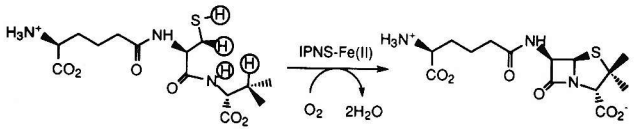
Crystal structure of isopenicillin *N* synthase is the first from a new structural family of enzymes

Peter L. Roach*, Ian J. Clifton†, Vilmos Fülöp†, Karl Harlos†, Geoffrey J. Barton†, Janos Hajdu†, Inger Andersson†‡, Christopher J. Schofield* & Jack E. Baldwin*§

* Dyson Perrins Laboratory and the Oxford Centre for Molecular Sciences, University of Oxford, South Parks Road, Oxford OX1 3QY, UK
† Laboratory of Molecular Biophysics and the Oxford Centre for Molecular Sciences, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

PENICILLIN antibiotics are all produced from fermentation-derived penicillins because their chemical synthesis is not commercially viable. The key step in penicillin biosynthesis, in which both the β-lactam and thiazolidine rings of the nucleus are created, is mediated by isopenicillin *N* synthase (IPNS), which binds ferrous iron and uses dioxygen as a cosubstrate. In a unique enzymatic step, with no chemical precedent, IPNS catalyses the transfer of four hydrogen atoms from its tripeptide substrate to dioxygen forming, in a single reaction, the complete bicyclic nucleus of the penicillins[1]. We now report the structure of IPNS complexed with manganese, which reveals the active site is unusually buried within a 'jelly-roll' motif and lined by hydrophobic residues, and suggest how this structure permits the process of penicillin formation. Sequence analyses indicate IPNS, 1-aminocyclopropane-1-carboxylic acid oxidase and many of the 2-oxo-acid-dependent oxygenases contain a conserved jelly-roll motif, forming a new structural family of enzymes.

The story of penicillin is one replete with surprises, from its serendipitous discovery[2], its unusual chemical structure and efficacy as an antibiotic[3,4], and more recently that its biosynthesis is dependent upon iron and dioxygen[5]. The key step in penicillin biosynthesis is the transformation of the linear tripeptide *L*-δ-(α-aminoadipoyl)-*L*-cysteinyl-D-valine (ACV) into isopenicillin *N* (IPN) by the loss of four hydrogen atoms in a desaturative ring closure with concomitant reduction of dioxygen to water.



L-δ-(α-Aminoadipoyl)-*L*-cysteinyl-*D*-valine (ACV) Isopenicillin N (IPN)

Thus, the full four-electron oxidizing power of dioxygen is used in this IPNS-catalysed desaturative step, unlike other non-haem ferrous iron-dependent oxygenases and oxidases which require electron donors or oxidizable cosubstrates[6].

§ To whom correspondence should be addressed.
‡ Permanent address: Swedish University of Agricultural Sciences, Uppsala Biomedical Centre, Department of Molecular Biology, PO Box 590, S-75124, Uppsala, Sweden.

TABLE 1 Data collection and phasing statistics

Compound	Resolution (Å)	Number of observations	Unique reflections	Completeness (%)	R_{merge}* (%)	Isomorphous difference† (%)	Number of sites	Phasing power‡ (acentric/centric)
Native	2.5	136,252	36,720	98.1	10.9			
EMP§	3.0	79,890	21,416	97.9	14.2	20.6	4	1.43/1.13
Baker's di-mercurial	3.5	44,943	13,961	94.4	12.5	30.2	4	1.56/1.22

