

We thank Professor M. A. Mokulskii for useful discussions, Dr Yu. T. Struchkov and Dr A. I. Yanovskii for making available CSD, Dr E. P. Smetanina for assistance in selecting bibliographic data on the structures not presented in CSD, and Mr D. A. Agrachev for his help in preparing the manuscript.

References

- ALEXEEV, D. G., LIPANOV, A. A. & SKURATOVSKII, I. YA. (1987). *Nature (London)*. In the press.
- ALTONA, C. & SUNDARALINGAM, M. (1972). *J. Am. Chem. Soc.* **94**, 8205-8212.
- Cambridge Structural Database (CSD) (1984). Univ. Chemical Laboratory, Lensfield Road, Cambridge, England.
- DAVIES, D. B. & DANYLUK, S. S. (1975). *Biochemistry*, **14**, 543-554.
- DE LEEUW, H. P. M., HAASNOOT, C. A. G. & ALTONA, C. (1980). *Isr. J. Chem.* **20**, 108-126.
- DUNITZ, J. D. (1972). *Tetrahedron*, **28**, 5459-5463.
- LESLIE, A. G. W., ARNOTT, S., CHANDRASEKARAN, R. & RATLIFF, R. L. (1980). *J. Mol. Biol.* **143**, 49-72.
- LEVITT, M. & WARSHEL, A. (1978). *J. Am. Chem. Soc.* **100**, 2607-2613.
- LUDEMANN, H. D. & WESTHOF, E. (1978). In *Nuclear Magnetic Resonance Spectroscopy in Molecular Biology*, edited by B. Pullman, p. 41. Dordrecht: D. Reidel.
- LUGOVSKOI, A. A. & DASHEVSKII, V. G. (1972). *Mol. Biol. (USSR)*, **6**, 440-448.
- MERRITT, E. A. & SUNDARALINGAM, M. (1985). *J. Biomol. Struct. Dyn.* **3**, 559-578.
- MURRAY-RUST, P. & MOTHERWELL, S. (1978). *Acta Cryst.* **B34**, 2534-2546.
- OLSON, W. K. (1982). *J. Am. Chem. Soc.* **104**, 278-286.
- PATTABIRAMAN, N. & RAO, M. J. (1982). *Int. J. Biol. Macromol.* **4**, 91-98.
- PEARLMAN, D. A. & KIM, S.-H. (1985). *J. Biomol. Struct. Dyn.* **3**, 85-98.
- VOROBYEV, YU. N. (1981). *Mol. Biol. (USSR)*, **15**, 517-525.
- WESTHOF, E. & SUNDARALINGAM, M. (1980). *J. Am. Chem. Soc.* **102**, 1493-1500.
- ZHURKIN, V. B., LYSOV, YU. P., FLORENTIEV, V. L. & IVANOV, V. I. (1982). *Nucleic Acids Res.* **10**, 1811-1830.

Acta Cryst. (1987). **B43**, 280-295

High-Resolution Neutron and X-ray Refinement of Vitamin B₁₂ Coenzyme, C₇₂H₁₀₀CoN₁₈O₁₇P.17H₂O

BY H. F. J. SAVAGE,* P. F. LINDLEY AND J. L. FINNEY

Department of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, England

AND P. A. TIMMINS

Institut Laue-Langevin, 156X, 38042 Grenoble CEDEX, France

(Received 4 June 1986; accepted 1 December 1986)

Abstract

High-resolution neutron and X-ray diffraction data for vitamin B₁₂ coenzyme have been obtained with the objective of elucidating the water organization in this crystalline system. Here, details of the data collection and refinement of the individual models are described. Neutron data: $M_r = 1939$ (assuming exchange of 19 coenzyme H atoms by D atoms and 17 D₂O solvent molecules) $P_12_12_12_1$, $a = 27.849$ (6), $b = 21.736$ (4), $c = 15.368$ (3) Å, $U = 9303$ (2) Å³, $Z = 4$, $D_m = 1.381$ (15), $D_x = 1.360$ Mg m⁻³, pyrolytic-graphite monochromator, $\lambda = 1.67$ Å, $\mu = 0.024$ mm⁻¹, $F(000) = 34.13$, $T = 279$ (1) K, final $R = 0.085$ for 5601 significant reflections. X-ray1 data: $M_r = 1939$, $P_21_21_21$, $a = 27.701$ (7), $b = 21.608$ (6), $c = 15.351$ (4) Å, $U = 9189$ (2) Å³, $Z = 4$, $D_m = 1.381$ (15), $D_x = 1.401$ Mg m⁻³, Cu $K\alpha$ radiation, $\lambda =$

1.5418 Å, $\mu = 25.9$ mm⁻¹, $F(000) = 4024$, $T = 277.0$ (5) K, final $R = 0.088$ for 4390 significant reflections. X-ray2 data: $M_r = 1939$, $P_21_21_21$, $a = 27.809$ (7), $b = 21.712$ (6), $c = 15.333$ (4) Å, $U = 9258$ (2) Å³, $Z = 4$, $D_m = 1.381$ (15), $D_x = 1.401$ Mg m⁻³, Cu $K\alpha$ radiation, $\lambda = 1.5418$ Å, $\mu = 25.9$ mm⁻¹, $F(000) = 4024$, $T = 277.0$ (5) K, final $R = 0.136$ for 5621 significant reflections. The orientation of the coenzyme molecule in these refined models is rotated in the unit cell by approximately 5° (about an axis close to the Co atom) with respect to the orientation observed in the original structure determination [Lenhart (1968), *Proc. R. Soc. London Ser. A*, **303**, 45-84]. One of the side chains of the corrin ring (c side chain) is disordered between two extreme positions. All the H- and D-atom positions for the coenzyme molecule and approximately 65% of the solvent D atoms were located from the neutron difference Fourier maps. Of the eleven methyl groups present, six are well ordered and five disordered. An acetone molecule (with partial occupancy) was

* Author to whom correspondence should be addressed. Present address: Medical Foundation of Buffalo, Inc., 73 High Street, Buffalo, NY 14203, USA.

located in the solvent regions in both the neutron and X-ray analyses. 109, 57 and 55 solvent sites were included in the neutron and two X-ray models respectively. A brief description of the solvent analysis and an example of the solvent structural interpretation are given.

Introduction

The interactions between water and macromolecular systems such as proteins are not well understood, especially at the molecular level. Many different techniques (diffraction, spectroscopy, calorimetry, computer simulations, quantum mechanics and accessibility calculations) have been used in a concerted effort to try to gain an understanding of these interactions (Finney, Goodfellow, Howell & Vovelle, 1986; Edsall & McKenzie, 1983; Cooke & Kuntz, 1974; Kuntz & Kauzmann, 1973). Despite this, our knowledge is still poor and we are strongly in need of high-quality structural data to help improve our understanding.

The hydrated crystal structure of vitamin B₁₂ coenzyme provides an excellent system of intermediate size in which the distribution of the solvent molecules and their general interactions at the interface can be studied to atomic resolution using diffraction methods. 17–18 water molecules per coenzyme molecule (68–72 per unit cell) were estimated to be present from density measurements in the original structure determination (Lenhart & Hodgkin, 1961; Lenhart, 1968). To formulate an accurate model of the water O atoms and their associated H atoms, a detailed neutron structural analysis of this system has been undertaken to below 1.0 Å resolution. To assist in the interpretation, a parallel X-ray study was also undertaken. This latter study was necessary as the earlier work was at lower resolution (approximately 1.2 Å) and certain structural differences were observed in the crystals grown from D₂O rather than H₂O.

In this article, we describe the details of the collection and analysis of the neutron and X-ray diffraction data and subsequent refinement of the respective structural models. A brief summary of the solvent refinement is also given: a detailed analysis of the solvent structure in terms of the relative positions, water networks, water–water and water–coenzyme interactions, disorder and possible movements of the solvent molecules is given elsewhere (Savage, 1986).

The coenzyme molecule consists of a corrin nucleus with a Co atom situated at the centre forming bonds to four N atoms of the inner nucleus (see Fig. 1). Several side chains of biological interest in the context of water–protein and water–DNA interactions are attached to the outer atoms of the nucleus. These include: three acetamides, three propionamides, one propionic acid group and eight methyl residues. The

propionic acid is linked to a phosphate group which in turn forms part of the nucleotide, containing a benzimidazole base. This latter group is linked through the NB3 position to the Co atom. On the opposite side of the Co atom, a 5'-deoxyadenosine nucleoside is situated with the C5' position of the ribose moiety forming a direct link to the Co atom.

