Spring 1980

A MOLECULAR ORBITAL STUDY OF 5-FLUOROURACIL AND RELATED COMPOUNDS

FRANK BLOCK

BLOCK, FRANK

A MOLECULAR ORBITAL STUDY OF 5-FLUOROURACIL AND RELATED COMPOUNDS

University of New Hampshire PH.D. 1980

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ✓.

1. Glossy photographs _____
2. Colored illustrations _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Print shows through as there is text on both sides of page _____
6. Indistinct, broken or small print on several pages ✓
7. Tightly bound copy with print lost in spine _____
8. Computer printout pages with indistinct print _____
9. Page(s) _____ lacking when material received, and not available from school or author
10. Page(s) _____ seem to be missing in numbering only as text follows
11. Poor carbon copy _____
12. Not original copy, several pages with blurred type _____
13. Appendix pages are poor copy _____
14. Original copy with light type _____
15. Curling and wrinkled pages _____
16. Other _____

University Microfilms International

300 N. ZEEB RD., ANN ARBOR MI 48106 (313) 761-4700

A MOLECULAR ORBITAL STUDY OF 5-FLUOROURACIL AND RELATED COMPOUNDS

BY

Frank Block
B.A., Brooklyn College, 1940
M.A., Columbia University, 1951

DISSERTATION

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy
in
Chemistry

May, 1980

This dissertation has been examined and approved.

Frank L. Pilar

Dissertation Director, Frank L. Pilar
Professor of Chemistry

Paul R. Jones

Paul R. Jones, Professor of Chemistry

Charles W. Owens

Charles W. Owens, Professor of Chemistry

James H. Weber

James H. Weber, Professor of Chemistry

Gerald L. Klippenstein

Gerald L. Klippenstein, Professor of Biochemistry

10 April 1980

Date

ACKNOWLEDGEMENT

I wish to express my deep appreciation for the guidance given me by my research and thesis advisor, Dr. Frank L. Pilar. I also wish to extend my thanks to Dr. Masahiro Inatome for the use of a CNDO/2 program deck and for his assistance in entering the data into the computer. Additional thanks are due to Mr. Robert E. Amos for the fine cooperation I received at the computer facility. I am also much indebted to my good friend, Dr. George T. Davis, for the many helpful discussions on reaction mechanisms and for his encouragement and constructive criticism. Finally, I wish to thank Mrs. Monica Tyler for having provided valuable secretarial assistance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

ABSTRACT

A MOLECULAR ORBITAL STUDY OF

5-FLUOROURACIL AND RELATED COMPOUNDS

by

FRANK BLOCK

University of New Hampshire, May, 1980

5-fluorouracil is a chemotherapeutic agent which is used in the treatment of disseminated colon and breast cancers. As an antimetabolite it has been found to inhibit the mechanism in which the normal substrate reacts with the enzyme and a cofactor in nucleic acid synthesis. A number of mechanisms have been summarized describing the conversion of deoxyuridylate (dUMP) to deoxythymidylate (dTMP), elucidating the role of the involved enzyme, thymidylate synthetase, and the co-factor, N^5, N^{10}-methylenetetrahydrofolate. The effect of replacing the hydrogen bonded to the 5-position of the pyrimidine ring by a fluorine atom is also mechanistically described. A molecular orbital study employing the CNDO/2 approximation was performed on 5-fluorouracil, uracil, and thymine and the results support the mechanisms and chemical effects drawn from the literature which center around the bond polarization of the 5,6 double bond of uracil. The electronic effects of fluorine substitution (in 5-fluorouracil) are also consistent with formation of a more stable enzyme-substrate complex and with favorable kinetic competition of 5-fluorouracil with uracil for the enzymatic nucleophilic site.

SECTION I

INTRODUCTION

Since nucleic acids play a principal role in the division and growth of the neoplastic cell, scientific research has addressed itself to the discovery of chemical substances which alter or inhibit the mechanisms controlling malignant cell division. Such substances are antimetabolites - structural analogs of normal metabolites for specific enzymes. The normal metabolite becomes displaced so the enzyme cannot carry out its normal function in nucleic acid synthesis. Antimetabolites may also interfere with normal cellular metabolism by fraudulent incorporation as a building unit.

SECTION II

BIOCHEMICAL CONSIDERATIONS OF 5-FLUOROURACIL AND RELATED COMPOUNDS

Introduction

In 1954 Rutman and co-workers[1] reported an increased utilization of uracil (Fig. 1) for nucleic acid biosynthesis in a rat liver tumor. Two years later Cohen and Barner[2] conducted studies of "thymineless death" in E. coli and suggested the use of compounds in cancer chemotherapy which would inhibit deoxyribonucleic acid (DNA) synthesis. Thymineless death refers to cell death caused by unbalanced cell growth due to DNA starvation. DNA fails to be synthesized when deoxythymidylic acid (Fig. 1) is not produced; however the synthesis of protein and ribonucleic acid (RNA) are not affected. Consequently, cellular growth is not inhibited, but normal cell division is impaired by absence of DNA. In 1957 Heidelberger[3] demonstrated that 5-fluorouracil (Fig. 1) inhibited the conversion of C^{14} -labelled formate into the methyl group in the biosynthesis of thymine (Fig. 1). The resulting paucity of thymine was then shown by the same investigator[4] to have a significant effect toward inhibiting animal tumor growth. He established that 5-fluorouracil (5-FU) was incorporated in the nucleic acids but at that time had not yet determined whether it was in the DNA or the RNA. It was later demonstrated that 5-FU was incorporated into RNA (and not into DNA) in both mouse tissues and a human neoplasm[5].

Theory and Method

A. Structural Characteristics of the Antimetabolite

The rationale for choosing 5-FU was based upon the great stability of the carbon-fluorine bond. The small Van der Waal's radius of the fluorine atom (1.35 Å), which is nearly that of hydrogen, produces a molecule nearly isosteric with uracil[6]. The 5-position was chosen for substitution on the pyrimidine ring because of its structural similarity to thymine. It was felt that these characteristics would facilitate the ability of the antimetabolite to occupy active sites on enzyme and that it might interfere with either the synthesis of thymine or its incorporation into the DNA.

B. Inhibitory Effect of 5-Fluorouracil and 5-Fluoro-2'-deoxyridine

Further studies corroborated the finding that 5-FU inhibited both the conversion of uracil 2-C^{14} into DNA thymine and, to a lesser extent, RNA uracil in mice-bearing transplants of Ehrlich ascites carcinoma[7]. It also completely inhibited the conversion of C^{14} formate into the methyl group of DNA thymine in spleen and tumor.

Since 5-FU and 5-fluoro-2'-deoxyuridine (Fig. 1) both inhibit the incorporation of uracil 2-C^{14} into DNA thymine - but not significantly into RNA uracil when administered in chemotherapeutic doses to intact animals bearing malignant tumors - it follows that the inhibitory effect of fluorinated pyrimidines on the synthesis of DNA is responsible for the antimetabolic behavior[8].

C. Metabolism of 5-Fluorouracil

The biosynthesis of deoxyuridylic acid (dUMP) and its conversion to deoxythymidylic acid (dTMP) is amply described in biochemistry texts[9,10].

The methylation takes place in the presence of thymidylate synthetase and N^5, N^{10}— methylenetetrahydrofolate (which is oxidized to the dihydrofolate). However, when 5-FU is introduced and competes with the normal metabolite (Fig. 2), the synthesis of DNA is inhibited. 5-FU is converted to 5-fluorouridine (5-FUR) (Fig. 1) followed by phosphorylation to the mono-di-, and tri-phosphates (Fig. 2). This incorporation into the nucleic acid pool is the source of the aberrant RNA. The monophosphate (5-FURMP) is also reduced to 5-fluoro-2' - deoxyuridylic acid (FUdRMP)*. The latter may also be formed directly by phosphorylation of 5-fluoro-2' - deoxyuridine (FUdR). FUdRMP does not phosphorylate any further. It is the FUdRMP which is considered to be responsible for the antineoplastic activity of 5-FU by inhibiting DNA synthesis via thymidylate synthetase blockage[11, 12]. The catabolism of 5-FU is similar to that of uracil (Fig. 2). Consequently, normal degradation of FUdR and 5-FU occur by processes analagous to uracil (Fig. 2), so that effects from the breakdown of these materials are not implicated in the observed antineoplastic activity.

D. Unique Potential for Thymidine Interdiction

The novelty and uniqueness of thymidine nucleotide as opposed to the other nucleic acids reside in its singular pathway of formation. That is, in contrast to all other deoxribonucleotides of DNA, it is not synthesized by phosphorylation and reduction of the ribonucleoside monophosphate. Rather, it is synthesized by "modification of the pyrimidine base". Therefore, successful research thrust in cancer chemotherapy has centered around analogs of precursors to deoxythymidylic acid. Further, the detailed findings related to the mechanisms of action

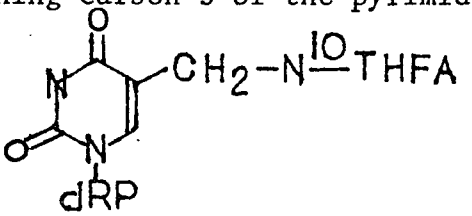
---

*The reaction proceeds via a multi-step pathway involving a reduction to the deoxyribose form at the diphosphate level.

of the chemotherapeutic precursor antimetabolite fluorouracils are substantially explained in the observation concerning the mechanism of thymidine-forming reactions catalyzed by thymidylate synthetase.

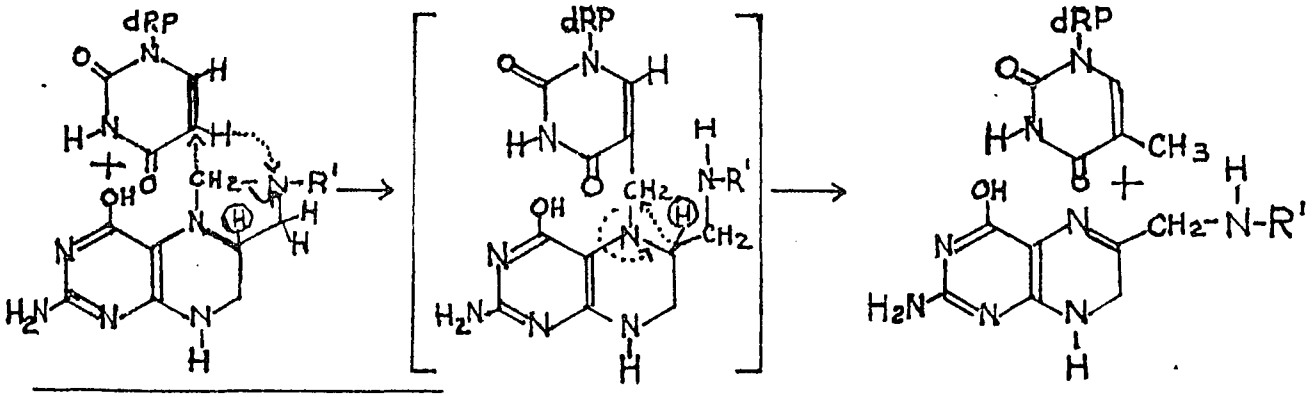
Results

Aside from the results of quantitative studies of enzyme kinetics and radioactive tracers[12-25] a number of mechanisms are available in the literature and may be briefly summarized as follows: Friedkin and Kornberg[18] postulated a mechanism involving the formation of an intermediate containing a methylene bridge between deoxyuridylate and tetrahydrofolate joining carbon 5 of the pyrimidine and nitrogen 10 of the pteridine*



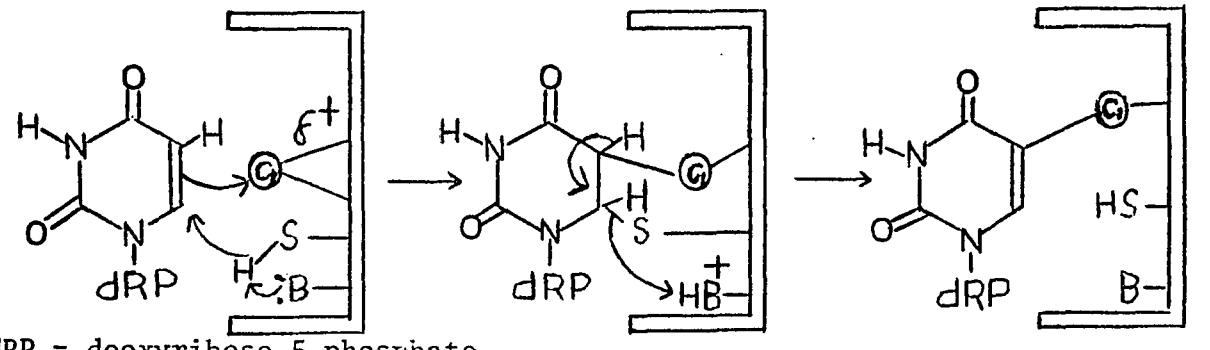
dRP= deoxyribose-5-phosphate

This would be followed by an intramolecular rearrangement and reductive cleavage, giving rise to dihydrofolate and thymidylate. Pastore and Friedkin[19] then established experimentally that tritium is transferred from tetrahydrofolate to the methyl carbon of the pyrimidine. Consistent with Friedkin's mechanism and based on labeling experiments, Lorenson et al[20] suggested an intramolecular hydride ion transfer from the 6 position of the pteridine.



*In a later paper (see ref. 25) the methylene bridge was postulated to be between carbon 5 of the pyrimidine and nitrogen 5 of the pteridine.