Recombinant IPNS from *Aspergillus nidulans*[8] was crystallized in the presence of 2.5 mM $MnCl_2$ (ref. 9). The crystals (0.5–1.0 mm overall dimensions) belong to space group $P2_12_12_1$ with unit cell dimensions of $a = 59.2$ Å, $b = 127.0$ Å, $c = 139.6$ Å. The asymmetric unit contains a dimer, and the solvent content of the crystals is 60%. The structure of the enzyme was solved by the method of multiple isomorphous replacement (MIR) from two heavy-atom derivatives. The derivatives were prepared by soaking native crystals at 20–21 °C in 0.5 mM solutions of the respective heavy-atom reagents for 16–20 h. High-resolution native data were collected using 0.893 Å radiation at beamline 4 of the ESRF (30 cm MAR research image plate detector). All other data were collected in-house on an 18 cm MAR Research image plate detector using CuK_α radiation from a Rigaku rotating anode generator operating at 60 kV, 70 mA. The data were processed with the DENZO program[19] and the MIR phases were improved by non-crystallographic symmetry averaging, solvent flattening and histogram matching using the CCP4 Suite of Programs (Program DM)[20]. The two derivatives had two sites in common. Electron density maps were interpreted with the program O[21]. Skeletonized maps[22] were calculated for chain tracing. Intermediate models of the enzyme were refined by simulated annealing using the program X-PLOR[23] with non-crystallographic symmetry restraints imposed. During 7 cycles of refinement and rebuilding, 328 residues (Gly 2 and Ser 3 were not visible in each subunit) could be fitted to the density in both molecules of the asymmetric unit. In the final cycle, the positions of all the atoms in the asymmetric unit including 194 water molecules, two active-site-bound manganese ions, and two additional managese ions (bound to Glu 81 and His 82 in each subunit) were refined without non-crystallographic symmetry restraints. At present, the crystallographic R factor (defined as $R_{cryst} = \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}| \times 100$) in the resolution range of 8–2.5 Å is 22.0% for all observations and 21.5% using 33,229 reflections with $F_{obs} > 2\sigma(F_{obs})$. The free R value[24] is 26.5% based on 1,449 randomly selected reflections (4% of the total) and 26.4% for 1,425 reflections with $F_{obs} > 2\sigma(F_{obs})$. The r.m.s. deviation of the bond lengths, bond angles and torsion angles from standard values is 0.01 Å, 2° and 24°, respectively. The mean coordinate error is 0.45 Å based on the SIGMAA method[25]. The r.m.s. difference between subunits is 0.32 Å for C_α atoms and 0.72 Å for all atoms. Gln 330, which is coordinated to the active-site metal, is the only Ramachandran outlier in each chain (chain A, $\phi = 120°$, $\psi = 100°$; chain B, $\phi = 112°$, $\psi = 90°$).

* $R_{merge} = \sum_j \sum_h |I_{h,j} - \langle I_h \rangle| / \sum_j \sum_h \langle I_h \rangle \times 100$, where $I_{h,j}$ is the intensity of an individual measurement, and $\langle I_h \rangle$ is the mean intensity of that reflection.

† Per cent isomorphous difference is calculated using the expression $\sum |F_{PH}| - |F_P| / \sum |F_P| \times 100$, where F_P and F_{PH} refer to the native and the derivative structure factor.

‡ The phasing power of a derivative is defined as the ratio of the amplitude of the r.m.s. heavy-atom scattering factor to the r.m.s. lack of closure.

§ Ethyl mercury phosphate.

