

boxamide, $C_5H_5N_3O$, an almost planar structure containing a six-membered ring (Rø & Sørum, 1972).

Experimental and results

Generation of signs and symbols with the \sum_2 formula was carried out by the program *SYMBOL* (Hjortås, 1969). It permits up to 4 symbols to be used in a starting set in addition to the origin-defining reflexions. Only reflexions with E 's greater than some preset value E_{start} are considered first. When no more new signs can be found the program lowers the E limit by 0.1 and repeats the search for triplets. *SYMBOL* thus works in cycles, gradually lowering the E limit until a preset lower limit E_{stop} is attained. Acceptance of new reflexions is based on the common probability measure $P(+)$ (Cochran & Woolfson, 1955) which is used in a rather cautious manner: when different symbolic indications of the sign for a reflexion are obtained, $P(+)$ is calculated by assuming that all indications other than the dominant one are wrong.

Significance of intersymbolic relations

In order to assess the possible value of relations obtained between symbols in $P\bar{I}$, signs and symbolic relations were developed from 14 different starting sets for structure *A* and five sets for structure *B*. Data sets of 2–300 E 's were used in both cases. For each starting set the number (N_c) of satisfied intersymbolic relations with correct signs being substituted for the symbols was compared with the maximum number of satisfied relations obtained for any sign combination, except the trivial one. The relative number N_c was found to range from 0 to 100% with an average of 50.9% for the 19 starting sets. The average for 14 sets with 4 or more intersymbolic relations was 44.1%. The results show that a relation like, for example, $a=b$ is not more likely to be true than $a=-b$.

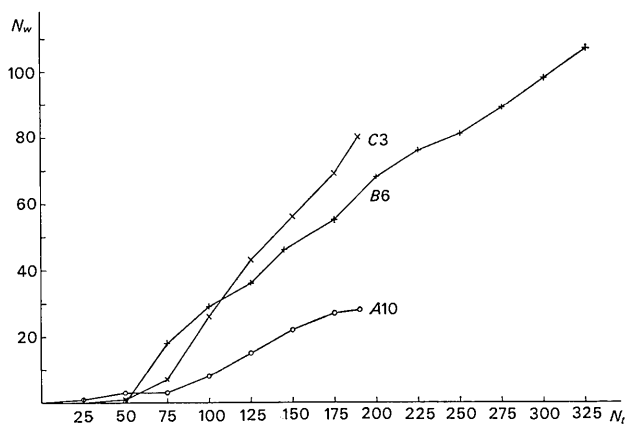


Fig. 1. Number of incorrect signs, N_w , versus total number of signs, N_t , as developed from the correct sign combination in starting sets *A10*, *B6* and *C3*.

E maps calculated with few terms

Structure factors that come out with a specific indication of sign, + or -, during the symbolic-addition procedure call for special attention. If many of these 'directly determined' signs are wrong the correct structure may never be found with this starting set. Both the stage at which errors show up and the kind of reflexions involved are, of course, decisive for the total number of incorrect signs. Sometimes the acceptance of a few or just one wrong sign may set off a fatal chain of errors in the sign propagation. The work on structure *B* provides an excellent example of this kind of trouble. One of the starting sets was *B6*: $111 = -$, $3\bar{7}0 = +$, $4\bar{5}4 = +$, $222 = +$ (from \sum_1) and $10\bar{1} = a$. With the correct sign + substituted for a there were no errors in the first 50 signs, among these 21 directly determined. The next reflexion to be accepted was 002 but the sign + (directly determined) is wrong for this particular choice of origin. Since 002 enters into a large number of \sum_2 relations the error spreads very quickly through the remainder of the data. The development of wrong signs for *B6* is illustrated in Fig. 1. This figure also shows the sign development for starting sets *A10* (structure *A*): $2, \bar{1}, 33 = -$, $0, 3, 20 = -$, $2, \bar{3}, 14 = +$, $0, 2, \bar{1}9 = a$, $4, \bar{2}, 30 = b$, $0, 1, \bar{1}8 = c$ and $2, \bar{3}, \bar{2}3 = d$, where $a=b=d = -$ and $c = +$ * and *C3* (structure *C*):

* Because of pseudosymmetry in structure *A*, only half the cell (halving of a axis) need be considered.

Table 2. Listing of strongest maxima in reduced and full E maps

Maxima corresponding to atomic positions have been marked according to the system used in Table 3.

