

Technical Note

R-SAPI-85

A COMPUTER PROGRAM FOR AUTOMATIC SOLUTION OF CRYSTAL STRUCTURES FROM X-RAY DIFFRACTION DATA

FAN Hai-fu

Research Group on Crystal Structure Analysis, Institute of Physics,
Academia Sinica, Beijing, China

R-SAPI-85 is a Rigaku version of SAPI-85 for use with a microVAX computer. The name SAPI is an abbreviation of "Structure Analysis Programs with Intelligent control", it may also so be read inversely as "Institute of Physics, Academia Sinica".

SAPI-85 is based on MULTAN-80 [1] and differs from which by the following:

1. The program can automatically handle diffraction data of structures having pseudo-translational symmetry and lead directly to a correct solution in favourable cases.

2. The program can recognize pseudo-centrosymmetric solutions when dealing with non-centrosymmetric structures. In addition, the program can break the enantiomorphous ambiguity given a group of atoms, each of which comes from either of the enantiomorphs.

3. Instead of the phase permutation of MULTAN-80, random starting phases are assigned to all but the origin and enantiomorph fixing reflections. A speed-up scheme is optional to use the early figure of merit with floatable cutoff value in order to save computing time.

4. The program possesses the ability of intelligent control on the path of phase development according to the complexity of the structure and the early stage result during the phase development.

5. The programs EXFFT and SEARCH of MULTAN-80 have been extended to include the calculation and peak search for Patterson and minimum functions. This enables the effective combination of Patterson and direct methods.

6. The program can output density maps, sections or composite sections, as "half-tone graphs" on a conventional lineprinter. The size and contrast of the map output can be controlled automatically or manually.

7. The program for the interpretation of space group symbols written by Burzlaff and Houtas [2] has been modified and included.

For the theoretical background on the treatment

of pseudo symmetry, the reader is referred to Fan Hai-fu & Zheng Qi-tai [3], Fan Hai-fu [4] and the references therein. For the detail of the random starting procedure used in this program system, see Yao Jia-xing [5].

The complete system of SAPI-85 is made up of the following:

"PREPAR"— perform the preliminary processing of F(obs) data for later use.

"PHASE"— solve the phase problem using direct methods.

"EXFFT"— fast Fourier transform program.

"SEARCH"— interpret the output of "EXFFT".

"MAP"— output density maps.

Included in the program "PREPAR", there is a subroutine "AUTOGP". It scans the E values and searches for pseudo-systematic extinction. Once this is found, reflections will be grouped automatically according to the pseudo-systematic-extinction rule and the program will print out message like this:

```
...PSEUDO-TRANSLATIONAL SYMMETRY HAS  
      BEEN FOUND BY THE PROGRAM...  
...NORMALIZATION IS TO BE RESTARTED...  
      REFLECTIONS WILL BE RESCALLED  
      ACCORDING TO 6 INDEX GROUPS  
      GROUP 1      H=2N  
                   K=3N  
      GROUP 2      H=2N  
                   K=3N+1  
      GROUP 3      H=2N  
                   K=3N+2  
      GROUP 4      H=2N+1  
                   K=3N  
      GROUP 5      H=2N+1  
                   K=3N+1  
      GROUP 6      H=2N+1  
                   K=3N+2
```

After that, Wilson statistics will be performed independently for each reflection group. This is particularly important for the automatic determination of

structures having pseudo-translational symmetry, such as superstructures.

The program "PHASE" uses the RANTAN procedure [5] for phase development. In addition, it possesses a number of special features:

1. With default controlling parameters, the program will run as follows:

(1) Generate a minimum number of phase sets. This number equals 10 or the number of independent atoms, whichever is the larger.

(2) After finished phase development for these sets, if none of them has a PSI(ZERO) figure of merit less than 1.6 and a RESIDUAL figure of merit smaller than 25, the program will continue to generate new phase sets up to a total number of 100 or three times the minimum number of phase sets determined according to (1), whichever is the smaller.

(3) The ordinary weighted tangent refinement will be used at first. After finished phase development for the first 10 sets, if any one of the following conditions is fulfilled, the weighting scheme will change to 'SWTR' [6] and the phase development process will be restarted.

- The averaged absolute figure of merit greater than 1.1 and the averaged deviation among the absolute figures of merit less than 0.1;
- The averaged absolute figure of merit greater than 1.25;
- The averaged deviation among the absolute figures of merit less than 0.02.

