CHEMISTRY LETTERS, pp.551-554, 1974. Published by the Chemical Society of Japan

HYDROGEN-BONDED COMPLEX BETWEEN NUCLEOTIDE BASE AND AMINO ACID, CRYSTAL STRUCTURE OF 5-BROMOCYTOSINE:N-TOSYL-L-GLUTAMIC ACID

Minoru OHKI, Akio TAKENAKA, Hirotaka SHIMANOUCHI, and Yoshio SASADA
Laboratory of Chemistry for Natural Products, Tokyo Institute of Technology, Ookayama, Meguro-ku, Tokyo 152

A new interaction mode between cytosine and the carboxyl side group of amino acid has been found in the crystal of 5-bromocytosine: N-tosyl-L-glutamic acid complex by X-ray analysis.

A survey of complex formation of nucleotide base with amino acid and the X-ray crystallographic studies of the complexes have been made to reveal a fundamental stereochemical basis for the interaction between nucleic acid and protein. In some glycine:cytosine complexes previously reported, the carboxylate group in the amino acid is hydrogen-bonded with the protonated N(3) atom and the amino group in the base.[1,2] We have succeeded in obtaining single crystals of the title complex. It is interesting to examine which of α- or γ-carboxyl group of glutamic acid is involved in binding to cytosine.

Crystals of the complex were obtained from an aqueous solution containing equimolar 5-bromocytosine and N-tosyl-L-glutamic acid. Crystal data are; $C_4H_4N_3OBr \cdot C_{12}H_{15}NO_6S$, space group P1, two formula units per unit cell with dimensions of $a=10.843(1)$, $b=11.690(2)$, $c=10.879(2)$ Å, $\alpha=122.15(1)$, $\beta=121.38(1)$, and $\gamma=69.93(1)^\circ$; calculated and measured densities are 1.644 and 1.66 $g \cdot cm^{-3}$, respectively. Intensities of the reflexions recorded on equi-inclination Weissenberg photographs (using Cu$K\alpha$) were measured by the TV densitometer.[3] Corrections were made for Lorentz and polarization factors and for spot shape, but not for absorption. A total number of the independent reflexions was 4259, of which zero-reflexions numbered 803. The approximate structure was obtained by the heavy-atom method. The block-diagonal least-squares refinement with anisotropic thermal parameters decreased an R factor to 0.068.

The atomic coordinates are given in Table, where two crystallographically independent molecules are designated as A and B. Bond lengths and angles in the two independent molecules are in good agreement; the averaged values are given in Fig. 1. They are normal in view of their standard deviations. The molecules are arranged in layers parallel to the $(1\bar{2}0)$ plane, as shown in Fig. 2. The complexes are linked by hydrogen bonds in the layer and come in van der Waals contact with those in adjacent layers. The hydrogen bonding scheme in the layer is shown in Fig. 3. The 5-bromocytosine molecules form a ribbon, elongated along [001], by the hydrogen bonds, N(1A)H···O(2B) 2.91, N(1B)H···O(2A) 2.93, N(4A'')H···N(3B) 3.13, and N(4B)H··· N(3A'') 3.12 Å. This hydrogen-bonded arrangement of the cytosine molecules is different from those found in cytosine,[4] cytosine monohydrate,[5] and cytosine 5-

Table. Fractional atomic coordinates in the 5-bromocytosine:N-tosyl-L-glutamic acid complex.

