Supplementary information

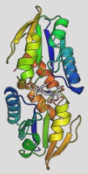
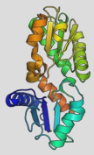
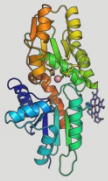
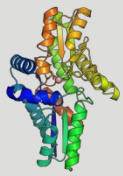
Structure of sirohydrochlorine ferrochelatase SirB: The last of the structures of the class II chelatase family

Takashi Fujishiro,* Yukino Shimada,† Ryosuke Nakamura,† Miho Ooi

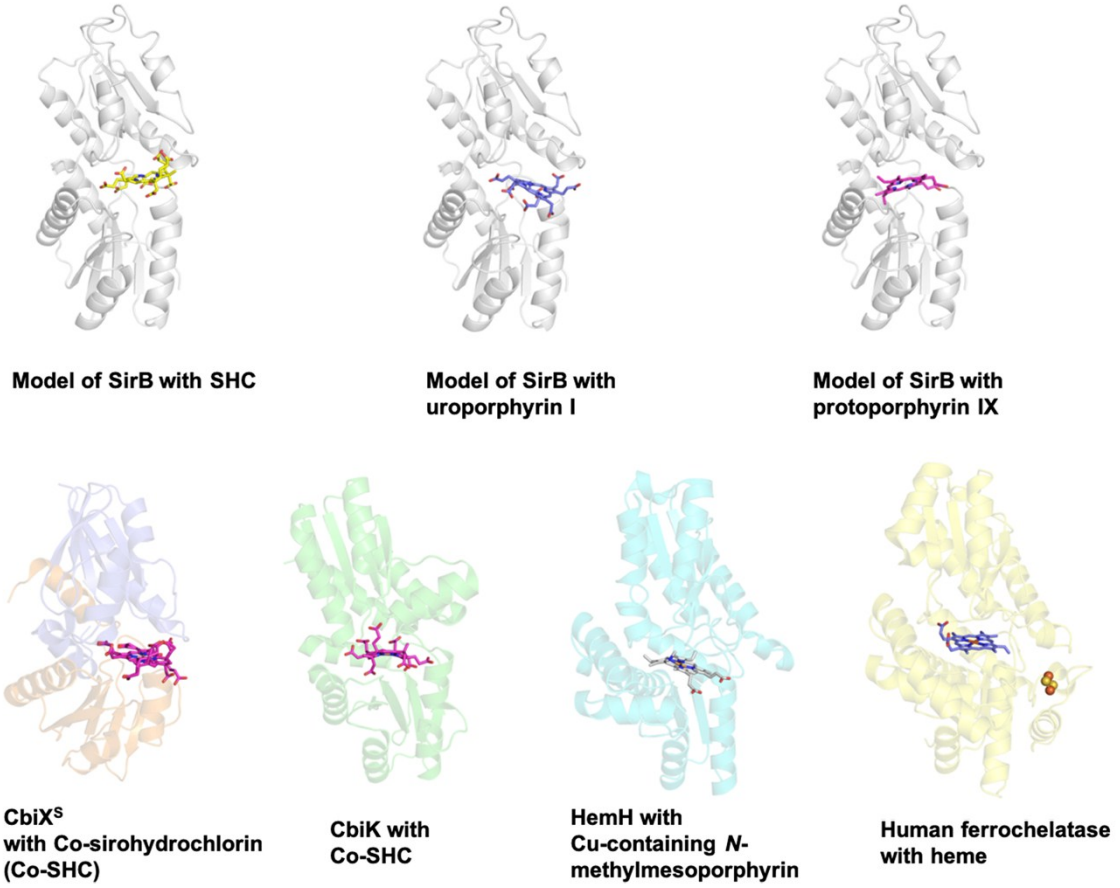
Department of Biochemistry and Molecular Biology, Graduate School of Science and Engineering, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama 338-8570, Japan

* Corresponding author (T. Fujishiro)
Tel: +81-48-858-9293
Fax: +81-48-858-4310
E-mail tfujishiro@mail.saitama-u.ac.jp (T. Fujishiro)