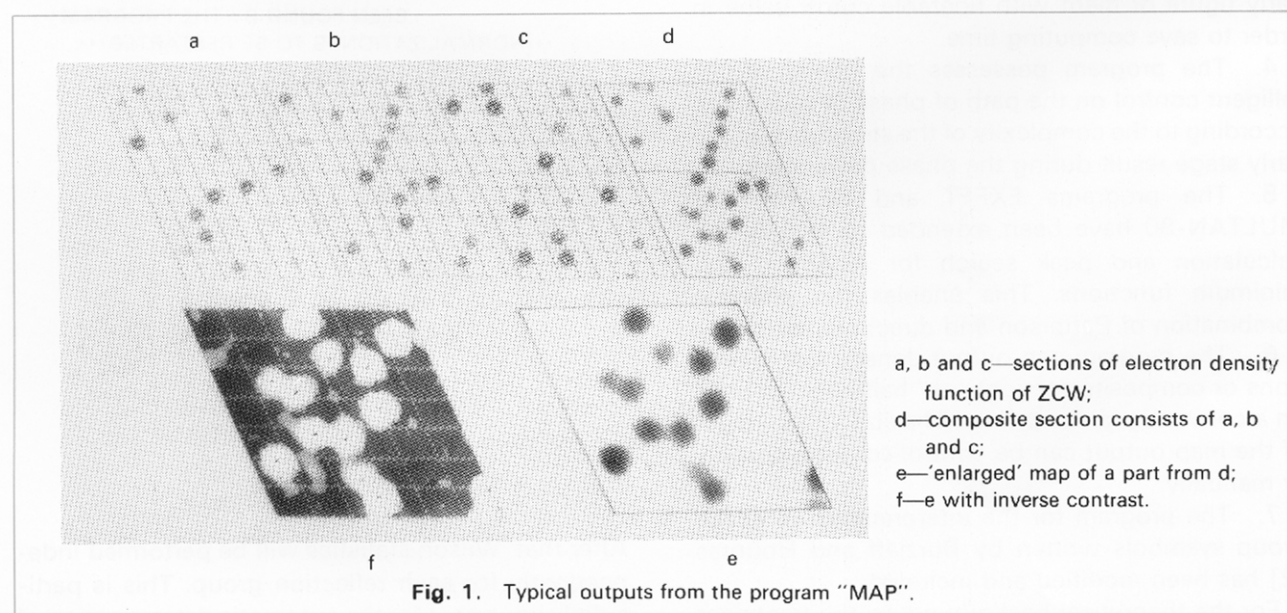
2. "PHASE" can automatically handle a structure having pseudo-translational symmetry like this:

All triplets composed of three 'weak' reflections will be eliminated in the SIGMA2 list and the phases development will be performed in two stages. During the first stage only phases of the 'strong' reflections are developed, while in the second stage the phases of 'weak' reflections are derived by the phase relationships of the type 'weak-strong-weak'. This procedure can lead directly to a complete solution of a superstructure in favourable cases without the need of knowing the averaged (subcell-) structure.

3. In dealing with noncentrosymmetric structures "PHASE" will check each set of the final phases to see whether the averaged deviation of the three-phase invariants from 0 or π is smaller than 15 degrees. If so, the corresponding set will be recognized as a pseudo-centrosymmetric solution and warning message will be printed out.

4. If a structure model containing both enantiomorphs is input to "PREPAR", "PHASE" can break the ambiguity using the method proposed by Fan Hai-fu [4]. The result is an E-map with different peak heights for the two enantiomorphs.

Apart from E-map and Fourier-map calculations, the program "EXFFT" can calculate also Patterson maps with $E \cdot F$ or F^2 as coefficients. It can also perform superpositions of two Patterson maps (MPP) or superposition of an E-map and a Patterson map (MEP). MPP is particularly useful for extending a partial structure to the complete one, especially when the partial structure can not dominate the phases to give a good Fourier map. MEP can be used to eliminate the spurious peaks on an E-map making it more interpretable.