Starting set <i>A10</i>				
Reduced E map: 50 terms				
x	y	z	Relative peak height	
0.564	0.920	0.339	410	C(13)
0.645	0.710	0.418	400	C(16)
0.427	0.733	0.311	390	C(12)
0.518	0.723	0.364	380	C(14)
0.673	0.913	0.393	370	C(15)
0.305	0.753	0.257	350	C(10)
0.443	0.930	0.286	340	C(11)
0.327	0.940	0.233	340	C(9)
0.766	0.907	0.448	280	C(17)
0.755	0.700	0.470	280	C(18)
0.205	0.957	0.181	250	C(7)
0.205	0.767	0.203	240	C(8)
Full E map: 191 terms				
0.418	0.727	0.310	870	C(12)
0.441	0.923	0.285	810	C(11)
0.302	0.743	0.255	730	C(10)
0.325	0.940	0.233	730	C(9)
0.541	0.913	0.339	670	C(13)
0.523	0.717	0.365	570	C(14)
0.636	0.707	0.418	470	C(16)
0.659	0.907	0.394	410	C(15)
0.202	0.757	0.203	410	C(8)
0.209	0.953	0.181	360	C(7)
0.759	0.893	0.448	310	C(17)
0.736	0.690	0.473	280	C(18)

Table 2 (cont.)

Starting set B6					
Reduced <i>E</i> map: 50 terms					
<i>x</i>	<i>y</i>	<i>z</i>	Relative peak height		
0.030	0.445	0.035	530	C(8)	
0.090	0.325	0.073	520	C(7)	
0.145	0.210	0.123	460	C(6)	
0.210	0.093	0.175	430	C(5)	
0.265	-0.028	0.225	390	C(4)	
0.325	-0.140	0.268	310	C(3)	
0.400	-0.258	0.318	250	C(2)	
0.300	-0.425	0.625	230		
0.650	-0.458	0.318	220	O	
0.580	-0.345	0.283	200	C(1)	
0.253	-0.340	0.583	200		
0.445	-0.375	0.375	190		
Full <i>E</i> map: 352 terms					
0.023	0.455	0.043	1880	C(8)	
0.135	0.333	0.035	1120	C(7)	
0.515	0.493	0.560	630		
0.215	0.230	0.108	610	C(6)	
0.325	-0.258	0.350	590		
0.215	-0.205	0.505	480		
0.400	-0.405	0.578	480		
0.450	-0.395	0.343	420		
0.210	0.083	0.158	400	C(5)	
0.225	-0.025	0.285	400	C(4)	
0.130	-0.085	0.463	390		
0.060	-0.148	0.578	390		
Starting set C3					
Reduced <i>E</i> map: 50 terms					
<i>x</i>	<i>y</i>	<i>z</i>	Relative peak height		
-0.020	-0.100	0.046	690		
-0.067	-0.286	0.131	630	C(7)	
-0.113	-0.482	0.219	620	C(2)	
-0.193	-0.857	0.396	580	C(5)	
-0.253	-0.200	0.069	570	N(8)	
0.080	-0.557	0.277	510	C(3)	
-0.153	-0.664	0.315	490		
0.153	-0.193	0.096	480	O(9)	
0.120	-0.371	0.196	480		
-0.337	-0.579	0.250	430	N(1)	
0.203	0.000	0.012	430		
0.033	-0.750	0.369	400	N(4)	
Full <i>E</i> map: 189 terms					
-0.020	-0.079	0.050	1450		
-0.087	-0.450	0.230	1020	C(2)	
-0.227	-0.171	0.071	840	N(8)	
-0.207	-0.886	0.408	780	C(5)	
0.173	-0.171	0.112	660	O(9)	
-0.613	-0.086	0.069	610		
0.073	-0.564	0.269	550	C(3)	
0.283	-0.786	0.431	540		
0.127	-0.364	0.200	510		
-0.653	-0.586	0.335	500		
0.320	-0.593	0.342	500		
-0.147	-0.629	0.300	500		

The features in Fig. 1 suggest that *E* maps calculated with a smaller number of terms might prove useful. We have calculated *E* maps from the 50 first signs developed by *SYMBOL* (reduced *E* map) for the three cases and compared them with *E* maps based on all accessible reflexions with $E \geq E_{\text{stop}}$ (full *E* map); *i.e.* for structure *A*: 191 *E*'s ($E_{\text{stop}} = 1.20$), *B*: 352 *E*'s ($E_{\text{stop}} = 1.10$) and *C*: 189 *E*'s ($E_{\text{stop}} = 1.10$). In Table 2 are given the coordinates of the highest peaks in pairs of such maps. Peaks corresponding to atomic positions in Table 3 have been marked appropriately.