Kalman[26] suggested a mechanism consistent, but not compelling, with active participation of a sulfhydryl group in the reaction catalyzed by thymidylate synthetase. His reasoning was based upon catalyzed label scrambling at position 5 of deoxyuridylate, observed by the mock "sulfhydryl enzyme", glutathione. The action of the active site was pictured as follows:

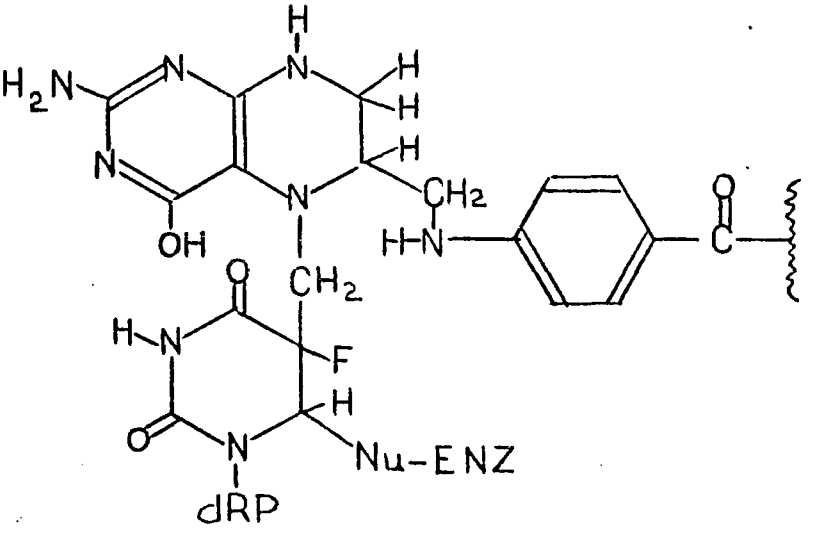


dRP = deoxyribose-5-phosphate

B - basic group of the enzyme

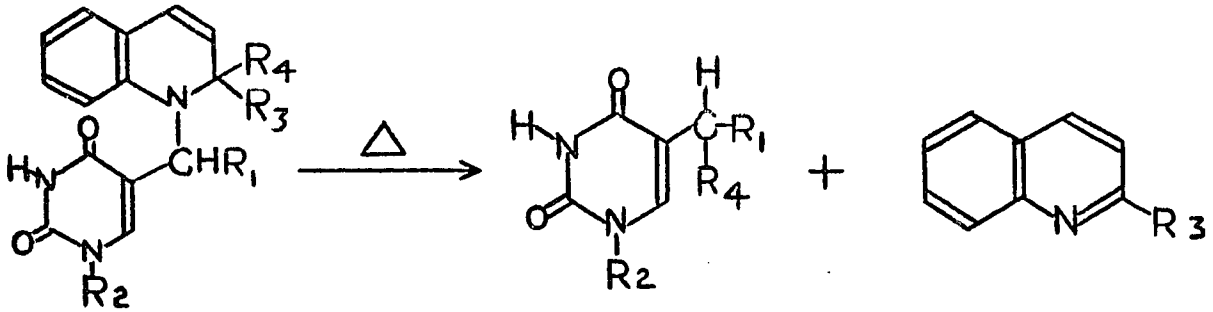
C = one carbon unit

Langenbach et al[27] concluded on the basis of: a. methylene-tetrahydrofolate requirement and b. extreme stability of the ternary enzyme-methylenetetrahydrofolate-fluorouridylate complex (indicating covalent bonds on the ternary complex) that a product of the following type exists:



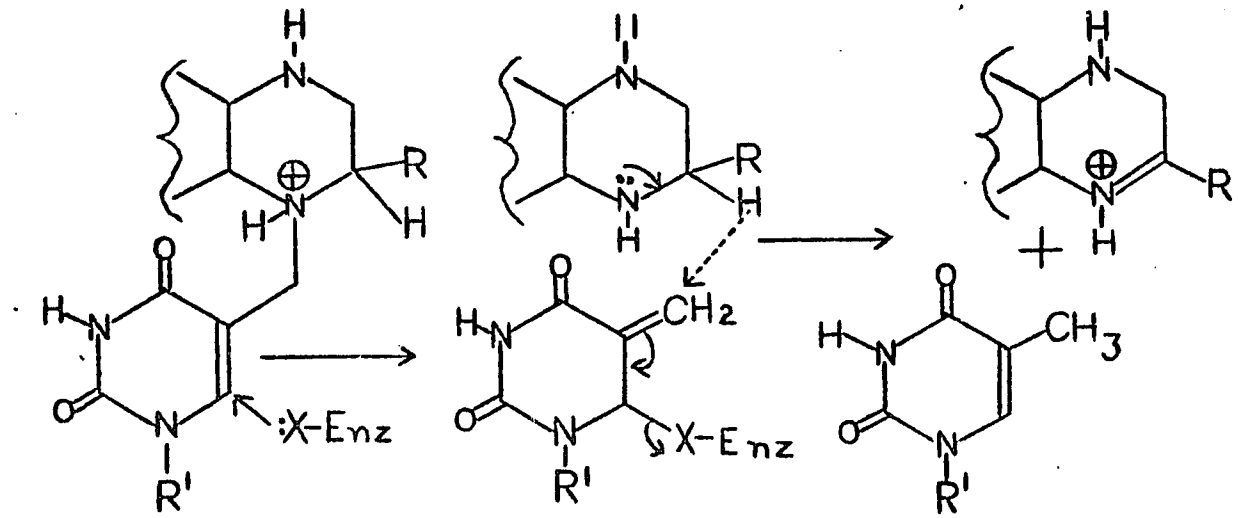
This type of intermediate was also postulated for the normal enzyme reaction.

Wilson and Mertes[28] worked with chemical analogs (models) of the binary complex of uridylic acid-methylenetetrahydrofolate. They showed that pyrolysis <u>in vacuo</u> (200°C) or reflux in high boiling solvents yielded thymine analogs by "hydride shifts", e.g.



1. $R_1=R_2=R_3=R_4=H$
2. $R_1=R_3=R_4=H;R_2=CH_3$
3. $R_1=R_2=H;R_3=R_4={}^2H$
4. $R_2=R_3=R_4=H;R_1={}^2H$
5. $R_1=R_2=R_4=H$
6. $R_1=R_4=H;R_2=CH_3$
7. $R_1=R_2=H;R_4={}^2H$
8. $R_2=H;R_1=R_4={}^2H$
9. $R_3=H$
10. $R_3={}^2H$

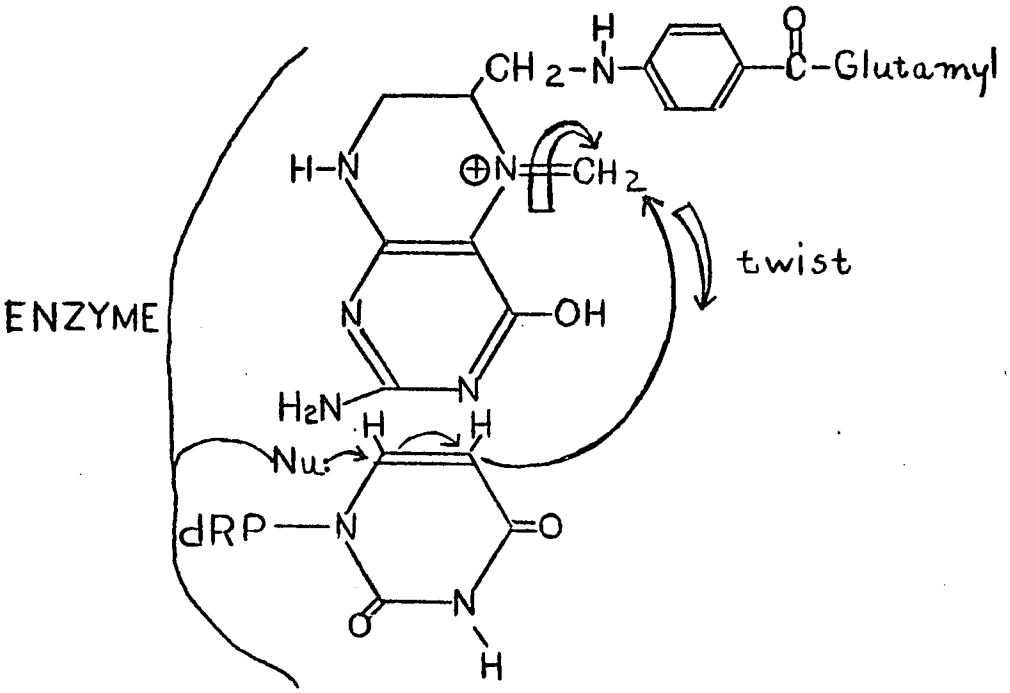
Consequently, they proposed a "hydride shift" mechanism for the action of thymidylate synthetase based on an enzyme assist to this "hydride transfer" on the preformed binary complex of uridylate and methylene-tetrahydrofolate as follows:



R = CH_2-NH-C_6H_4-CO-glutamyl. R' = 1-(5-phospho-2'-deoxyribosyl).

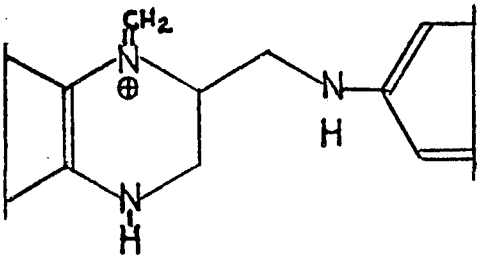
Labeling in their system gave results suggesting both intramolecular and intermolecular rearrangements. (The situation may, however, be different in the less thermal enzyme reaction).

They proposed an alternative, less credible mechanism for formation of the complex involving parallel aromatic π - complexing of an N--formiminium tetrahydrofolate ionic intermediate with dUMP followed by reaction with the enzyme to "rehybridize" the methylene on ternary complex formation, e.g.

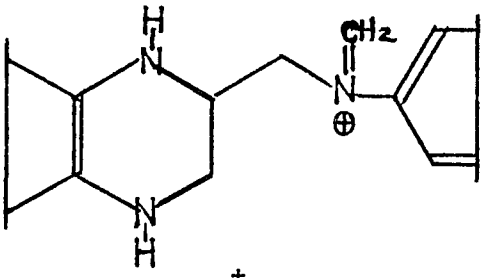


Sommer and Santi[29], based on amino acid analysis of an active site peptide from thymidylate synthetase, concluded that the nucleophilic functional group of the enzyme must be threonine hyroxyl or histidine nitrogen. They speculate poorly about the mechanism. If their results can be considered to be indicative, the histidine attack seems more likely, since the intermediate would then more readily decompose to products (histidine providing a more active intermediate for displacement than threonine). The prosthetic group sequence was Thr, His, Ala, Leu, Pro_2. However, their results do not seem to have received either prior or subsequent support.

Santi, McHenry and Sommer[30] stressed the possibility of FdUMP binding to enzyme which in some way "stimulates" build-up of an iminium cation, viz:

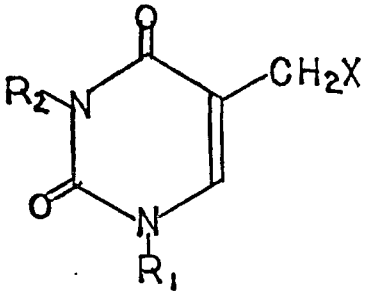


$5\text{-}CH_2 = {}^{+}FAH_4$



$10\text{-}CH_2 = {}^{+}FAH_4$

A mechanism in which a reactive exocyclic methylene group on the uracil heterocycle is formed as an intermediate has been postulated by Santi and Pogolotti[31]. They have described "carbonium ion" - like ionizations of methyl-substituted thymines of the following types:

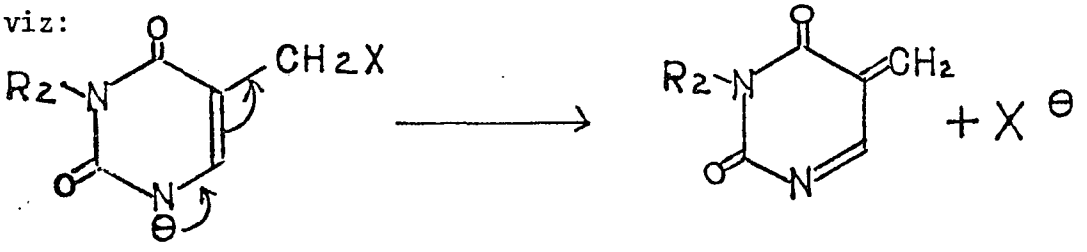


X = acetoxy or p-nitrophenoxy

R_1 = CH_3 or H

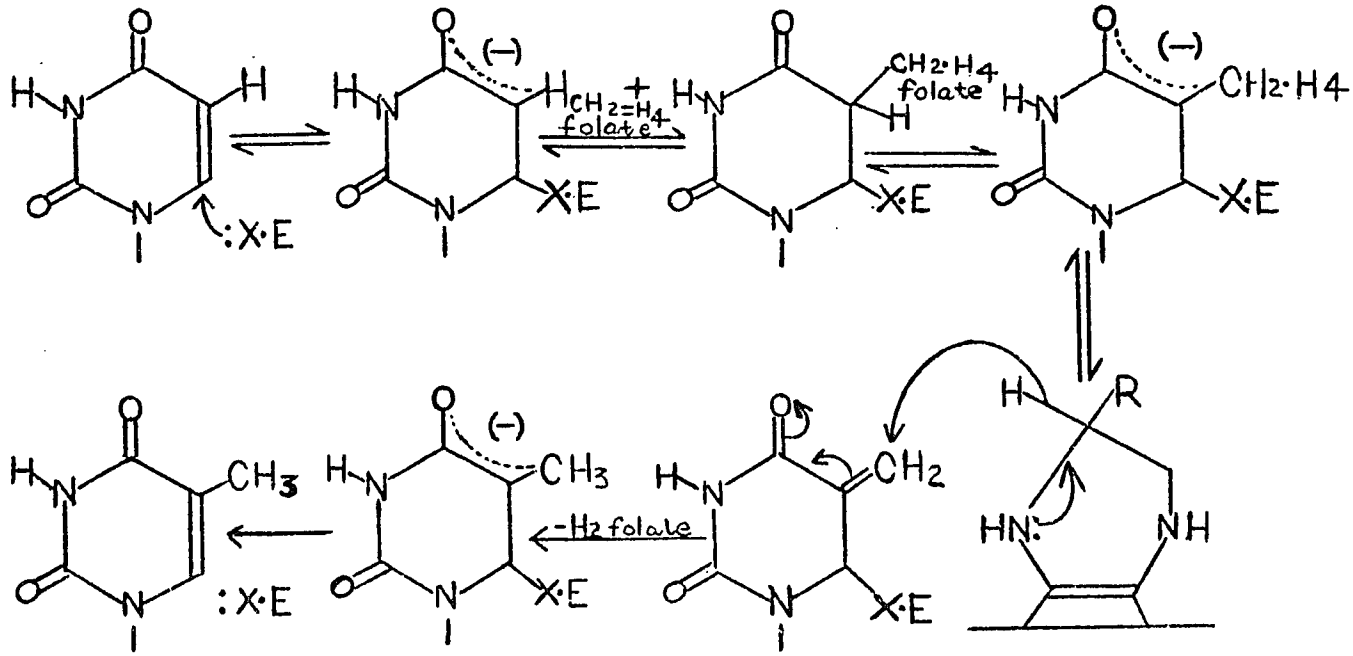
R_2 = CH_3 or H

When R_1 was H, the methoxide or borohydride-catalyzed displacement of X was facile, but when R_1 was methyl, the reaction was difficult. The substitution at R_2 was not critical. These results are consistent with ionization at the 1-position followed by internal nucleophilic ionization, viz:



