OBC
www.rsc.org/obc

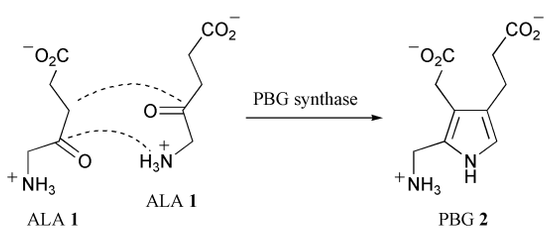
Stereochemistry and mechanism of the conversion of 5-aminolaevulinic acid into porphobilinogen catalysed by porphobilinogen synthase †

Catherine E. Goodwin and Finian J. Leeper *
University Chemical Laboratory, Lensfield Road, Cambridge, UK CB2 1EW.
E-mail: FJL1@cam.ac.uk

Received 5th March 2003, Accepted 26th March 2003
First published as an Advance Article on the web 2nd April 2003

(3*R*)- and (3*S*)-Deuteriated forms of 5-aminolaevulinic acid have been synthesised and the (3*R*)-form shows a significantly larger isotope effect when incubated with porphobilinogen synthase from bovine liver and from *Bacillus subtilis*; based on this and on available crystal structures, a modified mechanism for the enzymic reaction is proposed.

All the naturally occurring tetrapyrroles (haem, chlorophylls, vitamin B_{12}, *etc.*) are biosynthetically derived from the monopyrrole porphobilinogen (PBG, **2**).[1] PBG is made by the asymmetric condensation of two molecules of 5-aminolaevulinic acid (ALA, **1**) catalysed by PBG synthase (PBGS, also known as ALA dehydratase, EC 4.2.1.24).[2,3] The overall reaction catalysed by PBGS is shown in Scheme 1. It can be seen that one ALA molecule provides the acetate side-chain of PBG whereas the other ALA molecule provides the propionate side-chain. These ALA molecules must have separate binding sites in the active site of the enzyme, which are termed the A-site and the P-site respectively.

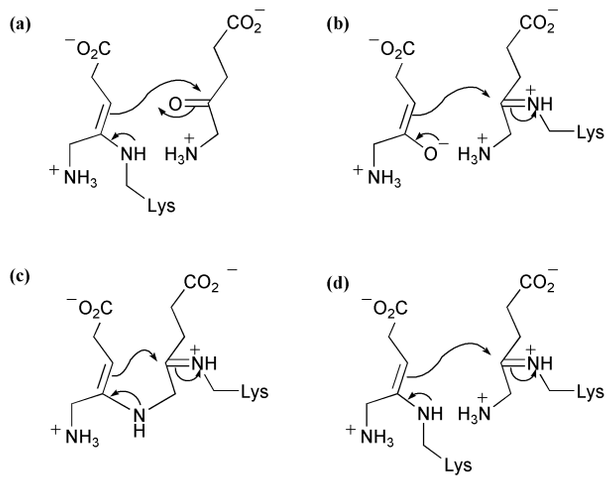


Scheme 1 Reaction catalysed by PBGS.

All PBGSs studied to date form a Schiff's base (*i.e.* an imine) between a conserved lysine residue and the keto-group of an ALA molecule. Treatment with $NaBH_4$ reduces the imine, which results in inactivation of the enzyme and allows identification of the lysine residue involved as Lys-246 (*E. coli* numbering). There has, however, been some uncertainty as to whether this Schiff's base is formed in the A-site (as in Scheme 2a[4]) or the P-site (as in Schemes 2b,[5] 2c[6] and 2d[2,7,8]).

Several crystal structures of PBGSs are now available, including ones from yeast,[9] *E. coli*,[10] and *Pseudomonas aeruginosa*.[11] These structures show that there are in fact two conserved lysine residues side-by-side in the active site. Very recently two crystal structures have been obtained in which *both* lysine residues have formed Schiff's bases.[8,12] These are shown in Fig. 1. As a result of these structures, it has been suggested[2,8] that the normal catalytic mechanism does in fact involve Schiff's bases in both the A-site and the P-site and the C–C bond forming step involves attack of an enamine in the A-site on an imine in the P-site (Scheme 2d).

