Structure of the Decarboxylated Derivative of the Chromophore from the Fluorescent Peptide produced by Iron-deficient *Azotobacter vinelandii*

By **Kyoyu Sasaki*** and **Yoshimasa Hirata**, Chemical Institute, Faculty of Science, Nagoya University, Chikusa, Nagoya, Japan

Iron-deficient cultures of *Azotobacter vinelandii* contain a yellow-green fluorescent peptide which, upon hydrolysis, yield several amino-acids and a chromophoric residue which retains the spectral properties of the parent peptide. The decarboxylated derivative of the chromophore (I), which retains the spectral behaviour, has been determined by *X*-ray crystallographic study. Crystals are monoclinic, space group $P2_1/c$, with four molecules of $C_{13}H_{12}N_3O_3Cl.2H_2O$ in the unit cell of dimensions $a = 10{\cdot}342(1)$, $b = 6{\cdot}986(1)$, $c = 21{\cdot}525(3)$ Å, $\beta = 115{\cdot}78(2)°$. The structure was solved by direct methods from diffractometer data and refined by block-diagonal least-squares techniques to R 8·0% for 1331 independent reflections. The molecules are piled one above the other between hydrogen-bonded spirals of chloride ions and water molecules.

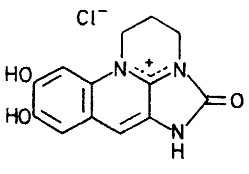
The presence of yellow-green pigments in cultures of *Azotobacter* was first reported by Beijerink[1] in his description of *Azotobacter agile*. The pigment was isolated from iron-deficient cultures of *Azotobacter vinelandii* and the principal component was shown to be a peptide with an attached yellow-green fluorescent chromophore.[2] Hydrolysis of the peptide yields eight amino-acids and the chromophoric substance which retains the spectral properties of the parent peptide. The chromophore has the formula $C_{14}H_{11}N_3O_5$, which upon pyrolysis yields the decarboxylated derivative $C_{13}H_{12}N_3O_3Cl$.

---

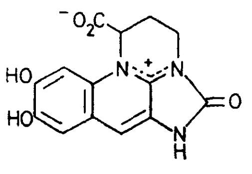
[1] M. W. Beijerink, *Zentr. Bakt.*, 1907, *II*, **2**, 561.

[2] W. A. Bulen and J. R. LeComte, *Biochem. Biophys. Res. Comm.*, 1962, **9**, 523.

Details of the isolation, properties, sequence of the peptide, and decarboxylation of the chromophore, as well as data on degradation products will be reported separately.[3] Recently, the crystal structure of the trimethylated derivative of the chromophore was established.[4] Crystal-structure analysis by X-ray diffraction has enabled elucidation of the structure of the decarboxylated derivative as (I), and the structure of the chromophore has been established as (II) by comparison of chemical and physical data.[3]



(I)



(II)

EXPERIMENTAL

The decarboxylated derivative (I) was crystallized from diluted hydrochloric acid solution as light yellow needles elongated along the b axis.

Crystal Data.—$C_{13}H_{12}ClN_3O_3,2H_2O$, $M = 329{\cdot}7$. Monoclinic, $a = 10{\cdot}342(1)$, $b = 6{\cdot}986(1)$, $c = 21{\cdot}525(3)$ Å, $\beta = 115{\cdot}78(2)°$. $U = 1400{\cdot}3$ Å^3, $D_m = 1{\cdot}56$ (by flotation in n-hexane–carbon tetrachloride), $Z = 4$, $D_c = 1{\cdot}56$, $F(000) = 688$. Space group $P2_1/c$ from systematic absences: $h0l$ for l odd, $0k0$ for k odd. Cu-K_α radiation, $\lambda = 1{\cdot}5418$ Å; μ(Cu-K_α) = 17·8 cm^{-1}.

Lattice constants and intensities were measured at 5 °C, by use of a Hilger and Watts automated four-circle diffractometer Y 290 controlled by a PDP 8 computer. Integrated intensities were measured for $\theta \leq 70°$ by the θ—2θ scan technique by use of nickel-filtered Cu-K_α radiation and a pulse-height analyser. Reflections for $\theta \leq 50°$ were integrated in 40 steps of intervals of 0·02° and those for $\theta > 50°$ were integrated in 50 steps of the same interval. In this way 1694 independent reflections above background were recorded, of which 1381 had $I > 3\sigma(I)$ and were considered observed.[5] All intensities were corrected for Lorentz and polarization factors, and structure amplitudes were derived. No absorption corrections were applied since the specimen was considered to be sufficiently small (0·15 × 0·15 × 0·3 mm), and normalized structure factors $|E|$ were derived.

