2. It was proved to be borneol by conversion into its phenylurethan, m. p. 140–141°, and its oxidation to camphor, m. p. 176°.

3. It was proved to be free from isoborneol.

CHAPEL HILL, NORTH CAROLINA

[CONTRIBUTION FROM THE DEPARTMENT OF CHEMISTRY, YALE UNIVERSITY]

RESEARCHES ON PYRIMIDINES. C111. THE DISCOVERY OF 5-METHYL-CYTOSINE IN TUBERCULINIC ACID, THE NUCLEIC ACID OF THE TUBERCLE BACILLUS[1]

BY TREAT B. JOHNSON AND ROBERT D. COGHILL[2]

RECEIVED AUGUST 5, 1925 PUBLISHED NOVEMBER 5, 1925

Introduction[3]

The problems of the greatest interest and importance to the organic chemist who is interested in the life processes of animal and plant cellular organisms are those dealing with a more accurate determination and definition of the composition and molecular structure of the fundamental, constructive constituents of the cell. The biological development of cellular organisms is dependent upon the interaction of characteristic organic combinations functioning in the living cell, and the normal life changes taking place during its growth are only revelations of profound and progressive alterations in chemical structures.

Of the many forms of bacterial cells which have been shown to exist in nature, that group which embraces the pathogenic type and operates as an enemy to mankind deserves the first consideration. Fortunately, the tubercle bacillus, which has probably received more attention than any other representative of this group, is an organism which can now be obtained easily in large quantity and thereby made available for accurate chemical research. Due to the progress that has been made in the artificial cultivation of this pathogenic organism, and through the assistance of the Research Committee of the National Tuberculosis Association, it has been made possible for us to obtain the interesting chemical results which we now report in this paper. The questions whether any constitutional differences exist between pathogenic and non-pathogenic organisms and whether such differences are detectable by application of the chemical method of research, are important ones and will receive attention during

[1] This research was supported and made possible by a special grant given by the National Tuberculosis Association, for which the writers hereby desire to express their thanks and sincere appreciation.

[2] Holder of the National Tuberculosis Association Research Fellowship, 1924–1925.

[3] A report of the results of this research was presented before the Organic Division of the American Chemical Society, at the Spring Meeting in Baltimore, April, 1925.

the progress of future investigations in this promising field of chemical bacteriology.

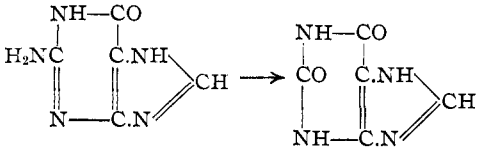
The Purine and Pyrimidine Constituents of Nucleic Acids.—So far as the writers are aware, no exact analysis revealing the actual distribution of nitrogen in animal, plant or bacterial cells has ever been made. Two important groups of nitrogen combinations are known to function in the life changes of all cells, namely, nucleic acids and proteins, and their presence can easily be shown, but we have a very limited knowledge of what proportion they represent of the total nitrogen content and regarding other forms (besides phosphatide combinations) this element actually functions in these organisms. Evidence is accumulating, however, as a result of recent research in this field, that a very large proportion of the nitrogen present in the protoplasmic structure of the cell is non-protein in nature. Of these two normal, nitrogenous constituents the nucleic acids occur in the smaller amount. Notwithstanding this fact, it has been possible to separate this combination in several cases in a high state of purity, and as a result of exhaustive research we have finally been enabled to formulate a fairly clear conception of their chemical constitution.[4]

It is now nearly 50 years since nucleic acids were first isolated from natural sources and attempts were made by Miescher, Kossel and others to establish their chemical structure by studying the products of hydrolytic decomposition. As a result of the progress already made by many workers in this field, we recognize today the following types of compounds as the constructive units of a nucleic acid molecule, namely, purines, pyrimidines, sugars (hexoses and pentoses) and phosphoric acid. Our present knowledge of the chemistry of nucleic acids prepared from micro-organisms has been acquired almost entirely from investigations dealing with only two acids, namely, the nucleic acid isolated from yeast cells, and tuberculinic acid,[5] which is a constituent of the tubercle bacillus.

