

- COLLINS, D. M. & HOARD, J. L. (1970). *J. Amer. Chem. Soc.* **92**, 3763.
- CROMER, D. T. & MANN, J. L. (1968). *Acta Cryst.* **24**, 321.
- FRIEDEL, E. (1925). *C. R. Acad. Sci., Paris*, **180**, 269.
- GALIGNÉ, J. L. & FALGUEIRETTES, J. (1968). *Acta Cryst.* **B24**, 1523.
- HOWELLS, E. H., PHILLIPS, D. C. & ROGERS, D. (1950). *Acta Cryst.* **3**, 210.
- HÜCKEL, E. (1921). *Phys. Z.* **22**, 561.
- International Tables for X-ray Crystallography* (1962). Vol. III. Birmingham: Kynoch Press.
- KAST, W. (1927). *Ann. Physik.* **83**, 418.
- KRIGBAUM, W. R., CHATANI, Y. & BARBER, P. G. (1970). *Acta Cryst.* **B26**, 97.
- LIPSON, H. & COCHRAN, W. (1966). *The Determination of Crystal Structures*, pp. 62–64. Ithaca: Cornell Univ. Press.
- MCPHAIL, A. T. & SIM, G. A. (1969). Private communication.
- NORTH, A. C. T., PHILLIPS, D. C. & MATTHEWS, F. S. (1968). *Acta Cryst.* **A24**, 351.
- SMYTH, C. P. (1955). *Dielectric Behavior and Structure*, p. 244. New York: McGraw-Hill.

*Acta Cryst.* (1971). **B27**, 1891

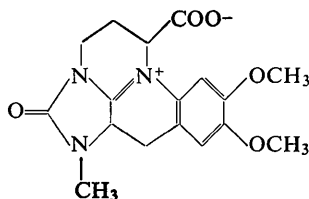
## Structure of the Chromophore from the Fluorescent Peptide produced by Iron-deficient *Azotobacter Vinelandii*

BY ISABELLA L. KARLE AND J. KARLE

Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, D.C. 20390, U.S.A.

(Received 9 October 1970)

Iron-deficient cultures of *Azotobacter vinelandii* contain a yellow-green fluorescent peptide which, upon hydrolysis, yields several amino acids and a chromophoric moiety that retains the spectral properties of the parent peptide. Crystal-structure analysis by X-ray diffraction has elucidated the molecular formula of the trimethylated derivative of the chromophore to be the zwitterion:



There appears to be an intramolecular attraction between  $N^+$  and  $O^-$  since the plane containing the carboxyl ion is perpendicular to the plane containing the three conjugated rings, thus making the  $N^+ \cdots O^-$  separation a minimum at 2.64 Å. The compound crystallizes with four molecules of water in the face-centered triclinic space group  $F\bar{1}$  with cell parameters:  $a = 9.12 \pm 0.02$ ,  $b = 19.94 \pm 0.04$ ,  $c = 21.17 \pm 0.04$  Å,  $\alpha = 93.0 \pm 0.3$ ,  $\beta = 92.8 \pm 0.3$ , and  $\gamma = 98.0 \pm 0.3^\circ$ . Visually estimated intensities were obtained from a crystal enclosed in a glass capillary. The final  $R$  index for anisotropic refinement was 9.0%. The structure is layered with sheets of hydrogen-bonded water molecules between layers of organic ions. Phases were determined directly from the structure factor magnitudes by the symbolic addition procedure.

### Introduction

The presence of yellow-green pigments in cultures of *Azotobacter* was first reported by Beijerinck (1907) 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 (Bulen & LeComte, 1962). Hydrolysis of the peptide yields the rare amino acids homoserine and  $\beta$ -hydroxyaspartic acid as well as 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 prolonged treatment with diazomethane yields the trimethylated derivative  $C_{17}H_{17}N_3O_5$  (Corbin, Karle & Karle, 1970). The purpose of this investigation was to establish the molecular formula of the chromophore. Crystals of the trimethylated derivative were provided by Dr J. L. Corbin of the Charles F. Kettering Research Laboratory. Details of the isolation, properties, and methylation of the chromophore, as well as data on degradation products will be published elsewhere (Corbin, 1970).

During the structure determination, it became apparent that a partial structure was obtained which

was misplaced with respect to the centers of symmetry. A procedure employing the tangent formula (Karle, 1968) is described wherein the partial structure is used in a noncentrosymmetric triclinic cell to develop the complete structure. The proper placement of the molecules with respect to the center of symmetry is then apparent.

### Experimental

Although the trimethylated chromophore crystallizes as transparent, well-formed prisms which are stable in a humid atmosphere, the crystals readily turn to powder when the humidity decreases. Accordingly, the diffraction data were collected from a wet crystal sealed in a capillary. The intensity data were recorded photographically by the multiple-film, equi-inclination Weissenberg technique and were estimated visually by comparison with a calibrated film strip. Layers 0-6 were recorded on the *a* axis. No cross layers were photographed; however, the individual layers were recorded under controlled conditions of exposure and photographic development. Corrections were made for spot size, Lorentz and polarization factors, and normalized structure factors  $|E|$  were derived. The average values of  $|E|^2$  and  $||E|^2 - 1|$  indicated that the triclinic cell was centrosymmetric. Cell parameters, derived from precession photographs, and other physical data are listed in Table 1.

Table 1. *Physical data*

M.F.	$C_{17}H_{17}N_3O_5 \cdot 4H_2O$	
M.W.	415.41 (calc. from formula)	
Color	Very light yellow	
Habit	Prismatic, elongated on <i>a</i>	
Size	$\sim 0.8 \times 0.3 \times 0.3$ mm	
Space group	$F\bar{1}$	$P\bar{1}$
<i>a</i>	$9.12 \pm 0.02$ Å	$9.12$ Å
<i>b</i>	$19.94 \pm 0.04$	11.52
<i>c</i>	$21.17 \pm 0.04$	11.73
$\alpha$	$93.0 \pm 0.3^\circ$	$79.8^\circ$
$\beta$	$92.8 \pm 0.3$	115.6
$\gamma$	$98.0 \pm 0.3$	121.1
<i>V</i>	3802.2	950.5
<i>Z</i>	8	2
Density (calc)	1.451 g.cm <sup>-3</sup>	
Radiation	Cu <i>K</i> $\alpha$ 1.5418 Å	
No. independent reflections	2758	
$\langle  E ^2 \rangle$	0.755	
$\langle  E ^2 - 1  \rangle$	1.046	

