Published in FEBS Letters 503, issues 2-3, 196-200, 2001
which should be used for any reference to this work

The X-ray structure of yeast 5-aminolaevulinic acid dehydratase complexed with two diacid inhibitors

P.T. Erskine[a], L. Coates[a], R. Newbold[b], A.A. Brindley[b], F. Stauffer[c], S.P. Wood[a], M.J. Warren[b], J.B. Cooper[a,*], P.M. Shoolingin-Jordan[a], R. Neier[c]

[a] *Division of Biochemistry and Molecular Biology, School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton, SO16 7PX, UK*
[b] *School of Biological Sciences, Queen Mary, University of London, Mile End Road, London, E1 4NS, UK*
[c] *Institut de Chimie, Université de Neuchâtel, Av. Bellevaux 51, Case postale 2, CH-2007 Neuchatel 7, Switzerland*

Abstract The structures of 5-aminolaevulinic acid dehydratase complexed with two irreversible inhibitors (4-oxosebacic acid and 4,7-dioxosebacic acid) have been solved at high resolution. Both inhibitors bind by forming a Schiff base link with Lys 263 at the active site. Previous inhibitor binding studies have defined the interactions made by only one of the two substrate moieties (P-side substrate) which bind to the enzyme during catalysis. The structures reported here provide an improved definition of the interactions made by both of the substrate molecules (A- and P-side substrates). The most intriguing result is the novel finding that 4,7-dioxosebacic acid forms a second Schiff base with the enzyme involving Lys 210. It has been known for many years that P-side substrate forms a Schiff base (with Lys 263) but until now there has been no evidence that binding of A-side substrate involves formation of a Schiff base with the enzyme. A catalytic mechanism involving substrate linked to the enzyme through Schiff bases at both the A- and P-sites is proposed.

Key words: 5-Aminolevulinate dehydratase; 4-Oxosebacic acid; 4,7-Dioxosebacic acid; Mechanism

1. Introduction

The enzyme 5-aminolaevulinate dehydratase (ALAD, also known as porphobilinogen synthase, EC 4.2.1.24) performs its reaction at the confluence of the C-5 and Shemin pathways in tetrapyrrole biosynthesis [1–5]. The enzyme binds two molecules of 5-aminolaevulinic acid and condenses them to form the pyrrole porphobilinogen (Fig. 1). Four porphobilinogen molecules are then joined together and cyclised by subsequent enzymes in the pathway to form uroporphyrinogen III. The pathway diverges to produce an array of tetrapyrroles including cobalamins, chlorophyll and haem which are fundamental to many biological processes.

The ALAD enzymes share a high degree of sequence identity, contain about 350 amino acids per subunit and are usually octameric [1–5]. There are differences between the ALAD enzymes isolated from various species in terms of their metal requirements and susceptibility to oxidation. In humans, hereditary deficiencies in ALAD give rise to the rare disease Doss porphyria [6] and the inhibition of the enzyme by lead ions [7] is one of the main symptoms of acute lead poisoning [8]. 5-Aminolaevulinic acid structurally resembles the neurotransmitter GABA [9] and its buildup may be responsible for some of the neurological symptoms which accompany the inhibition of ALAD in Doss porphyria and lead poisoning [10]. Elevated levels of ALA are also found in the hereditary disease type I tyrosinaemia [11] which is thought to stem from the accumulation of succinyl acetone, a breakdown product of tyrosine and a potent inhibitor of ALAD.

The X-ray structures of ALAD from several sources have shown that the enzyme has a large homo-octameric structure with each subunit adopting the $(\beta/\alpha)_8$ or TIM barrel fold with a 39 residue N-terminal arm forming inter-subunit contacts [10,12–15]. Each subunit has an N-terminal tail or arm region which allows pairs of monomers to associate with their arms wrapped around each other to form compact dimers. Four dimers, which also interact principally via their arm regions, form the octamer and create a large solvent filled cavity in the centre. The active site of each subunit is located in a pronounced cavity formed by loops at the C-terminal ends of the β-strands forming the TIM barrel. All eight active sites are oriented towards the outer surface of the octamer and appear to be independent. At the base of each active site are two lysine residues (210 and 263), one of which (Lys 263 in yeast ALAD) is known to form a Schiff base link to the substrate. X-Ray structure analysis of inhibitor complexes revealed that a large flap covering the active site (residues 215–235) undergoes a substantial increase in order upon formation of the Schiff base with the enzyme. In all ALADs, except for those from plants and some bacteria, there is a well defined metal binding site at the catalytic centre formed by three cysteine residues (133, 135 and 143 in yeast ALAD) and a solvent molecule. Some ALADs also possess a regulatory metal binding site which is located between each TIM barrel domain and the N-terminal arm of a neighbouring subunit in the octamer [13].