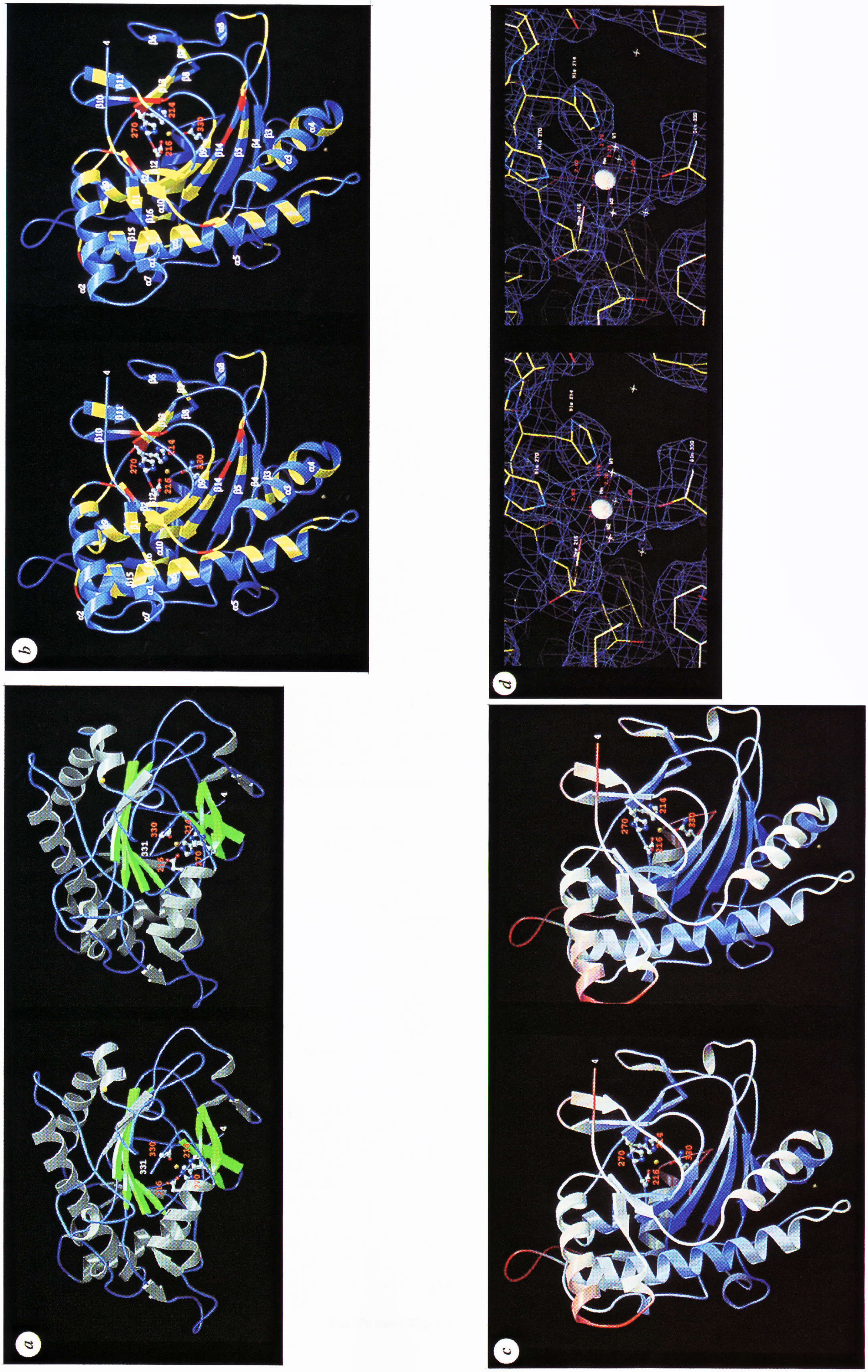
1,4-Diacetoxymercuri-2,3-dimethoxybutane.

Initially IPNS from *Cephalosporium acremonium* was crystallized, but the crystals obtained were of insufficient quality for structure determination[7]. Crystals of a recombinant *Aspergillus nidulans* IPNS[8] complexed with manganese[9] were then obtained and were used to determine the structure of IPNS at a resolution of 2.5 Å (Table 1, Fig. 1). These crystals contain manganese instead of iron at the active site and are relatively stable under aerobic conditions. The secondary structure of the enzyme consists of 10 helices and 16 β-strands (Fig. 1). Eight of the β-strands (β5, β8–14, Figs 1 and 2) fold to give a jelly-roll motif[10]. The jelly-roll motif is common among viral capsid proteins[10] and has been identified in enzymes[11,12]. Unlike other known jelly-roll proteins[10,12], in IPNS the jelly roll is not a completely closed structural unit and the active site is buried within the β-barrel. The side of the jelly roll consisting of β-strands 12, 9, 14 and 5 is extended at both ends by strands 1, 2, and 4, 5, respectively, to form a larger sheet (Fig. 1). The continuation of the β-sheet from β5 allows the C-terminal tail (324–331) extending from the final α-helix (α10) to enter between the two faces of the jelly-roll, allowing Gln 330 to ligate to the metal. This glutamine is present in all independently determined IPNS sequences. The other side of the jelly roll consists of strands 8, 13, 10 and 11. Strands 6 and 7 form a hairpin loop on the surface. The longest α-helix (α6) straddles β-strands 4, 5, 9, 12 and 14 on the outside of the jelly roll and is linked by a sharp turn containing the conserved Gly 165 to another strand of chain (166–178) containing α7 (Fig. 1*b*).