† Electronic supplementary information (ESI) available: purification and assay protocols and plots of enzymic activity *vs.* substrate concentration. See http://www.rsc.org/suppdata/ob/b3/b302509h/



Scheme 2 Four possible mechanisms for the C–C bond forming step.

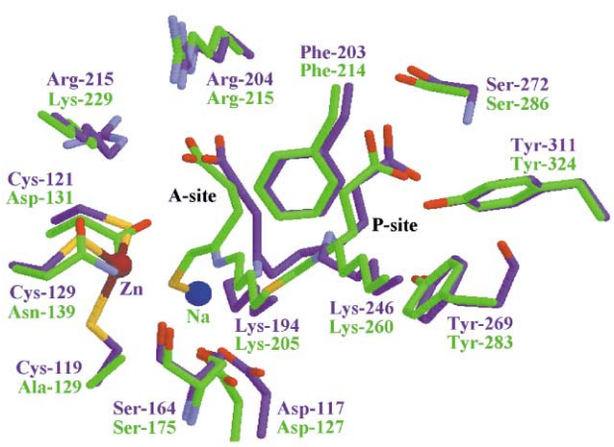


Fig. 1 An overlay of the three-dimensional active site structures of two different PBG synthases, both forming two Schiff's bases with substrate analogues. In purple is the Zn^{2+}-dependent *E. coli* PBGS complexed with 4,7-dioxosebacic acid (PBD code 1I8J)[12] and in green is the Mg^{2+}-dependent *Ps. aeruginosa* PBGS complexed with two molecules of 5-fluorolaevulinic acid and a Na^+ ion (PDB code 1GZG).[8] The two lysine residues which form the Schiff's bases are in front of the substrate analogue molecules and Phe-203/Phe-214 is behind. The close alignment of all the active site groups, apart from the differing metal ions and their ligands, suggests that the Zn^{2+}- and Mg^{2+}-dependent enzymes both use the same mechanism.

We have previously shown[13] that a substantial isotope effect is observed on the overall rate of reaction when [3,3-$^2H_2$]ALA is used, but no isotope effect is observed with [5,5-$^2H_2$]ALA. Furthermore it was deduced that the isotope effect with [3,3-$^2H_2$]ALA was largely due to the first deprotonation at C-3 rather than the second; this was because if the second deprotonation was rate-determining, one might expect the first

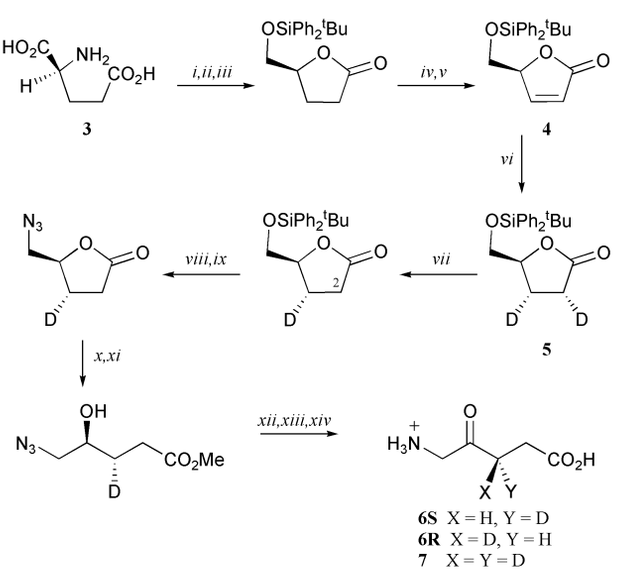
DOI: 10.1039/b302509h

Table 1 Kinetic parameters for the deuteriated and undeuteriated ALA molecules and the derived isotope effects

