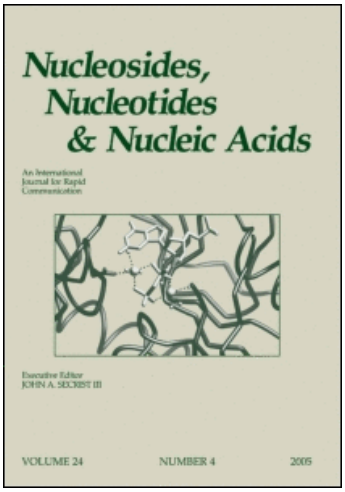


Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

Cis-trans Isomerization of the N-Formyl Amino Group in a Novel Modified Uracil Derivative: 6-Amino-3-methyl-5-(N-Formylamino)uracil

T. Srikrishnan[a]; R. Parthasarathy[a]; W. Pfleiderer[b]; Nimai C. De[a]; Girish B. Chheda[a]
[a] Biophysics Department, Roswell Park Cancer Institute, Buffalo, New York, USA [b] Department of Chemistry, Konstanz University, Konstanz, Germany

To cite this Article Srikrishnan, T. , Parthasarathy, R. , Pfleiderer, W. , De, Nimai C. and Chheda, Girish B.(1997) '*Cis-trans* Isomerization of the N-Formyl Amino Group in a Novel Modified Uracil Derivative: 6-Amino-3-methyl-5-(N-Formylamino)uracil', Nucleosides, Nucleotides and Nucleic Acids, 16: 3, 239 — 255
To link to this Article: DOI: 10.1080/07328319708001345
URL: http://dx.doi.org/10.1080/07328319708001345

NUCLEOSIDES & NUCLEOTIDES, 16(3), 239-255 (1997)

CIS-TRANS ISOMERIZATION OF THE N-FORMYL AMINO GROUP IN A NOVEL MODIFIED URACIL DERIVATIVE: 6-AMINO-3-METHYL-5-(N-FORMYLAMINO)URACIL

T. Srikrishnan+*, R. Parthasarathy+, W. Pfleiderer++, Nimai C. De** and Girish B. Chheda+**

+Biophysics Department, Roswell Park Cancer Institute, Buffalo, New York 14263, USA

and

++Department of Chemistry, Konstanz University, Konstanz, Germany

Abstract

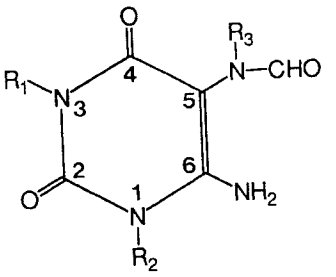
The modified nucleic acid base, 6-amino-3-methyl-5-(N-formylamino)uracil is one of three new modified uracil derivatives isolated and extracted from human urine in our laboratory. X-ray crystallographic investigation was undertaken to understand the cis-trans isomerization of the N-formylamino group about the C(ring)-N(formyl) bond. Crystals of the title compound are monoclinic, space group, $P2_1/n$, with unit cell dimensions a=16.932(2), b=10.565(1), c=4.298(1)Å, β=91.1(1)°, Z=4, D_{obs}=1.59 g.cm^{-3}, D_{calc}=1.591 g.cm^{-3}. Complete three dimensional data ($2\theta \leq 150°$ for CuKα) were collected on a CAD-4 diffractometer by the $\omega/2\theta$ scan method. The structure was solved by a straight forward application of multisolution techniques and refined by full-matrix least-squares method to a final R value of 0.044. The amino group at the 6-position is in the plane of the uracil ring, but the N-formyl amino group at the 5-position is twisted away from the uracil plane by ±82.2°, a conformation about half-way between the cis and trans positions of the N-formyl amino group. Consequently, the molecule in solution is able to be trapped as a mixture of cis-oid and trans-oid conformers giving rise to two peaks for aldehydric proton and two peaks in HPLC that show identical mass spectra. The uracil rings are base-paired across centers of inversion by a pair of N(1)-H(N1)...O(2) type of hydrogen bonding. Partial stacking of the bases is observed in the crystal structure.