Atom	x	y	z	Atom	x	y	z
Br5A	-0.4426	-0.1713	0.0272	Br5B	0.4443	0.1664	-0.0286
N(1)	-0.1349	-0.0535	0.0177	N-1B	0.1290	0.0681	-0.0127
C(2)	-0.0262	-0.0194	0.1687	C-2B	0.0219	0.0258	-0.1683
O(2)	0.0894	0.0109	0.1952	O-2B	-0.0907	-0.0069	-0.1931
N(3)	-0.0426	-0.0199	0.2829	N-3B	0.0423	0.0205	-0.2820
C(4)	-0.1628	-0.0583	0.2476	C-4B	0.1653	0.0572	-0.2458
N(4)	-0.1820	-0.0532	0.3613	N-4B	0.1832	0.0477	-0.3635
C(5)	-0.2745	-0.1063	0.0836	C-5B	0.2750	0.1069	-0.0858
C(6)	-0.2566	-0.0988	-0.0248	C-6B	0.2542	0.1089	0.0280
S--A	-0.7731	-0.3463	-0.1534	S--B	0.7321	0.4082	0.1623
O-3A	-0.8038	-0.2370	-0.1929	O-3B	0.6375	0.5041	0.2364
O-4A	-0.6476	-0.4398	-0.1599	O-4B	0.7527	0.2730	0.1335
N-5A	-0.7523	-0.2709	0.0306	N-5B	0.6643	0.4130	-0.0083
O-5A	-0.7982	-0.5315	0.0737	O-5B	0.8175	0.5212	-0.0594
O-6A	-0.9536	-0.3520	0.0532	O-6B	0.9717	0.3387	-0.0695
O-7A	-0.4039	-0.1407	0.6122	O-7B	0.4022	0.1531	-0.6054
O-8A	-0.2956	-0.0934	0.5210	O-8B	0.2910	0.1111	-0.5113
C-7A	-0.8315	-0.4126	0.0786	C-7B	0.8517	0.3986	-0.0805
C-8A	-0.7064	-0.3515	0.1175	C-8B	0.7323	0.3284	-0.1185
C-9A	-0.6289	-0.2644	0.2981	C-9B	0.6190	0.2873	-0.2897
C10A	-0.5091	-0.1991	0.3321	C10B	0.5052	0.2149	-0.3219
C11A	-0.3993	-0.1420	0.5035	C11B	0.3959	0.1587	-0.4930
C12A	-0.9255	-0.4370	-0.2715	C12B	0.9053	0.4726	0.2773
C13A	-0.9060	-0.5783	-0.3183	C13B	0.9157	0.6050	0.3339
C14A	-1.0298	-0.6455	-0.4053	C14B	1.0511	0.6533	0.4188
C15A	-1.1653	-0.5812	-0.4455	C15B	1.1770	0.5655	0.4381
C16A	-1.1825	-0.4423	-0.3939	C16B	1.1612	0.4307	0.3804
C17A	-1.0612	-0.3729	-0.3115	C17B	1.0256	0.3812	0.2957
C18A	-1.2985	-0.6600	-0.5482	C18B	1.3214	0.6169	0.5273