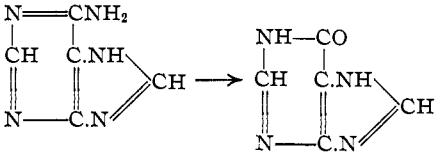
According to our present knowledge of the purine chemistry of nucleic acids it has been definitely established that only two purine combinations function as structural units of the nucleic acid molecule, namely, guanine (I) and adenine (III). Both of these are amino purines and are characterized by their behavior on hydrolysis, being easily de-aminized with formation of their corresponding oxypurines, namely, xanthine (II) and hypoxanthine (IV). This transformation is brought about not only artificially by hydrolysis with acids, but also normally during cell autolysis under the influence of specific enzymes which are present in the cell fluids. The generic relationship of these various purines derived from nucleic acid is expressed below.

[4] Walter Jones, "Nucleic Acids, Their Chemical Properties and Physiological Conduct," Longmans, Green and Co., New York and London, **1920.**

[5] Johnson and Brown, *J. Biol. Chem.*, **54,** 721 (1922); **57,** 199 (1923).



I Guanine II Xanthine



III Adenine IV Hypoxanthine

In the case of the pyrimidine compounds the biological relationship between the various combinations found among the products of hydrolysis has never been satisfactorily established. There has been great uncertainty as to which pyrimidines function as primary constituents. Thus far, all nucleic acids that have been carefully examined yield one or more of the three pyrimidine combinations, namely, uracil (VI), thymine (VIII) and cytosine (V). The observations made hitherto would seem to indicate that the nucleic acids of plant origin, such as tritico-nucleic acid of wheat and yeast nucleic acid, always yield uracil and cytosine on hydrolysis, while from nucleic acids of animal origin cytosine is found associated with thymine. In other words, the nature of the pyrimidines obtained is dependent upon the source of the nucleic acid examined. Whether this differentiation will hold after we have examined more nucleic acids and have developed more thoroughly our methods of pyrimidine analysis and identification remains to be established.

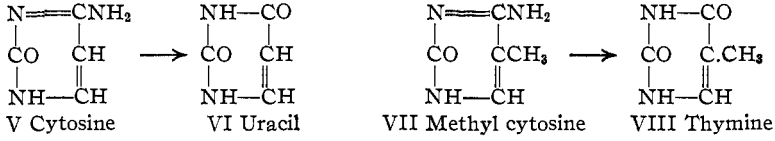
Notwithstanding the fact that it has been known for a long time[6] that cytosine can be deaminized easily by hydrolysis with acids with formation of uracil, there has never been agreement among workers in this field regarding the biological significance of this change. While good evidence has been presented pointing to the fact that the uracil obtained from a nucleic acid by hydrolysis is a secondary product,[7] the conclusions have never been convincing for two reasons: (1) no characteristic enzymes corresponding to the purine nucleotidases or nucleosidases have been shown to be present in the cell fluids, which have the power to stimulate pyrimidine de-aminization during cell autolysis,[8] and (2) a higher amino-pyrimidine corresponding in structure to cytosine (V), and capable of being transformed into thymine (VIII) by hydrolysis, has never been identified among the products formed by hydrolysis of nucleic acids. In other words, there has been a missing link in a chain of chemical and

[6] Kossel and Steudel, *Z. physiol. Chem.*, **38**, 49 (1903).

[7] Jones and Perkins, *J. Biol. Chem.*, **62**, 557 (1925).

[8] Ref. 4, p. 90.

biological evidence, which has prevented us from presenting a picture of nucleic acid-pyrimidine chemistry corresponding to that revealed in the case of the purine compounds. The introduction of a fourth pyrimidine, methyl-cytosine (VII) into our family of cell pyrimidines would enable us to complete our series and thereby express, as represented below, a generic relationship among these four compounds corresponding to that existing in the case of the cell purines. This relationship once having been established by discovery of the required pyrimidine (VII), the next problem



of immediate biochemical interest and importance would be that of proving whether these aminopyrimidines (V and VII) are primary structural units of the nucleic acid molecule, and whether the pyrimidines uracil and thymine, like the purines xanthine and hypoxanthine are secondary products resulting by de-aminization of their corresponding amino compounds (V and VII), respectively.

The Discovery of 5-Methyl-cytosine.—In 1904, Wheeler and Johnson contributed their fifth publication on pyrimidines[9] entitled—"Researches on Pyrimidines: 5-Methyl-cytosine," and in the introduction of that paper is recorded the following statement: "If uracil is to be reckoned as cytosine in the nucleic acids then 5-methyl-uracil or thymine possibly results by the cleavage of a corresponding base, namely, 5-methyl-cytosine or the isomeric 2-amino-5-methyl-6-oxypyrimidine. We therefore decided to prepare these new bases and to compare their properties with those of cytosine." In the paper referred to above is given not only a method of synthesis whereby this first aminopyrimidine (VII) can be obtained easily in quantity, but also a complete description of its chemical properties. In the following paper of the series on pyrimidine researches, Johnson and Clapp[10] described a practical synthesis of 2-amino-5-methyl-6-oxypyrimidine (X).