Data collection and processing

The crystals of coenzyme B₁₂ were grown using essentially the same method as for the original structure determination (Lenhart, 1968). The coenzyme was dissolved in water (D₂O) and volumes of between 0.6 and 0.8 ml of this solution were placed in 1 ml tubes which were then filled with acetone. The tubes were covered with dialysis tubing and placed in an acetone bath which was subsequently stored in a dark place to prevent light contact with the coenzyme. To reduce the large incoherent neutron-scattering component from the H atoms, the D isotope of water was used throughout all the crystal preparations. The density of the D₂O crystals was determined to be 1.381 (15) Mg m⁻³ at room temperature by flotation in an acetone/bromoethanol mixture. The measured density obtained in the original X-ray structural analysis was 1.355 (10) Mg m⁻³ where the crystals were grown in H₂O. Table 1 shows the results of the calculations of the expected number of solvent molecules for the H₂O and D₂O crystals from the density measurements assuming the bulk density of D₂O. The difference of 1–2 water molecules can be accounted for by the inclusion of an acetone molecule (found during the analysis of the solvent structure) in the D₂O solvent.

A summary of the data collection is given in Table 2.

1. Neutron data

Three-dimensional data were collected using a computer-controlled four-circle diffractometer (D8) installed on a thermal-neutron beam tube in the reactor hall of the Institut Laue-Langevin (Grenoble). One crystal with approximate dimensions 2.0 × 2.0 × 1.5 mm was used for the data collection. The crystal was maintained at a temperature of 279 (1) K by mounting it in a specially constructed constant-temperature device (Simms & Timmins, 1981) installed on the φ circle of the diffractometer. A wavelength of 1.67 Å was selected by using a pyrolytic-graphite monochromator crystal. A full complement of unique reflections up to 61° in θ was collected in a series of eight shells. $h = 0\text{--}29$, $k = 0\text{--}22$, $l = 0\text{--}16$. 1223 symmetry-equivalent reflections over the θ range 0–45° were also recorded. Up to 30° in θ , the intensities were measured by $\omega/x\theta$ scans (where $0 < x < 1$) of 34 steps, while for the remaining reflections $x = 2$ was

used. Each reflection required 4–5 min counting time, but for the θ range 55–61° the counting times were doubled in order to improve counting statistics. The upper limit of $\theta = 61^\circ$ approached the physical limitation of the instrument and the data collection was halted, although the quality of the data was still very good: 71% of the reflections with $55 < \theta < 61^\circ$ were observed above the 2.58σ significance level.

The intensities of two standard reflections were monitored every 30 reflections and no significant variations [$>2\sigma(I)$] were observed. 7994 intensities were recorded, of which 6339 were unique observations.

The intensity profiles were analysed by the algorithm of Lehmann & Larson (1974), incorporated in the *COLLSN* routine (Lehmann & Wilson, 1982) which was used for data reduction. Although empirical absorption measurements (North, Phillips

& Mathews, 1968) were recorded, the variations over the φ angle were less than $2\sigma(I)$ and consequently no absorption corrections were made. 7994 reflections were scaled and merged (Hamilton, Rollett & Sparks, 1964) giving a residual of 1.7%. The contamination of the intensities by the half-wavelength, $\lambda/2$, component was estimated from intensity measurements for reflections which were expected to be systematically absent and these measurements were compared with the values of their first-order harmonics: $I_0(hkl)(\lambda/2)/I_0(hkl)(\lambda)$; for example, the (300) ($\lambda/2$) and the (600) (λ). The $\lambda/2$ component was estimated to be 2.4%. The intensities were corrected and only relatively small changes, in most cases less than 0.1%, were produced in the resulting structure factors. The final neutron data set comprised 5994 reflections, of which 5601 had $F_o > 3\sigma(F_o)$.

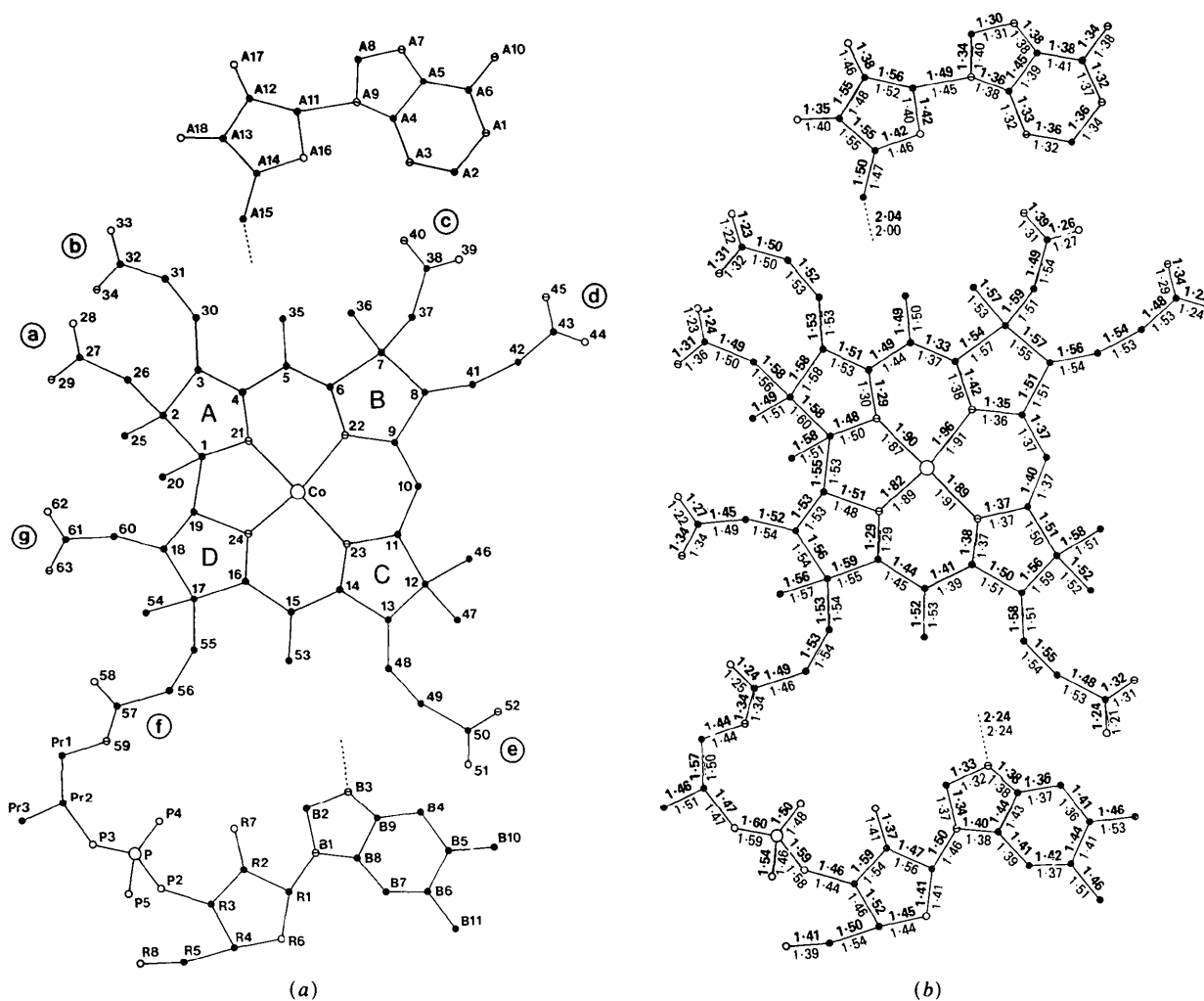


Fig. 1. (a) Atomic numbering and distribution of atom types in the vitamin B₁₂ coenzyme molecule: filled circles are C atoms, small open circles are O atoms, the two larger open circles are the Co and P atoms and the open circles with horizontal lines are N atoms; (b) interatomic distances (Å) for the coenzyme B₁₂ molecule: bold numbers (above bonds) refer to the neutron model while lighter numbers (below bonds) refer to the X-ray1 model.