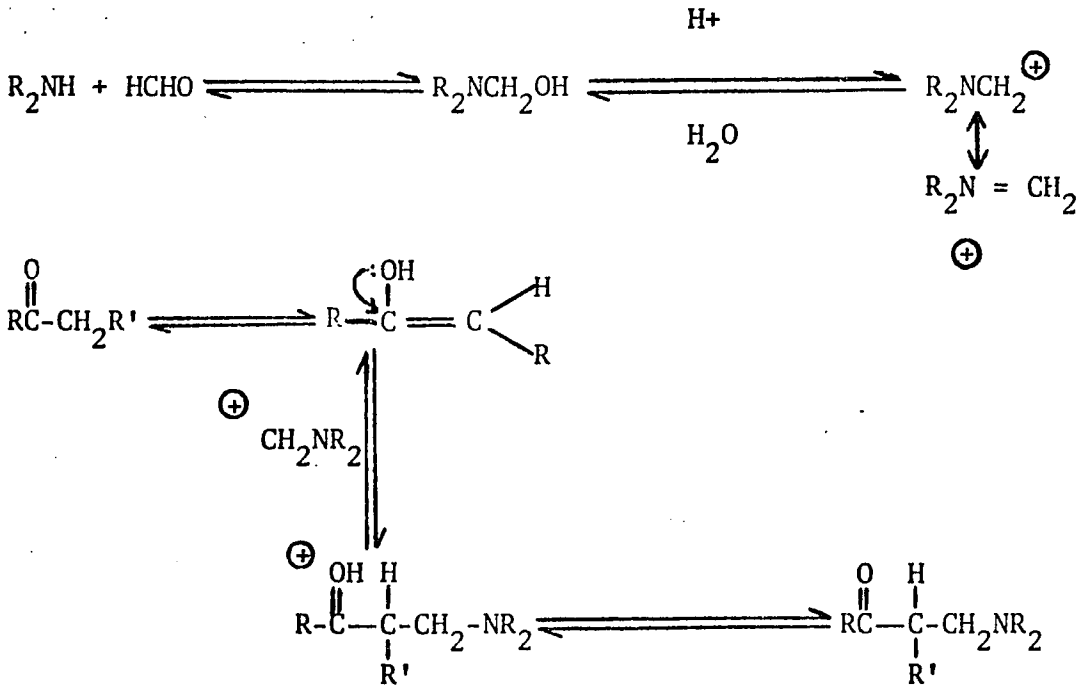
Although described as "carbonium-ion like", the reaction has virtually no indication of such an intermediate. The authors likened the more sluggish reaction (R_1=CH_3) to the thymidylate synthetase reaction with deoxyuridylate which proceeds by nucleophilic addition to the 6-position.

Several years later the same authors[32] again provided evidence for easy formation of exocyclic double bonds in thymine derivatives with a leaving group (p-nitrophenoxy) on the 5-methyl group. The necessity of the N-1 anion precludes this mechanism (or analogues) for natural occurrence in the thymidylate synthetase-catalyzed reactions of uridylic acid. Nevertheless, the authors idly speculated concerning the relationship of these studies to the thymidylate synthetase enzymatic nucleophilic attack to generate an active nucleophilic intermediate. The intermediate then attacks the induced methylene-iminium ion arising from methylenetetrahydrofolate. The retrogressive scission of the adduct produced an exocyclic methylene which was reduced by the eliminated tetrahydrofolate (hydride ion transfer), yielding the product carbanion which eliminates the enzyme by anionization.



All pyrimidine structures have a 1-(5-phospho-2'-deoxyribosyl) substituent and R=CH_2-NHC_6H_4COGlu.

The Mannich reaction mechanism has been proposed as a model describing the reaction of thymidylate synthetase, 5-fluorodeoxyuridylic acid and 5,10-methylenetetrahydrofolate[33]. In the Mannich reaction a primary or secondary amine reacts with a aldehyde (usually formaldehyde) and with a compound containing an active hydrogen. In acid solution it involves methylene iminium ion (referred to as "carbonium ion"). This lends some possible credence to postulated methylene iminium intermediates in both the dUMP and FdUMP reactions with thymidylate synthetase and methylenetetrahydrofolate. This mechanism lends support to the speculation of Pogolotti and Santi. The pertinent mechanism is as follows:



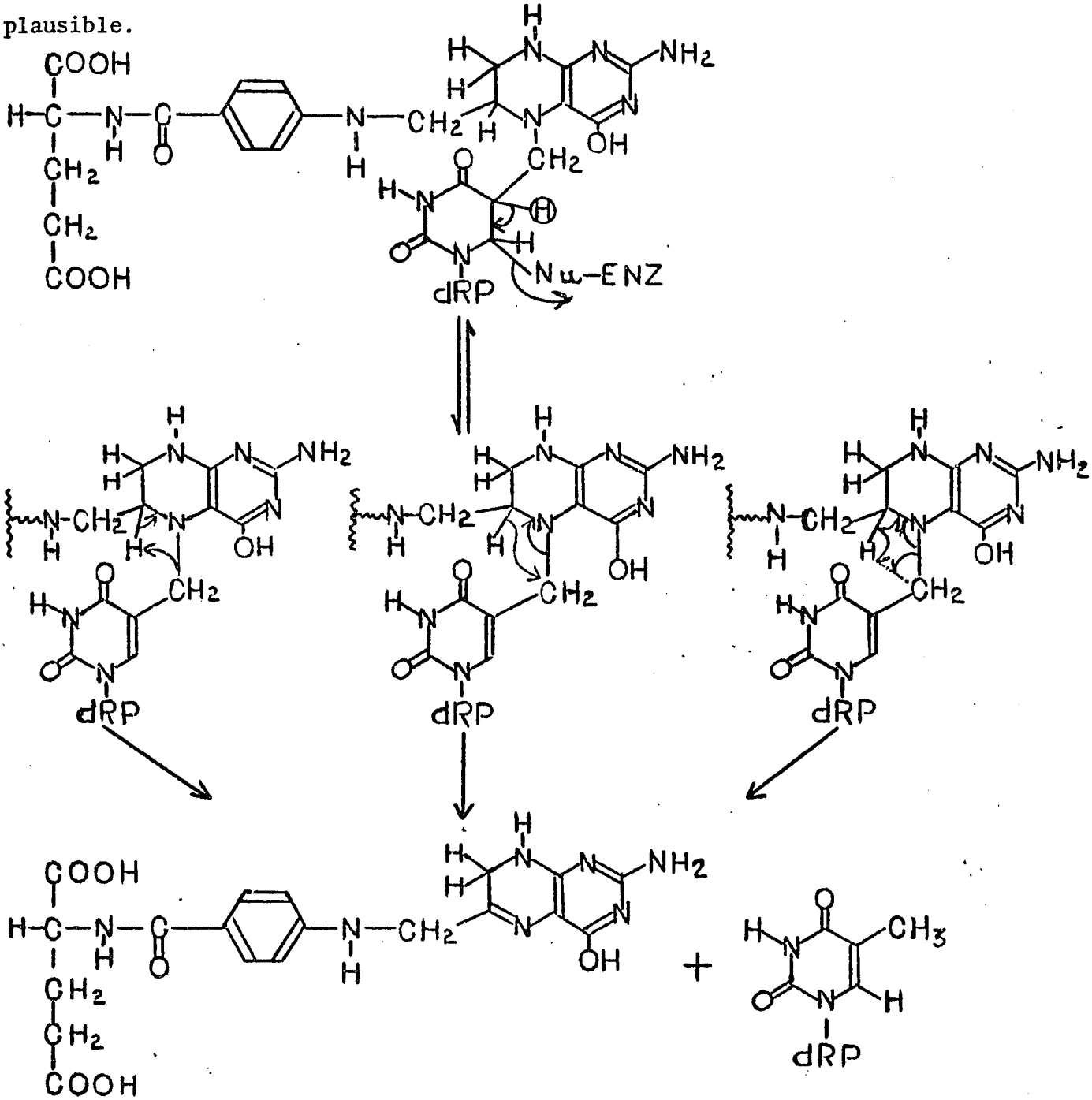
This postulation of this mechanism for the enzymatic reaction involves the questionable assumption that the intermediate is generated in a hyperacidic environment which is somehow "protected" from immediate destruction by the base, water.

In an effort to determine the extent of the reaction when 5-fluoro-2'-deoxyuridylic acid is substituted for 2'-deoxyuridylic acid, Dannenberg, Langenbach, and Heidelberger[34] obtained a difference ultraviolet spectrum upon formation of the ternary complex with thymidylate synthetase and 5,10-methylenetetrahydrofolate. The difference spectrum showed a decrease in the absorbance at 270nm providing evidence for saturation of the 5,6 double bond of the pyrimidine ring. An additional decrease in absorbance at 290nm is attributable to changes occurring in the tetrahydrofolate chromophore. When 10-methyltetrahydrofolate was substituted, the difference spectrum showed only an absorption decrease at 270nm. The data provides confirmation for reversible and nucleophilic addition to the 5,6 double bond which is responsible for the binding of 5-fluorodeoxyuridylic acid to the enzyme in the presence of the tetrahydrofolate cofactor. The amino acid residue essential for enzyme activity and inhibitor binding was shown to be cysteine. The tendency of enzyme to add to the 5,6 double bond was given by the extent of tritium exchange upon reversible addition of nucleophilic (-SH) to the double bond. Sharma and Kisliuk[35] showed that the tetrahydrofolate, upon reduction in the ternary complex with 5-fluoro-2'-deoxyuridylic acid gives an increase in the uv absorption at 335nm; by titration, the maximum change occurred upon 1:1 addition of FdUMP. Further addition of FdUMP to the second enzyme site caused the dihydrofolate to shift again to the tetrahydrofolate. The latter change has not been rationalized.

Discussion

From the above summary it is apparent that there is no dearth of speculation concerning mechanisms for the thymidylate synthetase-catalyzed reaction. Yet, no decisive information is available which

would clearly reveal a truly unique mechanism*. The evidence cited supports a covalently bound ternary complex of methylenetetrahydrofolate, deoxyuridylic acid (or 5-fluorodeoxyuridylic acid), and thymidylate synthetase. The following eclectic mechanism appears to be the most plausible.



*Two additional mechanisms without supportive evidence have been suggested by Wahba and Friedkin[25]. One involves first the formation of N^5-methyl-dihydrofolate: 5-10-Methylenetetrahydrofolate ⟶ N^5-methydihydrofolate $\xrightarrow{dUMP}$ dTMP + dihydrofolate. In the second mechanism, the thymidylate synthetase is first methylated: Enzyme + 5,10-methylenetetrahydrofolate ⟶ CH_3-enzyme + dihydrofolate. CH_3-enzyme + dUMP ⟶ dTMP + enzyme.

The first step of the reaction is the addition of enzyme sulfhydryl to FdUMP at the 6-position of the pyrimidine ring. This carbanionic nucleophilic intermediate at the 5-position then attacks methylene-tetrahydrofolate, generating the ternary adduct (complex). The decisive lesion of the normal process occurs in the next step. That step is widely accepted as the elimination of enzyme by an analog to the E_2 process. The final step in the normal process is the formation of the deoxythymidylate group by an intramolecular hydride ion, proton, or free radical shift from the 6-position of the tetrahydrofolate pteridine ring system. The latter step as pictured should be immune to solvent deuterium kinetic isotope effect. An experiment using D_2O would reveal any other existing factors contributing to the mechanism which presently is depicted entirely as an intramolecular process which only involves 4 atoms.

SECTION III

QUANTUM CHEMICAL CONSIDERATIONS OF 5-FLUOROURACIL AND RELATED COMPOUNDS

Introduction

There has been a steady increase in the literature in the application of molecular orbital theory to an understanding of chemical reactions*. In the field of biochemistry there has been an effort made to verify the location of reactive sites in the molecule in order to elucidate the reaction mechanisms involved**. Studies have been made involving molecular orbital calculations in the search for correlations between covalent bond formation and biological activity. Such studies include the investigation of antibacterial agents,[36-39] the correlation of ester hydrolysis rates of enzymes affected by drugs,[40-43] the effect of a class of compounds on plant growth,[44-46] and enzymic acetylation[47]. Reactions involving charge-transfer complexation have also been investigated quantum mechanically by studying electron-donor and electron-acceptor properties of molecules in a reaction through calculations of energy levels (E_{HOMO} {highest occupied molecular orbital energy} and E_{LEMO} {lowest empty molecular orbital energy}). Such studies included

*A knowledge of quantum theory is assumed by the author. An excellent reference text on the subject is "Elementary Quantum Chemistry", F. L. Pilar, McGraw-Hill, New York (1968). The following fine texts may also be consulted. "Quantum Theory of Molecular Electronic Structure", R. G. Parr, Benjamin, New York (1963). "Approximate Molecular Orbital Theory", J. A. Pople and D. L. Beveridge, McGraw-Hill, New York (1970).

** Texts related to the application of molecular orbital theory to biochemical problems are: "Quantum Biochemistry", B. Pullman and A. Pullman, Wiley (Interscience) New York (1963). "Molecular Orbital Theory in Drug Research", L. B. Kier, Academic Press, New York (1971). "Quantum Pharmacology", W. G. Richards, Butterworths, London (1971).

antimalarial drug reactions,[48,49] hallucinogenic reactions,[50] local anesthetics,[51] and carcinogenic hydrocarbons[52-57]. An attempt was made to explain carcinogenic activity of hydrocarbons on the basis of their electronic structure as determined by molecular orbital calculations[58]. The literature is abundant with studies in which molecular orbital theory has been applied to drugs in order to explain structure-activity relationships***. The examples cited represent only a small fraction of the papers published in this field and are meant to be illustrative rather than exhaustive.

