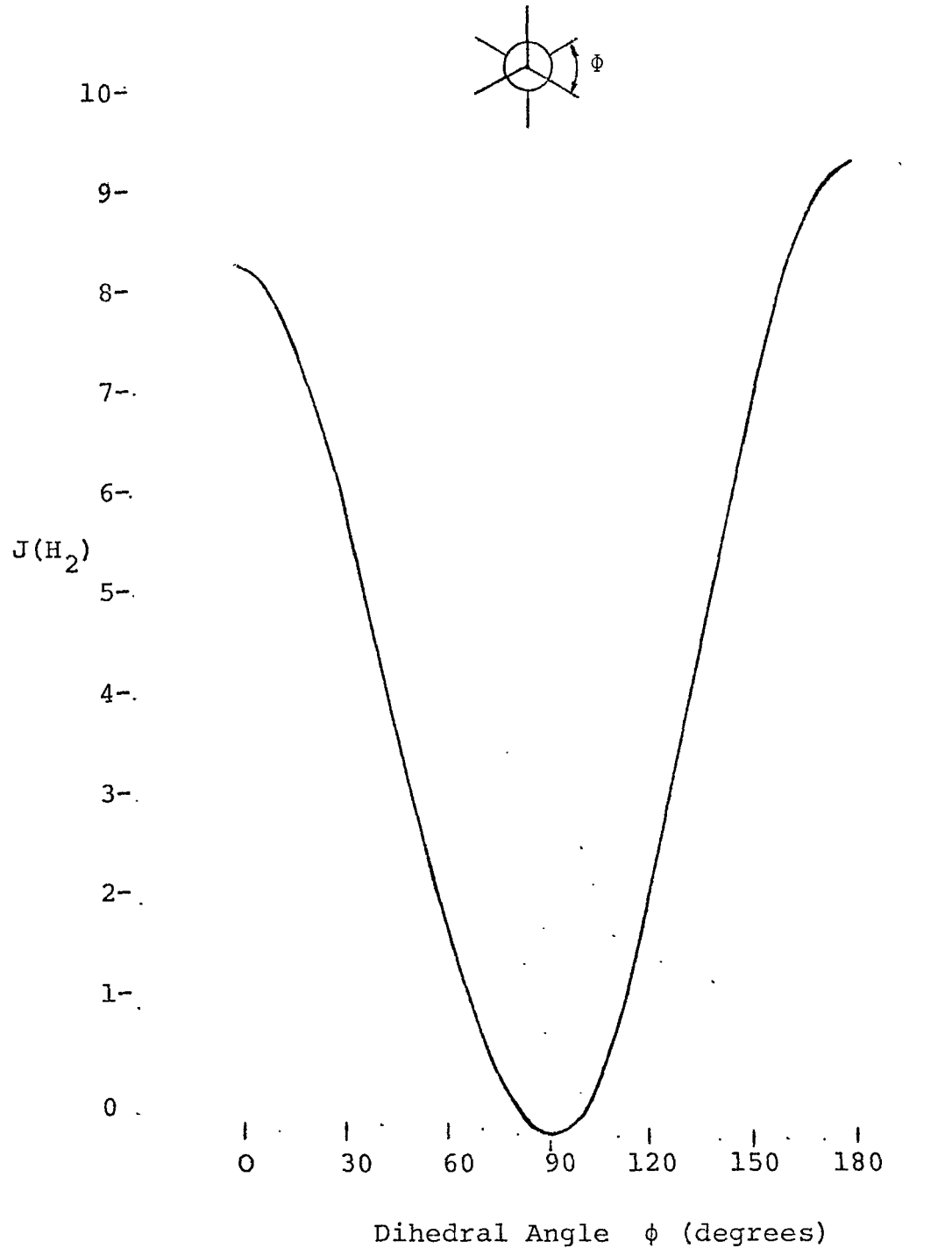


Figure 5. A Plot of the Karplus Equation

its reliability for approximation of dihedral angles in carbohydrates. Some evidence they have used to justify this will be presented in the discussion. The equations definitely provide justification for choosing between C(2')-endo or C(3')-endo if one rules out all other possibilities on the bases of X-ray data and molecular models. This may be an oversimplification, however.

C. Jardetzky has published a table of coupling constants expected for all 20 maximally puckered conformations as well as tables for the most common crystal conformations in less puckered pure C_s and C_2 conformations. These are still being used by current workers in the field.[35,36]

C. Jardetzky, who did the pioneering work in this area, has defined "maximally puckered" furanose rings in C_s conformation as having one atom 0.75Å out of plane, and in C_2 as having two atoms 0.4Å out of plane. Endo was defined as "on the same side as the CH_2OH group" or "up" when the molecule is drawn in the Haworth projection. The letters representing the type of substituent and the dihedral angles between types of substituents are listed below. These refer to the letters on Figure 3, and Figure 4, and the tables hold for both models as long as one assumes "maximal puckering," according to Jardetzky.[32]

When we attempted to apply Jardetzky's tables (e.g., Table I), we noticed several inaccuracies. For example, b-qe

TABLE I

DIHEDRAL ANGLES AND THEORETICAL COUPLING CONSTANTS OF MAXIMALLY PUCKERED CYCLOPENTANE RINGS ACCORDING TO C. JARDETZKY

a = axial
e = equatorial
qe = quasi-equatorial
qa = quasi-axial
b = bisectional

a-a	180°	9.2 Hz
b-b (trans)	120°	2.1
b-b (cis)	0°	8.2
e-e	60°	1.85
e-a (trans)	12-°	2.1
e-a (cis)	60°	1.85
a-qa	165°	8.6
a-qe	45°	3.9
b-qa (trans)	135°	4.6
b-qa (cis)	15°	7.6
b-qe (trans)	105°	0.4
b-qe (cis)	15°	7.6
e-qa	45°	3.9
e-qe	75°	0.4

cis is not the same for a twist conformation of 0.4Å puckering as for an enveloped puckered 0.75Å; it is actually the same for twist and envelope only when twist is puckered above and below the plane of the three reference atoms by the same distance that the envelope is puckered from the plane of the four reference atoms. While b-qa (cis) is 15° for the envloped puckered 0.4Å, it is 31° for the "maximally puckered" (0.75Å) envelope and 15° for the 0.4Å puckered envelope.

In spite of the above oversight, and a few other less significant inaccuracies, the pioneering work of Jardetsky has provided a convenient nomenclature of the substituents on five membered rings. We have developed a means of calculating the dihedral angles of any degree of puckering of envelope conformation, twist conformation, and all conformations between these extremes. This program and accurate tables are presented in the discussion section.

The angle Φ_{CN} of rotation about the bond between C(1') and the base is defined as 0° for pyrimidines when C(6) is over O(1'), and for purines when C(8) is over O(1'), when observed from N to C(1') along the glycosidic bond. When C(6) or C(8) of the pyrimidine or purine, respectively, is rotated to the left, the Φ_{CN} angle is negative. When the angle is negative, the conformation is termed "anti"; when positive, "syn". Chemical shifts of H(2'), H(3'), and H(4') have been correlated to anti conformation for pyrimidine

nucleosides not substituted at C(6), and the chemical shift of H(6) has been shown to correlate with anti conformation for both purine and pyrimidine 5'-nucleotides.[36,37]

Circular dichroism also has been used to provide evidence that anti conformations prevail for both the purine and the pyrimidine compounds in solution, and X-ray diffraction usually shows a negative Φ_{CN} in the crystals.[28,29]

Bovine Pancreatic Ribonuclease

Ribonucleases, enzymes which degrade ribonucleic acid, are widely distributed in nature. Thus far a ribonuclease enzyme has been found in every tissue or cell studied carefully.[40] The richest source of a ribonuclease is the digestive juice of the pancreas, especially in ruminants.[40,41]

Bovine pancreatic ribonuclease was first crystallized by M. Kunitz in 1940.[41] Martin and Porter fractionated the crystalline enzyme in 1951 and discovered two enzymically active components present in a ratio of about nine to one.[42] These have come to be known as bovine pancreatic ribonuclease A (the major component) and B, commonly abbreviated RNase A and RNase B, respectively. Most research involving a homogeneous ribonuclease has been done with RNase A, which is now commercially available in pure enough form for most kinetic studies.[43]

Ribonuclease catalyzes the depolymerization of RNA to nucleotides and oligonucleotides. The reaction proceeds in two distinct stages (Figure 6): (1) a fast transesterification reaction causing the depolymerization (or partial depolymerization) and forming 2',3'-cyclic phosphate derivatives of pyrimidine nucleosides (mainly) by cleaving the C(5')-O-P bonds preferentially after pyrimidine bases; followed by (2) relatively slow hydrolysis of the cyclic phosphate intermediates to form the 3'-phosphates of the nucleotide and oligonucleotide degradation products.[44] The second step occurs mostly at the level of mononucleotides and small oligonucleotides, and it is much slower than the transesterification depolymerization step.[45]

Pancreatic ribonuclease is commonly stated to be specific for O-P bonds[46] next to cytosine and uracil nucleosides but it is also reactive with the 2',3'-cyclic phosphates of the t-RNA nucleosides pseudouridine, dihydrouridine, N(6)-acetyl cytidine, iodomethyluridine, and ribothymidine, along with many unnatural nucleoside phosphates.[47,48,49,50,51] It was earlier thought that the enzyme was pyrimidine-specific, but it is now known that there is only a relative, but high, selectivity for cleavage of bonds next to pyrimidine nucleosides. For example, uracil nucleotides are about 400 times as reactive as adenine nucleotides.[51] There is also some

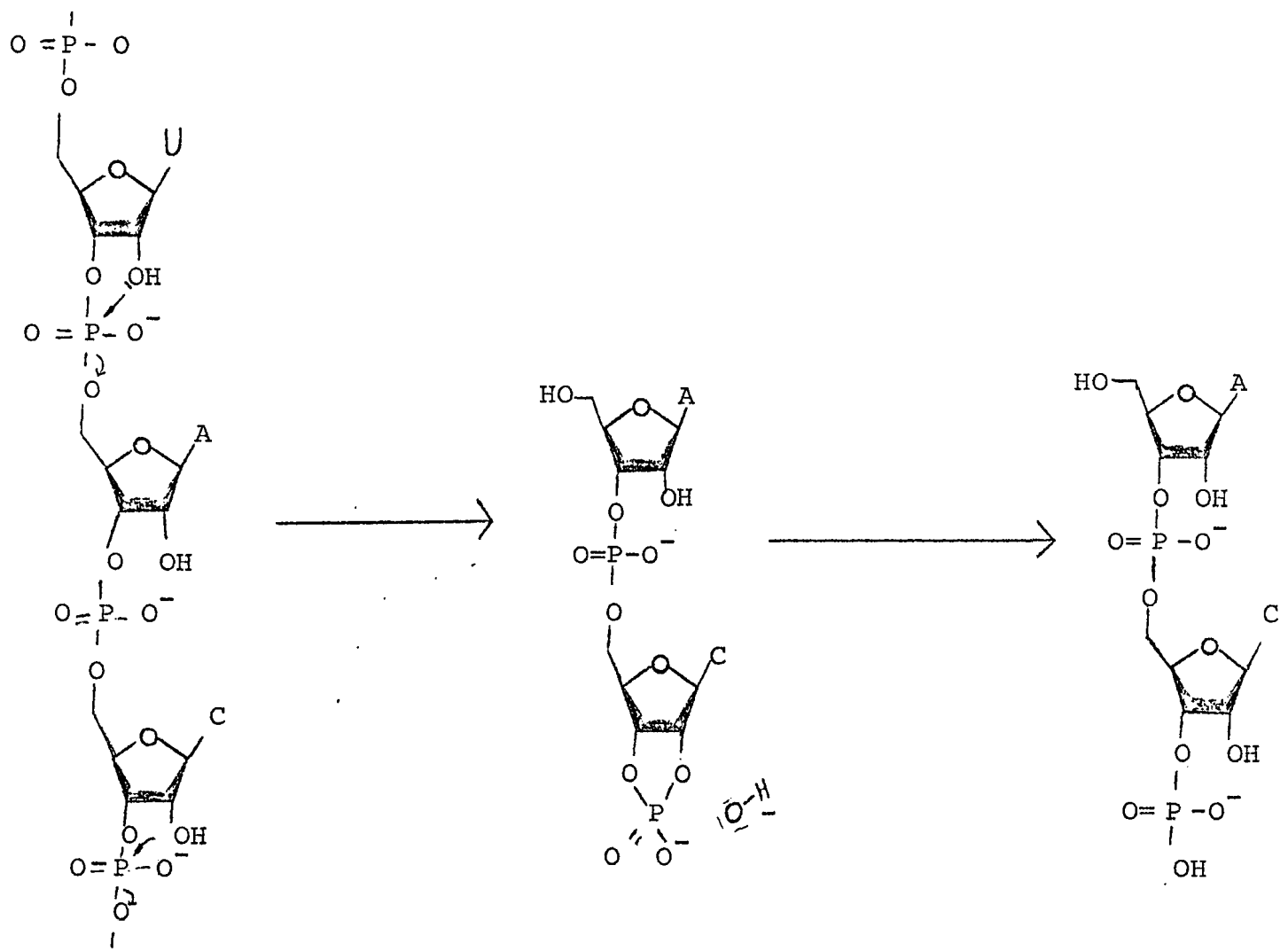


Figure 6. The Two-Step Hydrolysis Reaction Catalyzed by Ribonuclease

selectivity concerning the "nonspecific base" on the departing nucleotide following the one from which the cyclic phosphate is formed by the transesterification reaction. Relative cleavage rates may span a 300-fold range as the "nonspecific base" is varied.[52] No metal ions are required for the catalytic action.[53] Ribonuclease is inhibited by the final hydrolysis products, such as cytidine-3'-phosphate, as well as other nucleotides.[54,55]

The various areas of research directed at yielding information and ideas concerning all the intricate details of the mechanism of action of RNase A have not yet been merged into a unified, irrefutable picture. However, tremendous advances have been made in recent years and it may not be long before the enzyme is perfectly understood. The object of this discussion is to present some of the ideas being currently cogitated upon when postulating mechanisms. The discussion will begin with conclusions arrived at by G. G. Hammes and coworkers based on information from relatively new fast reaction kinetics techniques along with results from steady-state kinetics.[43,56,57]

The following details of the mechanism are well established (Eq. 10):

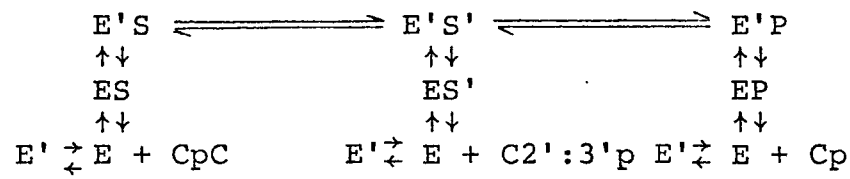
$$\begin{array}{c} E + S \rightleftarrows ES \rightleftarrows ES' \rightleftarrows EP \rightleftarrows E + P \\ \uparrow\downarrow \\ E + S' \end{array} \qquad \text{(Eq. 10)}$$

where E represents the enzyme, S represents an oligo- or polynucleotide or the model substrate CpC (cytidylyl-3',5'-cytidine), S' represents the cyclic phosphate intermediate phosphate such as cytidine-3',5'-cyclic phosphate, and P represents the final 3'-phosphate product.

The various enzyme and enzyme-substrate complex species in the above pathway are each in equilibrium with dissociating and reassociating hydrogen ions (e.g., ES may be represented as $EHS \rightleftarrows ES + H^+$). The extent of such ionization is different for each of the various enzyme and enzyme-substrate complex species. This allows study of the various steps of the mechanism by temperature jump kinetic techniques by following the change in extinction of a pH indicator upon perturbation of an equilibrium by a temperature jump. The mechanism has been divided into three steps for such study: (1) $E + S \rightleftarrows ES$, (2) $E + S' \rightleftarrows ES'$, (3) $E + P \rightleftarrows EP$. The first two "equilibria" were studied by stopped flow T-jump, while the third step was studied by T-jump techniques.

These fast reaction kinetics studies reveal a unimolecular isomerization of each enzyme-substrate complex following the bimolecular combination step and preceding the essentially irreversible steps, which are too slow to be observed. Each of the three steps is therefore of the form $E + X \rightleftarrows EX \rightleftarrows E'X$. There may be additional isomerizations too slow to be observed

by these fast reaction techniques. An isomerization of the enzyme alone is also observed, and the dynamics suggest that it is different from the isomerizations of the complexes, except for the ES' ⇄ E'S' complex isomerization. Assuming that the mechanism is sequential, the following minimal mechanism has been proposed on the basis of the above kinetic data (Eq. 11). (It should be mentioned that CpC was used as a model substrate for RNA, and that the second and third steps were studied using both cytidine and uridine mononucleotides. Earlier papers incorrectly interpreted the third step[57] to include two enzyme-substrate complex isomerizations instead of one).



(Eq. 11)

That His 12, His 119, and Lys 41 are involved in the catalysis is implied by numerous chemical studies and X-ray diffraction studies. Numerous mechanisms have been postulated to explain how the side chains of groups such as these might catalyze the reaction. Some of these are shown in the following pages. These mechanisms present some of the ideas one might want to consider before proposing a mechanism.

G. G. Hammes,[43,56,57] using both fast reaction and steady state kinetic techniques, has concluded that three

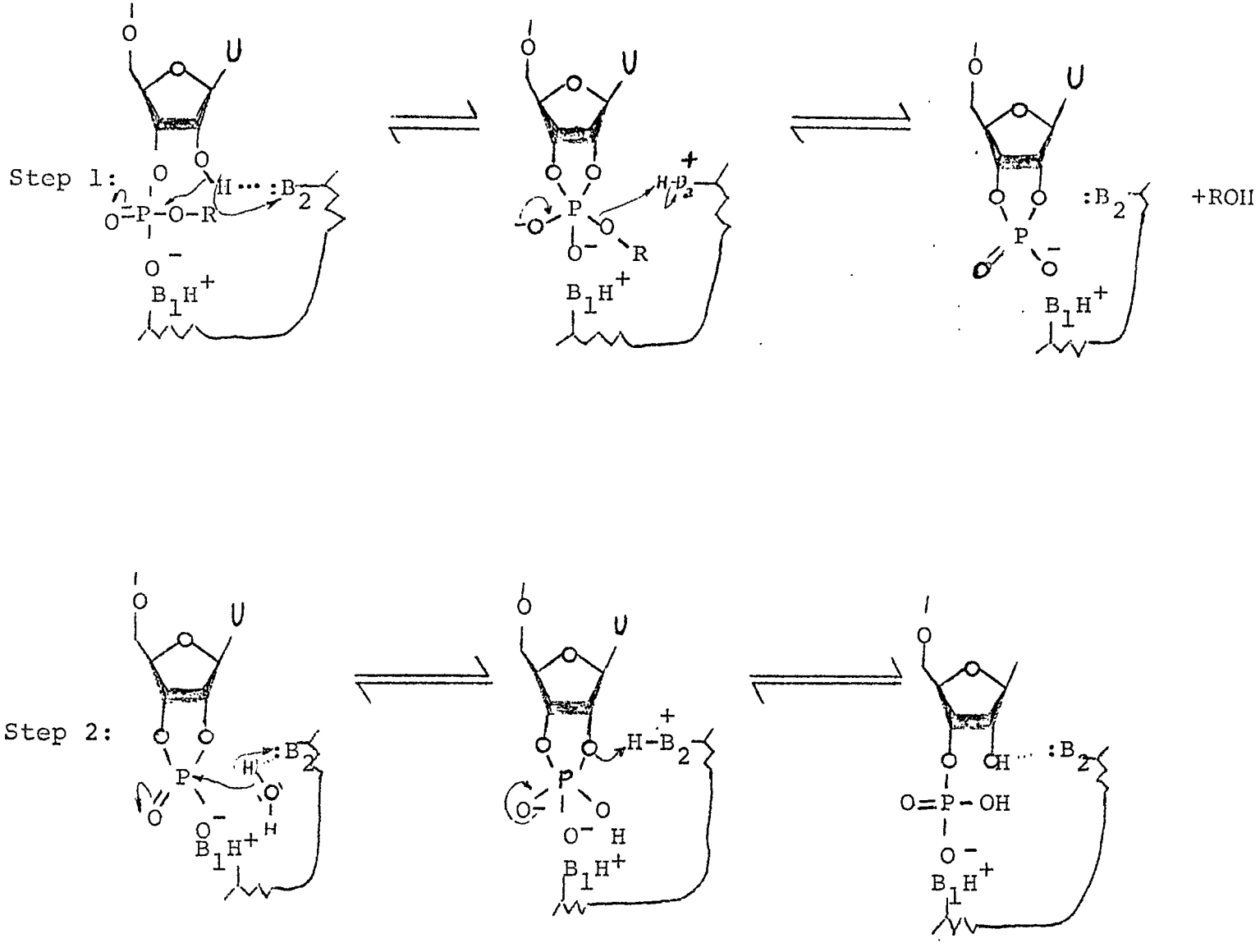
ionizing groups in the pK range 5 to 8 are implicated in the mechanism, probably His 119, His 12, and His 48. The third (probably His 48) is associated with the enzyme conformational change and is probably not in direct interaction with the substrate. Lys 41 is also probably involved. He also made use of the electron density maps of D. Harker, et al.[67] when making these proposals. He proposes the formation of an initial complex, primarily with histidines 12 and 119 in an active site located in a groove, followed by an isomerization with lysine 41 taking over part of the binding function so that His 12 and/or His 119 can take part in the catalysis. The isomerization is characterized by a slight closing of the groove (with assistance of His 48, possibly) in this proposal. He postulates no exact mechanism, but maintains these conclusions are consistent with previously proposed mechanisms, such as the one illustrated in Figure 7[56] where B_1 and B_2 do not necessarily represent only one group and where during the course of the reaction a group may move into or out of the function represented by B_1 or B_2.

Most of the recently proposed mechanisms, including the one illustrated above, involve one group acting first as a general base towards the 2'-OH and then as a general acid toward the departing 5'-OCH_2 of the next nucleotide in the cyclizing step (or in the above illustration, where B_2 is taken to be only one group in the present discussion). This

FIGURE 7.

MECHANISM OF RIBONUCLEASE

B_1H^+ represents one or two protonated base functions, while B_2 represents one base function. Base(s) B_1 perform binding duties only, while B_2 catalyzes the reaction and also may participate in binding the substrate.


Step 1:
Step 2:
+ROH

implies an adjacent type of displacement at the pentacovalent phosphorous intermediate (or transition state), and this suggests that there is an intermediate with a long enough life time to undergo pseudorotation at pentacovalent phosphorous (rather than a direct displacement mechanism), according to physical organic studies not involving enzymes.[58] Pseudorotation at pentacovalent phosphorous, which is trigonal bipyrimidal, is illustrated in the mechanism in Figure 8.[58]

His 119 and His 12 were implicated in the active site and in the enzymic catalysis a number of years ago, when it was shown that these residues could be specifically carbomethoxylated (at their histidine side chains), and that only one or the other group would react on each molecule. Either derivative was enzymically inactive, and any conformational change was judged inadequate to account for the loss of activity.[61] The reactions are greatly slowed by dianion inhibitors of enzymic activity, but not the monoanions.[62,63,64] These and many other aspects of carbomethoxylation reactions (reviewed by Barnard in 1969[53]), together with X-ray data implicate His 119 and His 12 in the catalysis.

Amino groups of lysine side chains can be selectively guanidinated by O-methylisourea, and when all of the lysine residues are guanidinated except Lys 41, the ribonuclease derivative still possesses enzymic activity. When Lys 41 is

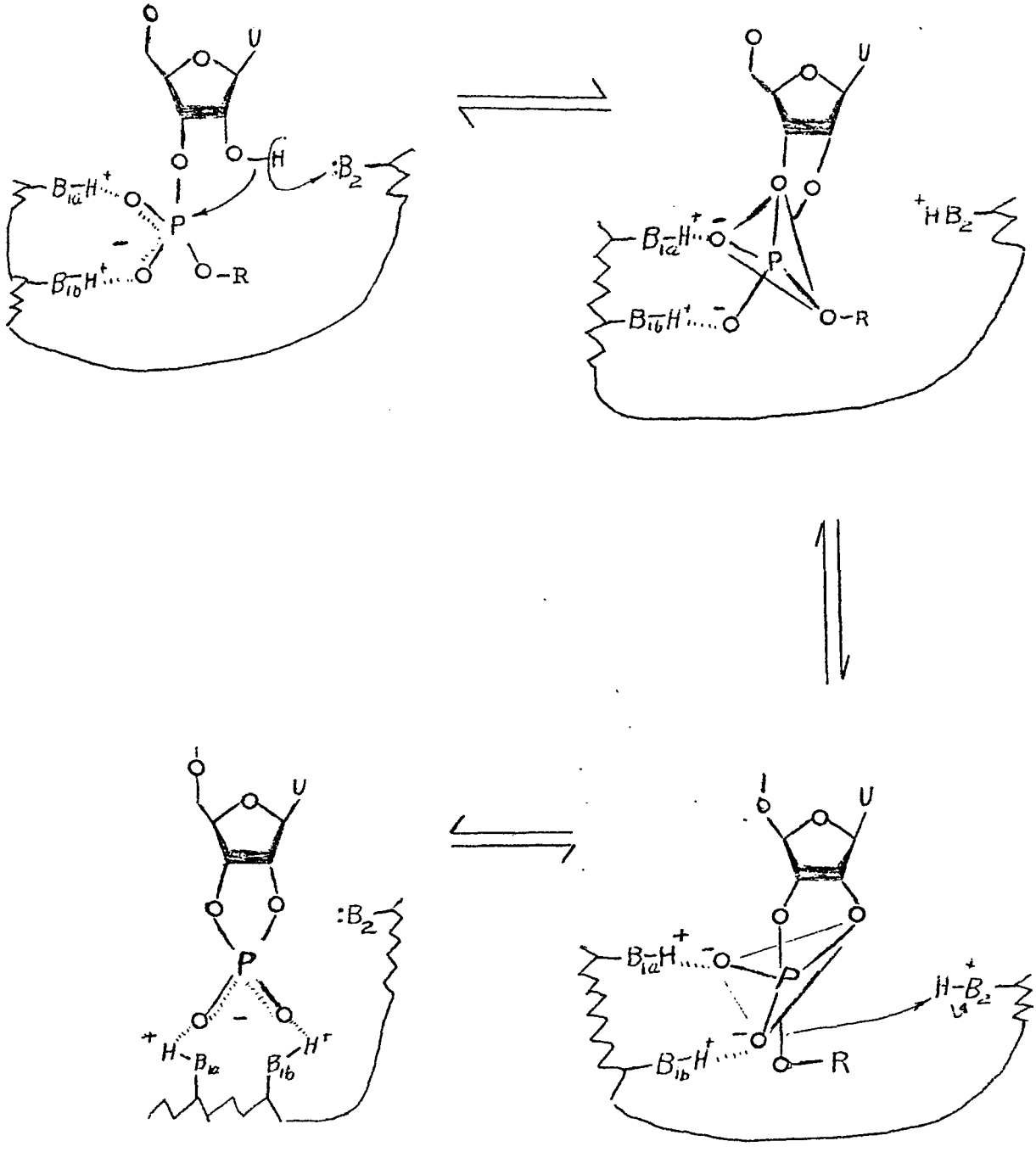
FIGURE 8

A MECHANISM OF RIBONUCLEASE INVOLVING PSEUDOROTATION AT PENTACOVALENT PHOSPHOROUS

B_{1a} = His 12

B_{1b} = LYS 41

B_2 = His 119



guanidinated there is no enzymic activity.[65] This is an example of the chemical evidence which suggests that Lys 41 is in the active site and is involved in the enzymic activity. Also Lys 41 rapidly reacts with 2,4-dinitrofluorobenzene at 15 degrees centigrade and pH 8.0 without disorganization of the structure, and the reaction is strongly inhibited by nucleotides (which also inhibit the enzymic activity competitively); the product has no enzymic activity.[53,61,66]

X-ray crystallographic study of the enzyme gradually progressed until in 1967 an electron density map at 2 angstroms resolution became available, and this revealed intricate details of the three-dimensional structure of the molecule, including location and shape and other details of the active site.[67]

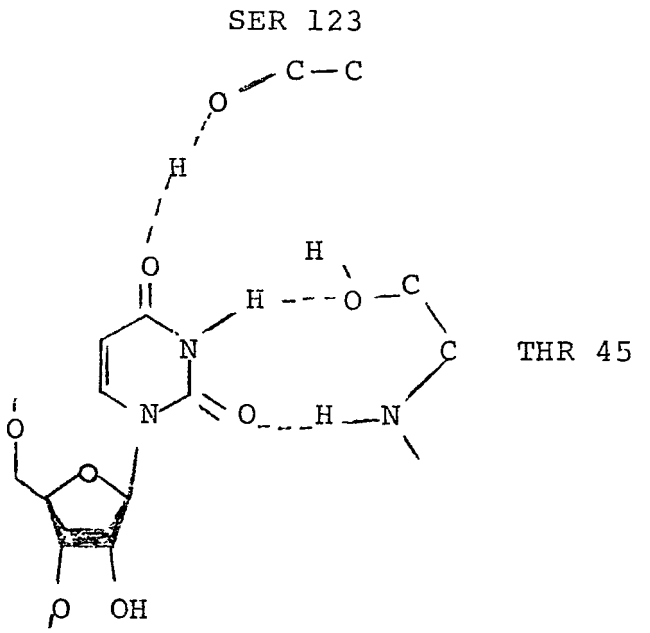
A crystallographic structure of RNase S also became available in 1967.[68] RNase S is the enzymically active product of cleavage of RNase A by the bacterial proteolytic enzyme subtilisin at the alanylseryl bond 20 to 21,[53,69] when the two fragments, S-Protein and S-Peptide, are both present and noncovalently bonded. The close similarity of the crystallographic structures of RNase S and RNase A indicates that probably the same protein conformation exists in aqueous solution, since two solvents of greatly different nature were used in the crystallizations.[53] Both crystals contained the highly stabilizing dianion ligands (such as phosphate and

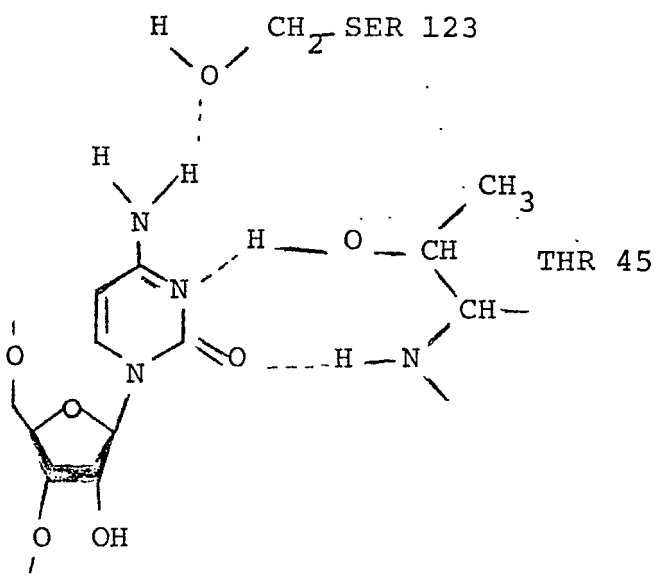
dianion nucleotides) at the active center, which were required for formation of suitable crystals. The species involved in enzymic reactions are monoanions.