The program "MAP" prints density maps in the form of 'half-tone graphs' with a conventional lineprinter. The program provides fully automatic control on the map size and contrast. However user's intervention on any of the controlling parameters is possible. Figure 1 shows some typical outputs of "MAP"

Examples:

(1) Cortison $C_{21}H_{28}O_5$

SG. $P2_12_12_1$

$a=10.040$, $b=23.649$, $c=7.784 \text{ \AA}$

The structure can be solved by default controlling parameters. The result is shown in Fig. 2.

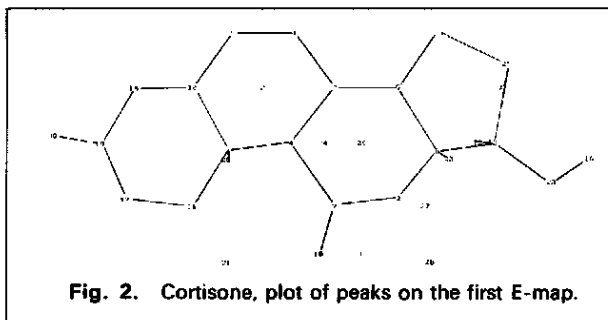


Fig. 2. Cortison, plot of peaks on the first E-map.

(2) OCP8 $C_{30}H_{22}O_9$

SG. C2

$a=20.622$, $b=7.715$, $c=17.94 \text{ \AA}$, $\beta=126.13^\circ$

This structure was originally solved by SAPI-85. The result is shown in Fig. 3.

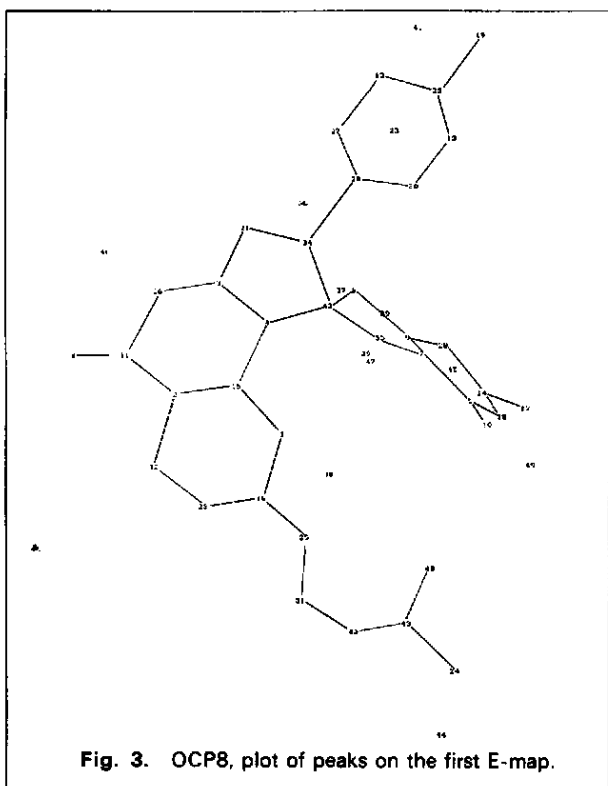


Fig. 3. OCP8, plot of peaks on the first E-map.

(3) Freielebenite $PbAgSbS_3$

SG. $P2_1/a$

$a=7.518$, $b=12.809$, $c=5.940 \text{ \AA}$, $\beta=92.25^\circ$

This is a superstructure with pseudo-translation vectors $t_1=a/2$ and $t_2=b/3$. The structure was originally solved by Ito & Nowacki [7]. It can be solved automatically by SAPI-85. The result is given in Table 1.

(4) HWHK $Ba_2Ce_2(CO_3)_4$

S.G. R 3

$a=5.070$, $c=38.408 \text{ \AA}$, $\gamma=120^\circ$

This is a superstructure with pseudo-translation vector $t=c/4$. It can also be solved automatically by SAPI-85. The result is given in Table 2.

(5) ZCW $C_{34}H_{47}O_{11}N \cdot HI$

SG. $P2_1$

$a=12.58$, $b=14.38$, $c=11.00 \text{ \AA}$, $\beta=114.6^\circ$

This is a noncentrosymmetric structure with heavy atoms in centrosymmetric arrangement. The structure was proved unsolvable by RANTAN-81. However, it can be solved by SAPI-85 in the following steps:

a. Patterson superposition—this results in a centrosymmetric map which includes both enantiomorphs.

b. Resolving the enantiomorphous ambiguity—this results in an E-map with different peak heights for the two enantiomorphs.

c. Superposition of the E-map and Patterson map—this cleans up the E-map by eliminating the spurious peaks not belonging to any one of the enantiomorphs. A peak list of the superposition result is given in Table 3, from which a fragment of the molecule can be found as shown in Fig. 4. The complete structure shown in Fig. 5 can then easily be obtained by Fourier recycling.