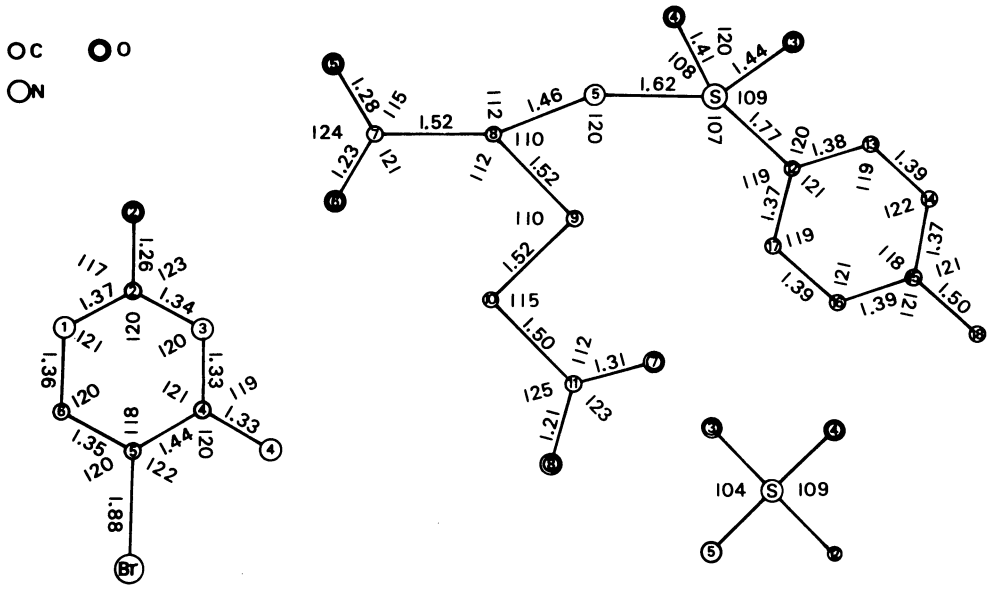

Fig. 1. The average bond lengths and angles of two independent 5-bromocytosine and two N-tosyl-L-glutamic acid molecules.

acetic acid.[6] On the other hand, the two crystallographically independent N-tosyl-L-glutamic acid molecules are associated to each other by the two hydrogen bonds between the α-carboxyl groups, O(5A')H···O(6B) 2.64 and O(5B)H···O(6A') 2.70 Å. These dimers are stacked, by overlapping the tosyl groups, in the same direction as the ribbon.

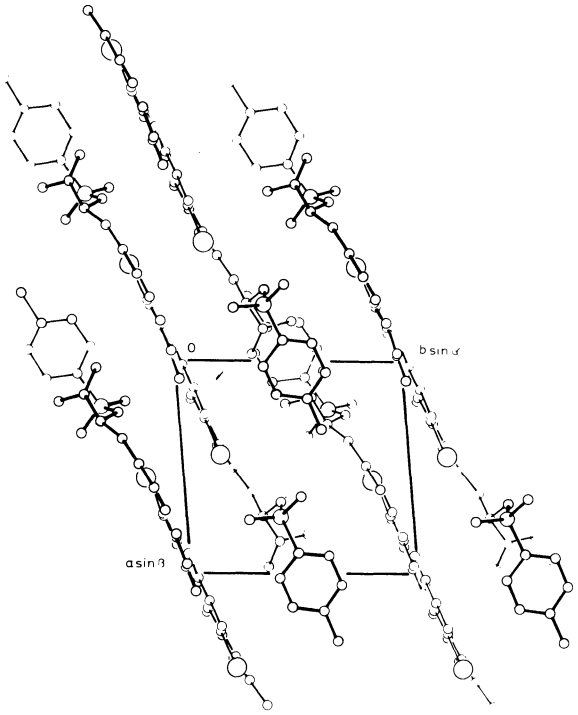

Fig. 2. A view of the structure down the c axis.

The cytosine ribbons and glutamic acid stacks are linked by the other hydrogen bonds in which the side chain of glutamic acid is participated. The O(8) atom is the donor to O(2) of 5-bromocytosine (O(8A'')H···O(2B) 2.58 and O(8B)H···O(2A'') 2.62 Å) and at the same time the acceptor from N(4) of the neighbouring base (N(4A'')H··· O(8A'') 2.87 and N(4B)H···O(8B) 2.89 Å). In addition, the O(7) atom seems to make a CH···O type interaction with C(6) of the third base (C(6A)H···O(7A'') 3.21 and C(6B)H···O(7B') 3.23 Å). The similar C(6)H···O contact is also found in the crystal structures of the two forms of cytidylic acid,[7,8] 1-methylcytosine hydrochloride,[9] and 1-(β-D-arabinofuranosyl)cytosine hydrochloride.[10] The protonation to N(3) of cytosine does not occur, probably due to the inductive effect of electronegative halogen substituent.

