






An evolutionary path to altered cofactor specificity in a metalloenzyme

Anna Barwinska-Sendra ¹, Yuritz M. Garcia², Kacper M. Sendra ¹, Arnaud Baslé¹, Eilidh S. Mackenzie ¹, Emma Tarrant¹, Patrick Card¹, Leandro C. Tabares³, Cédric Bicep¹, Sun Un³, Thomas E. Kehl-Fie ^{2,4,5}✉ & Kevin J. Waldron ^{1,5}✉

Almost half of all enzymes utilize a metal cofactor. However, the features that dictate the metal utilized by metalloenzymes are poorly understood, limiting our ability to manipulate these enzymes for industrial and health-associated applications. The ubiquitous iron/manganese superoxide dismutase (SOD) family exemplifies this deficit, as the specific metal used by any family member cannot be predicted. Biochemical, structural and paramagnetic analysis of two evolutionarily related SODs with different metal specificity produced by the pathogenic bacterium *Staphylococcus aureus* identifies two positions that control metal specificity. These residues make no direct contacts with the metal-coordinating ligands but control the metal's redox properties, demonstrating that subtle architectural changes can dramatically alter metal utilization. Introducing these mutations into *S. aureus* alters the ability of the bacterium to resist superoxide stress when metal starved by the host, revealing that small changes in metal-dependent activity can drive the evolution of metalloenzymes with new cofactor specificity.

¹Institute for Cell and Molecular Biosciences, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne NE2 4HH, UK. ²Department of Microbiology, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA. ³Department of Biochemistry, Biophysics and Structural Biology, Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif-sur-Yvette, France. ⁴Carl R Woese Institute for Genomic Biology, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA. ⁵These authors contributed equally: Thomas E. Kehl-Fie, Kevin J. Waldron. ✉email: kehlfie@illinois.edu; kevin.waldron@ncl.ac.uk

Metalloproteins are critical to all aspects of life. They are ubiquitous, with as many as half of all enzymes requiring a metal cofactor for function^{1,2}. Metalloenzymes are often highly specific for their cognate metal ion, exhibiting reduced activity with non-native metal cofactors in vitro and in vivo^{3–5}. As a result, understanding the principles that govern the specificity of metal-protein interactions is relevant to nearly all aspects of biology, medicine and biotechnology. Yet the underlying biochemical principles that determine metal cofactor specificity, and the evolutionary processes that drive the optimization of enzyme active sites to utilize a specific metal cofactor remain unclear. This in turn limits our ability to predict the metal used by proteins from their sequence, to produce synthetic metalloenzymes to perform novel reactions for industrial applications, and to rationally design metalloenzyme inhibitors for medicine.

Our limited appreciation of how metalloprotein architecture dictates metal specificity is exemplified by the superoxide dismutases (SODs)⁶. SODs detoxify superoxide, a reactive oxygen radical produced by aerobic metabolism and by the mammalian immune system in response to infection⁷. The biological importance of SODs is illustrated by their presence in most organisms, including obligate anaerobes, and the independent evolution of three distinct dismutase protein families. These SOD families are defined by the cofactor they utilize as nickel SODs, copper and zinc SODs, and manganese (Mn) or iron (Fe) SODs⁶. Mn/Fe-dependent SODs are the most common, and historically were thought to be strictly either manganese- or iron-specific. We recently established that a Mn/Fe SOD from the bacterium *Staphylococcus aureus* exhibits equal activity with either manganese or iron. These versatile enzymes are termed cambialistic SODs (camSOD)⁸. In addition to the camSOD (SodM), *S. aureus* also possesses a second, manganese-dependent SOD (SodA)^{8–10}. Although cambialistic SODs had previously been described^{6,11–15}, their biological importance was questioned. However, the *S. aureus* camSOD contributes to infection by enabling the bacterium to maintain a defense against superoxide when manganese starved by the host^{8,16,17}. All members of the Mn/Fe SOD family are related in sequence, exhibit identical protein folds, and

coordinate their metal ion using identical ligands⁶, making it unclear why some enzymes absolutely require manganese for catalysis (MnSOD), while others require iron (FeSOD), and still others show metal cofactor flexibility (camSOD).