The RNase A molecule is kidney shaped, and has a deep depression in the middle of one side. The binding site for substrate and inhibitors is in this depression, as deduced from locations of phosphate, arsenate, cytidine-2'-phosphate, uridine-5'-phosphate, and other dianion ligands.[67,68,70] This site contains His 12, His 119, and Lys 41. His 12 and His 119 are hydrogen bonded to the phosphate group of cytidine-3'-phosphate. (It is interesting to recall here that some of the proposed mechanisms show two cationic amino acid side chains hydrogen bonded to the usual monovalent anion substrates, and Hammes has proposed that the complex isomerization involves the closing of the groove and Lys 41 taking over part of the binding function. His 48 is in a suitable position for participation in closing the groove). The position of Lys 41 in the crystallographic structure of the complex with cytidine-3'-phosphate remained uncertain.[59] On the other hand, earlier crystallographic structures showed His 119 and Lys 41 bound to a phosphate dianion ligand.[53] Cytosine and uracil bases form three hydrogen bonds to two amino acid residues, Thr 45 and Ser 123, as illustrated (Figure 9).[59] Also there appears to be a hydrophobic stacking interaction with Phe 120.[53,60]

FIGURE 9

METHOD OF BINDING OF URACIL AND CYTOSINE SUBSTITUENTS OF RNA BY RIBONUCLEASE


SER 123
THR 45


SER 123
THR 45

CHAPTER II

EXPERIMENTAL

Analytical Methods

Analytical scale thin layer chromatography (tlc) was done with Brinkman pre-coated tlc glass plates, coated with silica gel GF-254 (calcium sulfate binder, indicator added which fluoresces under a short wave ultraviolet lamp). These plates were purchased in the 5 x 20 cm size and cut to about 2 1/2 x 5 cm with a glass cutter. These tiny tlc plates were developed along the 5 cm dimension. Spots were visualized by an ultraviolet lamp when possible; otherwise, a solution of potassium permanganate (5 mg per ml) in 2.5 N sodium hydroxide was applied.

Column chromatography on silica gel was done with Brinkman 70-230 mesh or with Baker Analyzed 60-200 mesh silica gel. The former was preferred when available because of a considerably smaller occupied volume and (liquid) column volume. Unless otherwise stated, fifty to one hundred grams of silica gel were used per gram of the substance being chromatographed. The length of the column of silica gel was at least fifteen times the diameter. Eluent was chosen by preliminary tlc experiments (silica gel G), using a liquid mixture which barely moved the desired compound (Rf about 0.1 to 0.2) and which separated it from other components, at least after multiple developments. Compounds had a larger "Rf" on silica gel columns than on the tlc plates.

Elution was done at a rate of one column volume per hour or a little slower.

Elemental analyses were performed by Galbraith Laboratories, Inc., Knoxville, Tennessee.

All melting points were taken on a modified Hershberg melting point apparatus equipped with a motor driven stirrer and a set of Anschutz thermometers in 50° ranges, and no correction was applied.

Nmr spectra at 100 MHz were obtained from a Varian HA-100 spectrometer. Nmr spectra at 60 MHz were obtained from a Varian T-60 spectrometer.

4,6-O-Ethylidene-D-glucopyronose(II)

This compound was prepared according to the method of Hockett, Collins, and Scattergood[71] with a few modifications of technique. A mixture of 90.0 g (0.500 mol) of reagent grade Baker anhydrous granular D-glucose and 77 ml (0.58 mol) of paraldehyde was placed in a 3-neck round bottom flask which was fitted with a mechanical stirrer. To this slurry was added 0.5 ml of concentrated sulfuric acid, the glass stopper was re-inserted, and the stirrer was immediately started. This procedure prevents formation of dark colored acid-catalyzed decomposition products of paraldehyde. The reaction was followed by the literature method of boiling samples of the reaction slush with an equal volume of p-dioxane, and the

reaction was considered to be complete when all went into solution. The reaction slush was scraped down from the side of the flask from time to time with a bent spatula to facilitate more thorough mixing, and it became necessary to add 10 ml of paraldehyde after 12 hours. The reaction was complete after 36 hours. Reaction time varied from preparation to preparation as in the literature, but was never as long as the shortest literature reaction time (48 hrs). Ethyl acetate (100 ml) was mixed in by means of the stirrer, and the reaction mixture was stored at 5° for several hours. The product was collected in a 9 cm Buchner funnel and washed with ethyl acetate (100 ml) and ether (100 ml) to give 92 g of crude product. The crude product was recrystallized from 550 ml of absolute ethanol containing 4 ml of concentrated ammonium hydroxide. The thoroughly pulverized crude product was boiled for no longer than 5 minutes in the recrystallization solvent to dissolve it without forming colored decomposition products, treated with Celite, and filtered through a 32 cm Whatman number 1 filter paper (fluted). The filtrate was weakly alkaline (alkacid test paper). Recrystallization from ethanol brought only to neutrality sometimes gave a product with a low melting point. The filtrate was stored for 24 hr at 5°, and the product was collected and washed with ether to give 59.7 g (58%) of a mixture of α- and β-4,6-O-ethylidene-D-glucopyranose, mp 177°-179°; 100 MHz nmr (D_2O)

δ 5.21 (d, 1/3 H, H(1α), J = 4 Hz), 4.83 (q, 1 H, H(7), J = 5 Hz), 4.56 (d, 2/3 H, H(1β), J = 8 Hz), 1.34 (d, 3H, CH_3, J = 5 Hz).

2,4-O-Ethylidene-D-erythrose(III)

A number of procedures for sodium metaperiodate oxidation of 2,4-O-ethylidene-D-glucose(II) have been published.[72,73,74,75] The following is a modification of the procedure of Rappoport and Hassid for preparation of the L-isomer of III.[76] A solution of sodium metaperiodate (24.2 g, 113 m mol) in water (334 ml) was added portionwise with stirring to a mixture of II (11.68 g, 56.8 m mol) and sodium bicarbonat (9.53 g, 113 m mol) in a one-liter flask cooled by a water bath. An ice bath was used briefly to keep the temperature below 35°, but overcooling causes precipitation of periodate. After one hour the solution was evaporated to dryness in vacuo at room temperature. The residue was extracted at room temperature with ethyl acetate (3 x 100 ml), and the combined ethyl acetate extracts were evaporated in vacuo at room temperature. To remove ethyl acetate from the resulting syrup, air was admitted and the syrupy solution was swirled. Careful evaporation in vacuo gave a white foam, which was left under vacuum overnight to give 8.1 g (98%) of III. Pure samples were obtained by precipitation from ethyl acetate at 5° to give an amorphous

white solid: 100 MHz nmr (D_2O) δ 5.14-5.17 (m, 1 H, H(1)), 4.78 (q, 1 H, H(5), J = 5 Hz), 1.30 (d, 3 H, CH_3, J = 5 Hz).

The unpurified product, including the foam, should be used as soon as possible after it is prepared. The amorphous solid (precipitated from ethyl acetate) is stable enough to be stored for a short period. Compound III s sensitive to heat and should be kept under 35°.

D-Erythrose(IV)

The following modification of literature procedures minimized decomposition.[72,73,74] A solution of 7.08 g (48.5 m mol) of 2,4-O-ethylidene-D-erythrose(III) in 0.1 N sulfuric acid (144 ml) in a 500 ml round bottom flask was heated in a 55° oil bath while dry nitrogen was bubbled through vigorously. The progress of the reaction was followed by tlc (silica gel G, ethyl acetate, visualized with $KMnO_4$). The faster moving spot is the reactant, and the product does not move in this solvent but can be checked for homogeneity by developing with 1:3 methanol-chloroform. When the reactant spot disappeared (about 4 hr) the reaction was cooled to room temperature and the solution was neutralized to pH paper by stirring for a minimum time with an excess of strong base ion exchange resin, OH form (Rexyn research grade) until weakly alkaline, followed by strong acid ion exchange resin, H form (Rexyn research grade), without

filtering off the basic resin. The resin was removed by suction filtration, and the filtrate was evaporated to dryness in vacuo at room temperature to give, after thorough vacuum drying, 5.24 g (90%) of a colorless syrup. Ordinarily, the viscous product was not thoroughly freed from solvent. It was dried by evaporation with pyridine and analyzed as the triacetate derivative(V).

D-Erythrose Triacetate(V)

This compound was prepared essentially according to the procedure of Murray and Pokop,[10] but the reaction time was reduced to 15 minutes in an ice bath and one hour at room temperature. This resulted in a purer, less viscous product. D-Erythrose was freed from water by evaporation with pyridine before use. After reaction with acetic anhydride in dry pyridine, the product was worked up as in the literature. The resulting syrup was evaporated with carbon tetrachloride to remove pyridine and stirred magnetically overnight under vacuum at room temperature to give a colorless syrup (when colorless erythrose was used) in yields of about 85%.

The crude syrup was warmed to 50° and transferred to the molecular still by means of a transfer pipet made from a length of 8 mm glass tubing, narrowed to 2 mm at one end. The crude product was molecularly distilled in 4 to 5 g quantities with competing decomposition (65-90°, .004 to .06 mm)

FIGURE 10

NMR SPECTRUM OF D-ERYTHROSE

TRIACETATE (V) IN CCl_4

AT 60 MHz

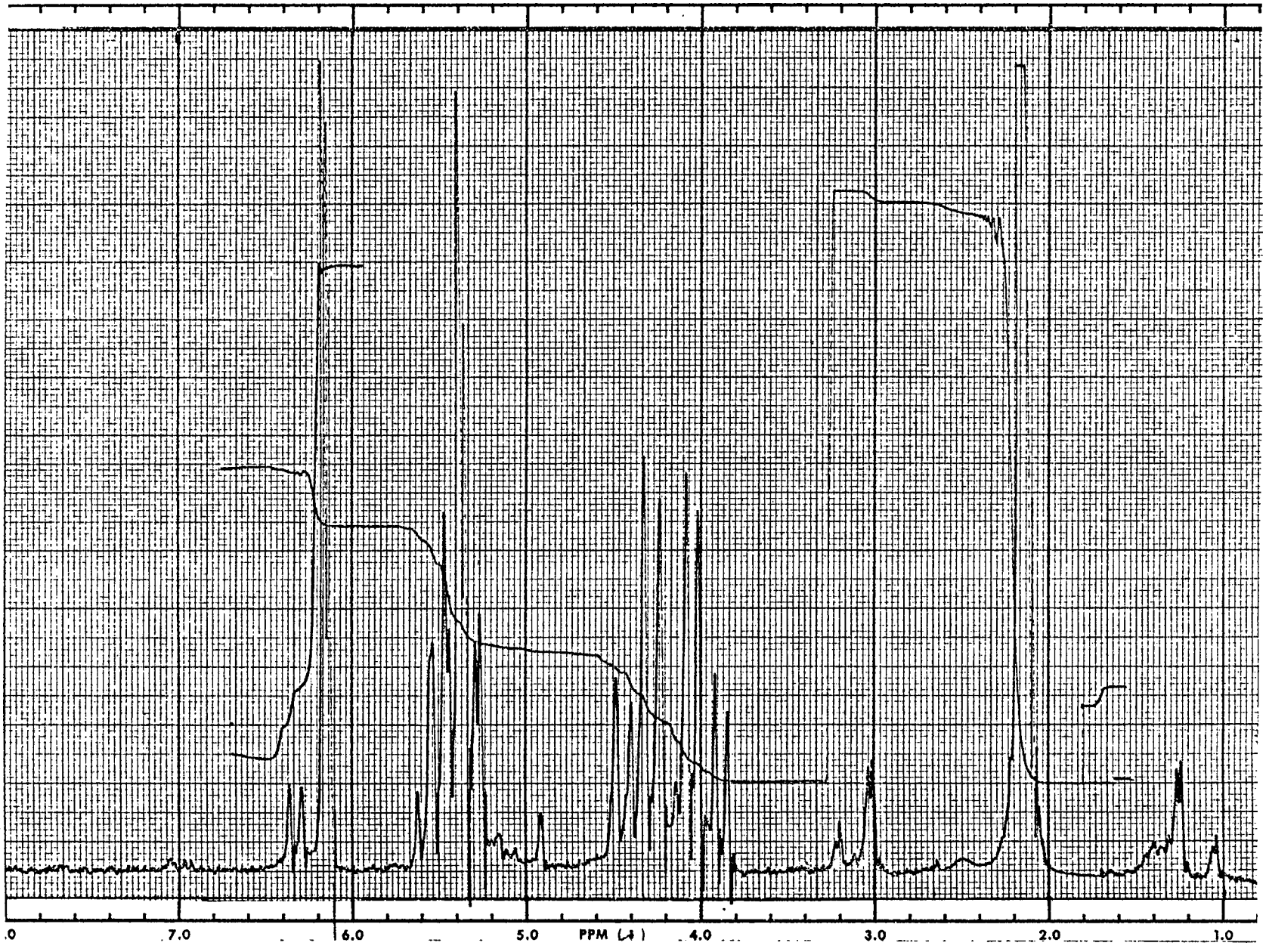
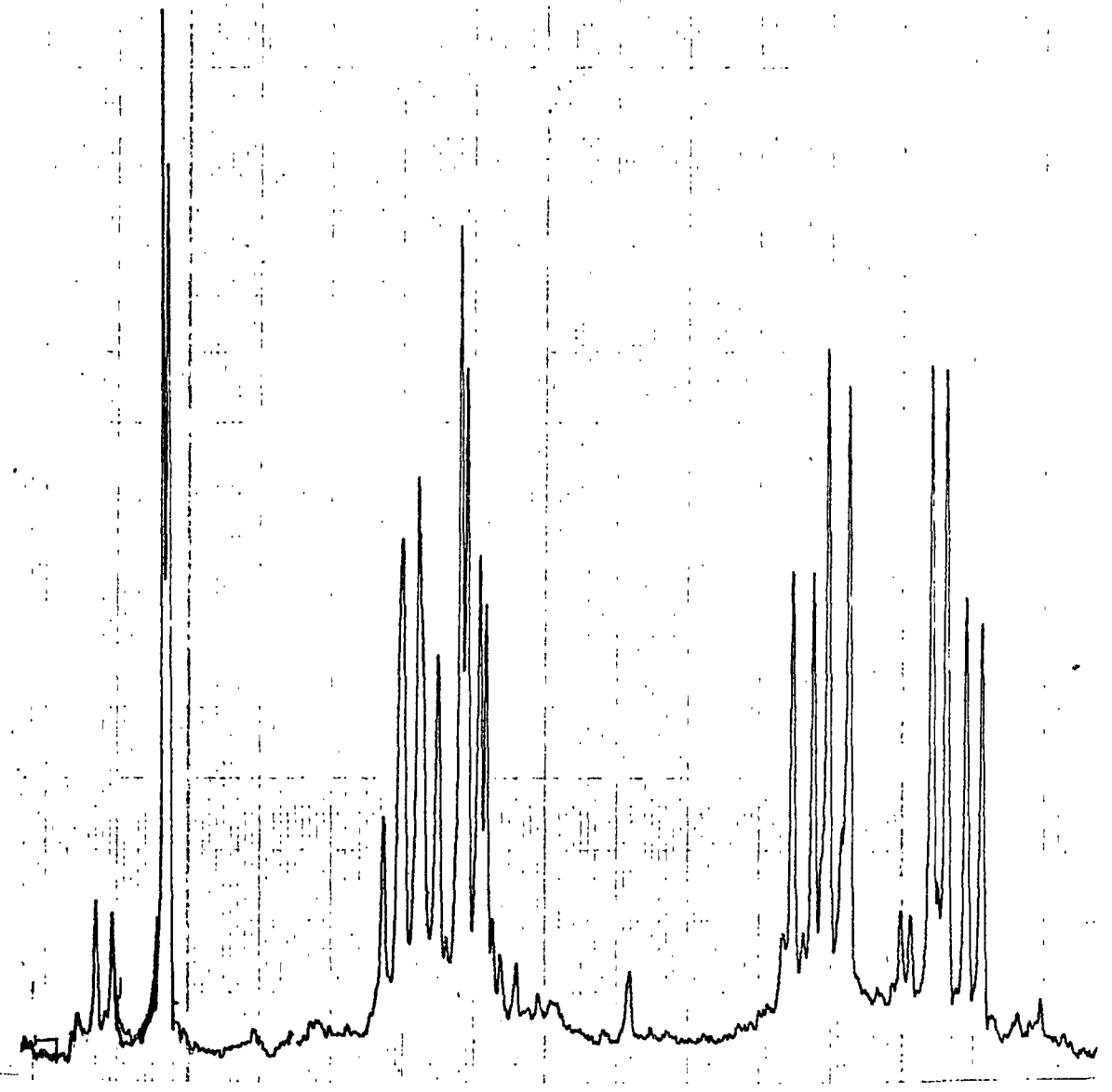

.0
7.0
6.0
5.0
PPM (δ)
4.0
3.0
2.0
1.0

FIGURE 11

NMR SPECTRUM OF D-ERYTHROSE

TRIACETATE (V) IN CCl_4

AT 100 MHz



in yields up to 60% of colorless oil which was predominantly the β-anomer; 100 MHz nmr (CCl_4) δ 6.24 (d, 0.15 H, H(1α), J = 4 Hz), 6.08 (d, 0.85 H, H(1β), J = 2 Hz), 5.39 (m, 1H, H(3β), J = 6.6 and 4 Hz), 5.21 (m, 1 H, H(2β), J = 2 and 6 Hz), 4.23 (m, 1 H, H(4bβ), J = 6 and 10 Hz), 3.85 (m, 1 H, H(4aβ), J = 4 and 10 Hz), 2.1 (9 H, CH_3).

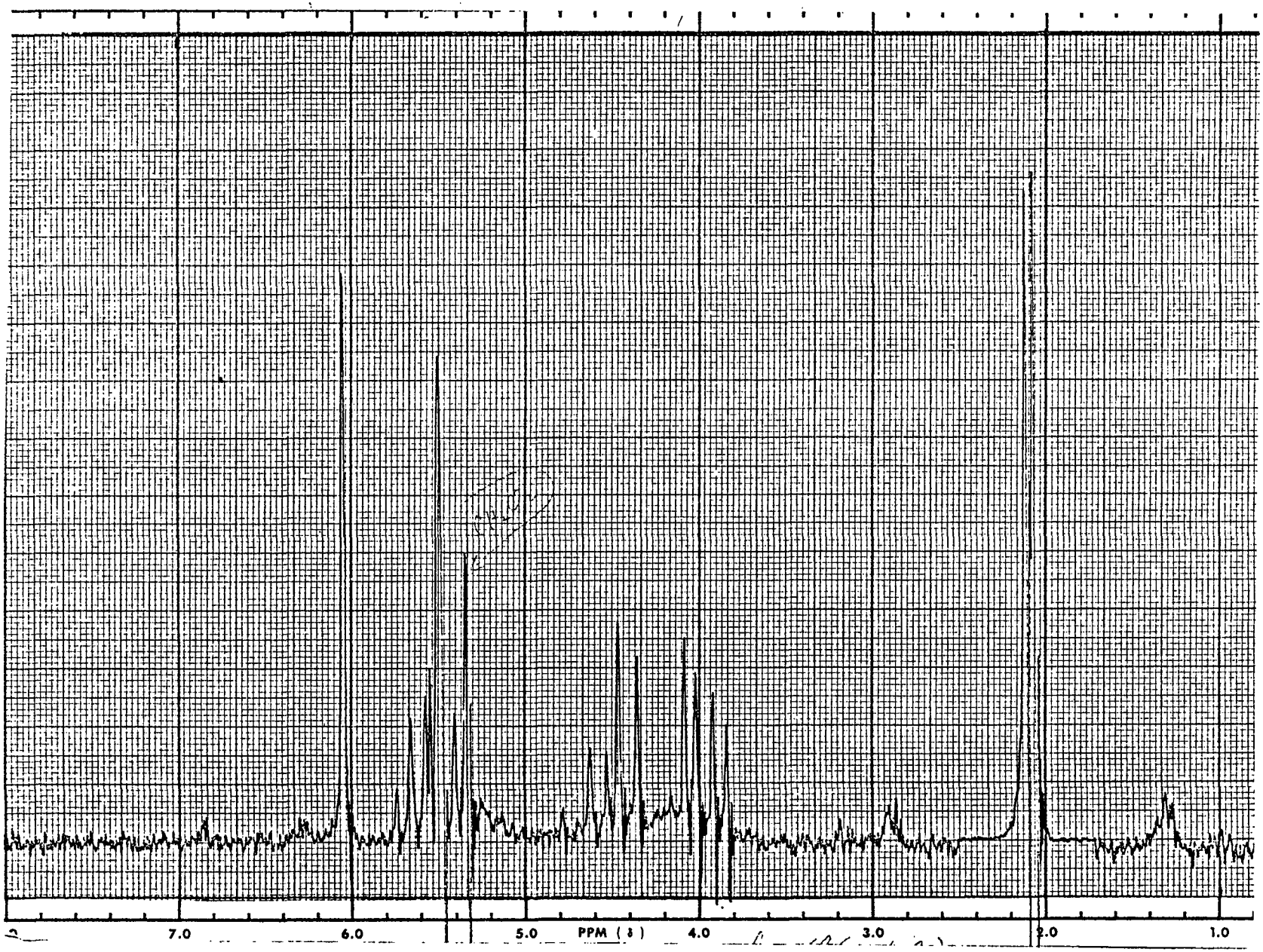
When molecular distillation was attempted under too vigorous conditions, decomposition products appeared in the distillate. Compound V, crude or distilled, was examined for purity by nmr (CCl_4) before use. The nmr in $CDCl_3$ was less satisfactory.

2,3-Di-O-Acetyl-β-D-erythrofuranosyl Cloride(VI)

Hydrogen chloride (Matheson reagent grade), passed through concentrated sulfuric aci;, silica gel and drierite, was bubbled into a one pint bottle of dry dichloromethane in an ice bath for ten minutes with minimum exposure to air. The hydrogen chloride-dichloromethane solution was protected from moisture and kept cold. A solution of molecularly distilled D-erythrose triacetate (2.08 g, 8.45 m mol) in dry dichloromethane in a 100 ml round bottom flask (14/35 S.T.) was cooled to 0°, and cold hydrogen chloride-dichloromethane (about 70 ml) was added with minimum exposure to air. The reaction was protected from moisture with a glass stopper and stored at -17° for 3 hr. The reaction vessel was then fitted

FIGURE 12

NMR SPECTRUM OF β-D-ERYTHROFURANOSYL CHLORIDE (VI) IN CCl_4 AT 60 MHz


7.0
6.0
5.0
PPM (δ)
4.0
3.0
2.0
1.0

with a distillation adaptor (with stopcock) and evaporated to dryness in vacuo at room temperature (water bath), protecting the vacuum pump with two dry ice-acetone baths followed by a tube filled with sodium hydroxide pellets. The last traces of acetic acid were removed by evaporation with dry carbon tetrachloride. The last bit of solvent was removed by warming the flask in a 30° water bath and rotating the flask while in a nearly horizontal position. The yield was 1.72 g(91%) of VI as a pale yellow oil: 60 MHz nmr (CCl_4) δ 6.1 (s, 1 H, H(1)), 5.4 to 5.7 (m, 2 H, H(2) and H(3)), 4.5 (m, 1 H, H(4b), J = 6 Hz and 10 Hz), 4.0 (m, 1 H, H(4a), J = 5 and 10 Hz), 2.1 (6 H, CH_3).

2,4-Bis(trimethylsilyloxy)pyrimidine

This compound was prepared essentially according to the method of Iwai, Nishimura, and Shimizu[77] from uracil (11.3 g), chlorotrimethylsilane (24.9 ml), and triethylamine (25.8 ml) in p-dioxane (200 ml total) except that the reaction time was extended to 24 hours. The product slush was suction filtered through a 150 ml medium porosity sintered glass filter funnel, protected from moisture. The filtrate was evaporated in vacuo without magnetic stirring in a water bath to give a crude solid product. The crude product was vacuum distilled (lit. 116°, 12 mm) to give 19.65 g (76% based on 11.3 g of uracil, lit. 72% based on

uracil not recovered) of a colorless oil which quickly crystallized. Alternatively, an oily crude product could be obtained by more rapid evaporation of p-dioxane with magnetic stirring. Vacuum distillation afforded a colorless oil which did not crystallize when stored several weeks.

This compound is rapidly hydrolyzed to uracil by atmospheric moisture, and should be stored in a desiccator. Transfer to reaction vessels was done by dried disposable pipets with minimum exposure to air.

Reaction of VI with 2,4-Bis(trimethylsilyloxy)pyrimidine

A mixture of distilled liquid 2,4-bis(trimethylsilyloxy)-pyrimidine (4.0 ml, 15.6 m mol) and VI obtained from 2.58 g of V (10.5 m mol) was heated in a glass-stoppered 14/20 S.T. 100 ml round-bottom flask for 24 hr at 80° and 1 hr at 180°. The cooled solution was stirred with absolute alcohol (50 ml) for 30 minutes, and the precipitated uracil was filtered off. The filtrate was evaporated to dryness in vacuo to give 1.9 g (60%) of a crude black syrup. From nmr spectra of the crude products of a number of preparations it was determined that two nucleosides were present in varying proportions, usually about 1:3. A portion was chromatographed on silica gel, eluting with ethyl acetate-chloroform (1:1), which gave the unseparated nucleosides as a colorless syrup of good purity by nmr. The minor product was tentatively

assigned the structure 1-(2,4-di-O-acetyl-β-D-erythrofuranosyl)-uracil(VIII), and this assignment was later confirmed by the alternate preparation on page

The predominant nucleoside product was tentatively assigned the structure 1-(2,4-di-O-acetyl-α-D-erythrofuranosyl)-uracil (VII): 100 MHz nmr ($CDCl_3$) δ 9.89 (broad, 1 H, N-H), 7.47 (d, 1 H, H(6), J = 8 Hz), 6.22 (d, 1 H, H(1'), H = 5 Hz), 4.22 (d, 2 H, H(4a') and H(4b'), J app = 4 Hz), 2.1 (6 H, CH_3).

In other preparations, nucleosides were formed under conditions as mild as room temperature. At temperatures usually used for Hilbert-Johnson reactions with trimethylsilyl-protected heterocycles (150° to 200°), VI was partially decomposed to give a dark colored reaction.

Ammonolysis of the Crude Mixture Containing VII and VIII

The crude product of the above reaction was dissolved in anhydrous methanolic ammonia previously saturated at 0° (25 ml per g) and allowed to stand for 2 days at -17°. The product was evaporated to dryness in vacuo to give a syrup.