Table 3. Relevant atomic coordinates in structures *A*, *B* and *C*

Coordinates of structure *A* correspond to the regular part of the hydrocarbon chain and are averaged values for two crystallographically independent molecules at an *R* index of 13.8%. Coordinates of *B* are as given by Hjortås (1972). Coordinates of *C* are the centrosymmetrical equivalents of those given by Rø & Sørum (1972).

<i>A</i>			
	<i>x</i>	<i>y</i>	<i>z</i>
C(7)	0.224	0.966	0.178
C(8)	0.203	0.748	0.203
C(9)	0.335	0.952	0.232
C(10)	0.309	0.743	0.256
C(11)	0.436	0.937	0.285
C(12)	0.418	0.725	0.310
C(13)	0.543	0.928	0.339
C(14)	0.522	0.716	0.363
C(15)	0.645	0.911	0.392
C(16)	0.631	0.715	0.416
C(17)	0.746	0.906	0.446
C(18)	0.732	0.700	0.471
<i>B</i>			
	<i>x</i>	<i>y</i>	<i>z</i>
C(1)	0.5763	-0.3527	0.2630
C(2)	0.4320	-0.2476	0.3069
C(3)	0.3671	-0.1443	0.2251
C(4)	0.2400	-0.0223	0.2443
C(5)	0.2033	0.0826	0.1563
C(6)	0.0920	0.2160	0.1656
C(7)	0.0835	0.3123	0.0653
C(8)	-0.0083	0.4535	0.0524
C(9)	0.3714	-0.2609	0.4368
C(10)	-0.0022	0.2439	0.2845
O	0.6683	-0.4438	0.3244
<i>C</i>			
	<i>x</i>	<i>y</i>	<i>z</i>
N(1)	-0.3414	-0.5757	0.2513
C(2)	-0.1116	-0.4882	0.2221
C(3)	0.0742	-0.5810	0.2824
N(4)	0.0375	-0.7570	0.3694
C(5)	-0.1910	-0.8460	0.3962
C(6)	-0.3760	-0.7535	0.3360
C(7)	-0.0638	-0.2885	0.1262
N(8)	-0.2539	-0.2153	0.0719
O(9)	0.1510	-0.1990	0.1006

$2\bar{3}6 = -$, $3\bar{2}5 = -$, $5\bar{2}2 = +$, $5\bar{1}0 = a$, $6\bar{3}4 = b$ and $3\bar{1}\bar{1} = c$, where $a = b = c = +$. In all three cases the number of wrong signs is reasonably low for the first 50–60 *E*'s. Thereafter an increasing fraction of wrong signs develops, the effect being most pronounced in case *C* and smallest in case *A*.

For structure *A* both full and reduced *E* maps give the regular part of the chain molecule. Maxima in the full *E* map are maybe slightly closer to refined values, but the small differences between the two maps have no practical consequences for the structure solution.

Structure *B* is bent in the main molecular plane and also twisted out of this plane (Hjortås, 1972). The irregularities of the chain are manifested by very diffuse maxima both in the Patterson and the *E* maps. All 9 atoms (C and O) in one half of the centrosymmetric chain [branching methyl groups not included, see Fig. 2(a)] are found among the 10 highest peaks in the reduced *E* map. Estimated positions of the atomic peaks are at an average distance of 0.29 Å from corresponding refined positions. A least-squares refinement (3 cycles) of coordinates for the 9 peaks forming a zigzag array followed by calculation of a difference map showed very clearly the missing atoms C(9) and C(10).

In the full *E* map 5 of the 10 strongest peaks are near true atomic sites. The average deviation from refined atomic coordinates is 0.40 Å, atom C(6) being considerably displaced, *ca.* 0.90 Å, from its refined position. A difference map was calculated after some refinement of the coordinates for the 5 peaks conforming to a chain. This map shows, in addition to many spurious peaks, two maxima of medium height somewhat shifted from the final positions of C(3) and C(2). Final positions of the 4 missing atoms are only slightly above background or on negative area. In the original analysis all trials to solve structure *B* from the full *E* map based on starting set B6 were unsuccessful.