For ease in data collection and indexing, a face-centered unit cell, space group  $F\bar{1}$ , with eight molecular units was chosen. In this cell the angles  $\alpha$ ,  $\beta$  and  $\gamma$  are between 90 and 100°. All atomic parameters, the Figures, and the subsequent discussion refer to the  $F\bar{1}$  cell. Cell parameters for a primitive cell listed in Table 1 are related to the face-centered cell by the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}.$$

### Structure analysis

Phases for the reflections were derived directly from the structure-factor magnitudes by the symbolic addition procedure (Karle & Karle, 1963, 1966). Although the sign determination for centrosymmetric crystals has become quite routine, the triclinic space group  $P\bar{1}$  (or  $F\bar{1}$ , in this case) often presents some difficulty. The nature of the difficulty stems from the fact that there are fewer  $\Sigma_2$  combinations for each *hkl* than in space groups with more symmetry, and that greater reliance must be placed on indications from single contributors with somewhat lower probabilities of being correct.

In the present investigation, the origin was specified by assigning + signs to 1,17,3, 0,14,16 and 224. These reflections are in the *ugu*, *gug*, and *ggu* categories when transformed to a  $P\bar{1}$  cell by the matrix given previously and are, therefore, proper for specifying the origin. In the triclinic system, all signs derived from the origin specification are positive, and to obtain negative signs at least one of the assigned symbols must be negative. Unknown symbols were assigned to four reflections in the course of the sign determination, 3,1,15 *c*, 4,2,8 *d*, 1,13,9 *g*, and 5,7,15 *h* to implement the  $\Sigma_2$  relationship. It soon became apparent that  $c = +$  and there were three indications that  $d = h$ . With two unknowns left, only three maps needed to be computed since the set containing all + signs was obviously incorrect. Each map contained one very large peak which *a priori* indicated that the phases were in error. Changing the sign of the symbol *h* from *d* to  $-d$  led to several maps having an anthracene-type structure which appeared in a different location in each map (Fig. 1). The structural formula was not known at this point; however the 14 atoms comprising the anthracene type fragment were used as a basis for a partial

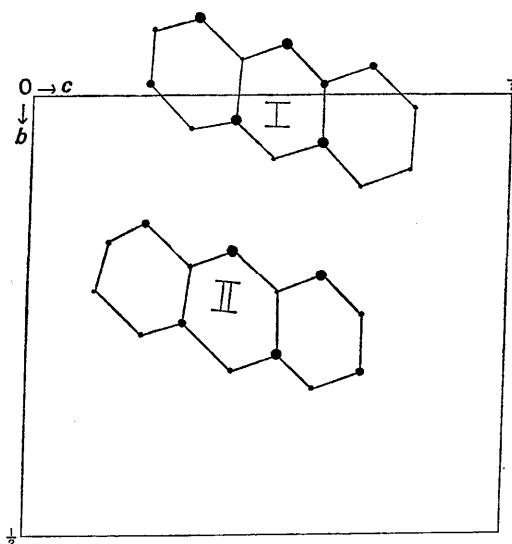


Fig. 1. Positions of maxima showing a probable fragment of the molecule for two different sets of phases.

Table 2. Observed and calculated structure factors

Columns represent the index *l*, |F<sub>0</sub>| and 10F<sub>o</sub>.

Table with 10 columns: Index l, |F0|, 10Fo, and 8 columns of observed and calculated structure factor values. The table contains a dense grid of numerical data points.

Table 2 (cont.)