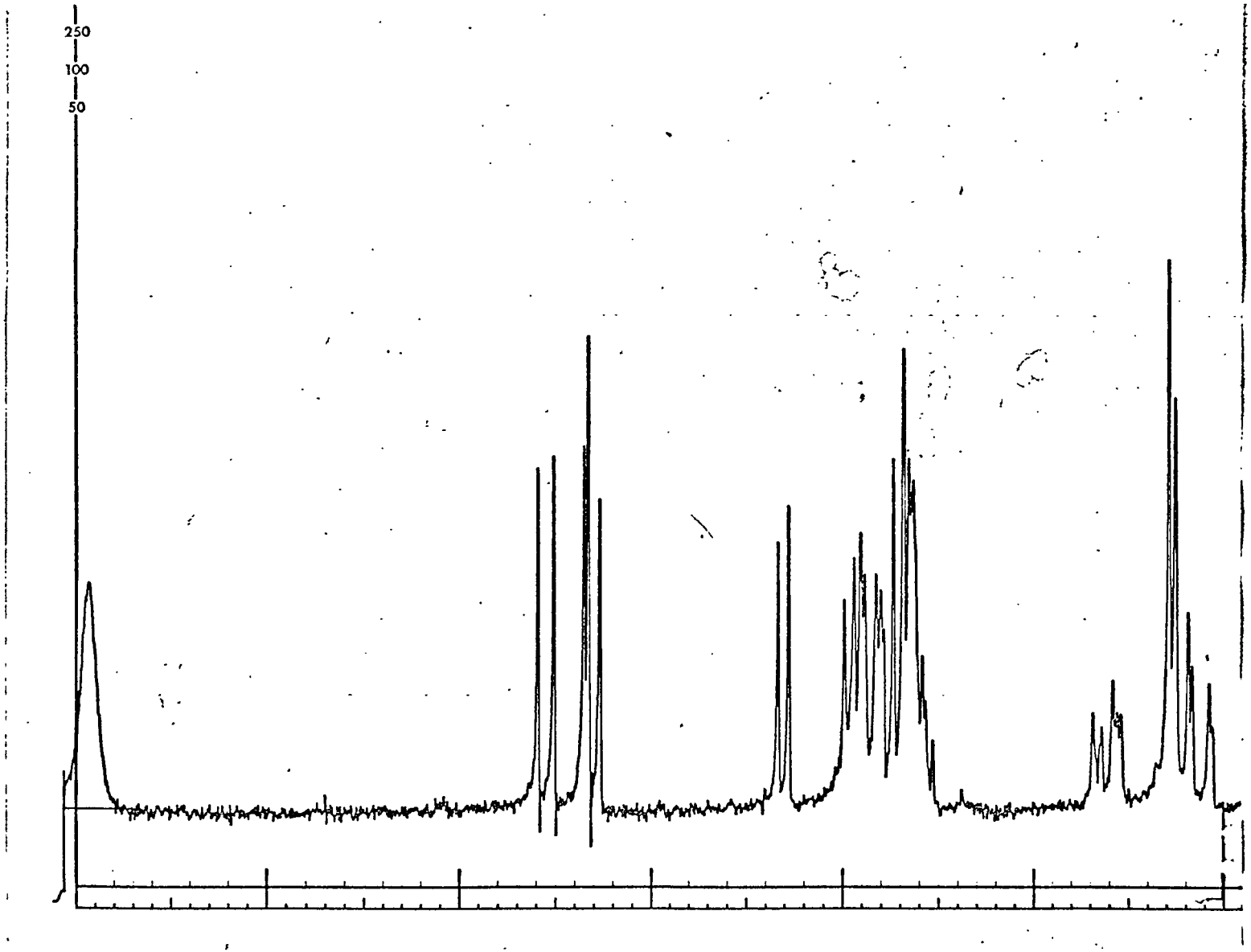
(a) 1-(α-D-Erythrofuranosyl)-uracil (IX)

The syrup obtained above was dissolved in water (2 ml per g) and allowed to stand at 5°, whereupon crystallization occurred to give pure 1-(α-D-erythrofuranosyl)-uracil. The

FIGURE 13

NMR SPECTRUM OF A 1:1 MIXTURE OF VII AND VIII

IN $CDCl_3$ AT 100 MHz


250
100
50

crystals were collected by suction filtration. The filtrate was evaporated to dryness in vacuo and the residue was chromatographed on silica gel. Elution with methanol-chloroform (1:10) yielded more IX, which began to appear after 3 column volumes: mp 204°-205.5°; 100 MHz nmr (D_2O) δ 7.87 (d, 1 H, H(6), J = 8 Hz), 6.14 (d, 1 H, H(1'), J = 5 Hz), 5.88 (d, 1 H, H(5), J = 8 Hz), 4.4 to 4.6 (m, 2 H, H(2') and H(3')), 4.12 (d, 2 H, H(4a') and H(4b'), J = 4 Hz).

Anal. Calcd for $C_9H_{13}O_6N_2$: C, 44.83; H, 4.70; N, 13.13; O, 37.33. Found: C, 44.70; H, 4.86; N, 12.91; O (by difference), 37.53.

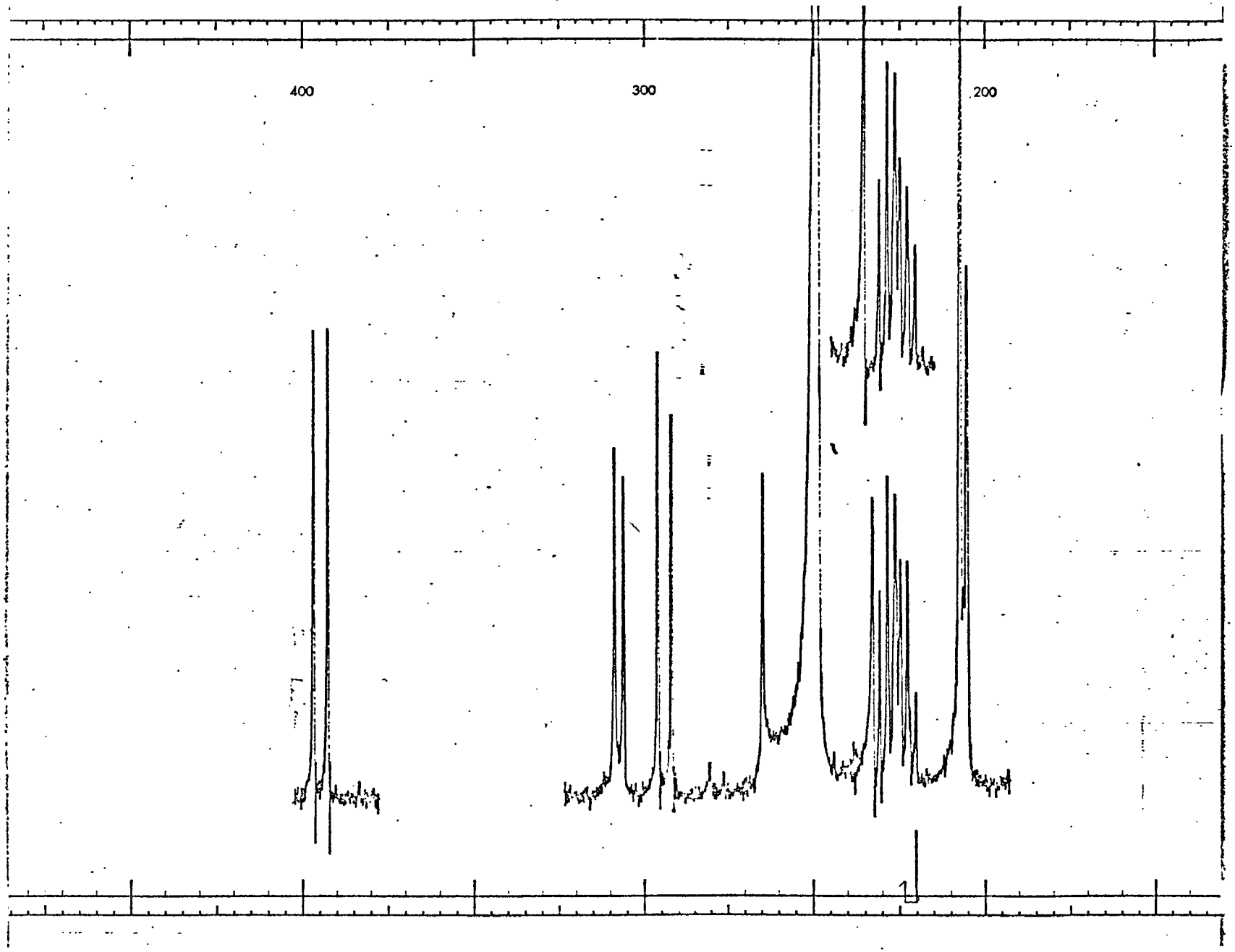
(b) 1-(β-D-Erythrofuranosyl)-uracil(X)

Further elution in the above chromatography with methanol-chloroform (1:10) gave 1-(β-D-erythrofuranosyl)-uracil, which began appearing after a total of 5 column volumes. Slow evaporation of the chromatography fractions gave X as a crystalline compound: mp 186.5°-187.5°; 100 MHz nmr (D_2O) identical to that of the principal product of the stannic chloride preparation below.

Anal. Calcd for $C_9H_{13}O_6N_2$: C, 44.83; H, 4.70; N, 13.13; O, 37.33. Found: C, 44.93; H, 4.82; N = 12.80; O (by difference), 37.45.

FIGURE 14

NMR SPECTRUM OF N(1)-(α-D-ERYTHROFURANOSYL)-URACIL (IX) IN D_2O AT 100 MHz


400
300
200

Reaction of D-Erythrose Triacetate with 2,4-Bis(trimethylsilyloxy)pyrimidine in the Presence of Stannic Chloride

Stannic chloride was distilled into a 25 ml graduated addition funnel containing 20.0 ml of dry dichloromethane until volume of liquid in the addition funnel was 25.0 ml. The addition funnel was fitted, with minimum exposure to air, to a round bottom 100 ml flask containing a solution of molecularly distilled D-erythrose triacetate (1.31 g, 5.31 m mol) and 2,4-bis(trimethylsilyloxy)pyrimidine in dry dichloromethane (50 ml). The 1:4 stannic chloride-dichloromethane solution (4.1 ml, 5.6 m mol $SnCl_4$) was added with cooling (water bath) and the reaction flask was swirled to mix the contents. The reaction was protected from moisture with a glass stopper and set aside at room temperature for 4 hours. The product was poured into a 500 ml erlenmeyer flask and diluted to 150 ml with dichloromethane. The dichloromethane solution was extracted with a saturated aqueous solution of sodium bicarbonate (100 ml) and separated.

Sometimes the layers were difficult to separate because of a precipitate, and dilution with water/and/or stirring facilitated the separation. The aqueous layer was extracted with dichloromethane (2 x 50 ml), and the organic layers were combined and extracted with water (100 ml), dried over anhydrous sodium sulfate, filtered through celite, and evaporated to dryness in vacuo to give 1.34 g of a white foam. This

crude product, which was resolved into two components by tlc (silica gel G, 1:49 methanol-chloroform), was subjected to column chromatography on silica gel.

(a) N(1)O(4)-Bis(2,3-di-O-acetyl-β-D-erythrofuranosyl)-uracil(XIII).

Elution with 1:99 methanol-dichloromethane afforded 310 mg (24%) of a white syrup (evaporation at atmospheric pressure) which had the same Rf as the front spot in the above tlc: 100 MHz nmr, ($CDCl_3$): δ 7.28 (d, 1 H, H(6), J = 9 Hz), 6.52 (d, 1 H, H(1'), J = 5 Hz), 5.5 to 6.0 (6 H; H(1"), H(5), H(2'), H(2"), H(3'), and H(3"), 4.4 to 4.7 (2 H, H(4b') and H(4b")), 3.9 to 4.1 (2 H, H(4a') and H(4a")), 2.06 to 2.13 (12 H, CH_3), and no N-H peak.

(b) 1-(2,3-Di-O-acetyl-β-D-erythrofuranosyl)-uracil(VIII)

Further elution in the above chromatography with methanol-chloroform (1:49) afforded 891 mg (56%) of a white syrup (evaporation at atmospheric pressure) which had the same Rf as the second spot in the tlc of the crude product. The 100 MHz nmr ($CDCl_3$) was identical to that of the minor product of the reaction of VI with 2,4-bis(trimethylsilyloxy)-pyrimidine: δ 10.3 (broad, 1 H, N-H), 7.38 (d, 1 H, H(6), J = 8 Hz), 5.91 (d, 1 H, H(1'), J = 5 Hz), 5.79 (d, 1 H, H(5),

J = 8 Hz), 5.6 (2 H, H(2') and H(3')), 4.56 (m, 1 H, H(4b')), 4.04 (m, 1 H, H(4a')), 2.14 (s, 3 H, CH_3), 2.08 (s, 3 H, CH_3).

(c) O(2)-(2,3-Di-O-acetyl-β-D-erythrofuranosyl)-uracil(XII) and O(4)-(2,3-Di-O-acetyl-β-D-erythrofuranosyl)-uracil(XI)

The later methanol-dichloromethane (1:49) fractions in the above chromatography afforded an unresolved mixture of VIII and another component by tlc (silica gel G, 1:49 methanol-dichloromethane, 5 developments). In another preparation using an 80% excess of 2,4-bis(trimethylsilyl-oxy)-uracil and crude V, enough of this third component was present in the crude unchromatographed product to be resolved by tlc (silica gel G, 1:49 methanol-dichloromethane, 5 developments). Chromatography performed as above afforded approximately 853 mg (24%) of the third component, 495 mg of which was completely separated from VIII. This third component was judged to be a mixture of two O-nucleosides, presumably XI and XII: 100 MHz nmr ($CDCl_3$) δ 9.76 and 9.71 (2 broad s, 2 H, N-H), 7.23 and 7.18 (2 d , 2 H, H(6), J = 8 Hz), 6.50 (d, 2 H, H(1'), J = 5 Hz), 6.02 (m, 2 H, H(2') $J_{1'2'}$ = 5 Hz, $J_{2'3'}$ = 5 Hz), 5.74 and 5.72 (2 d, 2 H, H(5), J = 8 Hz), 5.7 (m, 2 H, H(3'), hidden by H(5)), 5.59 (m, 2 H, H (4b'), J = 4 Hz and 10 Hz), 3.99 (m, 2 H, H(4a'), J = 3 Hz

FIGURE 15

NMR SPECTRUM OF XIII IN $CDCl_3$ AT

100 MHz, 3 to 7.7δ

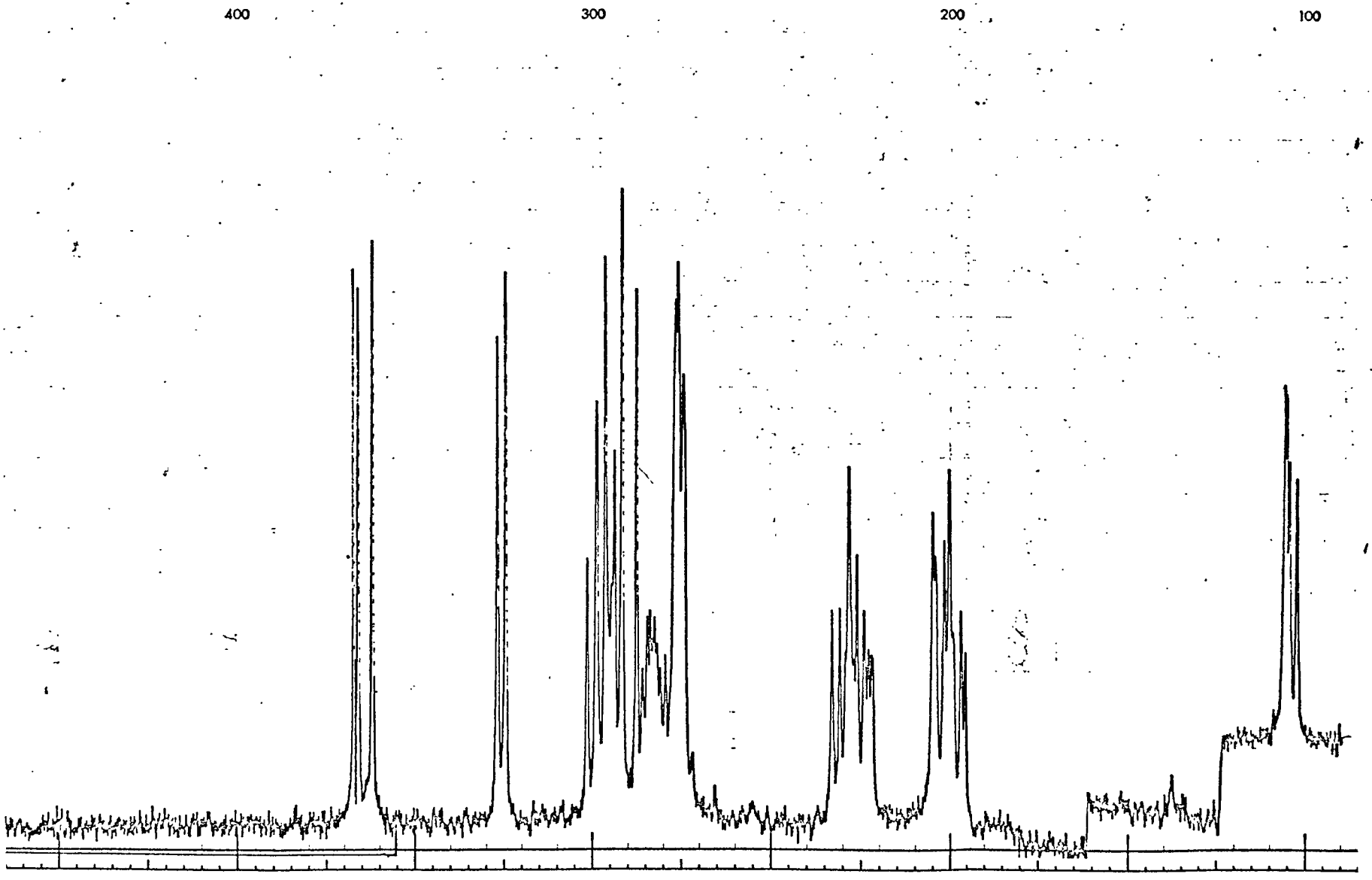

400
300
200
100

FIGURE 16

NMR SPECTRUM OF VIII FROM $SnCl_4$ REACTION

IN $CDCl_3$ AT 100 MHz

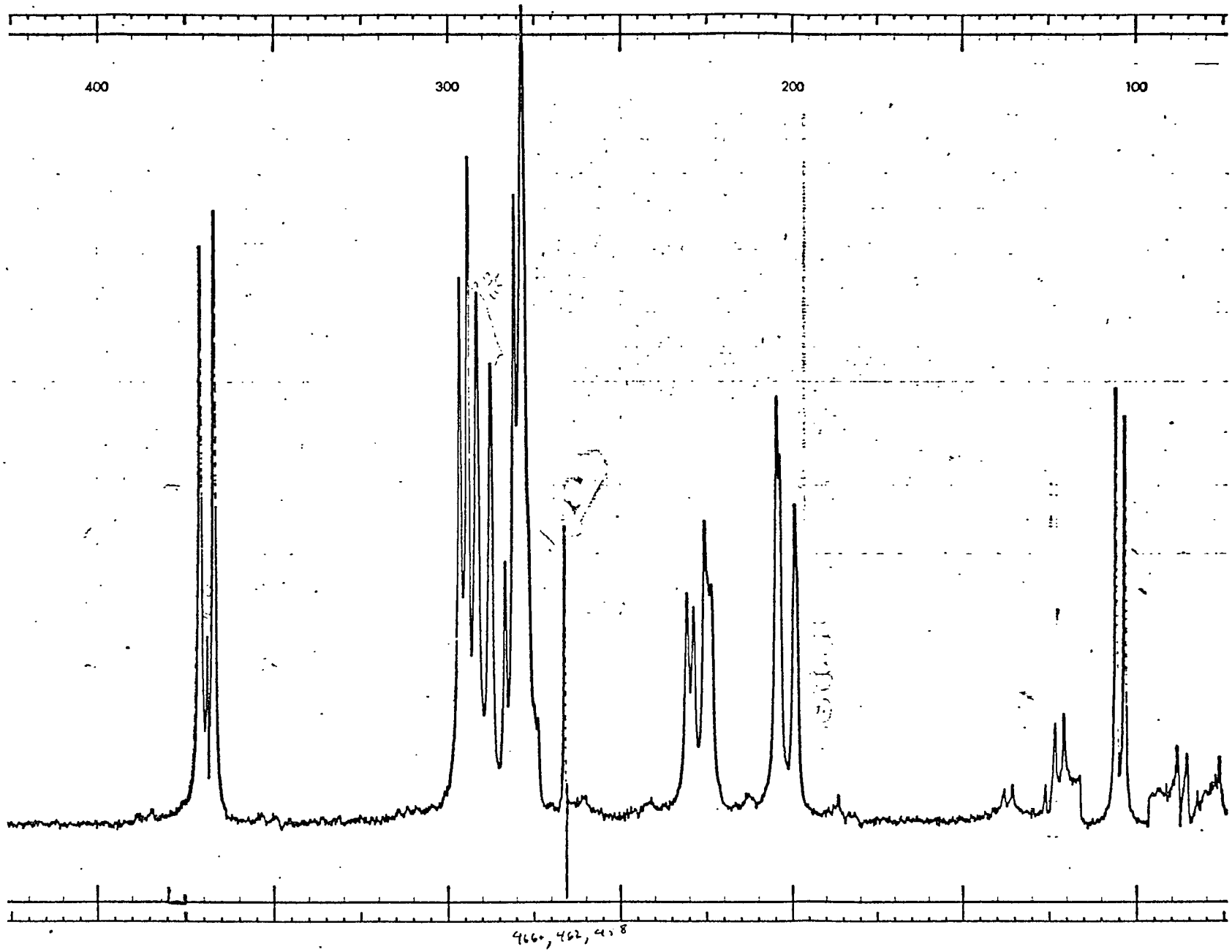

400
300
200
100
466, 462, 458

FIGURE 17

NMR SPECTRUM OF XI AND XII IN $CDCl_3$ AT

100 MHz, 1.9 to 5.9δ

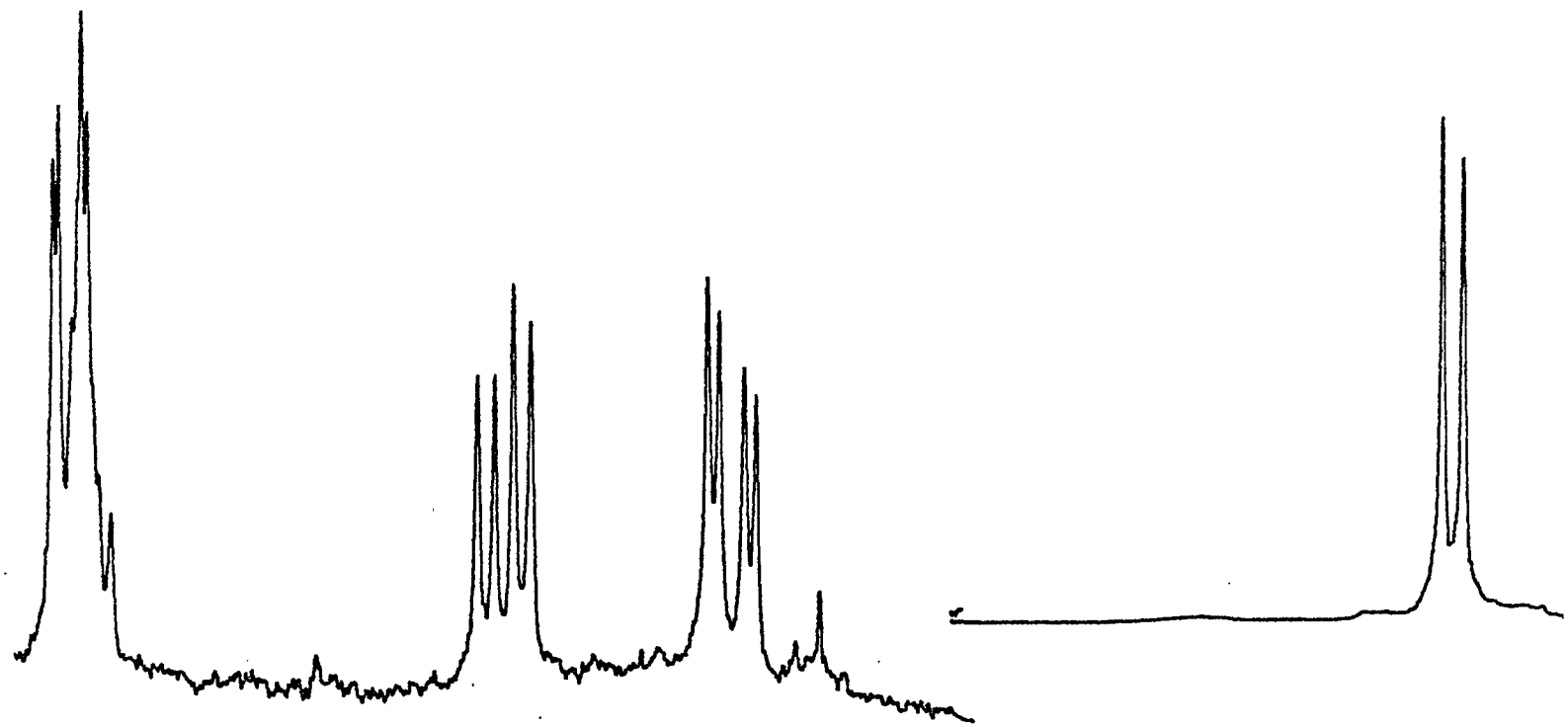
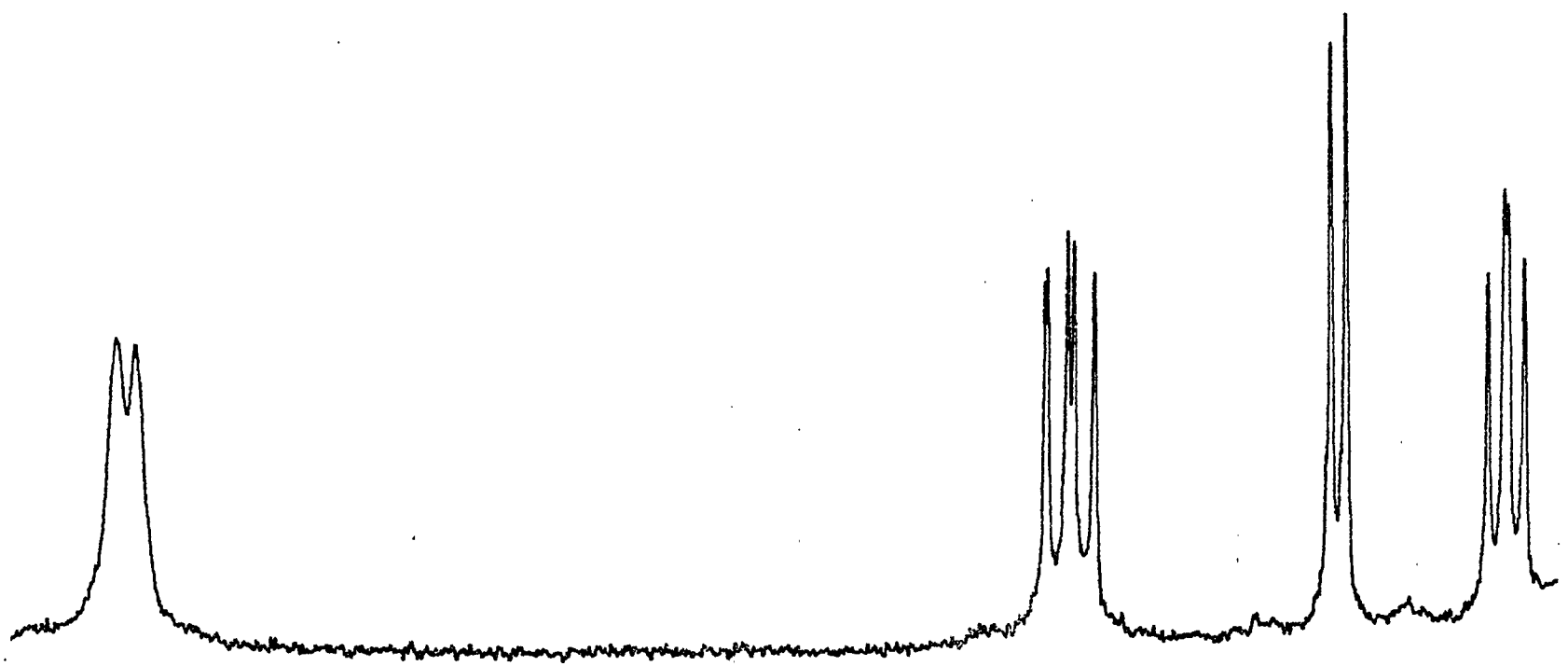


FIGURE 18

NMR SPECTRUM OF XI AND XII IN $CDCl_3$ AT

100 MHz, 5.9 to 9.9δ



and 10 Hz), 2.14 (s, 6 H, CH_3), 2.08 (s, 6 H, CH_3). Further elution afforded a trace of X.