The metal used by a protein is not permanently fixed, and can change in response to environmental pressures^{2,18}. For example, iron was readily soluble in the anaerobic oceans during life's early evolution and early organisms are thus thought to have been iron-philic^{18,19}. However, oxygenation by early photosynthetic organisms reduced the availability of iron^{18,19}. The resulting biological iron deficiency would have imposed selective pressure to adapt iron-dependent enzymes to use non-iron cofactors^{18–20}. While supported by bioinformatic analyses^{18,21}, no experimental evidence has been presented demonstrating the evolutionary process by which a change in metal specificity has evolved through iterative mutation²⁰.

Here, we exploit the close relationship between the staphylococcal SODs to understand how evolutionary changes in metal utilization occur. Genomic analysis shows the camSOD likely evolved from a manganese-specific predecessor that subsequently underwent neofunctionalization, a defined evolutionary process in which mutations rapidly accumulated in the duplicated gene during a period of functional redundancy, resulting in gain of a new beneficial function²². Integrated structural, biochemical, and electron paramagnetic resonance (EPR) studies reveal that two such mutations have altered amino acid residues in close spatial proximity to the SOD active site, driving the change in camSOD metal specificity. When these residues are reciprocally swapped, the metal specificities of the MnSOD and camSOD are largely interconverted. Remarkably, these residues possess non-polar sidechains located in the metal's secondary coordination sphere, and make no direct contacts to the metal-coordinating ligands. These subtle changes regulate the electronic structure and redox properties of the catalytic metal ion, dictating which metals the enzymes can use. Leveraging these findings reveals that small increases in iron-dependent catalysis by camSOD enhance the ability of *S. aureus* to overcome the immune response. Collectively, our data show how subtle changes to metalloenzyme architecture can dramatically alter the metal ion's reactivity and drive the evolution of isozymes with new cofactor specificity.

Table 1 Enzymatic activity of all SOD forms from this study.

Enzyme	Mutation	Mn activity ^a	Fe activity ^a	Cambialism ratio ^b
MnSOD	Wild type	1836 ± 79	4 ± 1	0.002
	F19I	1734 ± 61	5 ± 1	0.003
	G159L	942 ± 69	39 ± 4	0.041
	L160F ^c	1167 ± 92	19 ± 13	0.016
	G159L-L160F	173 ± 21	84 ± 12	0.486
	F19I-G159L-L160F	526 ± 49	57 ± 7	0.108
	camSOD	Wild type	251 ± 5	250 ± 10
I19F		216 ± 31	325 ± 1	1.505
L159G		808 ± 17	24 ± 3	0.030
F160L ^c		342 ± 5	111 ± 20	0.325
L159G-F160L		687 ± 33	11 ± 1	0.016
I19F-L159G-F160L		789 ± 54	39 ± 2	0.049

Enzymatic activities of all variants of *S. aureus* MnSOD and camSOD, in each of their metal-loaded forms, were assayed using a commercial SOD Activity Assay kit (Sigma). Each enzyme was assayed in triplicate using independent biological replicates, and error values quoted represent the standard deviation (SD) from the mean value given. All samples were assayed $n = 3$ except for three forms of the triple mutant variants that were assayed $n = 4$. Source data are provided as a source data file.

^aValues represent Units mg^{-1} protein \pm SD.

^bThe cambialism ratio (CR) is calculated by dividing the iron-dependent activity by the manganese-dependent activity, with a value of 1 representing a perfectly cambialistic enzyme. Enzymes that are cambialistic are defined herein as having a CR in the range 0.1–10.00.

^cWhereas all other samples were assayed simultaneously to circumvent noticeable day-to-day variability in the absolute activity detected using the commercial assay, the single-mutant variants with mutations at position 160 were assayed separately. To ensure comparability between the datasets, these variants were assayed simultaneously alongside samples of all wild-type forms, and small discrepancies in the absolute values were normalized by scaling using those wild-type activities as controls.