These hydrogen bonding scheme is quite different from that found in cytosine:N-benzoylglycine complex monohydrate[1] and in cytosine:N-formylglycine complex.[2] Thus, the present crystal provides the first example of interaction mode between cytosine and the acidic side group of amino acid. Cytosine, making the Watson-Crick type base-pair with guanine, could form another hydrogen bond through a remaining hydrogen atom in its amino group. In the present complex, this hydrogen atom does participate in hydrogen bonding with the O(8) atom. The present finding might

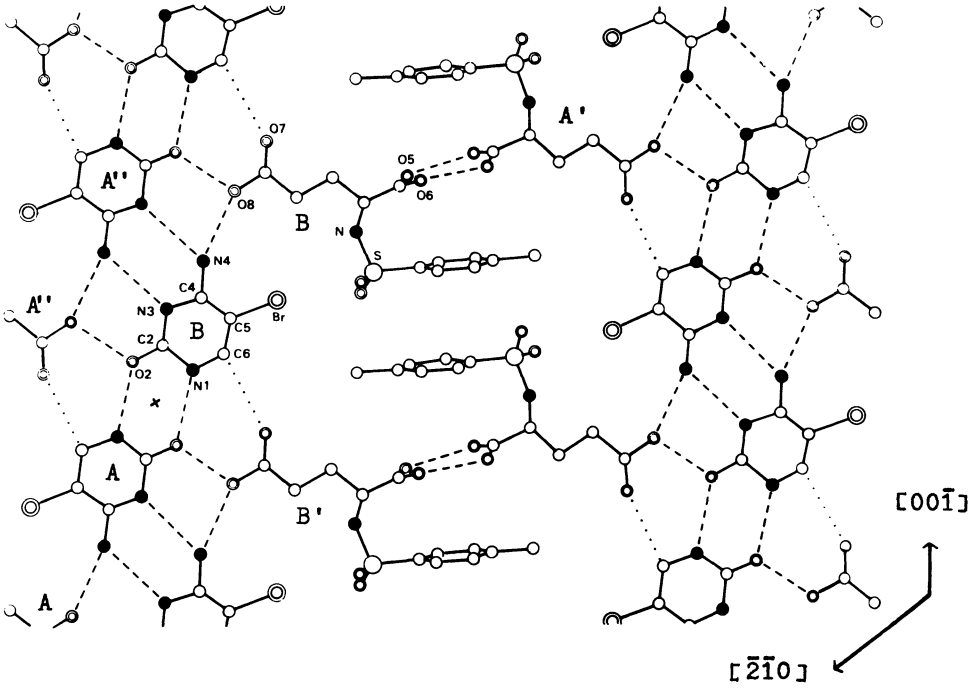

Fig. 3. Hydrogen-bonded network in the 5-bromocytosine:N-tosyl-L-glutamic acid complex. Crystallographically independent cytosine A and B, as well as glutamic acid A' and B, are related to each other by pseudo-inversions. Primed molecules are those generated by lattice translation.

provide a structural basis for a possible interaction mode between the base-paired cytosine in nucleic acid and the acidic side group of amino acid residues in protein.

REFERENCES

1). C. Tamura, T. Hata, S. Sato, & N. Sakurai, Bull. Chem. Soc. Japan, 45, 3254 (1971).
2). M. Ohki, A. Takenaka, H. Shimanouchi, & Y. Sasada, to be published.
3). T. Izumi, Japan. J. Appl. Phys., 10, 1724 (1971).
4). D. L. Barker & R. E. Marsh, Acta Crystallogr., 17, 1581 (1964).
5). G. A. Jeffrey & Y. Kinoshita, Acta Crystallogr., 16, 20 (1963).
6). R. E. Marsh, R. Bierstedt, & E. L. Eichhorn, Acta Crystallogr., 15, 310 (1962).
7). M. Sundaralingam & L. H. Jensen, J. Mol. Biol., 13, 914 (1965).
8). C. E. Bugg & R. E. Marsh, J. Mol. Biol., 25, 67 (1967).
9). B. L. Trus & R. E. Marsh, Acta Crystallogr., B28, 1834 (1972).
10). J. S. Sherfinski & R. E. Marsh, Acta Crystallogr., B29, 192 (1973).

(Received March 26, 1974)