LETTERS

Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis

Leah C. Blasiak[1], Frédéric H. Vaillancourt[2]†, Christopher T. Walsh[2] & Catherine L. Drennan[1]

Non-haem Fe(II)/α-ketoglutarate (αKG)-dependent enzymes harness the reducing power of αKG to catalyse oxidative reactions, usually the hydroxylation of unactivated carbons, and are involved in processes such as natural product biosynthesis, the mammalian hypoxic response, and DNA repair[1,2]. These enzymes couple the decarboxylation of αKG with the formation of a high-energy ferryl-oxo intermediate that acts as a hydrogen-abstracting species[2–4]. All previously structurally characterized mononuclear iron enzymes contain a 2-His, 1-carboxylate motif that coordinates the iron[1,2]. The two histidines and one carboxylate, known as the 'facial triad', form one triangular side of an octahedral iron coordination geometry. A subclass of mononuclear iron enzymes has been shown to catalyse halogenation reactions, rather than the more typical hydroxylation reaction[5,6]. SyrB2, a member of this subclass, is a non-haem Fe(II)/αKG-dependent halogenase that catalyses the chlorination of threonine in syringomycin E biosynthesis[5]. Here we report the structure of SyrB2 with both a chloride ion and αKG coordinated to the iron ion at 1.6 Å resolution. This structure reveals a previously unknown coordination of iron, in which the carboxylate ligand of the facial triad is replaced by a chloride ion.

The gene encoding SyrB2 is found in the non-ribosomal peptide synthetase gene cluster for production of the phytotoxin syringomycin E in *Pseudomonas syringae* pv. *syringae* B301D[7] (Fig. 1a). SyrB2 catalyses the Fe(II)-, αKG-, Cl^-- and O_2-dependent conversion of L-Thr to 4-Cl-L-Thr (ref. 5 and Fig. 1b), a modification essential for the antifungal activity of the natural product[8]. SyrB2 is active on L-Thr only when the amino acid is tethered to the phosphopantetheine prosthetic arm of the holo-SyrB1 thiolation domain[5]. Another class of halogenases, the flavin-dependent enzymes, have been structurally characterized and shown to catalyse halogenation at aromatic or electron-rich carbons[9]. Halogenation at aliphatic carbons, as in the natural products syringomycin E (ref. 7), a phytotoxin, barbamide[10], a molluscicidal compound, and victorin[11], a host-selective crop toxin, requires a more reactive catalyst. It has been shown that Fe(II)/αKG-dependent enzymes such as SyrB2 catalyse these more challenging halogenation reactions[5,6]. Other Fe(II)/αKG-dependent halogenases include BarB1/BarB2 and CmaB, which are involved in the biosynthesis of barbamide[10] and the phytotoxin coronatine[6], respectively.

The core of SyrB2 is composed of an eight-stranded anti-parallel β-sandwich 'jelly-roll' motif (comprising strands β2, β12, β5, β10 and strands β9, β6, β11, β3) that defines the cupin superfamily (Fig. 2a and Supplementary Fig. 1). A structural alignment search for SyrB2 using the DALI server[12] produced five matches with *Z*-scores >4, all members of the cupin superfamily. The closest structural relative, with a *Z*-score of 8.5, is human factor-inhibiting hypoxia-inducible factor-1 (FIH-1), an Fe(II)/αKG-dependent asparaginyl hydroxylase involved in the mammalian hypoxic response[13]. SyrB2 is the first structurally characterized mononuclear iron protein that does not contain the classic 2-His, 1-carboxylate iron coordination motif. The two histidine ligands (His 116 and His 235) are present, but the iron has no carboxylate ligand (Fig. 2b). Instead, this site is occupied by a chloride ion in our structure, as judged by an 8σ peak in the $F_o - F_c$ map.

