Acta Cryst. (1966). **21**, 860

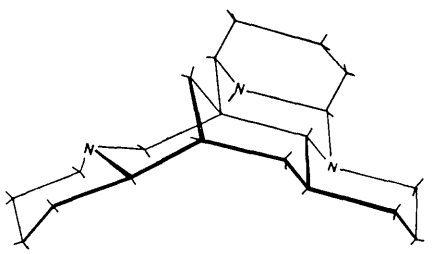
An Application of the Symbolic Addition Procedure to Space Group $P2_1$ and the Structure of the Alkaloid Panamine, $C_{20}H_{33}N_3$

BY ISABELLA L. KARLE AND J. KARLE
U.S. Naval Research Laboratory, Washington, D.C. 20390, *U.S.A.*

(*Received* 18 *February* 1966)

The symbolic addition procedure for obtaining phases directly from the structure factor magnitudes has been applied to the structure determination of the noncentrosymmetric crystal formed from the diperchlorate of the alkaloid panamine. The space group is $P2_1$ and the cell dimensions are $a=10{\cdot}91$, $b=8{\cdot}57$, $c=13{\cdot}58$ Å and $\beta=112{\cdot}7°$.

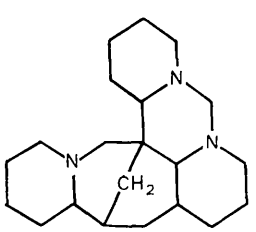
Panamine is an alkaloid isolated from the seeds of the *Ormosia* species. Its molecular configuration is found to be



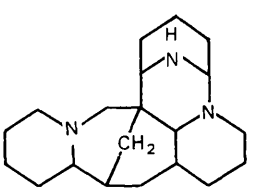
The molecule is composed of six six-membered puckered rings, five of which have the chair configuration and one the boat configuration. In the asymmetric unit only three hydrogen atoms are available to form hydrogen bonds with the eight oxygen atoms in the two perchlorate groups. The large thermal parameters found to be associated with the perchlorate groups are possibly a consequence of the limited numbers of hydrogen bonds available.

Introduction

A number of related alkaloids of unknown molecular formulae have been extracted from the seeds of *Ormosia panamensis* and *Ormosia jamaicensis* (see *e.g.* Lloyd & Horning, 1958, 1960; Lloyd, 1961). The molecular structure and stereoconfiguration of one of the alkaloids, jamine, $C_{21}H_{35}N_3$, which crystallizes in space group $P\bar{1}$ was elucidated recently (Karle & Karle, 1964*a*). On the basis of the structural formula of jamine and additional chemical and spectral data, a structural formula has been proposed by Wilson (1965) for the related alkaloid panamine, $C_{20}H_{33}N_3$:



jamine



panamine

The present investigation of the crystal structure of the diperchlorate salt of panamine confirms Wilson's proposed structural formula and establishes the stereoconfiguration of the molecule. The material crystallizes in space group $P2_1$ and therefore affords another opportunity for applying the symbolic addition procedure for phase determination to a noncentrosymmetric space group (Karle & Karle, 1964*b*). The details of the phase determination will be described.

Experimental

Crystals of panamine diperchlorate in the form of colorless stout prisms were kindly supplied by Drs H.A. Lloyd and P. Naegeli of the National Institutes of Health. Cell parameters were obtained from precession photographs taken with Cu $K\alpha$ radiation. The unit-cell dimensions are:

$$a=10.91\pm0.02 \text{ Å}$$
$$b=\ 8.57\pm0.02$$
$$c=13.58\pm0.02$$
$$\beta=112.7°\pm0.3°$$

For $Z=2$, the computed density is 1.459 g.cm^{-3}, whereas the density observed in mixed solvents is 1.447 g.cm^{-3}. The only systematic absences occurred among the $0k0$ reflections with k odd, hence the space group $P2_1$ was indicated.

Multiple-film equi-inclination Weissenberg photographs were taken along the b axis through the sixth layer and along the c axis through the ninth layer. The

intensities were estimated visually by comparison with a calibrated film strip. Corrections were made for spot-size, and Lorentz and polarization factors. The data from the two-different axes were cross correlated and placed on an absolute scale by means of a K curve. Structure factors $|F|$ and normalized structure factors $|E|$ were then computed. The above calculations were all made with one computer program for data reduction. The relationship between $|E|$ and $|F|$ for space group $P2_1$ is

$$|E_{\mathbf{h}}|^2 = |F_{\mathbf{h}}|^2/\varepsilon \sum_{j=1}^{N} f_{j\mathbf{h}}^2 \qquad (1)$$

where the $|F_{\mathbf{h}}|^2$ values are those from which the effects of thermal motion have been removed, $\varepsilon = 1$ except for the $0k0$ reflections where $\varepsilon = 2$, N is the number of atoms in the unit cell, and $f_{j\mathbf{h}}$ is the atomic scattering factor for the jth atom. The computed values for the statistical averages $\langle|E|\rangle$ and $\langle|E^2-1|\rangle$ were 0.82 and 0.77, respectively. They may be compared with 0.89 and 0.74 calculated for a noncentrosymmetric crystal with atoms in random positions. The total number of observed independent data was 2673.

