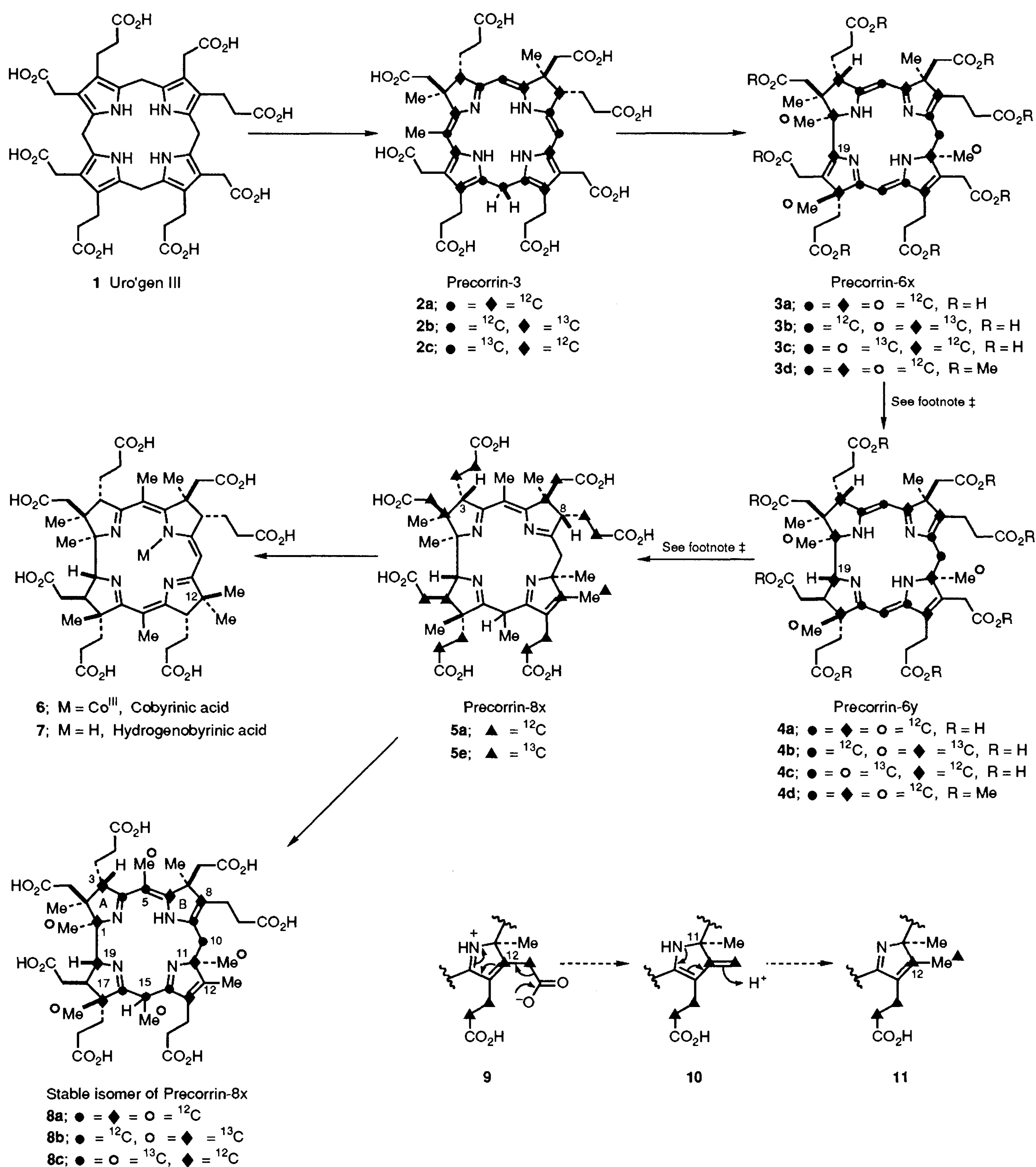
Biosynthesis of Vitamin B_{12}: Structural Studies on Precorrin-8x, an Octamethylated Intermediate and the Structure of its Stable Tautomer

