X-ray and neutron diffraction studies of the crystal and molecular structure of the predominant monocarboxylic acid obtained by mild acid hydrolysis of cyanocobalamin. Part I. X-ray diffraction studies of air-dried crystals

C E NOCKOLDS† and S RAMASESHAN*

†Electron Microscope Unit, The University of Sydney, N.S.W. 2006, Australia
Indian Institute of Science, Bangalore 560 012, India

Abstract. The crystal structure of air-dried crystals of the O-monocarboxylic acid of cyanocobalamin, E2, has been solved by x-ray analysis. The lattice constants are $a = 14.51$ A, $b = 17.09$ A, $c = 16.35$ A, $\beta = 103°$; space group P2$_1$. Phase angles were derived from observations on anomalous dispersion effects for 2567 of the measured 2875 pairs of reflections, hkl and $\overline{hkl}$, and used in the solution of the structure. The intensities of reflexions were eye estimated from Weissenberg photographs. The molecular arrangement is very different from that of other cyanocobalamins; the corrin rings of neighbouring molecules are at *ca* 90° to one another. It was not possible to recognise the one acid group among the amides with certainty.

Keywords. Vitamin B$_{12}$; air dried crystals of a-monoacid of B$_{12}$; anomalous dispersion method; crystal structure.

1. Introduction

Crystals of a-mono acid obtained by mild acid hydrolysis of vitamin B$_{12}$ were first given us in 1954 by Dr Lester Smith. X-ray photographs were taken of a chip from a large dry crystal by J Kamper who showed that the crystals were monoclinic. Serious structure analysis was not however contemplated until the acid appeared as a natural product in Dr Bernhauer's laboratory, in the fermentation liquors of *Propionibacterium shermanii*.

Both Dr Lester Smith and Prof. Bernhauer attempted to prepare derivatives of the acid for x-ray analysis, preferably substituted with a halogen atom. Only one that gave crystals was obtained, the ethylamide prepared by Dr Lester Smith. This crystallised from 90% acetone in fine needles which were poorly diffracting. Preliminary x-ray data from oscillation photographs showed that the crystals were orthorhombic, $a = 16.0$, $b = 21.4$, $c = 24.7$ A, space group P2$_1$2$_1$2$_1$.

The a-monoacid itself was given us in the form of very good, chunky, red monoclinic crystals. An experiment on recrystallisation of a sample from water precipitated by acetone produced a second modification in small orthorhombic needles with $a = 16.4$ A, $b = 22.0$ A, $c = 23.6$ A, space group P2$_1$2$_1$2$_1$, similar to the ethylamide and to cyanocobalamin. It was decided to study the monoclinic crystals, and the following data were determined.

---

* To whom all correspondence should be addressed.

2. Experimental

2.1 *Crystal data*

Formula probably $C_{63}H_{87}O_{15}N_{13}PCo \cdot 13H_2O$, monoclinic, $a = 14.51$ Å, $b = 17.09$ Å, $c = 16.35$ Å, $\beta = 103°$. Density 1.335 g/cm³, measured by flotation; calc. 1.335, $\mu = 26.2$ cm⁻¹. The cell dimensions were measured on oscillation and Weissenberg photographs. They were a little variable, probably owing to differences in crystal drying. Those recorded were measured on the crystal from which the main set of data was obtained.

Weissenberg photographs of crystals rotating about the a axis, and using CuKα radiation showed marked differences in the intensities of the Bijvoet pairs, $I(hkl)$ and $I(\overline{hkl})$: it was accordingly decided to use anomalous dispersion techniques to solve the structure (Peerdemann and Bijvoet 1956; Ramachandran and Raman 1956). Formula I illustrates the state of our knowledge of the structure of the acid at the outset of the investigation.

2.2 *Data collection*

Equi-inclination Weissenberg photographs were taken of crystals rotating about the a and c axes. The layers $0kl$ to $9kl$ were obtained from a fragment of a crystal of spherical shape with diameter approximately 0.7 mm. Care was taken to ensure that both the Bijvoet pairs appeared on the same side of the film to avoid errors from elongation and contraction of the spots. The multiple film technique was used and the intensities were estimated visually with the aid of a calibration strip. 6323 individual (hkl) and $(\overline{hkl})$ reflections were measured. These were made up of 2875 Bijvoet pairs of reflections and 573 reflections for which only one of the pairs was measured. The corrections for Lorentz and polarization factors were made. No correction for absorption was applied ($\mu r \simeq 9.2$). The absolute scale factor was determined using the Wilson (1942) method and an average temperature factor of 2.66 Å² was derived. The final scale factor derived from the structure factor calculation differed from the Wilson scale factor by less than 1%.

3. Determination of the heavy atom position

The Bijvoet pairs of reflections $|F(hkl)|^2$ and $|F(\overline{hkl})|^2$ were averaged and the Harker section at $y = \frac{1}{2}$ was calculated using the coefficients sharpened to point atoms. The highest peak was chosen as the cobalt-cobalt vector and the coordinates derived were:
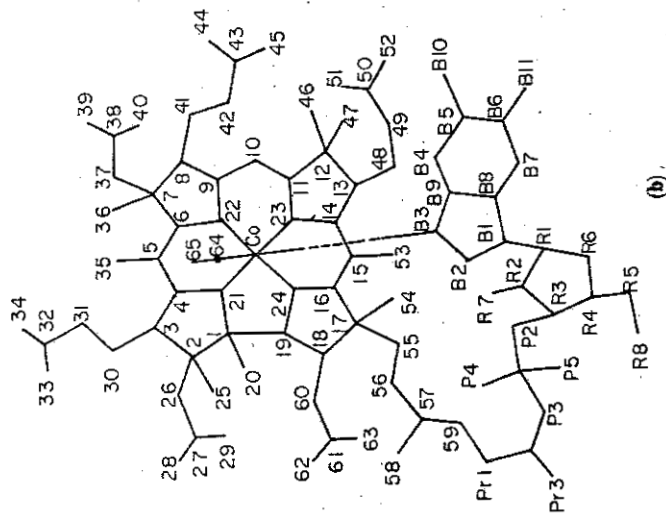
$$x = 0.6328, \qquad y = 0.2500, \qquad z = 0.1646.$$

After refinement the positional coordinates were $x = 0.6339$, $y = 0.2500$, $z = 0.1645$.
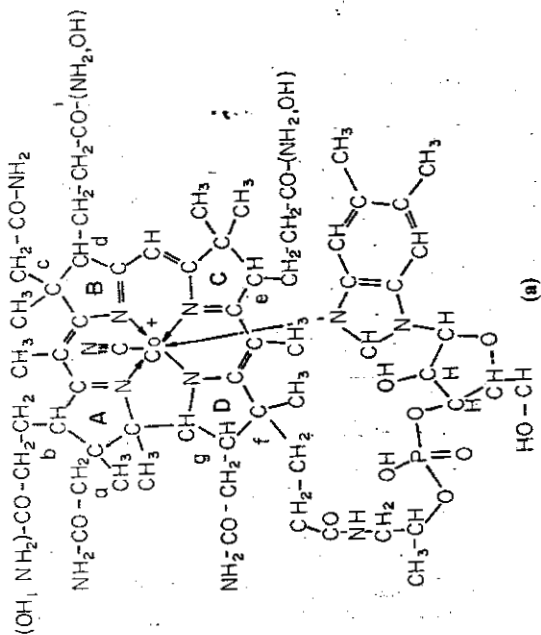
4. Calculation of the phase angles from the anomalous data

It has been shown (Peerdemann and Bijvoet, 1956; Ramachandran and Raman 1956) that the phase angle α, associated with a reflection, can be calculated from the relations:

$$\cos \psi = (|F(hkl)|^2 - |F(\overline{hkl})|^2)/4F'F''_A,$$

(a)

(b)

Formula 1 a) The structure of the mono acid; the acid group is probably at *b*, *d* or *e*. b) atom numbering used.

where ψ is the angle between the non-dispersive structure factor F' and the imaginary component F''_A. In our case

$$F''_A = F''_{Co} = 2f''_{Co} \cos 2\pi(hx_{Co} + ky_{Co} + lz_{Co}) \exp(-B \sin^2 \theta/\lambda^2),$$

and

$$F'^2 = \tfrac{1}{2}\{|F(hkl)|^2 + |F(\overline{hkl})|^2\} - |F''_{Co}|^2.$$

The two possible values of the phase angle associated with $F'(hkl)$ are $\alpha_{1,2} = \alpha_{Co} + \pi/2 \pm \psi$. Since in our case, the two cobalt atoms are centrically related, α_{Co} is $0°$ or $180°$ and the two alternative phase angles are α_1 and $180 - \alpha_1$. There is usually a strong bias towards the alternative phase angle nearest α_{Co} and this was chosen in the very successful structure determination of factor V_{1a}, $(C_{46}H_{66}O_9N_{11}CO \cdot 11 H_2O)$ (Venkatesan *et al* 1971). However in the case of factor V_{1a} the heavy atom to light atom ratio, $\Sigma f_H^2/\Sigma(f_L^2 + f_H^2)$, $r = 0.135$ while in our case $r = 0.084$. A program was therefore written to calculate F''_{Co}, $F'(hkl)$ and $\cos\psi = \Delta F^2/4|F''_{Co}||F'|$. The phase closer to the heavy atom phase was then selected by a routine which should place it in the appropriate quadrant of the phase circle,

If $\Delta F^2 > 0$ and $\alpha_{Co} = 0$, 1) between 0 and $\pi/2$,

If $\Delta F^2 > 0$ and $\alpha_{Co} = \pi$, 3) between π and $-\pi/2$,

If $\Delta F^2 < 0$ and $\alpha_{Co} = 0$, 4) between $-\pi$ and 0,

If $\Delta F^2 < 0$ and $\alpha_{Co} = \pi$, 2) between $\pi/2$ and π.