Phase determination

The procedure for phase determination was carried out in a manner very similar to that used for L-arginine dihydrate (Karle & Karle, 1964*b*). A general description of this procedure has recently been given (Karle & Karle, 1966). The formula for starting the phase determination is

$$\varphi_{\mathbf{h}} \approx \langle \varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}} \rangle_{\mathbf{k}_r}, \qquad (2)$$

which is an immediate consequence of the third general inequality, expression (34) of Karle & Hauptman (1950). The $\varphi_{\mathbf{h}}$ are the phases associated with the $|E_{\mathbf{h}}|$ and $\mathbf{k}_r$ implies that the average is taken only over those $\mathbf{k}$ corresponding to the larger $|E|$ values. Formula (2) is implemented by assigning three phases to specify the origin in space group $P2_1$ (Hauptman & Karle, 1956) and by assigning additional symbols to represent the phases of certain other reflections in a stepwise fashion as needed. Need is indicated if certain of the subsets of the reflections are inaccessible by the assignments already made or if the phase indications exceed a variance criterion described below. In all these assignments, $|E_{\mathbf{h}}|$ must have a large magnitude and $\mathbf{h}$ should enter into many combinations for the average in (2). Ordinarily contributors to (2) in which all three reflections, $\mathbf{h}$, $\mathbf{k}$, and $\mathbf{h}-\mathbf{k}$, have one or two indices zero ($h0l$, for example) are avoided since experience has shown that such combinations often indicate an incorrect phase. The phase assignments in the present investigation are shown in Table 1. The symbol a must be either 0 or π, whereas the symbols b and c can have values anywhere between $-\pi$ and π.

Table 1. *Phase assignments for specifying the origin and implementing equation* (2)

h	$\Phi_{\mathbf{h}}$	$\|E_{\mathbf{h}}\|$
$\bar{4}\,0\,15$	0	3·09
7 0 1	0	2·91
$\bar{9}\,1\,1$	0	2·72
8 0 2	a	3·43
$\bar{7}\,5\,9$	b	2·55
5 2 8	c	2·65

In the initial stage of phase determination, the average on the right-hand side of (2) often consists of a single term. In order for a single term to indicate reliably the phase of a new reflection, $E_{\mathbf{h}}$, $E_{\mathbf{k}}$ and $E_{\mathbf{h}-\mathbf{k}}$ should all have very large magnitudes. For centrosymmetric crystals, the probability of determining the signs correctly is evaluated by use of a well-known probability formula. In contrast for noncentrosymmetric crystals, the probable agreement of $\varphi_{\mathbf{h}}$ with the indications of $\langle \varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}} \rangle_{\mathbf{k}_r}$ is given in terms of a measure of the variance. A curve of the variance of $\varphi_{\mathbf{h}}$ as a function of the known pairs $\varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}}$ and the associated $|E|$ values can be found in the above-mentioned paper describing the symbolic addition procedure (Karle & Karle, 1966, Fig. 2). In the present application, a phase was not accepted unless the variance was smaller than 0.5 square radian, whether there were one or many terms in the average in (2).

In the course of application, for each new $\mathbf{h}$ considered, more and more terms can be included in the average in (2). It was soon apparent that the symbol a should have the value 0 in order to avoid a large number of exceptions among the contributors to (2). Furthermore, it was also apparent that $b \approx 2c$. These relationships occurred many times in the process of obtaining the initial 157 phases ($|E_{\mathbf{h}}| > 1.60$). At this stage, phases had not been determined for all $|E_{\mathbf{h}}| > 1.60$. There are always some reflections with fairly large $|E|$ values for which there are insufficient or seemingly ambiguous indications of phase using only relationship (2).