To verify that this iron ligand is a halide ion, we took advantage of the ability of SyrB2 to catalyse bromination[14] as well as chlorination[5]. Unlike chloride, which does not have a strong anomalous signal, bromide has an anomalous peak at 0.9195 Å that is well separated from the iron anomalous peak at 1.7397 Å. Bromide-soaked crystals were prepared and a dispersive difference Fourier map was calculated between the iron and bromide edges, confirming the location of the

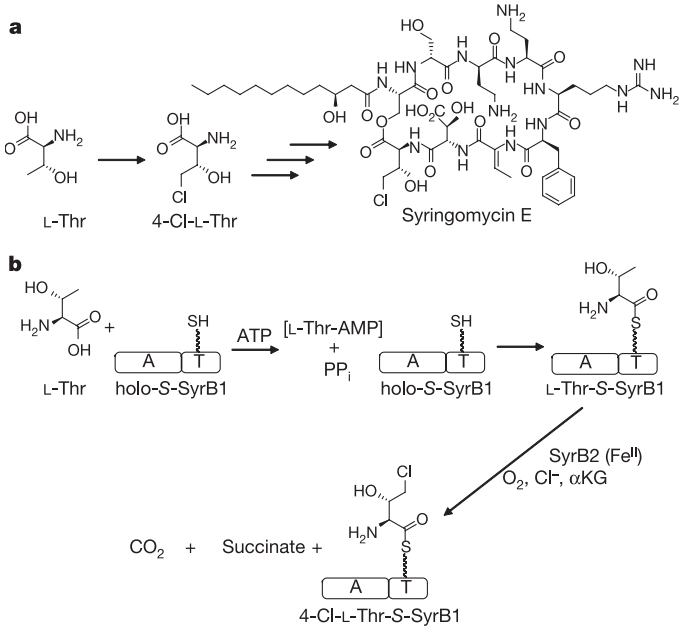

Figure 1 | SyrB2 activity in syringomycin E biosynthesis. a, SyrB2 catalyses the conversion of L-Thr to 4-Cl-L-Thr, which is then incorporated into syringomycin E. **b**, Scheme showing the adenylation of L-Thr by SyrB1, loading of L-Thr-AMP onto the SyrB1 phosphopantetheine group, and subsequent chlorination by SyrB2. The phosphopantetheine prosthetic group on SyrB1, indicated by a wavy line, is posttranslationally attached through a serine residue and functions as a 'swinging arm' for transfer of the covalently attached amino acid between enzyme active sites. A, adenylation domain; T, thiolation domain.

[1]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [2]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA. †Present address: Department of Biological Sciences, Research and Development, Boehringer Ingelheim (Canada) Ltd, Laval, Quebec H7S 2G5, Canada.

Table 1 | Activity and iron content of SyrB2 variants

	Wild-type SyrB2	A118D SyrB2	A118E SyrB2
Activity	+	−	−
Iron content	83.0%	65.3%	74.4%

iron and halide ion (Fig. 2d). Another bromide ion is located on the surface of each SyrB2 molecule, but it is unlikely to have biological significance.

In addition to the two histidines and one chloride ligand, the remaining iron coordination sites are occupied by the αKG and a water molecule. The water molecule binds opposite His 235 at the site where dioxygen would probably bind during catalysis (Fig. 2b). αKG coordinates iron in a bidentate fashion, with the C1 carboxylate opposite His 116, and the C2 ketone opposite the chloride ligand. The αKG is anchored in the active site through hydrogen bonds between the C5 carboxylate and Thr 113, Arg 248 and Trp 145 (Fig. 2c). Bidentate binding of αKG and the arginine and threonine hydrogen-bonding interactions are conserved across many members of the Fe(II)/αKG-dependent family[2], but the interaction of Trp 145 is unique to SyrB2. Phe 104 stacks above the planar αKG, whereas Arg 254 is situated 2.9 Å above the αKG and may be important in positioning the C1 carboxylate. A similarly placed arginine or asparagine is observed in many other family members[2].

The chloride ligand in SyrB2 sits in a largely hydrophobic pocket composed of residues Ala 118, Phe 121 and the β-carbon of Ser 231. However, it also forms interactions with two water molecules, which in turn form hydrogen bonds with Thr 143 and Asn 123 (Fig. 2c). The interactions between the chloride ligand and water may be involved in lengthening the Fe–Cl bond, which at 2.44 Å is slightly longer than the average six-coordinate Fe(II)–Cl bond length of 2.30 Å found in the Cambridge Structural Database[15] and other model compounds[16]. All other iron–ligand bond lengths are typical of Fe(II)/αKG-dependent enzymes.