In the calculation of F''_{Co}, B was assumed to be zero instead of the average 2.66 Å^2 found. This approximation had been adopted in the analysis of factor V1a and found to improve the anomalous phasing, perhaps because no absorption correction had been applied and the effects of absorption and temperature oppose one another.

Owing to inaccuracies in the intensity measurements, $\cos\psi = \Delta F^2/4|F''_{Co}||F'|$ was found to be greater than unity in a number of cases. Where the effect was small, for 81 reflections, for which $1.0 < \cos\psi < 1.2$, $\sin\alpha$ was set equal to 1.0. Where $\cos\psi > 1.2$, the reflection was rejected. The 265 reflections in this category were usually ones in which either F' or F'' or both were small.
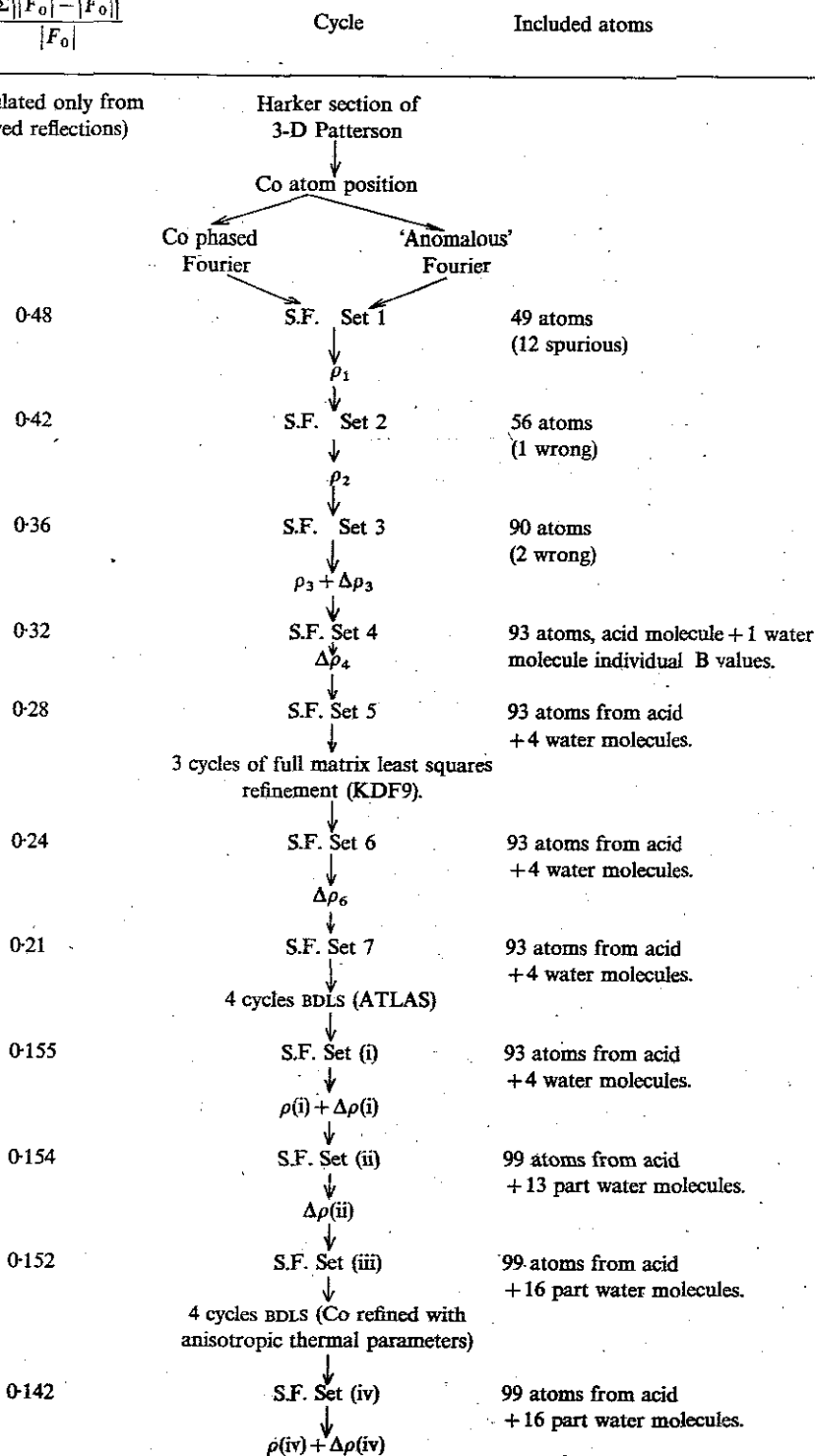
It was decided to compute two three-dimensional electron density distributions (i) the usual 'heavy atom' phased Fourier and (ii) one using the phases calculated from the anomalous differences. In this the *hol* reflections were omitted to reduce the effects of pseudo mirror symmetry, as well as the imperfectly phased terms described above, giving a total of 2567 reflections introduced into the calculation. These were weighted as suggested by Sim (1960) where the weighting fraction used in our calculation was $1 + \exp(-2x \sin\psi)$ where $x = 2|F||F_{Co}|/\Sigma f^2$.

Inadvertently, in the first phase angle calculation there was an error in the program which placed phase angles which should have been in quadrant (2) in quadrant (3) and vice versa, *i.e.* when α_{Co} was equal to π the program chose $-\alpha$ instead of α. This could have given a distribution ρ_a anom. similar to the cobalt phased map, ρ_{Co}, with a mirror plane at $y = \tfrac{1}{4}$. However the larger number of terms in the category $\alpha_{Co} = 0$ produced an inbalance more favourable to the actual structure solution though much less favourable than the correctly phased distribution evaluated later and used to compute ρ_b anom.

Flow diagram of structure refinement

$$R = \frac{\Sigma \left| |F_o| - |F_o| \right|}{|F_o|}$$

	Cycle	Included atoms
(R calculated only from observed reflections)	Harker section of 3-D Patterson	

Co atom position

Co phased Fourier ← → 'Anomalous' Fourier

R	Cycle	Included atoms
0·48	S.F. Set 1	49 atoms (12 spurious)
	ρ_1	
0·42	S.F. Set 2	56 atoms (1 wrong)
	ρ_2	
0·36	S.F. Set 3	90 atoms (2 wrong)
	$\rho_3 + \Delta\rho_3$	
0·32	S.F. Set 4	93 atoms, acid molecule + 1 water molecule individual B values.
	$\Delta\rho_4$	
0·28	S.F. Set 5	93 atoms from acid + 4 water molecules.
	3 cycles of full matrix least squares refinement (KDF9).	
0·24	S.F. Set 6	93 atoms from acid + 4 water molecules.
	$\Delta\rho_6$	
0·21	S.F. Set 7	93 atoms from acid + 4 water molecules.
	4 cycles BDLS (ATLAS)	
0·155	S.F. Set (i)	93 atoms from acid + 4 water molecules.
	$\rho(i) + \Delta\rho(i)$	
0·154	S.F. Set (ii)	99 atoms from acid + 13 part water molecules.
	$\Delta\rho(ii)$	
0·152	S.F. Set (iii)	99 atoms from acid + 16 part water molecules.
	4 cycles BDLS (Co refined with anisotropic thermal parameters)	
0·142	S.F. Set (iv)	99 atoms from acid + 16 part water molecules.
	$\rho(iv) + \Delta\rho(iv)$	

5. The determination of the light atom positions

The structure analysis was carried out through a series of calculations of electron density and difference maps, structure factor calculations and cycles of least squares as illustrated in the flow sheet. The coefficients of the Fourier series were used in the form $F' = [0 \cdot 5|F(hkl)|^2 + 0 \cdot 5|F(\overline{hkl})|^2 - |F''_{Co}(hkl)|^2]^{1/2}$. These strictly should be modified as the cobalt atom parameters are refined and a facility to do this was introduced when computing was moved to Atlas after set 7. At set 4 it was observed that the low order structure factors were calculating too low, suggesting extinction and these were corrected by the method of Pinnock et al (1956).