Substrate	V_{max}/units mg^{-1}	K_M/μM	$(V/K)/10^{-3}$	Isotope effect on V_{max}	Isotope effect on (V/K)
Bovine liver PBGS					
ALA **1**	0.683 ± 0.023	118 ± 15	5.77 ± 0.63	—	—
[3*S*-^{2}H]ALA **6S**	0.571 ± 0.014	110 ± 11	5.17 ± 0.42	1.20 ± 0.05	1.12 ± 0.15
[3*R*-^{2}H]ALA **6R**	0.328 ± 0.015	50 ± 13	6.63 ± 1.47	2.08 ± 0.12	0.87 ± 0.21
[3,3-$^2H_2$]ALA **7**	0.337 ± 0.009	72 ± 9	4.66 ± 0.51	2.03 ± 0.09	1.24 ± 0.19
B. subtilis PBGS					
ALA **1**	110.9 ± 1.8	273 ± 15	406 ± 17	—	—
[3*S*-^{2}H]ALA **6S**	56.7 ± 1.0	149 ± 10	381 ± 20	1.96 ± 0.05	1.07 ± 0.07
[3*R*-^{2}H]ALA **6R**	40.4 ± 0.4	90 ± 4	450 ± 17	2.75 ± 0.05	0.90 ± 0.05
[3,3-$^2H_2$]ALA **7**	30.5 ± 0.6	82 ± 7	371 ± 26	3.64 ± 0.09	1.09 ± 0.09

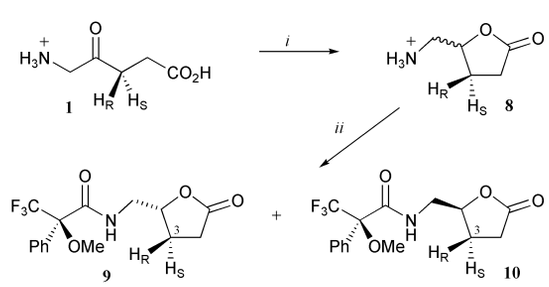
deprotonation to be reversible, resulting in release of ALA which had lost some of its deuterium, but in fact no such exchange could be detected. In this paper we report the synthesis of the two chirally monodeuteriated substrates, (3*R*)-[3-^{2}H]ALA **6R** and (3*S*)-[3-^{2}H]ALA **6S** as a method of probing the stereochemistry of the deprotonation steps.

The route used for the synthesis of **6S** is shown in Scheme 3. The synthesis of the unsaturated lactone **4** from L-glutamic acid **3** followed literature precedent.[14] The double bond of **4** was hydrogenated using deuterium gas over palladium on carbon. The ^{1}H NMR spectrum of the product **5** showed that the deuterium had come entirely from the opposite face to the bulky $CH_2OSi^tBuPh_2$ group. The deuterium at C-2 was largely removed by two rounds of deprotonation using LiHMDS followed by quenching the enolate with aq. NH_4Cl. ^{1}H NMR spectroscopy showed 8–14% deuterium remained at C-2 after this treatment but, as deuterium at C-2 is not expected to affect the rate of the PBGS-catalysed reaction, no further attempt to remove the last traces at this site was thought necessary. The remainder of the synthesis of **6S** is essentially as described in an earlier paper from this laboratory on the synthesis of 2-fluoroALA.[15] The enantiomeric **6R** was synthesised in the same way starting from D-glutamic acid.