Results

The two *S. aureus* SODs exhibit extensive similarity. Initially, we comprehensively characterized which metals the MnSOD and camSOD could use. The MnSOD was active only with manganese, while the camSOD was active with manganese or iron (Supplementary Fig. 1 and Table 1). As with other characterized cambialistic SODs^{12,14,23}, the activity of *S. aureus* camSOD is significantly lower than that of the MnSOD, suggesting that cambialism represents an evolutionary compromise, where catalytic activity is sacrificed at the expense of enabling cofactor flexibility. Alignment of the *S. aureus* SOD amino acid sequences demonstrated that the two proteins share 75% identity (Supplementary Fig. 2). Next, we examined the *S. aureus* SODs for structural differences using circular dichroism (CD) spectroscopy. The iron- and manganese-loaded forms of both proteins showed very similar CD spectra (Fig. 1a), consistent with both SODs containing comparable secondary structure composition in solution (Supplementary Table 1). We then determined three dimensional crystal structures of the iron-loaded forms of each SOD by X-ray diffraction and compared them with structures²⁴ of the manganese-loaded counterparts (Supplementary Table 2). All four forms adopted near-identical architectures (Fig. 1b and Supplementary Fig. 3) whose polypeptide backbones could be overlaid with only minor deviations (Supplementary Table 3).

Crucially, the metal-binding ligands adopted near-identical spatial positions in all four structures (Fig. 1c), with insignificant changes to metal-ligand bond lengths or angles within the crystallographic resolution (Supplementary Tables 4 and 5). This is consistent with our previous EPR study that showed the positions of the protons on the metal-coordinated solvent and histidine ligands are identical²⁴.

Identification of residues proximal to the active site. The structural similarity of the two SODs suggested that subtle differences must explain their different metal specificity. By overlaying the crystal structures of MnSOD and camSOD, all 50 amino acids that differed between the two proteins were spatially

localized (Fig. 2a). Few of these were at the dimerization interface (Supplementary Fig. 4), consistent with this region being highly conserved in SODs (Fig. 2b), and most were surface localized (Fig. 2c). However, we identified three residues (19, 159, and 160) internal to the protein that varied between MnSOD and camSOD that were spatially close (<10 Å) to the metal (Fig. 2d), where non-polar sidechains (Leu or Ile for Phe, or Gly for Leu) were swapped. Notably, these sidechains in the secondary coordination sphere made no direct contacts with the metal-coordinating sidechains in the primary coordination sphere.

Metal-specificity is determined by sequence positions 159/160. To study their role in metal specificity, we prepared variant forms

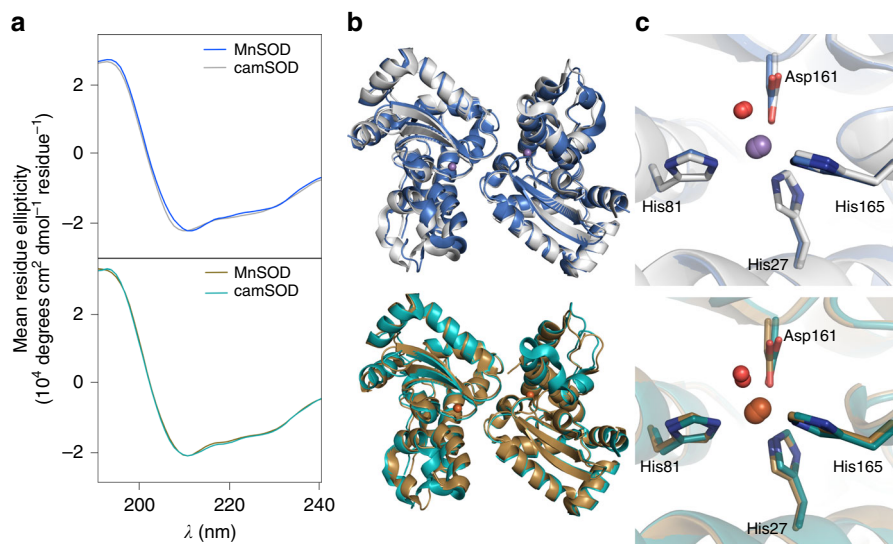


Fig. 1 Extensive structural similarity of *S. aureus* MnSOD and camSOD. **a** Overlaid CD spectra ($n = 1$) of the (upper panel) manganese-loaded (MnSOD in blue; camSOD in gray) and (lower panel) iron-loaded (MnSOD in gold; camSOD in teal) isoforms, demonstrating similar secondary structure content of the two isozymes in solution. **b** Superimposed crystal structure cartoons of (upper panel) manganese-loaded forms²⁴ and (lower panel) iron-loaded forms of each enzyme, with the proteins colored as per panel (a), and with manganese ions, iron ions, and waters shown as purple, orange, and red spheres, respectively. Analysis of all four structures show that the polypeptide backbones of all isoforms of the staphylococcal SODs are essentially identical to within the ~ 2 Å resolution of the structural data (Supplementary Tables 3–5). **c** No significant differences within the structural resolution were detectable in the metal coordination environment that could explain the disparate metal specificity of their catalysis.