In the reaction catalysed by ALAD, two identical molecules of 5-aminolaevulinic acid are bonded together to form an asymmetric pyrrole product porphobilinogen. The substrate molecules (along with each of their associated binding sites) are named according to which side of the product porphobilinogen they will go on to form, namely the acetate (A) or propionate (P) sides as shown in Fig. 1. The main features

*Corresponding author. Fax: (44)-2380-59 44 59.
E-mail address: j.b.cooper@soton.ac.uk (J.B. Cooper).

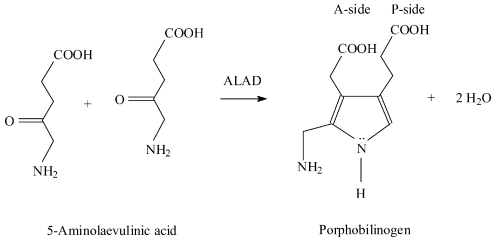

Fig. 1. The reaction catalysed by 5-aminolaevulinic acid dehydratase (ALAD). Two molecules of 5-aminolaevulinic acid are condensed to form the pyrrole porphobilinogen.

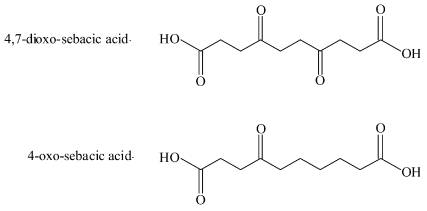

Fig. 2. The formulae of the two diacid ALAD inhibitors: 4-oxosebacic acid and 4,7-dioxosebacic acid.

of the received mechanism are Schiff base formation between one of the substrate molecules and an invariant lysine residue, the formation of an inter-substrate Schiff base and an aldol condensation. Single turnover experiments have shown that it is the substrate molecule which goes on to constitute the propionate side of the product which binds first and it has been established that this substrate forms a Schiff base with the enzyme [16–19]. Whether it is C–N or C–C bond formation between the two substrates which finally closes the pyrrole ring is still unclear. Whatever the mechanism, at some point during the reaction a long chain dicarboxylate is created as an intermediate. If C–C bond formation is first, the intermediate will take the form of a seven carbon chain diacid whilst if C–N bond formation occurs first, a 10 atom chain diacid will be formed [20–22]. As has been noted previously, dicarboxylic acids may be able to cross-link the carboxylic acid binding groups associated with the A and P substrate binding sites [23,24]. Because of this potential we have crystallised the ALAD from *Saccharomyces cerevisiae* in complex with the 10 carbon chain diacid irreversible inhibitors 4-oxosebacic acid and 4,7-dioxosebacic acid (formulae are shown in Fig. 2). These complexes have been subjected to X-ray structural analysis, the results of which are presented here.

This study shows that the diketo inhibitor binds by forming Schiff bases at the catalytic centre with both of the invariant lysine residues (263 and 210). Other studies have only provided evidence that Lys 263 forms a Schiff base with the P-site substrate [12–15]. The finding that the other invariant lysine at the active site is also capable of making such an interaction is in accord with a recent proposal for the catalytic mechanism which speculated that both the A- and P-side substrates bind by making Schiff bases with the enzyme [25]. Both inhibitors define interactions at the enzyme's A-site more clearly than previous studies on binding of inhibitors to yeast ALAD.

2. Materials and methods

2.1. Data collection and data processing

Co-crystals of yeast ALAD with the inhibitors were obtained using the hanging-drop method with the same conditions as those used for the native enzyme [26] except for the addition of the inhibitor to give a final concentration of 5 mM in each drop. Crystals of the complexes appeared within 3–4 weeks and were flash-cooled in liquid ethane using ~30% glycerol v/v as a cryoprotectant for cryostorage under liquid nitrogen.