The active-site structure (Fig. 1*d*) reveals the manganese ion, substituting for the ferrous ion at the metal binding site. It is attached by four protein ligands (His 214, Asp 216, His 270, Gln 330) and bears two water molecules, the latter occupying coordination sites directed into a hydrophobic cavity within the protein. It is likely that ACV and dioxygen bind to the coordination sites occupied by the water molecules and Gln 330 (Fig. 3). Such a structural characteristic (an iron-binding site within an unreactive hydrophobic substrate binding cavity) is probably a requirement for this class of enzyme, as it results in the isolation of the reactive complex and subsequent intermediates (Fig. 3) from the external environment. Thus, the reaction can be channelled along a single path, avoiding the many side reactions potentially open to the highly reactive species resulting from the reduction of dioxygen at the metal. The role of enzymes in such processes has been characterized as negative catalysis[13].

Previous sequence comparisons between IPNS, 1-aminocyclopropane-1-carboxylic acid oxidase and some 2-oxo-acid-dependent oxygenases have identified homologous regions, including two containing the active site histidines of IPNS (His 214 and His 270)[14]. For some other 2-oxoglutarate-dependent oxygenases convincing sequence alignments have not been achieved[14,15]. In the light of the structure of IPNS, sequence comparison[16,17] of IPNS with 1-aminocyclopropane-1-carboxylic acid oxidase and related 2-oxo-acid-dependent oxygenases (Figs 1 and 2) leads us to propose that many of them use the same basic structural platform and have thus evolved through a divergent process. In particular the jelly-roll core and the longest α-helix (α6) in IPNS are highly conserved. The presence of the α6 helix may be concerned with stabilization of the distorted jelly-roll motif (Fig. 1). These comparisons also indicate that the active site iron is likely to be coordinated by the side chains of the conserved aspartyl (Asp 216) and histidyl (His 214, His 270) residues for the other members of the family. The conservation of structural and active-site motifs throughout a number of these enzymes suggests that they operate through closely related mechanisms. Gln 330 which is also ligated to the manganese bound to the active site in the IPNS structure and is presumably involved in catalysis is conserved throughout all reported IPNS sequences, but is not conserved through other members of the family. It may be that this difference in coordination chemistry between IPNS and the other members reflects differences in their mechanism and substrate specificity, in particular the fact that IPNS does not use a 2-oxo-acid cosubstrate.

IPNS, 1-aminocyclopropane-1-carboxylic acid oxidase and the related 2-oxo-acid-dependent oxygenases constitute a new structural family which catalyses a wide range of oxidative chemistry, including hydroxylation, desaturation and desaturative