The next step was to reiterate the determination of values for the initial set of phases and to obtain additional phases by the use of the tangent formula (Karle & Hauptman, 1956)

$$\tan \varphi_{\mathbf{h}} \approx \frac{\sum_{\mathbf{k}} |E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}}| \sin (\varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}})}{\sum_{\mathbf{k}} |E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}}| \cos (\varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}})}. \qquad (3)$$

Since the numerator of (3) is proportional to $|E_{\mathbf{h}}| \sin \varphi_{\mathbf{h}}$ and the denominator of (3) is proportional to $|E_{\mathbf{h}}| \cos \varphi_{\mathbf{h}}$, the sum of the squares is proportional to $|E_{\mathbf{h}}|^2$. If this formula is calculated for many $\mathbf{h}$ values, then the scaling constant can be found from a comparison of $\sum_{\mathbf{h}} |E_{\mathbf{h}}|^2_{\text{calc}}$ and $\sum_{\mathbf{h}} |E_{\mathbf{h}}|^2_{\text{obs}}$ and hence comparisons can be made between individual $|E_{\mathbf{h}}|^2_{\text{calc}}$ and $|E_{\mathbf{h}}|^2_{\text{obs}}$. In the course of applying (3), particular reflections were rejected if their $|E_{\mathbf{h}}|_{\text{calc}}$ were very small, *i.e.* less than 0.3, or if their phases changed radically upon reiteration of (3), or if the number of contributors to the sums in (3) was very small, *i.e.* less than four.

The tangent formula (3) converges in one or two iterations if no new data are introduced.

From the application of (2), 157 phases were determined as 0, π, or functions of b and c where $b \approx 2c$. In order to apply (3), values had to be assigned to b and c. If c were near 0 or π, b would be near 0 and all phases would be essentially real. Since we are dealing with a noncentrosymmetric crystal, it is expected that a number of phases should be near $\pm\pi/2$. Consequently, as a first trial, c was assigned the value $+\pi/2$ with b necessarily equal to π. This assignment proved to be correct. If the first assignment had not been sufficiently good, then values of $+\pi/4$ and $+3\pi/4$ for c would have been used as subsequent trials. Only positive values need be considered for c, since in this manner one of the two possible enantiomorphs is specified. Once the origin is specified, the imaginary parts of all structure factors change sign in going from one enantiomorph to the other.

With the assignment of $c = +\pi/2$, the initial 157 phases were refined and an additional 143 were determined ($|E_h| > 1.45$) with the use of (3). Of these, 27 reflections were rejected for the reasons stated above. An E map computed with these 270 phases revealed the positions of the two Cl atoms and the 23 C and N atoms of panamine (Fig. 1). The eight O atoms belonging to the two perchlorate groups were not found in this map. A comparison of the phases determined by (2) and (3) with those computed from the refined structure, showed an average error of 22° for the 235 three-dimensional reflections. The distribution of the error is illustrated in Fig. 2.

Refinement

The approximate coordinates for the twenty-three C and N atoms of panamine and the two Cl atoms were obtained from the initial E map (Fig. 1). These coordinates were subjected to a preliminary least-squares refinement with the ORFLS program (Busing, Martin & Levy, 1962) which has been adapted in our laboratory to the CDC 3600 machine. The quantity minimized was $\Sigma (F_o - F_c)^2$. A difference map obtained with phases based on the known positions of twenty-five atoms, revealed the positions of the eight O atoms in the two perchlorate ions. The spread of the density representing each O atom indicated either a large thermal motion or some disorder in the ClO_4 groups which would account for the fact that the O atoms were not at all well-defined in the initial E map. Several cycles of refinement with isotropic temperature factors were performed and then several cycles with anisotropic temperature factors. For the anisotropic refinement, the computer was not large enough to refine all the parameters simultaneously, hence the parameters for the eight O atoms and the twenty-five remaining atoms were refined in alternate cycles. The refinement was terminated at $R = 11.9\%$ for the observed data. The observed and calculated structure factors, $|F_o|$ and $|F_c|$, and the phase φ are listed in Table 2.

Structure

The fractional coordinates and anisotropic thermal factors are listed in Table 3. Fig. 3 is a composite from the final electron density map and shows the molecule projected along the c axis. The unit cell contents projected along the b axis are illustrated in Fig. 4. The atoms have been labelled in a similar fashion, as far as possible, to those in jamine (Karle & Karle, 1964*a*).

Like jamine, the panamine molecule is composed of six six-membered puckered rings, five of which have the chair configuration. Only ring C (Fig. 5) is in the boat configuration. It is noteworthy that both rings E and F are in the chair configuration in panamine. The bond angles in these two rings are somewhat larger than in the rest of the molecule and have increased the potentially short distance between C(5) and C(10) to 3.15 Å, thus allowing sufficient distance between the

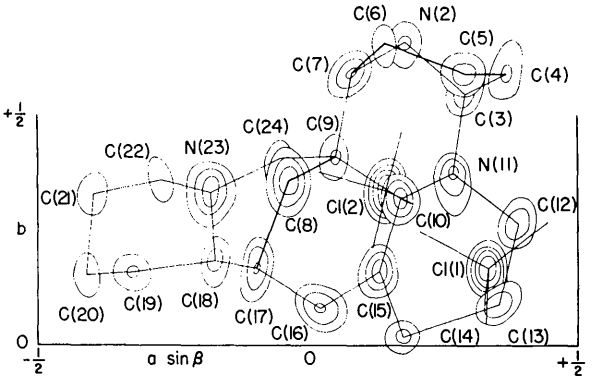


Fig. 1. Sections from a three-dimensional E-map projected down the c axis. The contours are at equal intervals on an arbitrary scale except that for the Cl atoms every other contour is missing. The E-map, a Fourier employing the E_h rather than the F_h as coefficients (Karle, Hauptman, Karle & Wing, 1958), was computed using only the 270 terms for which phases were determined using the symbolic addition procedure.