Structure Analysis.—The structure was solved by the Σ_2 relationship.[6] The origin was specified by assigning positive signs to $\bar{9},1,4$ $3,2,12$, and $\bar{7},3,1$, and a further three reflections, $\bar{3},4,12$, $2,0,14$, and $\bar{2},6,1$, were selected for starting sets. The most consistent set was taken as having the greatest value of an index C. Where C is defined as: $Z_h = \sum_k E_k E_{h-k}$ and $C = \sum_h |Z_h|$. After 7 cycles, the C values were printed for each set. A phase was not accepted if the probability for each sign indication was $< 0{\cdot}97$, and the calculations were only applied to 212 reflections with $|E| > 1{\cdot}6$. Table 1 shows a C value and a number of accepted phases for each starting set.

The Σ_2 relationship was computed by use of 444 reflections ($|E| > 1{\cdot}2$) with 206 phases in the most consistent set ($C = 143 \times 10^2$). An E map was calculated with the accepted 440 phases; 22 large peaks appeared in the map and were assigned as nonhydrogen atoms of (I), chloride ion, and two water molecules. The initial R value, based on the co-ordinates of the atoms as obtained from the E map and an overall isotropic thermal factor of 3·1 Å^2, was 32%. Refinements proceeded to R 11·1% with individual isotropic motion and 8% with 3 cycles of anisotropic refinement. The final co-ordinates and thermal parameters in the salt and the hydrated water molecules are given in Table 2. The final structure amplitudes are listed in Supplementary Publication No. SUP 20619 (3 pp., 1 microfiche).* After the final cycle of refinement, the shifts of all parameters were $< 0{\cdot}2\sigma$. Co-ordinates, thermal parameters, and the scale factor were refined by the block-diagonal approximation of the least-squares method, the expression minimized being $\Sigma w(|F_o| - |F_c|)^2$. The matrices used were 9 × 9 for co-ordinates and thermal parameters (4 × 4 for isotropic thermal motion) of each atom. In the refinement a constant weight was applied to all observed structure factors, and the scattering factors were taken from ref. 7.

TABLE 1

Comparison of starting sets ($|E| > 1{\cdot}5$)

Signs for reflections	1	2	3	4	5	6	7	8
$\bar{3}$ 4 12	+	+	+	+	−	−	−	−
2 0 14	+	+	−	−	+	+	−	−
$\bar{2}$ 6 1	+	−	+	−	+	−	+	−
C value ($\times 10^2$)	88	136	89	126	94	126	91	143
No. of accepted signs	181	200	185	199	196	192	188	206

RESULTS AND DISCUSSION

The present X-ray analysis has established the molecular structure of the decarboxylated derivative of the chromophore in the fluorescent peptide isolated from iron-deficient cultures of *Azotobacter vinelandii* to be as depicted in Figure 1. The unusual heterocyclic structure of the decarboxylated derivative is not found in naturally occurring systems. The molecule exists as a cation, where the charge is divided between the two nitrogen atoms [N(2) and N(13)] with the two N–C distances differing by 0·007 Å (Figure 1). This delocalization of positive charge affects the two N–C and the C=O bonds,

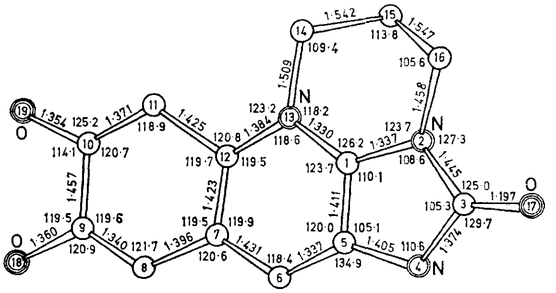


FIGURE 1 Bond lengths (Å) and angles (deg.). Standard deviations, based on the least-squares fit alone, are 0·012 Å for the bond lengths and 0·7° for the angles

---

* See note about Supplementary Publication in Notice to Authors No. 7, *J. Chem. Soc.* (*A*), 1970, Issue No. 20 (items less than 10 pp. are supplied as full size copies).

[3] M. Goto, personal communication, 1972.

[4] I. L. Karle and J. Karle, *Acta Cryst.*, 1971, *B*, **27**, 1891.

[5] K. Sasaki and Y. Hirata, *J. Chem. Soc.* (*B*), 1971, 1565.

[6] J. Karle and I. L. Karle, *Acta Cryst.*, 1966, **21**, 849.

[7] 'International Tables for X-Ray Crystallography,' vol. 3, Kynoch Press, Birmingham, 1968.