The two three-dimensional series ρ_{Co} and ρ_a anom first computed were compared in parallel. The 23 atoms in the corrin nucleus surrounding the cobalt atom were selected correctly, but further from the cobalt atom the situation was confused; of the next 26 selected, 11 proved wrong. These involved placing the benziminazole ring at right angles to the correct position linked with an incorrectly placed phosphorus atom. Essentially the correct solution of the atomic positions in the molecule was reached at $\rho 3$ and $\Delta \rho 3$. Additional atomic positions were sorted from later difference maps. The fact that two of the acetamide groups were disordered appeared at $\Delta \rho i$. Sixteen sites for water molecules, most of them only partly occupied, as shown in table 1, were gradually distinguished. At the final difference map, $\Delta \rho iv$, there were 45 peaks above $0 \cdot 5$ e/Å^3, the highest of $0 \cdot 8$ e/Å^3 at the cobalt atom position. Many of these peaks correspond with hydrogen atom positions; the others appear to be spurious.

The refinement of the atomic parameters proceeded throughout the analysis, at first derived from the gradients on the difference maps, particularly at $\Delta \rho 4$, later from the cycles of least squares. The first three cycles, on KDF9 at Oxford were full matrix least squares divided into two blocks. In the later two series, on Atlas at the Rutherford and Appleton laboratory, the block diagonal approximation was used and all parameters refined together. At first isotropic thermal parameters were refined for all atoms; in the

Table 1. Water molecules, occupancies and B values.

W	Occupation No.	Final B (Å^2)
1	1·0	12·8
3	1·0	16·3
4	1·0	9·0
2	0·57	8·2
5	0·43	5·4
6	0·70	10·5
7	0·50	9·0
8	0·30	7·9
9	0·50	10·6
10	0·50	18·9
11	0·50	18·1
12	0·50	13·5
13	0·34	9·3
14	0·34	9·2
15	0·50	13·3
16	0·30	8·0

last rounds, anisotropic thermal parameters were derived for the cobalt atom. The occupation numbers for the water molecules were derived empirically from the peak heights. They are listed in table 1. In the first three rounds of the least squares calculations the weighting function $\Sigma w(|F_0| - |F_0|)^2$ was used with w taken as unity. Later the weighting factor was changed to one due to Rollett and Mills (1961) with $w = |1/1 + \{(|F_0| - b)/a)^2\}^{1/2}$ $a = 5$ and $b = -50$. At the same time a fudge factor of 0·6 for positional and thermal parameters was used.

The positional and thermal parameters of the atoms are listed in table 2. Table 3 (Nockolds 1966) records the Bijvoet differences observed, F' observed and calculated, and α both determined from the Bijvoet differences and calculated from the structure found. The latest R recorded was 0·142. The scattering factors for cobalt, phosphorus, oxygen, nitrogen and carbon of Cromer and Weber (1965) were used. The scattering factors for both cobalt and phosphorus were corrected for the real part of the anomalous dispersion.

The estimated standard deviations of the atomic positions given in table 2 were derived by Cruickshank's (1959) formula and are probably underestimated since the least squares analysis depended upon the block diagonal approximation. They lead to estimated standard deviations for the bond lengths and bond angles of from 0·01–0·04 A and 1–2° respectively, also certainly underestimated.

The cobalt ion appears anisotropic with a marked vibration approximately in the direction of the a axis. It is possible that this appearance is an artefact of the data collection, since in the main series a was the axis of rotation and only two layers were measured around the subsidiary axis.

6. The use of the anomalous phasing

The observed phase angles listed in table 3 are the angles correctly derived from the Bijvoet differences. The density distribution ρ_b anom derived using these is much clearer than is ρ_a anom and would probably have led to the correct deduction of the positions of the atoms in the molecule at the first round. Table 4 summarises the distribution of the heavier peaks in ρ_b anom while table 5 lists the heights of peaks corresponding to real atomic positions in all three maps, ρ_a anom, ρ_b anom and ρ_{Co}. It is clear ρ_b anom is the best—only two atoms in the molecule have peak heights less than 1·0 e/A^3 and these give peaks of 0·8 and 0·9 e/A^3. On the other hand, it is not difficult to sort the atoms from the other maps in three rounds of calculation, as experience proved. The peaks are in general well resolved, ρ_a anom slightly better than ρ_{Co}.

Figure 1 shows that the phasing calculation was very successful. For 72 % of the reflections, the observed phase angle was within 45° of that finally calculated.

Figure 2 compares the differences between the final phase angles and those calculated from the Bijvoet differences with $B_{Co} = 0$ and 2·6 A^2. It is clear that $B = 0$ gives the better approximation. Further an examination of the effect of various weighting schemes proposed by Sim (1960) suggests that the simple empirical weighting $W = 0$ when the heavy atom contribution is very low and 1·0 otherwise would in practice have proved quite effective. Peaks appeared marginally higher in a distribution with terms so weighted rather than with the Sim weighted terms.

A comparison between the observed magnitude of the Bijvoet differences and theoretical estimates (Parthasarathy and Srinivasan 1964; Parthasarathy 1965) in-
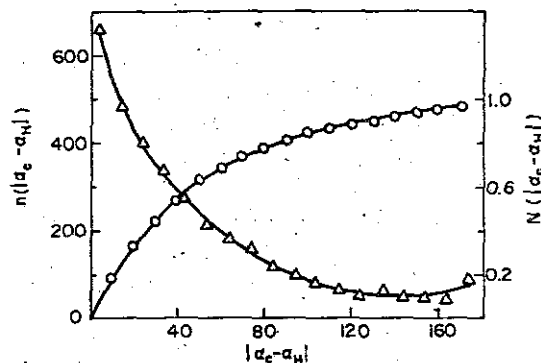
Figure 1. Phase analysis. Here n is the number of reflections in each $10°$ range and $N([\alpha_a - \alpha_c])$ is the normalised cumulative function for each value of $[\alpha_a - \alpha_c]$ where α_a is the phase (correctly calculated) from the Bijvoet effect, α_c the final phase calculated.
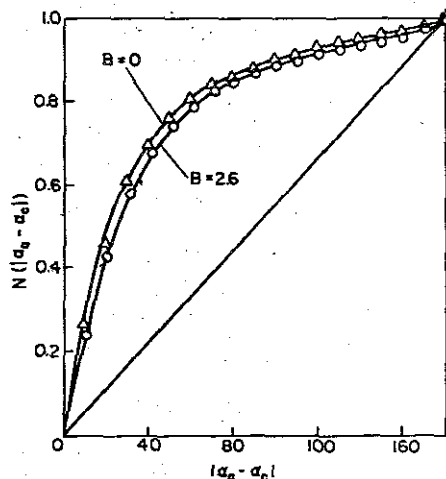

Figure 2. Effect of BCo on the calculations of α_a compared with α_c.

dicates that those observed were smaller than expected. This may be due to an approximation to centrosymmetry in the atomic distribution immediately surrounding the cobalt atom. At the same time, the effect of the heavy atom was greater than calculated, probably owing to the same feature.

7. The structure of the molecule

The geometry of the molecule and general distribution of bond lengths are very similar to those found in other derivatives of cyanocobalamin (Brink-Shoemaker *et al* 1964; Hodgkin *et al* 1962). The bond lengths and bond angles observed are recorded in figures 3 and 4 and those in the cobalt coordination sphere in figure 5. The cobalt-corrin-nitrogen distances are consistent with trivalent cobalt in essentially covalent binding.

Part I. X-ray diffraction studies of air-dried crystals

Table 2. Final positional and thermal parameters and the estimated standard deviations $\times 10^4$ in brackets.