Initial datasets were collected at the EMBL X11 beamline at DESY (Hamburg) to 2.1 Å resolution and these were used for the initial structure determination of the bound inhibitors. Datasets to slightly higher resolution (around 1.8 Å) were then collected at the ID14-2 and ID14-4 beamlines at ESRF (Grenoble) and these were used in the refinement. The data were processed using the program MOSFLM [27] and were merged with SCALA of the CCP4 suite [28]. The data collection and processing statistics are shown in Table 1. The complexes were isomorphous with the native enzyme crystals allowing the ligand molecules to be located by calculation of difference Fourier maps using the phases from the native yeast ALAD structure [10]. The structures were refined using SHELX-97-2 [29] initially by rigid body refinement to compensate for small changes in unit cell parameters. This was followed by stereochemically restrained least-squares refinement of atomic coordinates and isotropic temperature factors. The refinement was done by gradually increasing the resolution to achieve a greater radius of convergence. Graphical rebuilding was done using TURBO-FRODO (Bio-Graphics, Marseille) running on Silicon-Graphics computers. Various modifications to the restraint dictionaries were made to incorporate the ligand molecules and to ensure planarity of the Schiff bases formed with the enzyme. The coordinates of these complexes have been deposited with the Protein Data Bank (PDB) and the accession codes are shown in Table 1 (to be done on acceptance of the manuscript).

3. Results

Clear electron density for the bound ligands was found at the active site in the initial difference Fourier maps confirming that binding to the enzyme had occurred. The structures were refined to high resolution and the crystallographic *R*-factors and free *R*-factors are given in Table 1 along with other

Table 1
X-Ray data and refinement statistics for the yeast ALAD complexes

Ligand	4-Oxosebacic acid	4,7-Dioxosebacic acid
Unit cell		
a, *b* (Å)	103.5	104.3
c (Å)	168.2	167.2
Number of crystals	1	1
Statistics for entire dataset		
Resolution range (Å)	30.4–1.8	36.3–1.75
R-Merge (%)	5.9	11.4
Completeness (%)	94.6	95.5
Multiplicity	5.6	6.2
Mean $I/\sigma(I)$	4.9	4.6
Statistics for outer shell		
Resolution range (Å)	1.9–1.8	1.84–1.75
R-Merge (%)	42.5	36.4
Completeness (%)	87.7	98.5
Multiplicity	3.7	3.6
Mean $I/\sigma(I)$	1.5	1.6
Refinement		
R-Factor (%)	26.5	23.9
R-Free (%) (5% test set)	32.9	30.6
Totoal number of reflections	39 906	44 374
RMSD bond lengths (Å)	0.006	0.008
RMSD 1–3 distances (Å)	0.020	0.023
RMSD bumps (Å)	0.006	0.011
RMSD chiral volumes (Å^3)	0.035	0.044
RMSD planes (Å^3)	0.029	0.037

In both cases the space group was *I422*.