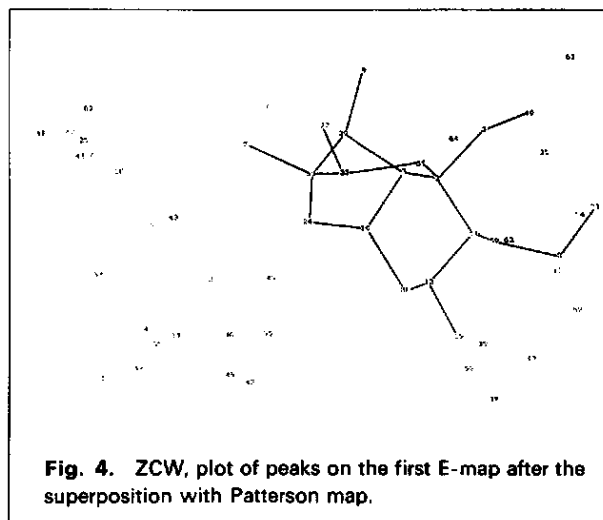
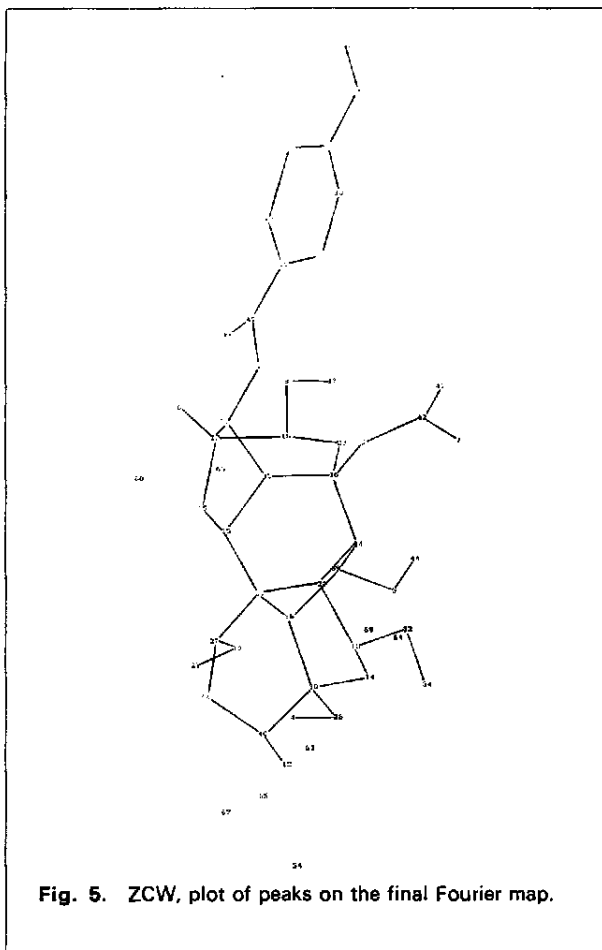