When performing molecular orbital calculations on an isolated molecule whose dimensions are based on x-ray crystallographic measurements, the data obtained are not a true representation of the in vivo situation since one is dealing with a dynamic biological system in which thousands of interdependent reactions are occurring simultaneously with solvent media and neighboring molecules exerting an effect. The calculations do not include the existence of these independent reactions****.

In the studies previously mentioned, the calculated molecular orbital data were consistent with experimental findings, which lends credence to the fact that these additional neighboring contributions were negligible.

Method and Theory

A. Complete Neglect of Differential Overlap (CNDO)

One of the molecular orbital methods which considers all-valence

---

***A comprehensive bibliography can be found in "Quantum Pharmacology", W. G. Richards. See previous notation on suggested reference texts.

****Molecular orbital calculations in general refer to the gas-phase molecule even though geometry of the molecule is usually obtained from x-ray crystallographic data.

electrons is the CNDO/2 SCF method[59]. This approximate method calculates self-consistent molecular orbitals for all valence-electrons (with 1s electrons considered as part of a nuclear core) in molecules where all atoms are first row elements. It is based on the complete neglect of differential overlap (CNDO) approximation in which the differential overlap distribution $\phi_m(1)\phi_n(1)d\tau_1$ $m \neq n$ of any two atomic orbitals is neglected in all electron repulsion integrals. The rationale is that a large number of these repulsion integrals in LCAO-SCF calculations are practically zero in value and are particularly difficult to evaluate if the atomic functions are centered on different atoms.

With the CNDO method the assumption is made that $(mn|\ell s) = \delta_{mn}\delta_{\ell s}(mm|\ell\ell)$ where δ_{mn} and $\delta_{\ell s}$ are Kronecker deltas and $(mn|\ell s)$ is the 2-electron repulsion integral over atomic orbitals defined as $\int\int\phi_m(1)\phi_n(1)\frac{1}{r_{12}}\phi_\ell(2)\phi_s(2)d\tau_1 d\tau_2$. This approximation thereby eliminates all multi-centered as well as one-center integrals where different atomic orbitals are involved for either of the two electrons. Even though the CNDO method is applied to all $\phi_m\phi_n$ atomic orbitals, $m \neq n$, with Roothan's[60] equation applying, the approximation is not invariant to cartesian axes rotation. However, this deficiency is corrected by making the remaining 2-electron repulsion integrals depend only on the atoms A & B on which ϕ_m and ϕ_ℓ are situated; $(mm|\ell\ell)$ are set equal to γ_{AB}, representing an average electrostatic interaction between an electron on A and one on B. Having neglected monatomic differential overlap, $\phi_m(1)\phi_n(1)d\tau_1$ $(m \neq n)$, the next consideration is that of the electrostatic interaction between an electron in an orbital of one atom and the cores of the other atoms. If V_B is the core potential at atom B, then the integral $(m|V_B|n)$ also equals zero by virtue of the

neglect of the monatomic differential overlap of the two unequal orbitals on atom A. This can be written as $(m|V_B|n) = \delta_{mn}V_{AB}$, which then leads to $H_{mm} = U_{mm} - \Sigma V_{AB}$ with ϕ_m on atom A. H_{mm} represents the diagonal core matrix element where $U_{mm} = (m|-\frac{1}{2}\nabla^2-V_A|m)$ is the one-center electron-core matrix element, and V_{AB} is the two-center valence-electron core matrix interaction contribution. Differential overlap is not neglected when ϕ_m and ϕ_n are on separate atoms because H_{mn} $[= (m|-\frac{1}{2}\nabla^2-V_A-V_B|n)]$ - the off-diagonal core matrix elements - describes bonding between A and B. It is necessary to retain differential overlap in H_{mn} by having it depend semi-empirically on overlap. H_{mn} is also called the resonance integral β_{mn}, and relates to the bonding capacity of the overlap; it is proportional to the overlap and can be written as $\beta_{mn} = \beta^o_{AB}S_{mn}$. β^o_{AB} depends only on the nature of atoms A and B. This satisfies the invariance requirements[61].

The CNDO/2 and CNDO/1 approximations are basically similar. The modifications of the CNDO/1 method will be presented after the parameterization is described. Slater functions representing the orbitals are used for the minimal basis set ϕ_m (1s for hydrogen and 2s, $2p_x$, $2p_y$, $2p_z$ orbitals for lithium to fluorine). Slater's rules are used to obtain the exponents. For hydrogen the effective nuclear charge is $\zeta = 1.2$ instead of 1.0. The overlap integrals are calculated with a computer. The repulsion integral is calculated as a two-center coulomb integral using the s orbital on atom A, (s_A). $\gamma_{AB} = \int s_A^2(1)(r_{12})^{-1} s_B^2(2)d\tau_1 d\tau_2$. The electron-core interaction parameter V_{AB} is also calculated using valence s functions. The core of atom B is treated

as a point charge at the nucleus so that $V_{AB} = Z_B \int s_A^2(1)(r_{1B})^{-1} d\tau_1$ where Z_B is the core charge and r_{1B} is the distance between electron 1 and the B nucleus. The bonding parameters β^o_{AB} are taken as the average of the bonding parameters of each atom. Thus, $\beta^o_{AB} = \frac{1}{2}(\beta^o_A + \beta^o_B)$ where β^o_A and β^o_B are selected empirically. Finally, the atomic matrix elements of the one-electron hamiltonian, U_{mm}, are obtained from experimental energy levels. The energy of an atomic core and the valence electrons are represented by an average of several states. The core integrals for the 2s and 2p orbitals are then related to either ionization energies or electron affinities related to these states. The equation is $U_{mm} = -\frac{1}{2}(I_m + A_m) - (Z_A - \frac{1}{2})\gamma_{AA}$ with ϕ_m belonging to atom A. I_m and A_m refer to the ionization energy and electron affinity, while Z_A is the core charge on A.

B. Method of Solution of LCAOSCF Equations

The LCAOSCF equations are solved in the following manner. The initial approximation to the molecular-orbital coefficients is obtained from a modified Hückel calculation in which the diagonal elements of the matrix representation of the Hartree-Fock operator, F_{mn}, are replaced by the atomic ionization energies from equations relating the ionization energies to core-and repulsion-integral parameters, U and γ. The off-diagonal elements of the Fock hamiltonian are replaced by resonance integrals $\beta^o_{AB} S_{mn}$. The electrons are assigned in pairs to molecular orbitals with lowest energy. The density matrix, whose elements are P_{mn}, is calculated from the occupied molecular orbital coefficients in order to form a new F_{mn}. Diagonalization of the new Fock matrix leads to a new set of orbital coefficients. The

iterations are repeated until self consistency is achieved.

C. Comparison of CNDO/1 and CNDO/2 Approximations

The CNDO/1 approximation has several shortcomings[62]. In calculations for diatomic molecules the equilibrium distance is too small and the corresponding dissociation energy is too large. This is due to electrons in one orbital of an atom penetrating the shell of another atom. The diagonal matrix element of the Fock hamiltonian can be written as: $F_{mm} = U_{mm} + (P_{AA} - \frac{1}{2}P_{mm})\gamma_{AA} + \sum_{B\ (\neq A)} [-Q_B \gamma_{AB} + (Z_B\gamma_{AB} - V_{AB})]$ where $P_{AA} = \sum_m^A P_{mm}$ is the total valence electron density at atom A, P_{mm} is the charge on orbital m, Q_B is the net charge on atom B, and $Z_B\gamma_{AB} - V_{AB}$ are the penetration integrals. The inclusion of the latter term results in a net attraction where there should be a net repulsion and also leads to incorrect bonding energies. In the CNDO/2 method, the penetration integrals are neglected and the shortcomings of CNDO in this regard are eliminated. Since the study is confined to a closed shell configuration, the following energy equations apply[61]. The total energy is composed of the monatomic energy, ξ_A, and diatomic energy ξ_{AB}: $\xi_A = \sum_m^A P_{mm}U_{mm} + \frac{1}{2}\sum_m^A \sum_n^A (P_{mm}P_{nn} - \frac{1}{2}P_{mn}^2)$

$$\xi_{AB} = \sum_m^A \sum_n^B (2P_{mn}\beta_{mn} - \tfrac{1}{2}P_{mn}^2\gamma_{AB}) + (Z_AZ_BR_{AB}^{-1} - P_{AA}V_{AB} - P_{BB}V_{BA} + P_{AA}P_{BB}\gamma_{AB})$$

For large internuclear distances, the last group of terms approximate to R_{AB}^{-1} and become $Q_AQ_BR_{AB}^{-1}$ with $Q_A = Z_A - P_{AA}$. In the CNDO/2 method the diagonal matrix elements of the Fock hamiltonian are now functions of the ionization energy, I, and the electron affinity, A, (an average of the two representing a one-center, one-electron term). The equations

for both the diagonal and off-diagonal elements are[63]:

$F_{mm} = \frac{1}{2}(I_m + A_m) + [(P_{AA}-Z_A)-\frac{1}{2}(P_{mm}-1)]\gamma_{AA} + \sum_{B\,(\neq A)} (P_{BB}-Z_B)\gamma_{AB}$

and $F_{mn} = \beta^{o}_{AB}S_{mn} - \frac{1}{2}P_{mn}\gamma_{AB}$ $(m \neq n)$

D. Mulliken Population Analysis

LCAO molecular orbital wave functions can be interpreted in terms of charge distributions in molecules. Mulliken's[64] population-analysis method defines numerical indices giving the number of electrons associated with each specific atom in the molecule. If ϕ is a normalized molecular orbital of a diatomic molecule it can be written as a linear combination of normalized atomic orbitals, χ_m and χ_n, centered on atoms A & B respectively, viz: $\phi = c_m\chi_m + c_n\chi_n$. The gross population at each atom consists of the net atomic population at that atom plus one-half of the overlap population between the two atoms. $N(A) = N(c_m^2 + c_m c_n S_{mn})$ and $N(B) = N(c_m c_n S_{mn} + c_n^2)$, where $S_{mn} = \int \chi_m \chi_n d\tau$ and N is the over-all total population of the electrons. In the general case (for any molecule) the equation for the gross atomic population on atom A is $N(i;m_A) = N(1)c_{im_A} \times (c_{im_A} + \Sigma_{B \neq A}\ c_{in_B} S_{m_A n_B})$

If the bonds between the atoms are polar, the gross charge on each atom can be obtained by subtracting the gross atomic population at the atom from the total number of electrons in the ground state of the free neutral atom.

E. Computer Program

A Fortran IV computer program for calculating CNDO and INDO molecular orbitals* was adapted for use in the Univac 1108 computer. The program

*This program was prepared by J. A. Pople, D. L. Beveridge, and P. A. Dobosh. It appears as Appendix A in J. A. Pople and D. L. Beveridge, "Approximate Molecular Orbital Theory", McGraw-Hill, NY (1970). Card copies of the program may be obtained from Quantum Chemistry Program Exchange (QCPE) Dept. of Chemistry, Indiana University, Bloomington, Ind. 47401.

calculates energies and other properties based on assumed wave functions for either open or closed shell molecules containing elements from H to Cl (CNDO) and H to F (INDO)*. The matrices can accommodate molecules containing up to 85 atoms or 80 basis functions (whichever is smaller). One atomic orbital basis function is allotted to hydrogen (1s), and one each of (2s, $2p_x$, $2p_y$, and $2p_z$) for Li thru F. Additional atomic orbital basis functions are available for the elements sodium through chlorine, but will not be used in this study.

After the method of approximation (CNDO or INDO) and the type of calculation (open or closed shell) have been selected, the number of atoms, charge and multiplicity are entered. This is followed by the cartesian coordinates of each atom. Following this, the coefficients which are to be used in the calculation of the overlap and coulomb integrals are assigned. The overlap and coulomb integral matrices are computed and the overlap integrals are then transformed to the molecular coordinate system by a rotation matrix. An extended Hückel-type approximation with zero differential overlap is made for the Fock matrix with F_{mm} formed from $-\frac{1}{2}(I + A)$ and F_{mn} formed from $(\beta_A^o + \beta_B^o) \times S_{mn}/2$. This matrix is then diagonalized followed by the construction of a density matrix. Repulsion terms are then added to the hamiltonian for either CNDO or INDO calculations since they are not included in the Hückel scheme.

The initial density matrix and the core hamiltonian are then used to form the Fock matrix. This matrix is diagonalized, forming a new density

---

*INDO (intermediate neglect of differential overlap) retains differential overlap in one-center integrals and consequently is better able to describe certain electronic states with open-shell configurations.

matrix, which then is used to form a new Fock matrix. This iteration is repeated until the electronic energy converges to 10^{-6}au, at which time the electronic energy is assumed to be minimized. The Fock matrix is diagonalized, the eigenvectors are printed out, and both the total energy and binding energy calculated. The binding energy is the difference between the minimum total energy of the molecule and the sum of the atomic energies of the component atoms. The program then computes the dipole moment, the monatomic and diatomic energies, and the electrostatic interaction between charged atoms in the molecule. The program also gives the Mulliken overlap population at each atom location.

Results

At the outset a calculation was carried out for formaldehyde and the results compared with the output of the QCPE CNDO/2 program for the same molecule. There was perfect agreement for all the parameters to the 4th decimal place or better. This substantiates the accuracy of the program. Calculations were then made for uracil, 5-fluorouracil, and thymine. The geometries* for these molecules were obtained from X-ray crystallography studies[65-67]. Also a CNDO calculation was made from the geometry obtained in an ab initio study of thymine molecule[68]. Finally, the bond angles and bond lengths of 5-fluorouracil were used to calculate the parameters of thymine by substituting methyl for fluorine and using the C-F bond length for C-CH_3. All of these calculations were

*The numbering system and the dimensions of each of the 2,4-diketopyrimidines which were obtained from crystal structure determinations and an ab initio study for use in the CNDO calculations are illustrated in Figures 3-6. Slight modifications in the bond angles were made whenever the literature values appeared incorrect. Estimated values for bond angles and lengths were introduced when they had been omitted. Figures 7 to 9 illustrate the charge at each atom in the molecules and Tables I-V provide the density matrix and overlap population data. The calculated dipole moments (compared with literature experimental values) together with total energies are listed in Table VI. Table VII summarizes the calculated charges at the 5 and 6 positions of uracil, 5-fluorouracil and thymine.

performed to determine the reactive centers in the uracil molecule, the effect of the fluorine atom in the 5-position, and to substantiate, if possible, that part of the proposed mechanism which deals with the electro- and nucleophilicity of the reactive sites.