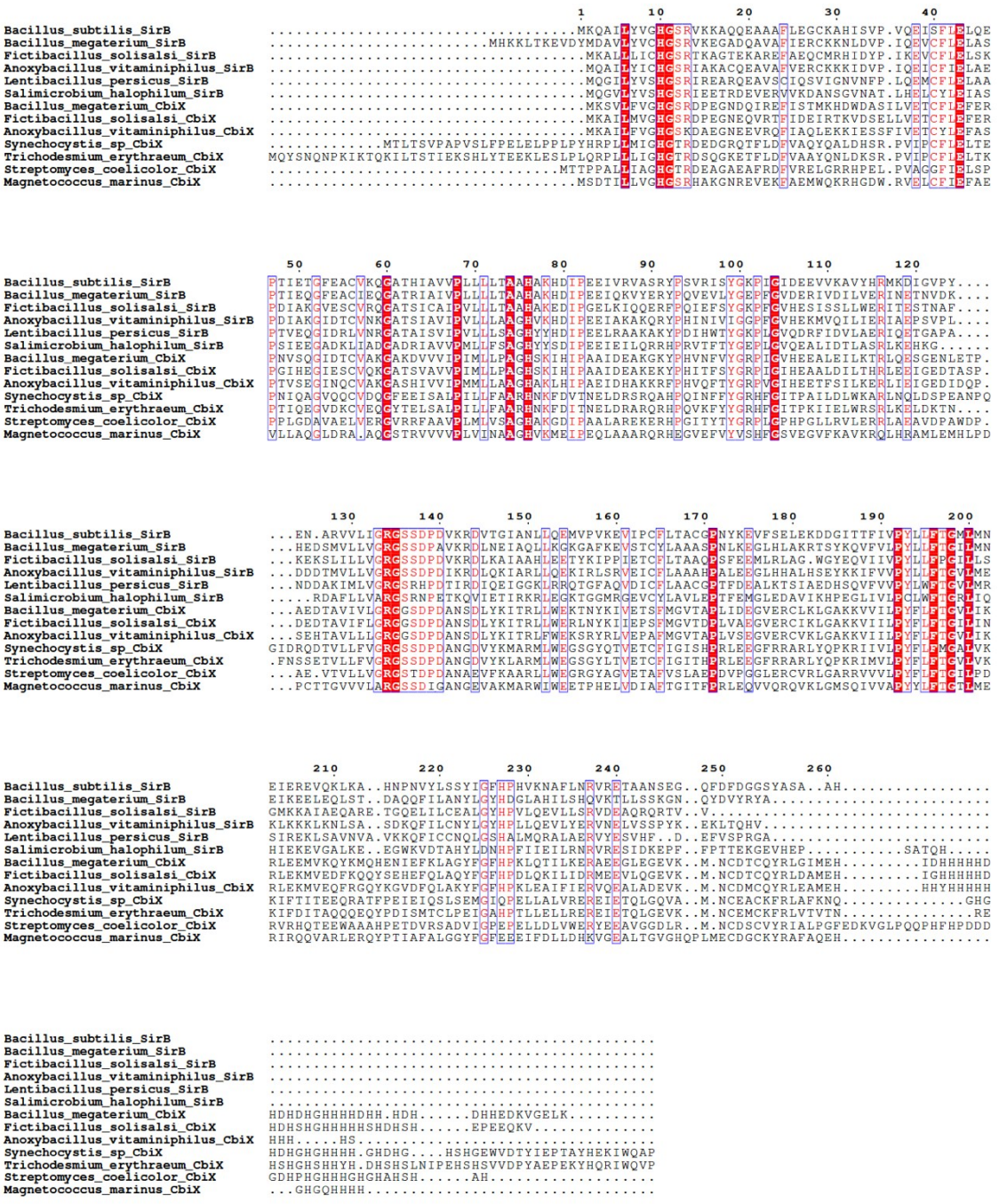
†These authors are equally contributed to this work.

name	"Ancestral" CbiX^{S} (or CfbA)	CbiX	SirB	CbiK	HemH
Structure					
Organization	Homodimer	monomer	monomer	monomer	monomer
Active site	Central dimer interface	N-terminal domain	N-terminal domain	C-terminal domain	C-terminal domain
Physiological function	Co-chelatase (Ni-chelatase)	Co-chelatase	Co-chelatase	Co-chelatase	Fe-chelatase