10 73 86	-18 0 31	14 0 40	+5 0 73	7 425 -344	-1 361 -369	-17 0 -40	15 953 940	5 11 L	-11 180 -202	0 950 510	0 718 686	14 243 -273	-12 216 -203
12 597 935	-16 406 435	10 282 266	+3 244 225	9 0 84	1 362 288	-15 903 982	17 247 223	-13 76 -126	+9 130 -118	10 0 -35	21511-1359	16 0 35	-10 147 -149
14 368 359	-14 282 143	4 15 L	-1 173 254	11 563 -443	-1 407 -468	-1 275 266	19 86 354	+13 484 468	-6 195 -218	12 0 75	6 604 387	18 79 -79	-8 420 -425
16 0 25	-16 239 -255	-16 238 -262	1 109 78	13 961 -552	5 185 87	-11 052 375	21 100 99	-11 297 -285	+5 813 874	14 0 -9	6 827 840	20 251 226	-6 114 -81
18 198 -199	-10 305 303	-16 0 +42	3 0 -11	15 200 -184	7 0 27	-9 595 647	5 9 L	-9 129 74	-3 342 -355	16 291 297	0 100 125	0 4 L	+4 226 205
20 158 -159	-8 959 -967	-14 0 -14	5 -5	17 145 11	7 145 11	-7 285 286	-3 377 366	+5 204 198	-11 008 -963	5 131 316	-6 137 196	-4 490 490	-12 310 318
22 277 -279	-6 150 -151	-12 217 296	7 0 53	10 0 36	11 0 107	-5 431 671	-13 0 27	-15 195 -189	1 314 341	-10 493 -484	12 483 -483	-10 395 -178	0 -237
24 4 17	-2 660 -595	-10 302 -313	9 185 -103	5 -11 L	13 99 -40	-3 275 -273	-11 166 177	-3 420 -473	3 130 118	-8 405 -331	14 0 7	-7 874 -462	2 366 -344
-22 112 -119	-6 192 -89	-8 97 135	11 407 -115	-17 563 -443	-13 514 144	-15 903 982	-11 166 177	-3 420 -473	3 130 118	-8 405 -331	14 0 7	-7 874 -462	2 366 -344
-20 452 407	2 220 -220	-8 171 -168	13 121 120	-15 316 17	0 33 707 692	-7 0 50	11 107 107	-11 166 177	-3 420 -473	3 130 118	-8 405 -331	14 0 7	-7 874 -462
-18 398 -395	4 142 -128	+4 782 741	15 149 -142	-13 215 -198	19 11 -18	3 556 +420	-5 0 0	5 798 876	0 -22 L	+2 607 -510	-12 352 -349	-2 248 -214	8 227 107
-16 445 -443	8 111 -105	-2 581 -598	9 -17 L	-11 162 -212	21 117 78	5 821 +912	-3 377 366	+5 204 198	-11 008 -963	5 131 316	-6 137 196	-4 490 490	-12 310 318
-14 91 -78	10 305 -314	2 385 -427	+7 223 223	-7 74 -132	-19 229 -259	9 0 6	1 064 931	11 348 341	0 191 -172	4 539 -521	-6 296 226	4 222 -145	14 227 143
-12 0 109	-12 710 -759	4 179 169	+5 988 +602	-9 415 -444	-17 323 354	11 118 -62	3 1007 +957	13 234 245	0 200 202	0 789 800	-6 297 -269	6 147 -196	16 12 L
-10 232 -192	14 207 -157	4 95 74	-3 403 463	-3 199 89	-18 97 -46	13 318 343	5 62 124	5 13 L	-6 20 L	8 368 330	-2 288 288	8 794 -771	-12 216 -203
-8 924 -855	10 208 177	8 183 -162	-1 499 -446	-1 0 -12	-13 102 238	15 803 +719	7 285 289	-15 195 -162	2 175 -124	10 456 -434	0 352 679	10 364 340	-10 110 -114
-4 257 140	18 78 44	10 385 380	1 624 823	-15 0 -79	17 0 18	9 349 335	-15 238 248	4 394 336	12 0 14	2 764 -733	12 162 164	-8 92 -194	-8 92 -194
-2 401 311	10 385 380	1 624 823	-15 0 -79	17 0 18	9 349 335	-15 238 248	4 394 336	12 0 14	2 764 -733	12 162 164	-8 92 -194	-8 92 -194	-8 92 -194
0 675 651	4 12 L	14 63 48	5 0 -47	5 0 -67	2 0 -66	21 0 66	13 474 466	-8 095 -277	8 313 274	16 104 -141	6 272 273	16 344 -375	-4 264 -292
2 253 +258	+20 0 -46	4 18 L	7 0 -68	7 215 201	-9 296 -299	23 281 -242	15 37 -32	15 37 -32	10 280 -308	6 0 8 L	0 295 238	18 92 -70	-2 523 -499
4 1240 1243	-18 148 198	-12 339 -334	9 0 61	9 393 386	-3 867 -866	5 1 -42	17 75 -58	5 231 229	6 -18 L	-10 201 178	10 107 -43	20 226 202	0 522 -828
6 980 1000	-16 186 190	-10 173 145	11 141 171	11 0 52	-11 269 -317	+15 254 -970	19 81 -79	-3 0 -48	0 218 -175	-8 637 -618	12 274 267	0 6 L	2 0 75
8 0 22	-14 139 161	-8 71 -81	13 0 -26	13 129 117	1 90 39	-13 203 214	5 7	-11 244 211	2 0 44	6 0 103	14 148 -134	-12 208 197	4 492 -497
10 346 298	-12 310 -315	-6 144 176	15 0 4	15 250 -266	5 225 193	-11 484 -939	-9 178 183	-9 214 332	0 244 -225	10 94 -95	2 546 -656	-15 869 912	-4 0 -55
12 400 512	-10 499 425	-4 407 425	-4 407 425	17 221 -191	17 73 -91	5 0 23	-9 324 274	5 7	-13 305 -332	5 7	-13 305 -332	5 7	-13 305 -332
14 987 496	-8 915 912	-2 398 -377	5 -15 L	17 90 222	-9 420 -420	5 0 23	-9 324 274	5 7	-13 305 -332	5 7	-13 305 -332	5 7	-13 305 -332
16 0 12	+6 869 811	0 0 -41	-15 319 287	21 292 -222	-5 0 0 L	13 115 -427	-3 314 1392	-7 311 -335	11 0 -46	12 0 -47	4 0 94	-11325-1320	-2 226 146
18 292 -241	-4 358 314	2 117 194	1 0 17	-11 248 -254	-17 0 67	13 0 41	-1 0 -54	-5 0 -54	14 244 212	6 802 781	-8 229 161	0 461 394	-12 421 342
20 292 -241	-4 358 314	2 117 194	1 0 17	-11 248 -254	-17 0 67	13 0 41	-1 0 -54	-5 0 -54	14 244 212	6 802 781	-8 229 161	0 461 394	-12 421 342
22 257 243	-4 358 314	2 117 194	1 0 17	-11 248 -254	-17 0 67	13 0 41	-1 0 -54	-5 0 -54	14 244 212	6 802 781	-8 229 161	0 461 394	-12 421 342
24 231 -201	4 299 274	10 440 -437	+5 306 394	+11 144 -195	19 0 -118	54465 1413	1 499 409	-11 0 -24	2 93 -44	12 593 -663	-2 891 -760	6 494 365	-8 0 -47
-22 202 237	-4 358 314	2 117 194	1 0 17	-11 248 -254	-17 0 67	13 0 41	-1 0 -54	-5 0 -54	14 244 212	6 802 781	-8 229 161	0 461 394	-12 421 342
-18 96 79	8 82 -114	4 20 L	1 0 -36	-7 446 489	5 -3	9 436 -340	5 125 140	-7 221 -244	6 130 -79	16 0 -91	+4 578 319	10 0 223	-4 113 139
-16 0 0	10 890 +916	-10 343 -348	1 0 -34	-5 602 -698	-10 233 -286	1131097-1038	7 76 -102	-5 567 605	8 0 77	6 0 6 L	4 327 -332	12 0 24	0 472 521
-14 232 -847	-14 138 150	3 89 -69	17 255 -156	11 279 -221	-3 0 -38	-15 192 162	-13 100 90	5 17 L	-10 320 -263	-12 175 -180	6 312 329	14 251 -847	2 0 -84
-12 688 -727	14 390 421	-4 358 314	5 244 440	-1 0 -24	17 0 18	9 349 335	-15 238 248	4 394 336	12 0 14	2 764 -733	12 162 164	-8 92 -194	-8 92 -194
-10 884 -926	16 114 -118	+4 358 314	7 623 -531	1 200 130	-13 0 89	17 516 465	13 81 +492	-1 689 -845	10 215 -159	-8 0 770	10 348 -367	18 0 46	6 414 -132
-8 668 -641	18 144 -128	-2 285 -227	9 126 132	3 421 402	-1 0 3	19 143 -185	15 0 -80	3 728 -758	16 363 -207	-6 444 369	12 320 246	20 243 -270	8 0 -4
-4 373 -349	6 154 L	11 194 -212	5 181 349	-9 861 613	5 0 22	12 0 131	5 0 22	6 -14 L	+4 289 -205	14 490 485	6 490 485	6 490 485	6 490 485
-4 466 -464	+20 308 -316	5 -23 L	13 76 76	7 341 369	-7 397 334	23 192 243	5 9 L	7 395 -407	-2 406 351	-2 329 -280	14 248 148	-12 244 -828	10 844 506
-2 122 131	-18 51 166	1 198 -117	15 241 -249	9 427 -343	-5 394 -608	5 3 L	-15 0 318	0 0 69	0 967 504	0 640 -668	18 399 354	-10 348 -306	-10 374 351
0 172 149	+14 148 150	3 89 -69	17 255 -156	11 279 -221	-3 0 -38	-15 192 162	-13 100 90	5 17 L	-10 320 -263	-12 175 -180	6 312 329	14 251 -847	2 0 -84
2 0 0	+10 213 213	5 0 -8	10 120 73	13 404 531	-1 73 35	-13 953 -890	-11 0 36	-15 296 315	6 1104 1179	0 288 274	-12319 1405	-4 429 401	-4 0 2
4 69 78	-12 338 -229	7 0 495	5 -13	19 988 633	1 288 -322	-11 324 -295	-9 0 -22	-13 115 -138	6 1104 1179	0 288 274	-12319 1405	-4 429 401	-4 0 2
8 0 94	-8 229 -229	5 -21 L	-15 983 693	10 64 -93	5 395 360	-8 417 -495	-9 437 389	-7 215 -190	10 0 81	10 296 -313	-2 230 203	0 246 236	0 187 -178
10 344 -421	-6 762 623	-5 288 -229	-13 394 -421	21 111 -100	7 689 -709	-5 782 783	-3 181 125	-7 0 -15	12 232 -210	12 321 -315	-6 287 241	2 330 406	2 287 273
12 82 -116	-4 358 314	-3 243 223	-11 0 -9	13 279 -291	15 496 900	-3 184 484	-1 955 938	-5 823 596	4 81 24	4 0 20	-3 361 291	4 180 -106	4 180 117
14 695 741	-2 305 333	-1 107 -108	-0 -13	-17 226 239	11 539 594	-11213 1158	1 481 148	-3 651 -765	16 0 28	16 0 28	16 0 28	16 0 28	16 0 28
16 208 222	0 244 -246	1 166 143	+7 521 616	-15 307 345	131082-1010	13398-1347	3 415 382	-1 146 142	6 -12 L	6 -4 L	0 306 -242	0 610 973	0 610 973
18 156 68	2 247 -222	3 92 46	-5 216 16	-13 279 -291	15 496 900	-3 184 484	-1 955 938	-5 823 596	4 81 24	4 0 20	-3 361 291	4 180 -106	4 180 117
20 126 -113	4 286 -279	5 345 294	-3 0 -36	-11 201 -184	17 0 11	5 254 44	7 942 466	3 0 -1	-2 198 196	-12 376 -366	4 306 432	12 0 27	-8 405 398
22 214 240	6 142 174	7 186 133	-1 379 -359	-9 293 -890	19 182 139	71119 -1152	9 286 -979	5 346 -419	0 445 443	8 0 132	6 498 -493	14 118 -157	-4 927 -500
4 10 L	8 300 477	9 244 -295	-8 389 -449	21 0 71	9 0 27	11 448 -432	7 0 -15	0 732	+6 282 -180	-6 282 -180	10 107 -43	20 226 202	-4 523 335
-22 197 -202	10 473 -706	11 504 312	3 331 -293	-5 990 -1178	5 -1	11822-1833	13 307 -246	9 219 -223	44226-1563	6 715 699	2 0 -193	12 350 -581	6 10 L
-20 197 177	12 154 -141	5 -19 L	3 442 440	-3 192 -199	-19 0 -18	131150 1116	15 208 -168	5 19 L	6 715 699	2 0 -193	12 350 -581	6 10 L	6 10 L