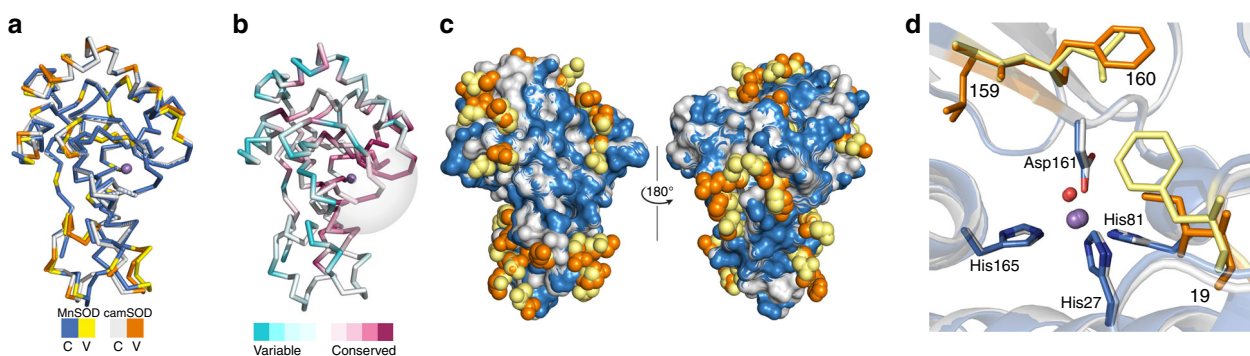


Fig. 2 Mutation of two residues located close to the active site inverts specificity. **a** Superimposed ribbon representation structures of the polypeptide backbones showing conserved (C) regions of the manganese-loaded forms of MnSOD (blue) and camSOD (gray), with all variable (V) residues highlighted (yellow and orange, respectively). **b** Analogous ribbon representation of the structure of *S. aureus* MnSOD, illustrating the localization of variable (cyan) and conserved (purple) regions of SOD sequence based on ConSurf analysis⁵⁹ of 500 SOD sequences (calculated conservation scores represented by the color spectrum key). Approximate dimerization interface is indicated (gray). **c** Superimposed space-filling models of MnSOD and camSOD, colored as in panel (a), illustrating that >85% of sidechains that vary between the *S. aureus* SODs are surface-exposed (2.5 \AA^2 surface exposure cutoff used). **d** Magnified view of the superimposed cartoon representations of MnSOD and camSOD structures, colored as in panel (a), showing the spatial location of the three residues targeted for mutagenesis in proximity to the manganese ion (purple).

of MnSOD and camSOD in which the amino acids at the three candidate positions were reciprocally swapped. The recognition that SOD metal specificity *in vitro* and *in vivo* is better described as a spectrum than as binary possibilities highlights the need for a quantitative measure of cofactor plasticity to be used alongside absolute activity. Herein, we define the term cambialism ratio (CR) as such a measure, being the ratio of iron-dependent activity to manganese-dependent activity detectable using a spectrophotometric assay⁸; this value is close to zero for MnSODs (CR = 0.002 for *S. aureus* MnSOD), close to unity for camSODs (CR = 0.996 for *S. aureus* camSOD), and has large values for FeSODs (CR > 800 for *E. coli* FeSOD²⁵).