---

* Author to whom correspondence should be addressed
**Deceased

Introduction

In our investigation of isolation and characterization of unusual metabolic substances present in human cancer urine, we have recently identified three new modified uracil derivatives; 6-amino-3-methyl-5-(N-formylamino)uracil (I), 6-amino-3-methyl-5-(N-formylmethylamino)uracil (II) and 6-amino-1-methyl-5-(N-formylmethylamino)uracil (III) (1,2). Previously, two modified uracils, 5-acetamido-6-amino-3-ethyluracil (3) and 1,3-dimethyl-6-amino-5-(N-formylmethylamino)uracil have been characterized from human urine (4).



I $R_1 = CH_3$; R_2 and $R_3 = H$

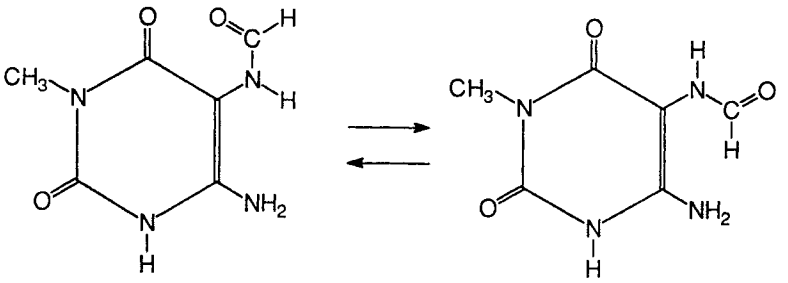
II $R_1 = CH_3$; $R_2 = H$; $R_3 = CH_3$

III $R_1 = H$; $R_2 = CH_3$; $R_3 = CH_3$

These modified uracils are derived from the methylated xanthines (5) which in turn may arise either from the dietary sources such as caffeine (6) or by the methylation of endogenously derived xanthines. These modified uracils are different from the methylated uracils present in tRNA (7) and have the potential of serving as tumor markers.

From structural considerations, these uracils show unusual resonance properties in nuclear magnetic resonance (NMR) spectra

and also exhibit unique behavior in high performance liquid chromatography. For example, the 6-amino-3-methyl-5-N-formylamino uracil in (I) in NMR spectra in DMSO-d_6 showed N^5 aldehyde proton and 6-amino protons as a pair of doublets. On heating these proton resonances for aldehyde and amino groups collapsed to singlets. Such temperature dependent behavior of these resonances and literature precedence (8) suggests that these compounds exist as cis and trans forms in solution.



In HPLC the highly purified amino uracil I exhibited two peaks. Each peak when injected separately generated the same two peaks. The ultraviolet spectra of each peak was identical. The LC-MS of these peaks also gave indistinguishable mass spectra, all indicating cis-trans isomerization of the N-formylamino group.

Because of the unusual behavior of these modified uracils in NMR spectrometry and high performance liquid chromatography, we directed our attention towards x-ray crystallographic studies to investigate the predominant form of the material, to determine the orientation of N-formylamino group to uracil ring (9) and to see if it participates in an internal C-H...O hydrogen bonding with the keto oxygen O(4). Furthermore, the 5-substituted uracils are an important class of compounds both biochemically and

pharmacologically, and as a result the x-ray crystallographic study may make a further contribution to the structure-activity relationships of these compounds.

The present paper deals with the crystal and molecular structure of the urinary uracil derivative, 6-amino-3-methyl-5-(N-formylamino)uracil (I).

Methods

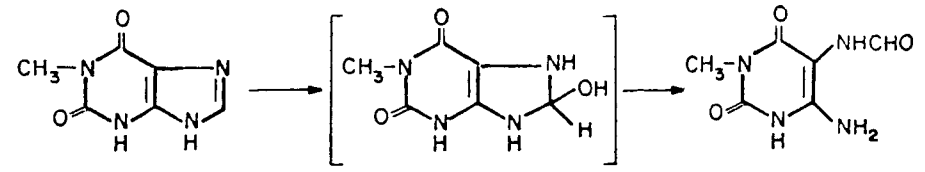
The 6-amino-3-methyl-5-(N-formylamino)uracil was synthesized according to the procedures reported by Pfleiderer (10). It was crystallized from methanol to give shiny transparent crystals.