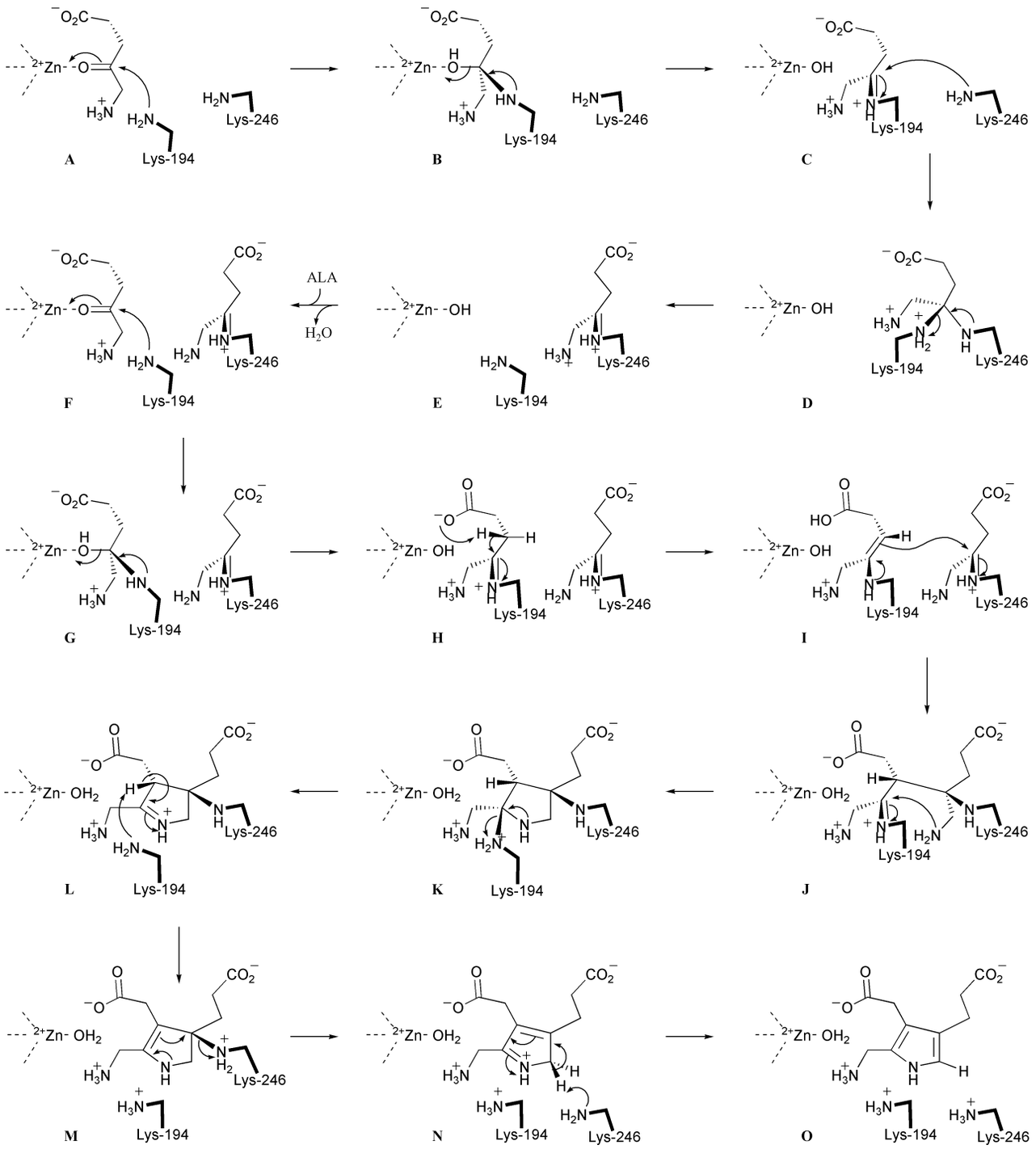
Scheme 3 Reagents: *i*, HNO_2; *ii*, $BH_3{\cdot}SMe_2$; *iii*, TBDPSCl, Et_3N, DMAP; *iv*, LiHMDS, PhSeBr; *v*, H_2O_2; *vi*, D_2, Pd/C; *vii*, LiHMDS then aq. NH_4Cl (twice); *viii*, TBAF; *ix*, Tf_2O then NaN_3; *x*, KOH; *xi*, Me_3SiCHN_2; *xii*, PCC; *xiii*, H_2, Pd/C; *xiv*, HCl.