Table 1. *Density calculations for coenzyme B₁₂ crystals grown from H₂O and D₂O solutions*

	H ₂ O (Lenhert, 1968)	D ₂ O
Measured density, D_m (Mg m ⁻³)	1.355 (10)	1.381 (15)
Molecular weight of coenzyme molecule	1580	1599
Molecular weight of coenzyme + solvent	1899	1934
Molecular weight of solvent in one asymmetric unit	319	335
Expected number of water molecules per asymmetric unit	17.7 (319/18)	16.7 (334/20)

Table 2. *Summary of data collection; space group = P2₁2₁2₁, Z = 4*

	Neutron	X-ray1	X-ray2
Number of crystals	1	1	2
Crystal volume (mm ³)	6.0	0.1	0.1
Temperature (K)	279 (1)	277 (1)	277 (1)
Wavelength (Å)	1.67	1.5418	1.5418
Resolution (Å)	0.95	1.10	0.92
θ_{\max} (°)	61	44	57
Total number of measurements	7994	7491	~13900
Number of independent reflections	6339	4390	~6100
Number of equivalent reflections	1233	~2500	~5000
Number of hkl with $F_o > 3\sigma(F_o)$	5601	4152	5621
Residual for symmetry-related hkl (%)	1.7	4.6	6.5

2. X-ray data

X-ray diffraction data were collected on a Hilger & Watts Y290 four-circle diffractometer using Ni-filtered Cu $K\alpha$ radiation. A cold-temperature apparatus mounted on the χ circle was used to maintain a stream of cold air at 277.0 (5) K over the crystal. 15 reflections with $\theta_{\max} = 40^\circ$ used for measuring lattice parameters. A complete set of data, X-ray1, extending to a resolution of 1.1 Å, and a partial set to 0.9 Å were measured from a single crystal of approximate dimensions 0.5 × 0.3 × 0.3 mm using the moving-window technique (Tickle, 1975). $h = 0-24$, $k = 0-22$, $l = 0-16$. For each reflection 60 steps were scanned (0.02° per step, 45 peak, 15 background, 1 s count time). Two equivalents were collected to $\theta = 30^\circ$, and one equivalent between $\theta = 30$ and 45° . Data collection was halted at this stage owing to a mechanical limitation in the design of the low-temperature device. After modifications to the device, a second set of data, X-ray2, was collected to a resolution of 0.92 Å. $h = 0-29$, $k = 0-23$, $l = 0-16$. Two crystals were required and it was found that their cell parameters differed significantly from each other and also from those of X-ray1, neutron and the original structure determination, 1968X-ray (Lenhert, 1968). This highlighted the problem of variable parameter size with different batches of crystal, despite vigorous attempts to use a uniform crystal growth and deuteration procedure. At a subsequent stage the changes in cell parameters could be explained in terms of small differences in the molecular packing as well as the solvent structure (see part 2 of *Results and Discussion*). For this high-resolution data, X-ray2, the cell parameters were averaged and the intensity data from both crystals combined. Two equivalents to $\theta = 57^\circ$ were collected from the two crystals: to $\theta = 45^\circ$ for the first crystal (51 steps of

0.02°, 34 steps for peak) and between $\theta = 40$ and 57° for the second (60 steps of 0.01°, 6 s stationary background counts).

Four standard reflections were recorded every 50 reflections and for X-ray1 their intensity sum decreased by less than 5% during data collection, while for X-ray2 the corresponding sum for the two crystals used was about 10%. Empirical absorption corrections were made for all three crystals using the method of North *et al.* (1968). Both sets of data, X-ray1 and X-ray2, were corrected for Lorentz, polarization, absorption and crystal decay. Symmetry equivalents were merged for each data set with residuals of 4.6 and 6.5%, to yield 4390 and 6100 independent reflections respectively.

Structure refinement

The initial phasing model used for the refinement of the neutron and X-ray1 models was based on the set of positional and thermal parameters obtained from the original structure determination (Lenhert, 1968). The refinement, based on F_o , was carried out using a combination of Fourier syntheses and reciprocal least-squares refinement techniques. The coenzyme B₁₂ molecule comprises 209 atoms (109 non-H and 100 H atoms) making it computationally difficult to refine using conventional full-matrix least squares. Thus a combination of two methods was used.

Stage 1: The models were treated as small macromolecules using the fast-Fourier-transform least-squares algorithm, FFTLS (Agarwal, 1978; Isaacs & Agarwal, 1978), in order to accelerate the convergence of the refinement. The structure factors were weighted using the expression $w = (2 \sin \theta / \lambda)^p$, where the weighting parameter, p , was initially chosen as -1.5 (to give weight to low-resolution data), but set to 0.0 near the end of the refinement (*i.e.*, unit weights). Only isotropic temperature factors were included at this stage.

Stage 2: The models were treated as large 'small molecules' using conventional partially blocked full-matrix least squares, BFMLS, in which the molecule was divided into overlapping segments. Each reflection was weighted during the refinement using the following scheme (Cruickshank, 1965): $w = [1 - \exp(-a \sin \theta / \lambda)] / (b + |F_o| + c|F_o|^2)$. The coefficients a , b and c were adjusted to give appropriate weighting analyses, and in the final refinement cycles were chosen as follows: X-ray, $a = 20$, $b = 100$, $c = 0.005$; neutron, $a = 17$, $b = 100$, $c = 0.0005$. During the latter cycles of refinement, anisotropic thermal vibration parameters were included for each of the coenzyme atoms; solvent atoms were treated isotropically. Corrections for extinction were made empirically using a factor g , in the expression $[1 + g|F_c|]^{-1/2}$; in the final refinement cycles values of $g = 2.196 \times 10^{-4}$ and 1.147×10^{-6} were used for the neutron and X-ray data respectively.

Table 3. *Course of refinements*

	Refinement method*	R factor
X-ray1		
Stage 1	Parallel refinement using: (1) FFTLS, DF's (2) BFMLS, DF's	0.370
0.430	19 cycles (1) 4 cycles (2)	↓
↓	Converged with r.m.s. deviation between models of 0.06 Å	↓
0.146	24 solvent sites assigned	0.149
Stage 2	BFMLS 4 cycles (anisotropic U's), DF's	0.151
	57 solvent sites assigned	0.088
X-ray2		
Stage 1	FFTLS, DF's	0.210
	26 solvent sites assigned	0.163
Stage 2	BFMLS, DF's	0.170
	3 cycles (isotropic U's) 3 cycles (anisotropic U's)	↓
	55 solvent sites assigned	0.136
Neutron		
Stage 1	FFTLS, DF's	0.408
(4088 data)	Parallel refinement starting from 2 different models: (1) 1968X-ray (2) X-ray1 FFT model	↓
	Converged with r.m.s. deviation of 0.08 Å	0.175
	90 H,D of coenzyme located 11 solvent sites assigned	↓
(5260 data)	All 100 H,D of coenzyme located 32 solvent sites assigned	0.143
Stage 2	BFMLS (2 cycles; isotropic U's), DF's	0.172
(5994 data)	4 cycles (anisotropic U's)	0.131
	Final model: 212 coenzyme positions 109 solvent positions	0.085

* FFTLS: fast Fourier least-squares; BFMLS: partially blocked full-matrix least squares; DF's: difference Fourier syntheses.

Table 3 shows the course of the refinements of the neutron and X-ray models. The X-ray scattering factors were obtained from *International Tables for X-ray Crystallography* (1974); values for neutral atoms were used except for Co which was corrected for the dispersion component of the anomalous scattering (Co^{3+} , $f' = -2.454 e$). Neutron scattering lengths were as follows: C = 6.65, N = 9.40, O = 5.80, Co = 2.80, P = 5.50, H = -3.74 and D = 6.67 fm (Bacon, 1980).

1. X-ray1

In order to test for the convergence and consistency of the two least-squares methods, FFTLS and BFMLS, the first X-ray model was initially refined using the two methods in parallel. In terms of computing CPU time per cycle of refinement, the FFTLS least-squares program required 2 min whereas the conventional BFMLS program required about 60 min on an IBM 370/165 computer. The two independent refinements (R factors: 0.146 for FFTLS and 0.149 for BFMLS) were seen to converge to almost the same model. The r.m.s. deviation between the two sets of

coordinates was 0.06 Å vindicating the usefulness of the Agarwal technique in high-resolution macromolecular refinement.

However, the r.m.s. deviations of these two models from the original phasing model (1968X-ray model) were 0.33 and 0.34 Å respectively (max. = 1.85 Å) indicating that the new X-ray structure (crystals grown from D₂O and at a lower temperature, 277 K) is somewhat different. The whole molecule appears to have rotated in the unit cell by approximately 5° (see later). The X-ray2 and neutron models also differ in a similar way from the 1968X-ray model.

Refinement was continued using BFMLS least squares with the inclusion of anisotropic thermal parameters. Several difference Fourier maps were calculated to analyse the solvent regions in which 53 water O sites and four acetone sites were assigned: the final R factor was 0.088, $wR = 0.114$, $S = 0.69$. Maximum LS shift $< 0.5\sigma$ and maximum and minimum heights in final difference map = 0.8 and $-0.9 e \text{ \AA}^{-3}$.

2. X-ray2

The model obtained from the X-ray1 refinement was used as the initial phasing model. The final R factor was 0.136 and 51 water O sites and four acetone sites were included in the model. $wR = 0.196$, $S = 1.47$. The higher R factor than that for the X-ray1 data may be largely due to the combination of data from the two crystals with slightly different molecular packings (see part 2 of *Results and discussion*). Maximum LS shift $< 0.5\sigma$ and maximum and minimum heights in final difference map = 0.8 and $-0.9 e \text{ \AA}^{-3}$.