Discussion

In contrast to the number of modified bases and nucleotides found in human urine that are derived from tRNA (7), these modified uracils are derived from the methylated xanthines suggesting that xanthines, once thought intractable, actually undergo metabolism to a significant degree to give ring-opened compounds.

Biochemically, it is of interest to note that apparently the process of N-demethylation occurs routinely to generate 1,3, 1,7 and 3,7-dimethylxanthines from caffeine. Mechanistically, the 1-methylxanthine undergoes enzyme-mediated hydration of the N^7-C^8 double bond to give 1-methyl-7,8-dihydrouric acid (2, Scheme 1), the latter compound then undergoes imidazole ring-opening to give the 6-amino-3-methyl-5-(N-formylamino)uracil (1, Scheme 1).

X-ray Crystallography: The title compound was crystallized from aqueous methanol as shiny thin transparent needles. Crystals of the modified uracil derivative are monoclinic, space group $P2_1/n_1$ with unit cell dimensions (at 22 ± 3°C) a = 16.932 (2), b = 10.565



Scheme 1

(1), c = 4.298 (1) Å, β = 91.1 (1)°, V = 768.8 Å^3, μ = 10.6 cm^{-1}, D_m = 1.59 g/c.c. (flotation in a mixture of bromoform and benzene), Dc = 1.591 g/c.c., F(000) = 384. Complete three-dimensional intensity data were collected using a crystal of dimensions 0.6 x 0.2 x 0.1 mm on a CAD-4 diffractometer by the $\omega/2\theta$ technique. The intensities of 1881 reflections (1298 with their $I \geq 3\sigma$) to the limit of $2\theta \leq 150°$ for CuKα (λ = 1.5418Å) were measured. The scan widths were calculated using the relation $(A + B \tan\theta)^0$ with A and B having values of 0.5 and 0.15 respectively. Aperture widths were determined using the relation $(4.0 + 1.2 \tan\theta)$mm. The maximum time spent on any reflection measurement was 100 seconds and the background count time was half the scan time. A faster scan was used for strong reflections. The intensities were monitored by measuring three reflections after every hour of x-ray exposure, and the variation of intensities was less than 5% during the complete data collection. The orientation matrix was checked every 100 refections. Out of 1881 reflections measured, 1298 were considered significant, based on the fact that the net count = peak -2 (left background and right background) was greater than 3σ (I). Lorentz and polarization corrections were applied to all the reflections.

The intensities of three reflections at $\chi \cong 90°$ were measured for all values of ϕ from 0° to 360° and the resultant curve of transmission as a function of ϕ was used to calculate the absorption for all the reflections. The average transmission factor was 0.89.

The structure was solved by a straight forward application of the multisolution technique (11) which yielded all the non-hydrogen atoms in the molecule. It was refined by full-matrix least-squares method, initially with isotropic thermal parameters. The reliability factor at the end of the isotropic refinement was 0.112. Further refinements were carried out with anisotropic thermal parameters for the non-hydrogen atoms. At the end of the anisotropic refinement the R factor was 0.082. A difference Fourier map was used to locate the atomic positions of all hydrogen atoms in the molecule. The structure was further refined using isotropic temperature factors for the hydrogen atoms and anisotropic thermal parameters for the non-hydrogen atoms and the final reliability index R defined as

$$\frac{\sum \| kF_o | - | F_c \|}{\sum | kF_o |}$$

was 0.044 for the observed 1298 reflections ($I \geq 3\sigma$). The weighted R factor was 0.081. The function minimized was $w(|F_o|^2-(1/k)|F_c|^2)$ where the weight $w=4|F_o|^2/\sigma(|F_o|^2)^2$ and $\sigma(|F_o|)^2=[\sigma^2(I)+p^2I^2]^{1/2}$ where $\sigma(I)$ is the standard deviation of intensity I based upon counting statistics; k is the scale factor and p is an ignorance factor used