Atom	x/a	y/b	z/c	B	Atom	x/a	y/b	z/c	B
Co	0.6339 (2)	0.2500 (0)	0.1645 (2)	*	C30	0.8162 (20)	0.4493 (17)	0.0369 (18)	5.2
C1	0.8184 (13)	0.2793 (10)	0.1349 (10)	6.1	C31	0.8480 (21)	0.4907 (18)	−0.0325 (18)	5.6
C2	0.8589 (16)	0.2692 (12)	0.0524 (12)	3.0	C32	0.1433 (17)	0.0844 (14)	0.0132 (14)	3.9
C3	0.7836 (16)	0.3629 (13)	0.0146 (13)	3.3	A34	0.2167 (18)	0.1150 (17)	0.0165 (16)	6.8
C4	0.7077 (17)	0.3407 (14)	0.0378 (14)	3.7	A33	0.5065 (18)	0.1053 (16)	0.0476 (15)	6.4
C5	0.6109 (15)	0.3635 (13)	−0.0027 (13)	3.1	C35	0.5885 (25)	0.5725 (22)	0.9067 (24)	7.8
C6	0.5779 (16)	0.3528 (14)	0.0262 (14)	3.6	C36	0.3904 (20)	0.4386 (18)	0.9314 (19)	5.8
C7	0.4291 (19)	0.3716 (16)	0.0902 (16)	4.7	C37	0.3874 (21)	0.2950 (18)	0.6393 (18)	5.8
C8	0.3812 (19)	0.3794 (16)	0.0704 (16)	4.9	C381	0.2637 (29)	0.2893 (25)	−0.0878 (24)	5.4
C9	0.4497 (18)	0.3774 (15)	0.1277 (15)	4.2	C382	0.2974	0.3314	−0.1276	6.3
C10	0.4187 (19)	0.2756 (16)	0.1892 (16)	5.0	N401	0.2438 (27)	0.3152 (25)	−0.1598 (26)	7.3
C11	0.4681 (19)	0.2123 (16)	0.2325 (16)	4.7	N402	0.3024	0.3307	−0.2054	2.6
C12	0.4251 (21)	0.2051 (18)	0.2782 (18)	5.7	O361	0.2149 (20)	0.2106 (18)	−0.0383 (18)	4.6
C13	0.5083 (15)	0.1112 (12)	0.3322 (12)	2.6	O92	0.2196	0.3457	−0.0860	5.2
C14	0.5911 (16)	0.1326 (12)	0.2860 (12)	3.0	C41	0.4052 (21)	0.4679 (19)	0.0989 (19)	6.0
C15	0.6773 (14)	0.0996 (11)	0.2985 (11)	2.1	C42	0.3717 (27)	0.4717 (25)	0.1900 (28)	9.2
C16	0.7503 (15)	0.1261 (12)	0.2574 (12)	2.6	C43	0.6382 (23)	0.0574 (20)	0.7878 (21)	6.6
C17	0.8434 (13)	0.0858 (9)	0.2594 (9)	1.4	N45	0.6367 (20)	0.1197 (18)	0.8420 (19)	0.8
C18	0.6017 (14)	0.1545 (11)	0.2238 (11)	2.1	O44	0.6275 (22)	0.0722 (20)	0.7008 (21)	9.3
C19	0.8184 (14)	0.1909 (10)	0.1624 (10)	1.8	C46	0.3986 (23)	0.0600 (21)	0.2120 (21)	7.0
N21	0.7201 (11)	0.2990 (8)	0.1601 (8)	1.6	C47	0.3536 (29)	0.1858 (27)	0.3295 (29)	6.4
N22	0.5336 (13)	0.3172 (11)	0.1090 (11)	3.7	C48	0.5470 (23)	0.1388 (17)	0.4259 (17)	5.4
N23	0.5583 (11)	0.1974 (9)	0.2286 (9)	2.4	C49	0.4870 (27)	0.4697 (24)	0.4923 (23)	7.4
N24	0.7342 (11)	0.1884 (8)	0.2057 (8)	1.6	C50	0.5032 (27)	0.0108 (23)	0.4936 (21)	6.9
C20	0.9860 (14)	0.3301 (11)	0.1211 (11)	2.2	A51	0.5368 (18)	0.4642 (16)	0.5576 (17)	7.1
C25	0.9615 (19)	0.3191 (16)	0.0070 (16)	5.0	A52	0.4268 (23)	0.4826 (20)	0.4471 (22)	9.6
C26	0.8457 (16)	0.2227 (13)	0.0644 (13)	3.3	C53	0.6914 (17)	0.0272 (14)	0.3612 (14)	3.9
C271	0.8549 (18)	0.2376 (14)	−0.0966 (14)	3.5	C54	0.1741 (16)	0.5135 (13)	0.7949 (13)	3.2
N281	0.9400 (15)	0.2241 (12)	−0.1082 (12)	3.8	C55	0.8888 (14)	0.0602 (11)	0.3471 (10)	2.0
O291	0.7906 (16)	0.2631 (13)	−0.1465 (13)	4.6	C56	0.9123 (16)	0.1217 (12)	0.4159 (12)	2.9

Table 2. *(Continued)*

Atom	x/a	y/b	z/c	B	Atom	x/a	y/b	z/c	B
C57	0·9416(14)	0·0842(11)	0·5026(11)	2·3	R7	0·8826(11)	0·3229(9)	0·4318(9)	3·9
O58	0·8860(12)	0·0438(10)	0·5326(11)	5·4	R8	0·7376(12)	0·2902(11)	0·6929(11)	5·2
N59	0·0310(13)	0·0922(11)	0·5391(11)	3·7	P1	0·0457(4)	0·2799(4)	0·6714(4)	3·4
C60	0·9773(18)	0·1247(14)	0·1855(13)	3·5	P2	0·9610(10)	0·2692(8)	0·5924(8)	3·3
C61	0·0693(16)	0·1098(14)	0·2474(14)	3·5	P3	0·0982(10)	0·1968(9)	0·6727(8)	3·4
O62	0·1020(13)	0·1598(12)	0·3018(12)	6·2	P4	0·0061(13)	0·2912(11)	0·7484(11)	5·7
N63	0·1077(18)	0·0432(16)	0·2437(17)	6·7	P5	0·1164(14)	0·3409(13)	0·6564(13)	6·8
C64	0·5991(16)	0·1798(12)	0·0772(12)	2·9	Pr1	0·0765(18)	0·0652(14)	0·6311(14)	3·9
N65	0·5881(14)	0·1389(12)	0·0165(12)	4·1	Pr2	0·0489(16)	0·1261(13)	0·6906(13)	3·2
B1	0·7001(11)	0·3680(9)	0·3919(9)	2·2	Pr3	0·0977(18)	0·1017(15)	0·7825(15)	4·2
B2	0·6906(14)	0·3049(10)	0·3428(11)	2·0	W1	0·1861(24)	0·1588(23)	0·4691(22)	12·8
B3	0·6672(12)	0·3294(9)	0·2618(9)	2·6	W3	0·0752(27)	0·3730(29)	0·4753(26)	16·2
B4	0·6443(16)	0·4624(12)	0·2040(12)	2·9	W4	0·7226(17)	0·4584(17)	0·6574(16)	9·0
B5	0·3526(16)	0·0458(13)	0·7827(14)	3·6	W2	0·9636(28)	0·4106(28)	0·8419(26)	8·2
B6	0·3332(16)	0·0687(13)	0·6993(12)	3·1	W5	0·0206(32)	0·4023(30)	0·8694(28)	5·4
B7	0·3140(16)	0·0155(12)	0·6304(12)	3·2	W6	0·5549(42)	0·2398(46)	0·6839(39)	10·5
B8	0·6824(13)	0·4365(10)	0·3446(10)	1·8	W7	0·4999(61)	0·3162(61)	0·6275(55)	9·0
B9	0·6591(14)	0·4102(11)	0·2635(10)	2·0	W8	0·4412(55)	0·3275(53)	0·6756(49)	7·9
B10	0·3639(19)	0·1076(16)	0·8550(17)	4·9	W9	0·3923(41)	0·3479(41)	0·7432(38)	10·6
B11	0·3369(19)	0·1579(16)	0·6766(16)	4·9	W10	0·2808(62)	0·2883(68)	0·5402(60)	18·9
R1	0·7454(16)	0·3626(12)	0·4957(13)	3·0	W11	0·3187(59)	0·3394(65)	0·4868(53)	18·1
R2	0·8558(14)	0·3617(11)	0·5032(11)	2·1	W12	0·8071(46)	0·0123(46)	0·6778(43)	13·5
R3	0·8771(14)	0·3111(11)	0·5818(11)	2·3	W13	0·1162(58)	0·3112(58)	0·2893(54)	9·3
R4	0·7952(15)	0·2567(14)	0·5688(12)	3·4	W14	0·5725(75)	0·2762(78)	0·7508(69)	9·2
R5	0·7615(16)	0·2250(12)	0·6460(12)	3·1	W15	0·1694(48)	0·3447(50)	0·1850(45)	13·3
R6	0·7128(10)	0·2960(8)	0·5171(8)	3·4	W16	0·1659(58)	0·2637(58)	0·0203(50)	8·0

* Final anisotropic temperature factor parameters for cobalt in the expression $\exp - (B_{11}h^2 + B_{22}k^2 + B_{33}l^2 + B_{12}hk + B_{13}hl + B_{23}kl)$ ($\times 10^5$).
$B_{11} = 418$, $B_{22} = 144$, $B_{33} = 118$, $B_{12} = 35$, $B_{13} = -233$, $B_{23} = -22$.