The position of Ala 118 in SyrB2 seems to have a role in the protein's halogenase activity by creating space in the active site for the chloride ligand to bind iron. Ala 118 is completely conserved among known mononuclear iron halogenases (Supplementary Fig. 2a). A structure-based sequence alignment of SyrB2 with non-halogenase Fe(II)/αKG-dependent enzymes shows that the active site residues are located at structurally equivalent positions (Supplementary Fig. 2b). Ala 118 in SyrB2 replaces the conserved aspartate or glutamate that serves as an iron ligand in other mononuclear iron enzymes. An alignment of SyrB2 with its most structurally similar relative, FIH-1, shows replacement of the aspartate by alanine, creating space for chloride to bind (Fig. 3a). If a glutamate is modelled in place of Ala 118 in the SyrB2 structure, it is ideally positioned for iron coordination. To verify that substitution of Ala 118 with aspartate or glutamate would abolish chlorination activity while retaining iron-binding ability, A118D and A118E variants were expressed, purified and reconstituted with Fe(II) and αKG. Both variants bound Fe(II), but the mutations completely abrogated chlorination activity (Table 1 and Supplementary Fig. 3).

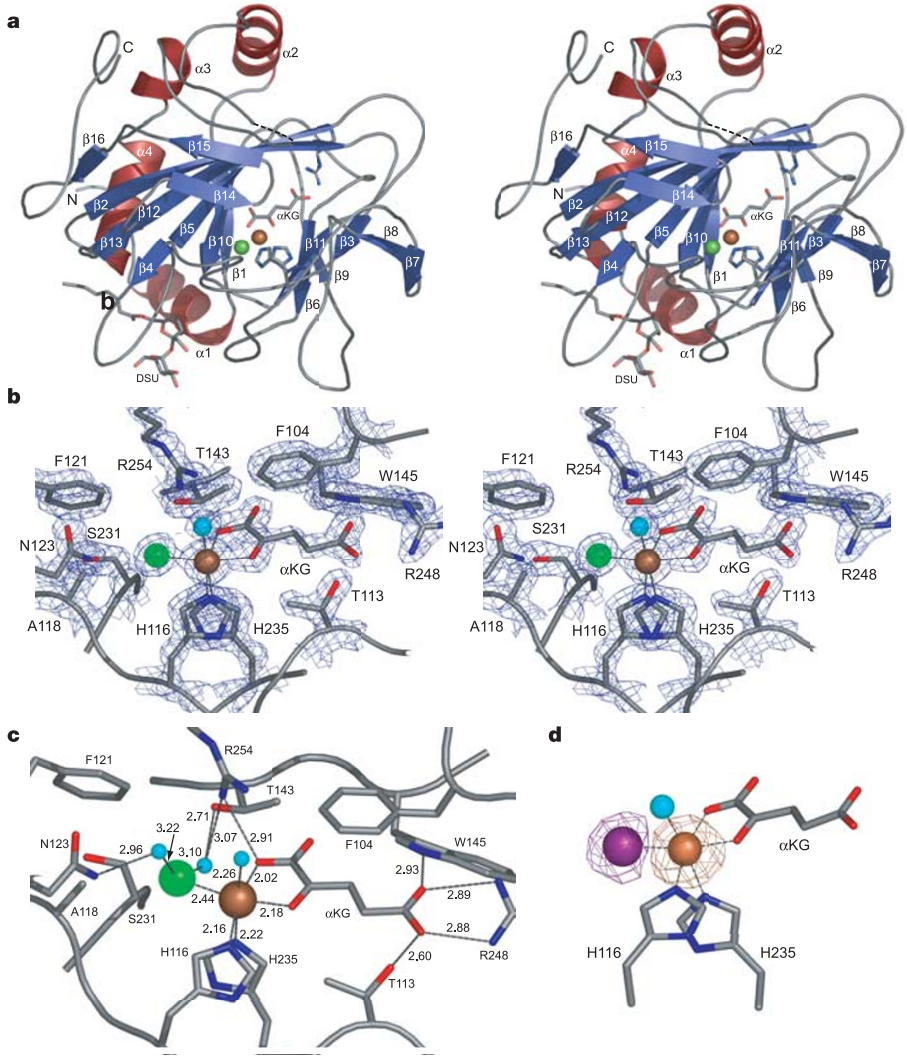

Figure 2 | Structure of SyrB2 with bound iron, chloride and αKG. Iron is shown in brown, chloride in green. **a**, Stereo view of the SyrB2 structure. Two protein ligands of iron (His 116 and His 235), αKG and Arg 248 are shown in stick representation. A disordered loop (residues 56–60) is indicated by the broken line. A detergent molecule (DSU) from the crystallization condition packs along the bottom surface of the structure. **b**, Stereo view of the SyrB2 active site. A $2F_o - F_c$ composite omit map contoured to 2σ is shown in blue mesh. **c**, SyrB2 active site distances. **d**, Dispersive difference Fourier map from bromide-containing crystals, calculated by subtracting data collected at 0.9197 Å (bromide edge) from data collected at 1.7340 Å (iron edge). The purple mesh ($+4\sigma$) confirms the location of the halide ion; the brown mesh (-4σ) confirms the location of the iron.