Ever since the publication of this synthesis of 5-methyl-cytosine in 1904, the writer has anticipated the discovery of this pyrimidine among the products of hydrolysis of a nucleic acid, but it was not until it was made possible for us to make an examination of the nucleic acid obtained from tubercle bacillus that we were successful in showing that this combination exists in nature as a structural unit of a nucleic acid molecule. We now find that this aminopyrimidine, 5-methyl-cytosine, is one of the products of hydrolysis of tuberculinic acid, and is present associated with

[9] Wheeler and Johnson, *Am. Chem. J.*, **31**, 591 (1904).

[10] Johnson and Clapp, *ibid.*, **32**, 130 (1904).

the pyrimidine cytosine in the base fraction precipitated by phosphotungstic acid.

The sources of the material from which we obtained our tuberculinic acid were desiccated tubercle bacilli grown on beef broth, which were furnished to us gratuitously by the Mulford Company, and by Parke, Davis and Company of Pennsylvania and Michigan, respectively. The tuberculinic acid was prepared from these bacilli according to the method previously described by Johnson and Brown,[5] and then subjected to hydrolysis with 25% sulfuric acid and the pyrimidines were isolated and identified in the customary manner. Cytosine was obtained in pure form as usual by crystallization from a concentrated aqueous solution. As 5-methyl-cytosine was known to be very similar in chemical properties to cytosine,[9] it was to be expected that this base would be found, if present, in the cytosine filtrate. That it would be difficult to separate in pure condition by direct crystallization was also realized on account of its extreme solubility, being 4.5 times as soluble in water as cytosine. This property very probably accounts for the fact that this base has not been identified earlier.

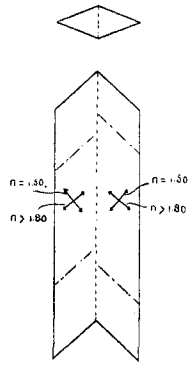


Fig. 1.—Diagrammatic representation of a crystal of cytosine picrate.

In order to separate and identify the pyrimidine, advantage was taken of its property of combining with picric acid to form a crystalline picrate. Accordingly, the aqueous filtrate, left after crystallization in the separation of cytosine, was treated with an excess of this acid and the resulting precipitate then purified by crystallization from hot water. After repeated fractional crystallizations from this solvent the salt was finally obtained in a satisfactory crystalline condition, but the crystals were always accompanied by a trace of cytosine picrate. These two salts crystallize in their own characteristic habits, which are easily recognized by means of the petrographic microscope. By a careful study of these crystalline picrates from the cytosine fraction of the tuberculinic acid and by comparison of their optical properties with those of picrate crystals prepared from the pure synthetic pyrimidines of known structure, we were able to establish conclusively the presence of 5-methyl-cytosine as a product of hydrolysis of tuberculinic acid. Cytosine picrate crystallizes in long needles which are twinned parallel to the length of the prism and perpendicular to its face. This unique property of the salt, together with the characteristic oblique terminations of the crystals, makes it very easily distinguished from 5-methyl-cytosine picrate which crystallizes in much smaller needles with square ends. The crystals also differ very markedly in their angles of extinction and their indices of refraction.

Specimens of the purified picrates prepared from the naturally occurring pyrimidine bases and also of the synthetic pyrimidines were submitted to Professor William E. Ford, Head of the Department of Mineralogy, Yale University, for a careful crystallographic investigation. After a thorough comparison of the several crystalline samples submitted for his examination, he has pronounced the crystalline picrate as identical in every respect with the synthetical 5-methyl-cytosine picrate. We desire to express here our appreciation of Professor Ford's assistance and coöperation in this important piece of research.