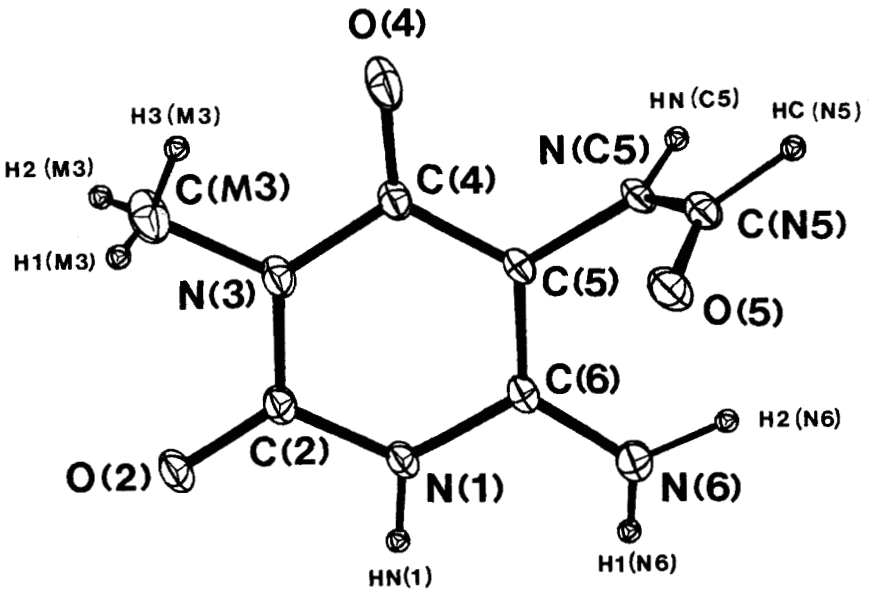


Fig. 1. An ORTEP drawing of the 3-methyl-5-(N-formylamino)uracil molecule showing the conformational details and the atomic numbering scheme used.

to down-weight intense reflections (p=0.05). The goodness of fit was 2.8. The final difference fourier was featureless with $|\Delta\rho|$= 0.2 e/Å^3. The calculations were done on the PDP 11/34 computer with the aid of the Enraf Nonius structure determination package (12). The atomic scattering factors were taken from the "International Tables for X-ray Crystallography" (13). The torsion angles were calculated using the program by Dr. S.T. Rao and the diagrams were prepared using the program ORTEP by Johnson (14).

The structure of the title compound is shown in Fig.1. Table 1 gives the final fractional atomic coordinates for all the atoms in the structure together with their standard deviations in parentheses. The bond distances and angles with their standard deviations are given in Table 2. The average standard deviation in

TABLE 1

Final Fractional Positional Parameters with estimated standard deviations given in Parentheses.

ATOM	X	Y	Z	B(Å²)
O(2)	-0.01359(8)	0.6505(1)	0.3847(4)	3.91(3)
N(6)	0.1885(1)	0.3918(2)	0.0946(4)	3.73(3)
O(5)	0.33810(8)	0.6269(2)	0.1715(3)	4.37(3)
N(1)	0.08626(8)	0.5251(2)	0.2278(4)	2.88(3)
N(C5)	0.25890(9)	0.5917(2)	-0.2505(4)	2.92(3)
N(3)	0.07717(8)	0.7338(2)	0.0606(4)	3.27(3)
O(4)	0.16886(8)	0.8142(2)	-0.2651(4)	4.39(3)
C(4)	0.1469(1)	0.7231(2)	-0.1121(4)	2.96(3)
C(2)	0.0468(1)	0.6375(2)	0.2307(4)	2.89(3)
C(5)	0.1862(1)	0.6045(2)	-0.0946(4)	2.75(3)

a bond distance involving the non-hydrogen atoms is 0.002Å and in a bond angle is 0.01°. The average standard deviation in a bond distance involving hydrogen atoms is 0.02Å. Table 3 lists the hydrogen bond distances and angles in the structure.