Reciprocal mutations introduced at position 19 had negligible effects on the metal specificity of either enzyme (Table 1). Conversely, exchanging the residues at position 159 significantly impacted metal specificity. In the MnSOD, conversion of Gly159 to Leu (MnSOD Gly159Leu) increased its Fe-dependent activity ~10-fold relative to wild-type MnSOD while decreasing its manganese-dependent activity ~2-fold (Table 1). While remaining a highly manganese-specific enzyme, this variant exhibited slightly increased cambialism (CR = 0.048). The reciprocal mutation in camSOD, converting Leu159 to Gly (camSOD Leu159Gly), diminished its iron-dependent activity >10-fold relative to wild-type camSOD while increasing its manganese-dependent activity >3-fold, resulting in an enzyme variant with a MnSOD-like activity profile (CR = 0.030) (Table 1). These results indicate that the residue at position 159 has a major effect on the ability of the *S. aureus* SODs to use iron and manganese, but the contribution of this residue alone was insufficient to fully explain the disparate metal specificities of the two *S. aureus* enzymes.

Interconversion of the residue at position 160 alone also influenced metal specificity (Table 1). The MnSOD variant (Leu160Phe) was similar to the MnSOD Gly159Leu variant, exhibiting diminished manganese-dependent activity and increased iron-dependent activity (CR = 0.016) but to a lesser extent than in Gly159Leu, whereas the camSOD variant (Phe160Leu) showed slightly modified activity with both metals but remained largely cambialistic (CR = 0.325). Crucially, combining this mutation with that at position 159 further inverted the enzymes' metal specificity. The MnSOD double mutant (Gly159Leu-Leu160Phe) showed a substantial difference in metal utilization from that of the wild type or the single-mutant variants. The iron-dependent activity of the MnSOD double mutant was increased >20-fold relative to wild-type MnSOD, resulting in a highly cambialistic enzyme that displayed activity with manganese and with iron that differed only ~2-fold (CR = 0.486). It's notable that, like in wild-type camSOD, this highly cambialistic variant exhibited diminished activity with manganese, illustrating that gaining cofactor flexibility requires a compromise with catalytic efficiency. It further demonstrates the importance of comparing both the absolute activity and the cambialism ratio when assessing metal specificity. Conversely, the camSOD double mutant (Leu159Gly-Phe160Leu) had greater activity with manganese and reduced activity with iron relative to the wild type or the single mutant (camSOD Leu159Gly). This shift in camSOD metal preference resulted in a highly manganese-specific enzyme (CR = 0.016) (Table 1). Cumulatively, these results demonstrate that just two mutations in the secondary coordination sphere largely interconvert the metal specificity of camSOD and MnSOD.

Next, triple mutant variants, in which residues 19, 159 and 160 were all swapped, were evaluated. Unexpectedly, the MnSOD triple mutant (Phe19Ile-Gly159Leu-Leu160Phe) exhibited greater activity with manganese and less activity with iron (CR = 0.108) than the double mutant (Table 1). The camSOD triple mutant (Ile19Phe-Leu159Gly-Phe160Leu) had increased activity

with both metals relative to the double mutant (camSOD Leu159Gly-Phe160Leu). These studies demonstrate that positions 159 and 160 are important in regulating the reactivity of the metal cofactor. Strikingly, mutations at these positions produced profound effects on metal specificity, yet did so with essentially no changes to the protein backbone structure, at least within the resolution of the crystal structures (Supplementary Fig. 5). The physical structures of the metals' primary ligand spheres were the same, as were the substrate access channels and hydrogen-bonding networks that are important for proton-coupled electron transfer⁶.