Table 3. The column headings represent the following terms.

$$dI/I = \frac{|F_{obs}^2 (hkl)|^2 - |F_{obs}(\overline{hkl})|^2}{0 \cdot 5|F_{obs}^2 (hkl)|^2 + 0 \cdot 5|F_{obs}(\overline{hkl})|^2} \times 100$$

$$F_0 = \{0 \cdot 5|F_{obs}^2 (hkl)| + 0 \cdot 5|F_{obs}^2 (\overline{hkl})| - F_{Co.}^{\prime 2} (hkl)\}^{1/2}$$

F_c is the calculated structure amplitude
$a(A)$ is the phase determined from the Bijvoet differences
$a(C)$ is the final calculated phase.

The table is available in the Radcliffe Library, Oxford.

Table 4. Peaks in ρ_b anom, a summary.

Peak Height (e/A^3)	Total	No. atom peaks	No. spurious peaks
> 3·0	12	11	1
2·5–3·0	12	11	1
2·0–2·5	26	20	6

There are marked differences in the cobalt-carbon and cobalt-nitrogen distances in the axial positions.

7.1 The corrin nucleus

The average bond length in the inner ring of thirteen atoms of the nucleus is 1·39 A, corresponding with the existence of the expected resonance system. Within this ring the variations are similar in character but less accurately determined than those measured in, for example, factor V_{1a} (Venkatesan *et al* 1971). The buckling of the ring, which is shown in figure 6 is similar to that found in other nucleotide-containing cobalamines. Each of the four five-membered rings has a slightly different conformation from the others. From the torsion angles, shown in figure 7, ring D is much closer to $C_s(m)$ symmetry than the rings A, B and C.

7.2 Conformations of the side chains attached to the corrin nucleus

The two acetamide groups on rings A and B both show disorder illustrated by the difference maps in figure 8. They appear to have been caught beginning to turn as one group does in the change from wet to dry vitamin B_{12}. Here all observed conformations are still nearer to the positions, turning outwards, which are shown in wet B_{12} rather than inwards in the corresponding air-dried B_{12} crystals. Figures 9 and 10 record the torsion angles in the major conformation of the groups on rings A, B, C and D. Two on rings A and B are essentially extended; that on C approximates to the gauche form.

7.3 The position of the acid

Formally it might be possible to recognise the single acid among the amides either from
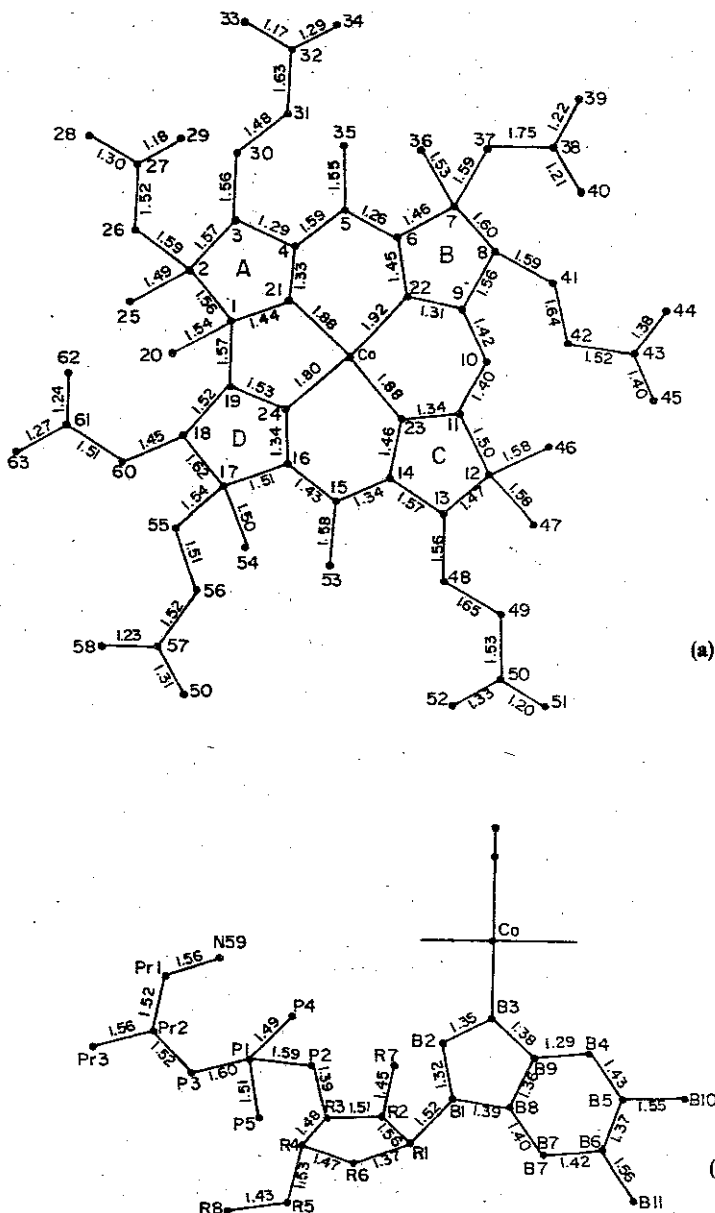
Figure 3. Interatomic distances (a) in the nucleus. Only one alternative for each of the acetamides in the A and B rings is shown. (b) in the propanolamine and nucleotide.

the interatomic distances, or the relative electron densities of the terminal atoms or their arrangement in the crystal.

Here the standard deviations of atomic positions are too large to permit differentiation by the distances within the terminal groups. The arrangement in the crystal is also inconclusive; there are at least two relatively short contacts with hydrophilic groups for
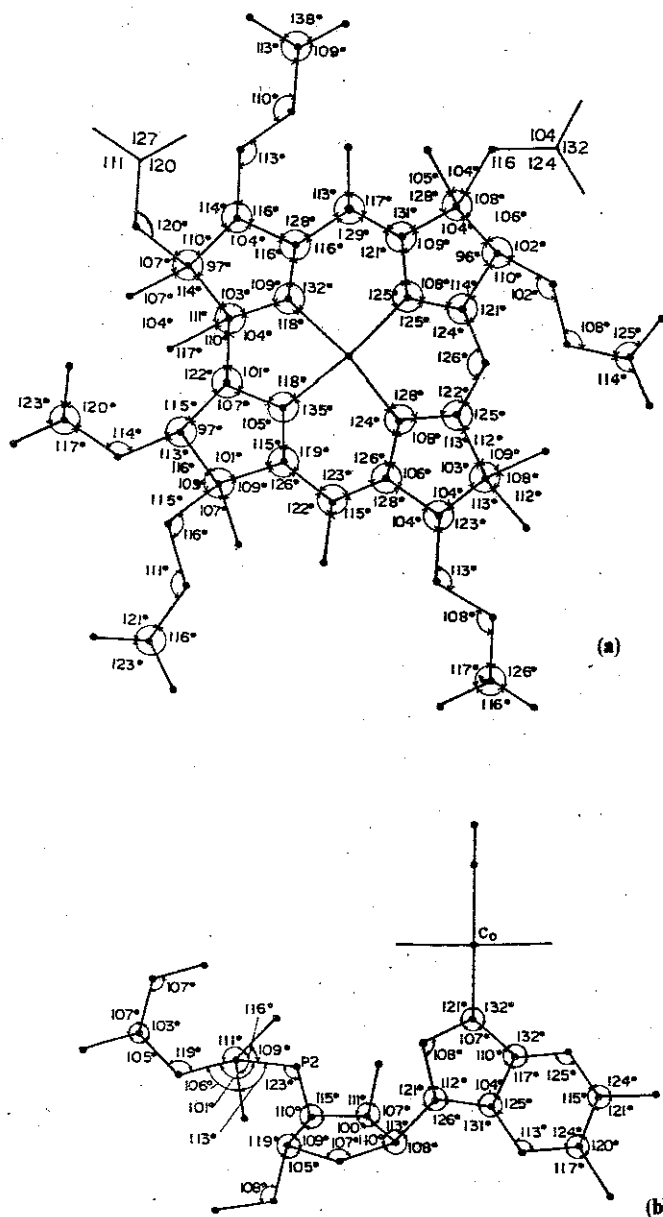
Figure 4. Bond angles (a) in the nucleus (b) in the propanolamine and nucleotide.

every one of the terminal atoms of the acetamide and propionamide or acid side chains. This is illustrated diagramatically in figure 11. There remains the electron densities which are listed in table 6, compared with the identifications suggested by the neutron study.