structure (Karle, 1968) in the *noncentrosymmetric* space group *F1*.

For convenience, in order to contain the coordinates of one molecule within the bounds of the unit cell, the *y* coordinates of the atoms in the partial structure I, Fig. 1, were arbitrarily shifted by  $\frac{1}{2}$ . The change in origin from that originally specified in the phase determination had the effect of changing the signs of all reflections which obey the condition:  $(-h+k)/2$  is odd. All further discussion of signs as well as the signs in Table 2 refer to the new origin.

Phases based on the partial structure were accepted if  $|F_c| > 0.25|F_o|$  for  $|E_o| > 1.5$ , and they were used as a starting set in the tangent formula (Karle & Hauptman, 1956). The resulting *E* map showed that two of the atoms assumed for the partial structure had disappeared. The 12 meaningful atoms (peak values of 200–400 on an arbitrary scale) were eventually identified as (1), (5), (6), (7), (8), (9), (10), (11), (12), N(13), O(20), and O(22), (see Fig. 2). Eight new peaks (peak values

170–250), which appeared in positions that could be part of the molecule, were used in a subsequent partial structure-tangent formula cycle; six additional peaks that could have been part of another molecule related by a center of symmetry were not used. The *E* map based on the partial structure of 20 atoms now revealed the 25 atoms of another molecule, related by a center of symmetry that did not correspond to an origin in the *F1* cell (Fig. 3). At this point, two water molecules were also found. Coordinates of the molecules were translated so as to place them in an *F1* cell, and a difference map revealed that each asymmetric unit had a total of four water molecules. Chemical information available to us indicated that the crystal contained at least two, and probably more, molecules of water crystallization for each chromophore molecule.

After refining the coordinates, the original phase determination was re-examined to determine the points of difficulty. The following exceptions to the  $\sum_2$  relationship were noted:

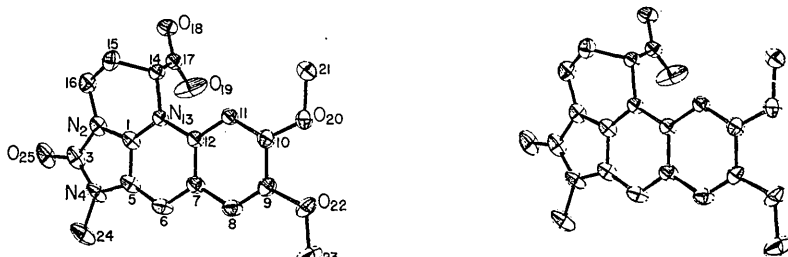


Fig. 2. A stereoscopic view of the molecule with thermal ellipsoids computed at a 50% probability level (Johnson, 1965).