The amino group at the 6-position of the uracil ring is in the plane of the uracil ring but the N-formyl amino group at the 5-position is twisted away from the uracil plane by ±82.2°. As a consequence there is no intramolecular C-H...O hydrogen bonding. The N-formyl substituent of the six-membered conjugated ring is

Table 1 Continued

C(6)	0.1554(1)	0.5046(2)	0.0708(4)	2.69(3)
C(N5)	0.3283(1)	0.6071(2)	-0.1085(4)	3.04(3)
C(M3)	0.0364(1)	0.8552(2)	0.0611(7)	4.88(5)
H1(N6)	0.166(1)	0.340(2)	0.177(5)	3.4(5)*
H2(N6)	0.231(2)	0.365(2)	-0.021(6)	4.8(6)*
H(N1)	0.066(1)	0.471(2)	0.326(5)	3.6(5)*
H(NC5)	0.262(1)	0.578(2)	-0.433(6)	4.0(5)*
H(CN5)	0.374(1)	0.597(2)	-0.243(5)	3.9(5)*
H1(M3)	0.021(3)	0.869(5)	0.22(1)	13.0(2)*
H2(M3)	-0.008(2)	0.854(3)	-0.053(8)	7.4(8)*
H3(M3)	0.075(2)	0.907(3)	-0.005(9)	8.3(9)*

STARRED B-VALUES WERE REFINED ISOTROPICALLY.
ANISOTROPIC B VALUES ARE GIVEN IN THE FORM OF THE ISOTROPIC EQUIVALENT DISPLACEMENT PARAMETER DEFINED AS: (4/3) * [A2*B(1,1) + B2*B(2,2) + C2*B(3,3) + AB(COS GAMMA)*B(1,2) + AC(COS BETA)*B(1,3) + BC(COS ALPHA)*B(2,3)].

expected to be in the plane of the 6-membered ring. However the neighboring substituents cause steric hindrance and causes the N-formyl group to rotate out of the plane, as in the present structure. Both the amino group and the oxygen play a role in causing the rotation of the N-formyl group. This rotation causes a lengthening of the N(formyl)-C(formyl) bond to 1.322(1)Å. The orientation of the N-formyl amino group about the C(ring)-N(formyl) bond is half-way between the cis and trans positions, enabling the molecule in solution to be trapped as a mixture of cis-oid and trans-oid conformers. The presence of these two conformers is

TABLE 2

Bond distances (Å) and Bond Angles (°) with estimated standard derivations in parentheses.

Bond Distances

N(1)-C(2)	1.362(1)	C(N5)-O(5)	1.230(2)
C(2)-N(3)	1.361(2)	C(6)-N(6)	1.321(2)
N(3)-C(4)	1.412(1)	N(1)-H(N1)	0.79(2)
C(4)-C(5)	1.420(2)	C(M3)-H1(M3)	0.74(5)
C(5)-C(6)	1.382(2)	C(M3)-H2(M3)	0.89(2)
C(6)-N(1)	1.379(1)	C(M3)-H3(M3)	0.90(3)
C(2)-O(2)	1.236(1)	N(C5)-H(CN5)	0.99(2)
N(3)-C(M3)	1.456(2)	N(6)-H1(N6)	0.75(2)
C(4)-O(4)	1.227(1)	N(6)-H2(N6)	0.93(2)
C(5)-N(C5)	1.418(1)		
N(C5)-C(N5)	1.324(2)		