Denis Thibaut,[a] Fumiyuki Kiuchi,[b] Laurent Debussche,[a] Francis Blanche,*[a] Masahito Kodera,[b] Finian J. Leeper[b] and Alan R. Battersby*[b]

[a] *Department Analyse, Centre de Recherche de Vitry-Alfortville, Rhône-Poulenc Rorer, 13 Quai Jules Guesde, BP14, F-94403 Vitry-sur-Seine Cedex, France*
[b] *University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, UK*

The structure of precorrin-8x, the octamethylated B_{12}-intermediate which follows precorrin-6y in *Pseudomonas denitrificans*, is largely determined by ^{13}C-labelling and NMR spectroscopy and the structure of its most stable tautomer is established.

The biosynthesis of vitamin B_{12} from uro'gen III **1** goes by way of cobyrinic acid **6** in *Propionibacterium shermanii* and probably *via* hydrogenobyrinic acid **7** in *Pseudomonas denitrificans*; the former is the Co^{III}-derivative of the latter. Formation of **6** and **7** from **1** involves the stepwise delivery of eight methyl groups from *S*-adenosylmethionine (SAM) onto carbon centres of **1**. The intermediates in this process are called precorrins and are distinguished by adding to the name a number corresponding to the number of *C*-methyl groups introduced from SAM.[1] Thus, the trimethylated interme-



diate[2–4] **2a** is precorrin-3. Recently, the isolation of a hexamethylated intermediate,[3] precorrin-6x† and proof[5,6] that its octamethyl ester has structure **3d** caused a major change in the direction of research on the biosynthesis of vitamin B_{12}. Then precorrin-6y,† which is the intermediate immediately following precorrin-6x, was isolated[7] and the structure of its ester **4d** was determined.[8] In addition, the reductase enzyme which saturates the C-18–C-19 double bond in **3a**‡ to give **4a**‡ has been found to be coded by the *cobK* gene in *P. denitrificans*[7] which has been overexpressed and the reductase isolated;[7] it is specifically dependent on NADPH (reduced nicotinamide adenine dinucleotide phosphate). Finally, the site to which the hydride equivalent is delivered from NADPH by the reductase has been shown[9] to be C-19 of **3a** and the complementary knowledge has been gained[10] that the reductase specifically transfers H_R from C-4 of NADPH.

† Occasionally there are sets of intermediates at the same level of methylation (*e.g.* precorrin-6). So the letters x, y etc. will eventually be replaced by A, B, C when the size of the set and the sequence within it are established.

‡ The double bonds of these octa-acids are shown in the same positions as proven for the corresponding esters but it cannot be excluded that the octa-acids are close tautomers of **3a** and **4a**.

Table 1 ^{13}C NMR data for stable isomer of precorrin-8x as labelled forms **8b** and **8c** dissolved in water pH *ca.* 2.7