<p>(1) <math>\begin{array}{cccc} 0 &amp; 8 &amp; 0 &amp; - \\ 1 &amp; \overline{19} &amp; 3 &amp; - \\ \hline 1 &amp; \overline{11} &amp; 3 &amp; - \end{array}</math></p> <p>(2) <math>\begin{array}{cccc} 0 &amp; \overline{18} &amp; 16 &amp; - \\ 5 &amp; 15 &amp; \overline{3} &amp; - \\ \hline 5 &amp; \overline{3} &amp; 13 &amp; - \end{array}</math></p> <p>(3) <math>\begin{array}{cccc} \overline{4} &amp; \overline{4} &amp; 8 &amp; + \\ 5 &amp; 15 &amp; \overline{3} &amp; - \\ \hline 1 &amp; 11 &amp; 5 &amp; + \end{array}</math></p>	<p>(4) <math>\begin{array}{ccc} 0 &amp; \overline{18} &amp; 16 &amp; - \\ 1 &amp; 13 &amp; 5 &amp; - \\ \hline 1 &amp; \overline{5} &amp; 21 &amp; - \end{array}</math></p> <p>(5) <math>\begin{array}{ccc} \overline{1} &amp; 19 &amp; \overline{7} &amp; - \\ 1 &amp; \overline{7} &amp; 23 &amp; \pm \\ \hline 0 &amp; 12 &amp; 16 &amp; + \end{array}</math></p> <p>(6) <math>\begin{array}{ccc} \overline{1} &amp; 5 &amp; 9 &amp; + \\ 1 &amp; 7 &amp; 7 &amp; = \\ \hline 0 &amp; 12 &amp; 16 &amp; + \end{array}</math></p>
--	---

Each reflection had an  $|E|$  value greater than 2.0, and the probability for each sign indication was in the range 0.94–0.97. Relationships (2)–(4) gave the three indications, each one incorrect, that symbol  $d=h$ . These were not the major sources of difficulty, however. Relationships (5) and (6) led to an incorrect sign for 0,12,16, which in turn caused a large subset of signs to be reversed. In spite of faulty sign indications, a sufficient fragment of the molecule appeared in an  $E$

map which could be used for a partial structure in a space group with less symmetry.

Identification of the N and O atoms, aside from chemical considerations, was made from the weight of the Fourier peaks. It was independently confirmed by the behavior of the thermal parameters in the least-squares refinement in the case where all the nonhydrogen atoms of the molecule were labeled as carbon atoms. Furthermore, a difference map computed after the first cycle of full-matrix refinement with anisotropic thermal factors was essentially featureless, except for the maxima ( $\sim 0.5$  to  $0.75$  e. $\text{\AA}^{-3}$ ) associated with the hydrogen atoms.

Isotropic refinement with individual scale factors for each layer and without hydrogen atoms led to an  $R$  index of 16.4%. The values of the individual scale factors were similar in magnitude, but they varied from each other in a random fashion. The data were adjusted to one scale by the values for the scale factors for the separate layers; subsequently, only one scale factor was varied in the anisotropic refinement. The weighting function used in the expression  $\sum w(|F_o|^2 - |F_c|^2)$  was:  $w = 0.5$  for  $F_o = 0$ ,  $w = 1.0$  for  $|F_o| < 40$  and  $w = \frac{40}{|F_o|}$  for  $|F_o| \geq 40$ . Sixteen hydrogen atoms were included with constant parameters in the final cycles of the anisotropic refinement. Their approximate coordinates as found in a difference map are listed in Table 3, and the thermal parameters used were the same as those of the atom to which each hydrogen atom is attached. The three hydrogen atoms on the methyl C(24) atom were not found, as well as the hydrogen atoms belonging to  $W(2)$ ,  $W(3)$ , and  $W(4)$ . The final  $R$  index for the observed data is 9.0%. Observed and calculated structure factors are listed in Table 2. Fractional coordinates and thermal parameters are shown in Table 4. Thermal parameter values should be considered as only approximate descriptions of the motions of the atoms, because of the way in which the data had to be collected.

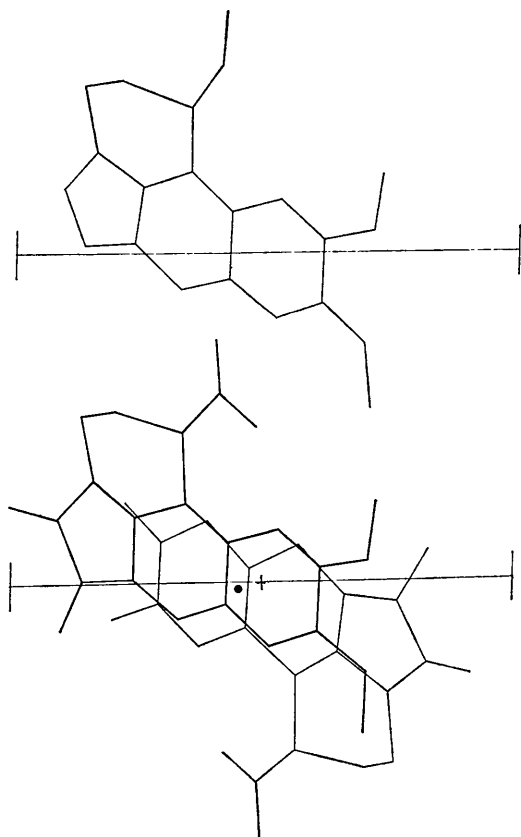


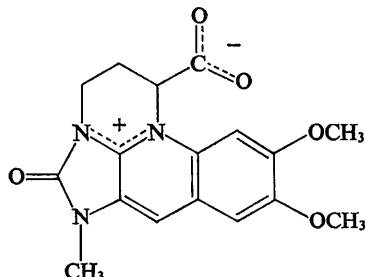
Fig. 3. Progress in obtaining positions of the atoms, assuming that the cell is noncentrosymmetric. Molecules were related by a center of symmetry, indicated by  $\bullet$ , and required translation in order that a center of symmetry in the cell, indicated by  $+$ , coincided with the center of symmetry between the two molecules.