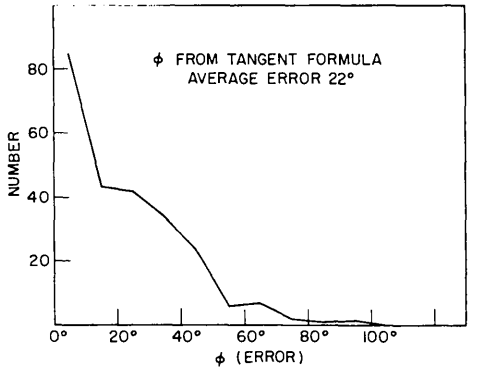


Fig. 2. The distribution of the difference between the values of the phases for the three-dimensional reflections as determined by the tangent formula and the final values calculated from the refined structure.

Table 2. *Structure factors for panamine*

The four columns give l, $|F_o|$, $|F_c|$, and the phase Φ (radians)

Table 2 (*cont.*)

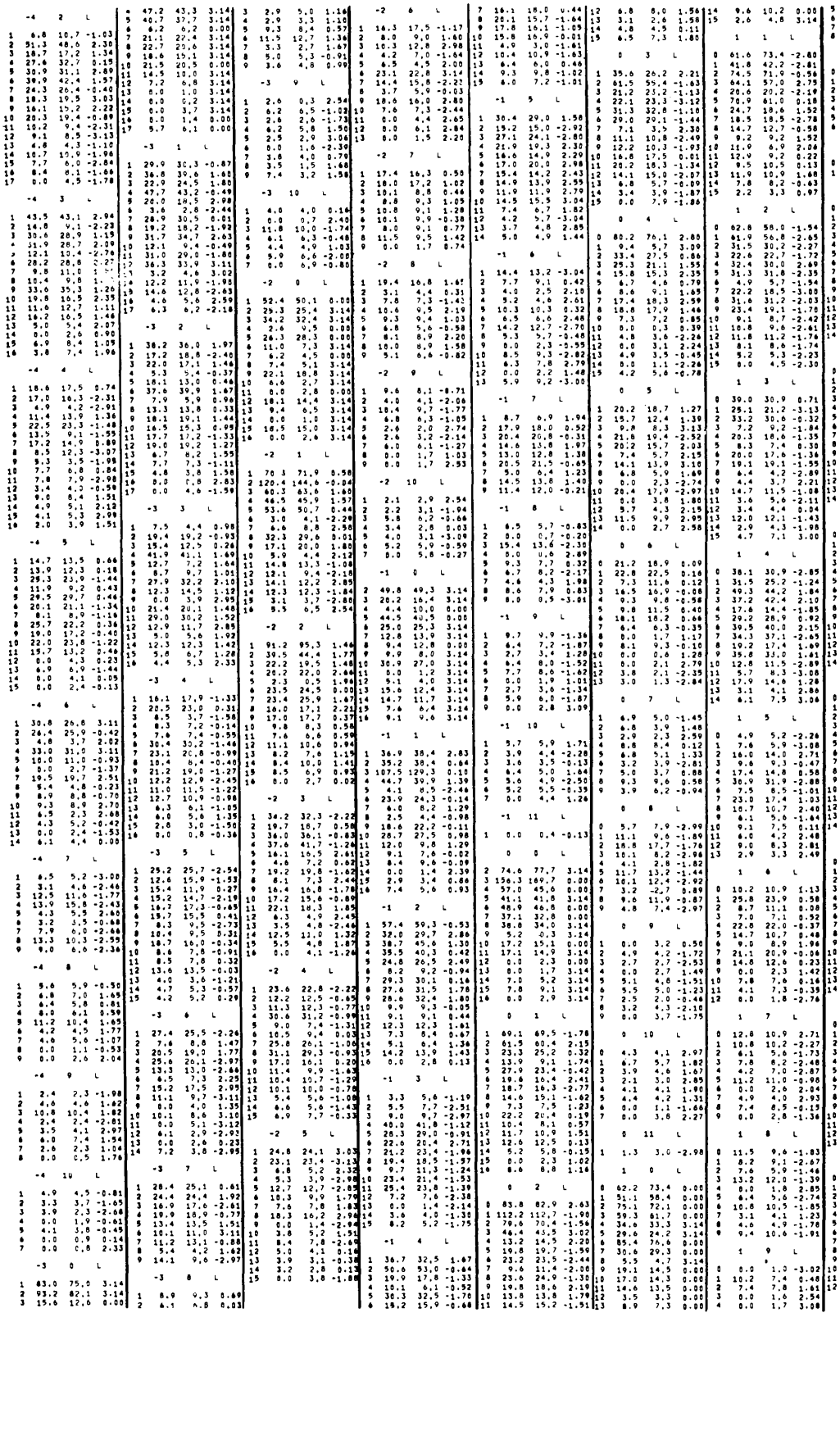


Table 2 (*cont.*)

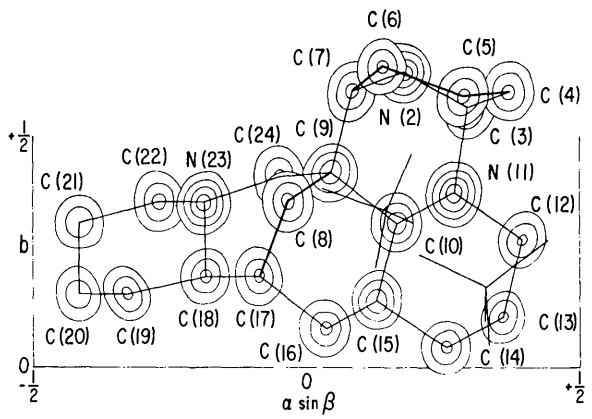


Fig. 3. Sections from the final three-dimensional electron density map projected down the *c* axis. For clarity, the density contours for the ClO_4 groups have been omitted. The contours are spaced at 2 e.Å^{-3} starting eith the 2 e.Å^{-3} contour.