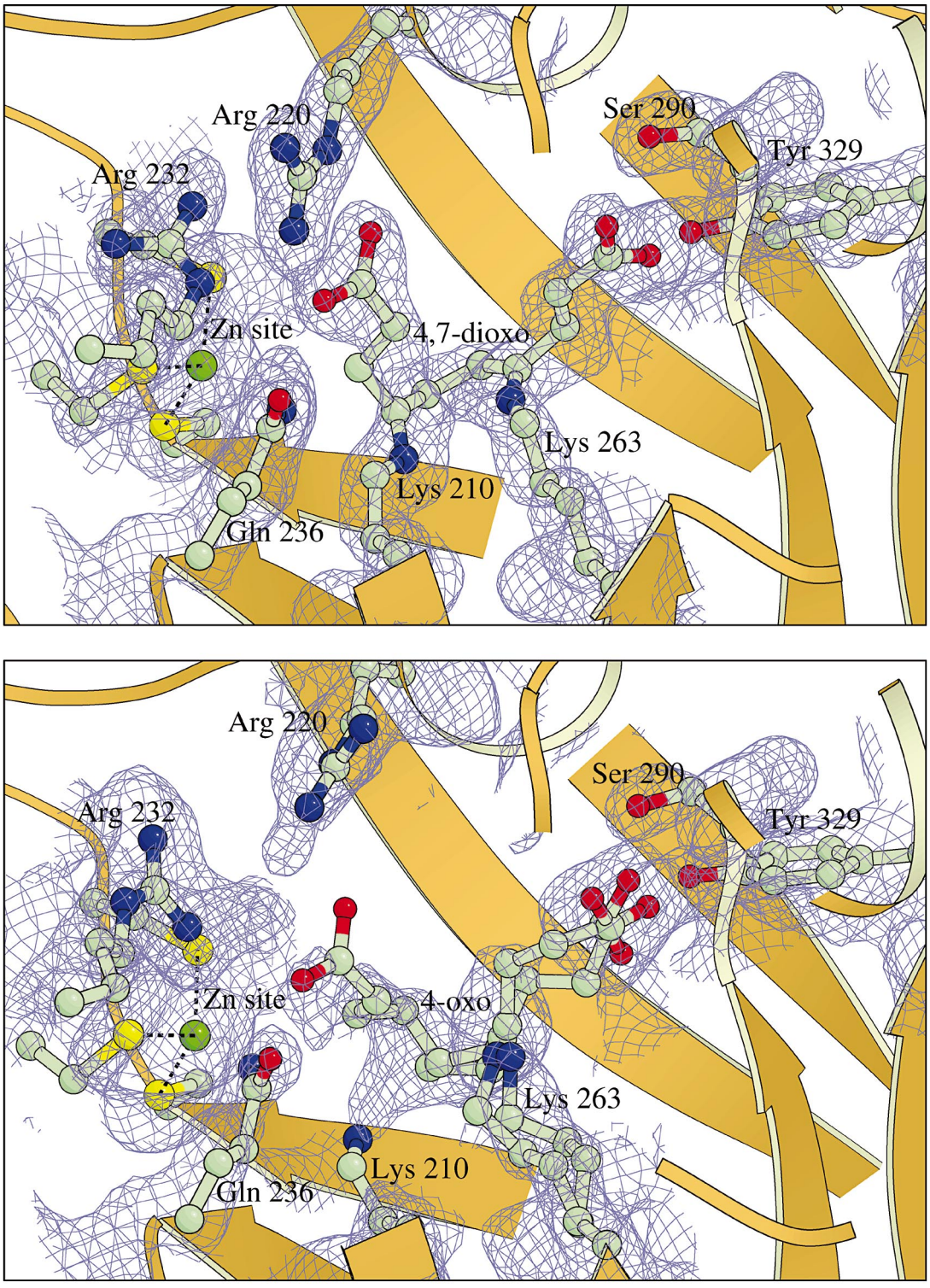

Fig. 3. The refined electron density maps (shown in purple) for the inhibitors 4,7-dioxosebacic acid at 1.75 Å resolution (top) and 4-oxosebacic acid at 1.8 Å resolution (bottom). The maps are contoured at 1.5 rms (top) and 1.2 rms (bottom).

refinement statistics. The refined structure of the 4,7-dioxosebacic acid complex has 91% of residues within the 'most favoured' regions of the Ramachandran plot by the criteria used in PROCHECK [30]. The 4-oxosebacic acid structure shows a somewhat higher degree of disorder especially in loops close to the active site. Accordingly, this structure has slightly poorer stereochemistry with 86% of residues within the 'most favoured' regions. For the better defined 4,7-dioxosebacic acid complex, 100% of the residues are within the so-called 'additional allowed' regions which compares with a value of 95% for the other complex.

The prefined electron density maps for each ligand are shown in Fig. 3. It appears that the 4-oxosebacic acid molecule is somewhat disordered and the part which binds in the

P-site has been modelled in two conformations. The same effect has been found with several other inhibitors which occupy the P-site [12,31]. In contrast, all of the 4,7-dioxosebacic acid molecule is defined well by the electron density map. The absence of disorder in this inhibitor complex may correlate with the ligand's more tightly bound state held by two Schiff bases with the enzyme.

Studies with substrate 5-aminolaevulinic acid, and the inhibitors laevulinic acid, succinyl acetone and 4-keto-5-aminohexanoic acid have helped to define the residues forming the P-site [12–15, 31]. Both ligands in the present study form a Schiff base with Lys 263 and bind with their P-site carboxyl groups making hydrogen bonds with Ser 290 and Tyr 329. These interactions are characteristic of all other P-site ligands which have been studied previously. These molecules fit in a hydrophobic pocket beneath the active site flap (215–235) which becomes substantially more ordered when the inhibitors bind. This ordering appears to be due to tight hydrophobic interactions between the P-site ligands and residue Phe 219 in the flap. Kinetic studies have suggested that the K_m of the P-site is lower than that of the A-site (4.6 μM as against 66 μM for the *Escherichia coli* enzyme) [21,22]. This correlates with the finding that all substrate analogues studied up till now bind predominantly in the P-site and the adjacent A-site appears to be occupied only by water molecules, and a number of the side chains forming this site are disordered. In consequence the residues which bind the A-side substrate have been more elusive to define.