a
b
c
d

◀ FIG. 1 *a–d*, The structure of IPNS. *a*, The three-dimensional structure of isopenicillin *N* synthase showing the conserved jelly-roll motif in green. The structure is viewed from the entrance to the active site. The active-site bound metal and its ligands are buried within the jelly roll. A second and weaker metal-binding site is located on the surface of the molecule in a non-conserved area with Glu 81 and His 82 as protein ligands. The secondary structure elements of the molecule are shown in grey except for the conserved β-strands of the jelly-roll motif (green). Residue 331 is the C-terminal threonine, and residue 4 is the first N-terminal residue that could be localized in the structure. *b*, Distribution of conserved residues in IPNS and related enzymes. The molecule is viewed from the back of the active site. Results from the sequence comparisons of Fig. 2 are mapped on the three-dimensional image of the molecule. The colouring scheme follows the convention of Fig. 2: sequence identities are shown in red and residues with similar physical–chemical properties in yellow. *c*, Structure coloured according to crystallographic temperature factors. Blue shows the smallest values and red the highest. Comparison with *c* shows that the distribution of low *B* factor residues in the structure largely follows the distribution of conserved residues in the family. *d*, Stereo-view of the active site region of IPNS. The figure is taken from the same orientation as *b* and *c*. Electron density was contoured at 1.4σ, where σ represents the r.m.s. electron density for the unit cell. The structure shows a distorted octahedral geometry around the manganese, with the ligands formed from two water molecules and the side chains of His 214, His 270, Asp 216 and Gln 330 (Fig. 2). The latter was unanticipated. An analysis of the protein geometry around the metal site indicates strain in the neighbouring polypeptide structure. The carboxylate of Thr 331 forms a H bond with Arg 87. The metal ligand Gln 330 is the only Ramachandran outlier. The coordination chemistry of IPNS is to our knowledge unique. Probably the most related coordination chemistry is found in the iron and manganese superoxide dismutases, where the metal ligands are provided by the side chains of an aspartyl and three histidyl residues, with a nearby glutaminyl residue which is not directly ligated to the metal. *a*, *b* and *c* were produced using MOLSCRIPT[26], *e* using O[21].