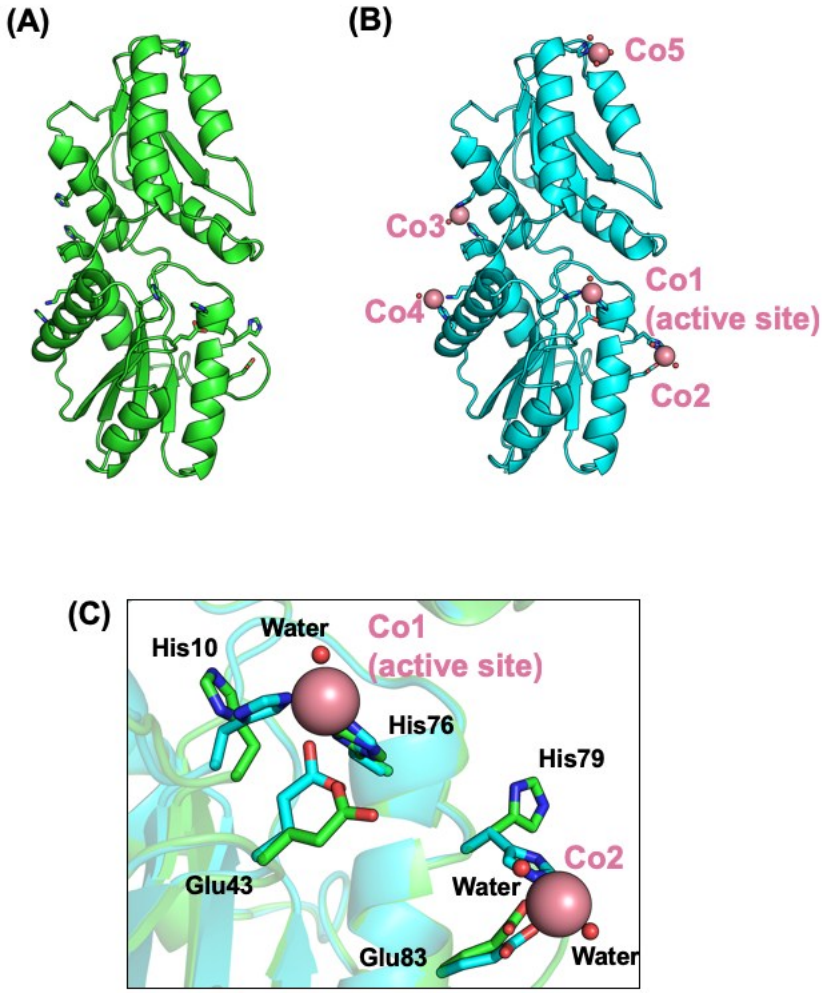
Supplementary Figure 1. Summary of evolutionarily related class II chelatase family proteins, and the SirB structure determined in this work. The cobalt-sirohydrochlorin chelatase CbiX^{S},[1, 2] which has recently been characterised as nickel chelatase CfbA[3, 4] involved in the biosynthesis of coenzyme F430 in methanogenic archaea, is regarded as an "ancestral" class II chelatase. SirB[5, 6] and CbiX[7] are also cobalt-chelatases catalysing the formation of cobalt-sirohydrochlorin, and are classified as class II chelatases having metal-binding active sites at their N-terminal domains. In contrast, HemH[8-11] and CbiK[2, 12, 13] have their metal-binding active sites at the C-terminal domains. These two chelatases, HemH and CbiK, however, play different roles: HemH, for the biosynthesis of heme, using iron and protoporphyrin IX; CbiK, for the biosynthesis of cobalt-sirohydrochlorin. The protein structures were depicted using PyMOL (Schrödinger, LLC) and rainbow-colored from the N-terminus (blue) to the C-terminus (red).