1-(β-D-Erythrofuranosyl)-uracil(X). Method II.

Best results were obtianed by the following method. Compound VIII was prepared by the stannic chloride method and separated from the other nucleoside products by chromatography as described above. The chromatography fractions containing VIII (1.51 g, syrup, evaporation at atmospheric pressure) were combined and evaporated to dryness in vacuo to yield a foam (1.48 g). The foam was dissolved in one-tenth saturated (0°) methanolic ammonia and stored at -17°. The reaction was followed by tlc (silica gel G, 1:9 methanol-dichloromethane), which resolved the reaction into three well-separated components: the fast-moving reactant, followed by a very faint transient intermediate (presumably monoacetate), and the slow-moving product spot. One day at -17° gave an incomplete reaction, but two additional days at 5° afforded a homogeneous product. This product was evaporated to dryness in vacuo and recrystallized from ethanol to give 759 mg (71%) of a white, microcrystalline compound: mp 186.5°-187.5°; 100 MHz nmr ($DMSOd_6$) δ 7.65 (d, 1 H, H(6), J = 8 Hz), 5.75 (d. 1H, H(1'), J = 6 Hz), 5.63 (d, 1 H, H(5), J = 8 Hz), 4.05 to 4.32 (3 H, H(4a'), H(3'), and H(2'), 3.71 (m, 1 H, H(4a'), J = 2.9 Hz); identical by nmr to X prepared by Method I.

FIGURE 19

NMR SPECTRUM OF N(1)-(β-D-ERYTHROFURANOSYL)-

URACIL (X) IN D_2O AT 100 MHz

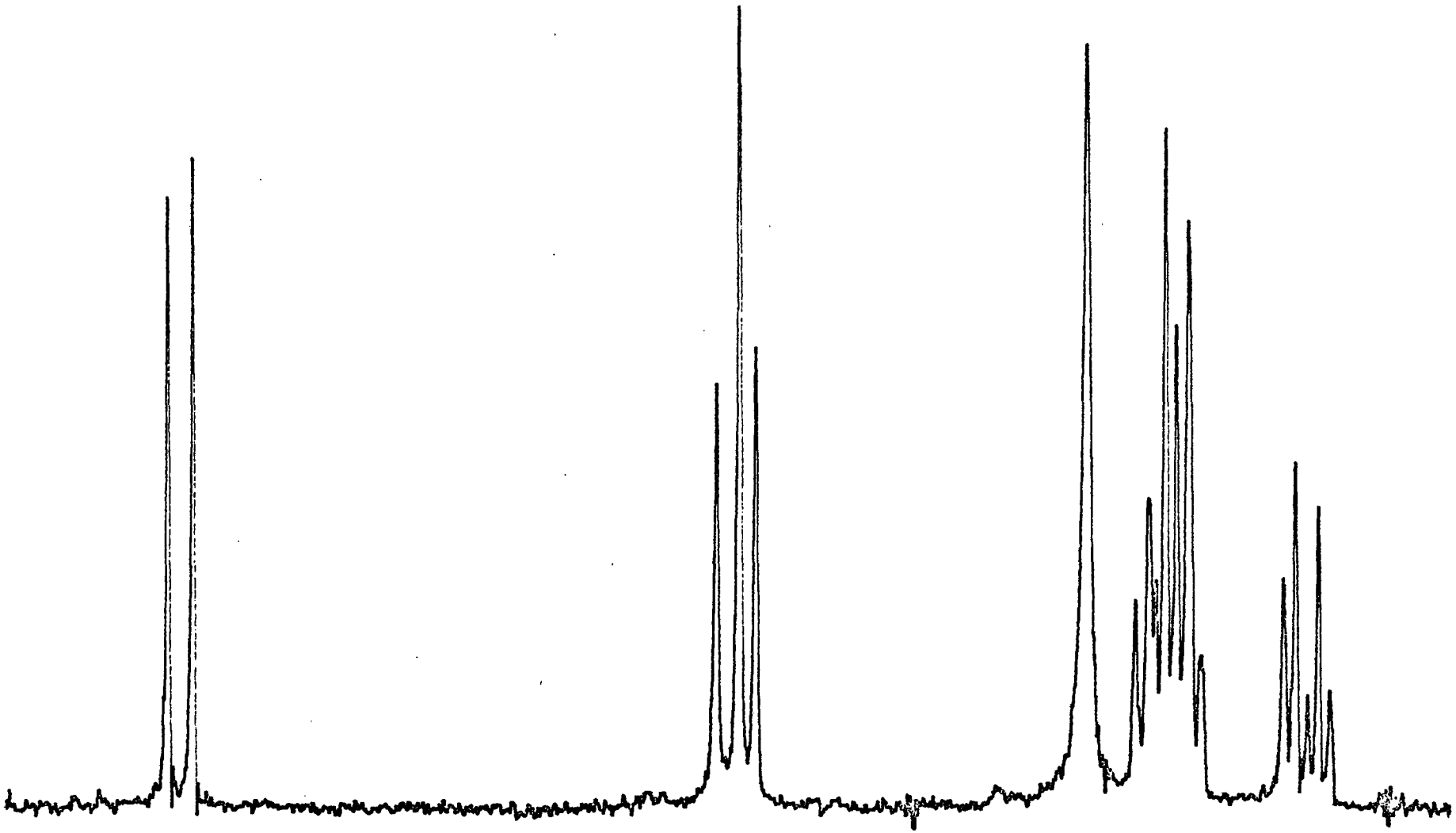


FIGURE 20

NMR SPECTRUM OF X IN D_2O AT 60 MHz

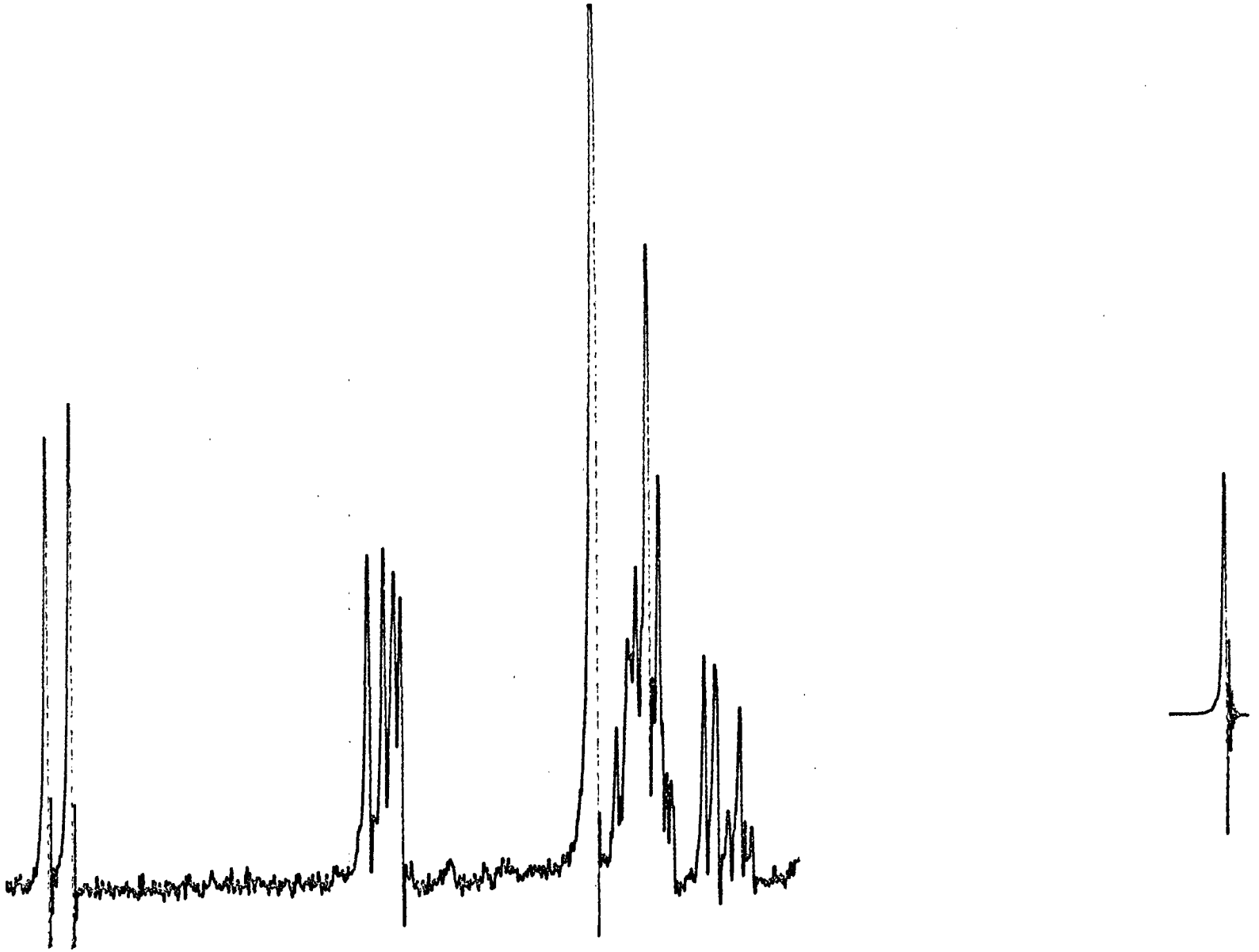
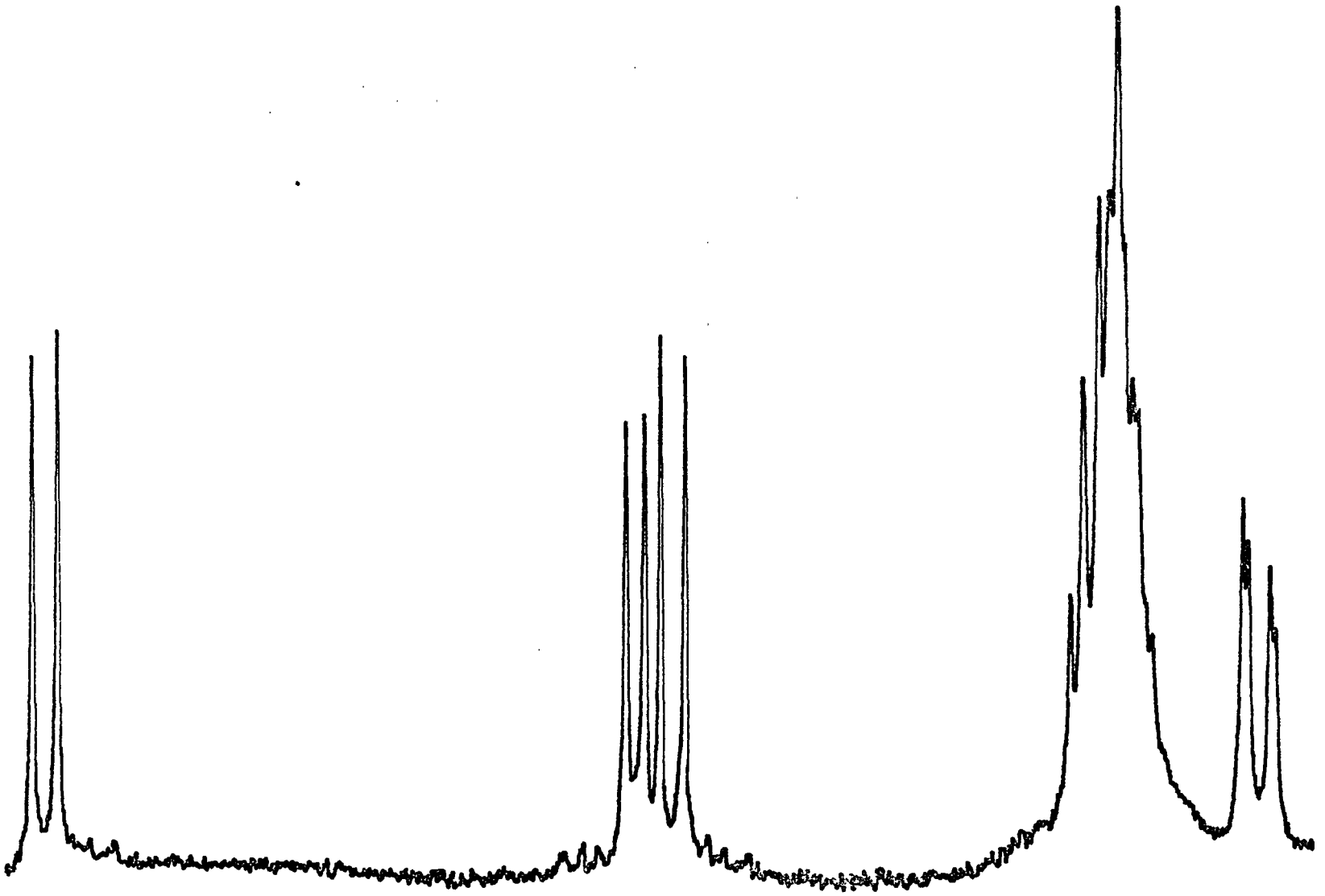


FIGURE 21

NMR SPECTRUM OF X IN DMSO-d_6 AT 100 MHz



N(4)-Acetylcytosine

To a magnetically stirred mixture of cytosine · 1/2 H_2O (4.50 g, 3.75 m mol) and dry pyridine (50 ml) in a 100 ml round bottom flask was added acetic anhydride (12 1/2 ml, 133 m mol). The two-phase reaction was stirred overnight at room temperature. The product was collected in a 5 cm buchner funnel and washed with dry pyridine (2 x 11 ml) and 100% ethanol (2 x 11 ml) to give 5.76 g (100%) of a white powder: 60 MHz nmr (CF_3COOH) is consistent with N(4)-acetylcytosine.

N(4)-Acetylcytosine Mercury.

This compound was prepared according to the procedure of Fox, Yung, Wempen, and Doerr.[78] From the N(4)acetylcytosine prepared above (3.06 g) was obtained 6.10 g (87%) of N(4)-acetylcytosine mercury: 60 MHz nmr (CF_3COOH) was similar to N(4)-acetylcytosine.

Attempted Reaction of N(4)-Acetylcytosine with VI

From a stirred mixture of mercuric cyanide (0.329 g, 1.30 m mol),and nitromethane (100 ml) was distilled 20 ml through a Vigreaux column to remove moisture. To the stirred, refluxing mixture was added VI (0.586 g, 2.76 m mol) in nitromethane dropwise over a period of 45 minutes. Solution became complete at this time, and reflux was continued for 15 minutes more. The cooled product was evaporated to dryness in vacuo,

and the residue was extracted with chloroform. The chloroform solution was washed with a 30% aqueous solution of potassium iodide and then with water, dried over anhydrous sodium sulfate, filtered, and evaporated to dryness in vacuo to yield a syrup. No nucleoside absorptions were detected by nmr.

Attempted Reaction of N(4)-Acetylcytosine Mercury with VI

To a magnetically stirred refluxing N(4)-acetylcytosine mercury (120 mg, 0.399 m mol) and toluene, dried by azeotropic distillation, was added a solution of 2,3-di-O-acetyl-β-D-erythrofuranosyl chloride (84 mg, .379 m mol) in toluene dropwise over a period of 30 minutes. Reflux was continued one hour longer, and the cooled reaction was filtered and evaporated to dryness in vacuo. The residue from this evaporation was extracted with chloroform, and the chloroform solution was filtered and evaporated to dryness in vacuo to give a syrup. This syrup was analyzed by 60 MHz nmr ($CDCl_3$), and no nucleoside was detected.

Attempted Reaction of N(4)-Acetylcytosine Mercury with D-Erythrose Triacetate(V) in the Presence of Titanium Tetrachloride

This preparation was attempted by the method of Murray and Prokop[10] which was successful when N(6)-benzoyl-

adenine mercurichloride was used. A mixture of N(4)-acetylcytosine mercury (2.20 g, 6.21 m mol), molecularly distilled D-erythrose triacetate (1.5206 g, 6.17 m mol), celite (2.95 g) and 1,2-dichloroethane (116 ml) was dried by azeotropic distillation of 22 ml of liquid. Titanium tetrachloride was distilled into a graduated addition funnel fitted to the distillation apparatus used in the above azeotropic distillation. (Caution: titanium tetrachloride reacts violently with atmospheric moisture, producing hydrogen chloride gas). Old bottles of titanium tetrachloride may contain hydrogen chloride under pressure, and the contents may be sealed by decomposition product and under pressure). The distilled titanium tetrachloride (0.8 ml, 9 m mol) was added to the azeotropically dried mixture, and the liquid phase became a very pale yellow color. The reaction was refluxed for 16 hr. Saturated aqueous sodium bicarbonate (6.4 ml) was stirred with the reaction, and the mixture was brought to neutrality with solid sodium bicarbonate. The product was dried over anhydrous sodium sulfate, filtered through celite, and evaporated to dryness in vacuo to give 644 mg of a black glass. This material was analyzed by 100 MHz nmr, which showed some unreacted starting material, but no nucleoside product. The physical appearance of the product was similar to that of products of unsuccessful attempts to prepare 2,3-di-O-acetyl-β-D-erythrofuranosyl chloride using titanium tetrachloride,

at reflux temperatures, and to decomposed samples of authentic 2,3-di-O-acetyl-β-D-erythrofuranosyl chloride.

N(1)-β-D-Erythrofuranosyluracil-2'(3')-phosphite(XIV)

To a solution of N(1)-β-D-erythrofuranosyluracil (107 mg, 0.500 m mol) in DMF (0.9 ml, 0.856 g) in a 6 ml vial (kept closed) was added trimethylphosphite (192.6 mg, 1.55 m mol) and trifluoroacetic acid (186 mg, 1.63 m mol). The reaction was followed by tlc (silica gel G, 1:1 methanol-chloroform). Two spots were seen: the fast moving reactant spot and the slower moving product spot. After 2 1/2 hr at room temperature, at which time a small amount of reactant was still present, 1 1/2 ml of concentrated ammonium hydroxide-water (1:4) was added, whereupon the reaction heated up and a precipitate dissolved. The reaction was transferred to a 50 ml round bottom flask, and another 11 ml of concentrated ammonium hydroxide-water (1:4) was added. The reaction was stored at room temperature for 2 1/2 hr, extracted with ether (3 x 10 ml), and evaporated to dryness in vacuo to give 539 mg of a syrup. This syrup was chromatographed on silica gel, eluting with methanol-chloroform (1:1) to give 103 mg (70%) of N(1)-β-D-erythrofuranosyl-2'(3')-phosphite, ammonium salt as a syrup. This compound was not analyzed, but was subjected to the following preparation.

N(1)-(β-D-Erythrofuranosyl)-uracil-2',3'-cyclic Monophosphate (XV)

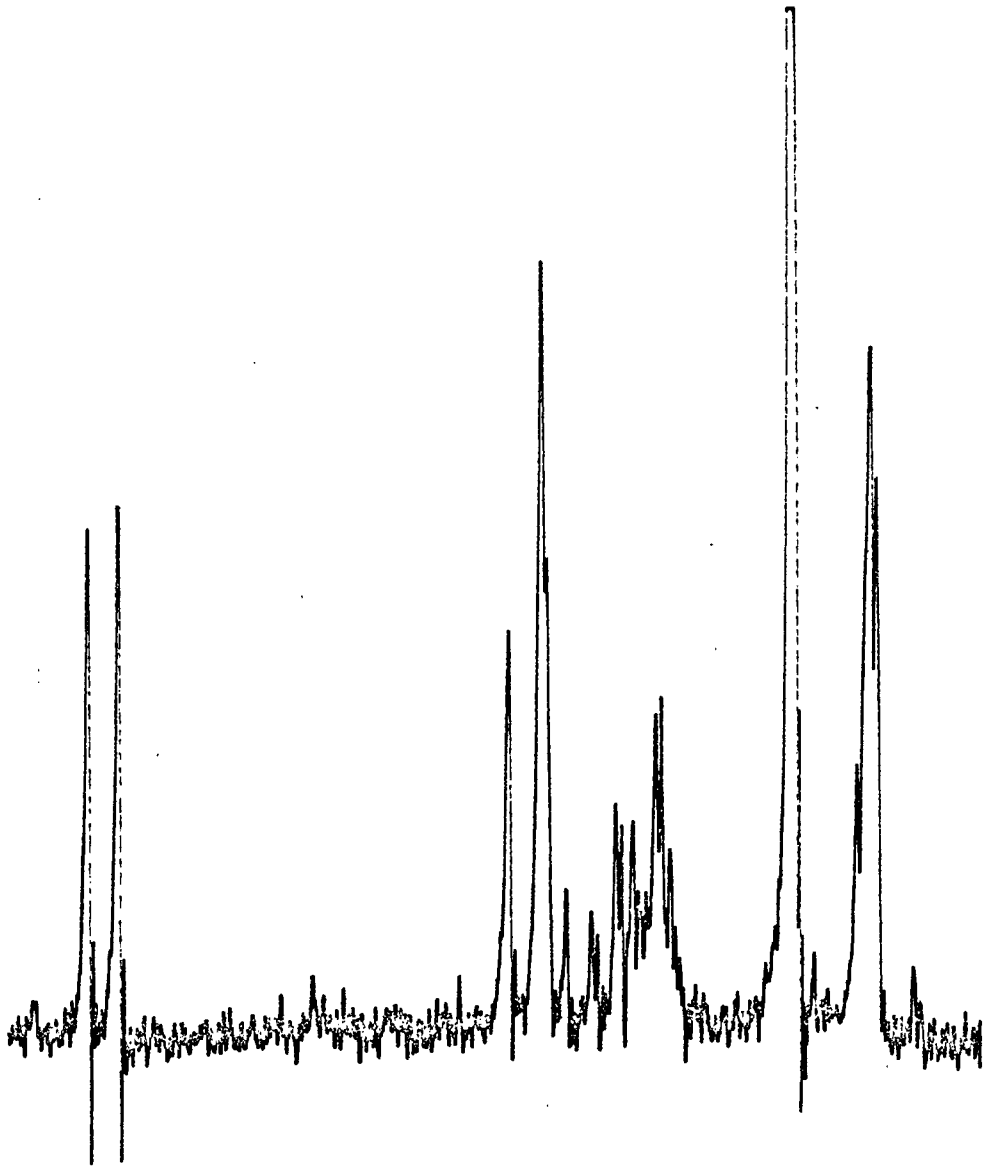
The ammonium salt of XIV (103 mg, 0.349 m mol) was treated with 1:9 triethylamine-absolute ethanol (70 ml) and heated 3 minutes in a steam bath; not quite all of the crude chromatography sample dissolved. The mixture was evaporated to dryness in vacuo, and the residue was evaporated once again with 1:9 triethylamine-ethanol and then with absolute ethanol. This triethylammonium salt was treated with acetonitrile (10 ml), hexachloroacetone (3.4 ml, 13.6 m mol), and pyridine (1.8 ml, 22.7 m mol). A side reaction between hexachloroacetone and pyridine was obvious. The solution was stirred for 8 hr at room temperature and then stored at 5° overnight. To this dark colored solution was added 1:4 concentrated ammonium hydroxide-water (50 ml), and the mixture was stirred for 2 hr at room temperature. The product was extracted with ether (3 x 25 ml) and evaporated to give over 3 g of residue. Chromatography of the entire crude product on silica gel (50 g) with 1:1 methanol chloroform afforded 0.324 g of partially purified nucleotide. The later chromatography fractions which contained nucleotide were combined to give 88 mg of the ammonium salt of 19 as an amorphous glass, pure enough for analysis by nmr spectroscopy: 60 MHz nmr (D_2O, t-Bu OD reference) δ 7.7 (d, H(6), J = 8 Hz), 6.9 (d, H(5), J = 8 Hz), 6.85 (s, 1 H, H(1')), 5.0-5.6 (m, H(2') and H(3')), 4.4 (m, H(4b') and H(4a')).

FIGURE 22

NMR SPECTRUM OF N(1)-(β-D-ERYTHROFURANOSYL)-

URACIL-2',3'-CYCLIC PHOSPHATE (XV)

IN D_2O AT 60 MHz



The 324 mg of partially purified ammonium salt of XV was rechromatographed on silica gel (30 g) with 47:53 methanol-chloroform to give 47 mg (46%) of a white glass.

Reaction of XI and XII with 0.2 N Sodium Hydroxide

Approximately 100 mg of the acetylated O-nucleotides obtained as a side product from the stannic chloride catalyzed reaction, evaporated to a foam from solution in a 100 ml round bottom flask (vacuum pump), was dissolved in 50 ml of 0.2 N sodium hydroxide. The acetate groups rapidly hydrolyzed to give a product with an Rf of 0.15 by tlc (silica gel G, 1:9 methanol-dichloromethane), presumably the mixture of deacetylated O-nucleosides. A solution of uracil in 0.2 N sodium hydroxide was also analyzed by tlc (silica gel G, 1:9 methanol-dichloromethane), and the Rf was 0.3. The basic solution of O-nucleosides was refluxed for 10 minutes, after which there was no change in the tlc chromatogram. Glucopyranosyl O-nucleosides are hydrolyzed to uracil under these conditions.[79]

Reaction of XI and XII with 0.1 N HCl

Approximately 100 mg of the acetylated O-nucleosides from the same sample used in the above experiment was dissolved in 50 ml of 0.1 N hydrochloric acid. The solution was neutralized with 0.2 N sodium hydroxide and analyzed by tlc

(silica gel G, 1:9 methanol-dichloromethane). The product had an Rf of 0.15, consistent with the deacetylated O-nucleosides above. Glucopyranosyl O-nucleosides are hydrolyzed to uracil under these conditions.[79]

Test for Activity Against Leukemia L1210 in Mice

Nineteen BDF1 female mice were injected with 1,000,000 leukemia L1210 cells each through the peritoneum using 22 gauge needles while under anaesthesia (day 1). The leukemic mice were divided into two groups of seven and one group of five. The group of five was injected with N(1)-(β-D-erythrosyl)-uracil once a day (500 mg/kg, previously determined to be non-toxic) for nine days beginning on day two. One group of seven was injected with arabinosyl cytosine (50 mg/kg for 6 days, 25 mg/kg for 3 days), beginning with day two. All drug injections were done with 26 gauge needles through the peritoneum. The third group was used as a control. No activity against leukemia was exhibited by N(1)-(β-D-erythrosyl)-uracil; the median survival time was approximately the same as the controls (ten days). All the mice treated with arabinosyl cytosine were quite healthy after all the other implanted mice had died. Median survival time of these mice was 17 days.

CHAPTER III

RESULTS AND DISCUSSION

Results and Discussion

The preparation of D-erythrose by the route presented in this dissertation is rather difficult. The most time-consuming step in the sequence of reactions is preparation of 4,6-O-ethylidene-D-glucopyranose (II). Fortunately, it can be prepared in large quantities and the product crystallizes readily to give a pure, relatively stable compound. There were still some problems encountered, however, and we have some helpful amendments to the literature preparation. The procedure followed in this laboratory is an adaptation of the very detailed literature preparation of Hockett, Collins, and Scattergood[71] in which we use a mechanical stirrer instead of a mechanical shaker. A considerable savings in time resulted; the literature reaction time was 48 to 96 hr, while the reaction time in this laboratory was only 12 to 36 hr. Since we did not run the reaction using a shaker, however, we cannot say with certainty that the shorter reaction time is attributable to the change in the method of stirring. It could have been some other factor such as particle size. Several preparations of II resulted in a recrystallized product of melting point below 150°, compared to 177°-179° for the pure compound. It was eventually discovered that recrystallization of this material from weakly alkaline ethanol, but not neutral ethanol, afforded a material with the proper melting point and removed some impurities from the nmr spectrum.

The nmr spectra of II in both DMSO-d_6 and D_2O showed the presence of both α- and β-anomers. The predominant anomer in the spectra had an H(1') absorption upfield from the other, as would be expected for the β-anomer because of the expected axial orientation of H(1')[80] and the diamagnetic anisotropy of the oxygen atom at the 2-position.[81] In addition, the coupling constant of the predominant anomer was large enough to be β(trans-diaxial), but the coupling constant of the minor component was too small to be a trans-diaxial coupling constant on a 6 membered ring.[82] Mutarotation may have taken place under conditions employed in nmr spectroscopy of II, so the anomeric mixture seen by nmr does not prove that the crystalline material is also a mixture.