Metal utilization correlates with the metal's electronic properties. Despite the similarity of the crystal structures, the mutations did have direct effects on the electronic structure and electrochemical properties of manganese centers. High-field electron paramagnetic resonance (HFEP) spectra of the Mn(II)-loaded *S. aureus* SOD wild-type forms²⁴ and the six variants were distinct (Fig. 3a). These spectra are determined by the *D* and *E* values of the zero-field interactions (ZFI) of the Mn(II) centers (Supplementary Fig. 6, Supplementary Eqs. 1–6, and Supplementary Note 1). Since the ZFI arises from the spin-orbit and magnetic interactions of the five unpaired Mn(II) electrons, this unambiguously demonstrated that the electronic structures of the metal centers were different. The sum, $|D|+E$, which measures the size of the Mn(II) ZFI, has been shown to correlate semi-quantitatively with manganese-dependent activity, with cambialistic SODs exhibiting intermediate values between the low values of Mn(II)-loaded FeSODs and the high values of MnSODs²⁶. As reported previously, the wild-type *S. aureus* proteins were consistent with this trend²⁴, but the current measurements shed greater light. The values of $|D|+E$ and the manganese-dependent enzymatic activities of seven of the enzyme forms (Supplementary Table 6) exhibited a significant linear correlation (Supplementary Fig. 7). The furthest departure from this correlation was camSOD Leu159Gly-Phe160Leu. This exception notwithstanding, the significance of this correlation can be interpreted from the results of studies on Mn(II/III)- and Fe(II/III)-containing, 4'-substituted 2,2':6',2''-terpyridine complexes. These compounds are structurally homologous and exhibit metal-specific or cambialistic SOD-like activity depending on the electron-donating/withdrawing strength of the 4' substituent, quantified by its Hammett σ values²⁷. The reduction potentials of both the manganese and iron complexes exhibit a linear dependence on σ , as do the ZFI of the Mn(II) complexes²⁶. Hence, the relationship between metal-specific activity, redox tuning through indirect secondary coordination sphere electronic effects, and Mn(II) ZFI in the enzymes is entirely consistent with the effects of subtle charge polarization in the simpler, better understood chemical systems.

Further evidence that the secondary coordination sphere residues influence the redox properties of the manganese cofactor comes from observations of their auto-oxidation properties. Mn(III) weakly absorbs in the 400–600 nm region^{25,28,29}. The spectra of the two wild-type SODs in their resting state (equilibrated with ambient oxygen) had similar shapes, but the MnSOD peak intensity was greater (Fig. 3b). This demonstrated the MnSOD was more susceptible to auto-oxidation under these conditions than the camSOD, whereas the converse was true for the double mutants (Fig. 3b). These trends in auto-oxidation correlated with manganese-dependent enzyme activity (Table 1). One half reaction catalyzed by SODs is the oxidation of superoxide to molecular oxygen, and auto-oxidation is formally the reverse of this reaction (Supplementary Eq. 7). If auto-oxidation proceeds by the reverse mechanism (Supplementary Eqs. 8–12), then any changes in metal reduction potential will affect this reaction in