Thus, in SyrB2, it seems that nature chooses alanine at this position instead of an aspartate or glutamate residue to create an iron coordination site for chloride binding.

Despite the requirement for a 'facial triad' of ligands in all other Fe(II)/αKG dependent hydroxylases, it seems unlikely that a carboxylate ligand could bind to iron in SyrB2 in any conformation or stage of catalysis. The closest candidates (Asp 114 and Asp 119) are >10 Å from the iron, pointing in the opposite direction (Fig. 3a). Asp 114 forms part of the β-sandwich core, and Asp 119 is not conserved. The iron coordination structure with only two direct protein ligands also helps to explain the low affinity of apo-SyrB2 for Fe(II) during reconstitution, which yields only ~35% incorporation, as opposed to reconstitution with αKG followed by Fe(II), which yields ~90% incorporation[5].

So far, SyrB2 crystallizes only in the presence of αKG, indicating that this cofactor provides structural stability to the protein. The structure of SyrB2 with only αKG bound has two channels leading to the active site (Fig. 3b). The loops covering the active site in this structure have higher *B*-factors (by 10–20 Å^2) than the average for the rest of the structure, indicating that they have some degree of mobility. Once iron and chloride bind, the protein closes down around the active site (Fig. 3c). This closure involves motions of two loops and a β-hairpin (Fig. 3d). The substrate of SyrB2, L-Thr, which is loaded on the phosphopantetheine arm of the 66-kDa SyrB1 protein, probably enters the active site through the remaining narrow tunnel above αKG (Fig. 3c). The distance from the iron to the edge of the tunnel is ~17 Å, similar to the length of the extended phosphopantetheine arm (~20 Å). Access through this tunnel is limited by the side chain of Phe 196 (Fig. 3c and Supplementary Fig. 4), which may be involved in gating the entrance to the active site. We cannot rule out the possibility, however, that SyrB2 undergoes a conformational change on interaction with SyrB1 that opens a different passageway to the active site.

In predictions of how the Fe(II)/αKG-dependent enzyme scaffold might be used for halogenation, a persistent issue was the location of the chloride. Without a structure, it was presumed that SyrB2 would contain the characteristic HX(D/E)X$_n$H motif for iron binding. The sequence of SyrB2 does contain a conserved HXD motif, but it is located >15 Å from the active site in our structure. If we assumed three protein ligands and bidentate αKG, only one coordination site for dioxygen would remain, leaving none for chloride binding. If this were true, chloride binding to iron would have to take place after decarboxylation of αKG (and subsequent release of succinate or CO_2) and formation of the ferryl-oxo intermediate. However, the high-energy ferryl-oxo species is expected to be very reactive; thus, it would be unfavourable to form this intermediate before chloride were in position to react. The missing carboxylate ligand in SyrB2 elegantly solves this mechanistic dilemma. Our crystallographic data confirm that halide ions bind directly to iron, before decarboxylation of αKG and formation of a reactive ferryl-oxo species.

Based on the structure of the resting state of SyrB2 and by analogy with the Fe(II)/αKG-dependent hydroxylation mechanism of TauD[17], we propose a working mechanism for halogenation by SyrB2 (Fig. 4). As observed crystallographically, αKG, chloride and water coordinate the iron in the resting Fe(II) state (Fig. 4, species A). L-Thr-*S*-SyrB1 binding would then exclude water from the active site and allow dioxygen to bind (Fig. 4, species B and C). Decarboxylation of αKG would lead to the formation of a high-energy ferryl-oxo intermediate (Fig. 4, species D), which would then abstract a hydrogen atom from the substrate. The substrate radical would abstract the chlorine atom (Fig. 4, species E), producing chlorinated L-Thr-*S*-SyrB1 and regenerating the reduced Fe(II) centre. After formation of the substrate carbon radical, however, some competition between transfer of Cl$^\bullet$ and OH$^\bullet$ would be expected. No threonine hydroxylation side reaction has been detected, indicating that Cl$^\bullet$ transfer is greatly favoured[5]. There are several possible explanations for this selectivity, including the positioning of the substrate in the active site during the reaction, and the lower potential of Cl$^\bullet$ versus OH$^\bullet$, as has been proposed to explain similar reactivity in model complexes[5, 18].