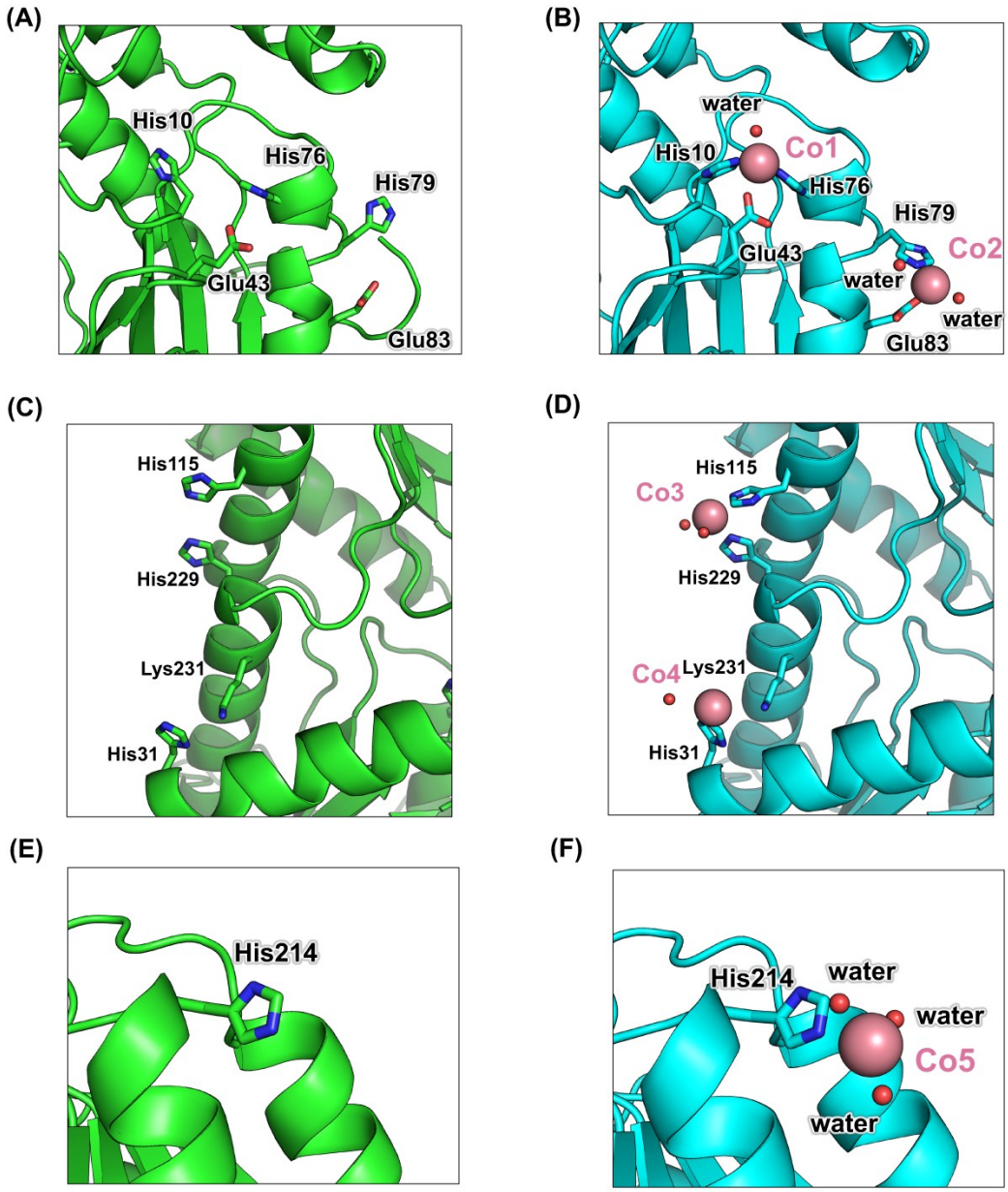
Supplementary Figure 2. Model structures of Co^{2+}-bound SirB with sirohydrochlorin (SHC), protoporphyrin IX, and uroporphyrin I. These docking models were compared with the crystal structures of $CbiX^S$ with cobalt-SHC (PDB ID: 2XWQ),[2] CbiK with cobalt-SHC (PDB ID: 2XWP),[2] HemH with copper-containing *N*-methyl mesoporphyrin (PDB ID: 1C9E),[14] and human ferrochelatase with heme (PDB ID: 3HCN).[14] Tetrapyrroles were shown as stick models. In human ferrochelatase, iron and sulphur of the [2Fe-2S] cluster are shown as yellow and brown spheres, respectively.