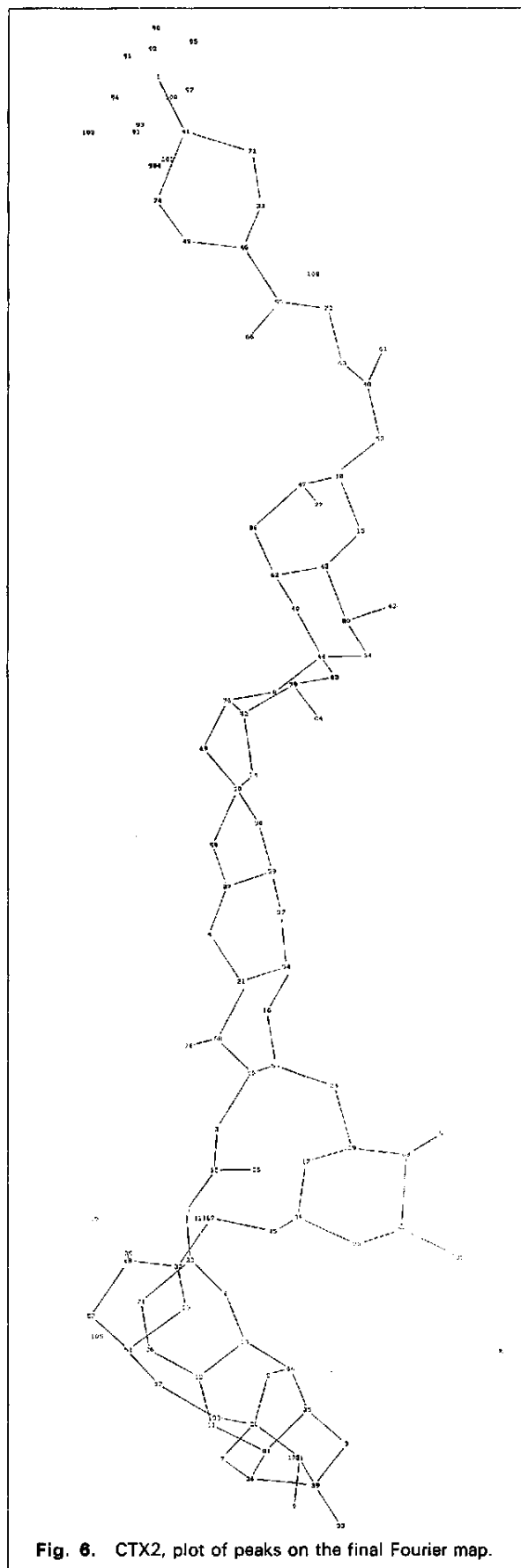
Fig. 4. ZCW, plot of peaks on the first E-map after the superposition with Patterson map.



(6) CTX2 $C_{67}H_{88}O_{22}Br$
 SG. $P2_1$

$a=44.654$, $b=9.264$, $c=9.494 \text{ \AA}$, $\beta=93.72^\circ$
 More than 20 independent atoms can be found on the first E-map. The complete structure can then be obtained by Fourier recycling. The result is shown in Fig. 6.

The author should like to express his gratitude to Professor M. M. Woolfson for helpful discussions and encouragement with this project.



**Table 1. Automatic Determination of Superstructure Freielebenite
Peak List of the Best E-map.**

*****						I	THIS PART IS NOT THE COMPUTER	
***						I	OUTPUT	
FREIELEBENITE						I		
P21/A						I		
***						I		
PEAK	HEIGHT	X	Y	Z	CLUSTER NUMBER	I	ATOM	
1	3573	0.3833	0.2517	0.2263	0	I	AG	
2	3451	0.6382	0.0849	0.7429	1	I	PB	
3	2959	0.1305	0.0839	0.7287	2	I	SB(AS)	
4	1025	0.3601	0.2205	0.6569	1	I	S1	
5	940	0.6487	0.0519	0.2686	0	I	S3	
6	826	0.1268	0.1215	0.1409	0	I	S2	
7	710	0.5437	0.1544	0.6586	1	I		
8	641	0.0400	0.1551	0.6599	2	I		
TOTAL NUMBER OF PEAKS IN EACH CLUSTER						3	2	ALL THE ATOMS HAVE THE CORRECT PARAMETER SHIFTS (SEE ITO, T., Z. KRIST., 137(1973)399)
*****						I		