Discussion

Given the step-wise picture and the nucleophilic character of the attacking enzyme - as considerable evidence indicates - uracil has the appropriate electrophilic center at C(6) as does fluorouracil and thymine. The presence of a negative charge at C(5) in both thymine and uracil favors development of high nucleophilicity in C(5) from any intermediate produced upon nucleophilic attack at C(6). High nucleophilicity at C(5) in uracil favors the second stage of postulated mechanisms of attack of activated methylene species which are electrophilic. High nucleophilicity at C(5) in thymine would tend to favor breakup of the thymine-enzyme complex through ejection of the enzyme as an anionic species, especially since the methyl group sterically favors dissociation of the cofactor (methylenetetrahydrofolate) leading to the localization of the negative charge. Fluorouracil also possesses the required electrophilic C(6) position for enzymatic attack and formation of the postulated adducts. However, the C(5) position has a decided electrophilic quality which will probably only stabilize or accomodate the negative charge from the enzyme, and may not assume sufficient nucleophilic character in the intermediate stage for any further step-wise reaction with an activated methylene group such as in the tetrahydrofolate co-factor. Because it is indicated by these calculations that the fluorouracil can accommodate the negative charge, it may be that the intermediate will be stabilized and will not easily dissociate the bound enzyme.

Kulakowska[69] and co-workers have calculated charge densities and

other properties of the same group of molecules treated in this study. They also used the CNDO/2 method as utilized in the present work. However, they assumed a standard fixed geometry for the uracil nucleus to be invariant to the set of compounds and the only bond adjustments made were empirical adjustments on the various substituent atoms. Their calculations exhibit the qualitative characteristics of the present work. They are in reasonably good accord with the present values, but do not precisely agree because the charge densities show some geometric sensitivity. The authors did not consider the implications of their work on biological systems, but attempted some rationalization of pK and dipole moment data. There is reasonably good agreement between their experimental dipole moment data and those calculated in the present work (Table XIX).

It is interesting at this point to mention a photochemical study[70] in which comparative CNDO/2 calculations of the ground and excited states of uracil, thymine, and cytosine were performed. Atoms C(5) and C(6) were found to be highly reactive, indicating that even in the excited state the chemistry of these pyrimidines is dominated by the 5,6 reaction sites.

SECTION IV

SUMMARY*

1. There is a greater negative (or less positive) charge at C(5) than at C(6). This substantiates a nucleophilic attack by the enzyme at the C(6) position rather than at C(5). Indirectly, it implicates eventual potentiation of a nucleophilic attack by the substrate on the co-factor.

2. There is a change in polarity at C(5) due to the presence of fluorine in 5-fluorouracil.

The change in polarity with the introduction of a fluorine atom demonstrates the electronegativity of the halogen atom. The positive charge at C(5) results in a stabilization of the enzyme adduct.

3. The C(5) - C(6) bond moment in 5-fluorouracil is very small. This suggests the presence of another factor (steric fit to the enzyme active site cavity) which may play a decisive role in determining positional attack.

Thus in the normal reaction (with uracil) steric fit may also be a contributing factor even though the molecule is electronically amenable to the observed selecivity.

---

*See Table VII

TABLE I

URACIL

DIAGONALS OF THE DENSITY MATRIX

.99804	.91398	.83971	.79805	1.15019	1.15156	1.17012	1.73385	.99667
.91054	1.02493	.88733	.98154	1.01112	1.01420	1.16349	1.18578	1.18394
1.14488	1.73400	1.00018	.86206	.96882	.80178	1.74544	1.80664	1.33115
1.41948	1.73164	1.26059	1.91797	1.46202	.85119	.99299	.94547	.84866

SUM APPROXIMATE OVERLAP POPULATION

3.549780	1.514635	.015458	-.009724	1.529212	.008543	.001299	1.821230	.013918	.005536
.000235	.006688								
1.514635	5.205728	1.569385	-.013042	-.022732	-.012831	.000032	-.024077	1.455581	-.012148
.003865	.003928								
.015458	1.569385	3.819470	2.135570	-.013601	.034456	.000950	.001340	.002837	1.507133
.004162	.000606								
-.009724	-.013042	2.135570	4.170345	-.013787	1.685136	-.018210	.000087	.005881	-.006878
1.534363	.006103								
1.529212	-.022732	-.013601	-.013787	5.248594	1.507561	-.024485	-.024918	.004205	.000032
.003337	1.396696								
.008543	-.012831	.034456	1.685136	1.507561	3.632847	1.732789	.001187	.000445	.005897
.004110	.011117								
.001299	.000032	.000950	-.018210	-.024485	1.732789	6.362717	-.000003	.000006	-.000052
-.003084	-.002420								
1.821230	-.024077	.001340	.000087	-.024918	.001187	-.000003	6.372216	-.002183	-.000058
.000006	-.002317								
.013918	1.455581	.002837	.005881	.004205	.000445	.000006	-.002183	.851187	-.011875
-.000243	-.000278								
.005536	-.012148	1.507133	-.006878	.000032	.005897	-.000052	-.000056	-.011875	.992988
-.010817	.000058								
.000235	.003865	.004162	1.534363	.003337	.004110	-.003084	.000006	-.000243	-.010817
.046474	-.000244								
.006688	.003928	.000606	.006103	1.396696	.011117	-.002420	-.002317	-.000278	.000058
-.000244	.848656								

TABLE II

5-FLUOROURACIL

DIAGONALS OF THE DENSITY MATRIX

1.00195	.90850	.83354	.80644	1.18800	1.12052	1.11930	1.76672	1.02506
.91332	1.00773	.95165	1.00291	.79000	1.00197	1.12703	1.84019	1.39908
1.98361	1.96681	1.21437	1.15245	1.12130	1.75225	1.01164	.87046	.96259
.81690	1.74250	1.78801	1.40187	1.38209	1.72425	1.28369	1.91041	1.43011
.85859	.96673	.85546						

SUM APPROXIMATE OVERLAP POPULATION

3.550430	1.470346	.010634	-.007293	.000037	1.465887	.003952	.001104	1.873243	-.005941
.003644	-.000773								
1.470346	5.194540	1.510916	-.004061	.000346	-.020619	-.008800	.000025	-.025734	1.258650
-.012080	.003000								
.010634	1.510916	3.897760	2.136219	-.012358	-.009028	.010697	.000635	.001331	-.005719
1.348590	.000259								
-.007293	-.004061	2.136219	3.921903	1.003340	-.006894	1.639067	-.015879	.000121	.004666
-.007407	.005247								
.000037	.000346	-.012358	1.003340	7.189689	.000349	-.009473	-.000649	.000000	-.000031
-.002337	-.000038								
1.465887	-.020619	-.009028	-.006894	.000349	5.240374	1.480912	-.022318	-.024041	.002914
.000032	1.245403								
.003952	-.008800	.010697	1.639067	-.009473	1.480912	3.661592	1.760470	.001101	.000186
.003851	-.001051								
.001104	.000025	.000635	-.015879	-.000649	-.022318	1.760470	6.314475	-.000002	.000003
-.000040	-.001607								
1.873243	-.025734	.001331	.000121	.000000	-.024041	.001101	-.000002	6.348461	-.001184
-.000042	-.001949								
-.005941	1.258650	-.005719	.004666	-.000031	.002914	.000186	.000003	-.001184	.058590
-.009205	-.000225								
.003644	-.012080	1.348590	-.007407	-.002337	.000032	.003851	-.000040	-.000042	-.009205
.966729	.000032								
-.000773	.003000	.000259	.005247	-.000038	1.245403	-.001051	-.001607	-.001949	-.000225
.000032	.855457								

TABLE III

THYMINE

DIAGONALS OF THE DENSITY MATRIX

.99046	.93897	.83898	.78206	1.18005	1.13720	1.13596	1.72713	1.03417
.83199	1.00257	.94520	.99127	.99841	.99900	1.10619	1.01454	1.01538
.98289	.98239	1.21285	1.16036	1.13081	1.74023	1.00479	.85598	.95180
.82435	1.72100	1.79909	1.44541	1.35059	1.74720	1.21288	1.93084	1.53128
.86211	.99120	.98932	.98513	.99910	.85889			

SUM APPROXIMATE OVERLAP POPULATION

3.550471	1.660184	.016154	-.007618	.000111	1.601980	.008406	.001262	1.709384	-.000619
.003774	.000011	.000006	.000002	.000715					
1.660184	5.180346	1.468193	-.013712	.001702	-.019785	-.011249	.000034	-.024978	1.251617
-.015738	.000023	-.000040	.000103	.003796					
.016154	1.468193	3.873929	2.079702	.006582	-.011790	.022682	.000804	.001515	-.005781
1.351088	-.002842	.003252	-.008703	.000418					
-.007618	-.013712	2.079702	4.094860	1.468553	-.016438	1.580602	-.019533	.000068	.004482
-.012554	.006046	.003500	.003794	.004720					
.000111	.001702	.006582	1.468553	3.995201	.001363	.006603	-.003112	.000001	-.000108
-.006950	1.342784	1.349302	1.348871	-.000102					
1.601980	-.019785	-.011790	-.016438	.001363	5.244258	1.422313	-.025796	-.024157	.003980
.000025	-.000015	.000118	-.000050	1.245633					
.008406	-.011249	.022682	1.580602	.006603	1.422313	3.636911	1.891972	.001226	.000294
.004259	.000058	-.006566	.003309	-.003008					
.001262	.000034	.000804	-.019533	-.003112	-.025796	1.891972	6.316085	-.000003	.000005
-.000041	.000038	.001212	.000007	-.001369					
1.709384	-.024978	.001515	.000068	.000001	-.024157	.001226	-.000003	6.422199	-.002273
-.000045	-.000000	-.000000	-.000000	-.002410					
-.000619	1.251617	-.005781	.004482	-.000108	.003980	.000294	.000005	-.002273	.862108
-.008652	.000004	.000008	-.000017	-.000282					
.003774	-.015738	1.351088	-.012554	-.006950	.000025	.004259	-.000041	-.000045	-.008652
.991198	.000122	.000053	.001597	.000037					
.000011	.000023	-.002842	.006046	1.342784	-.000015	.000058	.000038	-.000000	.000004
.000122	.989316	-.008193	-.008417	.000003					
.000006	-.000040	.003252	.003500	1.349302	.000118	-.006566	.001212	-.000000	.000008
.000053	-.008193	.985131	-.010592	-.000023					
.000002	.000103	-.008703	.003794	1.348871	-.000050	.003309	.000007	-.000000	-.000017
.001597	-.008417	-.010592	.999097	.000009					
.000715	.003796	.000418	.004720	-.000102	1.245633	-.003008	-.001369	-.002410	-.000282
.000037	.000003	-.000023	.000009	.858891					

TABLE IV

THYMINE*

DIAGONALS OF THE DENSITY MATRIX

1.73390	1.92603	1.26749	1.45247	1.73401	1.53926	1.68786	1.39346	1.18544
1.14058	1.12503	1.72789	1.21238	1.11868	1.16072	1.74457	.99970	.83741
.92500	.80341	1.02473	1.02432	.89153	.91388	.98383	.99348	.99144
1.14949	.98936	.98816	.84518	.81756	1.01698	.98317	1.00635	.99204
.84831	.85623	.99397	1.00617	.99582	.98371			

SUM APPROXIMATE OVERLAP POPULATION

6.379884	-.000003	-.025458	-.027384	1.802357	.001444	.000147	.001451	.000001	-.001746
-.001738	-.000050	-.000000	-.000000	-.000000					
-.000003	6.354593	.000039	-.022931	.001374	.001253	-.017774	1.787355	-.005246	-.000901
.000005	-.000059	.000103	.000106	.001900					
-.025458	.000039	5.178942	-.018261	1.506711	1.608475	-.013130	-.011622	.001740	.002642
1.203553	-.016788	.000117	.000029	-.000039					
-.027384	-.022931	-.018261	5.236347	1.524073	-.010014	-.014591	1.496698	.001381	1.206105
.002675	-.000014	-.000038	-.000013	.000089					
1.802357	.001374	1.506711	1.524073	3.565515	.022116	-.011074	.025412	.000197	-.003518
.002769	.004513	.000001	.000010	.000007					
.001444	.001253	1.608475	-.010014	.022116	3.854456	2.019271	.024225	.006142	.000258
-.015503	1.366307	-.008730	-.002543	.003184					
.000147	-.017774	-.013130	-.014591	-.011074	2.019271	4.117244	1.707802	1.441826	.004563
.004675	-.009682	.001605	.002611	.001467					
.001451	1.787355	-.011622	1.496698	.025412	.024225	1.707802	3.640252	.007299	-.012691
.000326	.004916	.003104	-.000133	-.005738					
.000001	-.005246	.001740	.001381	.000197	.006142	1.441826	.007299	3.988556	-.000098
-.000118	-.007717	1.350090	1.344577	1.349945					
-.001746	-.000901	.002642	1.206105	-.003518	.000258	.004563	-.012691	-.000098	.848309
-.000185	.000040	.000008	.000003	-.000020					
-.001738	.000006	1.203553	.002875	.002769	-.015503	.004675	.000326	-.000118	-.000185
.856230	-.006881	-.000018	.000004	.000008					
-.000050	-.000059	-.016788	-.000014	.004513	1.366307	-.009682	.004916	-.007717	-.000040
-.006881	.993972	.001933	.000178	.000117					
-.000000	.000103	.000117	-.000038	.000001	-.008730	.001605	.003104	1.350090	.000008
-.000018	.001933	1.006172	-.008480	-.009433					
-.000000	.000106	.000029	-.000013	.000010	-.002543	.002611	-.000133	1.344577	.000003
.000004	.000178	-.008480	.995818	-.007977					
-.000000	.001900	-.000039	.000089	.000007	.003184	.001467	-.005738	1.349945	-.000020
.000008	.000117	-.009433	-.007977	.983710					

*Geometric dimensions taken from ab initio calculations[70].