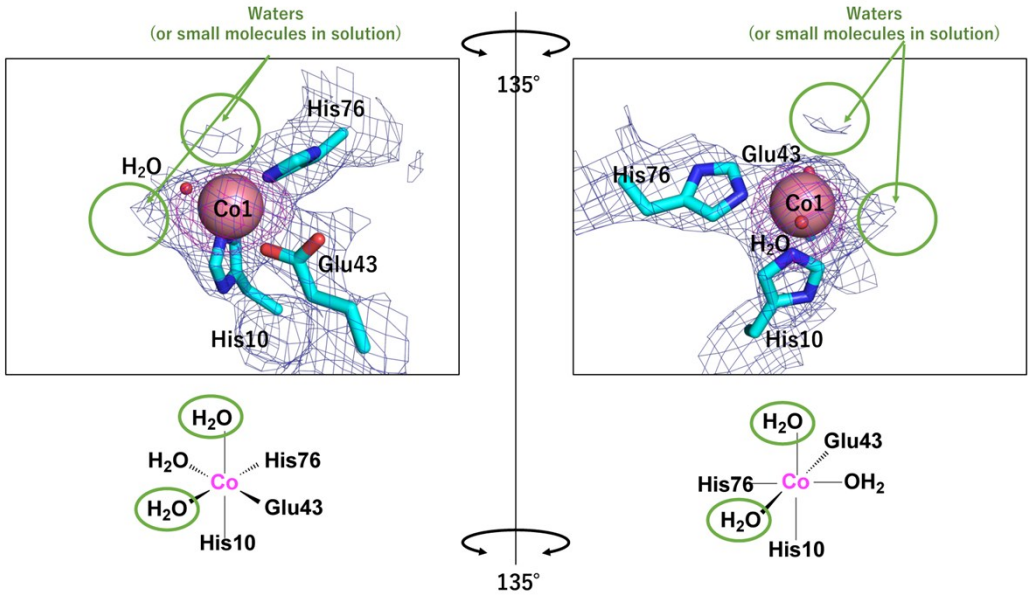
Supplementary Figure 3. Amino acid sequence alignments of SirB and CbiX, using Clustal omega.[15] The number of amino acid residues of *Bacillus subtilis* SirB are shown on the alignments. His10, Glu43, and His76 of *B. subtilis* SirB, which are ligands for cobalt/iron substrate binding, are strictly conserved in both SirB and CbiX. The figure was made using ESPript 3.[16] White letters on a red background indicate strictly conserved amino acid residues. Red letters in blue boxes indicate well-conserved amino acids.



Supplementary Figure 4. Overall structures of (A) metal-free SirB and (B) Co^{2+}-bound SirB. (C) Superimposition of active sites of metal-free and Co^{2+}-bound SirB. Co^{2+} and water molecules are shown in pink and small red spheres, respectively. The amino acid side chains ligating to the Co^{2+} ions are shown as stick models.



Supplementary Figure 5. Comparison of metal-binding sites between metal-free SirB and Co^{2+}-bound SirB. (A) Metal-free SirB and (B) Co^{2+}-bound SirB at the Co1 and Co2 sites; (C) Metal-free SirB and (D) Co^{2+}-bound SirB at the Co3 and Co4 sites; (E) Metal-free SirB and (F) Co^{2+}-bound SirB at the Co5 site.



Supplementary Figure 6. The geometry of the Co1 site. Green circles with arrows indicate plausible $2F_o$-F_c electron density corresponding to waters or solvents, although the density was too small to model anything. The geometry appears to be octahedral with slight distortion, not tetrahedral. The $2F_o$-F_c electron density and Co-anomalous difference maps were contoured at 1σand at 4σ, respectively.

Supplementary Table 1. The *B*-factor of each metal site of Co^{2+}-bound SirB. The occupancy of each metal was 1.0.

Co^{2+}-bound SirB	
Site	*B*-factor (Å^2)
Co1	68.5
Co2	128.3
Co3	82.8
Co4	102.8
Co5	119.8

Supplementary Table 2. List of primers. The *Nco*I and *Xho*I restriction sites are underlined.

Primer	**Sequence**
NcoI-BsSirA-F	5'-ATACCATGGGGAAAGTATATATTGTAGGAG-3'
XhoI-BsSirC-R	5'-GTGCTCGAGGATCCTTATTTCTTTCTTCTCATCTGATAAAGCC-3'
His6-BsSirB-F	5'-ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCATGAAACAAGCAATTTTATATGTC-3'
BsSirA-R	5'-TTTTACAACGCCTCGCTTAAATC-3'

Supplementary Table 3. Data collection and refinement statistics.