The two inhibitors used in the present study provide good definition of the A-site residues. The most striking finding is that the 4,7-dioxosebacic acid makes a second Schiff base through its C-7 keto function with Lys 210. This was clearly shown by the initial difference electron density and has been confirmed by subsequent refinement at 1.8 Å resolution. The A-site carboxyl group occupies a similar position in both ligands, hydrogen bonding with the amino group of Gln 236. However, in the more ordered 4,7-dioxosebacic acid complex it is notable that hydrogen bonds are formed between the A-site carboxyl group and the guanidinium groups of Arg 220 and Arg 232. Hence these residues are very likely to form a salt bridge with the carboxylate of A-side substrate. These arginines (along with Gln 236) are strongly conserved residues. In the less well defined 4-oxosebacic acid complex, the interactions with these arginine residues are more distant. Lys 210 is notably less well ordered in the 4-oxosebacic acid complex than it appears to be in the Schiff base link of the 4,7-dioxosebacic acid complex.

4. Discussion

Most mechanistic proposals for ALAD involve the formation of a Schiff base between the P-side substrate moiety and an active site lysine (Lys 263 in yeast ALAD). This stems from labelling studies which indicated that only the P-side substrate moiety becomes covalently trapped upon reduction with $NaBH_4$ [19]. It is likely that upon binding of the A-side substrate the ensuing reaction with P-side substrate is sufficiently fast to preclude trapping of a Schiff base at the A-site under the conditions used in these earlier studies. The present finding that the dioxo inhibitor forms Schiff bases with both of the invariant active site lysine residues (263 and 210 in yeast ALAD) suggests that both A- and P-side substrate

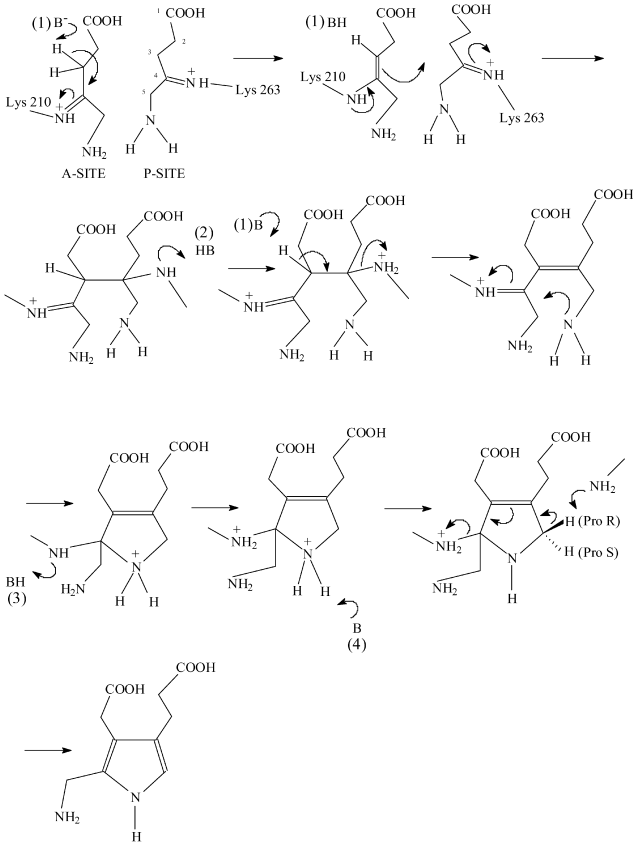

Fig. 4. A possible mechanism for ALAD in which both A- and P-side substrates form Schiff bases with the enzyme.

may form Schiff bases upon productive binding to the enzyme.