hydrogen atoms attached to C(5) and C(10). Rings *A*, *B*, *C*, *D*, and *E* are the same as in jamine. The two molecules differ in ring *F*. In jamine there is an additional $-CH_2$ group to form the six-membered ring *F* which is attached to N(2) rather than C(3). The in-

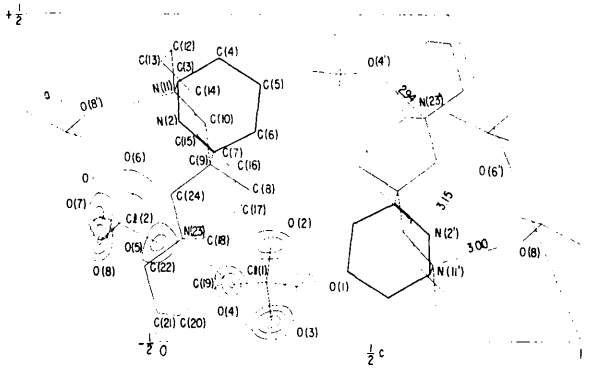


Fig. 4. Contents of the unit cell projected down the *b* axis. The density contours are indicated for the ClO_4 groups. The three possible NH···O bonds are shown.

Table 3. *Fractional coordinates and anisotropic temperature factors for panamine diperchlorate*

The thermal parameters are of the form $T=\exp[-(\beta_{11}h^2+\beta_{22}k^2+\beta_{33}l^2+2\beta_{12}hk+2\beta_{13}hl+2\beta_{23}kl)]$. Each thermal parameter is multiplied by 10^4.