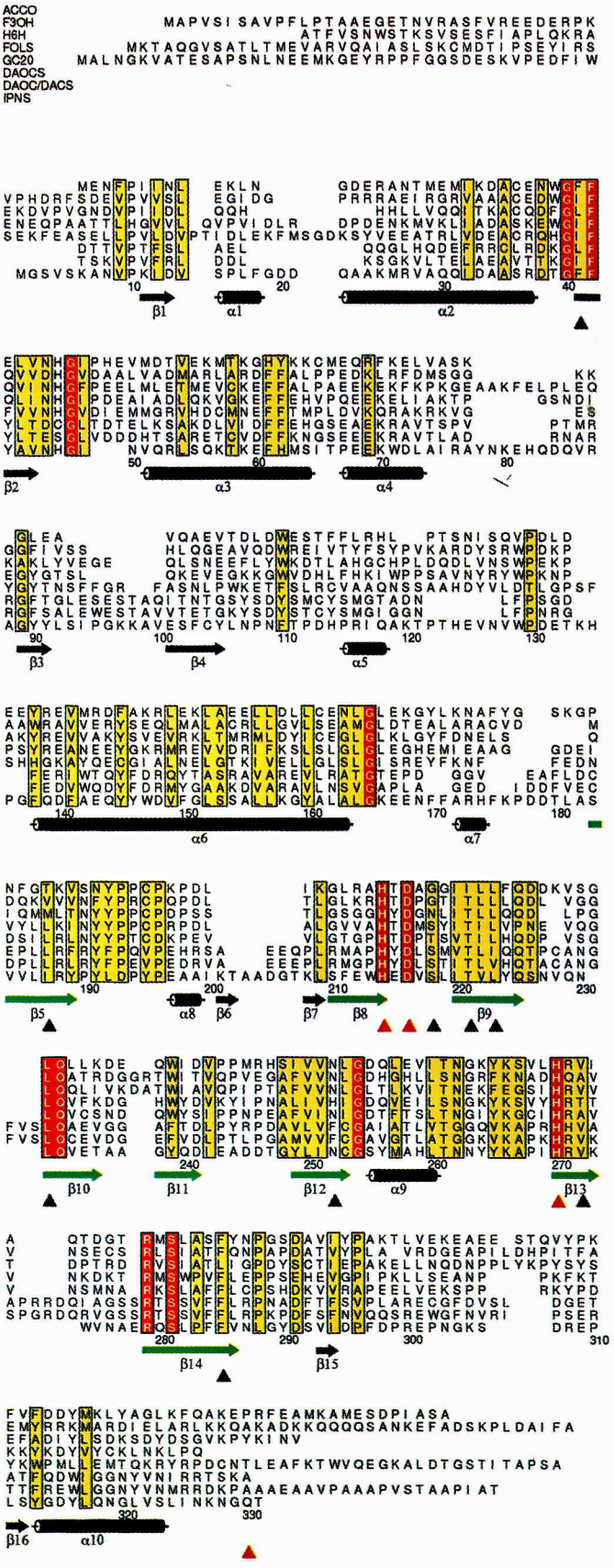


FIG. 2 Sequence comparisons between isopenicillin *N* synthase, 1-aminocyclopropane-1-carboxylic acid oxidase and related 2-oxo-acid-dependent oxygenases. The figure was prepared using ALSCRIPT[17]. Numbering follows the sequence of IPNS for which the secondary structures are shown as cylinders (helices) and arrows (β-strands). Identities are shown with red backgrounds, positions at which amino acids with similar physical–chemical properties are conserved are shown in yellow as in Fig. 1*b*. β-Strands 5 and 8–14 which comprise the jelly-roll motif are shown in green. The metal ligands are indicated by red triangles. Other residues with side chains within 8 Å of the metal are indicated by black triangles. ACCO, 1-Aminocyclopropane-1-carboxylate oxidase from tomato (accession number SwissProt P05116); F3OH, flavanone 3-β-hydroxylase from *Zea mays* (Genbank U04434); H6H, hyoscyamine 6-β-hydroxylase from henbane (SwissProt P24397); FOLS, flavanol synthase from *Petunia hybrida* (Genbank S67953); GC20, gibberellin C-20 oxidase from *Cucurbita maxima* (Genbank X73314); DAOCS, deacetoxycephalosporin C synthase from *Streptomyces clavuligerus* (SwissProt P18548); DAOC/DACS, deacetoxycephalosporin C synthase/deacetylcephalosporin C synthase, a bifunctional enzyme from *C. acremonium* (SwissProt P11935); IPNS from *A. nidulans* (SwissProt P05326). For some of the 2-oxo-acid-dependent oxygenases (such as clavaminic acid synthase[14,15]) preliminary sequence alignments did not indicate sufficient sequence similarity with IPNS to predict the presence of the jelly-roll motif with confidence (see also ref. 14). The conservation of structure throughout the 2-oxo-acid-dependent and related oxygenases will require further structural analyses. Pairwise comparison of the sequences followed by single linkage cluster analysis shows that the proteins all cluster with standard deviation scores >6.0. Scores over 6.0 indicate that the automatic multiple alignment will be correct within most of the core secondary structures[16]. The greatest variations in sequence and hence regions of lowest confidence in the alignment are between residues 1–8, 15–30, 72–137, 168–184, 198–207 and 298–331. It is anticipated that the main differences in tertiary structure between members of the family will be concentrated in these regions.