The step in the sequence which limits the scale of the preparation is the sodium metaperiodate oxidation of II. The product, compound III, is unstable to heat and also unstable in the presence of the side products in the reaction. For this reason the aqueous reaction solution must be evaporated quickly at temperatures below 30°. In our laboratory these considerations limited the scale of the reaction to about 0.1 mol. A large scale freeze-drying apparatus would greatly increase this upper limit. The dried residue must be immediately extracted with ethyl acetate, and the ethyl acetate extracts should be thoroughly freed from volatile impurities before hydrolysis to compound IV is attempted.

The hydrolysis of III to form D-erythrose is also complicated by decomposition. When quantity was more important than purity, the crude form was used. Most literature procedures used reflux temperatures in the hydrolysis.[74,75,76] This often gave an impure yellow product in our hands. It was much easier to obtain a pure, colorless product using our method. D-Erythrose is also slowly decomposed by heat, and evaporations were done at room temperature. The strong base resin was used because of convenience; no 4-carbon sugar isomerization product was ever noticed in the distilled triacetate, however. For maximum purity, crystalline III should be used in the hydrolysis, and the product should be neutralized by a column of weak base and weak acid ion exchange resins.

D-Erythrose triacetate (V) was prepared without difficulty according to the literature procedure,[10] except for the correction on the reaction time. Since carbohydrate impurities contaminated chromatography fractions containing acetylated nucleosides when the stannic chloride preparation was used, it was often desirable to purify V. This could be done in a sublimation apparatus when 4° coolant was used, or in a molecular still when tap water was used as the coolant.

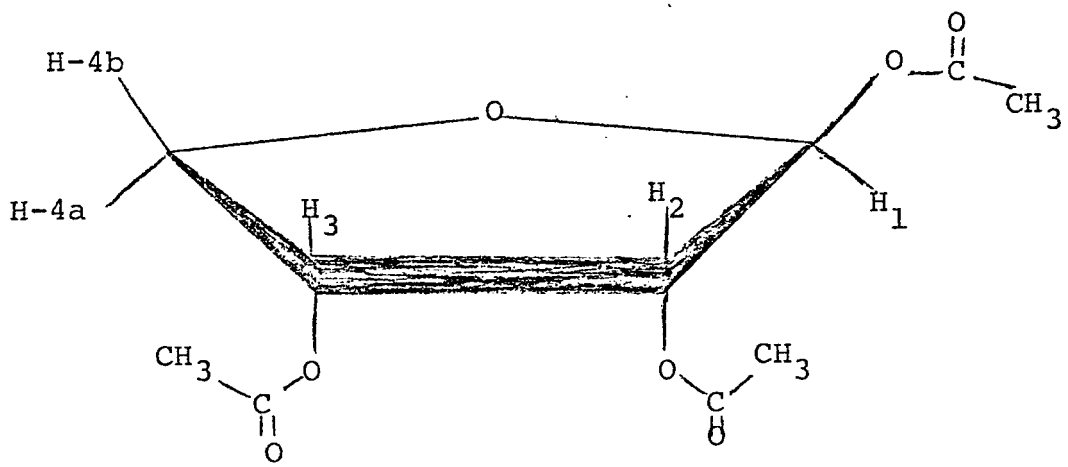
The nmr spectrum of V in carbon tetrachloride at 100 MHz was completely interpretable by first order rules.

There were two anomeric protons present in the spectrum, which is expected since in general both anomers are present in sugars acetylated in pyridine. The predominant anomer in both the crude and the molecularly distilled product was determined to have the trans-1,2-configuration on the basis of chemical shift and coupling constant. The H(1) absorption of this β-anomer is upfield from the H(1) of the minor component because of diamagnetic shielding from the non-bonded electrons of the neighboring oxygen atom at the 2-position.[83] The minor componet has a $J_{1,2}$ coupling constant in the range which could belong to either cis or trans protons. The more upfield part of the AB pair between 4 and 5 δ was tentatively assigned to the H(4)which is cis to the acetoxy group at C(3), on the basis of evidence provided for H(4) protons in the tables of Stevens and Fletcher.[83]

The assignments of H(2) and H(3) absorptions of the β-anomer in the nmr spectrum of V were made on the basis of coupling constants and confirmed by portionwise addition of $Eu(DPM)_3$ to the solution. Upon addition of the first portion of paramagnetic shift reagent, the H(1) absorption was moved well downfield with respect to the other absorptions in the spectrum. If the paramagnetic shift reagent coordinates nearest to the 1-position primarily in this rapidly pseudo-rotating molecule, then H(2) should move downfield faster than H(3). Accordingly, upon further addition of $Eu(DPM)_3$, the

H(2) and H(3) absorptions moved together and finally overlapped, at which point the neat AB pair formed by the methylene proton became two second order multiplets. All peaks moved downfield from TMS; H(1) moved fastest, followed in order by H(2) > H(3) > H(4b), H(4a) > $-O-\overset{O}{\overset{\|}{C}}-CH_3$. Of the acetyl groups, one [presumably the C(1) substituent] moved downfield a littler faster than the other two. The results of this experiment were taken to be evidence that the assignments are correct.

The methylene group forms an AB pair in the nmr spectrum of the β-anomer of V. Presumably the upfield H(4) proton, designated H(4a), is cis to the oxygen atom at C(3). Shielding from this oxygen should cause the proton on C(4) which is cis to it to absorb upfield from the other. It is worthwhile to note that H(4b) moves downfield slightly faster than H(4a).



D-Erythrose triacetate

The interpretation of the paramagnetic shift reagent experiment with V is not as straightforward as it was assumed to be in the above discussion. There are several places on the molecule where $Eu(DPM)_3$ could coordinate. The results indicate that it coordinates primarily to the acetyl group at the C(1) position. However, it is not unlikely that it also coordinates to the other two acetyl groups; if $Eu(DPM)_3$ preferentially coordinates to the O(3)-acetyl group more than the O(2)-acetyl group), it may or may not be the predominant factor in determining whether or not H(3) moves downfield faster than the H(2). However, there is no reason to presume that the O(3)-acetyl group would have a significantly greater affinity for the paramagnetic shift reagent.

The largest flaw in the preliminary interpretation is that the methyl groups are affected least of all. If the downfield shifts may be explained by the simplified model in which $Eu(DPM)_3$ coordinates to the O(1)-acetyl group exclusively, then the methyl group in that acetyl group should move downfield faster than it does; it should certainly be more greatly affected than H(4a). The answer may lie in an angular dependence factor in the paramagnetic shift phenomenon, but a more likely explanation is that the reagent coordinates appreciably to all three acetyl groups. Since H(4b) is about as far removed from the O(1)-acetyl group as H(3), it should move downfield about as rapidly; yet it moves much more

slowly. One answer to all these observations is that the fast moving protons are separated from the site of coordination by only one atom, and that $Eu(DPM)_3$ coordinates appreciably to the ether oxygen (or to the keto oxygen but on the ether oxygen side) of all three acetyl groups (with the greatest affinity for the O(1)-acetyl group). Since H(2) is near the O(2)-acetyl group and two additional sites of interaction and nearer to O(1), it would still be expected to move downfield faster than H(3). That H(4b) moves faster than H(4a) could be ascribed to the unpredictable result that greater affinity of the more distant O(1)-acetyl group outweighs greater proximity of the O(3)-acetyl group. That the methylene protons move a little faster than the methyl groups may reflect some affinity of the ring oxygen atom for $Eu(DPM)_3$.

The preparation of N-acetylcytosine mercury according to the literature method was straightforward. The α- and β-adenosine erythrosides were prepared by Murray and Prokop by the reaction of D-erythrose triacetate with N(6)-benzoyladenine mercurichloride in the presence of titanium tetrachloride. We ran the preparation under identical conditions using N-acetylcytosine mercury and met with no success. Presumably the cytosine mercury compound is less reactive than the adenosine mercury compound, and the intermediate chloride or dioxolenium ion formed from the sugar is unstable and decomposes rather than reacting with N-acetylcytosine mercury.

Alternatively, there may be an initial O-nucleoside formed with this base, and the unstable sugar is cleaved rather than undergoing rearrangement.

Preparation of 2,3-di-O-acetylerythrofuranosyl chloride was first attempted by the titanium tetrachloride method. No acylglycosyl chloride was observed in the nmr of the product of the reaction, which was attempted a number of times, at conditions as mild as room temperature. It is not unusual for this reaction to be unsuccessful for preparation of unstable acylglycosyl halides which can be prepared under milder conditions, such as the classical method of A. R. Todd using hydrogen chloride in ethyl ether.[85] Accordingly, D-erythrose triacetate was treated with anhydrous hydrogen chloride in dichloromethane at -17°, and moisture was avoided in the workup. This resulted in a good yield of an oil which was of good purity by nmr, when moisture was carefully excluded. The product was unstable, and was rapidly decomposed to a black tar in the presence of water. This contrasts to 2,3-O-isopropylidene-L-erythrofuranosyl chloride, which was prepared by Lerner.[86] The latter compound was reported to be very stable and gave a rather low yield of nucleoside in a reaction with N(6)-benzoyladenosine.

The 60 MHz nmr spectrum of 2,3-di-O-acetylerythrofuranosyl chloride (VI) is somewhat similar to that of β-D-erythrose triacetate. Only one anomer is present in the product.

This is not unusual; one anomeric form of acylglycosyl halides is generally somewhat much more stable than the other, and the compounds readily anomerize by way of a unimolecular isomerization mechanism so that only one anomer is present in the product.[87] The H(1) absorption of VI is a singlet, and this suggests a trans-1,2-configuration. The conclusion that VI is a β-anomer is consistent with the results of the Hilbert-Johnson reaction, and this will be discussed later. Interestingly, the J(1,2) coupling constant of 2,3-di-O-isopropylidene-L-erythrofuranosyl chloride is also in the range which strongly suggests trans-1,2-configuration.

The first reactions attempted with the chloride VI were the highly stereoselective (trans-1,2) reactions involving mercury salts. Thus, VI was treated with N-acetylcytosine in boiling nitromethane in the presence of mercuric cyanide, and also with N-acetylcytosine mercury in boiling toluene. Both reactions failed to give a high enough yield of nucleoside to be detectable by nmr. Presumably this was due to instability of VI or the dioxolenium ion formed from VI, just as in the titanium tetrachloride-acetylcytosine mercury reaction. Perhaps the reaction would work at milder temperatures suing an unusual solvent such as N,N-dimethylformamide to dissolve the mercury compound. Success with the trimethylsilylated pyrimidine base led us away from further experiments with mercury compounds at this time, however.

When 2,3-di-O-acetyl-β-D-erythrofuranosyl chloride was reacted with 2,4-bis(trimethylsilyloxy)pyrimidine without solvent, both α- and β- anomers of 2,3-di-O-acetylerythrofuranosyl uracil VII and VIII resulted. The α-anomer predominated, in varying ratios. The reaction went at temperatures as low as room temperature, contrasting to usual temperatures of 150° to 200° for this modification of the Hilbert-Johnson reaction. Under more typical conditions (185°, 30 minutes), much dark colored decomposition products resulted. The usual conditions used for synthesis of VII and VIII by the method was 60° to 80° for one to two days, and the α:β ratio (estimated by nmr integral) was typically about 3:1. This is not the first time a 2-acyloxy acylglycosyl halide has given predominately the α-anomer. The Hilbert-Johnson reaction is not as stereoselective as mercury salt reactions with this type of sugar derivative, and predominantly cis-1,2-nucleosides often result from trans-1,2-halides.

The nucleoside products of the reaction were tentatively identified from the nmr spectra of unseparated product mixtures, which fortunately consisted of varying ratios of the two nucleosides. Chromatography on silica gel, on first trial, afforded a purified but unseparated acyl nucleoside mixture, and was not further developed. There were two absorptions in the H(6) region of uracil nucleosides at both 60 and 100 MHz. The H(1') of one nucleoside was downfield

from the other nucleoside and the H(5), H(2'), and H(3') of both nucleosides were conglomerated. This downfield H(1') was from the predominant anomer and tentatively was presumed to be the α-anomer, on the basis of the assumption that diamagnetic shielding effect of O(2') resulted in H(1β) being upfield from H(1α).[83] This assignment could not be verified by coupling constant, since both were in the intermediate range not assignable to either cis or trans geometry.[83] The possibility that one of these was an O-nucleoside or an N(3)-glycoside was tentatively eliminated on the basis that a brief search of the literature revealed no such product ever reported from a trimethylsilylated pyrimidine base, and that mixtures of α- and β- anomers were frequently reported. Probably the steric hindrance of the trimethylsilyl substituent accounts for the lack of any reported O-nucleosides produced from them.

The mixture of acetylated nucleosides VII and VIII from the Hilbert-Johnson reaction were hydrolyzed together in methanolic ammonia, to give N(1)-(β-D-erythrofuranosyl)-uracil (IX) and N(1)-(β-D-erythrofuranosyl)-uracil (X). Compounds XI and X were correlated to their corresponding precursors VII and VIII by both product ratio and similarities in the nmr spectra. Compound IX readily crystallized from the mixture, and compound X was readily separated from remaining IX by chromatography and also crystallized easily. The final

yield of the desired β-pyrimidine nucleoside was disappointingly low, attributed to instability of VI and to the predominance of the α-anomer. Experimentation with 2,4-diethoxypyrimidine in polar solvents may improve the ratio of the β-anomer, but it is doubtful if any advantage in total yield would be gained over the reaction discussed next.

The recently introduced reaction of polyacylglycosides with trimethylsilylated pyrimidine nucleoside bases in the presence of stannic chloride at room temperatures reportedly gives excellent yields of N(1)-glycosides, and gives only trans-1,2-products when 2-acyloxy sugars are used. When the reaction was run with D-erythrose triacetate and 2,4-bis(trimethylsilyloxy)pyrimidine, a mixture of at least four acetylated nucleosides resulted. The predominant product, isolated in yields of up to 56% by column chromatography, was identical by nmr to the minor product of the above Hilbert-Johnson reaction. This reaction, then, furnishes a satisfactory entry into the erythrosyl pyrimidine nucleoside analogs. The minor products of the reaction are the O(2)-and O(4)-nucleosides, and the N(1), O(4)-bis(erythroside). The O(2)- and O(4) nucleosides were formed in higher proportion when the amount of base was increased. The N(1), O(4)-bis-glycoside was formed in higher proportion when less base was used. The structural assignments are discussed in the following paragraphs.

The predominant product of the stannic chloride reaction under all reaction conditions used had an identical nmr spectrum to the minor product of the Hilbert-Johnson reaction, and the spectra and melting point of their hydrolysis products were also identical. None of the major product of the Hilbert-Johnson reaction was ever detected in the product of the stannic chloride reaction, nor has a cis-1,2-nucleoside ever been produced from a 2-acyloxy acylglycoside using this reaction. This is excellent evidence that the identification of the products of the Hilbert-Johnson reaction was correctly made.

One product isolated in the chromatography of the stannic chloride reaction product was shown to be a mixture of two similar compounds by nmr. Two N-H, H(6), and H(5) absorptions were clearly visible, but all the sugar absorptions were degenerate. On the basis of the similarity of the spectra of the sugar moieties of the two compounds, and of the fact that none of the α-isomer VII was produced by the reaction, these chromatographically homogeneous compounds were judged to be the β-O(2)- and O(4)-nucleosides XI and XII. The chemical shift of H(1') of each of these O-nucleosides is the farthest downfield of any monoglycoside of this series prepared in this laboratory. The O-nucleosides XI and XII were unusually stable, and could not be hydrolyzed under conditions which rapidly hydrolyze other O-nucleosides.[79]

The other principal minor product was the fastest moving on silica gel, and was shown by nmr integration to be a bis(glycosyl)uracil. Examination of the absorptions of the erythroside moieties revealed that the spectra of the sugar moieties could be duplicated by superimposing the spectrum of VIII on that of XI or XII. Therefore, this must be an N,O-bisglycosyl uracil. The possibility of an N(3),O-bisglycoside was ruled out because the N(3)-nucleoside was not among the principal minor products, and it is doubtful that the steric hindrance at the N(3)-position is much less in the 2-glycosyloxy-4-trimethylsilyloxy pyrimidine than in 2,4-bis(trimethylsilyloxy)-pyrimidine. This leaves N(1),O(4)-bis(β-D-erythrofuranosyl)-uracil (XIII) as the probable structure of this bis(glycoside).

Compounds XI, XII and XIII are the first nucleosides other than N(1)-glycosides reported from the stannic chloride reaction as well as from a trimethylsilylated pyrimidine nucleoside base. Their production here might be attributed to the low steric hindrance of the dioxolenium ion formed from V to trans-1,2 approach, due to lack of a bulky substituent on that side of the ring. Also the high instability of this dioxolenium ion might be expected to result in decreased selectivity for attack at the N(1)-position.

Our efforts to find an entry into the pyrimidine erythroside series have resulted in much data that may help to shed light on the mechanisms of the various glycosidation reactions studied in the process. The mechanisms will be discussed here in terms of an S_N1 reaction in which the ionized acylglycosyl species may react in either or both of two ways; firstly as a closely associated ion pair, and secondly as a fully dissociated dioxolenium cation. There is little doubt about the participation of the latter species in many of the above glycosidation reactions; it is the only reasonable explanation for stereospecific formation of trans-1,2-nucleosides from trans-1,2-acylglycoside species by certain reactions which give an anomeric mixture when non-participating protecting groups are used. Varying participation of closely associated ion pairs is the best explanation of variation of anomeric product ratios of the Hilbert-Johnson reaction using alkyloxy pyrimidines when solvent polarity is varied.

The reactions involving mercury salts failed under temperature conditions in which the Hilbert-Johnson reaction gave fairly good total yields of nucleosides. The reacting species in all of the mercury salt reactions is probably the fully dissociated glycoside, which becomes a dioxolenium cation in the case of a compound like VI; that this is so is

indicated by the stereospecificity of the reaction for the trans-1,2-product when 2-acyloxy glycosides are used. The function of the mercury cation is to pull off the halide, causing complete dissociation so that the dioxolenium ion forms. That the dioxolenium cation may be less stable than the ion pair is suggested, but not proven, by the lower selectivity of the stannic chloride reaction (in which O-nucleosides are produced) compared to the Hilbert-Johnson reaction with the same heterocycle. Failure of the dioxolenium cation to react with the heterocycle at the temperatures involved in the mercuri and mercuric cyanide reactions may probably be attributed to rapid decomposition of the unstable species under these conditions. Conversely, the same dioxolenium cation produced in the stannic chloride reaction is stable enough at room temperature to survive long enough to react with 2,4-bis(trimethylsilyloxy)pyrimidine and form the trans-1,2-products.

One factor operating against the mercuri reactions is the low solubility of the mercury compounds. The failure of the titanium tetrachloride-acetylcytosine mercury reaction to produce a nucleoside under conditions that gave high yields with benzoyladenosine mercury may be at least partly due to lower solubility of acetylcytosine mercury. If this is the case, perhaps a naphthoyl protecting group would be helpful. The failure of this reaction might

also be due to lower reactivity of acetylcytosine mercury or to a side reaction of the heterocycle with titanium tetrachloride. Finally, acetylcytosine mercury may form nucleosides through an O-nucleoside intermediate; poly-o-acyl-erythrosides, like poly-o-acyl-deoxyribosides, have an unstable ionized form, and might also be expected to decompose rather than rearrange, just as poly-o-acyl-deoxyribosides do. The possibility of a titanium tetrachloride side reaction could obviously be eliminated as a cause of failure in the acetylcytosine mercury-acetylerythrosyl chloride reaction, but the other three possible factors remain. In the mercuric cyanide reaction, the most probable explanation for failure to produce a nucleoside where benzoyladenosine mercury succeeded is lower reactivity of the heterocycle. The success of the stannic chloride reaction at room temperature suggests that the mercuri and/or mercuric cyanide reactions may be successful at room temperature in a suitable solvent, such as N,N-dimethylformamide.

The failure of the stannic chloride reaction to produce the alpha nucleoside VII produced in the Hilbert-Johnson reaction unambiguously indicates that the reacting intermediate formed from the primarily beta polyacylglycoside is the completely dissociated dioxolenium cation. On the other hand, the failure of the Hilbert-Johnson reaction

to produce any of the bis or O-nucleosides even at high temperatures indicates that the beta nucleoside VIII produced in this reaction from the beta chloride was formed from a different, less reactive or more sterically hindered intermediate glycoside species rather than from the dioxolenium cation. These observations are consistent with the hypothesis that the acylglycosyl halide in the Hilbert-Johnson reaction in this case forms a closely associated ion pair, which reacts to form both α- and β- products. The dioxolenium ion is not formed to any significant extent. The α-anomer predominates in the product because of the closely associated chloride ion, which is predominantly if not exclusively on the β- side. The possibility of a concomittant S_N2 mechanism in the Hilbert-Johnson reaction has not been excluded by this product stury, but is not necessary to explain the results.

Ulbricht has postulated that the Hilbert-Johnson reaction of a trans-1,2-halide proceeds through an S_N2 displacement of halide and a competing S_N1 reaction in which double inversion of configuration takes place through a dioxolenium cation, giving both α- and β- products.[15] Our studies with VI have given evidence that an S_N1 reaction can take place with a 2-acyloxy glycosyl halide without formation of a dioxolenium cation, and that the closely associated ion pair is the most likely reacting species

in the Hilbert-Johnson reaction of VI with 2,4-bis-(trimethylsilyloxy)pyrimidine. We therefore favor the mechanism of Sorm in which the reactions of alkyloxy- and silyloxypyrimidines are discussed in terms of completely dissociated dioxolenium ions and closely associated ion pairs.[16,17] The Hilbert-Johnson reaction using trimethylsilylated pyrimidine nucleoside bases without solvent apparently proceeds through the ion pair only, with no participation by the dioxolenium ion, since no O-nucleosides were observed when the highly reactive 2,3-di-O-acetyl-β-D-erythrofuranosyl chloride was used in the reaction.

The success of the stannic chloride method in producing the desired nucleoside in satisfactory yields, when all of the more popular procedures failed in our hands, attests to the usefulness of this excellent new reaction.

The cyclic monophosphate of β-erythrosyl uracil was easily prepared in a variation of the trialkyl phosphite-hexachloroacetone method introduced by Holy and Smrt.[88] When we followed their procedure for the hexachloroacetone oxidation step, a very large amount of solid impurity was obtained because of a reaction of hexachloroacetone with pyridine (the co-solvent), which was not mentioned by these authors. We recommend the use of N,N-dimethylformamide as the solvent for this reaction, which they have done in more recent papers.[89,90]

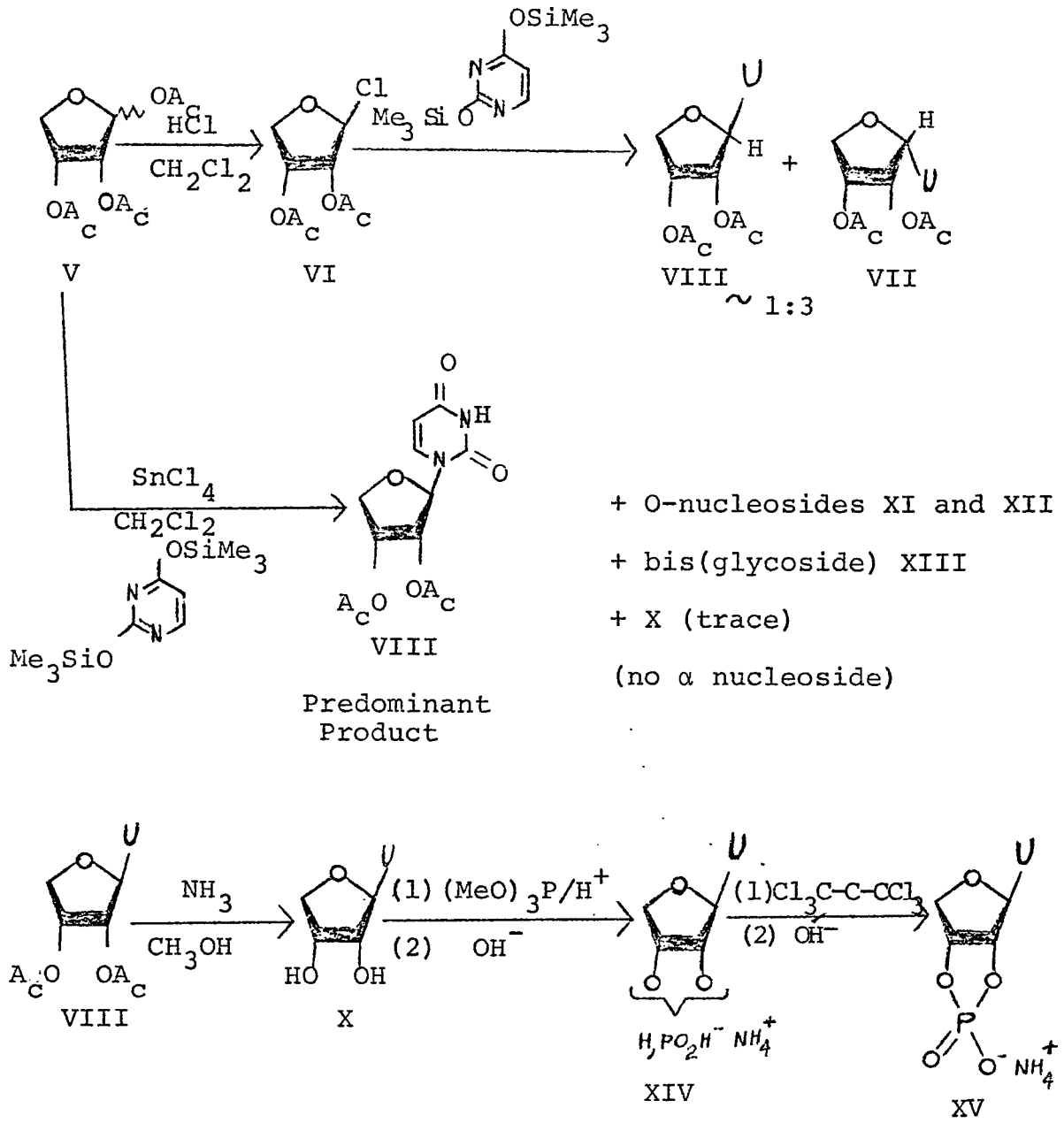


Figure 23. Schematic Summary of the Synthesis

CHAPTER IV

CONFORMATION OF FURANOSE RINGS BY NMR SPECTROSCOPY

Since study of nucleic acid conformations and nucleoside and nucleotide conformations is of current interest, we decided to analyze the conformations of the furanose rings of our erythrosyl nucleosides by nmr spectroscopy. It soon became apparent in our literature search that there was no universal agreement on the coupling constants assigned each conformation. In addition, there has been no attempt to assess the correlation of the Karplus equation and the experimentally determined coupling constants of nucleosides. We therefore undertood a thorough analysis of the method of determining conformation of furanose rings from nmr coupling constants.

Our first concern was the validity of the Karplus equation for nucleosides. Although Karplus himself cautioned against a quantitative application of this equation, since coupling constants should depend on other factors in addition to dihedral angles,[30] nucleoside chemists have repeatedly attempted to determine the conformations of their furanose rings by this method. Such faith has undoubtedly been influenced by the excellent correlation to the fixed conformation of the bicyclic vicinal diols of Anet.[91] Another early study also showed good correlation.[92] However, a recent study by Coxon of the nmr of a carbohydrate of fixed conformation showed errors of up to 24° in angles calculated from coupling constants when compared to angles measured on a molecular model.

In view of the results of this latter study, a more cautious approach would require that a range of possible coupling constants for each conformation should be tabulated, so that very unlikely conformations could be eliminated. A choice from among the remaining conformations might in some cases be made by studying the relative steric strain in these conformations, and also by considering which conformations have appeared in X-ray structures of various nucleosides.

One should also consider that a range of possible extents of puckering exists for each conformation. Most workers in the field seem to have ignored this factor. For example, Smith, Hruska, and Grey[35] considered only maximal (0.75Å) puckering, using the erroneous tables of Jardetsky (vide infra),[31] in arriving at the conclusion that uridine and β-pseudouridine are in conformational equilibrium (e.g., C(2')-endo⇄C(3')-endo) in solution. Actually, X-ray data shows most crystalline nucleosides to be puckered 0.5 to 0.6Å from the best four-atom plane.[28] If the table used by Prestregard and Chan[34,94] is used in these cases, the possibility of a preferred conformation of C(2')-endo cannot be as easily eliminated for uridine and β-pseudouridine.

The X-ray crystal structure of a bicyclic nucleoside of fixed conformation, 2',3'-isopropylidene-3,5'-cycloadenosine iodide,[95] has been published. The nmr spectrum of the very similar molecule, 2'3'-isopropylidene-3,5'-cycloguanosine

iodine, has also been published.[96] We took the opportunity to use this literature data to furnish a test for the applicability of the Karplus equation to nucleosides and carbohydrates. While the X-ray shows a C(4')-endo conformation, molecular models show an O(1')-exo conformation. This dichotomy was attributed to the unusual bond lengths and tetrahedral angles (90° to 132°) in the crystal structure of this highly strained molecule. This observation detracts from similar studies where there was no X-ray structure available. It should be noted that C(4')-endo, as well as C(1')-endo, C(1')-endo-O(1')-exo, and C(4')-endo-O-(1')-exo, conformations could be forced on the molecular model by twisting it into shape and holding it there. Hence, this procedure may furnish a test for additional possible conformations in molecular models of apparently fixed conformation.