Atom	*x*	*y*	*z*	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
N(2)	1·1708	0·6342	0·2592	64	90	50	−2	4	13
C(3)	1·2946	0·5534	0·2978	56	136	54	13	15	11
C(4)	1·3696	0·5908	0·4147	65	152	45	−20	−2	−18
C(5)	1·2854	0·5795	0·4877	62	210	40	11	2	−24
C(6)	1·1409	0·6417	0·4291	73	157	48	2	7	−23
C(7)	1·0787	0·5921	0·3108	48	120	36	13	9	−3
C(8)	0·9629	0·3566	0·3603	66	148	43	0	23	4
C(9)	1·0407	0·4167	0·2922	32	109	31	10	4	3
C(10)	1·1643	0·3110	0·3199	45	81	38	−9	12	6
N(11)	1·2654	0·3709	0·2768	37	117	45	7	14	−9
C(12)	1·3908	0·2733	0·3108	65	144	106	43	32	2
C(13)	1·3544	0·1049	0·2752	74	141	127	16	60	−19
C(14)	1·2553	0·0402	0·3202	53	99	100	7	11	8
C(15)	1·1299	0·1403	0·2887	42	87	52	0	6	−3
C(16)	1·0323	0·0826	0·3381	70	123	63	0	11	37
C(17)	0·9133	0·1925	0·3237	63	148	59	−3	23	32
C(18)	0·8162	0·1921	0·2029	41	118	61	2	8	−4
C(19)	0·6703	0·1592	0·1936	62	182	97	−27	30	10
C(20)	0·5808	0·1520	0·0740	37	270	122	−19	6	3
C(21)	0·5865	0·3113	0·0256	62	209	88	−10	−22	−32
C(22)	0·7318	0·3553	0·0364	45	233	43	6	−19	8
N(23)	0·8146	0·3531	0·1558	38	118	44	0	−5	4
C(24)	0·9509	0·4093	0·1705	28	123	37	−8	5	−5
Cl(1)	0·3272	0·1670	0·6827	60	160	66	−14	26	8
O(1)	0·3289	0·0473	0·6067	180	859	262	−38	113	−242
O(2)	0·2018	0·2459	0·6400	63	280	341	32	−19	102
O(3)	0·4418	0·2670	0·7159	85	252	173	−106	37	1
O(4)	0·3277	0·0918	0·7741	224	445	92	−203	5	67
Cl(2)	0·1409	0·3345	0·9778	105	137	39	−11	26	−9
O(5)	0·2054	0·3068	0·9065	334	403	209	−173	208	−136
O(6)	0·0169	0·3824	0·8988	289	544	239	182	−150	24
O(7)	0·1183	0·2088	1·0302	556	191	185	−24	182	39
O(8)	0·1859	0·4627	1·0472	332	377	150	−181	153	−194
Standard error									
C	0·0124	0·0161	0·0141	11	23	9	14	8	13
N	0·0092	0·0115	0·0100	8	15	6	9	6	9
Cl	0·0034	0·0040	0·0034	3	6	2	4	2	3
O	0·0169	0·0194	0·0184	23	45	17	26	16	24

Table 4. *Bond lengths and angles*

The standard deviations for the C–C and C–N bond lengths are of the order of 0·023 Å and for the Cl–O bond lengths they range from 0·015 to 0·030 Å. The standard deviations for the angles are near 1·3°.

Lengths		Angles	
N(2)–C(3)	1·425 Å	N(2)C(3)C(4)	109·6°
C(3)–C(4)	1·514	C(3)C(4)C(5)	115·6
C(4)–C(5)	1·593	C(4)C(5)C(6)	111·9
C(5)–C(6)	1·560	C(5)C(6)C(7)	112·9
C(6)–C(7)	1·543	C(6)C(7)N(2)	108·7
C(7)–N(2)	1·475	C(7)N(2)C(3)	115·4
		N(2)C(7)C(9)	110·3
		C(6)C(7)C(9)	114·7
		C(7)C(9)C(10)	112·0
		C(7)C(9)C(8)	113·2
C(7)—C(9)	1·554	C(7)C(9)C(24)	103·3
C(9)—C(10)	1·545	C(9)C(10)N(11)	113·7
C(10)–N(11)	1·523	C(10)N(11)C(3)	112·6
N(11)–C(3)	1·600	N(11)C(3)C(4)	112·6
		N(11)C(3)N(2)	108·3
		C(3)N(11)C(12)	113·0
		C(10)N(11)C(12)	113·5
N(11)–C(12)	1·514	N(11)C(12)C(13)	109·7
C(12)–C(13)	1·523	C(12)C(13)C(14)	110·2
C(13)–C(14)	1·545	C(13)C(14)C(15)	112·4
C(14)–C(15)	1·529	C(14)C(15)C(10)	111·2
C(15)–C(10)	1·528	C(14)C(15)C(16)	113·0
		C(15)C(10)N(11)	110·7
		C(15)C(10)C(9)	113·3
		C(10)C(15)C(16)	108·9
C(15)–C(16)	1·542	C(15)C(16)C(17)	115·7
C(16)–C(17)	1·553	C(16)C(17)C(8)	110·2
C(17)–C(8)	1·520	C(16)C(17)C(18)	108·9
C(8)—C(9)	1·563	C(17)C(8)C(9)	108·5
		C(8)C(9)C(10)	106·6
		C(8)C(9)C(24)	110·5
		C(10)C(9)C(24)	111·4
		C(8)C(17)C(18)	110·8
C(17)–C(18)	1·572	C(17)C(18)C(19)	109·7
C(18)–N(23)	1·518	C(17)C(18)N(23)	109·4
N(23)–C(24)	1·500	C(18)N(23)C(24)	112·7
C(24)–C(9)	1·563	N(23)C(24)C(9)	109·8
		N(23)C(18)C(19)	106·4
C(18)–C(19)	1·572	C(18)C(19)C(20)	107·6
C(19)–C(20)	1·540	C(19)C(20)C(21)	107·7
C(20)–C(21)	1·526	C(20)C(21)C(22)	112·3
C(21)–C(22)	1·580	C(21)C(22)N(23)	104·9
C(22)–N(23)	1·521	C(22)N(23)C(18)	111·5