Fig. 4 shows a proposal for the catalytic mechanism involving both ALA moieties held by Schiff bases with active site lysine residues. In this mechanism a basic group (shown as (1) in Fig. 4) is required to abstract a proton from C-3 of A-side ALA leading to an enamine intermediate. This nucleophilically attacks P-site ALA yielding the C–C bond linking the two substrates in the final pyrrole product. We have speculated that base (1) is a zinc-bound hydroxide [13]. This base could also catalyse a second deprotonation of the A-side C-3, leading to the formation of an inter-substrate C–C double bond. This could be followed by nucleophilic attack of the P-side N-5 on the A-side C-4, leading ultimately to the inter-substrate C–N bond as well as the cleavage of the Schiff base link between the A-side C-4 and Lys 210.

The proximity of the amino groups of the two active site lysine residues suggests that each could fulfil a role as a base to deprotonate the other at different stages in the reaction. Thus we speculate that base (2) which protonates the Schiff base nitrogen of Lys 263 may actually be the spatially adjacent Schiff base involving Lys 210. The latter Schiff base could easily be re-protonated since it is close to a number of polar side chains in the vicinity of the active site zinc ion. A basic group (3) is also required for deprotonation of the second active site lysine (Lys 210) and we speculate that this base may be Lys 263. In the reaction a base (4) assists the deprotonation of the N-5 of P-side ALA. We have speculated that two proximal active site residues, namely Asp 131 and Ser

179, may be involved in this process. The final deprotonation shown in Fig. 4 involves the C-5 of P-side ALA. This step, which has been shown to be stereospecific for the *pro-R* hydrogen [32], could also be catalysed by Lys 263 in view of its proximity. In the X-ray structure the *pro-R* hydrogen on C-5 is pointing in the direction of a polar pocket formed by the side chains of Asp 131, Ser 179, Tyr 207, Tyr 287 and Lys 263. Of these residues the lysine is well placed to catalyse a deprotonation. In contrast, the *pro-S* hydrogen points towards an aromatic pocket formed by Phe 46, Phe 89 and Phe 219. Thus the high resolution X-ray structure of the 4,7-dioxosebacic acid complex is also consistent with the experimentally observed stereochemistry of the enzyme catalysed reaction.

Early experiments on ALAD suggested a mechanism in which the substrate at the A-site was initially linked to the enzyme through a Schiff base to facilitate deprotonation at the C-3 position [33]. Subsequently, single turnover experiments demonstrated that it is not the A-site substrate but the P-side substrate that binds initially to the enzyme through a Schiff base [16–19]. A major step in the understanding of the mechanism stemmed from the X-ray structure analysis of yeast ALAD [10] which revealed two adjacent lysines in the active site (Lys 263 and Lys 210). This re-opened the possibility that two Schiff bases may be involved in the mechanism, one binding each substrate. This possibility has been further substantiated by the finding that the inhibitor 4,7-dioxosebacic acid binds by forming Schiff bases with both of the active site lysines.

Acknowledgements: We gratefully acknowledge the financial support of the BBSRC (UK). We also thank the staff at EMBL (Hamburg) and ESRF (Grenoble) for access to synchrotron beam time and associated travel funds. We thank Prof. M. Akhtar FRS and M. Montgomery (Southampton) for comments on the mechanism.