Bond Angles

N(1)-C(2)-N(3)	117.2(1)	C(4)-C(5)-C(6)	121.3(1)
N(1)-C(2)-O(2)	120.8(1)	C(4)-C(5)-N(C5	118.0(1)
N(3)-C(2)-O(2)	121.9(1)	N(C5)-C(5)-C(6)	120.7(1)
C(2)-N(3)-C(4)	123.5(1)	C(5)-C(6)-N(1)	117.7(1)
C(2)-N(3)-C(M3)	118.3(1)	C(5)-C(6)-N(6)	124.4(1)
C(M3)-N(3)-C(4)	118.2(1)	N(6)-C(6)-N(1)	117.9(1)
N(3)-C(4)-O(4)	118.9(1)	C(5)-N(C5)-C(N5)	122.7(1)
N(3)-C(4)-C(5)	116.1(1)	N(C5)-C(N5)-O(5)	125.0(1)
O(4)-C(4)-C(5)	125.0(1)		

TABLE 3

Hydrogen bond distances (in Å) and angles (°)

Donor	Hydrogen	Acceptor	Distances (in Å)			Angle (°)	Symmetry of the Acceptor atom
D	H	A	D-H	H...A	D...A	D-H...A	
N(1)	H(N1)	O(2)	0.79	2.01	2.795	173°	(-X 1-Y 1-Z)
N(C5)	H(NC5)	O(5)	0.80	2.21	2.870	158°	(X Y 1-Z)
N(6)	H1(N6)	O(5)	0.75	2.35	3.010	147°	(½-X Y-½ ½-Z)
N(6)	H2(N6)	O(4)	0.93	2.01	2.939	174°	(½-X Y-½ -Z-½)

indicated by the aldehydric proton which is a doublet and the 6-amino protons which also shown as doublet. On heating, these two doublets coalesce to singlets. The two peaks in HPLC with indistinguishable mass spectra as well as the fact that either peak generates the two peaks on loading to a column strongly indicate the presence of both cis-oid and trans-oid conformers.

Fig. 2 shows a stereoscopic view of the packing of the molecules in the unit cell. The planar uracil ring are base-paired across centers of inversion by a pair of N(1)-H(N1)...O(2) type of hydrogen bonding (Table 3). There is a partial stacking of the uracil bases in the crystal structure, with the distance between the planes of 3.41Å. The bases are stacked with N(1), N(3) and

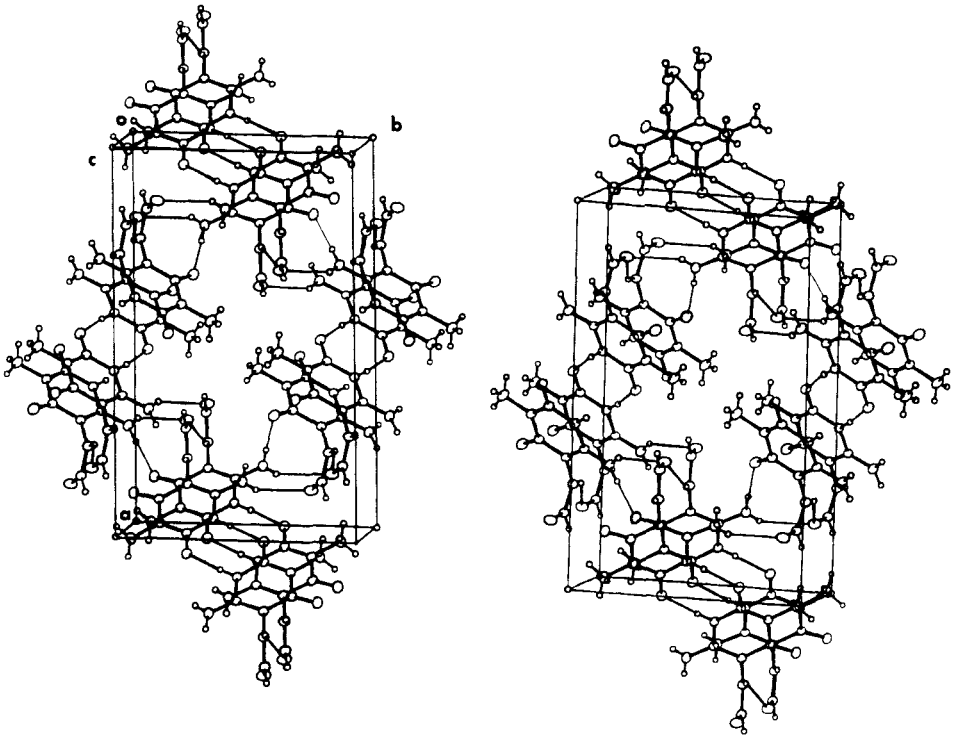


Fig. 2 A stereoscopic view of the packing of the molecules in the unit cell projected down c-axis showing the base stacking and the hydrogen bonds.