The electron densities are also clearly affected by the inaccuracies of the analysis and the process of crystal drying. The two terminal atoms actually nearest to one another in
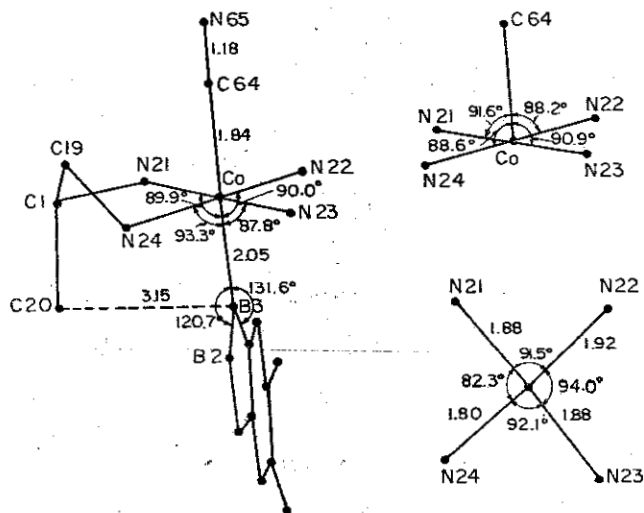
Figure 5. Interatomic distances and bond angles in the cobalt coordination sphere.
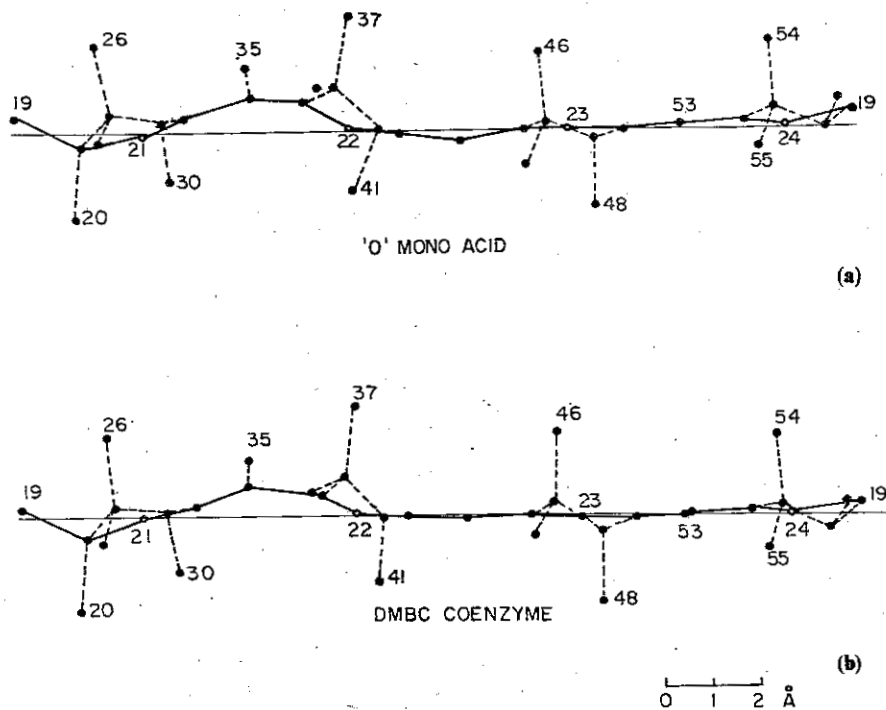
Figure 6. Atomic positions projected on a cylinder of 2·8 A around an axis perpendicular to the least squares plane of N21, 22, 23, 24. (a) O-mono acid (b) dimethylbenzimidazole cobamide coenzyme.
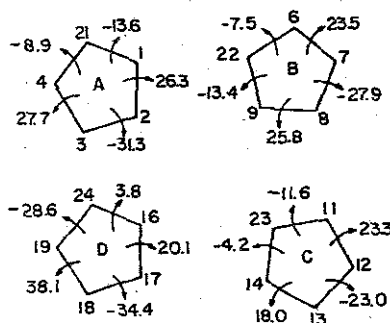
Figure 7. Observed torsion angles (in degrees) in rings, A, B, C and D.
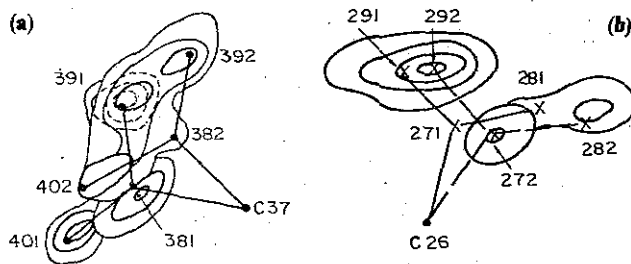
Figure 8. Difference electron density over (a) at the alternative positions for C38, O39, N40 after removal from the phasing calculations (b) in the region of the acetamide on ring A; C272, N282 and O292 not included in the phasing calculations.

electron density are the acetamide O, N 391–401 at 4·4 and 4·2 $e/\text{Å}^3$. As this group is certainly from other evidence an amide, the densities must be viewed as unreliable. Bearing this warning in mind, it is perhaps still worth noticing that on the density criterion the group at *b* should be the acid with 33 and 34, 5·1 and 5·4 $e/\text{Å}^3$ both oxygen atoms. This is the identification eventually found by neutron diffraction; other identifications are however different as indicated in the table.

7.4 *The nucleotide and the propanolamine*

The observed distances and angles in the nucleotide and propanolamine are shown in figures 3b and 4b. Their conformations are illustrated in figure 12. The benziminazole ring is sensibly planar, as would be expected; atom B10 deviates the most from this plane, (0·11 Å). The ribose ring has the envelope conformation with atom R2 out of the plane defined by the other four atoms. The hydroxyl group R8 is turned, as in air-dried B_{12}, towards the ring oxygen atom R6, 2·8 Å away.

When the whole of the nucleotide is projected onto the least squares plane through N21, N22, N23, N24, it can be seen to have shifted from the position it occupies in B_{12}. This seems to be due to a change in the angle between the benziminazole group and the plane of the corrin ring which carries the sugar and phosphate group about 0·5 Å nearer to ring D. The whole sugar, phosphate, propanolamine chain is rather flexible; this

Table 5. Peak heights in ρ_a anom, ρ_b anom and ρ_{Co} for the O mono acid.

| Atom	ρ_a anom	ρ_b anom	ρ_{Co}	Atom	ρ_a anom	ρ_b anom	ρ_{Co}
Co	28·0	40·0	59·2	O 39	0·7	1·4	1·0
C 1	1·6	2·1	0·9	C 41	1·4	2·3	1·1
C 2	1·2	1·6	1·6	C 42	1·0	1·3	0·5
C 3	0·9	1·4	1·2	C 43	1·4	1·2	0·1
C 4	0·6	1·5	1·5	N 45	1·0	2·2	1·2
C 5	1·6	1·3	1·2	O 44	1·0	1·4	0·7
C 6	1·1	1·3	1·2	C 46	1·3	1·4	0·4
C 7	0·6	1·3	0·5	C 47	1·6	2·0	0·7
C 8	0·8	1·3	1·0	C 48	0·8	1·6	1·3
C 9	0·9	2·0	2·0	C 49	1·6	1·4	0·6
C 10	2·0	3·0	2·1	C 50	2·0	1·8	2·0
C 11	1·2	1·9	1·5	A 51	0·5	1·7	1·0
C 12	1·2	1·4	1·4	A 52	1·4	1·9	1·3
C 13	1·4	1·8	1·4	C 53	1·2	1·7	1·1
C 14	1·0	1·7	1·1	C 54	0·9	2·4	1·6
C 15	0·5	1·9	0·8	C 55	1·8	1·8	0·8
C 16	0·8	2·5	1·8	C 56	0·5	2·5	1·2
C 17	0·7	1·7	1·4	C 57	1·4	2·9	2·2
C 18	1·8	1·8	2·0	O 58	1·6	2·8	2·0
C 19	1·5	0·9	1·0	N 59	1·6	2·5	1·3
N 21	2·0	3·2	2·6	C 60	0·8	2·5	1·4
N 22	1·5	2·4	2·2	C 61	1·0	1·4	0·9
N 23	2·0	3·8	2·1	O 62	1·3	1·0	0·7
N 24	2·5	3·6	2·7	N 63	1·6	2·4	0·4
C 20	2·6	3·1	2·5	C 64	0·8	1·1	1·0
C 25	1·3	1·7	1·1	N 65	0·2	1·4	1·5
C 26	1·6	1·8	1·8	B 1	1·2	2·2	1·1
C 271	0·9	1·4	1·0	B 2	1·3	1·2	1·0
N 281	1·2	2·1	1·4	B 3	1·0	2·6	1·7
O 291	1·8	2·3	1·4	B 4	0·6	1·3	0·5
C 30	1·3	2·0	0·8	B 5	1·4	3·2	1·0
C 31	2·0	1·2	1·1	B 6	1·2	1·5	1·0
C 32	0·9	2·3	0·9	B 7	1·2	1·4	*
A	0·3	1·1	0·6	B 8	1·5	1·9	1·1
A	0·8	1·6	1·0	B 9	1·1	1·8	0·9
C 35	0·4	1·5	0·5	B 10	0·9	1·4	0·3
C 36	1·4	1·4	1·2	B 11	1·1	2·2	1·6
C 37	0·4	1·0	1·3	R 1	1·1	2·4	1·6
C 381	0·7	1·5	0·7	R 2	0·8	1·6	1·8
N 401	0·6	0·8	0·6	R 3	1·3	3·0	2·0
R 4	2·6	3·5	4·4	W 1	1·0	1·4	1·2
R 5	1·1	2·4	1·3	W 3	0·8	1·4	0·9
R 6	1·0	2·6	1·7	W 4	1·1	2·6	1·6
R 7	1·8	3·6	2·2	W 2	0·8	2·0	1·4
R 8	1·3	2·4	1·2	W 5	1·1	1·0	0·6
P 1	3·1	4·8	4·3	W 6	1·2	1·3	0·4
P 2	2·7	3·3	3·5	W 7	*	*	0·8
P 3	1·0	2·6	1·8	W 8	*	0·8	0·0
P 4	1·0	2·7	1·0	W 9	0·9	1·9	1·0
P 5	1·5	2·2	2·0	W 10	0·9	1·5	*
Pr 1	1·6	1·5	0·8	W 11	0·9	0·6	0·0
Pr 2	0·8	1·0	0·7	W 12	1·2	0·7	*
Pr 3	0·9	2·0	0·4	W 13		1·3	*