TABLE V

THYMINE*

DIAGONALS OF THE DENSITY MATRIX

1.00239	.90873	.83506	.80382	1.19531	1.13655	1.11346	1.75788	1.02673
.89120	1.00122	.95319	.95046	1.04752	.96538	1.10303	.98741	1.04754
.96313	.96739	1.22090	1.16057	1.11751	1.74645	1.01529	.85594	.96335
.80564	1.74342	1.78717	1.40882	1.40057	1.72429	1.28528	1.91114	1.43693
.86861	.99079	1.00003	.99468	1.01710	.86512			

SUM APPROXIMATE OVERLAP POPULATION

3.549999	1.473852	.010200	-.008111	.000169	1.487141	.003419	.001099	1.870870	-.006045
.003645	.000020	.000012	.000004	-.000939					
1.473852	5.203196	1.507451	-.014375	.002623	-.019631	-.009915	.000018	-.025733	1.256875
-.015823	.000060	-.000065	.000209	.002933					
.010200	1.507451	3.872345	2.098320	.022458	-.010567	.015063	.000675	.001300	-.005516
1.353466	-.005129	.005380	-.014391	.000373					
-.008111	-.014375	2.098320	4.086393	1.731166	-.015321	1.599138	-.017540	.000098	.005223
-.012648	.013515	.009022	.008248	.005756					
.000169	.002623	.022458	1.731166	3.968463	.002274	.038091	-.005874	.000001	-.000155
-.008479	1.320967	1.333893	1.329440	-.000151					
1.487141	-.019631	-.010567	-.015321	.002274	5.245434	1.477821	-.023362	-.024010	.002856
-.000040	-.000018	.000404	-.000113	1.243245					
.003419	-.009915	.015063	1.599138	.038091	1.477821	3.640220	1.761177	.001087	.000265
.004142	-.000060	-.012703	.006418	-.000303					
.001099	.000018	.000675	-.017540	-.005874	-.023362	1.761177	6.339980	-.000002	.000004
-.000034	.000028	.003401	.000061	-.001485					
1.870870	-.025733	.001300	.000098	.000001	-.024010	.001087	-.000002	6.357642	-.001318
-.000044	-.000000	-.000000	-.000000	-.002032					
-.006045	1.256875	-.005516	.005223	-.000155	.002856	.000265	.000004	-.001318	.868607
-.008524	.000006	.000014	-.000032	-.000203					
.003645	-.015823	1.353466	-.012648	-.008479	-.000040	.004142	-.000034	-.000044	-.008524
.990786	.000111	.000043	.002379	.000040					
.000020	.000066	-.005129	.013515	1.320967	-.000018	-.000060	.000028	-.000000	.000006
.000111	1.000030	-.010288	-.010247	.000005					
.000012	-.000065	.005380	.009022	1.333893	.000404	.012703	.003401	-.000000	.000014
.000043	-.010288	.994683	-.015221	-.000070					
.000004	.000209	-.014391	.008248	1.329440	-.000113	.006418	.000061	-.000000	-.000032
.002379	-.010247	-.015221	1.017102	.000020					
-.000939	.002933	.000373	.005756	-.000151	1.243245	-.000903	-.001485	-.002032	-.000203
.000040	.000005	-.000070	.000020	.865119					

*Geometric dimensions are those of 5-fluorouracil with the methyl group substituted for fluorine.

TABLE VI

	Uracil	5-Fluorouracil	Thymine	Thymine[a]	Thymine[b]
Sum of EA[c]	-77.98	-104.80	- 84.19	- 84.19	- 84.19
Sum of EAB[d]	-13.24	- 13.68	- 15.99	- 15.99	- 16.00
Total energy	-91.22	-118.48	-100.18	-100.18	-100.19
Binding energy[e]	- 5.69	- 6.04	- 7.21	- 7.20	- 7.22
Dipole moment	4.78[f]	4.24[g]	4.31[h]	3.98[h]	4.07[h]

a. Used geometric dimensions of 5-fluorouracil, substituting CH_3 for F.

b. Geometric dimensions obtained from ab initio calculation of the molecule[69].

c. Monatomic energy in electron volts.

d. Diatomic energy in electron volts.

e. Energy in atomic units.

f. Literature[71] experimental value 4.16 $\pm$ 0.04 debyes.

g. Literature[71] experimental value 4.11 $\pm$ 0.05 debyes.

h. Literature[71] experimental value 4.13 $\pm$ 0.03 debyes.

TABLE VII

POLARIZATION OF 5-6 BOND

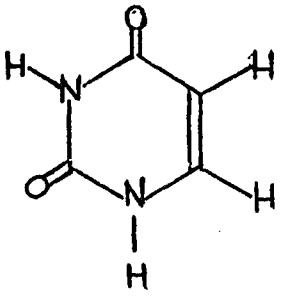
	Uracil	5-Fluorouracil	Thymine	Thymine[a]	Thymine[b]
C(5)	-.1703	.0781	-.0949	-0864	-.1172
C(6)	.1805	.1022	.1262	.1277	.1455
Charge Difference	.3508	.0241	.2211	.2141	.2627

a. 5-Fluorouracil data

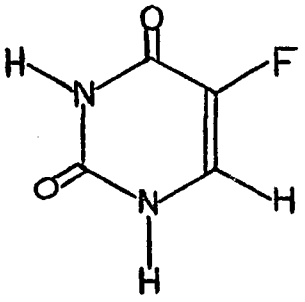
b. See * footnote in Figure 8.

FIGURE 1

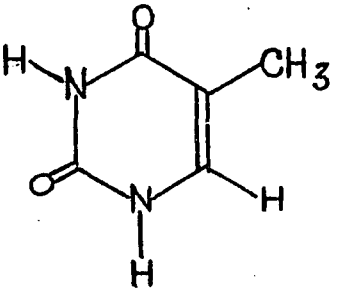
STRUCTURES



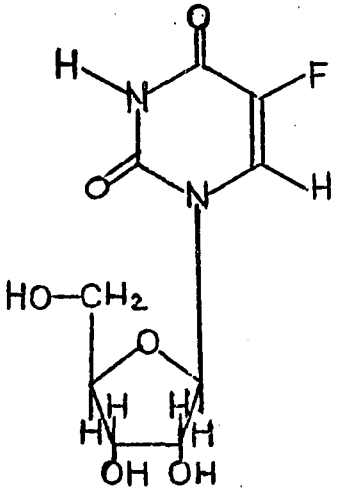
Uracil



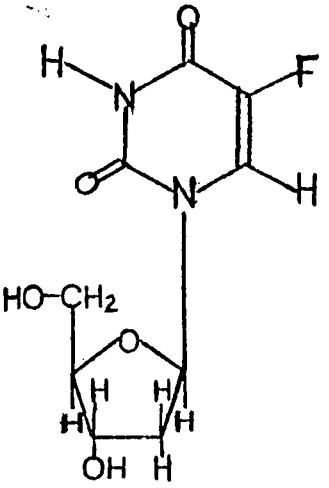
5-Fluorouracil



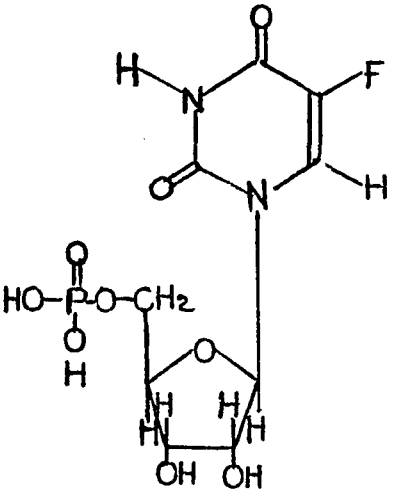
Thymine



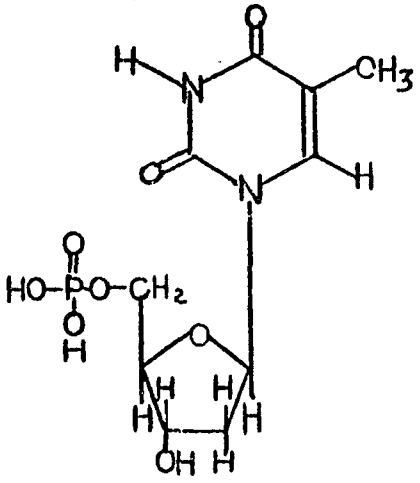
5-Fluorouridine



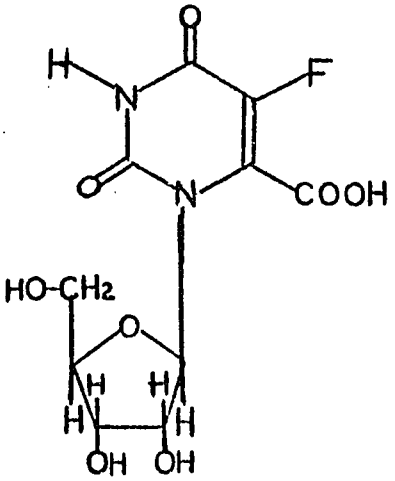
5-Fluoro-2'-deoxyuridine



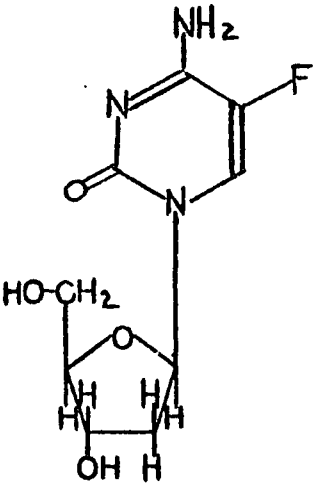
5-Fluorouridylic acid



Deoxythymidylic acid



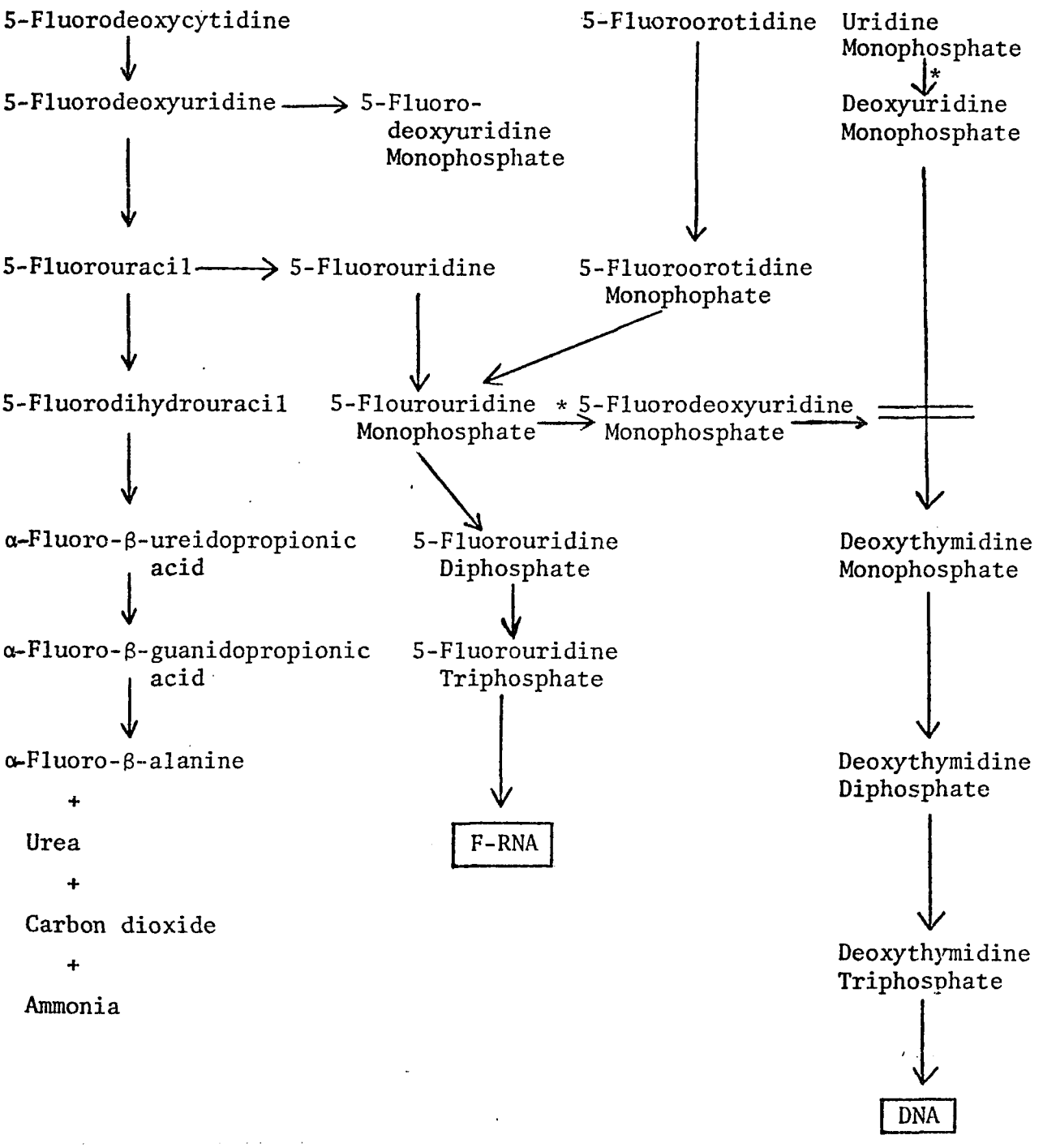
5-Fluoroorotidine



5-Fluorodeoxycytidine

FIGURE 2

METABOLISM OF 5-FLUOROURACIL

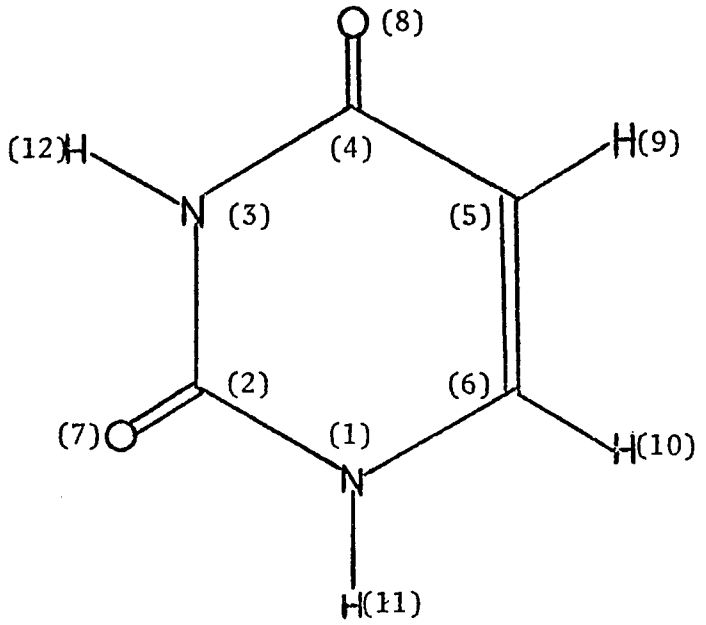


*Via a multi-step pathway involving reduction to the deoxyribose form at the diphosphate level.