N(6) of one base overlapping on C(2), C(M3) and C(6) of another base related by a c-translation. Bases separated by a c-translation are also linked by a N-H...O hydrogen bonding involving the formyl amino nitrogen atom N(C5) of one base and O(5) of another base. The amino group N(6) takes part in two hydrogen bonds to O(4) and O(5) of two different bases related by the n-glide operation.

Energy Minimization and study of the cis-trans isomerization of the N-formyl amino group: We used the molecular modeling software SYBYL 6.2 version from the Tripos Associates, St. Louis, Mo (15) to study

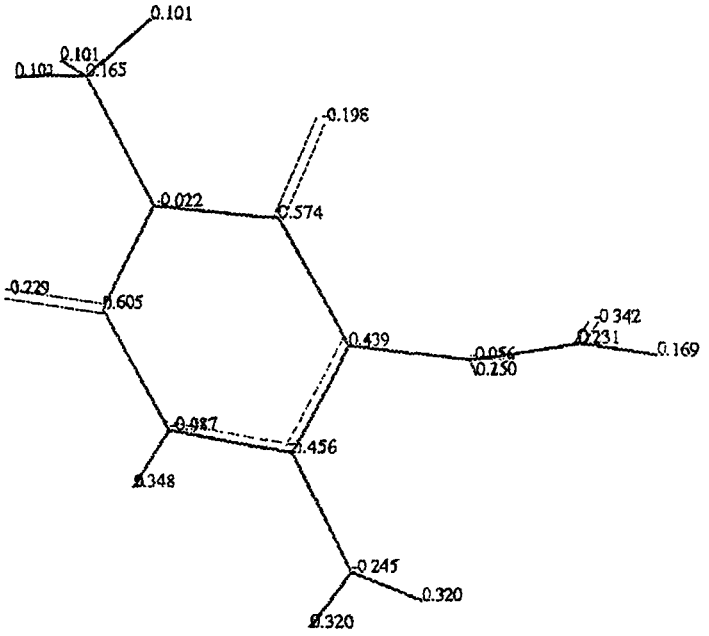


Fig. 3 Schematic of the molecule giving the residual charges on the atoms as calculated by the Kollman method.

the energetics of the cis-trans isomerization of the N-formyl amino group in this compound. We chose the following three torsion angles 1) τ_1 (C6-C5-NC5-CN5), 2) τ_2 (C5-NC5-CN5-05) and τ_3 (C5-C6-N6-H2N6) (see Fig.1 for atom names and numbering). A grid search was performed using SYBYL 6.2β version from the TRIPOS Associates Inc. The above three torsion angles were varied from 0° to 360° in increments of 60° increments. During this search, each of the 216 resulting conformations was energy minimized using the TRIPOS force field and the Kollman charges. Fig. 3 gives the residual charges on the atoms in the molecule calculated using the Kollman method. Table 4 gives the lowest 36 conformations with their energies.