* Atom position in negative region.
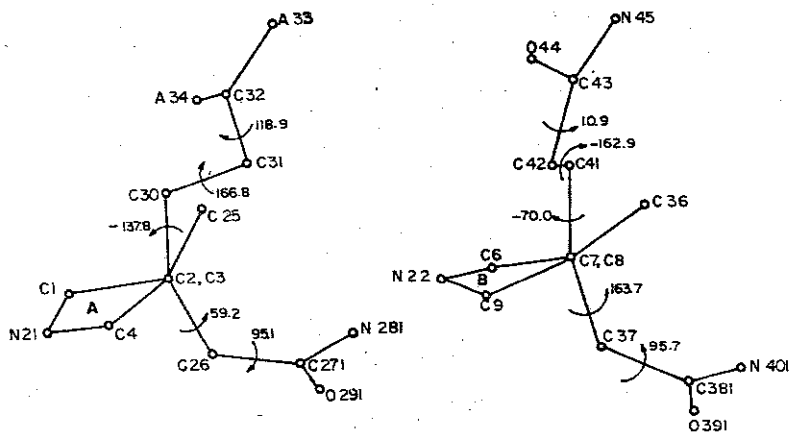
Figure 9. Conformation of acetamide and propionic acid side chains of rings A and B.
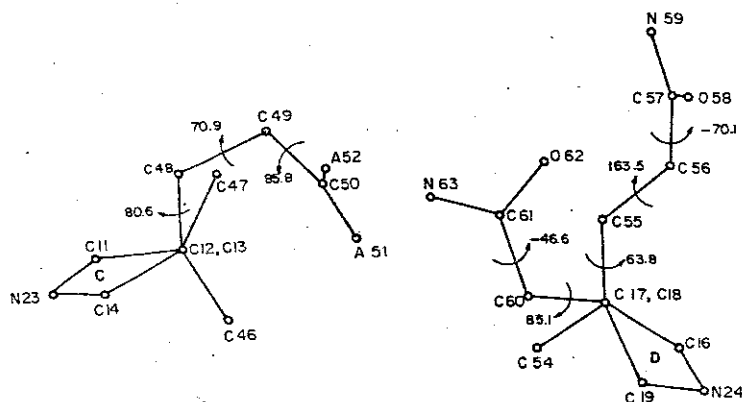

Figure 10. Conformation of acetamide and propionic side chains of rings C and D.

Table 6. Electron density over peaks at terminal side chain atoms.

Position	No.	Peak height	Identification from neutron analysis
a	281	6·0	N
	291	6·6	O
b	33	5·4	O
	34	5·1	O
c	391	4·2	O
	401	4·4	N
d	44	4·1	O
	45	4·7	N
e	51	4·7	(O)
	52	3·7	(N)
f	62	5·9	O
	63	5·4	N
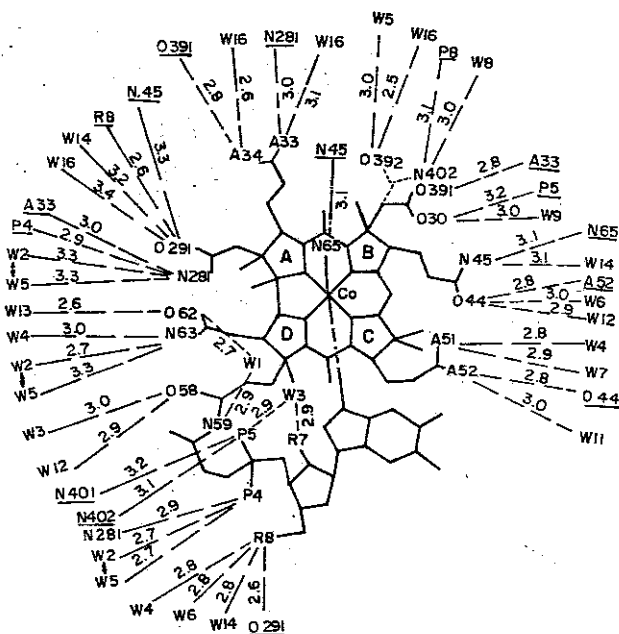
Figure 11. Diagram of intermolecular contact distances less than 3·4 Å.
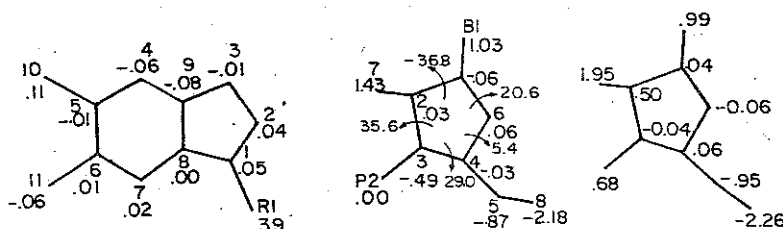
Figure 12. Projection on to least squares planes (a) through B1–11, (b) R, 1, 2, 4 and 6, R1346 defining the benzimidazole and ribose rings.

change may well be a consequence of packing conditions in this particular crystal structure (figures 13 and 14).