3. Neutron

The refinement proceeded in three steps due to processing of the data at different times: (1) using 4088 data and FFTLS; (2) using 5260 data and FFTLS; and finally (3) 5994 data (all data to 0.95 Å) and BFMLS (see Table 3).

In step (1) two independent starting models were used: (a) the 1968X-ray model and (b) the X-ray1 model (from stage 1). The resulting two models were seen to converge with a r.m.s. deviation of 0.08 Å ($R = 0.175$). 90 of the 100 coenzyme H and D atoms were located. In step (2), all the missing H atoms (attached to methyl groups) were located along with 32 solvent sites. In step (3), anisotropic thermal parameters were included for all the coenzyme atoms and from the difference maps 109 solvent sites were assigned: final $R = 0.085$, $wR = 0.109$, $S = 2.04$. Maximum LS shift $< 0.5\sigma$ and maximum and minimum heights in final difference map = 0.28 and -0.35 fm.

Results and discussion

The final set of coordinates and thermal parameters (including solvent sites) obtained from the refinement of the neutron and X-ray1 data are listed in Table 4.*

1. Bond lengths and angles

The bond lengths and angles calculated from the neutron and X-ray1 coordinates are given in Figs. 1 and 2. The values obtained from the X-ray2 model (deposited) are, within experimental error, the same

* A complete listing of the structure factors, atomic coordinates, thermal parameters, occupancies, together with all the individual bond distances and angles for the three models, neutron, X-ray1 and X-ray2, have been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 43601 (127 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

as those for the neutron and X-ray1 models; the e.s.d. ranges (\AA) for the models are as follows:

	Range	Average
Neutron (non-H atoms)	0.007-0.021 \AA	0.012 \AA
Neutron (H atoms)	0.015-0.065	0.027
X-ray1	0.007-0.019	0.014
X-ray2	0.012-0.041	0.024

Table 5 lists the corresponding bond lengths and angles around the Co atoms in the three refined models (and also the 1968X-ray model), and the statistics of the chemically equivalent bonds for the neutron model; the bond lengths were not corrected for thermal motion.

Generally, the mean lengths obtained for the standard bond types were seen to be within 2σ of their accepted values. However, some large deviations of up to 0.08\AA (6σ) were observed for several individual bond lengths, particularly for some of the single C-C

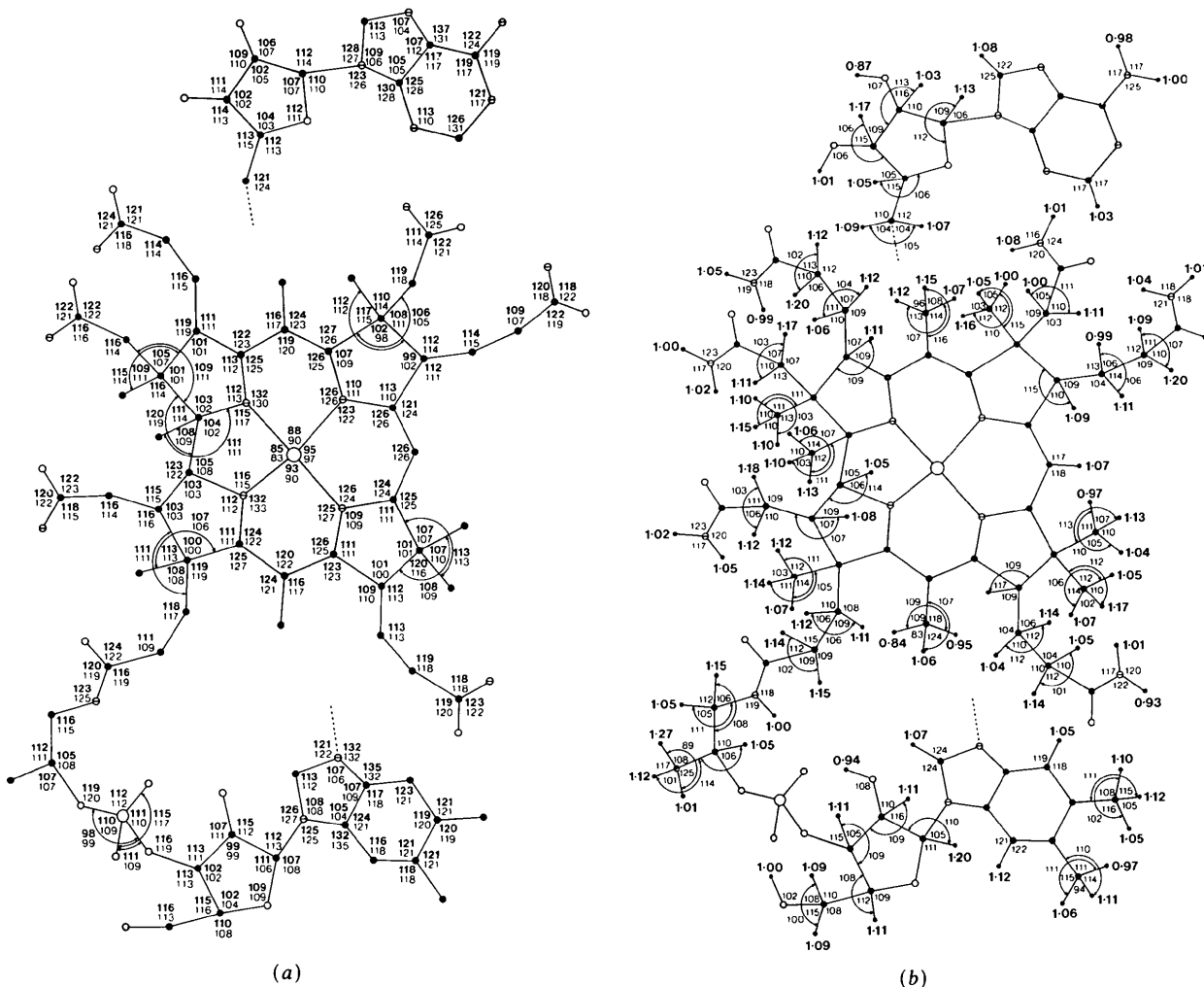


Fig. 2. (a) Interatomic angles in the coenzyme B_{12} molecule: bold numbers refer to the neutron model and lighter numbers to the X-ray1 model; (b) interatomic distances (bold numbers) and angles (lighter numbers) involving the H and D atoms of the neutron model. The distances are in \AA and the angles in degrees.

Table 5. Bond lengths (Å) and angles (°) around the Co atom and coenzyme bond-length statistics (Å) for the neutron model

See text for average e.s.d.'s for bond distances. Average e.s.d.'s of bond angles around the Co atom are: $\sim 0.8^\circ$ for neutron, $\sim 0.5^\circ$ for X-ray1, $\sim 1.0^\circ$ for X-ray2 and $\sim 2.0^\circ$ for 1968X-ray.

		Neutron	X-ray1	X-ray2	1968X-ray
Co-N21		1.90	1.87	1.82	1.92
Co-N22		1.96	1.91	1.87	1.91
Co-N23		1.89	1.91	1.87	1.97
Co-N24		1.82	1.89	1.87	1.98
Co-NB3		2.24	2.24	2.22	2.24
Co-CA15		2.04	2.00	1.98	2.05
N21-Co-N22		88	90	89	91
N21-Co-N23		176	172	172	170
N21-Co-N24		85	83	83	81
N21-Co-CA15		92	92	95	95
N22-Co-N23		95	97	97	98
N22-Co-N24		171	172	172	172
N22-Co-CA15		82	84	83	86
N23-Co-N24		93	90	91	90
N23-Co-CA15		91	92	91	89
N24-Co-CA15		95	92	95	93
N21-Co-NB3		91	91	90	93
N22-Co-NB3		88	89	89	88
N23-Co-NB3		86	85	85	85
N24-Co-NB3		96	94	94	95
CA15-Co-NB3		169	173	171	170

Bond type		Number	Mean length (Å)	R.m.s. deviation (Å)	Average e.s.d. from refinement (Å)	Max. deviation from mean (Å)
Co-N	Rings A, D	2	1.860	0.059	0.017	0.042
	Rings B, C	2	1.925	0.050	0.017	0.035
N-C(<i>sp</i> ²)	Rings A, D	2	1.289	0.004	0.009	0.003
	Rings B, C	4	1.379	0.032	0.008	0.045
C(<i>sp</i> ²)-C(<i>sp</i> ²)	Inner ring	6	1.407	0.056	0.010	0.078
C(<i>sp</i> ²)-C(<i>sp</i> ³)	Corrin nucleus	8	1.520	0.031	0.011	0.030
	Side chains	9	1.480	0.017	0.014	0.025
C(<i>sp</i> ³)-C(<i>sp</i> ³)	Corrin nucleus	20	1.556	0.027	0.011	0.061
	Side chains	27	1.540	0.036	0.013	0.083
C=O	Amide groups	6	1.243	0.017	0.016	0.026
C-N	Amide groups	6	1.335	0.029	0.014	0.058
C-O	Ribose	9	1.417	0.037	0.015	0.064
C(<i>sp</i> ²)-C(<i>sp</i> ²)	Phenyl	6	1.414	0.027	0.013	0.054
O-D	Hydroxyl	4	0.992	0.037	0.029	0.054
N-D	Amide	15	1.012	0.034	0.022	0.086
C-H(<i>sp</i> ² , <i>sp</i> ³)		20	1.087	0.045	0.023	0.111
C-H	Methylene	28	1.108	0.052	0.025	0.090
C-H	Methyls	33	1.082	0.078	0.034	0.241
CH, CH ₂ , CH ₃		81	1.092	0.063	0.028	0.241