Table 3. Approximate coordinates for hydrogen atoms as found in a difference map

	$x$	$y$	$z$
H(6)	0.338	0.578	0.188
H(8)	0.512	0.610	0.280
H(11)	0.608	0.388	0.283
H(14)	0.533	0.318	0.213
H(15-1)	0.503	0.322	0.100
H(15-2)	0.442	0.253	0.122
H(16-1)	0.292	0.310	0.038
H(16-2)	0.193	0.278	0.102
H(21-1)	0.840	0.388	0.345
H(21-2)	0.695	0.370	0.395
H(21-3)	0.817	0.397	0.428
H(23-1)	0.540	0.657	0.377
H(23-2)	0.733	0.682	0.410
H(23-3)	0.693	0.688	0.338
H $W(1-1)$	0.092	0.295	0.313
H $W(1-2)$	-0.105	0.288	0.307

## Discussion

The structural formula of the trimethylated derivative of the chromophore in the fluorescent peptide isolated from iron-deficient cultures of *Azotobacter vinelandii* has been established by X-ray analysis to be:



The molecule is shown in the stereodiagram in Fig. 2. The unusual heterocyclic structure of the chromophore does not appear to fit any naturally occurring system (Corbin, Karle & Karle, 1970). The molecule exists as a zwitterion, where the negative charge is divided between the two oxygen atoms in the carboxyl group with the two C—O distances differing by 0.02 Å (Fig.

4); the positive charge is divided between the two nitrogen atoms, N(2) and N(13), with N(13) probably possessing most of the charge since the C(1)—N(13) bond length is more than 0.04 Å shorter than the C(1)—N(2) bond length. All four C—N bonds in the five-membered ring are adjacent to double bonds. They assume a definite double-bond character, evidenced by their foreshortened bond lengths, 1.34 to 1.38 Å, as compared to 1.47 to 1.50 Å for the single bonds N(4)—C(24), N(2)—C(16), and N(13)—C(14). The molecule is characterized by two planes whose equations are:

$$7.07485x + 2.67101y - 13.3834z = 1.53679 \quad (A)$$

$$5.61748x - 2.52317y + 16.0298z = 4.70301 \quad (B)$$

All atoms (excluding H atoms), except the carboxyl ion and C(15), lie essentially in plane A, the greatest deviations being  $-0.068$  Å for C(23) in one of the methoxy groups and  $+0.127$  Å and  $-0.120$  Å for C(16) and C(14), respectively. Atom C(15) is 0.66 Å above this plane. The greatest deviation for any of the atoms comprising the three collinear rings is  $+0.034$  Å for C(3). The carboxyl ion and the bonded C(14) atom, coplanar to within 0.003 Å, lie in plane B which

Table 4. Fractional coordinates and thermal parameters

Thermal parameters are of the form  $T = \exp[-\frac{1}{4}(B_{11}h^2a^{*2} + B_{22}k^2b^{*2} + B_{33}l^2c^{*2} + 2B_{12}hka^{*}b^{*} + 2B_{13}hla^{*}c^{*} + 2B_{23}klb^{*}c^{*})]$ . The  $B_{ij}$  values are in Å<sup>2</sup> units.

	<i>x</i>	<i>y</i>	<i>z</i>	<i>B</i> <sub>11</sub>	<i>B</i> <sub>22</sub>	<i>B</i> <sub>33</sub>	<i>B</i> <sub>12</sub>	<i>B</i> <sub>13</sub>	<i>B</i> <sub>23</sub>	<i>B</i> <sup>†</sup>
C(1)	0.3386	0.4202	0.1486	1.13	2.48	2.85	-0.28	-0.15	0.66	2.0
N(2)	0.2614	0.3806	0.1014	1.36	2.79	2.67	-0.71	-0.76	0.33	2.3
C(3)	0.1844	0.4200	0.0639	1.98	4.19	3.20	-0.71	-0.91	1.16	3.1
N(4)	0.2093	0.4845	0.0923	2.34	3.26	3.56	-0.05	-1.06	1.49	2.8
C(5)	0.3036	0.4869	0.1442	1.33	2.64	3.38	-0.07	-0.30	1.16	2.5
C(6)	0.3681	0.5356	0.1882	2.37	2.02	4.10	0.14	-0.15	0.99	2.6
C(7)	0.4637	0.5169	0.2357	2.37	1.86	3.03	-0.21	-0.22	0.33	2.2
C(8)	0.5371	0.5661	0.2822	2.66	1.86	3.56	0.07	-0.15	0.16	2.5
C(9)	0.6297	0.5496	0.3277	2.43	2.02	3.03	-0.49	-0.53	-0.11	2.3
C(10)	0.6629	0.4812	0.3295	2.53	2.17	2.49	-0.02	-0.38	0.14	2.2
C(11)	0.5915	0.4323	0.2858	2.27	1.70	2.49	-0.02	-0.45	0.16	1.9
C(12)	0.4977	0.4493	0.2386	2.01	1.70	2.49	-0.35	-0.38	0.16	1.9
N(13)	0.4264	0.4011	0.1924	1.65	1.70	2.49	-0.35	-0.53	0.33	1.7
C(14)	0.4380	0.3270	0.1913	2.24	1.70	2.49	-0.07	-0.60	0.08	1.9
C(15)	0.4269	0.2999	0.1214	2.73	2.64	2.67	-0.35	0.02	0.16	2.4
C(16)	0.2842	0.3101	0.0878	3.28	2.95	3.03	-0.49	-0.91	-0.09	2.7
C(17)	0.3129	0.2903	0.2297	2.50	2.02	2.67	-0.14	-0.45	0.33	1.9
O(18)	0.2988	0.2283	0.2245	2.95	1.70	4.27	-0.56	-0.03	0.49	2.5
O(19)	0.2392	0.3266	0.2609	6.66	2.64	9.62	-0.35	5.70	0.16	4.9
O(20)	0.7559	0.4711	0.3774	2.89	2.48	3.03	0.21	-1.52	-0.16	2.4
C(21)	0.7865	0.4031	0.3848	5.39	2.79	3.56	1.20	-1.52	0.14	3.2
O(22)	0.7049	0.5931	0.3738	4.29	2.02	4.10	-0.28	-1.29	-0.83	3.0
C(23)	0.6677	0.6615	0.3752	3.99	2.17	6.06	-0.14	-0.53	-0.99	3.8
C(24)	0.1409	0.5427	0.0688	3.31	4.81	5.34	1.06	-1.06	2.49	4.0
O(25)	0.1053	0.4006	0.0166	3.73	5.28	3.38	-0.85	-2.13	1.16	3.7
W(1)	0.0085	0.2901	0.3328	2.99	4.81	3.74	-0.28	0.07	0.33	3.7
W(2)	-0.0439	0.5571	0.4722	3.93	3.88	4.10	-0.71	-0.91	0.06	3.6
W(3)	0.0635	0.3183	0.4629	9.45	4.97	4.63	0.78	-1.59	-0.49	5.1
W(4)	0.3869	0.2918	0.4738	14.52	9.47	10.87	0.49	-1.67	-0.83	9.4
Standard deviations										
C	0.0006	0.0002	0.0002	0.32	0.22	0.20	0.16	0.16	0.16	
N	0.0005	0.0002	0.0002	0.26	0.16	0.16	0.14	0.14	0.13	
O	0.0004	0.0002	0.0002	0.26	0.14	0.14	0.14	0.14	0.11	
W(1-3)	0.0005	0.0002	0.0002	0.25	0.16	0.16	0.14	0.14	0.13	
W(4)	0.0010	0.0004	0.0004	0.65	0.48	0.48	0.42	0.42	0.32	