In order to determine dihedral angles from literature X-ray data two computer programs were written in Fortran V based on the vector equation (Eq. 12):

$$\text{(Eq. 12)} \qquad \cos(\text{A-B-C-D}) = \frac{(\vec{BC} \times \vec{BD}) \cdot (\vec{BC} \times \vec{BA})}{|\vec{BC} \times \vec{BD}| \quad |\vec{BC} \times \vec{BA}|}$$

These programs were named XRAY and XRING, and are printed beginning on the next page. Program XRAY gives the dihedral angle when the x, y, z coordinates of any four atoms are known, and XRING gives all five ring dihedral angles when the x, y, z coordinates of the five atoms of any five-membered

ring are known. The input of each program consists of the unit cell coordinates (a, b, c) and the x/a, y/b, z/c coordinates x(i), y(i), and z(i) of adjacent atoms. These programs do not give the sign of the dihedral angle, but this can be determined from Newman projections easily.

It was found convenient to record conformations of saturated five-member rings in terms of "intermediate conformation" (between envelope and twist) and "total pucker." A three atom plane was chosen for each X-ray structure so that the two out-of-plane atoms (puckered atoms) were adjacent and on opposite sides of the three-atom plane. One of the puckered atoms was the atom puckered from the best four-atom plane. The sum of the perpendicular distances of the two puckered atoms from the three-atom plane was designated total pucker. The tables of the theoretical angles for cyclopentane conformations presented later show that the largest dihedral angle of the atoms of the five-membered ring (ring dihedral angles) is about the same for conformations of the same total pucker.

The displacements of the puckered atoms are easily calculated from the dihedral angles obtained from program XRING and the bond angles of ring atoms with adjacent ring atoms (ring bond angles) by the equation given below (Eq. 13). The ring bond angles are nearly always given in journal articles on crystal structures.

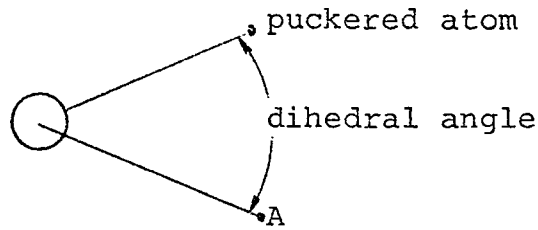
XRING

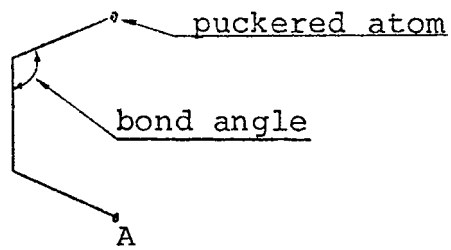
```
      DIMENSION X(6),Y(6),Z(6),XV(5),YV(5),ZV(5),ANGLE(5)
    1 FORMAT()
      READ(5,1)A,B,C
    2 READ(5,1,END=900)(X(I),Y(I),Z(I),I=1,5)
      N=1
    7 DO 10 I=1,5
      XV(I)=(X(I)-X(2))*A
      YV(I)=(Y(I)-Y(2))*B
      ZV(I)=(Z(I)-Z(2))*C
   10 CONTINUE
      XC34=YV(3)*ZV(4)-YV(4)*ZV(3)
      YC34=ZV(3)*XV(4)-ZV(4)*XV(3)
      ZC34=XV(3)*YV(4)-XV(4)*YV(3)
      XC31=YV(3)*ZV(1)-YV(1)*ZV(3)
      YC31=ZV(3)*XV(1)-ZV(1)*XV(3)
      ZC31=XV(3)*YV(1)-XV(1)*YV(3)
      DOT=XC34*XC31+YC34*YC31+ZC34*ZC31
      S4=SQRT(XC34**2.+YC34**2.+ZC34**2.)
      S1=SQRT(XC31**2.+YC31**2.+ZC31**2.)
      ANGLE(N)=ACOS(DOT/(S1*S4))*180./3.1415927
      IF(DOT.GE.0.) GO TO 3
      ANGLE(N)=-1.*ANGLE(N)
    3 IF(N.EQ.5)GO TO 5
      N=N+1
      X(6)=X(1)
      Y(6)=Y(1)
      Z(6)=Z(1)
      DO 12 I=1,5
      X(I)=X(I+1)
      Y(I)=Y(I+1)
      Z(I)=Z(I+1)
   12 CONTINUE
      GO TO 7
    5 WRITE(6,1)(ANGLE(N),N=1,5)
      GO TO 2
  900 STOP
      END
```

XRAY

```
      DIMENSION X(4),Y(4),Z(4),XV(4),YV(4),ZV(4)
    1 FORMAT()
      READ(5,1)A,B,C
    2 READ(5,1,END=900)(X(I),Y(I),Z(I),I=1,4)
      DO 10 I=1,4
      XV(I)=(X(I)-X(2))*A
      YV(I)=(Y(I)-Y(2))*B
      ZV(I)=(Z(I)-Z(2))*C
   10 CONTINUE
      XC34=YV(3)*ZV(4)-YV(4)*ZV(3)
      YC34=ZV(3)*XV(4)-ZV(4)*XV(3)
      ZC34=XV(3)*YV(4)-XV(4)*YV(3)
      XC31=YV(3)*ZV(1)-YV(1)*ZV(3)
      YC31=ZV(3)*XV(1)-ZV(1)*XV(3)
      ZC31=XV(3)*YV(1)-XV(1)*YV(3)
      DOT=XC34*XC31+YC34*YC31+ZC34*ZC31
      S4=SQRT(XC34**2.+YC34**2.+ZC34**2.)
      S1=SQRT(XC31**2.+YC31**2.+ZC31**2.)
      ANGLE=ACOS(DOT/(S1*S4))*180./3.1415927
      IF(DOT.GE.0.) GO TO 3
      ANGLE=-1.*ANGLE
    3 WRITE(6,1)ANGLE
      GO TO 2
  900 STOP
      END
```

(Eq. 13) d = bond length x sin (dihedral angle) x cos (bond angle)





The conformation of 2',3'-O-isopropylidene-N(3),5'-cycloguanosinium iodide was found to be C(4')-endo 0.46Å, C(3')-exo 0.08Å. The dihedral angles are tabulated below (Table II).

The tables show that the dihedral angle calculated from J(2',3') by the Karplus equation is about 12° above the value predicted from the furanose ring conformation (compare the numbers with asterisks). This is within the range of error observed in B. Coxon's experiment. More important, the substituent dihedral angles obtained from the X-ray data indicate that errors of at least the range seen in B. Coxon's

TABLE II

CRYSTAL STRUCTURE AND NMR DATA FOR 2',3'-O-ISOPROPYLIDENE-N(3),5'-CYCLOGUANOSINIUM(CYCLOADENOSINE) IODIDE

	Dihedral Angle	Deviation	Predicted Angle of Hydrogens
C(1')-C(2')-C(3')-C(4')	19.6°	-	19.6°*
C(2')-C(2')-C(3')-O(4')	6.6°	-13.0°	6.6°
C(2')-C(3')-C(4')-O(1')	31.3°	-	88.7°
O(3')-C(3')-C(4')-C(5')	127.8°	-23.5°	112.2°
O(1')-C(1')-C(2')-C(3')	4.8°	-	115.2°
N(9)-C(1')-C(2')-O(2')	147.8°	23.0°	92.2°
C(3')-C(4')-O(1')-C(1')	43.6°	-	-
C(4')-O(1')-C(1')-C(2')	29.7°	-	-
	Coupling Constant		
H(1')-C(1')-C(2')-H(2')	0		75°-105°
H(2')-C(2')-C(3')-H(3')	5.8		32°*
H(3')-C(3')-C(4')-H(4')	0		75°-105°

experiment can be explained on the basis of bond angle distortion alone. In this case, the deviation is towards a more staggered orientation of the hdyrogen atoms, which would be nearly eclipsed at the ring dihedral angle of 19°. Attempts to correlate the Karplus equation to dihedral angles of such strained bicyclic molecules using molecular models alone are therefore unreliable, except as a means of providing us with a maximum value of the expected inherent error in the Karplus equation. It is necessary to have available X-ray structures in which the coordinates of the hydrogen atoms have been accurately determined in order to make a quantitative evaluation of the Karplus equation.

In order to determine whether or not such large angle distortions occur in nucleic acid constituents of non-rigid conformation, the dihedral angles of furanose ring atoms and substituents of fourteen nucleosides and nucleotides were calculated from literature data using the programs XRAY and XRING already described. The results are given in the tables on the following pages (Table III). In these tables ΦA'B' represents the various dihedral angles between atoms attached to atom A' and atom B', sighting along the A'-B' bond. The letter H beside an angle indicates it is a dihedral angle between hydrogen atoms, while the letter S indicates the dihedral angle includes at least one non-hydrogen substituent on the ring atom A' or B'. No letter beside an angle means it is the

TABLE III

CRYSTAL STRUCTURE CONFORMATIONS AND DIHEDRAL ANGLES OF NUCLEOSIDES AND NUCLEOTIDES

Compound	Φ1',2'	Δ1',2'	Φ2',3'	Δ2',3'	Φ3',4'	Δ3',4'	Conformation	
5-methyluridine[97]	25.7°	-	38.5°	-	38.2°	-	C(3')endo	.56Å
	147.7°(S)	+2.0°	47.1°(S)	+8.6°	82.1°(S)	0.3°	C(2')exo	.06Å
	100.3°(H)	+6.0	30.7°(H)	-7.8°	161.7°(H)	3.5°	total	.62Å
A-(2')-p-(5')-U	28.3°	-	38.2°	-	35.6°	-	C(3')endo	.46Å
uridine residue[98]	152.3°(S)	+4.0°	48.6°(S)	+10.4°	83.8°(S)	-0.6°	C(2')exo	.18Å
	88.3°(H)	-3.4°	36.3°(H)	-1.9°	158.6°(H)	3.0°	total	.64Å
Cytidine[99]	28.0°	-	37.4°	-	34.7°	-	C(3')endo	.45Å
	128.4°(S)	-19.6°	42.9°(S)	+5.5°	83.0°(S)	-2.3°	C(2')exo	.16Å
	90.6°(H)	-1.4°	42.7°(H)	+5.3°	155.0°(H)	0.3°	total	.61Å
2,4-dithiouridine[100]	25.9°	-	37.2°	-	35.8°	-	C(3')endo	.51Å
	150.7°(S)	+4.8°	41.4°(S)	+4.2°	79.3°(S)	4.1°	C(2')exo	.10Å
	94.8°(H)	+0.7°	40.6°(H)	+3.4°	166.2°(H)	10.4°	total	.61Å

TABLE III CONTINUED

CRYSTAL STRUCTURE CONFORMATIONS AND DIHEDRAL ANGLES OF NUCLEOSIDES AND NUCLEOTIDES

Compound	Φ1',2'	Δ1',2'	Φ2',3'	Δ2',3'	Φ3',4'	Δ3',4'	Conformation	
5-chlorouridine[101]	32.2°	-	34.5°	-	23.2°	-	C(2')endo	.45Å
	-	-	35.7°(S)	+1.2°	-	-	C(3')exo	.13Å
	153.9°(H)	+1.7°	35.1°(H)	+0.6°	98.8°(H)	+4.1	total	.58Å
A-(2')-p-(5')-U	43.3°	-	38.1°	-	20.0°	-	C(2')endo	.52Å
Adenosine residue[98]	78.1°(S)	+1.4	41.2°(S)	+2.9°	148.1°	+8.1°	C(1')exo	.14Å
	154.3°(H)	-9.0°	29.5°(H)	-8.6°	92.8°(H)	-7.2°	total	.66Å
Thymidine[102]	27.8°	-	36.9°	-	33.2°	-	C(3')exo	.43Å
	154.7°(H)	+6.9°	47.7°(H)	+10.8°	163.9°(S)	+10.7°	C(2')endo	.18Å
	22.9°(H)	-4.9°	75.7(H)	-7.4°	87.6°(H)	0.8°	total	.61Å
			44.9°(S)	+11.0°				

TABLE III CONTINUED

CRYSTAL STRUCTURE CONFORMATIONS AND DIHEDRAL ANGLES OF NUCLEOSIDES AND NUCLEOTIDES

Compound	Φ1',2'	Δ1',2'	Φ2',3'	Δ2',3'	Φ3',4'	Δ3',4'	Conformation	
Deoxycytidine HCL[103]	23.4°	-	31.3°	-	29.7°	-	C(3')endo	.29Å
	70.5°(H)	-26.1°	55.6°(H)	25.3°	159.7°(H)	10.0°	C(2')exo	.22Å
	49.0°(H)	25.6°	166.2°(H)	14.9°	77.4°(S)	-12.9°	total	.51Å
			46.4°(S)	15.1°				
α-D-2'-amino-2'-deoxyadenosine[104]	22.3°	-	29.5°	-	26.7°	-	C(3')endo	.34Å
	98.0°(S)	0.3°	36.7°(S)	7.2°	148.8°(S)	2.1°	C(2')exo	.15Å
	25.9°(H)	3.6°	46.2°(H)	16.7°	97.8°(H)	5.5°	total	.49Å
α-Pseudouridine[105]	34.1°	-	34.3°	-	22.0°	-	C(2')exo	.56Å
	34.2°(S)	+0.1°	34.6°(S)	0.3°	97.1°(S)	-0.9°	C(1')endo	.01Å
	38.9°(H)	4.8°	49.6°(H)	15.3°	143.7°(H)	1.7°	total	.57Å

TABLE III CONTINUED

CRYSTAL STRUCTURE CONFORMATIONS AND DIHEDRAL ANGLES OF NUCLEOSIDES AND NUCLEOTIDES

Compound	Φ1',2'	Δ1',2'	Φ2',3'	Δ2',3'	Φ3',4'	Δ3',4'	Conformation	
3'-CMP, orthorhombic[106]	36.6°	-	38.8°	-	28.1°	-	C(2')endo	.49Å
	77.8°(S)	-5.6°	48.3°(S)	+9.5°	152.0°	+4.1°	C(2')exo	.15Å
	149.3°(H)	-7.3°	62.0°(H)	+23.2°	81.0°(H)	-10.9°	total	.64Å
3'-CMP, monoclinic[107]	23.4°	-	30.9°	-	21.3°	-	C(2')endo	.39Å
	61.5°(S)	-35.1°	44.9°(S)	+14.0°	141.3°(S)	0.0	C(3')exo	.13Å
	144.2°(H)	+ 0.8°	45.5°(H)	14.6°	75.4°(H)	-23.3°	total	.52Å
5-Fluorouracil-deoxyriboside[108]	37.1°	-	40.3°	-	30.1°	-	C(3')exo	.50Å
	33.7°(H)	-3.4°	17.1°(H)	-23.2°	76.8°(H)	-13.3°	C(4')endo	.16A
	121.2°(H)	-35.9°	32.8°(S)	-7.5°	154.3°(S)	+4.2°	total	.66Å
			97.1°(H)	18.1°				

TABLE III CONTINUED

CRYSTAL STRUCTURE CONFORMATIONS AND DIHEDRAL ANGLES OF NUCLEOSIDES AND NUCLEOTIDES

Compound	Φ1',2'	Δ1',2'	Φ2',3'	Δ2',3'	Φ3',4'	Δ3',4'	Conformation	
2'-deoxyadenosine monohydrate[109]	23.7°	-	32.6°	-	33.2°	-	C(3')exo	.48Å
	27.34°(H)	+3.6°	37.8(H)	+5.2°	91.0°(H)	+4.2°	C(2')endo	.06Å
	142.5°(H)	+1.2°	73.3°(H)	-9.1°	156.5°(S)	+3.3°	total	.54Å
			49.0°(S)	16.4°				

dihedral angle between furanose ring atoms. It is assumed that the dihedral angle of substituents should be approximately equal to the ring dihedral angle, or 120° plus or minus the ring dihedral angle. Any deviation is recorded in the column headed ΔA',B' beside the distorted angle.

TABLE IV

DIHEDRAL ANGLE DISTORTIONS FOR HYDROGEN ATOMS OF NUCLEOSIDES

Number of Cases	$\Delta H_A', H_B'$
20	0-5°
14	5-10°
6	10-15°
3	15-20°
3	20-25°
3	25-30°
0	30-35°
1	35-40°
50	

The fourteen nucleosides and nucleotides provided fifty dihedral angles between hydrogen substituents (Table IV). Sixteen of these (32%) deviated by more than ten degrees from the ideal values predicted from bond angles of 120° in the

Newman projection. The 90% boundary of deviation was 23.3°, and 98% of the dihedral angles deviated by 26.1° or less. Of course, since the ring bond angles must be less than 109.5°, the Newman Projection bond angles must deviate from 120°. Ideally this deviation should be such that all bond angles except the ring bond angle (the bond angle involving three ring atoms) are equal. A computer program is presented later for calculating the ideal Newman projection angles on this basis. It will be shown that this idealized deviation results in only a very minor correction to the deviations of dihedral angles in the above tables, never over two degrees. The larger deviations in dihedral angles are therefore a result of some distorting force, presumably steric strain.

In the correlations of the Karplus equation to dihedral angles of carbohydrates and nucleosides already discussed, the largest deviation is within the 98% boundary of 26.1° for distortion of dihedral angles in X-ray structures of nucleosides and nucleotides. On this basis it is reasonable to allow a tolerance of ±26° on each dihedral angle predicted from the Karplus equation. The X-ray data shows that one large deviation is likely to be accompanied by others in the same molecule, so that the occurrence of two or three large deviations (but within the 26° limit) in the same conformation cannot be used as evidence that the conformation is statistically improbable.

None of the known assessments of the validity of the Karplus equation show an error greater than 24° for any dihedral angle, so it was decided that the 26° tolerance of the X-ray structures was an adequate tolerance for dihedral angles calculated from coupling constants. We do not mean to imply that distorted bond angles are the only sources of error in this method. For comparison, it has been estimated that $J_{1,2}$ of sugar acetates might be about 1.5 Hz too small because of the two electronegative substituents at C-1;[83] it has also been estimated that when an electronegative substituent such as OH is anti-periplanar to a proton on a neighboring carbon atom, the vicinal coupling constant between protons on the two carbons might be about 2.0 Hz too small;[81] finally, an error of up to 10% (up to 1 Hz) may be caused by incorrect choice of the constant J_0 in the Karplus equation. Our 26° tolerance allows a maximum error in theoretical coupling constant of about 4 Hz.

With an idea of the range of expected deviation from the Karplus equation in hand, the next project was to test available angles measured from models against dihedral angles obtained from X-ray structures. After extending Jardetsky's 0.4Å and 0.6Å data into full tables for all envelope conformations, large discrepancies of b-qa and b-ae angles with analogous angles on X-ray structures were noticed. Because of this and the paucity of data for twist conformations

of various degrees of puckering, a method was developed for calculating the dihedral angles mathematically.

To calculate theoretical coupling constants for each conformation, it was assumed that all bond lengths were 1.54Å and all projected angles on the Newman projection were 120°. The nomenclature of C. Jardetsky was adopted. The ring carbon atoms containing bisectional substituents are referred to as "bisectional atoms"; those showing qe and qa substituents are "quasi atoms", and those having e and a substituents are "puckered atoms" of the five-membered ring. The angles b-b cis and b-b trans remain 0° and 120°, respectively. To calculate the rest of the angles, it was necessary to calculate the dihedral angles of the furanose ring atoms. The dihedral angles for substituents were then calculated from Newman projections.

Calculation of the angles b-qa and b-qe (cis and trans) was quite straightforward. The equations for the dihedral angles of the furanose atoms related to these substituents is applicable to envelope, twist, and all conformations between these. The equation was derived in the following way. In the figure below (Figure 24), we sighted along the bond of two carbon atoms sharing bisectional and quasi-substituents, respectively, sighting from bisectional to quasi. The line C(3)-C(4) makes an angle of (∠A-90°) away from the plane AB, which is perpendicular to the line of sight. It follows

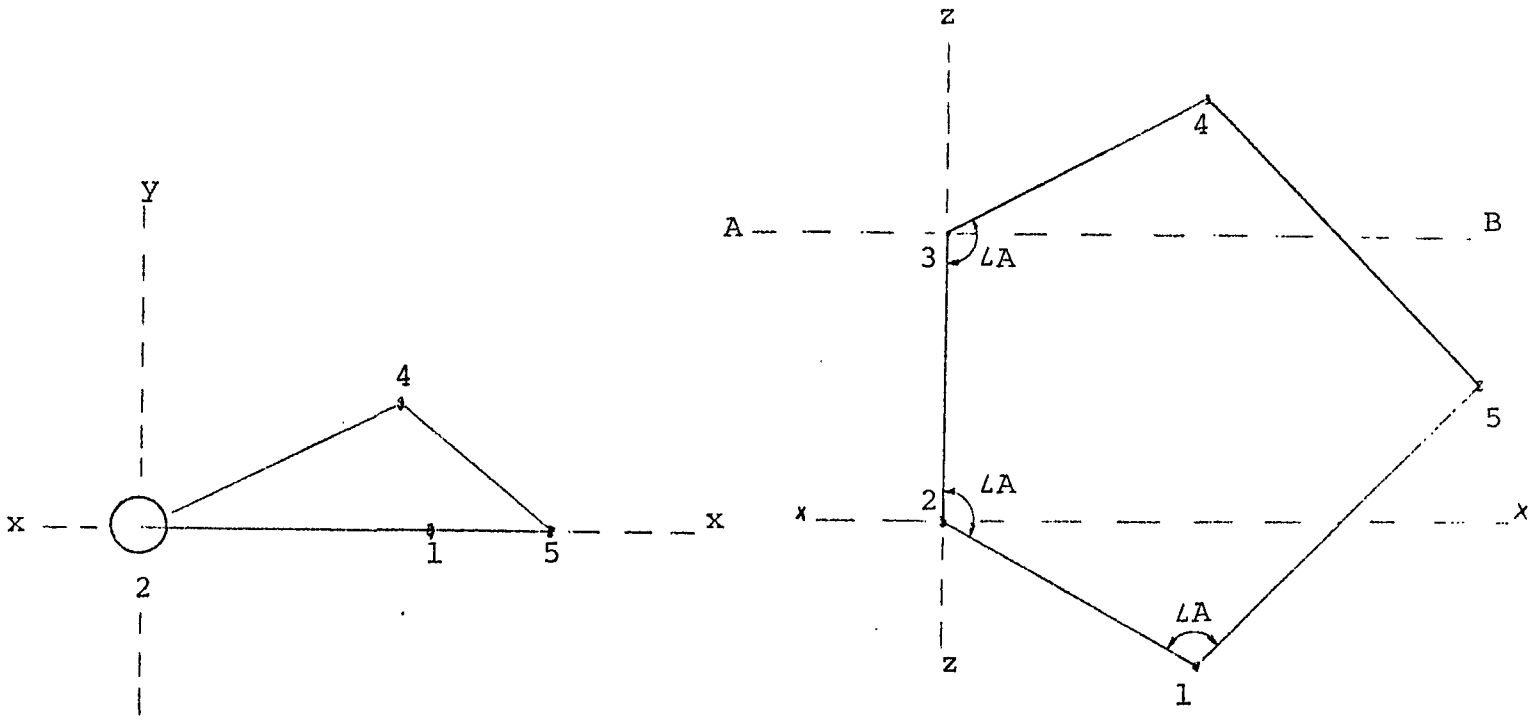


Figure 24. Mathematical Model Used in Program DIHEDR

that the dihedral angle C(1)-C(2)-C(3)-C(4) (α) can be obtained from the equation (Eq. 14):

$$\text{(Eq. 14)} \quad \sin \alpha = d \div 1.54 \cos(\angle A - 90°)$$

where d is distance from C(4) to the plane containing C(1), C(2), and C(3). From this equation it is easily seen that the various bisectional-quasi angles are the same for an envelope conformation of 0.4Å above plane C(1)C(2)C(3), as for a twist conformation of 0.4Å pucker above and below the plane C(1)C(2)C(3) (0.8Å "total pucker").

To calculate the angles qe-e, qe-a, qa-e, and qa-a (puckered-quasi angles) it was necessary to define the locus of each ring atom in terms of Cartesian coordinates (x,y,z). To do this, the line through C(3) and C(2) of Figure 24 was designated the Z axis. The line AB is parallel to the X axis, and the Y axis is perpendicular to the C(1)-C(2) bond in the Newman projection. Thus the locus of C(2) is (0,0,0), C(3) is (o,0,1.54), and C(1) is (1.54 cos(∠A-90°), 0,0).

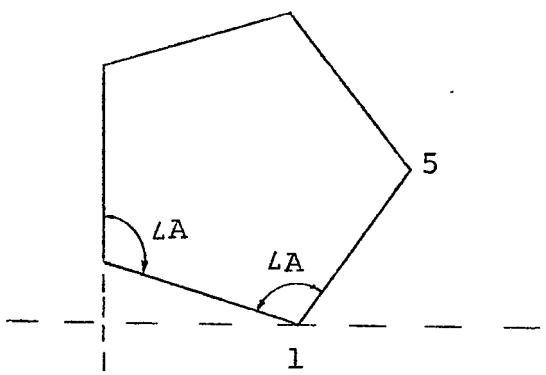


Figure 25. Top View of the Mathematical Model of Cyclopentane

From Figure 25 it can be seen that the X-coordinate of the C(5) atom is 1.54 cos(∠A-90°) + 1.54 cos(270°-2∠A), and its Z-coordinate of C(4) is 1.54 + 1.54 sin(∠A-90°).

The projection of the line C(3)-C(4) on the AB plane of the former figure is 1.54 cos (∠A-90°), and from this we obtain an X-coordinate for C(4) of 1.54 cos(∠A-90°) x cos α.

The coordinates thus obtained are tabulated below (Table V):

TABLE V

COORDINATES USED IN PROGRAM DIHEDR

	X	Y	Z
(1)	1.54 cos(∠A-90°)	0	-1.54 sin(∠A-90°)
(2)	0	0	0
(3)	0	0	1.54
(4)	1.54 cos(∠A-90°)αcos	d	1.54 + 1.54 sin(∠A-90°)
(5)	1.54 cos(∠A-90°) + 1.54 cos(270°-2∠A)	0	-1.54 sin∠(A-90°) + 1.54 sin(270°-2∠A)

To obtain the angle C(2)-C(3)-C(4)-C(5)(∠β) in which C_4 is the puckered atom of an envelope conformation, the following vector equation was used (Eq. 13):

$$(\text{Eq. 13}) \quad \cos \beta = \frac{(\vec{C_3C_4} \times \vec{C_3C_5}) \cdot (\vec{C_3C_4} \times \vec{C_3C_2})}{|\vec{C_3C_4} \times \vec{C_3C_5}| \quad |\vec{C_3C_4} \times \vec{C_3C_2}|}$$

The dihedral angles of the substituent were obtained from the following Newman projections (Figure 26).

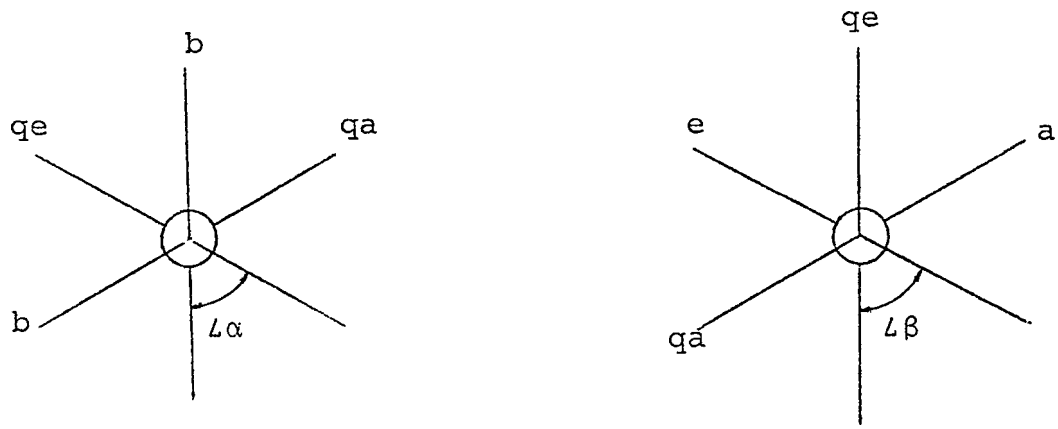


Figure 26

Newman Projections Showing Relationship of Substituent Angles to Ring Bond Angles

The angles ∠α and ∠β are all that are necessary to calculate all the dihedral angles for the envelope conformation, under the assumptions made (Table VI).