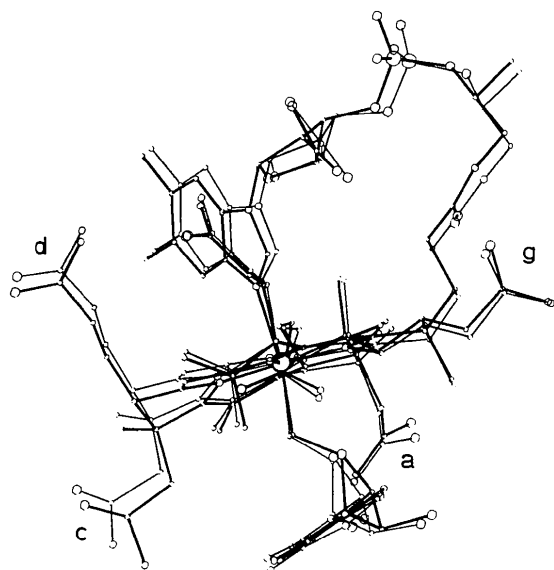
C-H length of 1.09 Å. The largest deviations are associated with the more disordered groups of C35, C53 and CPR3. As shown in Fig. 2(b), the angles in the ND₂ groups are all around 120(4)°, with the D-N-D angle slightly smaller (by 2-4σ) than the other two angles.

2. Rotation of the coenzyme molecule in the unit cell

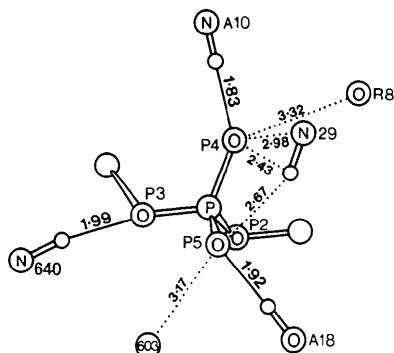
When compared with the 1968X-ray model crystallized from H₂O at room temperature, the three refined models grown in D₂O at 277 K were seen to be rotated by approximately 5° about an axis which is almost parallel to the *c* axis of the unit cell and passing close to the Co atom (Fig. 3a). The r.m.s. deviations (Å) of the positional parameters between the four models are as follows (maximum deviations in brackets):

	X-ray1	X-ray2	Neutron
1968X-ray	0.38 (1.68)	0.33 (1.63)	0.27 (1.68)
X-ray1		0.08 (0.17)	0.09 (0.20)
X-ray2			0.07 (0.15)

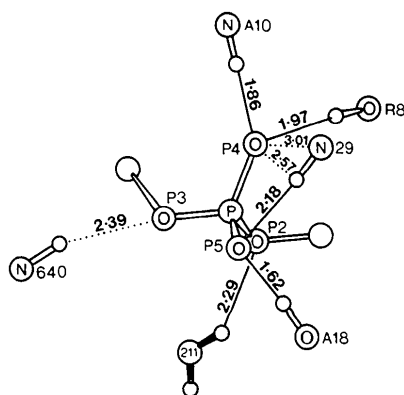
The main consequence of the reorientation in the unit cell appears to be a redistribution of the hydrogen bonds around the phosphate group. In the 1968X-ray model, the phosphate O atom, OP4, is hydrogen bonded to NA10 (Fig. 3b) and also forms weaker contacts to N29 and OR8, whilst OP5 is bonded to OA18 and water W603. The N29 atom appears to form two weak hydrogen-bond contacts to two of the phosphate O atoms, one to OP4, H...O = 2.43 Å (although N29...OP4 = 2.98 Å), and the other to OP5, H...O = 2.67 Å (N29...OP5 = 3.67 Å). In the other three models (Fig. 3c), OP4 forms strong hydrogen bonds to NA10 and OR8, while N29 again forms a 'three-centred interaction' with OP4 and OP5. However, this time it forms a shorter H...O contact with OP5 (2.18 Å) than with OP4 (2.57 Å) even though the N29...OP4 distance (3.01 Å) is shorter than N29...OP5 (3.17 Å). OP5 makes stronger hydrogen bonds to N29 and OA18 but a weaker contact to the water at position W603 (OP5...W603 = 3.40 Å). N29 appears to form a bifurcated hydrogen-bond



(a)



(b)



(c)

Table 6. *Regional comparisons of the temperature factors, \bar{u}^2 (\AA^2), for the neutron and three X-ray models [B values (\AA^2) in parentheses]*

Model	Inner corrin ring	Outer corrin ring	Inner side chains	Outer side chains
X-ray1	0.014 (1.1)	0.022 (1.7)	0.028 (2.2)	0.045 (3.6)
X-ray2	0.020 (1.6)	0.036 (2.8)	0.042 (3.4)	0.059 (4.6)
Neutron (non-H only)	0.024 (1.9)	0.034 (2.7)	0.041 (3.2)	0.061 (4.8)
1968X-ray	0.041 (3.2)	0.049 (3.9)	0.058 (4.6)	0.077 (6.1)
Neutron (H and D only)	0.043 (3.4)	0.047 (3.7)	0.064 (5.1)	0.089 (7.0)

arrangement to OP4 and OP5 and small movements of the amide H atom appear to be permitted without large disruptions in the overall hydrogen-bonding structure.

The H \cdots O hydrogen-bond length between OP3 and N640 of the *c* side chain is increased from 1.99 Å in the 1968X-ray model to 2.39 Å in the neutron model. This side chain in the latter model is disordered (see § 5 below), which probably results from the weakening of the H bond.

3. Thermal parameters

The thermal parameters of the individual atoms were generally seen to increase from the centre of the corrin ring outwards to the more peripheral side chains. Table 6 lists the average temperature factors over the various regions of the molecule for all four models. The isotropic equivalent temperature factors are plotted in Fig. 4. The thermal ellipsoids at the 50% probability level for the individual atoms in the neutron model are drawn (Johnson, 1976) in Fig. 5.

Six of the eleven methyl groups appear to be well ordered (Fig. 6a) and possess reasonably shaped ellipsoids although some are flattened with their smallest axis of vibration parallel to the C–H direction. These ordered methyls are attached to tetrahedral C atoms and deviations of up to 20° occur in the torsion-angle value expected for an ideal staggered conformation. This is probably due to close repulsive H \cdots H interactions which appear to be minimized with respect to maximizing the local H \cdots H distances at the expense of the torsional contacts, X \cdots H (of X–C–C–H).

The remaining five methyl groups were seen to have elongated neutron density maxima (Fig. 6b), indicating extensive thermal libration or disorder. Four of these methyls, C35, C53, CB10 and CB11, are attached to planar groups and can undergo small rotations until they encounter short H \cdots H contacts. The fifth methyl, CPr3, is located next to a symmetry-related position of the disordered *c* side chain. (Although not shown, the CPr3 position is located ~3.3 Å above N640 in Figs. 3b and 3c.) When the latter group occupies the alternative N640 position (see below), CPr3 must rotate to prevent a very close D \cdots H–(CPr3) *ca* 1.6 Å.

Fig. 3. (a) Projection of the non-H atoms of the X-ray1 model (heavy line) and 1968X-ray model (fine line) onto the *ab* plane of the unit cell. Some side-chain labels are included. (b) and (c) rearrangement of the hydrogen-bonding network around the phosphate group of the coenzyme B₁₂ molecule: (b) 1968X-ray model; (c) neutron model. The H-atom positions in the 1968X-ray model are calculated positions (X–H = 1.0 Å). Distances in Å.

Inspection of an X-ray1 difference map, in which only the coenzyme non-H atoms were phased, revealed electron density maxima of between 0.2 and 0.5 e Å⁻³. The majority of these peaks were located at the same positions as the H and D sites in the neutron model, corresponding to 74 (62 for the X-ray2 model) of the 100 coenzyme H and D atoms: including 25 of the 33 methyl H atoms. Fig. 6(c) shows the electron density over methyl C20 and C46.