8. The crystal structure

Figures 15 and 16 are projections of the crystal structure along the b and a axes respectively. They show that the arrangement of the molecules is very different from that found in cyanocobalamin. Here the plane of the corrin nucleus is nearly parallel with the a axis but inclined at about 45° to the b axis and consequently at about 90° to the planes in neighbouring molecules. There is a distinct gap between layers of molecules in the (10$\bar{1}$) plane from one side of the crystal to the other. This gap is
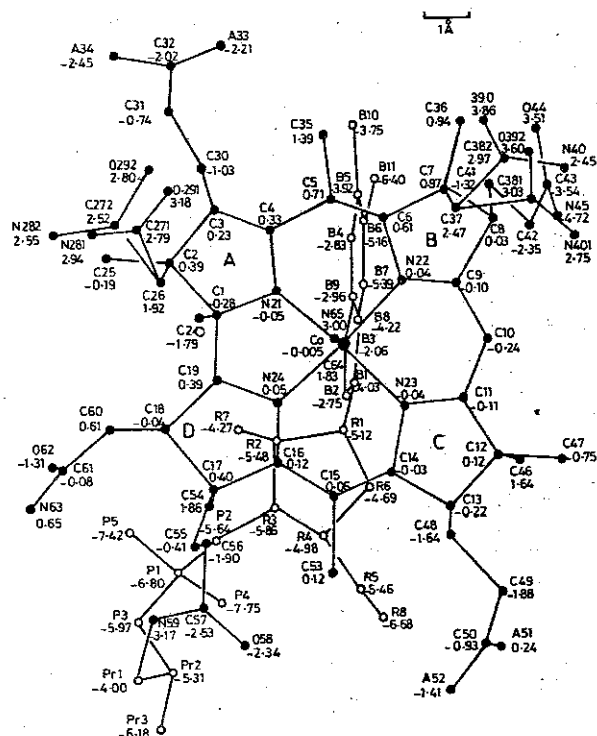
Figure 13. Projection of the positions of the atoms in the mono acid molecule on to the least squares plane through Co, N21, 22, 23, 24. The figures show the heights of the atoms in Å above and below the plane.
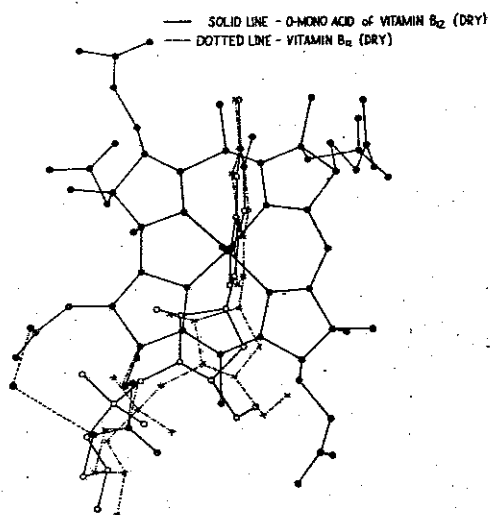
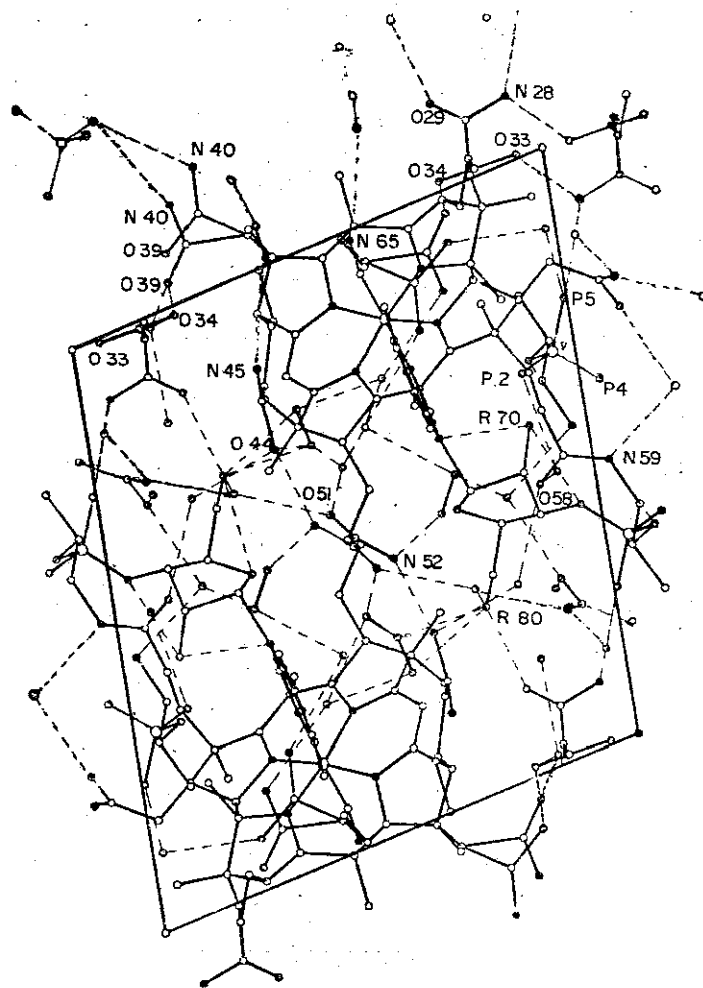**Figure 14.** Projection of the atomic positions on to the least square plane through Co, N21, N22, N23 and N24.

Figure 15. Atomic positions projected along *b*. Hydrogen bonds dotted.

occupied by water molecules and constitutes a strong cleavage plane in the crystals.

Between the molecules there are eight direct contacts between hydrophilic groups in different molecules in addition to contacts through single water molecules which appear strongly bound. These contacts are marked in figure 11. Only one of the two water molecules found held within a single molecule in many other B_{12} derivatives is present; that usually found between A52 and O58 is missing. It is however present in the wet crystal structure.

It is indeed easy to correlate the sixteen sites for water molecules found in this crystal structure with thirteen of the water molecules found in the wet crystal structure, though the relation is not 1:1. Some of the sites in the air-dried crystals clearly represent alternative positions to which a single water molecule in the wet crystal could move. Three of the water molecules in the wet crystals have no corresponding sites in the dried crystals and represent the three lost in the transformation suggested by the density. To
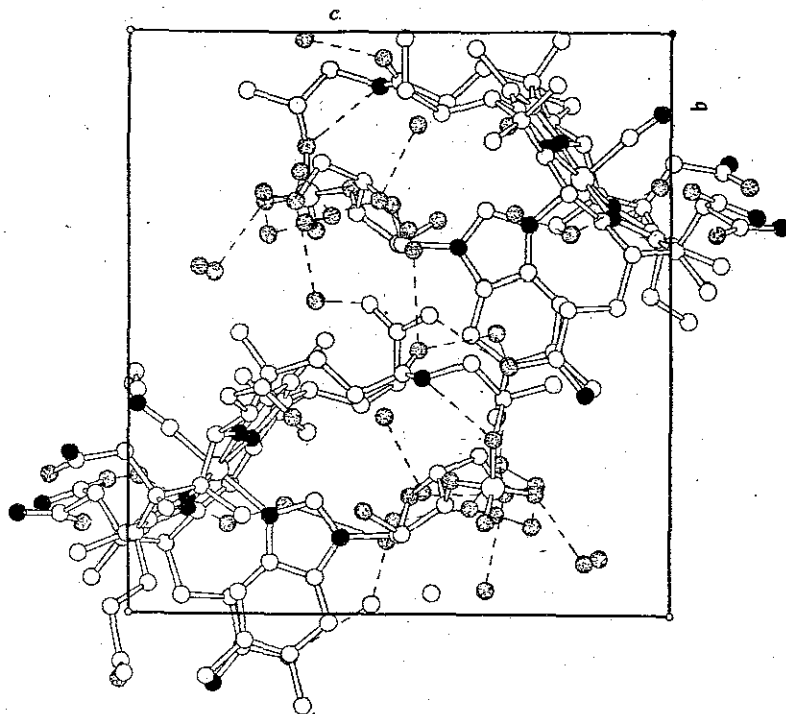
Figure 16. Atomic positions projected along *a*.

confirm that there was no loss of water actually during data collection a zero layer Weissenberg photograph was taken after the main axis photographs had been collected. There was no observable difference in the photograph taken at the end of the experiment and that taken before.

In conclusion it is difficult to see any reason why the introduction of the acid group at *b*, the position found through neutron diffraction, should favour such a radical change in molecular packing from that of cyanocobalamin itself as is observed in this crystal structure.

Acknowledgements

References

Brink-Shoemaker C, Cruickshank D W J, Hodgkin D C, Kamper M J and Pilling D 1964 *Proc. R. Soc.* (*London*) **A278** 1

Cruickshank D W J 1959 *Computing methods in crystallography* (Oxford: Pergamon Press) Ch. 14
Cromer D T and Weber J T 1965 *Acta Crystallogr.* **18** 104
Hodgkin D C, Lindsey J, Sparks R A, Trueblood K N and White J G 1962 *Proc. R. Soc. (London)* A266 494
Moore F H, Willis B T M and Hodgkin D C 1967 *Nature (London)* **214** 130
Nockolds C K 1966 *An x-ray analytical study of structures related to vitamin B_{12}*, Thesis for D.Phil. degree, Oxford
Nockolds C E, Ramaseshan S, Waters J M, Waters T N M and Hodgkin D C 1967 *Nature (London)* **214** 129
Parthasarathy S 1965 *Acta Crystallogr.* **18** 1028
Parthasarathy S and Srinivasan R 1964 *Acta Crystallogr.* **17** 1400
Peerdemann A F and Bijvoet J M 1956 *Acta Crystallogr.* **9** 1012
Peerdemann A F, Bommel A J V and Bijvoet J M 1951 *Proc. Acad. Sci. Amsterdam* **54** 16
Pinnock P R, Taylor C A and Lipson H 1956 *Acta Crystallogr.* **9** 175
Ramachandran G N and Raman S 1956 *Curr. Sci.* **25** 348
Sim G A 1959 *Acta Crystallogr.* **12** 813
Venkatesan K, Dale D, Hodgkin D C, Nockolds C E, Moore F H and O'Connor B H 1971 *Proc. R. Soc. (London)* A323 455
Wilson A J C 1942 *Nature (London)* **130** 152