	SirB (metal-free)
Data collection	
Temperature (K)	100
Wavelength (Å)	1.000
Resolution range[(a)]	39.6– 2.0 (2.1–2.0)
Space group	$P2_1$
Cell dimensions	
a, *b*, *c* (Å)	69.7 53.8 80.4
α, β, γ (°)	90 105.4 90
Total reflections[(a)]	258734 (24477)
Unique reflections[(a)]	38718 (3817)
Redundancy[(a)]	6.7 (6.4)
Completeness (%)[(a)]	99.7 (99.1)
I/σ_I[(a)]	20.2 (2.55)
R_{sym} (%)[(a), (b)]	7.5 (77.2)
$CC_{1/2}$[(a)]	1.00 (0.84)
Refinement	
Resolution range (Å) [(a)]	31.4– 2.0 (2.1–2.0)
R_{work} (%)[(a), (c)]	18.5 (26.4)
R_{free} (%)[(a), (d)]	21.6 (30.1)
Rmsd bond length (Å)[(e)]	0.005
Rmsd bond angle (°)[(e)]	1.06
Clashcore, all atoms	4.9 (98th percentile)
Average *B*-factor (Å^2)	38.0
PDB ID	5ZT8

[a]The values in parentheses are for the highest resolution shell. [b]$R_{sym} = \sum\sum|I - \langle I \rangle|/\sum I$, where I is the intensity of each reflection. [c]$R_{work} = \sum||F_o| - k|F_c||/\sum|F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively. [d]R_{free} was calculated as the R_{work} for 5% of the reflections that were not included in the refinement. [e]rmsd, root mean square deviation.

	Co^{2+}-bound SirB	Co^{2+}-bound SirB (Co peak)
Data collection		
Temperature (K)	100	100
Wavelength (Å)	1.000	1.605
Resolution range[(a)]	41.8– 2.9 (3.0–2.9)	41.8– 3.1 (3.2–3.1)
Space group	$P2_1$	$P2_1$
Cell dimensions		
a, *b*, *c* (Å)	65.0 52.1 75.5	65.0 52.2 75.6
α, β, γ (°)	90 112.1 90	90 112.1 90
Total reflections[(a)]	66609 (6910)	42871 (390)
Unique reflections[(a)]	10156 (1024)	7555 (250)
Redundancy[(a)]	6.6 (6.7)	5.7 (1.5)
Completeness (%)[(a)]	99.8 (100.0)	86.9 (29.7)
I/σ_I[(a)]	12.68 (1.8)	13.7 (1.1)
R_{sym} (%)[(a), (b)]	11.0 (102.0)	9.0 (57.9)
$CC_{1/2}$[(a)]	1.00 (0.76)	1.00 (0.66)
Refinement		
Resolution range (Å) [(a)]	32.4– 2.9 (3.0–2.9)	
R_{work} (%)[(a), (c)]	24.5 (35.9)	
R_{free} (%)[(a), (d)]	27.6 (41.3)	
Rmsd bond length (Å)[(e)]	0.007	
Rmsd bond angle (°)[(e)]	1.53	
Clashcore, all atoms	12.0 (96th percentile)	
Average *B*-factor (Å^2)	88.7	
PDB ID	5ZT7	

[a]The values in parentheses are for the highest resolution shell. [b]$R_{sym} = \sum\sum|I - \langle I \rangle|/\sum I$, where I is the intensity of each reflection. [c]$R_{work} = \sum||F_o| - k|F_c||/\sum|F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively. [d]R_{free} was calculated as the R_{work} for 5% of the reflections that were not included in the refinement. [e]rmsd, root mean square deviation.

	Co^{2+}-bound SirB (Co inflection)	Co^{2+}-bound SirB (Co remote)
Data collection		
Temperature (K)	100	100
Wavelength (Å)	1.606	1.566
Resolution range[(a)]	41.8– 3.1 (3.2–3.1)	41.8– 3.1 (3.2–3.1)
Space group	$P2_1$	$P2_1$
Cell dimensions		
a, *b*, *c* (Å)	65.0 52.2 75.6	65.0 52.2 75.6
α, β, γ (°)	90 112.1 90	90 112.1 90
Total reflections[(a)]	42871 (390)	42987 (433)
Unique reflections[(a)]	7555 (257)	7613 (283)
Redundancy[(a)]	5.7 (1.5)	5.6 (1.5)
Completeness (%)[(a)]	87.0 (30.5)	87.4 (33.4)
I/σ_I[(a)]	13.7 (1.1)	13.8 (1.1)
R_{sym} (%)[(a), (b)]	9.0 (57.9)	9.2 (62.9)
$CC_{1/2}$[(a)]	1.00 (0.66)	1.00 (0.58)
Refinement		
Resolution range (Å) [(a)]		
R_{work} (%)[(a), (c)]		
R_{free} (%)[(a), (d)]		
Rmsd bond length (Å)[(e)]		
Rmsd bond angle (°)[(e)]		
Clashcore, all atoms		
Average *B*-factor (Å^2)		
PDB ID		