The final step in the synthesis of **6S** and **6R** is an acid-catalysed hydrolysis of the methyl ester (the acid is important as ALA dimerises readily under alkaline conditions). However acidic conditions could cause enolisation of the ketone and thus loss of stereochemical purity. It was important, therefore, to develop a method to assay the stereochemical integrity of the deuterium labelling. This was achieved, as shown in Scheme 4, by initial reduction of ALA with $NaBH_4$ to give, after acidic work-up, the racemic aminomethyl-lactone hydrochloride **8**. This reduction avoids any further risk of enolisation of the ketone. The amine was then derivatised by acylation with (*R*)-Mosher's acid to give the diastereoisomers **9** and **10** in a 1 : 1 ratio. It was not necessary to separate these diastereoisomers as the protons at C-3 appeared at different chemical shifts in the 500 MHz ^{1}H NMR spectrum. Thus starting with undeuteriated ALA **1** four separate signals (all dddd, two of them overlapping) were seen in the δ 1.80 to δ 2.35 region for the protons attached to C-3 of **9** and **10**. Starting with **6S**, however, two of the four signals (δ 1.85 and δ 2.31) were present and starting with **6R** the other two signals (δ 1.94 and δ 2.29) were present. In each case a trace of the other isomer (*ca.* 5%) could just be detected. This demonstrates that the samples of **6S** and **6R** were very nearly stereochemically pure and no significant amount of racemisation had occurred that would interfere with the enzymic experiments.



Scheme 4 Reagents: *i*, $NaBH_4$ then HCl; *ii*, MTPA, EDCI, HOBt, Et_3N.

The samples of chirally deuteriated ALA **6S** and **6R** and also the 3,3-dideuteriated ALA **7**[13] were then tested as substrates for bovine liver PBGS (purchased from Sigma) and *Bacillus subtilis* PBGS (purified from an expressing strain of *E. coli*). A full range of substrate concentrations was used, so as to get accurate values for both V_{max} and K_M. The resulting curves are given in the supplementary material† and the derived kinetic parameters are given in Table 1. It can be seen from the table that the 3*R*-deuteriated ALA **6R** shows a significantly greater isotope effect on V_{max} than 3*S*-deuteriated ALA **6S**. This is true for PBGS from both sources but the difference is more pronounced for the bovine liver enzyme, although the magnitude of the isotope effects is smaller for this enzyme. For neither PBGS is there any significant isotope effect on V/K for any of the deuteriated ALA molecules.‡ This suggests that a step with a higher energy transition state occurs before the first deprotonation and so molecules are already committed to proceeding on to products by this stage.



Scheme 5 Proposed mechanism for the PBGS-catalysed reaction.

Because the greater isotope effect is observed with 3*R*-deuteriated ALA **6R**, we feel it is more likely that removal of the pro-*R* hydrogen atom attached to C-3 is the first deprotonation step in the enzymic mechanism. Armed with this information, one can look again at the crystal structures of this enzyme to try to determine what base is effecting the deprotonation and in which side of the active site.

To try to distinguish between the different mechanisms shown in Scheme 2, we have inspected the available crystal structures and have concluded that the mechanism of Scheme 2d is the most probable. This is because (a) there does not appear to be any base appropriately located to remove the 3*R* proton (nor the 3*S* proton) from an ALA molecule attached to Lys-246; (b) furthermore, the conformation of the ALA molecule bound to Lys-246 is inappropriate for the formation of a planar double bond between C-3 and C-4, with a dihedral angle (N=C4–C3–C2) of close to 90°; (c) if the oxygen atom of the enolate in Scheme 2b were coordinated to the Zn^{2+} ion then the required perpendicular attack from the double bond of the enolate on the iminium ion with Lys-246 would not be possible; (d) stereoelectronic reasons also eliminate the mechanism of Scheme 2c as it is a 5-*endo*-tet cyclisation which is disfavoured according to Baldwin's rules.[16] The mechanism in Scheme 2d, however, gives an enamine attached to Lys-194 which is very well oriented for attack on the iminium ion between ALA and Lys-246.