4. Conformation of the side chains

All the side chains (apart from the *c* acetamide) adopt the same relative conformations in the four different models. When the six terminal atoms of each amide group are included in the calculation of the least-squares plane, the largest deviation, 0.141 Å, from the respective planes occurs for D198 in the *d* side chain. The r.m.s. deviation for the seven amide

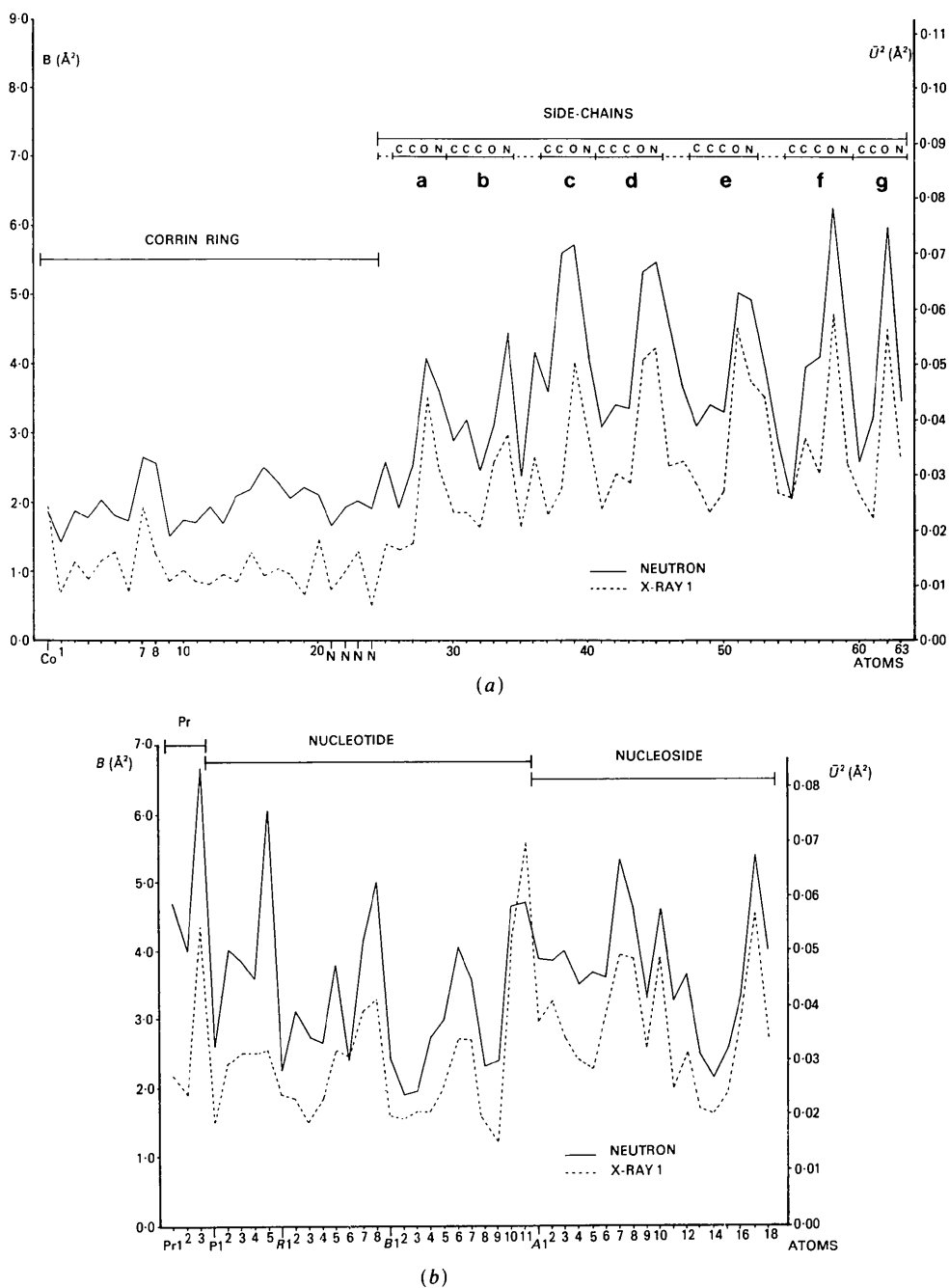


Fig. 4. Plot of the individual equivalent isotropic thermal parameters for the atoms in the coenzyme B_{12} molecule for the neutron and X-ray1 models: (a) corrin ring and side chains, and (b) nucleotide and nucleoside moieties.

planes is 0.044 Å, which is three to four times that of the positional e.s.d.'s. However, when only the four non-H atoms are included, there is a general improvement in the planarity of the amide groups: the average r.m.s. deviation is 0.014 Å, comparable to the average positional e.s.d. of 0.013 Å. Several of the D atoms deviate significantly from their respective amide planes (non-H atoms), for example, D192 and D198 deviate by 0.110 and 0.324 Å and appear to be displaced in the direction of their hydrogen-bonded acceptor atoms.

5. Disorder of the *c* side chain

In the X-ray1 and X-ray2 models, the *c* side chain was found to occupy an alternative conformation to that in the 1968X-ray structure. In the neutron model, however, both conformations were present with the largest difference occurring in the position of the amide group: two extreme positions were assigned, N40 and N640, which were *ca* 1.8 Å apart. The latter position corresponds to the conformation present in the 1968X-ray model. Fig. 7 shows neutron density

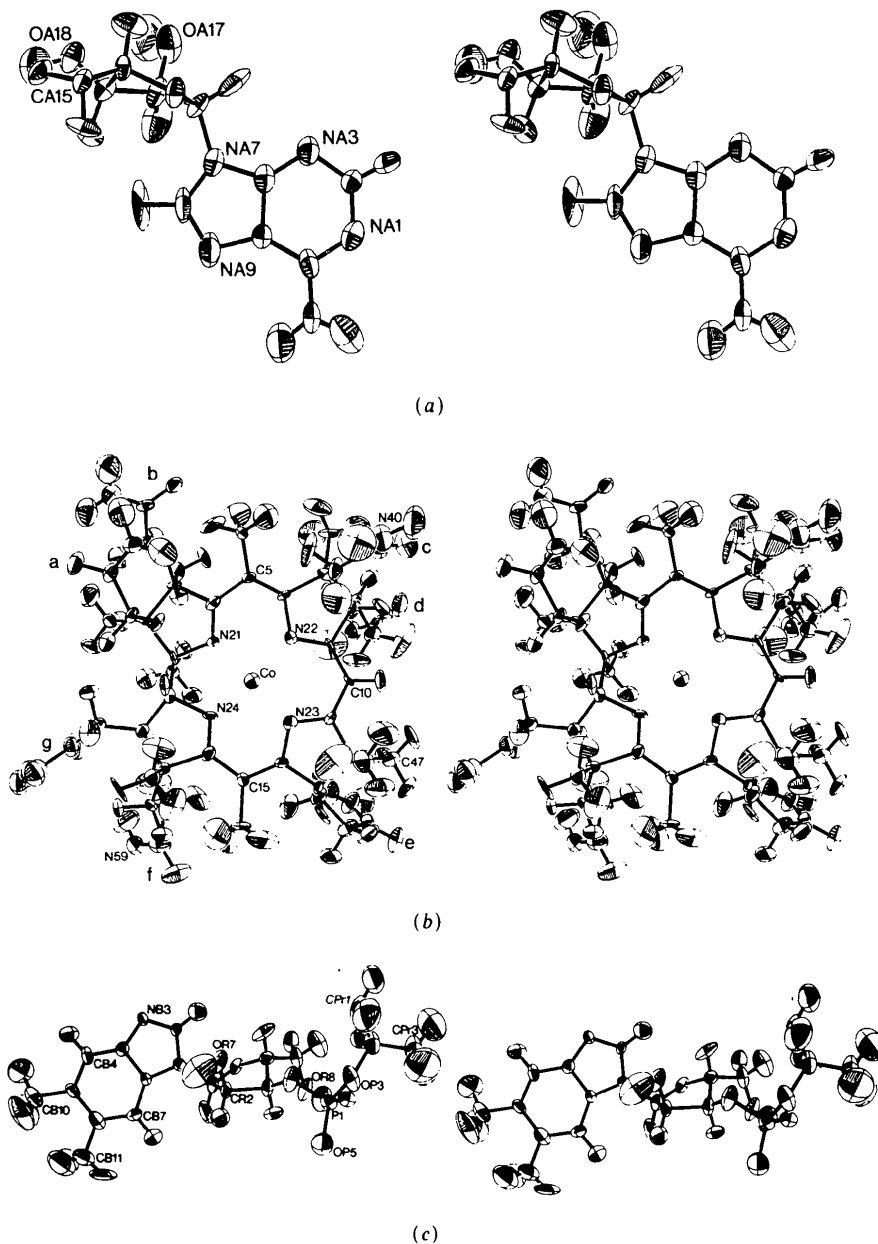


Fig. 5. Stereoscopic drawings (*ORTEP*; Johnson, 1976) of the coenzyme B₁₂ molecules for the neutron model; ellipsoids are drawn to enclose 50% probability: (a) nucleoside moiety; (b) corrin ring and side chains; (c) nucleotide moiety. The Co atom and three of the H atoms (H161 of C20, H167 of C35 and H141 of CR2) form non-positive-definite matrices and are included as isotropic atoms.