TABLE 4

TORSION ANGLES AND THE ASSOCIATED ENERGIES OF SOME OF THE CONFORMERS

SERIAL NO.	TORSION ANGLES(°)			ENERGY (KCAL/MOL)
	τ_1	τ_2	τ_3	
1	60	120	0	4.55
2	120	120	0	4.57
3	60	-60	0	4.68
4	60	0	0	4.75
5	60	180	0	4.75
6	60	-120	0	5.38
7	0	120	0	7.83
8	-120	180	0	8.42
9	-120	0	0	8.55
10	-60	120	0	9.12
11	180	120	0	9.26
12	-120	120	0	12.54
13	180	180	0	15.71
14	0	0	0	18.25
15	120	-120	0	55.06
16	120	180	0	61.95
17	-120	60	0	79.48
18	120	0	0	82.64
19	-60	180	0	83.31
20	-60	60	0	83.54

Table 4 Continued

21	-60	0	0	85.96
22	120	-60	0	86.08
23	0	60	0	94.07
24	120	60	0	95.70
25	60	60	0	97.91
26	0	-60	0	99.01
27	-120	-60	0	100.34
28	-60	-60	0	101.99
29	0	-120	0	104.19
30	-120	-120	0	104.72
31	-60	-120	0	105.47
32	0	180	0	112.0
33	180	0	0	128.12
34	180	60	0	136.00
35	180	-60	0	149.77
36	180	-120	0	163.49

In the crystal, the three torsion angles τ_1, τ_2 and τ_3 have the values of -83.8, 4.1 and -174.7. The energy of the crystal is 13.48 kcal/mol. This indicates that the molecule takes a conformation which is in between the cis and the trans conformations.

Acknowledgment: Our thanks are due to the referees for their very constructive criticisms which improved this revised version of the paper. Their thanks are also due to one of the referees who

suggested inclusion of a section on the energetics of the system, which substantially improved the quality of this revised version. T.S. and R.P. thank Ms. J. Mann for excellent technical assistance. We would like to thank Dr. Robert Wollman for assistance in the energy calculation and minimization using the TRIPOS system of programs. This work was support by NIH-GM-24864, CA-14185 and the New York State Department of Health.

REFERENCES

1) De, N.C., Mittelman, A., Jenkins, E.E., Crain, P.F., McCloskey, J.A. and Chheda, G.B., J. Carbohydrates, Nucleosides, Nucleotides 7, 113 (1980).

2) De, N.C., Mittelman, A., Dutta, S.P., Edmonds, C.G., Jenkins, E.E., McCloskey, J.A., Blakeley, C.R., Vestal, M.L. and Chheda, G.B., J. Carbohydrates, Nucleosides, Nucleotides 8, 363 (1981).

3) Finks, K., Adams, W. and Pfleiderer, W., J. Biol. Chem. 239, 4250 (1964).

4) Rao, G.S., Khanna, K.L. and Cornish, Experientia 29, 953 (1973).

5) Arnaud, M.J. and Wilsch, C., J. Agric. Food Chem. 27, 524 (1979).

6) Arnaud, M.J., Biochemical Medicine 16, 67 (1976).

7) Chheda, G.B., in Handbook of Biochemistry and Molecular Biology, Vol. 1 (ed. G. Fasman), Chemical Rubber Co., Cleveland, OH, 1975, pp.251.

8) Pople, J.A., Schneider, W.G. and Bernstein, H.J., "High Resolution Nuclear Magnetic Resonances", Chapter 13, McGraw Hill, New York, 1959.

8a) Boiteux, S., Belleney, J., Rogues, B.P. and Laval, J., Nucleic Acid Research 12, 5429 (1984).

9) Srikrishnan, T., Parthasarathy, R., De, Nimai C. and Chheda, G.B., Amer. Cryst. Assn. Meetings, Series 2, Vol. 9, p. 34 (1981).

10) Pfleiderer, W., Liebigs Ann. Chem. 747, 111 (1971).

11) Germain, G., Main,P.and Woolfson, M.M., Acta Cryst. A27, 368-376(1971).

12) Frenz, B.A., "The Enraf Nonius Structure Determination Package: Programs for PDP 11/34", 1975.

13) "International Tables for X-ray Crystallography", Vol. IV, Kynoch Press, Birmingham, AL, p. 72, 1974).

14) Johnson, C.K., ORTEP Report ORNL-3792, Oak Ridge National Laboratory,Tennessee, 1965.

15) SYBYL : Molecular modelling Software (1994) TRIPOS INC., St.Loius., MO.

Received March 22, 1994
Accepted December 16, 1996