For these reasons we have concentrated on the double-Schiff's base mechanism first proposed by Neier.[7] We have based our detailed mechanism on the conformations adopted by the two 5-fluorolaevulinic acid molecules which form Schiff's bases with the two active-site lysine residues in the crystal structure of the Mg^{2+}-dependent PBGS from *Ps. aeruginosa*

(see Fig. 1, green).[8] However the close similarity between the active sites of the Zn^{2+}- and Mg^{2+}-dependent PBGSs (shown in Fig. 1) leaves little doubt that both follow the same mechanism. In the *Ps. aeruginosa* structure the 5-fluorolaevulinic acid molecule attached to Lys-205 (equivalent of Lys-194 in *E. coli*) has an ideal conformation for formation of the enamine with a dihedral angle for N=C4–C3–C2 of 173°. The nearest atom to the 3*R* proton is one of the oxygen atoms of the carboxylate group of the same ALA molecule, with an oxygen to carbon distance of 2.77 Å. Again the dihedral angles are close to perfect for deprotonation by this carboxylate group—the dihedral angle for C4–C3–C2–C1 is −88° (−90° would be ideal) and the dihedral angle for C3–C2–C1–O is −8° (0° would be ideal). Therefore in Scheme 5 we tentatively show this carboxylate group acting as the base for the first deprotonation. However it should be noted that intramolecular general base catalysis is not usually very efficient when it involves as small a ring as this.[17] Intramolecular deprotonation of C-3 of laevulinic acid derivatives by the carboxylate group has been observed but the reported effective molarity is only about 0.1 M.[18] Another possibility is that the deprotonation may be directly effected by a nearby enzymic group, *e.g.* a hydroxide ion ligated to the zinc ion in Zn^{2+}-dependent enzymes or the aspartate residue which replaces the zinc in Mg^{2+}-dependent enzymes.

The full mechanism that we propose for the PBGS-catalysed reaction is shown in Scheme 5. The zinc ion appears ideally placed to act as a Lewis acid catalyst in the formation of the Schiff's base between the keto group of an ALA molecule in the A-site and Lys-194 (see **A**). In contrast there does not appear to be any acidic group well placed to catalyse attack of Lys-246 on the keto group of an ALA molecule in the P-site. What we propose, therefore, is that the first ALA molecule initially forms a Schiff's base with Lys-194 in the A-site (**C**) but then imine exchange occurs and it is transferred to Lys-246 in the P-site (**E**). This would be consistent with earlier pulse-labelling results.[6,19] Following this the second ALA molecule binds in the A-site (**F**) and forms the Schiff's base with Lys-194 (**H**). Deprotonation of C-3 of the second ALA molecule, possibly by its own carboxylate group, then gives the enamine (**I**). Attack of the enamine on the iminium ion in the P-site then occurs to form the C–C bond (**J**). At this stage the second deprotonation at C-3 might seem a possibility but it is doubtful whether a planar C3–C4 double bond is attainable due to the position of the carboxylate binding sites and the attachment to the two lysine residues. What seems more likely, therefore, is that another imine exchange reaction occurs first, forming a five-membered Δ^1-pyrroline ring (**L**). Now the second deprotonation at C-3 seems much more feasible and furthermore Lys-194, which has just been released, would be well placed to be the base (molecular modelling suggests that the conformational change necessitated by the formation of the five-membered ring means that the proton is not well placed to be removed by the ALA carboxylate group, unlike in the first deprotonation step). The resulting enamine (**M**) could then expel the amino group of Lys-246 giving a pyrrolenine (**N**). The final step would be loss of a proton from C-5 to generate the pyrrole ring of PBG (**O**). It is known[20] that the proton at C-5 which is lost is the pro-*R* proton and the amino group of Lys-246 would be ideally placed to effect this deprotonation.

We feel that the mechanism in Scheme 5 is fully consistent with the stereoelectronic requirements of the various steps as well as with the stereochemical and other experimental data presented here and published previously.

Notes and references

‡ In our preliminary communication[13] we said that there was an isotope effect on *V*/*K* for the *B. subtilis* enzyme. However, the enzymic assays in that work did not extend to low enough substrate concentration to get an accurate measure of the K_M value. In this work the assays were performed with a much greater range of substrate concentrations and the values for K_M, and hence *V*/*K*, are therefore much more accurate.