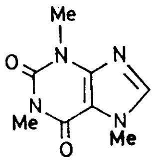
TABLE 2

Final fractional atomic co-ordinates and anisotropic thermal parameters (all × 10^4),* with estimated standard deviations in parentheses

Atom	x	y	z	β_{11}	β_{22}	β_{33}	β_{12}	β_{23}	β_{31}	$B/\text{Å}^2$ †
C(1)	7701(7)	3316(11)	4303(4)	46(8)	168(21)	22(2)	31(22)	12(12)	36(7)	2·4
N(2)	6353(6)	3789(10)	3873(3)	73(8)	215(20)	19(2)	43(21)	15(10)	41(7)	2·8
C(3)	5596(8)	4304(12)	4276(4)	75(10)	203(25)	30(3)	25(26)	13(14)	54(9)	3·1
N(4)	6544(6)	3943(10)	4949(3)	50(7)	200(20)	19(2)	45(20)	29(10)	26(6)	2·6
C(5)	7886(7)	3385(11)	4993(4)	73(9)	145(21)	21(2)	20(23)	−1(11)	49(8)	2·5
C(6)	9159(7)	2952(11)	5508(4)	57(9)	164(24)	22(2)	−19(22)	−12(12)	36(7)	2·6
C(7)	10313(7)	2447(11)	5342(3)	72(9)	130(19)	17(2)	−25(23)	−1(11)	37(7)	2·3
C(8)	11672(7)	2024(11)	5862(4)	64(9)	152(24)	22(2)	28(22)	4(11)	34(8)	2·8
C(9)	12776(7)	1586(12)	5720(4)	63(9)	189(22)	17(2)	7(24)	−7(12)	28(7)	2·6
C(10)	12555(7)	1511(12)	5004(4)	67(9)	143(21)	25(2)	69(23)	24(12)	46(8)	2·6
C(11)	11236(7)	1906(11)	4477(4)	51(8)	139(21)	22(2)	6(21)	−13(11)	38(7)	2·5
C(12)	10085(7)	2313(12)	4641(4)	62(9)	146(20)	19(2)	2(23)	1(12)	28(7)	2·5
N(13)	8737(5)	2831(9)	4125(3)	63(7)	131(18)	15(2)	13(27)	5(12)	24(8)	2·2
C(14)	8392(8)	2817(13)	3368(4)	85(10)	251(28)	15(2)	9(27)	2(12)	32(8)	3·2
C(15)	6763(8)	2540(14)	2938(4)	75(10)	264(26)	19(2)	29(23)	−6(18)	27(8)	3·3
C(16)	5835(8)	3947(14)	3127(4)	88(10)	302(30)	18(2)	123(29)	26(14)	26(8)	3·6
O(17)	4381(5)	4860(10)	4043(3)	63(7)	318(20)	29(2)	78(20)	0(10)	44(6)	3·9
O(18)	14086(5)	1144(9)	6233(3)	71(7)	265(18)	18(2)	56(18)	19(9)	21(5)	3·3
O(19)	13742(5)	1022(9)	4925(3)	67(6)	266(18)	22(2)	46(18)	−2(9)	46(6)	3·1
Cl(20)	1717(2)	2630(4)	2527(1)	145(3)	282(7)	22(1)	−20(9)	−7(4)	39(2)	4·5
O(21)	3283(5)	−198(10)	3704(3)	94(8)	325(20)	21(2)	75(21)	−23(10)	41(6)	3·8
O(22)	8409(6)	1717(11)	1820(3)	143(9)	349(22)	30(2)	−43(24)	9(11)	73(7)	4·7

* In the form: $T = \exp[-(\beta_{11}c^2 + \beta_{22}k^2 + \beta_{33}l^2 + \beta_{12}hk + \beta_{23}kl + \beta_{31}lh)]$. † Thermal factors for the isotropic refinement.

since N(2)–C(3) is more than 0·071 Å longer than N(4)–C(3). The corresponding bond lengths of the trimethylated derivative (III)[4] are N(2)–C(3) 1·380 and N(4)–C(3) 1·375 Å. In a caffeine[8] the two N–C bond lengths adjacent to the C=O group are 1·35 and 1·42 Å, and the C=O bond length is then 1·19 Å. We consider that an N–C bond adjacent to a C=O group is longer than usual when the C=O bond length is shorter than usual. The C=O bond length of the decarboxylated derivative is 1·197 Å, and the corresponding distance in the trimethylated derivative (III) is 1·221 Å. The mean values of the various bonds [C–C(single) bonds 1·545, C–C(aromatic) 1·399, and C–O(aromatic) 1·357 Å] are as expected.