Case C: in the reduced *E* map 8 of the strongest maxima correspond to true atomic positions. The only missing heavier atom, C(6), which is needed to close the hexagon of the pyrazine ring [see Fig. 2(b) for atomic numbering] is found as a peak of lower relative height 240.

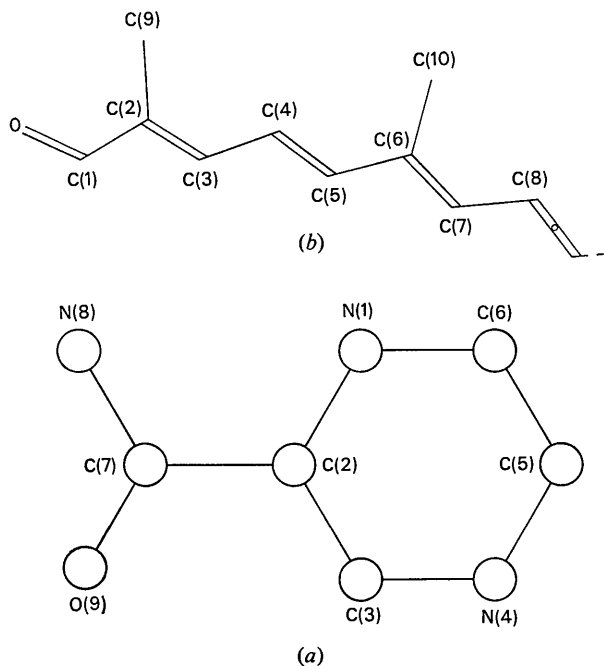


Fig. 2. (a). Atomic numbering in structure *B*. (b) Atomic numbering in structure *C*.

Among the 12 strongest maxima in the full *E* map only 5 correspond to true atomic positions. This map is inferior to the reduced map for other reasons also. The true hexagonal pattern of atoms is not easily recognized here since both N(4) and C(6) fall in minima. [N(1) and C(7) are on low maxima.] However, hexagons in various wrong positions are readily established by using other peaks, and in the work of Rø & Sørum (1972) the structure could not be found from the full *E* map based on starting set C3. It is also seen from Table 2 that of the true atomic peaks found in both maps, *i.e.* C(2), C(3), C(5), N(8), O(9), those in the reduced map are closer to the refined coordinates.

Among the false peaks, there are two common to both maps: the strongest one and another at approximately $-0.150, -0.660, 0.315$. Both peaks can be identified as false features when the true position of the hexagon has been established, as it is in the reduced *E* map.

Aids for selecting most probable models

In many instances the correct structure model is not easily recognized just by comparing different *E* maps. This is very often found with the kind of highly regular structures considered here. It is also a matter of experience that refinements of structures like hydrocarbon chains often converge very slowly, even when most coordinates are fairly well known at the outset. Therefore, various aids for the exclusion of less probable models are valuable. Large variations in peak heights and the presence of ghost maxima in the *E* map are known to be indications of wrong models. The correct model of *e.g.* structure *A* was found in maps with a fairly even distribution of peak heights and there were only two rows of maxima corresponding to a single zigzag pattern in the hydrocarbon chain. Spurious maxima beyond true chain ends were weak. They are caused mainly by the pronounced translational symmetry within the molecule as noted by *e.g.* Bürgi & Dunitz (1971). These authors suggest extension of the structure-factor set to include smaller *E* values as a means of reducing the level of such detail. In the $P\bar{1}$ case, however, we would like to emphasize the dangers involved in such a procedure. A better way of suppressing 'translational' peaks appears to be the commonly used method of normalizing the structure factors in groups, *e.g.* according to parity.

Case *A* also provides an example of the possible use of structural knowledge as an aid to the application of direct methods. There is a sequence of $00l$ reflexions in the data with $34 \leq l \leq 44$ corresponding to *d* spacings in the range 1.32–1.02 Å. The molecular axis in *A* is approximately parallel to [201], and the distribution of signs in the $00l$ sequence must conform to the repeating pattern of peaks in a hydrocarbon chain. Table 4 gives the $00l$ signs for various sign combinations in starting set A1: $2\bar{1}\bar{5} = -, 2\bar{1}, 3\bar{2} = +, 20\bar{3} = +, 4, \bar{2}, \bar{1}\bar{0} = +$ (from Σ_1), $2, \bar{2}, 3\bar{3} = a, 2, \bar{3}, \bar{2}\bar{3} = b$ and $0, \bar{3}, \bar{1}\bar{8} = c$. Combinations VI and VII both give the sign $-$ for

the whole set of 00 l reflexions and can be ruled out. I and IV give alternating signs for consecutive reflexions in the set and were considered unlikely; moreover, IV corresponds to the trivial solution. Only V and VIII of the four combinations left gave models with a satisfactory distribution of peak heights in the E maps and no ghost peaks. The model from VIII proved to be correct.