[a]The values in parentheses are for the highest resolution shell. [b]$R_{sym} = \sum\sum|I - \langle I \rangle|/\sum I$, where I is the intensity of each reflection. [c]$R_{work} = \sum||F_o| - k|F_c||/\sum|F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively. [d]R_{free} was calculated as the R_{work} for 5% of the reflections that were not included in the refinement. [e]rmsd, root mean square deviation.

References

1 A. A. Brindley, E. Raux, H. K. Leech, H. L. Schubert and M. J. Warren, *J. Biol. Chem.*, 2003, **278**, 22388–22395.

2 C. V. Romão, D. Ladakis, S. A. Lobo, M. A. Carrondo, A. A. Brindley, E. Deery, P. M. Matias, R. W. Pickersgill, L. M. Saraiva and M. J. Warren, *Proc. Natl. Acad. Sci. USA*, 2011, **108**, 97–102.

3 K. Zheng, P. D. Ngo, V. L. Owens, X. P. Yang and S. O. Mansoorabadi, *Science*, 2016, **354**, 339–342.

4 S. J. Moore, S. T. Sowa, C. Schuchardt, E. Deery, A. D. Lawrence, J. V. Ramos, S. Billig, C. Birkemeyer, P. T. Chivers, M. J. Howard, S. E. Rigby, G. Layer and M. J. Warren, *Nature*, 2017, **543**, 78–82.

5 H. K. Leech, E. Raux-Deery, P. Heathcote and M. J. Warren, *Biochem. Soc. Trans.*, 2002, **30**, 610–613.

6 E. Raux, H. K. Leech, R. Beck, H. L. Schubert, P. J. Santander, C. A. Roessner, A. I. Scott, J. H. Martens, D. Jahn, C. Thermes, A. Rambach and M. J. Warren, *Biochem. J.*, 2003, **370**, 505–516.

7 H. K. Leech, E. Raux, K. J. McLean, A. W. Munro, N. J. Robinson, G. P. Borrelly, M. Malten, D. Jahn, S. E. Rigby, P. Heathcote and M. J. Warren, *J. Biol. Chem.*, 2003, **278**, 41900–41907.

8 M. Hansson and L. Hederstedt, *J. Bacteriol.*, 1992, **174**, 8081–8093.

9 S. Al-Karadaghi, M. Hansson, S. Nikonov, B. Jonsson and L. Hederstedt, *Structure)*, 1997, **5**, 1501–1510.

10 M. D. Hansson, T. Karlberg, M. A. Rahardja, S. Al-Karadaghi and M. Hansson, *Biochemistry*, 2007, **46**, 87–94.

11 D. Lecerof, M. N. Fodje, R. Alvarez Leon, U. Olsson, A. Hansson, E. Sigfridsson, U. Ryde, M. Hansson and S. Al-Karadaghi, *J. Biol. Inorg. Chem.*, 2003, **8**, 452–458.

12 S. A. Lobo, A. A. Brindley, C. V. Romão, H. K. Leech, M. J. Warren and L. M. Saraiva, *Biochemistry*, 2008, **47**, 5851–5857.

13 H. L. Schubert, E. Raux, K. S. Wilson and M. J. Warren, *Biochemistry*, 1999, **38**, 10660–10669.

14 D. Lecerof, M. Fodje, A. Hansson, M. Hansson and S. Al-Karadaghi, *J. Mol. Biol.*, 2000, **297**, 221–232.

15 F. Sievers and D. G. Higgins, *Cur. Protoc. Bioinformatics*, 2014, **48**, 3.13.11–16.

16 X. Robert and P. Gouet, *Nucle. Acids Res.*, 2014, **42**, W320–324.