(III)

This molecule is characterized by a plane of equation: $0{\cdot}2629x + 0{\cdot}9634y - 0{\cdot}0666z = 3{\cdot}7045$. All atoms except four [C(3), C(15), C(16), and O(17)] are coplanar to within 0·02 Å; the carbonyl group atoms C(3) and O(17) are above this plane by 0·10 and 0·18 Å, and the deviations of atoms C(15) and C(16) are −0·58 and +0·09 Å.

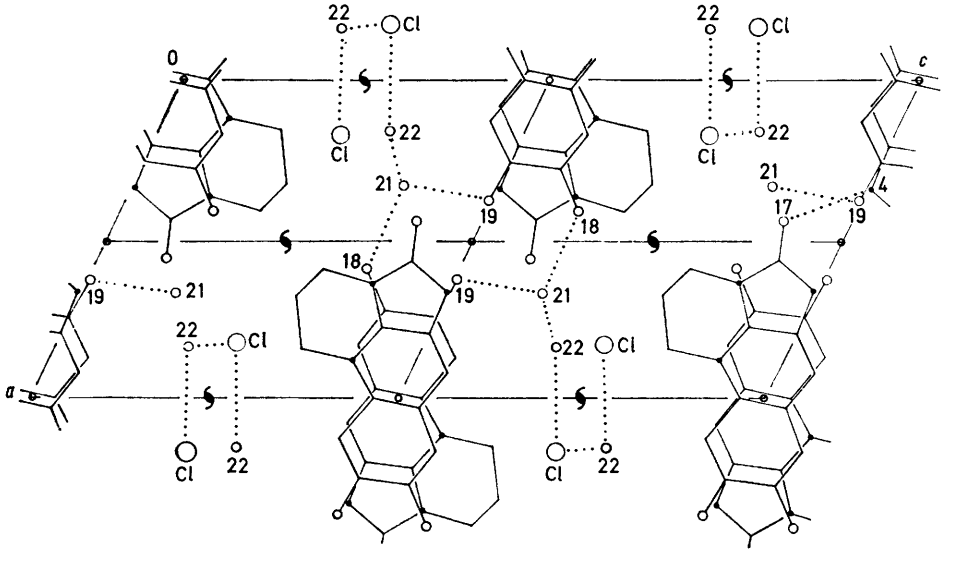


FIGURE 2 A packing diagram viewed along the *b* axis. Some short contacts are shown by dotted lines

[8] D. J. Sutor, *Acta Cryst.*, 1958, **11**, 456.

In the crystal, the presence of two water molecules and a chloride ion per organic molecule influence the packing of the molecules. The chloride ion and one water molecule [O(22)] form spirals by hydrogen bonding around one of the two-fold screw axes (at $x = 0$, $z = \frac{1}{4}$) in each unit cell, and another water molecule [O(21)] is linked to the spiral water [O(22)] and two hydroxy-groups [O(18) and O(19)] of the organic molecule by hydrogen bonds. The lengths of these hydrogen bonds are: O(21) $\cdots$ O(22′), 2·70 O(21) $\cdots$ O(19′) 2·60, and O(21) $\cdots$ O(18′) 2·75 Å. Organic molecules are piled one above the other between spirals of water–chloride-ion hydrogen bonds. Lengths of hydrogen bonds are shown in Table 3, and the arrangement of the molecules in the crystal, as viewed along the *b* axis, is shown in Figure 2. The shortest hydrogen bond distance between organic molecules is 2·85 Å [N(4) $\cdots$ O(17′)].

TABLE 3

Possible hydrogen bonds; distances (Å)

O(21) $\cdots$ O(19^{I})	2·60	O(22) $\cdots$ Cl(20^{II})	3·21
Cl(20) $\cdots$ O(22^{I})	3·15	O(21) $\cdots$ O(18^{III})	2·75
O(21) $\cdots$ O(22^{II})	2·70	N(4) $\cdots$ O(17^{IV})	2·85
O(22) $\cdots$ O(17^{II})	2·97	O(22) $\cdots$ C(6^{V})	3·25

Roman numerals as superscripts denote the following equivalent positions relative to the reference molecule at x, y, z:

I $x - 1, y, z$

II $-x + 1, y - \frac{1}{2}, -z + \frac{1}{2}$

III $-x + 2, -y, -z + 1$

IV $-x + 1, -y + 1, -z + 1$

V $x, -y + \frac{1}{2}, z - \frac{1}{2}$

We thank Dr. M. Goto for providing the crystals. All computations were performed on a FACOM 230–60 at Nagoya University Computation Centre using our programs.

[2/1270 *Received, 5th June*, 1972]