8b				**8c**			
δ	Coupling, *J*/Hz	Assign-ment	Chemical shift of coupled protons δ_H (assignment)	δ	Coupling, *J*/Hz	Assign-ment	Chemical shift of coupled protons δ_H (assignment)
160.7	d, 5	C-6	3.11 (7-CH_2), 2.12 (5-Me), 1.38 (7-Me)	192.9	d, 53	C-4	*a*
133.8	s	C-13	4.15 (H-15), 3.70, 2.91 (13^1-CH_2), 2.15 (12-Me)	183.9	d, 43	C-14	1.65 (15-Me), 2.91, 3.70 (13^1-CH_2)
99.0	d, 5	C-8	3.11 (7-CH_2), 2.4–2.6 (8-CH_2CH_2), 1.38 (7-Me)	182.0	d, 45	C-16	1.22 (17-Me), 4.39 (H-19)
81.4	d, 39	C-19	1.57 (1-Me), 4.38 (H-19)	175.8	dd, 49, 9	C-9	2.47, tentatively 2.22 (8^1-CH_2)
79.0	t, 39	C-1	2.97 (2-CH_B), 1.57 (1-Me), 1.11 (2-Me)	119.7	br t, *ca.* 45	C-5	*a*
73.8	d, 32	C-11	2.15 (12-Me), 1.60 (11-Me), 3.61–3.65 (H-10)	37.4	dt, 31, 44	C-15	*a*
58.3	s	C-3	3.51 (2-CH_A), 1.11 (2-Me)	33.5	d, 49	C-10	1.60 (11-Me)
57.5	d, 35	C-17	2.73 (18-CH_2), 1.22 (17-Me)	26.4	s	11-Me	1.60 (11-Me)
26.4	d, 32	11-Me	1.60 (11-Me)	19.4	s	17-Me	See assignments for **8b**
19.4	d, 35	17-Me	1.22 (17-Me)	19.2	s	1-Me	
19.2	d, 39	1-Me	1.57 (1-Me)	18.9	d, 31	15-Me	
18.9	s	15-Me	4.15 (H-15), 1.65 (15-Me)	18.4	d, 41	5-Me	
18.4	s	5-Me	2.13 (5-Me)				

[a] Correlations for these centres are not available because H-3, H-10 and H-15 exchanged for ^{2}H in the D_2O solvent; correlations to C-5 and to 5-Me in the ^{1}H dimension were obscured by MeCN, the internal standard.

Major advances in the molecular biology of B_{12}-biosynthesis[11] have allowed isolation of the enzyme which converts precorrin-6y **4a** into the next biosynthetic intermediate, precorrin-8x. This enzyme, coded by the *cobL* gene, introduces the last two *C*-methyl groups and effects decarboxylation of the C-12 acetate residue. This communication outlines our studies on the structure of the enzymic product precorrin-8x.† Precorrin-8x heptamethyl ester, though briefly handleable for mass spectrometry,[12] underwent oxidative lactone formation during attempted purification. Accordingly, all the labelling and spectroscopic studies have been carried out on the hepta-acid with 0.3–0.5 mg being obtained in each labelling experiment.

It must surely be true that the *cobL* enzyme produces one substance, precorrin-8x, for this is the nature of enzymic processes. But during isolation, it changes into a mixture of at least five closely related forms which by HPLC[12] in an acidic solvent yielded five fractions A to E. Fractions A and B were homogeneous, the former being stable whereas the latter in water (pH 2.7 due to residual acid from HPLC) changed to A during 3–5 days at *ca.* 18 °C. Fractions C–E also changed to A under these conditions though more slowly. This phenomenon will be explained after discussing the studies which support structure **8a** for the stable isomer A of precorrin-8x.

The unlabelled form **8a** of this stable isomer A showed *m/z* 881.4133 by fast atom bombardment mass spectrometry corresponding to monoprotonated $C_{45}H_{60}N_4O_{14}$ which requires 881.4184. ^{13}C-Labelled precorrin-3 **2b** prepared biosynthetically[3] from 5-amino[4-^{13}C]laevulinic acid, [4-^{13}C]ALA, was converted enzymically[12] by way of the hexamethylated systems **3b** and **4b** into precorrin-8x using [*methyl*-^{13}C]SAM as the source of the last five *C*-methyl groups. Form A **8b** was isolated and studied using (*a*) 1D ^{13}C NMR, (*b*) 1D ^{1}H NMR, (*c*) 2D ^{1}H–^{13}C one-bond correlation, (*d*) 2D ^{1}H–^{13}C two-three bond correlation and (*e*) ^{1}H–^{1}H COSY. Table 1 collects the information from (*a*), (*c*) and (*d*) and the results from (*b*) and (*e*), though not detailed here, interlock with those in Table 1. Similarly, [5-^{13}C]ALA and [*methyl*-^{13}C]SAM were used for enzymic production[12] of precorrin-8x *via* **2c**, **3c** and **4c** and the stable form A **8c** was isolated for study by NMR using methods (*a*), (*c*) and (*d*) above; Table 1 collects the results.