TABLE VI

DIHEDRAL ANGLES FOR THE ENVELOPE CONFORMATION

a-qe = e-qa = $\angle\beta$	b-qa cis = b-qe cis = $\angle\alpha$
a-qa = 120 + $\angle\beta$	b-qa trans = 120 + $\angle\alpha$
a-qe = 120 - $\angle\beta$	b-qe trans = 120 - $\angle\alpha$

The above mathematical model of cyclopentane was incorporated into a Fortran V computer program named DIHEDR. The input of this computer program consists of the distance of C(4) from the four atom plane and the angle $\angle A$. The printed output consists of $\angle\alpha$, $\angle\beta$, and the various substituent angles and the corresponding coupling constants calculated from the form of the Karplus equation given in the Introduction. Also, the bond length C(4)-C(5) is printed. The input bond angle must be adjusted so that this bond length is equal to 1.5400Å. This is conveniently done from the teletype facilities available in this laboratory; alternatively, the program could be easily altered so that this is done by the computer. The program is printed on the next page.

While the above program is convenient where only rough approximations are desired, it suffers from the drawback that four ring bond angles must be equal, and the ring bond angle of the puckered atom is always smallest (except in planar conformation). In nucleosides, X-ray crystal structures show

```
      DIMENSION PHI(6),RJ(6)
      D=1.54
    1 FORMAT()
    2 READ(5,1,END=900)B
      READ(5,1)ANGLE
      PSI=(ANGLE-90.)*3.14159/180.
       U=COS(PSI)
      V=SIN(PSI)
      PSI2=(270.-2.*ANGLE)*3.14159/180.
      X=COS(PSI2)
      Y=SIN(PSI2)
      C=D*V+D
      E=D*U
      F=-D*V
      A=E*SQRT(1.-(B/E)*(B/E))
       G=E+D*X
      H=F+D*Y
      RI=-B*D
      RJ2=A*D
      RK=B*H-B*D
      RL=A*D-A*H+G*C-G*D
      RM=B*G
      W1=RI*RK+RJ2*RL
      W2=SQRT(RI*RI+RJ2*RJ2)
      W3=SQRT(RK*RK+RL*RL+RM*RM)
      ANGLE1=ACOS(W1/(W2*W3))*180./3.14159
      ANGLE2=ASIN(B/E)*180./3.14159
      WRITE(6,1)ANGLE1,ANGLE2
      PHI(1)=ANGLE1
      PHI(2)=120.+ANGLE1
      PHI(3)=120.-ANGLE1
      PHI(4)=ANGLE2
      PHI(5)=120.+ANGLE2
      PHI(6)=120.-ANGLE2
      WRITE(6,1)PHI
      DO 10 I=1,6
      IF(PHI(I).LE.90.) GO TO 4
      RJ(I)=9.5*(COS(PHI(I)*3.14159/180.))**2.-0.28
      GO TO 10
    4 RJ(I)=8.5*(COS(PHI(I)*3.14159/180.))**2.-0.28
   10 CONTINUE
      WRITE(6,1)RJ
      TESTBD=SQRT((G-A)*(G-A)+B*B+(H-C)*(H-C))
      WRITE(6,1)TESTBD
      GO TO 2
  900 STOP
      END
```

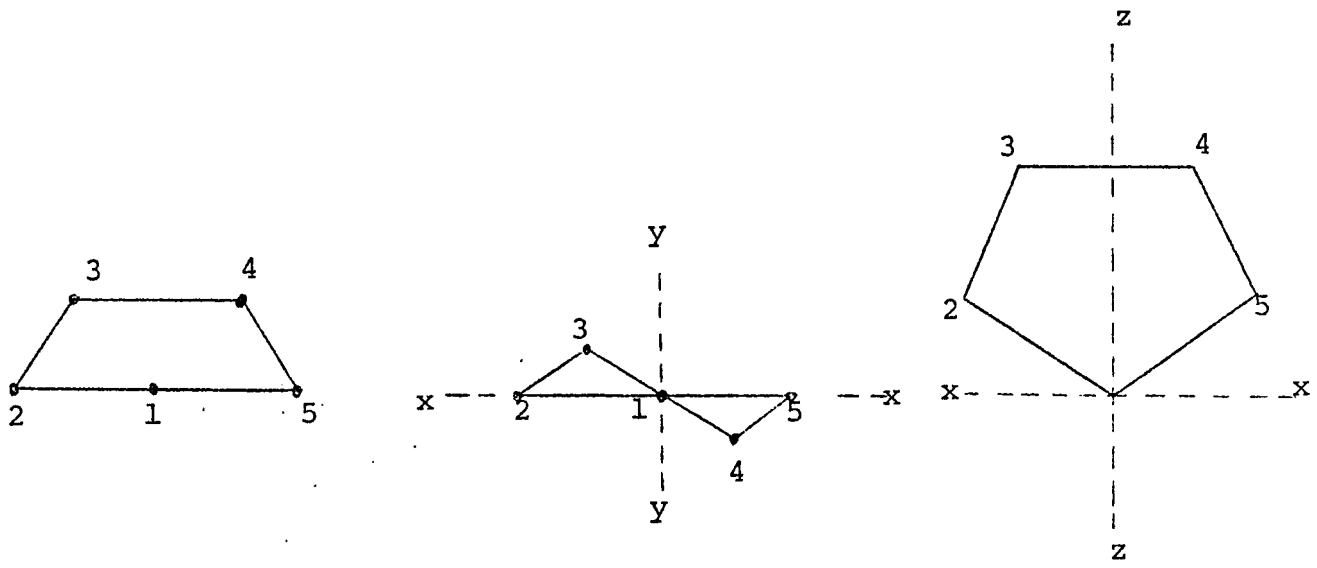


Figure 27. The Mathematical Model Used in Program CYCLO

that the ring bond angles are smallest at C(2') and C(3'), followed by C(1') and C(4'), and then by the O(1') ring bond angle, which is always largest. To more accurately approximate the real molecules, another program was developed. This program can be used to predict dihedral angles of envelope, twist, or intermediate conformations.

In this model of cyclopentane (Figure 27), C(1) is assigned the x,y,z coordinates (0,0,0). The positions of C(2) and C(5) are easily calculated from the only dictated bond angle, ∠1. The x-coordinates of C(3) and C(4) in the envelope conformation are simply -1.54/2 and 1.54/2, respectively. Their y-coordinates are input variables, and the z-coordinates of these points are obtained from the equation (Eq. 14):

$$\text{(Eq. 14)}\quad z_b = z_a + \sqrt{1.54^2-(x_b-x_a)^2-(y_b-y_a)^2}$$

For twist conformation, the x-coordinates are obtained from the equation (Eq. 15):

$$\text{(Eq. 15)}\quad x = \pm\sqrt{(1.54/2)^2-d^2}$$

For intermediate conformations, x(3) and x(4) are input variables and must be adjusted so that the bond length C(3)-C(4) is the desired length.

The x, y, z coordinates for twist conformation are given in the table below (Table VII).

TABLE VII

COORDINATES OF COMPUTER PROGRAM CYCLO

	x	y	z
(1)	0	0	0
(2)	A	0	B
(3)	C	D	E
(4)	F	G	H
(5)	I	0	B

$A = 1.54 \cos\left(\frac{180-\angle 1}{2}\right)$ $\quad E = B + \sqrt{1.54^2-D^2-(C-A)^2}$ $\quad C = \sqrt{.77^2-D^2}$

$B = 1.54 \sin\left(\frac{180-\angle 1}{2}\right)$ $\quad F = -\sqrt{.77^2-G^2}$ $\quad G = -D$

C = input variable $\quad H = B + \sqrt{1.54^2-G^2-(F-I^2)}$ $\quad I = -A$

The mathematical model was incorporated into a Fortran V computer program designed for teletype use. The program is printed on the following pages.

The program consists of three mathematical models incorporated into one program. For pure cyclopentane envelope conformations, the distance of atoms C(3) and C(4) from the plane of the atoms C(1), C(2) and C(5) must be typed twice (one for each atom), then ∠1 must be typed. For pure cyclopentane twist

```
  1 FORMAT()
  2 READ(5,1,END=900)D
  4 READ(5,1)G
    READ(5,1)ANGLE
    IF(ANGLE.LE.10.) GO TO 204
    PSI=(ANGLE-90.)*3.1415927/180.
    U=COS(PSI)
    PSI3=((180.-ANGLE)/2.)*3.1415927/180.
    X=COS(PSI3)
    Y=SIN(PSI3)
    A=1.54*X
    B=1.54*Y
    RI=-A
    IF(G.LE.0.) GO TO 20
    C=0.77
    E=B+SQRT(-1.*(C-A)*(C-A)-(D*D)+1.54*1.54)
    F=-0.77
 89 H=B+SQRT(-1.*(F-RI)*(F-RI)-(G*G)+1.54*1.54)
    RJ=(G-D)*(H-B)
    RK=(E-H)*G
    RL=(H-E)*(F-RI)
    RM=(C-F)*(H-B)
    RN=(H-E)*(F-A)
    P=G*(F-C)
    Q=(D-G)*(F-RI)
    S=(A-F)*(G-D)
    O=G*(H-B)-H*G
    R=H*(F-RI)+F*(B-H)
    T=F*G+G*(RI-F)
    W1=(RJ+RK)**2.+(RL+RM)*(RN+RM)+(P+Q)*(P+S)
    W2=SQRT((RJ+RK)*(RJ+RK)+(RL+RM)*(RL+RM)+(P+Q)*(P+Q))
    W3=SQRT((RJ+RK)*(RJ+RK)+(RN+RM)*(RN+RM)+(P+S)*(P+S))
    W4=O*(RJ+RK)+R*(RL+RM)+T*(P+Q)
    W5=SQRT(O*O+R*R+T*T)
    A2345=ACOS(W1/(W2*W3))*180./3.14159
    A3451=ACOS(W4/(W5*W2))*180./3.14159
    QA=A*(B-E)-B*(A-C)
    QB=-D*(B-H)+G*(B-E)
    QC=(B-E)*(A-F)-(A-C)*(B-H)
    QD=(A-C)*(-G)+D*(A-F)
    W6=-D*B*QB+QA*QC+D*A*QD
    W7=SQRT(D*D*B*B+QA*QA+A*A*D*D)
    W8=SQRT(QB*QB+QC*QC+QD*QD)
    W9=(B-2.*A)*QA
    W10=B-2.*A
    A5123=ACOS(W9/(W10*W7))*180./3.1415927
    A5123=180.-A5123
    A1234=ACOS(W6/(W7*W8))*180./3.1415927
    BOND34=SQRT((C-F)**2.+(D-G)**2.+(E-H)**2.)
    TL13=SQRT(C*C+D*D+E*E)
```

```
      A13=ASIN(TL13/(2.*1.54))*180./3.1415927*2.
     TL35=SQRT((C+A)*(C+A)+D*D+(E-B)*(E-B))
     A35=ASIN(TL35/(3.08))*180./3.1415927*2.
     TL41=SQRT(F*F+G*G+H*H)
     A41=ASIN(TL41/3.08)*180./3.1415927*2.
     TL52=SQRT((RI-A)*(RI-A))
     A52=ASIN(TL52/3.08)*180./3.1415927*2.
     TL24=SQRT((A-F)*(A-F)+G*G+(B-H)*(B-H))
     A24=ASIN(TL24/(3.08))*180./3.1415927*2.
     WRITE(6,1)A1234,A2345,A3451,A5123
     WRITE(6,1)BOND34
     IF(ANGLE.LE.10.) GO TO 205
  91 WRITE(6,1)A13,A24,A35,A41,A52
     IF(ANGLE.LE.10.)GO TO 2
     IF(G.LE.0.) GO TO 2
     AD=((180.-ANGLE)/2.)*3.1415927/180.
     PD3=1.54*SIN(AD)
     PD4=1.54*COS(2.*ASIN(TL24/3.08)-3.1415927/2.)
     D4=PD3*D/PD4
     A43=ASIN(D/PD4)*180./3.1415927
     WRITE(6,1)D4,A43
     GO TO 2
  20 C=SQRT(0.77*0.77-D*D)
     E=B+SQRT(-1.*(C-A)*(C-A)-(D*D)+1.54*1.54)
     F=SQRT(0.77*0.77-G*G)*-1.
     GO TO 89
 204 F=ANGLE
     READ(5,1)C
     READ(5,1)ANGL
     PSI=(ANGL-90.)*3.1415927/180.
     U=COS(PSI)
     PSI3=((180.-ANGL)/2.)*3.1415927/180.
     X=COS(PSI3)
     Y=SIN(PSI3)
     A=1.42*X
     B=1.42*Y
     RI=-A
     E=B+SQRT(-1.*(C-A)*(C-A)-(D*D)+1.54*1.54)
     GO TO 89
 205 TL13=SQRT(C*C+D*D+E*E)
     A13=ASIN(TL13/(2.96))*180./3.1415927*2.
     TL35=SQRT((C+A)*(C+A)+D*D+(E-B)*(E-B))
     A35=ASIN(TL35/(3.08))*180./3.1415927*2.
     TL41=SQRT(F*F+G*G+H*H)
     A41=ASIN(TL41/2.96)*180./3.1415927*2.
     TL52=SQRT((RI-A)*(RI-A))
     A52=ASIN(TL52/2.84)*180./3.1415927*2.
     TL24=SQRT((A-F)*(A-F)+G*G+(B-H)*(B-H))
     A24=ASIN(TL24/(3.08))*180./3.1415927*2.
     GO TO 91
 900 STOP
     END
```

conformations, the distance of atoms C(3) and C(4) from the plane of the other three atoms must be typed, and they must be equal in magnitude and opposite in sign. The angle ∠1 must also be typed. The output consists of the four ring dihedral angles C(1)-C(2)-C(3)-C(4), C(2)-C(3)-C(4)-C(5), C(3)-C(4)-C(5)-C(1) and C(5)-C(1)-C(2)-C(3); then the bond length C(3)-C(4), which is always 1.5400 for these two cases (no adjustments required, an improvement over DIHEDR); then all five angles ∠1, ∠2, ∠3, ∠4, and ∠5 are printed. For envelope conformations the distance of C(1) from the four atom plane of the other atoms is also printed, along with the angle between the planes. By altering ∠1, the angles ∠2 = ∠5 and ∠3 = ∠4 may be adjusted to any order of relative size. For twist conformations designed for nucleoside applications, make ∠1 > ∠2 = ∠5 > ∠3 = ∠4. For envelope, the opposite order should be used (the puckered atom usuallyu has the smallest bond angle in X-ray structures).

The same program also gives dihedral angles for intermediate, envelope and twist conformations of tetrahydrofuran in which the oxygen atom is at the origin in Figure 27. After Y(3) and Y(4) are typed as input, X(3) and X(4) must also be typed, followed by ∠1. Output is the same as for cyclopentane twist conformation, and bond length C(3)-C(4) must be normalized to 1.54Å by adjusting X(3), X(4), and ∠1. Any relative order of magnitude of bond angles may be produced. The program was named CYCLO.

The ring dihedral angles were calculated for cyclopentane analogs of the various furanose ring conformations seen in X-ray structure of nucleosides and nucleotides. These calculations were done using the program CYCLO for pure envelope and twist conformations only. The hydrogen dihedral angles were calculated by assuming 120° bond angles in the Newman projections, and the tables on the following pages were prepared. These tables show the hydrogen dihedral angles for all common nucleoside conformations of total pucker 0.71Å, 0.5Å, and 0.3Å (Table VIII, Table IX, and Table X, respectively).

The average total pucker in the X-ray structures examined earlier was 0.59Å and the range for the fourteen examples was 0.49 to 0.67 Å total pucker. Therefore, the 0.50Å and 0.71Å tables are probably suitable for defining the approximate upper and lower limits of nucleoside conformation in solution. It is recommended that these tables should be used by calculating the theoretical dihedral angles of the molecule being studied by nmr, using the Karplus equation; each dihedral angle should be considered to have a possible error of 26°, a total range of 52°. Unlikely conformations can then be eliminated. When the tables are used in this way, reliable conclusions can be reached concerning conformations of five-membered saturated rings in solution. For quick comparison to experimental data and for accurate

TABLE VIII A

NUCLEOSIDE CONFORMATIONS OF 0.71Å TOTAL PUCKER

	Φ(1'2')	Φ(2'3')	Φ(3'4')	cis Φ(1'2")	trans Φ(2"3')	cis Φ(3'4")
C(2')endo	165°	45°	92°	45°	75°	28°
C(2')endo-C(1')exo	165°	36°	108°	45°	84°	14°
C(2')endo-C(3')exo	156°	45°	84°	36°	75°	36°
C(3')endo	92°	45°	165°	28°	75°	45°
C(3')endo-C(4')exo	108°	36°	165°	14°	84°	45°
C(2')exo	75°	45°	148°	45°	165°	45°
C(2')exo-C(1')endo	75°	36°	134°	45°	84°	14°
C(3')exo	148°	45°	75°	45°	165°	45°
C(3')exo-C(4')endo	134°	36°	75°	14°	84°	45°
C(3')endo-C(2')exo	84°	45°	156°	36°	165°	36°

TABLE VIII B

NUCLEOSIDE CONFORMATIONS OF 0.71Å TOTAL PUCKER

	J(1'2')	J(2'3')	J(3'4')	cis J(1'2")	trans J(2"3')	cis J(3'4")
C(2')endo	8.6 Hz	4.0 Hz	-0.27 Hz	4.0 Hz	0.29 Hz	6.3 Hz
C(2')endo-C(1')exo	8.6	5.2	0.63	4.0	-0.19	7.7
C(2')endo-C(3')exo	7.6	4.0	-0.19	5.2	0.29	5.2
C(3')endo	-0.27	4.0	8.6	6.3	0.29	4.0
C(3')endo-C(4')exo	0.63	5.2	8.6	7.7	-0.19	4.0
C(2')exo	0.29	4.0	6.6	4.0	8.6	4.0
C(2')exo-C(1')endo	0.29	5.2	4.3	4.0	-0.19	7.7
C(3')exo	6.6	4.0	0.29	4.0	8.6	4.0
C(3')exo-C(4')endo	4.3	5.2	0.29	7.7	-0.19	4.0
C(3')endo-C(2')exo	-0.19	4.0	7.6	5.2	8.6	5.2

TABLE IX A

NUCLEOSIDE CONFORMATIONS OF 0.5Å TOTAL PUCKER

	Φ(1'2')	Φ(2'3')	Φ(3'4')	cis Φ(1'2')	trans Φ(2'3')	cis Φ(3'4')
C(2')endo	152°	32°	100°	32°	88°	20°
C(2')endo-C(1')exo	152°	26°	110°	32°	94°	10°
C(2')endo-C(3')exo	146°	32°	94°	26°	88°	26°
C(3')endo	100°	32°	152°	20°	88°	32°
C(3')endo-C(4')exo	110°	26°	152°	10°	94°	32°
C(2')exo	88°	32°	140°	32°	152°	32°
C(2')exo-C(1')endo	88°	26°	130°	32°	94°	10°
C(3')exo	140°	32°	88°	32°	152°	32°
C(3')exo-C(4')endo	130°	26°	88°	10°	94°	32°
C(3')endo-C(2')exo	94°	32°	152°	26°	152°	26°

TABLE IX B

NUCLEOSIDE CONFORMATIONS OF 0.5Å TOTAL PUCKER

	J(1'2')	J(2'3')	J(3'4')	cis J(1'2")	trans J(2"3')	cis J(3'4')
C(2')endo	7.1 Hz	5.8 Hz	-0.01 Hz	5.8 Hz	-0.27 Hz	7.2 Hz
C(2')endo-C(1')exo	7.1	6.6	0.83	5.8	-0.23	8.0
C(2')endo-C(3')exo	6.3	5.8	-0.23	6.6	-0.27	6.6
C(3')endo	-0.01	5.8	7.1	7.2	-0.27	5.8
C(3')endo-C(4')exo	0.83	6.6	7.1	8.0	-0.23	5.8
C(2')exo	-0.27	5.8	5.3	5.8	7.1	5.8
C(2')exo-C(1')endo	-0.27	6.6	3.6	5.8	-0.23	8.0
C(3')exo	5.3	5.8	-0.27	5.8	7.1	5.8
C(3')exo-C(4')endo	3.6	6.6	-0.27	8.0	-0.23	5.8
C(3')endo-C(2')exo	-0.23	5.8	7.1	6.6	7.1	6.6

TABLE X

NUCLEOSIDE CONFORMATIONS OF 0.3Å TOTAL PUCKER

	Φ(1'2')	Φ(2'3')	Φ(3'4')	cis Φ(1'2')	trans Φ(2'3')	cis Φ(3'4')
C(2')endo	139°	19°	108°	19°	101°	12°
C(2')endo-C(1')exo	139°	15°	114°	19°	105°	6°
C(2')endo-C(3')exo	135°	19°	105°	15°	101°	15°
C(3')endo	108°	19°	139°	12°	101°	19°
C(3')endo-C(4')exo	114°	15°	139°	6°	105°	19°
C(2')exo	101°	19°	132°	19°	139°	19°
C(2')exo-C(1')endo	101"	15°	126°	19°	105°	6°
C(3')exo	132°	19°	101°	19°	139°	19°
C(3')exo-C(4')endo	126°	15°	101°	6°	105°	19°
C(3')endo-C(2')exo	105°	19°	135°	15°	139°	15°

analysis of conformational equlibria, tables of coupling constants are also presented.

The values in the following tables were obtained by assigning the puckered atom or atoms the smallest ring bond angle, followed by the adjacent atoms, and then the most remote atom or atoms, with the largest ring bond angle. The differences between the sizes of the angles were about the same. This approximates the situation in crystal structures of nucleosides.

Most nucleotides are in an intermediate conformation in crystalline form. To study the relationship between various intermediate conformations and pure envelope and twist conformations, the following table (Table XI) was prepared using the tetrahydrofuran mathematical model of the computer program CYCLO. The table shows the three dihedral angles for the furanose ring of a ribonucleoside in C(2')-endo conformations ranging from pure envelope to pure twist conformation, all with 0.6Å total pucker.

It can be seen that there is a continuous variation in each dihedral angle as the conformation moves from envelope to twist, and the difference between envelope and twist for C(2')-endo-C(3')-exo conformations is not large. Since conformations in which the less puckered atom is out of plane by more than half the distance of the puckered atom are rare (only one example among the fourteen crystal structures studied

TABLE XI

CALCULATED C(2')-ENDO CONFORMATIONS OF FURANOSE RINGS OF 0.6Å TOTAL PUCKER

D(Å)	G(Å)	Φ(1',2)	Φ(2',3')	Φ(3',4')
0.6	0.0	157°	36°	96 1/2°
0.5	-0.1	154 1/2°	35 1/2°	94 1/2°
0.45	-0.15	153°	35°	93 1/2°
0.4	-0.2	152°	35°	92°
0.3	-0.3	150°	35°	90°

in this project), a value of about two-thirds of the way from the twist to the envelope value might be chosen for each dihedral angle in order to simplify the theoretical tables. The only errors of greater than three degrees would arise if C(1') or C(4') were one of the puckered atoms, and if it were puckered by more than one-fifth of the distance of the more puckered C(2') or C(3') atom. Only one example of this appears among the fourteen X-ray structures studied.

For comparison of dihedral angles of idealized cyclopentane and tetrahydrofuran models of C(2')-endo conformation, the following table (Table XII) of cyclopentane model angles is presented (compare these to Table XI).

TABLE XII

CALCULATED C(2')-ENDO CONFORMATIONS OF 0.6Å TOTAL PUCKER, USING A CYCLOPENTANE MODEL

Conformation	Φ(1',2')	Φ(2',3')	Φ(3',4')
0.6Å Envelope	158°	38°	96°
0.6Å Twist	150°	38°	90°

The correlation is quite satisfactory. Crystal structures show nucleosides are between the two extremes of these models: the average carbon-oxygen bond length of the furanose

ring is greater than 1.24Å, while carbon-carbon bond lengths are less than 1.54Å. The difference between the two models and between envelope and twist conformations becomes less as the total pucker is decreased.

The ring bond angles of the five membered saturated rings in all of the computer models are less than 109.5° on the average and are different from the ring bond angles of one or both adjacent ring atoms. In order to determine the error this causes in the dihedral angles of the above tables, computer programs GEM, JNVLOP, and JTWIST were written.

Computer program GEM prints the Newman projection bond angle between substituents on a ring atom (substituent bond angle), A120S, when fed the "tetrahedral" ring bond angle, A215, and the "tetrahedral" bond angle A314 which must be adjusted so that the calculated bond angles A315 and A312, which are printed along with A120S, are equal. The above angles are named according to the diagram below (Figure 28).

FIGURE 28

The Bond Angles of Computer Program GEM

GEM

```
  1 FORMAT()
  2 READ(5,1,END=900)A215
    READ(5,1)A314
    BA2=((360.-A215)/2.-90.)*3.1415927/180.
    RAD=3.1415927/180.
    A215=A215*RAD
    A314=A314*RAD
    A90=3.1415927/2.
    A=-1.54*COS(A215-A90)
    B=-1.54*SIN(A215-A90)
    C=1.54*COS(A314/2.)*COS(BA2)
    D=1.54*SIN(A314/2.)
    E=-1.54*COS(A314/2.)*SIN(BA2)
    RDA315=ASIN(SQRT(C**2.+D**2.+(E-1.54)**2.)/3.08)
    A315=2.*RDA315/RAD
    A312=2.*ASIN(SQRT((C-A)**2.+D*D+(E-B)**2.)/3.08)/RAD
    A120S=2.*ASIN(D/(1.54*COS(2.*RDA315-A90)))/RAD
    WRITE(6,1)A315,A312
    WRITE(6,1)A120S
    GO TO 2
900 STOP
    END
```

JNVLOP

```
      DIMENSION A(8),RJ(8)
    1 FORMAT()
    2 READ(5,1,END=900)B
      READ(5,1)Q
      READ(5,1)AEG
      READ(5,1)QAQE
       READ(5,1)BB
      ARB=(360.-BB)/2.
      ARP=(360.-AEG)/2.
      ARQ=(360.-QAQE)/2.
      A(1)=Q+ARQ-ARP
      A(2)=Q+ARP-ARQ
      A(3)=QAQE-A(2)
      A(4)=QAQE+A(1)
      A(5)=B+ARQ-ARB
      A(6)=B+ARB-ARQ
      A(7)=BB-A(5)
      A(8)=BB+A(6)
      DO 6 I=1,8
      IF(A(I).LE.90.) GO TO 4
      RJ(I)=9.5*COS(A(I)*3.1415927/180.)**2.-0.28
      GO TO 6
    4 RJ(I)=8.5*COS(A(I)*3.1415927/180.)**2.-0.28
    6 CONTINUE
      WRITE(6,1)(A(I),I=1,4)
      WRITE(6,1)(RJ(I),I=1,4)
      WRITE(6,1)(A(I),I=5,8)
      WRITE(6,1)(RJ(I),I=5,8)
      GO TO 2
  900 STOP
      END
```

JTWIST

```
      DIMENSION A(8),RJ(8)
    1 FORMAT()
    2 READ(5,1,END=900)P
      READ(5,1)Q
      READ(5,1)AEG
      READ(5,1)QAQE
      ARP=(360.-AEG)/2.
      ARQ=(360.-QAQE)/2.
      A(1)=Q+ARQ-ARP
      A(2)=Q+ARP-ARQ
      A(3)=QAQE-A(2)
      A(4)=QAQE+A(1)
      A(5)=AEG+A(2)
      A(6)=P
      A(7)=AEG-A(6)
      A(8)=AEG+P
      DO 6 I=1,8
      IF(A(I).LE.90.) GO TO 4
      RJ(I)=9.5*COS(A(I)*3.1415927/180.)**2.-0.28
      GO TO 6
    4 RJ(I)=8.5*COS(A(I)*3.1415927/180.)**2.-0.28
    6 CONTINUE
      WRITE(6,1)(A(I),I=1,4)
      WRITE(6,1)(RJ(I),I=1,4)
      WRITE(6,1)(A(I),I=5,8)
      WRITE(6,1)(RJ(I),I=5,8)
      GO TO 2
  900 STOP
      END
```

The angles A120S obtained from GEM are named AEG, QAQE, or BB depending whether they are calculated for puckered, quasi, or bisectional ring atoms, respectively, for a particular envelope of twist conformation. For envelope cyclopentane conformations, these three angles along with B, the second largest ring dihedral angle, and Q, the largest ring dihedral angle, are the data required for the computer program JNVLOP. This program gives all the dihedral angles and coupling constants except those between substituents on both bisectional atoms. These latter angles, of course, never change from 0° or 120°.

Similar data for twist cyclopentane conformations is obtained from JTWIST. In this case, P is the largest ring dihedral angle. Bisectional-quasi substituent dihedral angles may be obtained by substituting A for P, B for Q, QAQE for AEG, and BB for QAQE.

The programs GEM, JTWIST, and JNVLOP may be used if higher accuracy is desired. They were used here only to determine whether the approximations used in constructing the tables of substituent dihedral angles were valid. It was found they were indeed valid to within two degrees even for 0.71Å total pucker using the tetrahydrofuran computer model, where the largest deviation occurred.