† Thermal factors for the isotropic refinement.

is perpendicular to the plane of the rings. This configuration results in a minimum intramolecular distance between  $N^+(13)$  and  $O^-(19)$  of only 2.64 Å.

This investigation affords a further insight into the preferred orientation of  $OCH_3$  groups in enol or phenol ethers. The two  $OCH_3$  are *ortho* to each other, in the plane of the adjacent aromatic ring, and are disposed in opposite directions. The  $C(11)\cdots C(21)$  and  $C(8)\cdots C(23)$  separations are only 2.82 and 2.77 Å,

respectively. Hydrogen atoms on the  $CH_3$  groups are oriented so that one H atom is in the plane of the heavy atoms and directed away from the ring, whereas the other two H atoms are symmetrically disposed above and below the plane maximizing the distance of closest approach between an H atom and  $C(8)$  or  $C(11)$  on the ring. In a photoproduct from a mescaline derivative (Yonemitsu, Nakai, Kanaoka, Karle & Witkop, 1969; Karle & Karle, 1970),

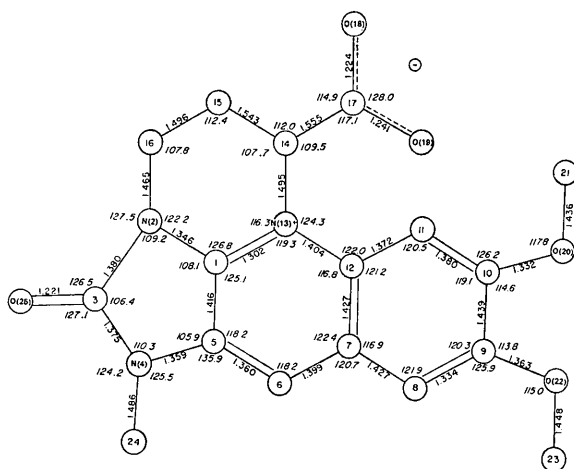
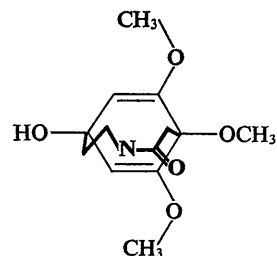


Fig. 4. Bond lengths and angles. Standard deviations based on the least-squares fit alone are 0.007 Å for the bond lengths and 0.4° for the angles.



the two  $OCH_3$  groups adjacent to the double bonds lie so

that each  $C=O$  group is planar, and the  $CH_3$  groups are *cis* to the double bonds. In the present investigation, this type of preferred orientation is reflected in the value of C–C bond lengths in the adjacent aromatic ring where  $C(8)–C(9)$  and  $C(10)–C(11)$ , the bonds *cis* to the  $O–CH_3$  bonds, are 0.10 and 0.06 Å shorter than the  $C(9)–C(10)$  bond. A similar observation has been made in the

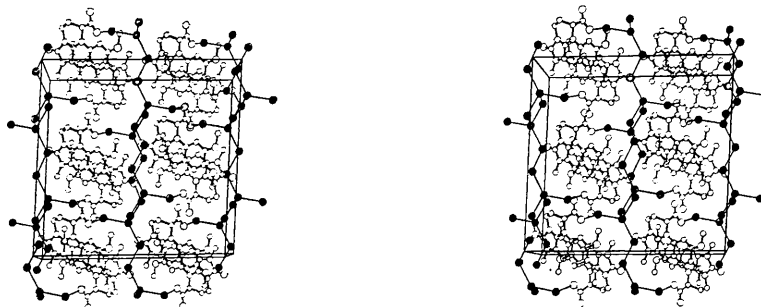


Fig. 5. Packing in the  $C_{17}H_{17}N_3O_5 \cdot 4H_2O$  crystal. Shaded atoms represent the O atoms of the  $H_2O$  molecules, and the lines radiating from them indicate possible hydrogen bonds. Axial directions are  $b \downarrow$ ,  $c \rightarrow$  and  $a$  out from the plane of the paper.

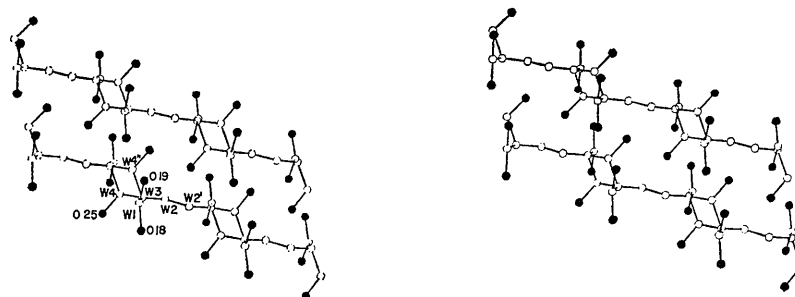
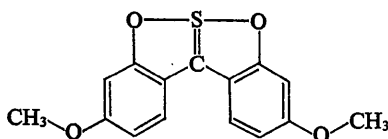
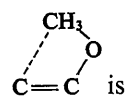