from difference maps over (a) the whole side chain, (b) the C38, N40 and N640 regions in more detail and (c) the amide planes of the two main conformations. The occupancy values for the different N-atom positions among the four models were assigned as follows (refined in the neutron model):

Model	N40	N640	Temperature (K)
X-ray1	1.00	—	277
X-ray2	1.00	—	277
Neutron	0.74	0.26	279
1968X-ray	—	1.00	294

No significant disorder was observed for the carbonyl O39, which would otherwise be expected if the disorder simply involved a rotation around the C37–C38 bond. Inspection of the region around O39 revealed that, if it moved as a result of a C–C rotation, the new position would approach the adjacent CPr3 methyl group (symmetry related) to within 3.0 Å (less than the van der Waals contact, *ca* 3.5 Å). The thermal parameters of atoms C7 and C8 of the B pyrrole

ring, to which the side chain is attached, are higher than expected when compared with those of similar positions on the remaining pyrrole rings (see Fig. 4a). Thus, instead of a C–C bond rotation, the disorder of this side chain appears to be more complicated with the O39 position remaining almost fixed, while the movement of the N atom between two extreme positions is accommodated by small changes in the positions and torsion angles of the rest of the side chain and pyrrole ring (Fig. 7a). These movements (a) maintain the planarity of the amide group and (b) prevent the carbonyl O atom from approaching the nearby methyl group too closely.

Continuous neutron density is seen between the N40 and N640 sites (Fig. 7b) indicating that the side chain is disordered and may be undergoing librational motion. A series of sites can be placed across the continuous density (N840, 841, 842) to represent alternative conformations. A better model might be achieved by the inclusion of anharmonic temperature/scattering factors for the individual atoms.

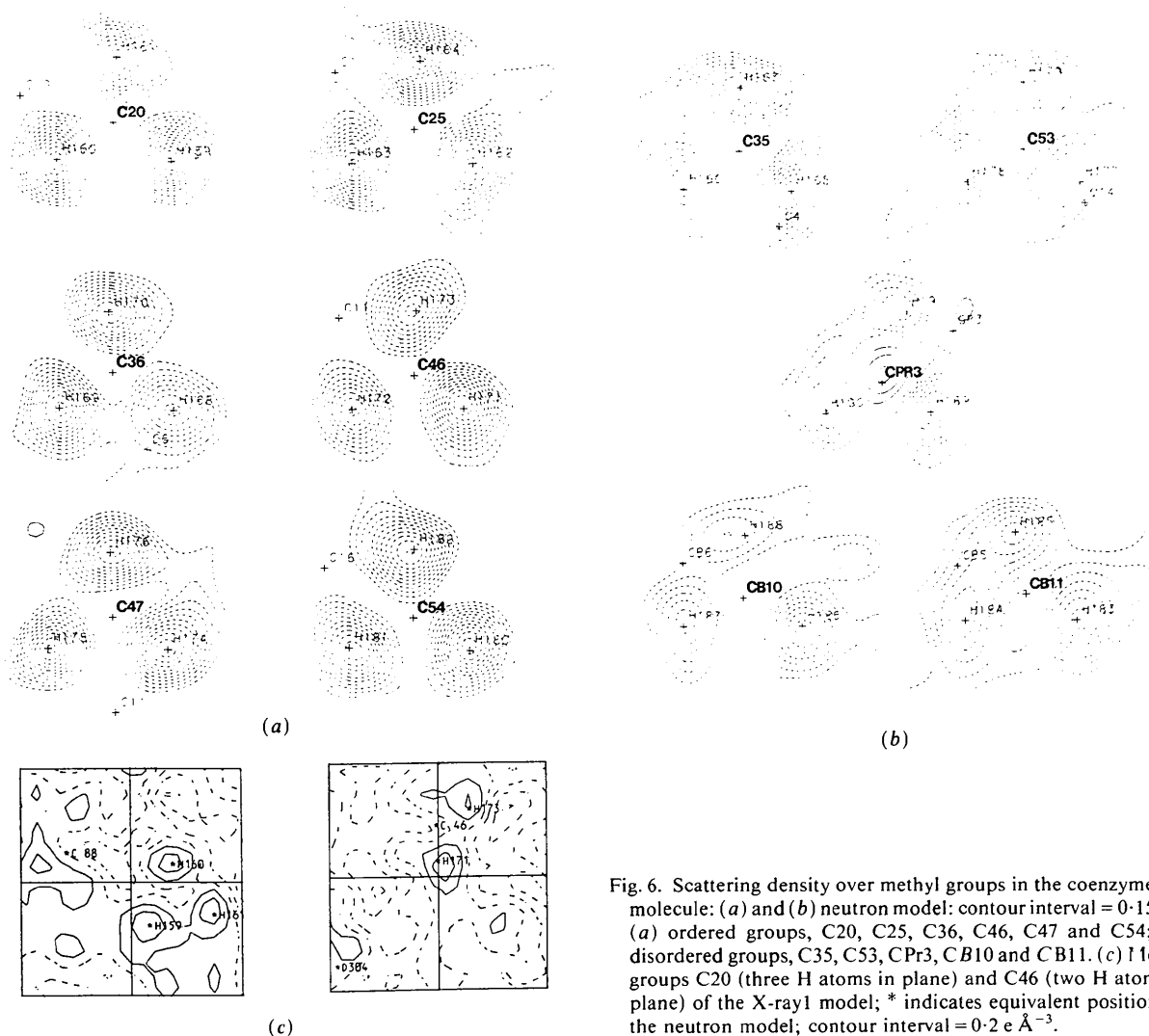


Fig. 6. Scattering density over methyl groups in the coenzyme B_{12} molecule: (a) and (b) neutron model: contour interval = 0.15 fm; (a) ordered groups, C20, C25, C36, C46, C47 and C54; (b) disordered groups, C35, C53, CPr3, CB10 and CB11. (c) Methyl groups C20 (three H atoms in plane) and C46 (two H atoms in plane) of the X-ray1 model; * indicates equivalent positions in the neutron model; contour interval = $0.2 \text{ e} \text{ \AA}^{-3}$.

6. Hydrogen-deuterium exchange

In order to obtain an estimate of the H/D exchange, which from the D₂O/coenzyme concentration was calculated to be about 99% (10 mg coenzyme B₁₂ per 1 ml of 99.8% D₂O), the thermal and occupancy parameters for the D-atom sites were varied in several cycles of least-squares refinement. The average e.s.d. of the occupancies was 0.07 and although several large deviations from unit occupancy were observed for D200(0.78), D201(0.88) and D209(0.83), only D200 deviates by more than 3σ. The average value of the refined occupancies for the 19 D-atom sites is 0.95 (7), which is in good agreement with the concentration estimate.

7. Solvent and hydrogen bonding

The details of the solvent analysis, refinement, formulation of water networks, structure and possible movements are described in a separate publication (Savage, 1986). However, to illustrate the complexity of the associated disorder problem, an example of a possible interpretation of one region of solvent density is given here, along with a brief summary of the solvent analysis and refinement.