The NMR assignments in Table 1 were made using the same analysis reported[5,8] for precorrin-6x **3** and precorrin-6y **4**; it will be given in our full paper. All the data are self-consistent and lead to structure **8a** for the stable isomer of precorrin-8x. This structure, apart from the stereochemistry at C-3 and C-15, also depends on the knowledge that precorrin-8x is enzymically converted into **7**, of firmly established structure, and is formed enzymically from **2a** *via* **3a** and **4a**.

Attention now focussed on which of the original fractions contained the true biosynthetic intermediate, precorrin-8x. Fraction A, **8a** was not converted enzymically into **7** whereas B gave **7** rapidly (*ca.* 100 nmol h^{-1} mg^{-1} of pure *cobH* enzyme[12]). Under the same conditions, fractions C to E yielded **7** much more slowly (C, 12; D, 8; E, 18% of the rate for B). Fraction B is thus precorrin-8x and it was found to be stable at room temperature in 0.1 mol dm^{-3} aqueous NaOH longer than was needed for NMR spectroscopy; fractions C–E reverted to B under these conditions (yield $\geqslant$ 50%).

It was essential to know whether the carbon attached to C-12 of precorrin-8x is a methylene or a methyl group; see part structures **9** → **10** → **11** for the probable mechanism of the decarboxylation process. Accordingly, [2,3-$^{13}C_2$]ALA was converted into precorrin-3 and precorrin-8x was enzymically derived[12] from it as before. Fraction B was studied under the foregoing alkaline conditions by NMR using 1D ^{13}C and ^{1}H spectra and 2D ^{1}H–^{13}C two-three bond correlation spectroscopy. The resultant data in Table 2 show that precorrin-8x carries a methyl group at C-12, confirmed by off-resonance spectrum, and C-12 itself is an sp^2 centre as in structure **5e**. The eight ^{13}C-labels in the propionate side-chains yielded useful corroborating evidence by the ^{1}H–^{13}C correlation (Table 2). The ^{1}H NMR spectrum showed two 1 H signals not split by directly attached ^{13}C and of chemical shift appropriate for H-3 and H-8 (δ 3.11 and 2.80). That C-8 is sp^3 hybridised in precorrin-8x is supported by a ^{13}C-spectrum of the sample from [4-^{13}C]ALA and [*methyl*-^{13}C]SAM taken *quickly* in the original mildly acidic solvent before appreciable change from form B to the stable A form **8b** had occurred. This spectrum showed almost no signal around δ 99 where C-8 of **8b** occurs and this signal only grew as the change to **8b** took place.

The ^{1}H NMR spectrum also showed two methyl groups on double bonds, one unlabelled (δ 1.85, 5-Me) and one enriched in ^{13}C (δ 1.84, 12-Me). The other six methyl groups were all on saturated carbons ($\delta < 1.5$) and one of them (δ 1.42) appeared as a doublet, shown to be coupled to a quartet at δ 3.71. This is in agreement with C-15 carrying a hydrogen as in the stable tautomer **8b**. Also evident in the ^{1}H NMR spectrum was a pair of coupled doublets at δ 2.87 and 3.15 (*J* 19 Hz), the latter showing a strong correlation with C-12; thus these signals must be due to a CH_2 at position 10.

Table 2 ^{13}C NMR data for precorrin-8x labelled form **5e** dissolved in 0.1 mol dm^{-3} aqueous NaOH