1 A. R. Battersby, *Nat. Prod. Rep.*, 2000, **17**, 507.
2 P. M. Shoolingin-Jordan, P. Spencer, M. Sarwar, P. E. Erskine, K. M. Cheung, J. B. Cooper and E. B. Norton, *Biochem. Soc. Trans.*, 2002, **30**, 584.
3 E. K. Jaffe, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2000, **56**, 115; A. R. Battersby and F. J. Leeper, *Top. Curr. Chem.*, 1998, **195**, 143.
4 D. L. Nandi and D. Shemin, *J. Biol. Chem.*, 1968, **243**, 1236.
5 P. M. Jordan, in *Biosynthesis of Tetrapyrroles*, ed. P. M. Jordan, Elsevier, Amsterdam, 1991, p. 1.
6 P. M. Jordan and J. S. Seehra, *FEBS Lett.*, 1980, **114**, 283; P. M. Jordan and J. S. Seehra, *J. Chem. Soc., Chem. Commun.*, 1980, 240.
7 R. Neier, *Adv. Nitrogen Heterocycles*, 1996, **2**, 35.
8 F. Frère, W. D. Schubert, F. Stauffer, N. Frankenberg, R. Neier, D. Jahn and D. W. Heinz, *J. Mol. Biol.*, 2002, **320**, 237.
9 P. T. Erskine, E. M. H. Duke, I. J. Tickle, N. M. Senior, M. J. Warren and J. B. Cooper, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2000, **56**, 421; P. T. Erskine, L. Coates, R. Newbold, A. A. Brindley, F. Stauffer, S. P. Wood, M. J. Warren, J. B. Cooper, P. M. Shoolingin-Jordan and R. Neier, *FEBS Lett.*, 2001, **503**, 196; P. T. Erskine, R. Newbold, A. A. Brindley, S. P. Wood, P. M. Shoolingin-Jordan, M. J. Warren and J. B. Cooper, *J. Mol. Biol.*, 1999, **312**, 133; P. T. Erskine, N. Senior, S. Awan, R. Lambert, G. Lewis, L. J. Tickle, M. Sarwar, P. Spencer, P. Thomas, M. J. Warren, P. M. Shoolingin-Jordan, S. P. Wood and J. B. Cooper, *Nat. Struct. Biol.*, 1997, **4**, 1025; P. T. Erskine, R. Newbold, J. Roper, A. Coker, M. J. Warren, P. M. Shoolingin-Jordan, S. P. Wood and J. B. Cooper, *Protein Sci.*, 1999, **8**, 1250.
10 P. T. Erskine, E. Norton, J. B. Cooper, R. Lambert, A. Coker, G. Lewis, P. Spencer, M. Sarwar, S. P. Wood, M. J. Warren and P. M. Shoolingin-Jordan, *Biochemistry*, 1999, **38**, 4266.
11 N. Frankenberg, P. T. Erskine, J. B. Cooper, P. M. Shoolingin-Jordan, D. Jahn and D. W. Heinz, *J. Mol. Biol.*, 1999, **289**, 591.
12 E. K. Jaffe, J. Kervinen, J. Martins, F. Stauffer, R. Neier, A. Wlodawer and A. Zdanov, *J. Biol. Chem.*, 2002, **277**, 19792; J. Kervinen, E. K. Jaffe, F. Stauffer, R. Neier, A. Wlodawer and A. Zdanov, *Biochemistry*, 2001, **40**, 8227.
13 D. Appleton and F. J. Leeper, *Bioorg. Med. Chem. Lett.*, 1996, **6**, 1191.
14 S. Hanessian and P. J. Murray, *Tetrahedron*, 1987, **43**, 5055.
15 F. J. Leeper and M. Rock, *J. Fluorine Chem.*, 1991, **51**, 381.
16 J. E. Baldwin, *J. Chem. Soc., Chem. Commun.*, 1976, 734.
17 A. J. Kirby, *Adv. Phys. Org. Chem.*, 1980, **17**, 183.
18 R. P. Bell and A. D. Covington, *J. Chem. Soc., Perkin Trans. 2*, 1975, 1343.
19 P. M. Jordan and P. N. B. Gibbs, *Biochem. J.*, 1985, **227**, 1015.
20 M. M. Abboud and M. Akhtar, *J. Chem. Soc., Chem. Commun.*, 1976, 1007.