Table 4 (*cont.*)

Lengths		Angles	
		C(22)N(23)C(24)	106·8
Cl(1)–O(1)	1·459	O(1)Cl(1)O(2)	108·4
Cl(1)–O(2)	1·432	O(1)Cl(1)O(3)	113·3
Cl(1)–O(3)	1·437	O(1)Cl(1)O(4)	107·9
Cl(1)–O(4)	1·400	O(2)Cl(1)O(3)	115·1
		O(2)Cl(1)O(4)	104·7
		O(3)Cl(1)O(4)	106·9
Cl(2)–O(5)	1·419	O(5)Cl(2)O(6)	96·3
Cl(2)–O(6)	1·424	O(5)Cl(2)O(7)	117·4
Cl(2)–O(7)	1·364	O(5)Cl(2)O(8)	116·8
Cl(2)–O(8)	1·406	O(6)Cl(2)O(7)	108·1
		O(6)Cl(2)O(8)	105·4
		O(7)Cl(2)O(8)	110·7

dividual bond lengths and bond angles are listed in Table 4. The average of twenty C–C bond lengths is 1.548 Å and the average of eight C–N bond lengths is 1.509 Å. All the angles in the panamine molecule are near the tetrahedral value. The entire molecule is quite rigid as shown by the isotropic thermal parameters which are in the range 2.2–3.8 except for C(19), C(20) and C(21) where they are between 4.0 and 5.0.

Unlike the organic molecule, the two perchlorate groups have very large thermal parameters which made accurate determination of the coordinates of the O atoms rather difficult. The positions of the ClO_4 groups may be rather flexible because there is little stabilization from hydrogen bonding. Only three hydrogen bonds can be formed with the eight available O atoms. The possible hydrogen bonds are N(2)–O(6) at 3.15 Å, N(11)–O(8′) at 3.00 Å, and N(23)–O(4) at 2.94 Å (Fig. 4); hence the two extra protons must attach themselves to N(11) and N(23) to form a doubly charged ion.

The packing is illustrated in Fig. 4. In the *c* direction, the perchlorate ions distribute themselves between the organic ions. The nearest intermolecular approaches are between O(4) and O(5) at 3.20 Å and between O(6) and O(7′) at 3.47 Å. The only other near approach is between N(2) and C(14) in the molecule above (*b* direction) at 3.61 Å. All other intermolecular distances are larger than 3.8 Å.

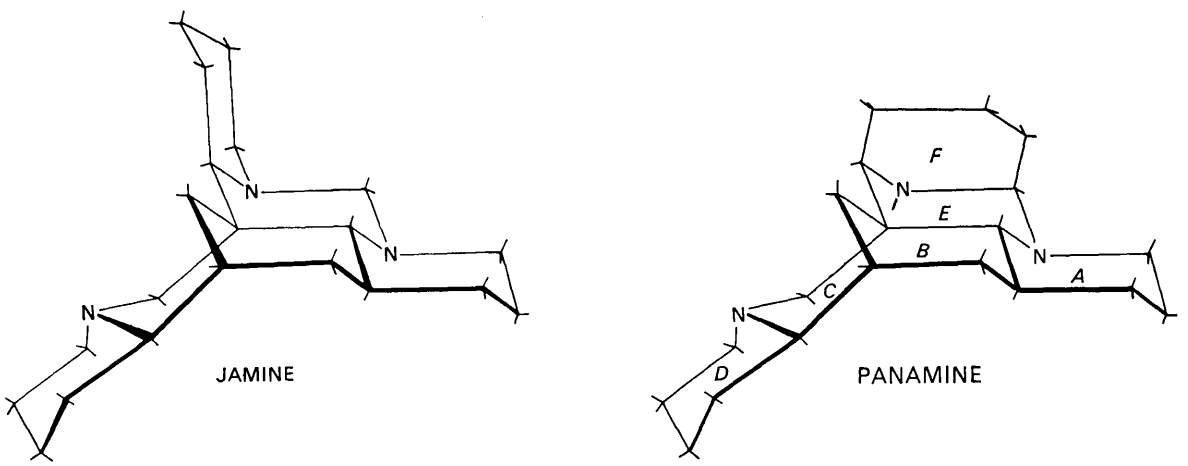


Fig. 5. Stereoconfiguration of panamine and jamine.

We wish to acknowledge the invaluable assistance of Mr S. Brenner and Dr D. Mitchell, who performed all the calculations on high-speed computers, and of Mrs K. Dragonette who read the intensities of the photographic films.