FIGURE 3

Bond Distances and Angles of Uracil (Stewart)[67]

	Å
N (1) - C (2)	1.371
C (2) - N (3)	1.376
N (3) - C (4)	1.371
C (4) - C (5)	1.430
C (5) - C (6)	1.340
C (6) - N (1)	1.358
C (2) - O (7)	1.215
C (4) - O (8)	1.245
C (5) - H (9)	0.931
C (6) - H (10)	0.957
N (1) - H (11)	0.836
N (3) - H (12)	0.877



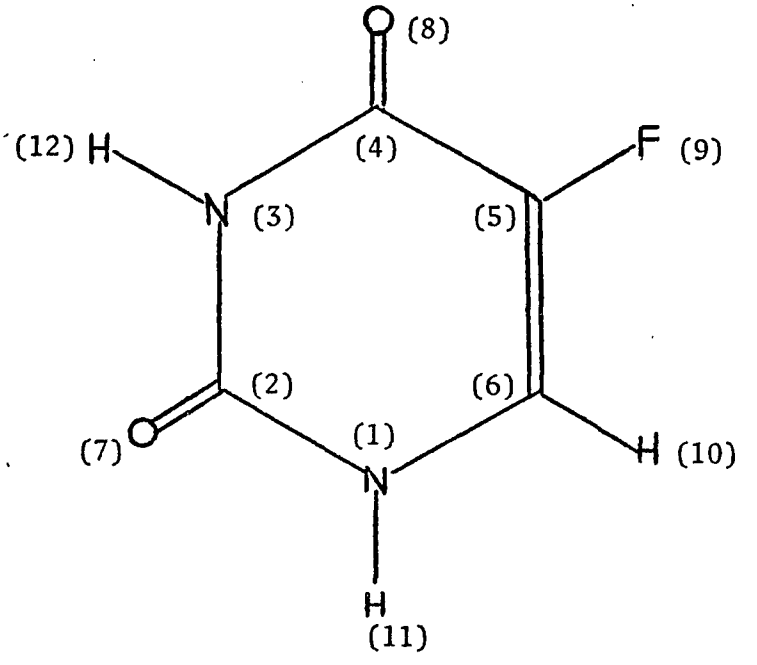
N (1) - C (2) - N (3)	114.0°	C (5) - C (4) - O (8)	125.3°
C (2) - N (3) - C (4)	126.7	C (4) - C (5) - H (9)	118.1
N (3) - C (4) - C (5)	115.5	C (6) - C (5) - H (9)	123.0
C (4) - C (5) - C (6)	118.9	C (5) - C (6) - H (10)	132.2
C (5) - C (6) - N (1)	122.3		
C (6) - N (1) - C (2)	122.6*	N (1) - C (6) - H (10)	114.5
N (1) - C (2) - O (7)	123.7	C (6) - N (1) - H (11)	122.2*
N (3) - C (2) - O (7)	122.3	C (2) - N (1) - H (11)	115.2
N (3) - C (4) - O (8)	119.2	C (2) - N (3) - H (12)	117.8
		C (4) - N (3) - H (12)	115.5

* Corrected values

FIGURE 4

Bond Distances and Angles of 5-Fluorouracil (Fallon)[68]

	Å
N (1) - C (2)	1.40
C (2) - N (3)	1.40
N (3) - C (4)	1.39
C (4) - C (5)	1.46
C (5) - C (6)	1.35
C (6) - N (1)	1.39
C (2) - O (7)	1.20
C (4) - O (8)	1.24
C (5) - F (9)	1.36
C (6) - H (10)	1.10*
N (1) - H (11)	1.00*
N (3) - H (12)	1.00*



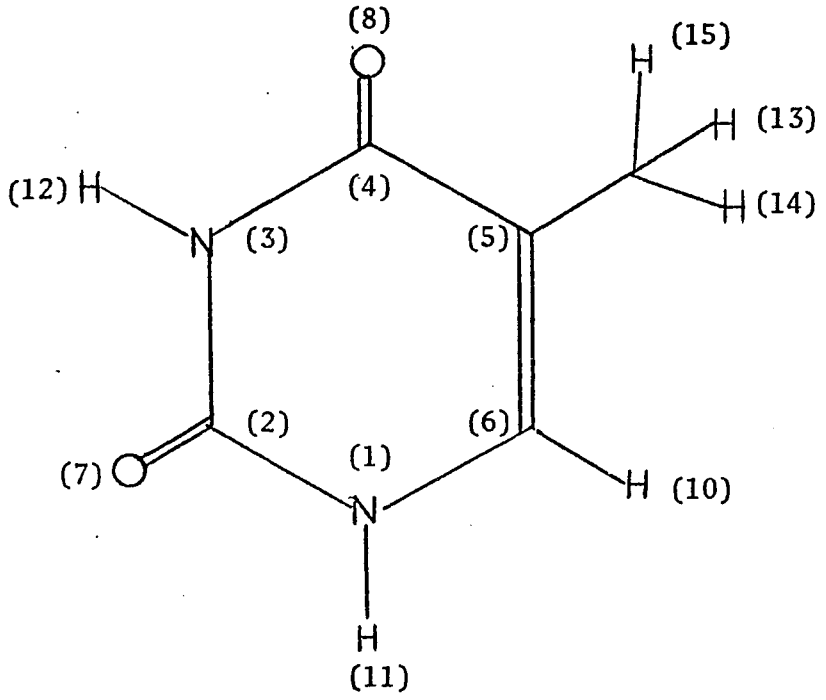
N (1) - C (2) - N (3)	116°	C (5) - C (4) - O (8)	126°
C (2) - N (3) - C (4)	127	C (4) - C (5) - F (9)	113
N (3) - C (4) - C (5)	112	C (6) - C (5) - F (9)	122
C (4) - C (5) - C (6)	125	C (5) - C (6) - H (10)	121*
C (5) - C (6) - N (1)	118	N (1) - C (6) - H (10)	121*
C (6) - N (1) - C (2)	122	C (6) - N (1) - H (11)	119*
N (1) - C (2) - O (7)	121	C (2) - N (1) - H (11)	119*
N (3) - C (2) - O (7)	123	C (2) - N (3) - H (12)	116.5
N (3) - C (4) - O (8)	122	C (4) - N (3) - H (12)	116.5

*Estimated

FIGURE 5

Bond Distances and Angles of Thymine (Ozeki et al)[69]

	Å
N (1) - C (2)	1.314
C (2) - N (3)	1.345
N (3) - C (4)	1.413
C (4) - C (5)	1.476
C (5) - C (6)	1.369
C (6) - N (1)	1.408
C (2) - O (7)	1.246
C (4) - O (8)	1.193
C (5) - C (9)	1.522
C (6) - H (10)	1.100
N (1) - H (11)	1.000
N (3) - H (12)	1.000
C (9) - H (13)	1.090
C (9) - H (14)	1.090
C (9) - H (15)	1.090



N (1) - C (2) - N (3)	118°	C (5) - C (4) - O (8)	125°
C (2) - N (3) - C (4)	126	C (4) - C (5) - C (9)	119
N (3) - C (4) - C (5)	114	C (6) - C (5) - C (9)	122**
C (4) - C (5) - C (6)	119	C (5) - C (6) - H (10)	120*
C (5) - C (6) - N (1)	120	N (1) - C (6) - H (10)	120*
C (6) - N (1) - C (2)	123	C (6) - N (1) - H (11)	118.5*
N (1) - C (2) - O (7)	121	C (2) - N (1) - H (11)	118.5*
N (3) - C (2) - O (7)	121	C (2) - N (3) - H (12)	117*
N (3) - C (4) - O (8)	121	C (4) - N (3) - H (12)	117*

*Estimated
**Corrected

FIGURE 6

Bond Distances of Thymine (Clementi et al)[70]

	Å
N (1) - C (2)	1.38
C (2) - N (3)	1.38
N (3) - C (4)	1.38
C (4) - C (5)	1.41
C (5) - C (6)	1.38
C (6) - N (1)	1.35
C (2) - O (7)	1.22
C (4) - O (8)	1.22
C (5) - C (9)	1.54
C (6) - H (10)	1.10
N (1) - H (11)	1.00
N (3) - H (12)	1.00

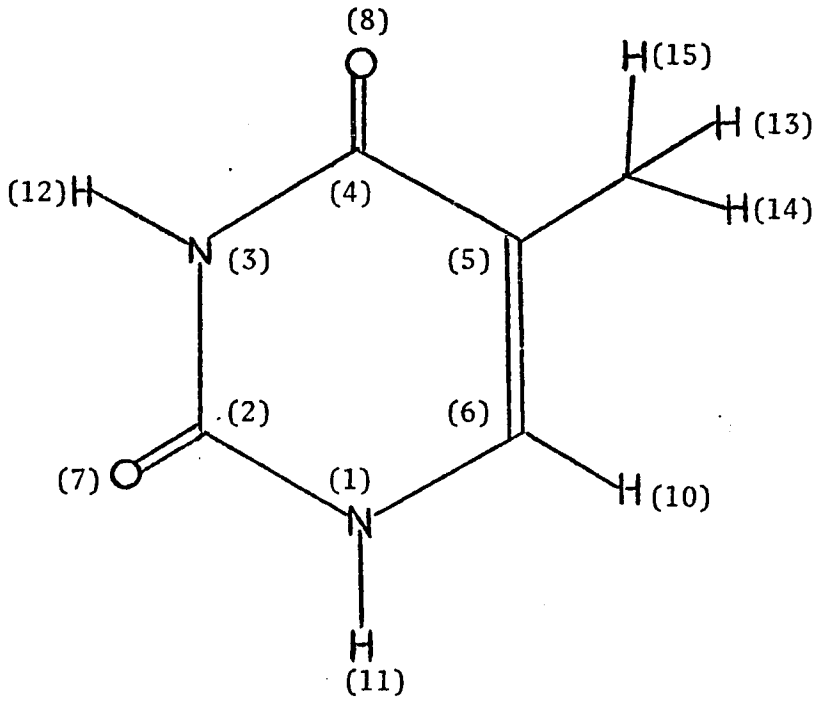
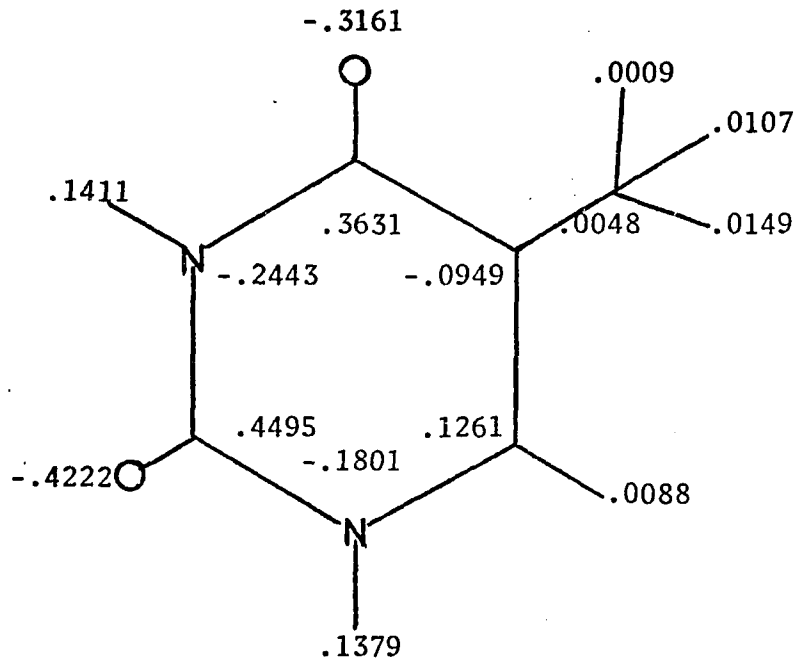