The crystal structure of SyrB2 reveals an unexpected iron scaffold that was not predicted from the protein sequence and provides insight into the function of Fe(II)/αKG-dependent halogenases. SyrB2 could serve as an excellent model system for studying the

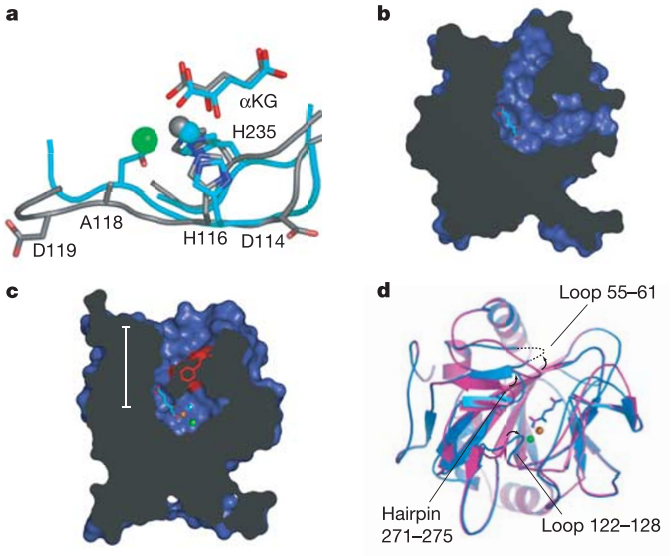

Figure 3 | Structural comparisons. **a**, View of the active site from a structural alignment of SyrB2 (grey) and FIH-1 (cyan). Alignment was done with LSQMAN[29]. Ninety-nine residues (33% of SyrB2) align with a root mean square deviation of 1.95 Å. Iron and chloride from SyrB2 are shown in grey and green, respectively; iron from FIH-1 is shown in cyan. Labelled amino acids correspond to SyrB2. **b**, Cut-away view of SyrB2 with bound αKG and no iron. Two channels lead to the active site. Phe 196 is not visible owing to clipping planes. **c**, Cut-away view of Cl–Fe(II)–αKG–SyrB2 in a similar orientation to **b**, but the molecular surface is clipped to show Phe 196 (red) clearly. SyrB2 closes down around the bound iron (brown) and chloride (green), leaving only one opening to the active site. The water molecule located at the predicted dioxygen-binding site is shown in cyan. Scale bar, ~15 Å. **d**, αKG–SyrB2 (magenta) and Cl–Fe(II)–αKG–SyrB2 (blue). Arrows show change in conformation on binding of iron and chloride.

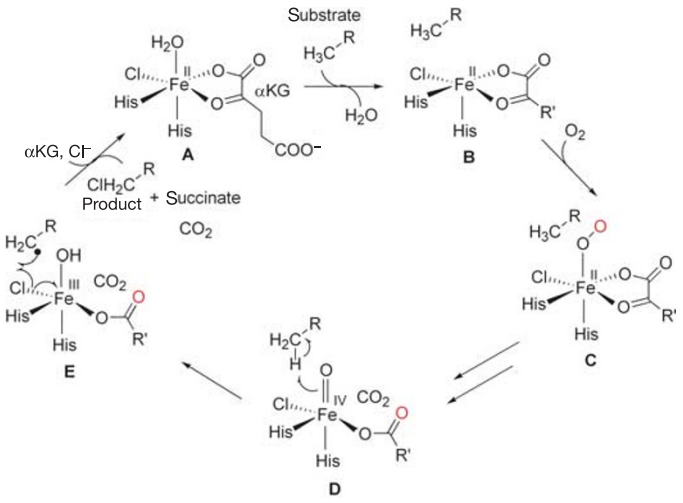

Figure 4 | Proposed mechanism for the reaction catalysed by SyrB2. See text for a description of the proposed steps. The crystallized form of SyrB2 is labelled A. The oxygen atom inserted into αKG is shown in red.

reactivity of iron halides and the detailed mechanism of halogenation by mononuclear iron enzymes.

METHODS

Protein purification, crystallization, data collection, site-directed mutagenesis and assays of the A118D and A118E variants were done as described in Supplementary Methods.