**Table 2. Automatic Determination of Superstructure HWHK.
Peak List of the Best E-map.**

*****						I	THIS PART IS NOT THE COMPUTER
***						I	OUTPUT
HWHK						I	
BACEE(CO3)2						I	
R3 Z=6						I	
I(PSEUDO)=C/4						I	
***						I	
PEAK	HEIGHT	X	Y	Z	CLUSTER NUMBER	I	ATOM
1	4808	1.0050	0.0109	0.6669	1	I	CE2
2	4662	0.6717	0.3445	0.7497	1	I	BA1
3	4434	1.3384	-0.3232	0.8251	1	I	CE1
4	4392	1.0059	0.0114	0.9095	1	I	BA2
5	534	1.3373	-0.3021	0.6859	1	I	C2
6	525	1.3373	-0.3144	0.7639	1	I	F1
7	516	0.9979	0.0030	0.8059	1	I	C3
8	514	0.5419	-0.9423	0.8763	1	I	O4
9	513	1.0048	0.0188	0.7689	1	I	C2A
10	493	1.0036	-0.0180	0.7297	1	I	E2
11	476	1.0045	0.0094	0.8478	1	I	F1A
12	465	0.6698	-0.6654	0.8160	1	I	E2A
13	460	0.6598	-0.6589	0.8930	1	I	C4
14	457	1.4005	0.2002	0.8176	1	I	O1
15	435	1.0727	-0.4657	0.6982	1	I	O2
16	350	1.5402	0.0600	0.6253	1	I	O1A
17	344	1.6727	0.3447	0.6099	1	I	C1
18	332	1.3295	-0.3283	0.7233	1	I	C4A
19	324	0.6716	-0.6740	0.8514	1	I	C2A
20	318	0.8784	-0.2552	0.7998	1	I	O3
21	317	1.0614	-0.1274	0.9342	1	I	DIFFRACTION RIPPLE
22	313	1.6889	0.3302	0.6748	1	I	E2A
23	312	1.3402	-0.3446	0.6496	1	I	F1A
24	300	1.0666	-0.1301	0.6859	1	I	DIFFRACTION RIPPLE
25	295	1.6683	0.3334	0.5684	1	I	F1A
26	291	1.0171	-0.4195	0.8363	1	I	DIFFRACTION RIPPLE
27	284	1.1884	0.0641	0.8927	1	I	DIFFRACTION RIPPLE
28	280	0.7462	-0.1281	0.7838	1	I	O2A
29	279	1.0832	-0.2399	0.6678	1	I	DIFFRACTION RIPPLE
30	275	1.2875	0.0618	0.6575	1	I	DIFFRACTION RIPPLE
31	261	1.1879	-0.2113	0.7088	1	I	O3A
32	255	0.3959	-0.8050	0.8733	1	I	O1A
33	252	0.7503	0.0998	0.7516	1	I	DIFFRACTION RIPPLE
34	250	1.0898	0.3673	0.8322	1	I	
35	246	1.6680	0.3345	0.6490	1	I	
36	246	1.1374	-0.3269	0.8057	1	I	A TRANSLATION IMAGE OF AN ATOM
37	245	0.9545	-0.6030	0.7410	1	I	
38	232	0.8417	-0.4796	0.7684	1	I	ALL HEAVY ATOMS HAVE THE CORRECT PARAMETER SHIFTS ALONG C AXIS
39	231	0.6599	-0.3234	0.7852	1	I	ALL LIGHT ATOMS HAVE PEAK HEIGHTS LARGER THAN THEIR TRANSLATION IMAGES (SEE FAN, YAO, MAIN & WOODLSON (1983) ACTA CRYST. A39, 566-569)
40	229	0.6562	0.1184	0.7235	1	I	
TOTAL NUMBER OF PEAKS IN EACH CLUSTER						40	
*****						I	

Table 3. Resolving the Enantiomorphous Ambiguity of ZCW.
Peak List of the Best E-map after the Superposition
with Patterson Map.