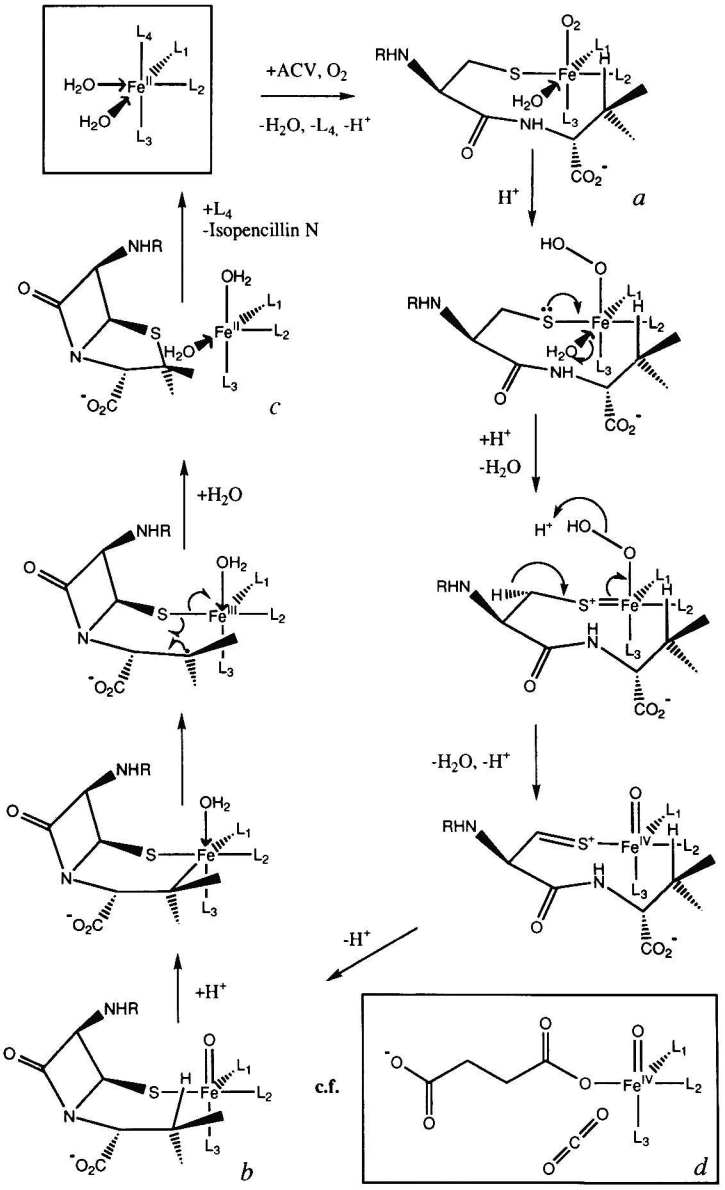


FIG. 3 Proposed scheme for the overall reaction catalysed by IPNS. L_{1-4}, Protein ligands (His 214, Asp 216, His 270, Gln 330); R, *L*-δ-(α-aminoadipoyl). The scheme shows (*a*), the two substrates bound to the iron, *b*, the proposed β-lactam ferryl intermediate and *c*, the enzyme–product complex with IPN and two water molecules. Mechanistic studies on the conversion of ACV to IPN, using kinetic isotope effects and substrate analogues, have led to the proposal of an enzyme-bound β-lactam intermediate directly attached to the iron centre, at the oxidation sate of IV (*b*). This complex may then react to complete carbon–sulphur bond formation by a free-radical type mechanism[5,18]. Spectroscopic investigations on the nature of the metal ligands at the active site of IPNS have indicated that the side chains of 2 or 3 histidyl and an aspartyl residue act as metal ligands[5,27–30]. EPR measurements have been interpreted as showing the presence of a water molecule coordinated to the ferrous iron in an IPNS/ACV/NO complex analogous to *a*[30]. Support for the proposed iron–sulphur bond (in *a*) has come from UV-VIS[27] and EXAFS spectroscopy[28]. Indeed, the active site surrounding the metal-binding site of IPNS is a predominantly hydrophobic cavity, consistent with the generation of a highly reactive intermediate, ligated to the metal by residues that are highly conserved throughout the family, His 214, His 270 and Asp 216 (Fig. 2). The intermediacy of a ferryl species such as *b* implies a protein-bound ligand (L_4), presumably Gln 330, is released from the iron before its formation. Initial modelling studies indicate substrate binding may induce substantial reorganization at the active site. The presently proposed mechanism for the 2-oxo-acid-dependent oxygenases involves initial formation of an enzyme/iron/2-oxo acid/dioxygen complex, from which a ferryl intermediate *d* structurally related to *b* is subsequently formed[6]. The three protein ligands to the metal in this putative complex are probably also the conserved aspartyl and two histidyl residues.

cyclization reactions. Some members have been shown to have a remarkably lax substrate specificity[15,18]. The determination of the structure of IPNS will help investigations into their catalytic mechanisms and will allow systematic engineering to attempt the enzymatic synthesis of new types of otherwise inaccessible β-lactams and improve the biosynthetic route to antibiotics of medicinal importance. Studies on the mechanism may also lead to the design of new non-protein catalysts capable of catalysing in a stereoselective manner reactions that are currently synthetically impossible. □