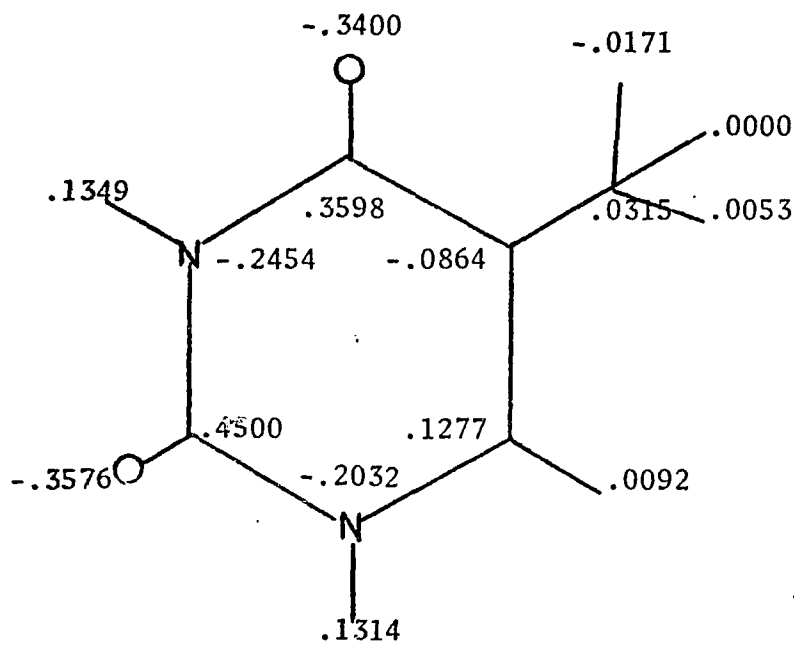
FIGURE 7

CHARGE DISTRIBUTION

THYMINE



THYMINE *

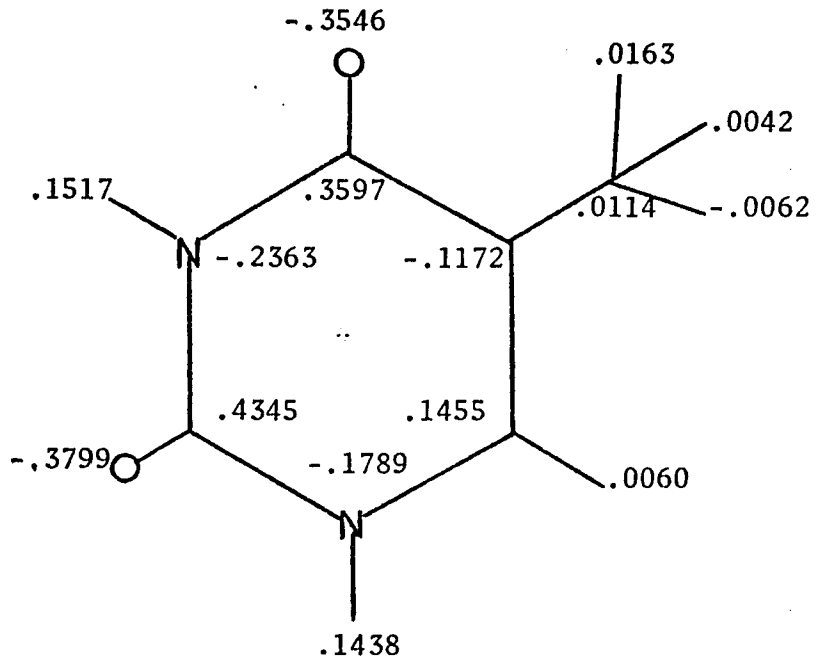


*The bond angles and bond lengths are those of 5-fluorouracil with the methyl group substituted for fluorine.

FIGURE 8

CHARGE DISTRIBUTION

THYMINE *

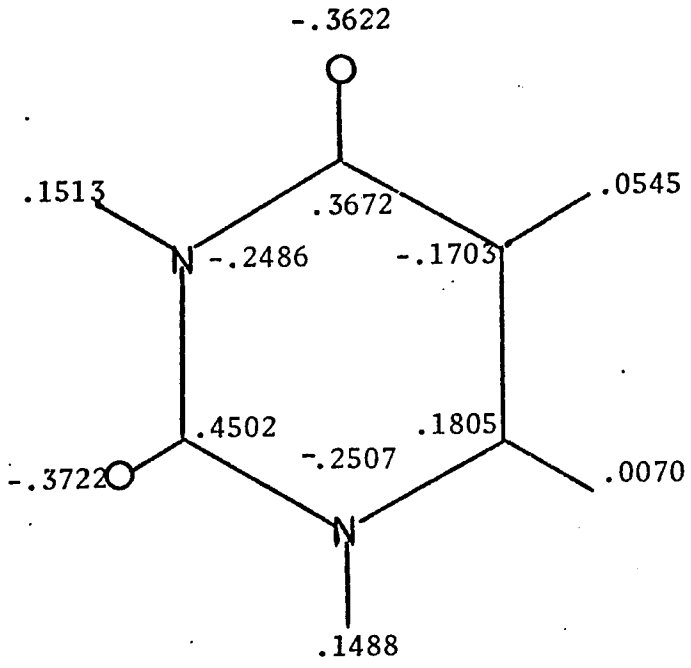


*The molecular dimensions are those obtained by an ab initio calculation[70]. The special arrangement of the hydrogen atoms of the methyl group is the same as that used in the ab initio treatment.

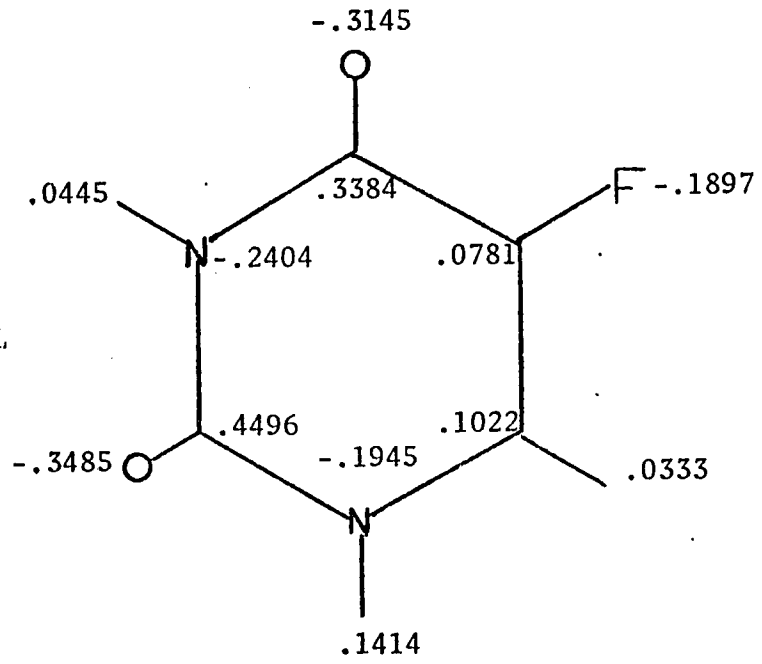
FIGURE 9

CHARGE DISTRIBUTION

URACIL



5-FLUOROURACIL



SECTION V

REFERENCES

1. Rutman, R. J., Cantarow, A., and Paschkis, K. E.; Cancer Res. 14, 119 (1954).

2. Cohen, S. S., and Barner, H. D.; Bacteriol. 71, 588 (1956).

3. Heidelberger, C., Chaudhuri, N. K., Dannenburg, P., Mooren, D., Griesbach, L., Duschinsky, R., Schnitzer, R. J., Pleven, E., and Scheiner, J.; Nature 179, 663 (1957).

4. Heidelberger, C., Griesbach, L., Cruz, O., Schnitzer, R. J., and Grunberg, E.; Proc. Soc. Exptl. Med. 97, 470 (1958).

5. Chaudhuri, N. K., Montag, B. J., and Heidelberger, C.; Cancer Res. 18, 318 (1958).

6. Heidelberger, C., Griesbach, L., Montag, B. J., Mooren, D., and Cruz, O.; Cancer Res. 18, 305 (1958).

7. Dannenberg, P. B., Montag, B. J., and Heidelberger, C.; Cancer Res. 18, 329 (1958).

8. Heidelberger, C., Ghobar, A., Baker, R. K., and Mukherjee, D. L.; Cancer Res. 20, 897 (1970).

9. Lehninger, A. L., "Biochemistry: The Molecular Basis of Cell Structure and Function", 2nd Ed., pp 736-739, Worth Publishers, Inc., New York, N. Y. (1975).

10. Davidson, J. N., "The Biochemistry of Nucleic Acids", 7th Ed., p. 222 Academic Press, New York, N. Y. (1972).

11. Cohen, S. S., Flaks, J. G., Barner, H. D., Loeb, M. R. and Lichtenstein, J., Proc. Natl, Acad. Sci. 44, 1004 (1958).

12. Hartmann, K. U., and Heidelberger, C., J. Biol. Chem. 256, 3006 (1961).

13. Chaudhuri, N. K., Mukherjee, K. L., and Heidelberger, C., "Biochemical Pharmacology", Vol. 1. pp. 328-341, Pergamon Press, Ltd., Gt. Britain (1958).

14. Mukherjee, K. L., and Heidelberger, C., J. Biol. Chem. 235, 433 (1960).

15. Mathews, C. K., and Cohen, S. S., J. Biol. Chem. 238, 367 (1963).

16. Blakley, R. L., J. Biol. Chem. 238, 2113 (1963).

17. Reyes, P., and Heidelberger, C., Mol. Pharmacol. 1, 14 (1965).

18. Friedkin, M., and Kornberg, A., in N. D. McElroy and B. Glass (Editors), "The Chemical Basis of Heredity", p. 609, Johns Hopkins Press, Baltimore, MD. (1957).

19. Pastore, E. J. and Friedkin, M., J. Biol. Chem. 237, 3802 (1962).

20. Lorenson, M. Y., Maley, G. F. and Maley, F. J., Biol. Chem. 242, 3332 (1967).

21. Capco, G. R., Krupp, J. R. and Mathews, C. K., Arch. Biochem. Biophys. 158, 726 (1973).

22. Dunlap, R. B., Harding, N. G. L. and Huennekens, F. M., Biochemistry 10, 88, (1971).

23. Phear, E. A. and Greenberg, D. M., J. Am. Chem. Soc., 79, 3737 (1951).

24. Greenberg, D. M., Nath, R. and Humphreys, G. K., J. Biol. Chem. 236, 2271 (1961).

25. Wahba, A. J. and Friedkin, M. J., Biol. Chem. 237, 3794 (1962).

26. Kalman, T. I., Ann. N. Y. Acad Sci. 186, 166 (1971).

27. Langenbach, R. J., Danenberg, P. V. and Heidelberger, C., Biochem. Biophys. Res. Commun. 48, 1565 (1972).

28. Wilson, R. S. and Mertes, M. P., Biochemistry 12, 2879 (1973).

29. Sommer, H. and Santi, D. V., Biochem. Biophys. Res. Commun. 57, 689 (1974).

30. Santi, D. V., McHenry, C. S. and Sommer, H., Biochemistry 13, 471 (1974).

31. Santi, D. V. and Pogolotti, A. L., J. Heterocycl. Chem. 8, 265 (1971).

32. Pogolotti, A. L. and Santi, D. V., Biochemistry 13, 456 (1974).

33. Cummings, T. F. and Shelton, J. R., J. Org. Chem. 25, 419 (1960).

34. Danenberg, P. V., Langenbach, R. J. and Heidelberger, C., Biochemistry 13, 926 (1974).

35. Sharma, R. K. and Kisliuk, R. L., Fed. Proc. Fed. Amer. Soc. Exp. Biol. 32, 591 (1973).

36. Ikeda, N., Yakugaku Zasshi 75, 1073 (1955).

37. Fukui, K., Immamura, A. and Nagata, C., Bull. Chem. Soc. Japan 33, 122 (1959).

38. Dodd, M. C. and Stillman, W. B., J. Pharmacol. Exp. Ther. 82, 11 (1944).

39. Yoneda, F. and Nitta, Y., Chem. Pharm. Bull. Japan 12, 1264 (1964).

40. Ban, T. and Nagata, C., Jap. J. Pharmacol. 15, 32 (1965).

41. Neely, W. B., Mol. Pharmacol. 1, 137 (1965).

42. Pullman, B. and Valdemoro, C., Biochim. Biophys. Acta 43, 548 (1960).

43. Farkas, M. and Kruglyak, J. A., Nature (London) 223, 523 (1969).

44. Fukui, K., Nagata, C. and Yonezawa, T., J. Amer. Chem. Soc. 80, 2267 (1958).

45. Cocordano, M., Ricord, J. and Julg, A., J. Theor. Biol. 12, 129 (1966).

46. Cocordano, M., C. R. Acad. Sci. 226, 897 (1968).

47. Perault, A. and Pullman, B., Biochim. Biophys. Acta 66, 86 (1963).

48. Singer, J. A. and Purcell, W. P., J. Med. Chem. 10, 754 (1962).

49. Purcell, W. P. and Sundarem, K., ibid, 12, 18 (1969).

50. Snyder, S. H. and Merril, C. R., Proc. Nat. Acad. Sci. U.S. 54, 258 (1965).

51. Yoneda, F. and Nitta, Y., Chem. Pharm. Bull. Japan 13, 579 (1965).

52. Pullman, A. and Pullman B., "Cancerisation par les Substances Chimiques et Structure Moleculaire", Masson, Paris (1955).

53. Mason, R., Nature (London) 181, 820 (1958).

54. Chalvet, O. and Mason, R., ibid, 192, 1070 (1961).

55. Pullman, A. and Pullman, B. ibid, 196, 228 (1962).

56. Pullman, A. and Pullman, B. ibid, 199, 4892 (1963).

57. Pullman, A., "Quantum Aspects of Polypeptides and Polynucleotides" (M. Weissbluth, ed.) p. 47, Wiley (Interscience) New York (1963).

58. Pullman, A., C. R. Soc. Biol. 139, 1956 (1945).

59. Pople, J. A. and Segal, G. A., J. Chem. Phys. 44, 3289 (1966).

60. Roothan, C. C. J., Rev. Mod. Phys. 23, 69 (1951).

61. Pople, J. A., Santry, D. P., and Segal, G. A., J. Chem. Phys. 43, S. 129 (1965).

62. Pople, J. A. and Segal, G. A., ibid, 43, S 136 (1965).

63. Santry, D. P. and Segal, G. A., ibid, 47, 158 (1967).

64. Mulliken, R. S., J. Chem. Phys. 23, 1833, 1841, 2338, 2343 (1955).

65. Stewart, R. F., Acta Cryst. 23, 1102 (1967).

66. Fallon, L. III, ibid, B 29, 2549 (1973).

67. Ozeki, K., Sakabe, N., and Tanaka, J., ibid. B 25, 1038 (1968).

68. Clementi, E., André, M. Cl., Klint, D., and Hahn, D., Acta Phys. Acad. Sci. Hung, 27, 493 (1969).

69. Kulakowska, I., Geller, M., Lesyng, B., and Wierchowski, K. L., Biochim. Biophys. Acta 361, 119 (1974).

70. Danilov, V. I., Dopov. Akad. Nauk Ukr. RSR, Ser. A 1, 61 (1976), [Chem. Abstr., 84: 104854g (1976)].