PEAK	HEIGHT	X	Y	Z	CLUSTER NUMBER	ATOM
1	3819	0.9071	0.2534	0.1933	1	I
2	1271	0.3085	0.4290	0.1035	2	I
3	1177	0.4247	0.3760	0.2023	2	I
4	1162	-0.3047	0.5745	-0.1062	2	I ENANTIOMORPH OF PEAK 2 (ALSO 08)
5	935	0.0638	0.5380	0.2613	2	I
6	929	0.9624	0.3666	0.5014	0	I
7	868	0.2401	0.3722	0.1349	2	I
8	849	0.3640	0.2569	-0.1142	2	I
9	800	0.3092	0.4418	0.3710	2	I
10	780	0.1074	0.3704	0.0208	2	I
11	775	0.8464	0.1871	0.1272	1	I
12	769	0.1819	0.4444	-0.1249	2	I
13	756	0.4507	0.5352	-0.1466	2	I X
14	686	-0.3818	0.5219	-0.2000	2	I
15	669	0.1764	0.4610	-0.2556	2	I
16	637	-0.2444	0.4479	0.2948	2	I X
17	635	-0.2401	0.6294	-0.1442	2	I
18	631	0.9674	0.4689	0.7351	0	I ENANTIOMORPH OF PEAK 7
19	597	0.3357	0.3101	0.4726	0	I ENANTIOMORPH OF PEAK 5 (ALSO 06)
20	579	0.2024	0.4074	0.2500	2	I
21	559	-0.4195	0.6242	-0.2050	2	I
22	550	-0.1059	0.6296	-0.0190	2	I
23	545	0.2188	0.5716	0.1633	2	I
24	543	0.0473	0.4630	0.0532	2	I
25	540	0.5375	0.4630	0.1344	2	I ENANTIOMORPH OF PEAK 13
26	540	-0.1852	0.5556	0.1222	2	I ENANTIOMORPH OF PEAK 12
27	536	0.2605	0.6383	0.2735	2	I
28	534	-0.1740	0.5513	0.2564	2	I ENANTIOMORPH OF PEAK 15
29	508	-0.1748	0.6296	0.3250	2	I
30	504	0.9730	0.1435	0.5025	0	I ENANTIOMORPH OF PEAK 6
31	476	0.5111	0.0185	0.4250	0	I
32	470	0.3355	0.1894	0.4735	0	I
33	458	0.4667	0.2198	-0.0250	2	I
34	445	0.4667	0.2943	-0.0250	2	I
35	437	0.5077	0.3889	0.7000	0	I
36	437	-0.2189	0.4364	-0.1535	2	I
37	430	0.2190	0.5926	-0.4000	2	I
38	428	0.1111	0.3704	-0.1250	2	I
39	423	0.2444	0.5511	-0.2815	2	I
40	419	0.5365	0.4074	0.2250	2	I
41	416	0.9135	0.1852	0.7000	2	I X
42	415	-0.2590	0.3655	-0.2698	2	I
43	403	0.3556	0.6667	-0.3250	2	I X
44	402	-0.1333	0.7963	0.3018	2	I
45	400	-0.2028	0.5926	-0.2500	2	I
46	399	-0.0455	0.5370	-0.0512	2	I
47	396	-0.1591	0.7037	0.3534	2	I
48	390	-0.1733	0.7963	0.3698	2	I X
49	379	-0.1111	0.6296	0.1250	2	I
50	378	0.2843	0.2963	-0.0582	2	I
51	349	-0.3156	0.5583	-0.3731	2	I
52	343	0.5086	0.1111	0.7000	0	I
53	338	0.1275	0.3704	-0.3289	2	I
54	338	0.9303	0.3148	0.7000	2	I
55	334	-0.1333	0.5000	-0.1750	2	I X
56	324	0.3111	0.4074	-0.0250	2	I
57	322	0.4648	0.5926	-0.2250	2	I
58	316	0.1333	0.5000	0.1750	2	I
59	303	0.3333	0.5185	0.1500	2	I
60	289	0.4889	-0.0185	0.5750	4	I
61	287	0.6444	0.3333	0.3250	2	I
62	275	0.2849	0.2037	-0.0571	2	I
63	275	-0.2029	0.4074	0.4000	2	I
64	255	0.4000	0.4630	0.2000	2	I
65	253	-0.3333	0.4815	-0.1500	2	I
66	212	0.8444	0.3148	0.1497	1	I
67	193	-0.3111	0.5926	0.0250	2	I
TOTAL NUMBER OF PEAKS IN EACH CLUSTER						3 53 2 2
						I * ENANTIOMORPH OF AN ATOM WITH PEAK HEIGHT GREATER THAN THAT OF THE TRUE ATOM

References

[1] P. Main, S. J. Fiske, S. E. Hull, L. Lessinger, G. Germain, J. P. Declercq & M. M. Woolfson (1980), MULTAN-80: A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data. Univ. of York, England and Louvain, Belgium.

[2] H. Buzlaff & A. Hountas (1982). J. Appl. Cryst. **15**, 464-467.

[3] Fan Hai-fu & Zheng Qi-tai (1982). Acta Phys. Sin. **31**, 191-198.

[4] Fan Hai-fu (1984). Rigaku Journal Vol. 1 No. 2, 15-21.

[5] Yao Jia-xing (1981). Acta Cryst. **A37**, 642-644.

[6] S. E. Hull & M. J. Irwin (1978). Acta Cryst. **A34**, 863-870.

[7] T. Ito & W. Nowacki (1974). Z. Krist. **139**, 85-102.