A series of difference Fourier maps was calculated for the neutron, X-ray1 and X-ray2 models, using the

Table 7. Neutron hydrogen-bond geometries involving the D atoms of the coenzyme B₁₂ molecule

	X-D (Å)	X-D...Y (°)	D...Y (Å)	X...Y (Å)
N29-D192...NA1	1.02 (2)	168 (2)	2.14 (2)	3.14 (1)
-D193...OP5	1.00 (2)	170 (2)	2.18 (2)	3.17 (2)
N34-D194	0.99 (2)			
-D195...O44	1.05 (2)	166 (2)	1.88 (2)	2.90 (2)
N40-D196...O212	1.08 (2)	149	2.03	3.01
-D197...O211	0.99 (3)	146 (4)	2.10 (5)	2.98 (6)
N640-D796...NA7	1.14 (8)	164 (6)	2.41 (8)	3.53 (4)
-D797...OP3	0.87 (8)	164 (5)	2.39 (5)	3.24 (4)
N45-D198...O28	1.04 (2)	164 (2)	1.82 (2)	2.83 (2)
-D199	1.01 (3)			
N52-D200...O214	1.01 (2)	152	2.09	3.02
-D201...OR6	0.93 (2)	162 (2)	2.14 (2)	3.03 (1)
N59-D136...O211	1.00 (2)	152 (3)	2.29 (5)	3.21 (4)
N63-D202...O33	1.05 (2)	160 (2)	1.93 (2)	2.94 (1)
-D203...O51	1.02 (2)	165 (2)	1.84 (2)	2.84 (1)
NA10-D204...OP4	0.98 (3)	174 (2)	1.86 (2)	2.84 (1)
-D205...O39	1.00 (2)	163 (3)	2.06 (2)	3.03 (2)
OR7-D206...O216	0.94 (3)	167 (2)	1.95 (3)	2.87 (3)
OR8-D207...OP4	1.00 (3)	152 (2)	1.97 (3)	2.89 (2)
OA17-D208...O222	0.87 (3)	162 (3)	1.83 (4)	2.67 (4)
OA18-D209...OP5	1.01 (2)	179 (2)	1.62 (2)	2.63 (2)

fully refined parameters of the coenzyme molecule. Initially, all the previously assigned solvent sites were omitted from the phasing and as various portions of the solvent regions were interpreted they were included in the refinement. 109, 57 and 55 solvent sites were assigned in the neutron, X-ray1 and X-ray2

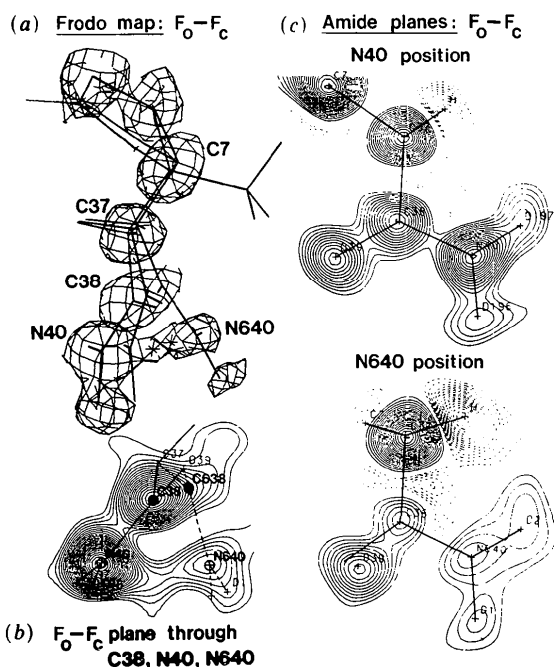


Fig. 7. Disorder of the *c* side chain (acetamide): (a) Frodo electron density map ($|F_o| - |F_c|$) over the side chain and C7 pyrrole atoms showing three possible conformations; N40, N840 and N640. (b) Contoured section ($|F_o| - |F_c|$) through the C38, N40, N640 and N640 region (interval 0.2 fm). (c) Contoured section ($|F_o| - |F_c|$) through the amide planes of the side chain in the N40 and N640 conformations (interval 0.2 fm).

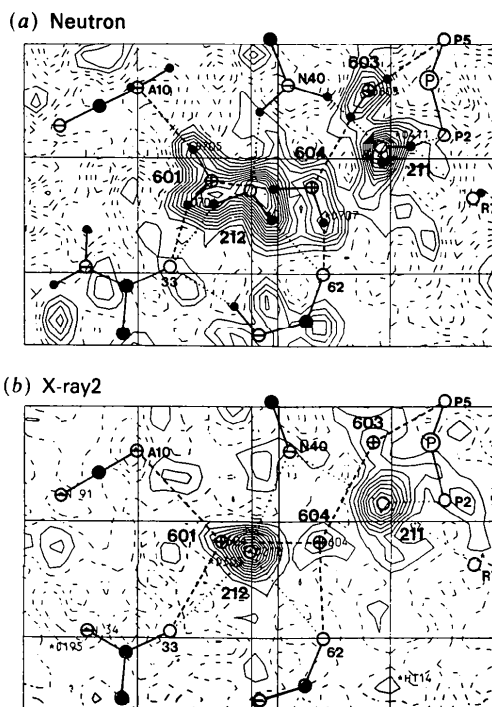


Fig. 8. Contoured sections ($|F_o| - |F_c|$ difference Fourier map) through the solvent region around the disordered *c* side chain (alternative N640 position is not shown; it lies ~ 1.8 Å behind the N40 position). Two water networks are present: network A contains waters 211 and 212, and network E contains waters 601, 603 and 604; (a) neutron map, interval 0.07 fm and (b) X-ray2 map, interval $0.1 e^{-3}$.

model, respectively, *all with partial occupancies* ranging from 0.95 down to 0.12. Four of the solvent sites in each of these models were assigned as an acetone molecule with occupancies of approximately 0.5, 0.6 and 0.6 respectively.

Fig. 8 shows the interpretation of the complex region of solvent density (from an $F_o - F_c$ difference map) situated around the disordered *c* side chain. Two different water networks are observed in the overlapping density and these correspond to the two extreme disordered positions for the amide N atom:

	Network	Water sites
N40	A	211, 212
N640	E	601, 603, 604

The N640 position is not shown in Fig. 8, but lies ~ 1.8 Å behind the N40 position.

The majority of the numerous hydrogen bonds in this crystal involve water molecules and a complete listing of all the values obtained for the hydrogen-bonding parameters, $D \cdots Y$, $X-D \cdots Y$ angles *etc.*, is given in Savage (1986). The values for the hydrogen-bond geometries involving only the coenzyme D atoms are listed in Table 7.

Use of the facilities of the Institut Laue-Langevin (Grenoble, France) is gratefully acknowledged, together with the technical assistance of J. Allibon during the collection of the neutron data. Dr M. Lehmann advised and assisted at various stages of the data collection and initial data processing. Drs G. and E. Dodson gave much assistance in the application of the fast Fourier transform least-squares refinement technique (Agarwal, 1978), for which we are extremely grateful. We thank the SERC for the award of a studentship to HFJS and for funding to carry out the neutron data collection and initial data processing at the ILL. The use of computing facilities at Daresbury Laboratory (SERC), the SERC Interac-

tive Graphics Workstation at Birkbeck College, and the University of London Computer Centre are gratefully acknowledged.

The encouragement of Professor Dorothy Hodgkin was instrumental in the initiation and execution of this project.

References

- AGARWAL, R. C. (1978). *Acta Cryst.* **A34**, 791-809.
- BACON, G. E. (1980). Updated table of neutron scattering lengths in *Neutron Diffraction*. Oxford: Clarendon Press.
- COOKE, R. & KUNTZ, I. D. (1974). *Ann. Rev. Biophys. Bioeng.* **3**, 95-126.
- CRUICKSHANK, D. W. J. (1965). In *Computing Methods in Crystallography*, edited by J. S. ROLLETT. Oxford: Pergamon Press.
- EDSALL, J. T. & MCKENZIE, H. A. (1983). *Annu. Rev. Biophys. (Tokyo)*, **15**, 53-183.
- FINNEY, J. L., GOODFELLOW, J. M., HOWELL, P. L. & VOVELLE, F. (1986). *J. Biomol. Struct. Dyn.* **3**, 599-622.
- HAMILTON, W., ROLLETT, J. S. & SPARKS, R. A. (1964). *Acta Cryst.* **18**, 129-130.
- International Tables for X-ray Crystallography*. (1974). Vol. IV. Birmingham: Kynoch Press. (Present distributor D. Reidel, Dordrecht.)
- ISAACS, N. W. & AGARWAL, R. C. (1978). *Acta Cryst.* **A34**, 782-791.
- JOHNSON, C. K. (1976). *ORTEPII*. Report ORNL-5138. Oak Ridge National Laboratory, Tennessee.
- KUNTZ, I. D. & KAUFMANN, W. (1973). *Adv. Prot. Chem.* **28**, 239-345.
- LEHMANN, M. S. & LARSON, F. K. (1974). *Acta Cryst.* **A30**, 580-584.
- LEHMANN, M. S. & WILSON, S. (1982). *College V Data Reduction System: Treatment of Reflection Profiles*. ILL Tech. Rep. 74LIIT. Institut Laue-Langevin, Grenoble, France.
- LENHART, P. G. (1968). *Proc. R. Soc. London Ser. A*, **303**, 45-84.
- LENHART, P. G. & HODGKIN, D. C. (1961). *Nature (London)*, **192**, 937-938.
- NORTH, A. C. T., PHILLIPS, D. C. & MATHEWS, F. S. (1968). *Acta Cryst.* **A24**, 351-359.
- SAVAGE, H. F. J. (1986). *Biophys. J.* **50**, 947-965, 967-980.
- SIMMS, P. & TIMMINS, P. A. (1981). ILL Internal Tech. Rep. 82 SIO9T. Institut Laue-Langevin, Grenoble, France.
- TICKLE, I. J. (1975). *Acta Cryst.* **B31**, 329-331.