Received 24 January; accepted 26 April 1995

1. Pang, C.-P. *et al. Biochem. J.* **222,** 789–795 (1984).
2. Fleming, A. *Brit. J. exp. Path.* **10,** 226–236 (1929).
3. Florey, H. W. *et al. Antibiotics* Vol. 2 (Oxford Univ. Press, London, New York, 1949).
4. Crowfoot, D., Bunn, C. W., Rogers-Low, B. W. & Turner-Jones, A. *The Chemistry of Penicillin* (eds Clark, H. T., Johnson, J. R. & Robinson, R.) 310–321 (Princeton Univ. Press, Princeton, New Jersey, 1949).
5. Baldwin, J. E. & Schofield, C. J. in *The Chemistry of β-lactams* (ed. Page, M. I.) Chapter 1 (Blackie, London, 1992).
6. Feig, A. L. & Lippard, S. J. *Chem. Rev.* **94,** 759–805 (1994).
7. Fujishima, Y. *et al. J. molec. Biol.* **242,** 712–714 (1994).
8. Baldwin, J. E., Blackburn, J. M., Sutherland, J. D. & Wright, M. C. *Tetrahedron* **47,** 5991–6002 (1991).
9. Roach, P. L., Schofield, C. J., Baldwin, J. E., Clifton, I. J. & Hajdu, J. *Prot. Sci.* **4,** 1007–1009 (1995).
10. Stuart, D. I. *Curr. Opin. struct. Biol* **3,** 167–174 (1993).
11. Norris, G. E., Sullivan, T. J., Anderson, B. F. & Baker, E. N. *Structure* **2,** 1049–1059 (1994).
12. Srinivasan, N. *et al. Structure* **2,** 1017–1027 (1994).
13. Retey, J. *Angew. Chem.* **29,** 355–361 (1990).
14. Prescot, A. G. *J. exp. Bot.* **44,** 849–861 (1993).
15. Marsh, E. N., Chang, M.D-T. & Townsend, C. A. *Biochemistry* **31,** 12648–12657 (1992).
16. Barton, G. J. *Meth. Enzym.* **183,** 403–428 (1990).
17. Barton, G. J. *Prot. Engng* **6,** 37–40 (1993).
18. Baldwin, J. E. & Bradley, M. *Chem. Rev.* **90,** 1079–1088 (1990).
19. Otwinowski, Z. in *Data Collection and Processing* (eds Sawyer, L., Isaacs, N. W. & Bailey, S.) 55–62 (DL: SCI/R34, Daresbury Laboratory, Warrington, U.K., 1993).
20. Cowtan, K. *Acta crystallogr.* D**50,** 760–763 (1994).
21. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr.* A**47,** 110–119 (1991).
22. Jones, T. A. & Thirup, S. *EMBO J.* **5,** 819–822 (1986).
23. Brünger, A. T., Kuriyan, J. & Karplus, M. *Science* **235,** 458–460 (1987).
24. Brünger, A. T. *Nature* **355,** 472–474 (1992).
25. Read, R. J. *Acta crystallogr.* A**42,** 140–149 (1986).
26. Kraulis, P. J. *J. appl. Crystallogr.* **24,** 946–950 (1991).
27. Orville, A. M. *et al. Biochemistry* **31,** 4602–4612 (1992).
28. Randall, C. R. *et al. Biochemistry* **32,** 6664–6673 (1993).
29. Scott, R. A., Wang, S., Eidsness, M. K., Kriauciunas, A. & Chen, V. J. *Biochemistry* **31,** 4596–4601 (1992).
30. Chen, V. J. *et al. J. biol. Chem.* **264,** 21677–21681 (1989).

ACKNOWLEDGEMENTS. We thank Z. H. Zhang for helpful discussions, J. W. Keeping, J. Pitt, M. Sami, A. Hassan and W. J. Sobey for expert technical assistance, B. Rasmussen and A. Åberg for the outstanding facilities at beam line 4 of the ESRF (Grenoble, France) and the MRC, BBSRC, EPSRC and Zeneca via a DTI Link Scheme for financial support of this work.