The literature tables of C. D. Jardetzky, measured on molecular models[31,32,33,34] all contain large errors. For

example, the b-qa cis angle for maximally puckered envelope conformations are 28° rather than the literature 15°. Maximally puckered envelope conformation in our tables have 0.71Å, which makes C. D. Jardetzky's values for e-qe angles (and related angles) equal to ours. Using their value of 0.75Å puckering, the error for these and all related angles would be still larger. Also, the a-e angles of twist conformations of 0.8Å total pucker should be about 50° instead of 60°. The recently published tables for envelope conformation measured by Smith, et al.[110] by an improved method are somewhat more accurate, and agree well with an interpolated value of about 0.55Å pucker from our tables. We have the only complete tables for conformations of all degrees of puckering, the only correlation data to actual nucleosides, and the only accurate tables for twist and intermediate conformations.

From the angles in Table XIII, the substituent dihedral angles and coupling constants for envelope and twist conformations may be obtained by means of Newman projections (Table XIV) or by means of the computer programs JTWIST or JNVELOP, using AEG = QAQE = BB = 120.0°.

In view of the errors of the Jardetzky tables, some of the literature conclusions based on them were reviewed. The dihedral angles calculated from literature coupling constants of α-pseudouridine are given below.[111]

TABLE XIII

RING DIHEDRAL ANGLES OF CYCLOPENTANE IN VARIOUS TOTAL PUCKERS. SEE TABLE XIV.

	Envelope		Twist		
Total pucker	Q	B	P	Q	B
0.80	-	-	50.1°	40.4°	15.5°
0.71	45.5°	28.3°	44.5°	36.0°	13.8°
0.60	38.3°	23.8°	37.7°	30.5°	11.7°
0.50	31.7°	19.7°	31.5°	25.5°	9.7°
0.40	23.4°	15.8°	25.2°	20.4°	7.8°
0.30	18.9°	11.7°	18.9°	15.3°	5.9°

TABLE XIV

CALCULATION OF SUBSTITUENT DIHEDRAL ANGLES FROM RING BOND ANGLES

See Figure 3 and Figure 4

a-e = P	a-qe = 120 - Q
a-a = 120 + P	b-qa cis = B
e-e = 120 - P	b-qe cis = B
qa-e = Q	b-qa trans = 120 + B
qe-a = Q	b-qe trans = 120 - B
a-qa = 120 + Q	

α-Pseudouridine (α - ψ)

$\Phi(H_1',H_2') = 50°$ $\Phi(H_2',H_3') = 43°$ $\Phi(H_3',H_4') = 158°$

The large cis dihedral angles direct our attention to the table for maximal (0.71Å) puckering. Four pure conformations are within the 26° tolerance for all three angles: C(2')-exo and C(3')-endo envelope conformations, and C(2')-exo-C(1')endo and C(3')endo-C(2')exo twist conformations (Table XV).

TABLE XV

POSSIBLE 0.71Å CONFORMATIONS OF α-ψ

	$\Phi(H_1',H_2')$	$\Phi(H_2',H_3')$	$\Phi(H_3',H_4')$
C(2')exo	45°	45°	148°
C(2')exo-C(1')endo	45°	36°	165°
C(3')endo	28°	45°	165°
C(3')endo-C(2')exo	36°	45°	156°

As the above table shows, the closest correlation is for C(2')-exo. Interestingly, the crystal structure of α-pseudouridine has a C(2')-exo conformation.[105] A cautious conclusion drawn from the above data is that the conformation of α-pseudouridine probably greatly favors C(2')-exo and/or

C(3')-endo conformations (probably with a neighboring atom also puckered to some extent) in its conformational equilibrium, and may exist in one of these conformational ranges exclusively. For 0.5Å total pucker, only C(2')-exo and C(3')-endo are within the 26° tolerance of the dihedral angles of α-pseudouridine determined from nmr coupling constants. The literature conclusion for this nucleoside, based on erroneous tables of dihedral angles, was that no particular conformation was consistent with the nmr data, and that the molecule must therefore be in a rapid conformational equilibrium. This example very effectively illustrates the superiority of the methods recommended in this work.

The dihedral angles calculated from coupling constants (literature data) of β-pseudouridine, uridine, and their 3'-phosphates do not correlate to any conformations in our 0.71Å and 0.5Å tables, even with our liberal tolerance, and these tables provide no grounds to doubt the conclusion of I. C. P. Smith, et al., that these molecules are in conformational equilibrium primarily involving C(3')- and C(2')-puckered conformations.

The approximate coupling constants of 1-(diacetyl-erythrosyl)-uracil in $CDCl_3$ (a.) and 1-(β-D-erythrosyl)-uracil in DMSO-d_6 (b.) and D_2O (c.) are tabulated below along with the corresponding cis or trans dihedral angle calculated from the Karplus equation (Table XVI).

TABLE XVI

COUPLING CONSTANTS AND CALCULATED DIHEDRAL ANGLES OF VIII AND X

	J(1',2')	ϕ(1',2')	J(3',4' cis)	ϕ(3',4' cis)	J(3',4' trans)	ϕ(3',4' trans)
a.	5	138°	4	45°	2	120° or 60°
b.	6	145°	-	-	2	120° or 60°
c.	6	145°	-	-	-	-

For each of these molecules a conformation of closest correlation to the theoretical table can be named. For (a) it is C(3')exo-C(4')endo of 0.71Å total pucker, and for (b.) it is C(2')endo-C(1')exo of 0.5Å total pucker. It would be too presumptive, however, to conclude that these are the conformations of these molecules in solution. The data definitely suggest, however, that a shift in conformational equilibrium occurs when the CH_2OH grouping is eliminated from uridine. The conformation or conformational equilibrium of erythrose nucleosides, like ribonucleosides, probably involves C(2') and C(3')-puckered conformations primarily, to relieve steric strain of the cis-diol moiety. The smaller coupling constant for $J_{3',4'\ trans}$ of erythrosyl uracil and the larger $J_{1',2'}$ are both consistent with a shift of the conformational equilibrium towards C(2')-endo and/or C(3')-exo puckered conformations (probably with an adjacent atom puckered to some extent also).

An interesting correlation can be drawn for conformation of the nucleosides discussed here in solution to the probable preferred orientation of the bulky C(1') and C(4') substituents. Steric strain of these substituents with substituents on next nearest ring atoms would cause conformations in which they occupied equatorial or quasi-equatorial positions to be favored. Thus C(2')-endo and C(3')endo conformations would be favored over C(2')-exo and C(3')-exo conformations

for β-ribonucleosides. Similarly, C(2')-exo and C(3')-endo conformations could be preferred for α-ribonucleosides on these ground (the eliminated conformations would place one of the two bulkiest substituents of the molecule in a quasi-axial position). Since erythrose nucleosides lack the bulky C(4') substituent, C(3')-exo would now not be disfavored by the suggested interactions. Furthermore, C(2')-endo might be favored over all other conformations since the heterocycle would be quasi-equatorial and no hydroxyl group would be axial, and C(2')-exo would be disfavored because of the quasi-axial heterocycle and the axial hydroxyl group. All of the above is consistent with the conclusions drawn from nmr coupling constants about conformations of δ-pseudouridine, β-erythrosyl uracil, uridine, -pseudouridine, and the 3'-phosphates of the latter two molecules. A study of the direction of bond angle distortion in crystal structure of the fourteen nucleosides studied in this project did not indicate that such steric interactions with substituents on next-nearest ring atoms were an important factor; however, the situation may be quite different in solution since these bulky substituents may freely rotate. The commonly observed C(3')-exo conformation in crystal structures of deoxyribonucleosides might be explained by relief of steric strain between H(6) and H(3') in the preferred anti conformation, together with the decreased overall steric strain in the molecules (compared to ribonucleosides).

The tables of theoretical angles were used to check the assumption used in our structural assignments from nmr data that a coupling constant of 3 Hz or less suggests, but does not prove, trans configuration of the protons. The largest angle for cis substituents is 45.5° and is from the data for envelope conformations of 0.71Å pucker. The Karplus equation yields a corresponding coupling constant of 3.9 Hz, the smallest theoretical coupling constant for cis- substituents. Thus the Karplus equation must be erroneous or the dihedral angle must be distorted in order for a cis coupling constant to be 3 Hz or less. Applying the 26° tolerance to the maximum dihedral angle of 45.5° gives a minimum probable value of 0.7 Hz for a cis coupling constant. The minimum angle which can give a 3 Hz coupling constant after allowing a 26° tolerance is 26°. Thus a cis coupling constant of less than 3 Hz can be occasionally expected, and a small coupling constant is not final proof of trans configuration.

A re-examination of the nmr spectra of a series of molecules prepared in this laboratory by J. P. Cleveland was appropriate for this project. The compounds are the 3,4-0,0-cyclic carbonate, sulfite, borate, and isopropylidene derivatives of cis-3,4-tetrahydrofurandiol. The nmr spectra of these compounds are qualitatively very similar. A downfield second order multiplet is unambiguously assigned to H(3) and H(4) on the basis of electronegativity considerations and

on a relatively large change in chemical shift as the substituents are varied. The remainder of the spectrum of each molecule consists of a second order AB pair; the downfield part of this pair exhibits a coupling constant near zero, and the upfield portions exhibit a coupling constant in the medium range (these are in addition to the large geminal coupling constant). On the basis of criteria used in the synthetic part of this work, the downfield part of the AB pair should be assigned to the H(2) and H(5) protons trans to the H(3) and H(4) protons, respectively based on coupling constant; on the basis of chemical shift, however, the downfield position of the AB pair should be assigned to the cis H(2) and H(5) protons. It is difficult to resolve the ambiguity, and this is an example which demonstrates why neither criteria alone was accepted as proof of structure in the experimental section. The spectrum of the carbonate has been computer analyzed by Cleveland and the data below was obtained from his thesis:

$J(3,4) \doteq 7.0 \pm 0.2$ Hz

$J(2,3 \text{ trans}) = J(3,4 \text{ trans}) = 0.35 \pm 0.2$ Hz

$J(2,3 \text{ cis}) = J(3,4 \text{ cis}) = 3.5 \pm 0.2$ Hz

The reinterpretation of the data is as follows. Pucker at the C(3) and C(4) positions were ruled out for steric reasons. Possible conformations are given below. A large

TABLE XVII A

CALCULATED DIHEDRAL ANGLES FOR LIKELY CONFORMATIONS OF THE BICYCLIC DERIVATIVES OF CIS-3,4-TETRAHYDROFURANDIOL

	Φ(2,3 cis)	Φ(2,3 trans)	Φ(3,4)
	0.71Å		
C(2)-endo ⇄ C(5)-endo	23°	143°	28°
C(2)-exo ⇄ C(5)-exo	23°	97°	28°
O-endo	28°	92°	0°
O-exo	28°	148°	0°
	0.50Å		
C(2)-endo ⇄ C(5)-endo	16°	136°	20°
C(2)-exo ⇄ C(5)-exo	16°	104°	20°
O-endo	20°	100°	0°
O-exo	20°	140°	0°
	0.30Å		
C(2)-endo ⇄ C(5)-endo	9°	129°	12°
C(2)-exo ⇄ C(5)-exo	9°	111°	12°
O-endo	12°	114°	0°
O-exo	12°	126°	0°
Carbonate	48° (75°-105°?)	75°-105° (129°?)	22°

extent of puckering was originally ruled out on the basis of the 0.31Å total pucker of the 2',3'-O,O-cyclophosphorothioate of uridine in the crystal structure.[113] Data for likely conformations of 0.3Å and 0.5Å and 0.71Å pucker are given below.

Comparison of the spectrum of the carborate quickly ruled out the possibility of the 0.35 Hz coupling constant belonging to cis protons, since the minimum deviation would be 47°. The conformation appears to be either O-endo or C(2)-exo⇄C(5)-exo of at least 0.5Å pucker (with a neighboring atom also puckered to some extent, probably). Molecular models and the x-ray structure of the nucleoside cyclophosphorothioate both favor an O-endo interpretation. A distortion of 22° from the eclipsed theoretical orientation of H(2) and H(3) and H(4) and H(5) is not at all unexpected on the basis of X-ray structures of nucleosides, and might also be explained by the likelihood of an intermediate conformation closely related to O-endo. The actual conformation is probably an intermediate conformation somewhere between O-endo and C(2)-exo⇄C(5)-exo.

Since it is probably more accurate to take the average of the coupling constants than of the dihedral angles to get the theoretical coupling constant of a conformational equilibrium, the equilibrium conformations of the above tables were reanalyzed on that basis in the tables below. The table based on average coupling constant does not lead to any change in the conclusion drawn on the tables based on average dihedral angles in conformational equilibria.

TABLE XVII B

DIHEDRAL ANGLES OF POSSIBLE CONFORMATIONAL EQUILIBRIA OF BICYCLIC DERIVATIVES OF CIS-3,4-TETRAHYDROFURANDIOL CALCULATED FROM AVERAGE COUPLING CONSTANT

	Φ(2,3 cis)	Φ(2,3 trans)	Φ(3,4)
		0.71Å	
C(2)-endo $\rightleftarrows$ C(5)-endo	6.0 (31°)	5.4 (142°)	28°
C(2)-exo $\rightleftarrows$ C(5)-exo	6.0 (31°)	1.2 (113°)	28°
		0.50Å	
C(2)-endo $\rightleftarrows$ C(5)-endo	7.0 (22°)	4.6 (136°)	20°
C(2)-exo $\rightleftarrows$ C(5)-exo	7.0 (22°)	1.2 (113°)	20°
		0.30Å	
C(2)-endo $\rightleftarrows$ C(5)-endo	7.8 (12°)	3.6 (130°)	0°
C(2)-exo $\rightleftarrows$ C(5)-exo	7.8 (120°)	1.1	0°

The following suggestion is presented to explain the unusual downfield position of the protons cis to O(3) and O(4). In the bicyclic tetrahydrofuran-3,4-diol derivatives, O-endo conformation places the H-1 and H-4 protons which are trans to O-4 and O-3, respectively, in quasi-axial positions (Figure 29). Similarly, the conformational equilibrium 1-exo ⇄ 4-exo places these protons in axial ⇄ bisectional positions. Perhaps in the relatively fixed conformations of these bicyclic derivatives, ring-bond diamagnetic anisotropies exceed the shielding effect of cis-oxygen substituents, as they do in pyranose carbohydrates. At any rate, these quasi-axial protons of the tetrahydrofurandiol derivatives are more shielded than the quasi-equatorial protons, just as axial protons of pyranose carbohydrates are more shielded than equatorial protons.

The cyclic monophosphate derivative of β-D-erythrosyl uracil also exhibits an anomaly in the absorptions of H-4a and H-4b. In this case these protons have nearly the same chemical shift, while in β-D-erythrosyl uracil and its diacetate H-4a is upfield from H-4b. This compound might be expected to have O-exo conformation (analogous to O-endo in the tetrahydrofurandiol derivatives) on the basis of similarity to uridine-2',3'-cyclophosphorothioate and on its small $J_{1',2'}$. Thus the quasi-axial protons are more shielded in

this fixed conformation relative to the quasi-equatorial protons than in the non-cyclic nucleotides and nucleosides.

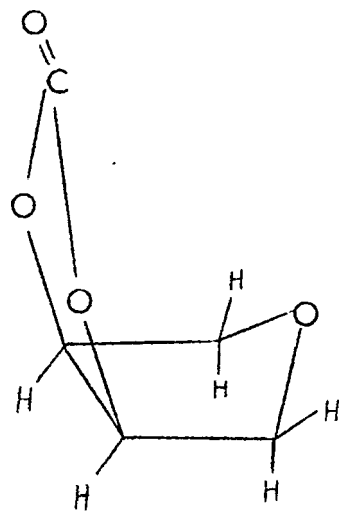


FIGURE 29

CIS-3,4-TETRAHYDROFURANDIOL-3,4-CYCLIC CARBONATE

IN O-ENDO CONFORMATION

BIBLIOGRAPHY

1. J. J. Fox and I. Wempen, Advances in Carbohydrate Chemistry, 14, 283-380 (1959).

2. J. A. Montgomery and H. J. Thomas, Advances in Carbohydrate Chemistry, 17, 301-370 (1962).

3. J. Pliml and M. Prystas, Advances in Heterocyclic Chemistry, 8, 115-142 (1967).

4. W. W. Zorbach and R. S. Tipson, editors, Synthetic Procedures in Nucleic Acid Chemistry, Volume 1, Wiley, New York (1968).

5. G. L. Matlock and G. O. Phillips, J. Chem. Soc., 1836 (1956).

6. T. L. V. Ulbricht and G. T. Rogers, J. Chem. Soc., 6125 (1965); ibid., 6130 (1965).

7. J. A. Montgomery and H. J. Thomas, J. Org. Chem., 31, 1411 (1966).

8. G. T. Rogers, R. S. Shadbolt, and T. L. V. Ulbricht, J. Chem. Soc., 203 (1969); ibid., 209 (1969).

9. B. R. Baker, R. E. Schaub, and H. M. Kissman, J. Am. Chem. Soc., 77, 5911 (1955).

10. D. H. Murray and J. Prokop, J. Pharm. Sci., 56, 867 (1967).

11. K. A. Watanabe and J. J. Fox., J. Heterocyclic Chem., 6, 109 (1969).

12. G. T. Rogers and T. L. V. Ulbricht, J. Chem. Soc., 2450 (1969).

13. See reference 4 for several examples of reaction conditions.

14. B. Capon, I. C. Collins, A. A. Lewy, and W. G. Overend, J. Chem. Soc., 3243 (1964).

15. G. T. Rogers and T. L. V. Ulbricht, Tetrahedron Lett., 1025 (1968).

16. M. Prystas and F. Sorm, Collection Czech. Chem. Commun., 30, 2960 (1965).

17. M. Prystas and F. Sorm, Collection Czech. Chem. Commun., 31, 1035 (1966).

18. G. Wagner and H. Pischel, Arch. Pharm., 295, 373 (1962); ibid., 296, 699 (1963).

19. L. Birkofer, A. Ritter, and H. P. Kühlthan, Chem. Ber., 97, 934 (1964).

20. L. Birkofer and A. Ritter, Angew. Chem., 77, 414 (1965).

21. J. Davoll, B. Lythgoe, and A. R. Todd, J. Chem. Soc., 833 (1946); V. M. Clark, A. R. Todd, and J. Zussman, J. Chem. Soc., 2952 (1951).

22. J. Farkas, L. Kaplan, and J. J. Fox, J. Org. Chem., 29, 1469 (1964).

23. T. Y. Shen, H. M. Lewis, and W. V. Ruyle, J. Org. Chem. 30, 835 (1965).

24. U. Niedballa and H. Vorbrüggen, Angew. Chem. Internat. Edit., 9, 461 (1970).

25. D. H. Warnock, K. A. Watanabe, and J. J. Fox, Carbohydrate Research, 18, 127 (1971).

26. P. deMayo, Molecular Rearrangements, Wiley, New York, 710-711 (1964).

27. W. E. Willy, G. Binsch, and E. L. Eliel, J. Am. Chem. Soc., 92, 5394 (1970).

28. M. Sundaralingam, Biopolymers, 7, 821-860 (1969).

29. J. Kraut and L. H. Jensen, Acta Cryst., 16, 79 (1963).

30. M. Karplus, J. Am. Chem. Soc., 85, 2870 (1963).

31. C. D. Jardetzky and M. Smith, J. Molecular Spectroscopy, 28, 70 (1968).

32. C. D. Jardetzky, J. Am. Chem. Soc., 84, 62 (1962).

33. C. D. Jardetzky, J. Am. Chem. Soc., 83, 2919 (1961).

34. C. D. Jardetzky, J. Am. Chem. Soc., 82, 229 (1960).

35. I. C. P. Smith, F. E. Hruska, and A. A. Grey, J. Am. Chem. Soc., 92, 4088 (1970).

36. A. A. Grey, I, C. P. Smith, and F. E. Hruska, J. Am. Chem. Soc., 93 1765 (1971).

37. M. P. Schweizer, J. T. Witkowski, and R. K. Robin, J. Am. Chem. Soc., 93, 277 (1971).

38. M. P. Schweizer, A. P. Broom, P. O. P. Ts'O, and D. P. Hollis, J. Am. Chem. Soc., 90, 1042 (1968).

39. D. W. Miles, M. J. Robins, R. K. Robins, M. W. Winkley, and H. Eyring, J. Am. Chem. Soc., 91, 824 (1969); ibid., 91, 831 (1969).

40. Eberhard Harbors, Introduction to Nucleic Acids, Reinhold, New York, 212-215 (1968).

41. M. Kunitz, J. Gen. Physiol., 24, 15 (1940).

42. A. J. P. Martin, R. R. Porter, Biochem. J., 49, 215 (1951).

43. E. J. del Rosario and G. G. Hammes, J. Am. Chem. Soc., 92, 1750 (1970).

44. D. M. Brown and A. R. Todd, in E. Chargoff and J. N. Davidson, editors, The Nucleic Acids, Academic Press, New York (1955).

45. H. Witzel and E. A. Barnard, Biochem., Biophys Res. Commun., 7, 295 (1962).

46. R. J. Hilmoe, L. A. Heppel, S. S. Springhorn, and D. E. Koshland, Jr., Biochem. Biophys. Acta, 53, 214 (1961).

47. F. Melchers, D. Dutting, and H. G. Zachau, Biochem. Biophys. Acta, 108, 182 (1965).

48. H. Feldman, D. Dutting, and H. G. Zachau, Z. Physiol. Chem., 347, 236 (1966).

49. R. W. Holley, G. A. Everett, J. T. Madison, and A. Zamir, J. Biol. Chem., 240, 2122 (1965).

50. F. A. Neelon, M. Molinaro, H. Ishikura, L. B. Sheiner, and G. L. Cantoni, J. Biol. Chem., 242, 4515 (1967).

51. D. W. Verwoerd and W. Zillig, Biochem. Biophys. Acta, 68, 484 (1963).

52. H. Witzel and E. A. Barnard, Biochem. Biophys. Res. Commun., 7, 295 (1962).

53. E. A. Barnard, Ann. Rev. Biochem. 38, 677 (1969).

54. W. Gamble and L. D. Wright, Biochem. J., 97, 340 (1965).

55. M. Yamasaki and K. Arima, Biochem. Biophys. Acta, 139, 202 (1967).

56. G. G. Hammes, Accounts of Chem. Res., 1, 321 (1968).

57. G. G. Hammes and F. G. Walz, Jr., J. Am. Chem. Soc., 91, 7179 (1969).

58. D. A. Usher, Proc. Natl. Acad. Sci. U. S., 62, 661 (1969).

59. D. A. Blow and T. A. Steitz, Ann. Rev. Biochem., 39, 63 (1970).

60. G. Kartha in E. A. Barnard, Ann. Rev. Biochem., 38, 718 (1969).

61. H. A. Scheraga and J. A. Rupley, Adv. in Enzymology, 24, 161 (1962).

62. A. M. Crestfield, W. H. Stein, and S. Moore, in H. A. Scheraga and J. A. Rupley, Adv. Enzymology, 24, 214 (1962).

63. A. M. Crestfield, W. H. Stein, and S. Moore, J. Biol. Chem., 238, 2413 (1963).

64. C. A. Ross, A. P. Mathias, and B. R. Rabin, Biochem. J., 85, 145 (1962).

65. W. A. Klee and F. M. Richards, J. Biol. Chem., 229, 489 (1957).

66. C. H. W. Hirs, M. Halmann, and J. H. Kycier in T. W. Goodwin and O. Lindberg, editors, Biological Structure and Function, Vol. I, Academic Press, 41 (1961).

67. G. Kartha, J. Bello, and D. Harker, Nature, 213 (1967).

68. H. W. Wyckoff, K. D. Hardman, N. M. Allewell, T. Inagami, L. N. Johnson, and F. M. Richards, J. Biol. Chem., 242, 3984 (1967).

69. F. Haurowitz, The Chemistry and Function of Proteins, Academic Press (1963).

70. H. P. Avey, M. O. Boles, C. H. Carlisle, S. A. Evans, S. J. Morris, R. A. Palmer, B. A. Woolhouse, and S. Shall, Nature, 213, 557 (1967).

71. R. C. Hockett, D. V. Collins and A. Scattergood, J. Am. Chem. Soc., 73, 599 (1951).

72. E. J. Bourne, G. T. Bruce and F. L. Wiggins, J. Chem. Soc., 2708 (1951).

73. A. C. Neish, Can. J. Chem., 32, 334 (1954).

74. A. S. Perlin, Methods in Carbohydrate Chemistry, Vol. I, Academic Press, New York, 64 (1962).

75. R. Schaffer, J. Am. Chem. Soc., 81, 2838 (1959).

76. D. A. Rappoport and W. Z. Hassid, J. Am. Chem. Soc., 73, 5524 (1951).

77. I. Iwai, T. Neshimura and B. Shimigu, in W. W. Zorbach and R. S. Tipson, editors, Synthetic Procedures in Nucleic Acid Chemistry, Vol. I, Wiley, New York, 388 (1969).

78. J. J. Fox, N. Yung, I. Wempen and I. L. Doerr, J. Am. Chem. Soc., 79, 5060 (1957).

79. T. L. V. Ulbricht and G. T. Rogers, J. Chem. Soc., 6125 (1965).

80. R. U. Lemieux, R. K. Kullnig, H. J. Bernstein and W. G. Schneider, J. Am. Chem. Soc., 80, 6098 (1958).

81. J. F. Stoddart, Stereochemistry of Carbohydrates, Wiley, New York (1971).

82. R. U. Lemieux, R. K. Kullnig, H. J. Bernstein and W. G. Schneider, J. Am. Chem. Soc., 80, 6098 (1958).

83. H. G. Fletcher, Jr. and J. D. Stevens, J. Org. Chem., 33, 1799 (1968).

84. E. Pascu, Ber., 61, 1508 (1928).

85. L. J. Haynes and F. H. Newth, Advances in Carbohydrate Chem., 10, 207-256 (1955).

86. L. M. Lerner, J. Org. Chem., 34, 101 (1969).

87. R. U. Lemieux, Advances in Carbohydrate Chem., 9, 1 (1954).

88. A. Holy and J. Smrt, Collection Czech. Chem. Commun., 30, 1529 (1965).

89. A. Holy, J. Smrt, and F. Sorm, ibid., 32, 2987 (1967).

90. A. Holy, J. Smrt, and F. Sorm, ibid., 33, 3818 (1968), and numerous additional papers in that journal.

91. F. A. L. Anet, Can. J. Chem., 39, 789 (1961).

92. R. J. Abraham, L. D. Hall, L. Hough and K. A. McLauchlan, J. Chem. Soc., 3699 (1962).

93. B. Coxon, Carbohydrate Research, 13, 321 (1970).

94. S. I. Chan and J. H. Prestegard, J. Am. Chem. Soc., 91, 2843 (1969).

95. J. Zussman, Acta Crystallogr., 6, 504 (1953).

96. R. E. Schirmer, J. H. Noggle, J. P. Davis and P. A. Hart, J. Am. Chem. Soc., 92, 3266 (1970).

97. D. J. Hunt and E. Subramanian, Acta Crystallogr., B25, 2144 (1969).

98. E. Shefter, M. Barlow, R. A. Sparks and K. N. Trueblood, Acta Crystallogr., B25, 895 (1969).

99. S. Furberg, C. S. Peterson, and C. Romming, Acta Crystallogr., 18, 313 (1965).

100. G. H. Lin and M. Sundaralingam, Acta Crystallogr., B27, 961 (1971).

101. S. W. Hawkinson and C. L. Coulter, Acta Crystallogr., B27, 34 (1971).

102. D. W. Young, P. Tollin and H. R. Wilson, Acta Crystallogr., B25, 1423 (1969).

103. E. Subramanian and D. J. Hunt, Acta Crystallogr., B26, 303 (1970).

104. D. C. Rohrer and M. Sundaralingam, J. Am. Chem. Soc., 92, 4956 (1970).

105. D. C. Rohrer and M. Sundaralingam, J. Am. Chem. Soc., 92, 4950 (1970).

106. M. Sundaralingam and L. H. Jensen, J. Mol. Biol., 13, 914 (1965).

107. C. E. Bugg and R. E. Marsh, J. Mol. Biol., 25, 67 (1967).

108. D. R. Harris and W. M. Macintyre, Biophys. J., 4, 203 (1964).

109. D. G. Watson, D. J. Suter and P. Tollin, Acta Crystallogr., 19, 111 (1965).

110. T. Schleich, B. J. Blackburn, R. D. Lapper and I. C. P. Smith, Biochemistry, 11, 137 (1972).

111. A. A. Grey, I. C. P. Smith and F. E. Hruska, J. Am. Chem. Soc., 93, 1765 (1971).

112. J. P. Cleveland, Master's Thesis, Georgia Tech (1965).

113. W. Saenger and F. Eckstein, J. Am. Chem. Soc., 92, 4712 (1970).