References

HAUPTMAN, H. & KARLE, J. (1956). *Acta Cryst.* **9**, 45.
KARLE, I. L., HAUPTMAN, H., KARLE, J. & WING, A. B. (1958). *Acta Cryst.* **11**, 257.
KARLE, I. L. & KARLE, J. (1964*a*). *Acta Cryst.* **17**, 1356.
KARLE, I. L. & KARLE, J. (1964*b*). *Acta Cryst.* **17**, 835.
KARLE, J. & HAUPTMAN, H. (1950). *Acta Cryst.* **3**, 181.
KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635.
KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849.
LLOYD, H. A. & HORNING, E. C. (1958). *J. Amer. Chem. Soc.* **80**, 1506.
LLOYD, H. A. & HORNING, E. C. (1960). *J. Org. Chem.* **25**, 1959.
LLOYD, H. A. (1961). *J. Org. Chem.* **26**, 2143.
WILSON, E. M. (1965). *Tetrahedron*, **21**, 2561.

Acta Cryst. (1966). **21**, 868

Transformations of Variance-Covariance Tensors

BY DONALD E. SANDS

Department of Chemistry, University of Kentucky, Lexington, Kentucky, U.S.A.

(*Received* 19 *May* 1966)

The variances and covariances of atomic coordinates are elements of second-rank contravariant tensors. The effects of transformations of axes, symmetry operations, and shifts of origin have been determined. Formulas have been derived for calculating the variances and covariances of bond lengths and angles. The principal axes of variance–covariance ellipsoids have been determined, and a method has been obtained for calculating the minimum variance weighted mean of correlated variables.

Introduction

The variance of a function f of a set of atomic parameters x^i is given by

$$\sigma^2(f)=\Sigma \frac{\partial f}{\partial x^i}\frac{\partial f}{\partial x^j}\operatorname{cov}(x^i, x^j) \qquad (1)$$

where the summation is over all parameters included in f. The quantity $\operatorname{cov}(x^i, x^j)$ is the covariance of x^i and x^j if these are different parameters; it is the variance of x^i if $i=j$. This equation is strictly valid for linear functions, but it is still useful for slowly varying non-linear functions (Arley & Buch, 1950). The v.c. (variance-covariance) matrix can be obtained from the least-squares inverse matrix (Hamilton, 1964).

If f is a function of crystallographically equivalent atoms, certain parameters will be linearly related to others, and the corresponding covariances will be non-zero. The covariances between the parameters of a single atom may also be non-zero; this will be the case if symmetry imposes a relationship between the parameters, but it is also true of an atom in a general position in an oblique coordinate system (Templeton, 1959). Even crystallographically independent parameters in an orthogonal system may not be statistically independent, and the covariances will not be exactly zero.

If f is a scalar function, the derivatives in (1) with respect to contravariant vector components are the covariant* components of a vector. The invariance of $\sigma^2(f)$ with respect to linear transformation is proof that the v.c. matrix of the atomic coordinates is a second-rank contravariant tensor (*International Tables for X-ray Crystallography*, 1959). That is, v.c. matrices transform in the same way as products of coordinates.

Among the functions of crystallographic interest are bond lengths and bond angles. Frequently, the parameters used in computing these distances and angles are not those obtained directly in the structure determination, for which the complete v.c. matrix is presumed known, but are generated from these parameters by application of certain symmetry operators of the space group. The variance of f may be computed from the original v.c. matrix if the derivatives in (1) are evaluated with respect to the original parameters (Busing, Martin & Levy, 1964). Alternatively, the derivatives can be evaluated with respect to the transformed coordinates; in this case the transformed v.c. matrix must be used. These transformations are among the topics discussed in this paper.

A generalization of (1) is

$$\operatorname{cov}(f_k, f_l)=\Sigma \frac{\partial f_k}{\partial x^i}\frac{\partial f_l}{\partial x^j}\operatorname{cov}(x^i, x^j) \qquad (2)$$

where the summation is over all parameters x^i and x^j included in the functions f_k and f_l. The transformation properties of v.c. matrices are implicit in this formula. In matrix notation

$$\mathbf{F}=\mathbf{D}\,\mathbf{V}\,\bar{\mathbf{D}} \qquad (3)$$

where $\operatorname{cov}(f_k, f_l)$ is the klth element of $\mathbf{F}$, $\partial f_i/\partial x^j$ is the ijth element of $\mathbf{D}$, $\mathbf{V}$ is the v.c. matrix of the parameters, and $\bar{\mathbf{D}}$ is the transpose of $\mathbf{D}$.

* *Covariance* is both a differential geometry term which describes transformation properties and a statistical term for a measure of correlation. The intended meaning will be apparent from the context.