References

[1] Jordan, P.M. (1991) New Compr. Biochem. 19, 1–65.
[2] Jordan, P.M. (1994) Curr. Opin. Struct. Mol. Biol. 4, 902–911.
[3] Warren, M.J. and Scott, A.I. (1990) Trends Biochem. Sci. 15, 486–491.
[4] Jaffe, E.K. (1995) J. Bioenerg. Biomembr. 27, 169–179.
[5] Jaffe, E.K. (2000) Acta Crystallogr. Sect. D Biol. Crystallogr. 56, 115–128.
[6] Doss, M., Von-Tieperman, R., Schneider, J. and Schmid, H. (1979) Klin. Wochenschr. 57, 1123–1127.
[7] Simons, T.J.B. (1995) Eur. J. Biochem. 234, 178–183.
[8] Warren, M.J., Cooper, J.B., Wood, S.P. and Shoolingin-Jordan, P.M. (1998) Trends Biochem. Sci. 23, 217–221.
[9] Brennan, M.J.W. and Cantrill, R.C. (1979) Nature 280, 514–515.
[10] Erskine, P.T., Senior, N., Awan, S., Lambert, R., Lewis, G., Tickle, I.J., Sarwar, M., Spencer, P., Thomas, P., Warren, M.J., Shoolingin-Jordan, P.M., Wood, S.P. and Cooper, J.B. (1997) Nat. Struct. Biol. 4, 1025–1031.
[11] Mitchell, G., Larochelle, J., Lambert, M., Michaud, J., Grenier, A., Ogier, H., Gauthier, M., Lacroix, J., Vanasse, M., Larbrisseau, A., Paradis, K., Weber, A., Lefevre, Y., Melancon, S. and Dallaire, L. (1990) New Engl. J. Med. 322, 432–437.
[12] Erskine, P.T., Newbold, R., Roper, J., Coker, A., Warren, M.J., Shoolingin-Jordan, P.M., Wood, S.P. and Cooper, J.B. (1999) Protein Sci. 8, 1250–1256.
[13] Erskine, P.T., Norton, E., Cooper, J.B., Lambert, R., Coker, A., Lewis, G., Spencer, P., Sarwar, M., Wood, S.P., Warren, M.J. and Shoolingin-Jordan, M.J. (1999) Biochemistry 38, 4266–4276.
[14] Frankenberg, N., Erskine, P.T., Cooper, J.B., Shoolingin-Jordan, P.M., Jahn, D. and Heinz, D.W. (1999) J. Mol. Biol. 289, 591–602.
[15] Mills-Davis, N.L. (2000) PhD thesis, University of Southampton.
[16] Jordan, P.M. and Seehra, J.S. (1980) J. C. S. Chem. Comm. 1243, 240–242.
[17] Jordan, P.M. and Seehra, J.S. (1980) FEBS Lett. 114, 283–286.
[18] Jordan, P.M. and Gibbs, P.N.B. (1985) Biochem. J. 227, 1015–1020.
[19] Gibbs, P.N.B. and Jordan, P.M. (1986) Biochem. J. 236, 447–451.
[20] Jarret, C., Babalova, J., Stauffer, F., Henz, M.E. and Neier, R. (1998) in: Electronic Conference on Heterocyclic Chemistry 98 (ECHET98) (Rzepa, H.J. and Kappe, O., Eds.), Imperial College Press, Article 103.
[21] Jarret, C., Stauffer, F., Henz, M.E., Marty, M., Luond, R.M., Bobalova, J., Schurmann, P. and Neier, R. (2000) Chem. Biol. 7, 185–196.
[22] Neier, R. (2000) J. Heterocycl. Chem. 37, 487–508.
[23] Senior, N., Thomas, P.G., Cooper, J.B., Wood, S.P., Erskine, P.T., Shoolingin-Jordan, P.M. and Warren, M.J. (1996) Biochem. J. 320, 401–412.
[24] Stauffer, F., Zizzari, E., Engeloch-Jarret, C., Faurite, J.-P., Babalova, J. and Neier, R. (2001) ChemBioChem 2, 343–354.
[25] Neier, R. (1996) Adv. Nitrogen Heterocycles 2, 35–146.
[26] Erskine, P.T., Senior, N., Maignan, S., Cooper, J., Lambert, R., Lewis, G., Spencer, P., Awan, S., Warren, M., Tickle, I.J., Thomas, P., Wood, S.P. and Shoolingin-Jordan, P.M. (1997) Protein Sci. 6, 1774–1776.
[27] Leslie, A.G.W. (1992) Recent changes to the MOSFLM package for processing film and image plate data, Joint CCP4+ESF-EAMCB Newsl. Protein Crystallogr. No. 26.
[28] CCP4 (1994) Acta Crystallogr. Sect. D Biol. Crystallogr. 50, 760–763.
[29] Sheldrick, G.M. (1998) in: Crystallographic Computing, Vol. 7 (Watenpaugh, K. and Bourne, P.E., Eds.), Oxford University Press, Oxford.
[30] Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) J. Appl. Crystallogr. 26, 283–291.
[31] Erskine, P.T., Newbold, R., Brindley, A.A., Wood, S.P., Shoolingin-Jordan, P.M., Warren, M.J. and Cooper, J.B. (2001) J. Mol. Biol., in press.
[32] Chaudhry, A.G. and Jordan, P.M. (1976) Biochem. Soc. Trans. 4, 760–761.
[33] Nandi, D.L. and Shemin, D. (1968) J. Biol. Chem. 243, 1236–1242.