Structure determination. Phases were determined from a SAD experiment on seleno-methionine-labelled SyrB2 with αKG. Sixteen selenium sites (eight per molecule) were found in SOLVE[19] and refined in SHARP[20], with an acentric figure of merit of 0.431 to 2.3 Å resolution. To improve the quality of the electron density map, non-crystallographic averaging (over the two molecules in the asymmetric unit) and solvent flattening were done in DM[21]. The initial model was built manually using XFIT[22] and refined in CNS[23] with non-crystallographic symmetry (NCS) restraints applied. Several rounds of iterative model building and refinement without NCS restraints followed. When the *R*-factors dropped below 25.5% (work) and 28.0% (free), the model was refined against the high-resolution αKG–SyrB2 and Fe(II) data sets. The same set of test reflections (extended for higher resolution structures) was used to calculate R_{free} for all models. Final rounds of refinement for all structures were done with the program REFMAC5 (ref. 24). Simulated annealing composite omit maps calculated in CNS were used to verify structures throughout refinement.

The final model of αKG–SyrB2 contains residues 2–310 in molecule A, and residues 3–55, 61–133 and 139–310 in molecule B out of 310 residues. The final models of Cl/Br–Fe(II)–αKG–SyrB2 contain residues 3–56 and 61–310 in molecule A, and residues 4–53 and 62–310 in molecule B. All structures contain one detergent molecule from the crystallization condition (*n*-decanoyl sucrose; DSU), bound on the surface of molecule A. The topology and parameter files for αKG were obtained from the REFMAC5 (ref. 24) dictionary, and the DSU dictionary was generated in the program SKETCHER[25]. Planar restraints on the sugar rings were removed for refinement. A summary of the final refinement statistics is available in Supplementary Table S1. Assessment of the models in PROCHECK[26] found no unfavourable geometries and >90% of residues were in the 'most favoured' conformation. A dispersive difference Fourier map was calculated in CNS by subtracting the bromide inflection data from the iron inflection data using phases calculated from a model that contained only protein atoms. The observed negative dispersive difference electron density confirmed the presence of one iron site per asymmetric unit, and the positive dispersive difference electron density indicated the presence of two bromide sites (one bound to iron in the active site and one on the surface) per molecule. Secondary structure assignment was done in DSSP[27], and figures were generated in PyMOL[28].

Received 3 November; accepted 22 December 2005.