Crystallographic Report

Cytosine Picrate.[11]—This picrate occurs in very slender, blade-like crystals having a brilliant golden-yellow color. The crystals are several times longer in proportion to their breadth than as represented in Fig. 1. Under the microscope it is seen that the slender crystals are twinned individuals—the trace of the twinning plane dividing them parallel to their length. The end of a crystal is commonly terminated by two sloping planes, the traces of which make approximately angles of 47° with the prism edges. In some cases a re-entrant angle—formed by the same planes—is to be observed at the other end of the crystal. Cleavages also exist parallel to these terminal planes. The prism faces are vertically striated, but a crystal mounted on the goniometer gave 42° as the prism angle. The vibration direction of the faster ray makes an angle of about 38° with the prism edge. The crystal, therefore, extinguishes with crossed nicols at approximately the 45° position—the two halves of the twin, however, extinguish at about 14° to each other. The measured extinction angles of the faster ray with the prism edge vary from 34° to 48°. This is due apparently to the fact that the crystals do not lie on the slide on a pinacoid face but rather upon one or the other of the prism planes. The crystals show a very high birefringence, the lower index being measured at 1.50 while the higher one exceeds 1.80 in value. They exhibit a slight pleochroism, showing an orange-yellow for the vibration direction of the slower ray and a pure yellow at right angles to this. The crystal system is monoclinic, the twinning plane being the orthopinacoid (100). The law of twinning is similar to that commonly observed in the case of the crystals of gypsum.

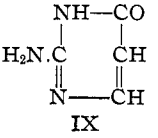
5-Methyl-cytosine Picrate.—This salt occurs in exceedingly minute, lath-like crystals. They possess a most brilliant luster and are a golden-yellow. Between crossed nicols they show parallel extinction. They show a high birefringence, the two indices that could be determined being 1.44 and greater than 1.80. The faster ray vibrates parallel to the elongation of the crystal and the slower ray across it. The crystals are orthorhombic in their symmetry.

The Unknown Picrate from Tuberculinic Acid.—This salt occurs in mats of very slender prisms, with an occasional prism detached. They agree crystallographically in every detail given above for 5-methyl-cytosine picrate prepared from synthetic 5-methyl-cytosine, and in my judgment must be identical with it. It is very easy to distinguish them from the crystals of cytosine picrate.

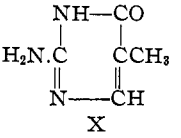
In the light of the results of our research it will be very important, hereafter, in all work dealing with a study of the hydrolytic products of nucleic acids to search carefully for the two aminopyrimidines, isocytosine or 2-amino-uracil (IX)[11] and 2-amino-5-methyl-6-oxypyrimidine (X).[10] Both

[11] Wheeler and Johnson, *Am. Chem. J.*, **29**, 492 (1903).

compounds undergo deaminization under the hydrolytic influence of acids giving, respectively, uracil (VI) and thymine (VIII).



IX



X

In our work on tuberculinic acid we have already obtained some evidence indicating that another basic compound, besides cytosine (V) and 5-methyl-cytosine (VII), is present in the basic pyrimidine fraction. The study of these various pyrimidines will be continued this coming year in the Sterling Laboratory.

Summary

1. 5-Methyl-cytosine has been identified as a product of hydrolysis when tuberculinic acid, the nucleic acid present in the tubercle bacillus, is digested with sulfuric acid.

2. This pyrimidine base has been isolated in the form of its characteristic picrate, and the latter shown to be identical with the picrate of 5-methyl-cytosine previously synthesized by Wheeler and Johnson in 1904.

3. The discovery of this compound increases the number of pyrimidines functioning in the life changes of a cell to four, namely, cytosine, 5-methyl-cytosine, uracil and thymine.

NEW HAVEN, CONNECTICUT

NEW BOOKS

The Spectroscopy of X-Rays. By MANNE SIEGBAHN, Professor in the University of Upsala. Translated with the author's additions, by GEORGE A. LINDSAY, Assistant Professor of Physics in the University of Michigan. Oxford University Press, American Branch, New York, 1925. xii + 287 pp. 118 figs. 24 × 16 cm. Price $6.00 net.

In extolling the German edition of Professor Siegbahn's classic work [THIS JOURNAL, **46,** 2130 (1924)] this reviewer expressed the following proof of appreciation: "Publication in English speedily is greatly to be desired." Happily Dr. Lindsay has felt the urge to make this great monograph a part of the scientific literature in the English language, so that it might be readable and understandable by all those very numerous scientific folks who must struggle with German to such an extent that the substance is lost in the form, and an eminently satisfactory translation is the result of this service.

Further comment upon Professor Siegbahn's treatment of the subject matter would be merely repetition, even though this has seemed to us to gain in the English form. This new edition has been sponsored by the author and additions have been made to bring it completely up to the time