Fig. 6. A layer of  $H_2O$  molecules parallel to the  $ab$  plane. Shaded atoms represent O(18), O(19), and O(25) from the organic moiety. The lines indicate possible hydrogen bonds. Axial directions are  $a \uparrow$ ,  $b \rightarrow$  and  $c$  into the plane of the paper.



molecule (Gilardi & Karle, 1971), where the C–C bonds *cis* to the O–CH<sub>3</sub> bonds are 0.04–0.05 Å shorter than the C–C bonds *trans* to the O–CH<sub>3</sub> bonds. In each of the aforementioned molecules, as well as in reserpine

(Karle & Karle, 1968), the C···C distance in  is only 2.77 to 2.82 Å. Observations of preferred orientation, by means of the nuclear Overhauser effect, have also been made on enol and phenol ethers in solution; the methoxy and vinyl protons *cis* to the ethylenic bond show long-range coupling (Woods, Miura, Ogiso, Kurabayashi & Mishima, 1968). If a methoxy group is not coplanar with an attached unsaturated ring, such as the central OCH<sub>3</sub> group in 3,4,5-trimethoxyphenyl derivatives, the chemical reactivity of the nonplanar OCH<sub>3</sub> is enhanced (Brossi & Teitel, 1969).

The high-water content in this crystal, four H<sub>2</sub>O molecules per organic molecule, is interesting considering that the organic moiety participates in only three hydrogen bonds, O(25)···HW(4), O(18)···HW(1), and O(19)···HW(1). Fig. 5 (stereodiagram) illustrates the packing of the molecules and shows the layered structure in which layers of organic molecules are interleaved with layers of H<sub>2</sub>O molecules, hydrogen-bonded to each other. The H<sub>2</sub>O molecules are easily lost from the lattice, destroying the crystal in the process. The depiction in Fig. 6 is at right-angles to that in Fig. 5 and it shows a portion of the layer of H<sub>2</sub>O molecules parallel to the *ab* plane. Thermal parameters for W(4) are rather large, probably an indication of a reduced occupancy. Occupancy parameters were not refined. It may be at this site that the lattice begins to lose its water content.

Lengths for the hydrogen bonds are shown in Table

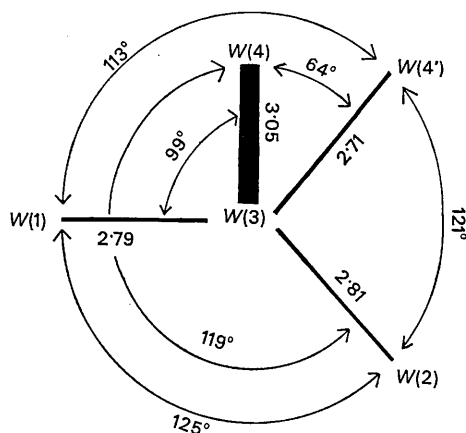


Fig. 7. Geometry about the W(3) atom.

5. Hydrogen atoms on three water molecules were not located, and there are several different ways in which they could be directed. A difficulty arises in presuming that a hydrogen bond exists between W(2) and W(2'), formed across a crystallographic center of symmetry. Since a hydrogen bond requires one H atom between two O atoms, and a center of symmetry implies either two H atoms or none, there may be a small deviation from a centrosymmetric space group. Another question arises as to whether there is hydrogen bonding between W(3) and W(4) (see Figs. 5, 6 & 7). There are a sufficient number of protons for one to exist between W(3) and W(4); the W(3)···W(4) distance of 3.05 Å, although somewhat long for an OH···O bond, has been observed in other hydrates; and the coordination about W(3) is unusual, being a distorted trigonal bipyramid with one position vacant. However, a trigonal bipyramid type of coordination about a water molecule has been established in at least one inorganic crystal, Na<sub>4</sub>XeO<sub>6</sub>·8H<sub>2</sub>O (Ibers, Hamilton & MacKenzie, 1964).

Table 5. Possible hydrogen bonds

	Lengths (Å)	Symmetry operation on primed atoms
O(18)···W(1')	2.97	$\bar{x}, \frac{1}{2}-y, \frac{1}{2}-z$
O(19)···W(1)	2.70	
W(1)···W(3)	2.79	
W(2)···W(2')	2.81	$\bar{x}, 1-y, 1-z$
W(2)···W(3')	2.81	$\bar{x}, 1-y, 1-z$
W(3)···W(4')	2.71	$\frac{1}{2}-x, \frac{1}{2}-y, 1-z$
W(3)···W(4)	3.05	
W(4')···O(25)	2.81	$-\frac{1}{2}+x, y, -\frac{1}{2}+z$

#### References

- BEIJERINCK, M. W. (1907). *Zentr. Bakteriolog. Parasitenk.* II, 2, 561.
- BROSSI, A. & TEITEL, S. (1969). *Organic Preparations and Procedures*, 1(3), 171.
- BULEN, W. A. & LECOMTE, J. R. (1962). *Biochem. Biophys. Res. Comm.* 9, 523.
- CORBIN, J. L. (1970). Private communication.
- CORBIN, J. L., KARLE, I. L. & KARLE, J. (1970). *Chem. Commun.* 186.
- GILARDI, R. D. & KARLE, I. L. (1971). *Acta Cryst.* B27, 1073.
- IBERS, J. A., HAMILTON, W. C. & MACKENZIE, D. R. (1964). *Inorg. Chem.* 3, 1412.
- JOHNSON, C. K. (1965). *ORTEP, A Fortran Thermal-Ellipsoid Plot Program for Crystal Structure Illustrations*. Report ORNL-3794, Oak Ridge National Laboratory, Oak Ridge, Tennessee.
- KARLE, I. L. & KARLE, J. (1963). *Acta Cryst.* 16, 969.
- KARLE, I. L. & KARLE, J. (1968). *Acta Cryst.* B24, 81.
- KARLE, I. L. & KARLE, J. (1970). *Acta Cryst.* B26, 1276.
- KARLE, J. (1968). *Acta Cryst.* B24, 182.
- KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* 9, 635.
- KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* 21, 849.
- WOODS, M. C., MIURA, I., OGISO, A., KURABAYASHI, M. & MISHIMA, H. (1968). *Tetrahedron Letters*, p. 2009.
- YONEMITSU, O., NAKAI, H., KANAOKA, Y., KARLE, I. L. & WITKOP, B. (1969). *J. Amer. Chem. Soc.* 91, 4591.