δ	Coupling, *J*/Hz	Assign-ment	Chemical shift of coupled protons δ_H (assignment)
166.7	d, 45	C-12	3.15 (H-10), 1.84 (12-Me), 1.11 (11-Me)
51.0	d, 34	C-2	3.11 (H-3), 1.28 (1-Me), 0.90 (2-Me)
47.2	d, 33	7-CH_2	2.42, 1.95 (7-CH_2), 1.22 (7-Me)
45.0	d, 33	C-7	1.22 (7-Me)
42.8	d, 34	2-CH_2	0.90 (2-Me)
42.5	d, 37	18-CH_2	4.33 (H-19), 3.63 (18-CH_2)
41.1	d, 37	C-18	
38.5	d, 33	17^2-CH_2	
37.5	d, 33	13^2-CH_2	
37.2	d, 34	8^2-CH_2	
37.1	d, 35	3^2-CH_2	3.11 (H-3)
33.6	d, 33	17^1-CH_2	
24.8	d, 34	13^1-CH_2	
24.1	d, 33	8^1-CH_2	
22.6	d, 35	3^1-CH_2	3.11 (H-3)
11.2	d, 45	12-Me	1.84 (12-Me)

The foregoing evidence is best accommodated by structure **5a** for the biosynthetic intermediate, precorrin-8x, a structure which can be expected to change to **8a** so gaining stabilisation[13] from the extended amidine system of rings A and B. There are other tautomers possible for the two separated chromophores in addition to **5a** and **8a** and epimerisation could occur at C-3, C-8 and C-15. These changes can explain the formation noted above of several structural forms all eventually dropping into the thermodynamic well represented by structure **8a**. All the main features of structure **5a** are secure but the final details, *e.g.* some points of stereochemistry, will require a further major effort in the future. It should be noted that precorrin-8x has all the *C*-methyl groups of **6, 7** and vitamin B_{12} attached to the macrocycle and only rearrangement of the 11-Me to position 12 is necessary to yield hydrogenobyrinic acid **7**, a rearrangement catalysed by the enzyme encoded by *cobH*.[12]

Grateful acknowledgement is made to J. Lunel, P. E. Bost and J.-C. Brunie for their interest and help, also to the SERC and to Rhône-Poulenc Rorer for financial support.

Received, 16th April 1992; Com. 2/01986H

References

1 H. C. Uzar, A. R. Battersby, T. A. Carpenter and F. J. Leeper, *J. Chem. Soc., Perkin Trans. 1*, 1987, 1689.
2 For reviews of earlier literature, see F. J. Leeper, *Nat. Prod. Rep.*, 1989, **6**, 171 and references cited therein.
3 D. Thibaut, L. Debussche and F. Blanche, *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 8795.
4 M. J. Warren, C. A. Roessner, S.-I. Ozaki, N. J. Stolowich, P. J. Santander and A. I. Scott, *Biochemistry*, 1992, **31**, 603.
5 D. Thibaut, F. Blanche, L. Debusschc, F. J. Leeper and A. R. Battersby, *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 8800.
6 F. Blanche, M. Kodera, M. Couder, F. J. Leeper, D. Thibaut and A. R. Battersby, *J. Chem. Soc., Chem. Commun.*, 1992, 138.
7 F. Blanche, D. Thibaut, A. Famechon, L. Debussche, B. Cameron and J. Crouzet, *J. Bacteriol.*, 1992, **174**, 1036.
8 D. Thibaut, F. Kiuchi, L. Debussche, F. J. Leeper, F. Blanche and A. R. Battersby, *J. Chem. Soc., Chem. Commun.*, 1992, 139.
9 G. W. Weaver, F. J. Leeper, A. R. Battersby, F. Blanche, D. Thibaut and L. Debussche. *J. Chem. Soc., Chem. Commun.*, 1991, 976.
10 F. Kiuchi, D. Thibaut, L. Debussche, F. J. Leeper, F. Blanche and A. R. Battersby, *J. Chem. Soc., Chem. Commun.*, 1992, 306.
11 F. Blanche, A. Famechon, D. Thibaut, L. Debussche, B. Cameron and J. Crouzet, *J. Bacteriol.*, 1992, **174**, 1050.
12 D. Thibaut, M. Couder, A. Famechon, L. Debussche, B. Cameron, J. Crouzet and F. Blanche, *J. Bacteriol.*, 1992, **174**, 1043.
13 A. Eschenmoser, *Angew. Chem., Int. Ed. Engl.*, 1988, **27**, 5.