1. Koehntop, K. D., Emerson, J. P. & Que, L. Jr. The 2-His-1-carboxylate facial triad: a versatile platform for dioxygen activation by mononuclear non-heme iron(II) enzymes. *J. Biol. Inorg. Chem.* **10,** 87–93 (2005).
2. Hausinger, R. P. Fe(II)/α-ketoglutarate-dependent hydroxylases and related enzymes. *Crit. Rev. Biochem. Mol. Biol.* **39,** 21–68 (2004).
3. Price, J. C., Barr, E. W., Hoffart, L. M., Krebs, C. & Bollinger, J. M. Jr. Kinetic dissection of the catalytic mechanism of taurine:α-ketoglutarate dioxygenase (TauD) from *Escherichia coli*. *Biochemistry* **44,** 8138–8147 (2005).
4. Price, J. C., Barr, E. W., Tirupati, B., Bollinger, J. M. Jr & Krebs, C. The first direct characterization of a high-valent iron intermediate in the reaction of an α-ketoglutarate-dependent dioxygenase: a high-spin Fe^{IV} complex in taurine/α-ketoglutarate dioxygenase (TauD) from *Escherichia coli*. *Biochemistry* **42,** 7497–7508 (2003).
5. Vaillancourt, F. H., Yin, J. & Walsh, C. T. SyrB2 in syringomycin E biosynthesis is a non-heme Fe^{II} α-ketoglutarate and O_2 dependent halogenase. *Proc. Natl Acad. Sci. USA* **102,** 10111–10116 (2005).
6. Vaillancourt, F. H., Yeh, E., Vosburg, D. A., O'Connor, S. E. & Walsh, C. T. Cryptic chlorination by a non-haem iron enzyme during cyclopropyl amino acid biosynthesis. *Nature* **436,** 1191–1194 (2005).
7. Guenzi, E., Galli, G., Grgurina, I., Gross, D. C. & Grandi, G. Characterization of the syringomycin synthetase gene cluster. A link between prokaryotic and eukaryotic peptide synthetases. *J. Biol. Chem.* **273,** 32857–32863 (1998).
8. Grgurina, I. *et al.* Relevance of chlorine-substituent for the antifungal activity of syringomycin and syringotoxin, metabolites of the phytopathogenic bacterium *Pseudomonas syringae* pv. *syringae*. *Experientia* **50,** 130–133 (1994).
9. Dong, C. *et al.* Tryptophan 7-halogenase (PrnA) structure suggests a mechanism for regioselective chlorination. *Science* **309,** 2216–2219 (2005).
10. Chang, Z. *et al.* The barbamide biosynthetic gene cluster: a novel marine cyanobacterial system of mixed polyketide synthase (PKS)–non-ribosomal peptide synthetase (NRPS) origin involving an unusual trichloroleucyl starter unit. *Gene* **296,** 235–247 (2002).
11. Walton, J. D. Host-selective toxins: agents of compatibility. *Plant Cell* **8,** 1723–1733 (1996).
12. Holm, L. & Sander, C. Mapping the protein universe. *Science* **273,** 595–603 (1996).
13. Dann, C. E., Bruick, R. K. & Deisenhofer, J. Structure of factor-inhibiting hypoxia-inducible factor 1: an asparaginyl hydroxylase involved in the hypoxic response pathway. *Proc. Natl Acad. Sci. USA* **99,** 15351–15356 (2002).
14. Vaillancourt, F. H., Vosburg, D. A. & Walsh, C. T. Dichlorination and bromination of a threonyl-*S*-carrier protein by the nonheme Fe^{II} halogenase SyrB2. *ChemBioChem* (in the press).
15. Allen, F. H. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr. B* **58,** 380–388 (2002).
16. Goldsmith, C. R., Jonas, R. T., Cole, A. P. & Stack, D. P. A spectrochemical walk: single-site perturbation within a series of six-coordinate ferrous complexes. *Inorg. Chem.* **41,** 4642–4652 (2002).
17. Grzyska, P. K. *et al.* Steady-state and transient kinetic analyses of taurine/α-ketoglutarate dioxygenase: effects of oxygen concentration, alternative sulfonates, and active-site variants on the Fe^{IV}-oxo intermediate. *Biochemistry* **44,** 3845–3855 (2005).
18. Kojima, T., Leising, R. A., Yan, S. & Que, L. Jr. Alkane functionalization at nonheme iron center. Stoichiometric transfer of metal-bound ligands to alkane. *J. Am. Chem. Soc.* **115,** 11328–11335 (1993).
19. Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr. D* **55,** 849–861 (1999).
20. La Fortelle, E. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement in MIR and MAD methods. *Methods Enzymol.* **276,** 472–494 (1997).
21. Cowtan, K. DM: an automated procedure for phase improvement by density modification. *CCP4/ESF-EACBM Newslett. Protein Crystallogr.* **31,** 34–38 (1994).
22. McRee, D. E. XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125,** 156–165 (1999).
23. Brunger, A. T. *et al.* Crystallography and NMR system (CNS): a new software system for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
24. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53,** 240–255 (1997).
25. Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D* **59,** 1131–1137 (2003).
26. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK—a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26,** 283–291 (1993).
27. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22,** 2577–2637 (1983).
28. Delano, W. L. The PyMOL Molecular Graphics System ⟨http://www.pymol.org⟩ (2002).
29. Kleywegt, G. J. & Jones, T. A. A super position. *CCP4/ESF-EACBM Newslett. Protein Crystallogr.* **31,** 9–14 (1994).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. W. Nix and T. Doukov for help with data collection; and M. Fischbach for critically reading the manuscript. The Advanced Light Source and Stanford Synchrotron Radiation Laboratory are supported by the US Department of Energy. The SSRL Structural Molecular Biology Program is also supported by the National Institutes of Health. This work was supported in part by grants from the NIH (to C.L.D., to C.T.W. and to L.C.B.), a Merck-sponsored Fellowship of the Helen Hay Whitney Foundation (to F.H.V.) and a Natural Sciences and Engineering Research Council of Canada Postdoctoral Fellowship (to F.H.V.).

Author Information The coordinates and structure factors for αKG–SyrB2, Cl–Fe(II)–αKG–SyrB2 and Br–Fe(II)–αKG–SyrB2 have been deposited in the Protein Data Bank with accession codes 2FCT, 2FCU and 2FCV, respectively. Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to C.L.D. (cdrennan@MIT.EDU).