Table 4. Signs of some 00 l reflexions for various sign combinations in starting set A1

	I	II	III	IV	V	VI	VII	VIII
	$a=+$	$a=+$	$a=+$	$a=+$	$a=-$	$a=-$	$a=-$	$a=-$
	$b=+$	$b=+$	$b=-$	$b=-$	$b=+$	$b=+$	$b=-$	$b=-$
00 l	$c=+$	$c=-$	$c=+$	$c=-$	$c=+$	$c=-$	$c=+$	$c=-$
$l=34$	+	+	+	+	-	-	-	-
35	-	+	+	-	+	-	-	+
36	+	-	-	+	+	-	-	+
37	-	-	-	-	-	-	-	-
38	+	-	-	+	+	-	-	+
39	-	+	+	-	+	-	-	+
40	+	+	+	+	-	-	-	-
41	-	-	-	-	-	-	-	-
42	+	-	-	+	+	-	-	+
43	-	+	+	-	+	-	-	+
44	+	+	+	+	-	-	-	-

Conclusions

Intersymbolic relations are of no significance in selecting the correct sign combination in a chosen starting set in $P\bar{1}$.

When incorrect signs have been included in the data, they may cause a very rapid propagation of more errors

during the symbolic-addition procedure. It may therefore prove advantageous to break off the Σ_2 process at an early stage and calculate E maps with a small number of terms. For three cases examined, E maps were calculated with the 50 structure factors signed in the first stages by Σ_2 . They were found to contain as much or even significantly more correct information than maps based on all E 's above some arbitrary limit, e.g. 1.2. It is implied that if incorrect signs enter into the data at a very early stage, even reduced E maps may contain too many erroneous features.

Discrimination between probable and less probable sign models may be aided by the use of structural information. A rather crude application of structural knowledge is shown as an example from the work on a chain structure. More refined methods based on these principles could certainly be of great value, in particular with structures giving heavy overlap in Patterson space.

References

- BÜRGI, H. B. & DUNITZ, J. D. (1971). *Acta Cryst.* **A27**, 117-119.
 COCHRAN, W. & WOOLFSON, M. M. (1955). *Acta Cryst.* **8**, 1-12.
 HAUPTMAN, H. & KARLE, J. (1953). *Solution of the Phase Problem. I. The Centrosymmetric Crystal*, A.C.A. Monograph No. 3. Pittsburgh: Polycrystal Book Service.
 HJORTÅS, J. (1969). Tek. Rapp. 07-RII-69. Inst. for Røntgenteknikk, NTH, Trondheim, Norway.
 HJORTÅS, J. (1972). *Acta Cryst.* **B28**, 2252-2259.
 MO, F. (1971). Unpublished results.
 RØ, G. & SØRUM, H. (1972). *Acta Cryst.* **B28**, 1677-1684.

Acta Cryst. (1973). **A29**, 362

A Complete Catalogue of Polyhedra with Eight or Fewer Vertices

BY DOYLE BRITTON* AND J. D. DUNITZ

Laboratory for Organic Chemistry, Federal Institute of Technology, 8006 Zurich, Switzerland

(Received 29 January 1973; accepted 30 January 1973)

All non-isomorphic convex polyhedra with 4, 5, 6, 7 and 8 vertices are listed. The relationships within each class are described.

In the course of an attempt to describe in a systematic way the coordination of eight ligand atoms around a central atom with no symmetry restrictions, we encountered the problem of enumerating all possible non-isomorphic convex polyhedra with eight vertices. According to Alexandrow (1958) the number $N(n)$ of

polyhedra with n vertices is: $N(4)=1$, $N(5)=2$, $N(6)=7$, $N(7)=34$, $N(8)=257$, but we were unable to find any publication in which these polyhedra are described. Grace (1965) has determined by computer search all polyhedra with up to eleven faces with the restriction that only three edges meet at each vertex. The duals of these polyhedra are the polyhedra with up to eleven vertices with the restriction that all faces are triangular; these, however, are only a small fraction of the

* Permanent address: Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455, U.S.A.