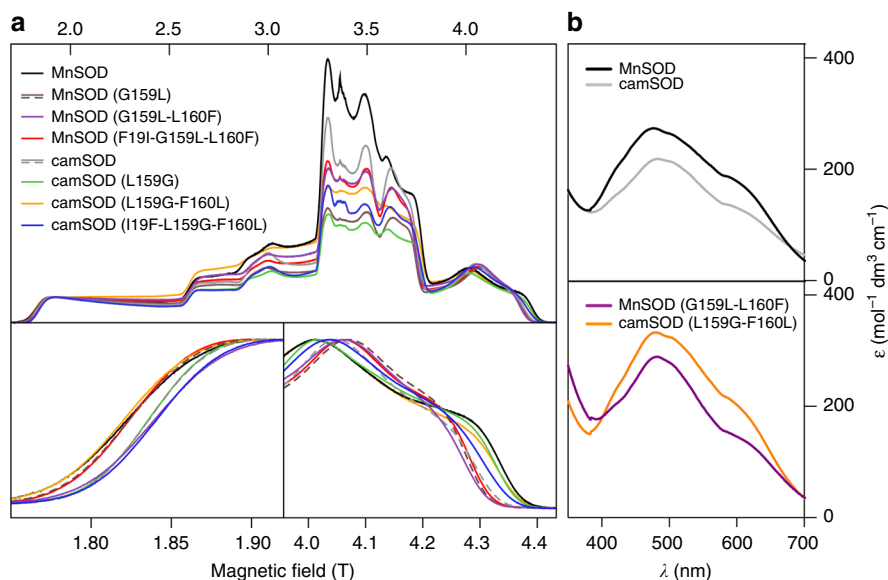


Fig. 3 Differences in electronic structure and auto-oxidation of the SODs. **a** The 94 GHz Mn(II) field-swept echo HFEPR spectra (upper panel) of each of the dithionite-reduced, manganese-loaded forms of the *S. aureus* wild type and variant SODs (labeled in the color key, top left), acquired at 5–6 K ($n = 1$). The lower panels show expanded views of the (left) lower and (right) upper field shoulders, respectively, which are indicated by arrows in the upper panel. SOD mutations gave rise to detectable differences in the enzymes' HFEPR spectra and their ZFI (Supplementary Figs. 6–7 and Supplementary Table 6), demonstrating altered electronic structure and redox properties caused by the mutations. **b** UV/visible absorption spectra of the manganese-loaded forms of the (upper) wild type and (lower) double mutant variants of MnSOD (black and purple, respectively) and camSOD (gray and orange, respectively). MnSOD exhibits greater auto-oxidation at rest than camSOD, illustrated by increased absorbance of Mn(III) at ~480 nm, a trend which is reversed in the double mutant variants that also exhibited reciprocal changes in metal specificity (Table 1). All samples (800 μM) were equilibrated with ambient aerobic conditions in 100 mM phosphate buffer, pH 7.5, 100 mM KCl, 1 mM EDTA before spectra were acquired. Spectra were collected on independent biological replicates ($n = 2$), mathematically converted to extinction coefficient (ϵ) using the Beer-Lambert law, and a representative spectrum is shown for each variant.

both directions (Supplementary Note 2). There are other factors that could differentially affect the two reactions, namely substrate accessibility and proton availability. However, within the limits of the resolution of the crystal structures, we found no evidence of changes in structure that were likely to affect superoxide/dioxygen motion to/from the metal ion or proton availability, with the positions of the residues around the active site remaining unchanged (Supplementary Fig. 5). While our structural analysis suggested these alternative possibilities were unlikely, we also measured the susceptibility of the oxidized form of the manganese-loaded SODs to reduction by absorption spectroscopy. Each enzyme was oxidized with permanganate, and then titrated with increasing concentrations of dithionite while monitoring the disappearance of the oxidized Mn(III) species using UV/visible detection (Supplementary Fig. 8). The camSOD was fully reduced at lower concentrations of dithionite than the MnSOD, consistent with the proposed difference in reduction potentials between the wild-type isozymes (Supplementary Eqs. 13–17). Importantly, this trend was also reversed in the double mutant variants, with MnSOD G159L-L160F becoming reduced at lower concentrations of dithionite than camSOD L159G-F160L. The SOD metal site is only accessible via a narrow solvent channel, making small redox-active molecules inefficient electrochemical mediators for SODs, thus precluding direct reduction potential measurements³⁰.

Altogether, the spectroscopic data indicate that the two wild-type *S. aureus* SODs and the two double mutants differ in their manganese redox properties. The HFEPR spectra demonstrated that the electronic structure of their Mn(II) centers were significantly different. The auto-oxidation data and dithionite reduction data both suggest that the two wild-type SODs have different reduction potentials. These differences are inverted in

the double mutants in both assays, which correlates with the inversion of manganese-dependent activity in these variants (Table 1). We conclude therefore that the altered metal specificity caused by the mutations is related to changes in reduction potential of the metal ions. The double mutation in MnSOD resulted in an altered electronic structure and re-tuning of the manganese reduction potential to more closely match that of camSOD and more catalytically active with iron, whereas the camSOD mutations altered reduction potential to make it more similar to MnSOD and more catalytically active with manganese.

***S. aureus* camSOD likely arose after a duplication of MnSOD.**

A challenge in studying metalloenzyme neofunctionalization is a lack of knowledge regarding the properties of the predecessor from which it evolved. To evaluate the evolutionary relationship between the *S. aureus* SODs, we performed a maximum likelihood phylogenetic analysis using 2691 aligned Mn/Fe SOD sequences from bacteria, archaea and eukaryotes (Supplementary Data 1). The overall topology of the resulting phylogenetic tree was consistent with that derived from a previous analysis of a more limited sequence library¹⁹ (Fig. 4a). Combining amino acid correlation analysis (Supplementary Data 2) with the phylogenetics (Supplementary Fig. 9) showed that the analyzed SOD sequences could be grouped into two distinct subtypes based on the presence of highly conserved GGH or AAQ motifs, as well as a third, diverse group lacking either motif (Fig. 4a, b and Supplementary Fig. 10). We annotated known SOD metal specificities onto the tree, including only examples where activity with both manganese and iron cofactors were definitively tested, which demonstrated that the GGH subgroup contains predominantly manganese-dependent SODs^{14,31} and the AAQ subgroup contains primarily iron-dependent SODs^{23,25}.

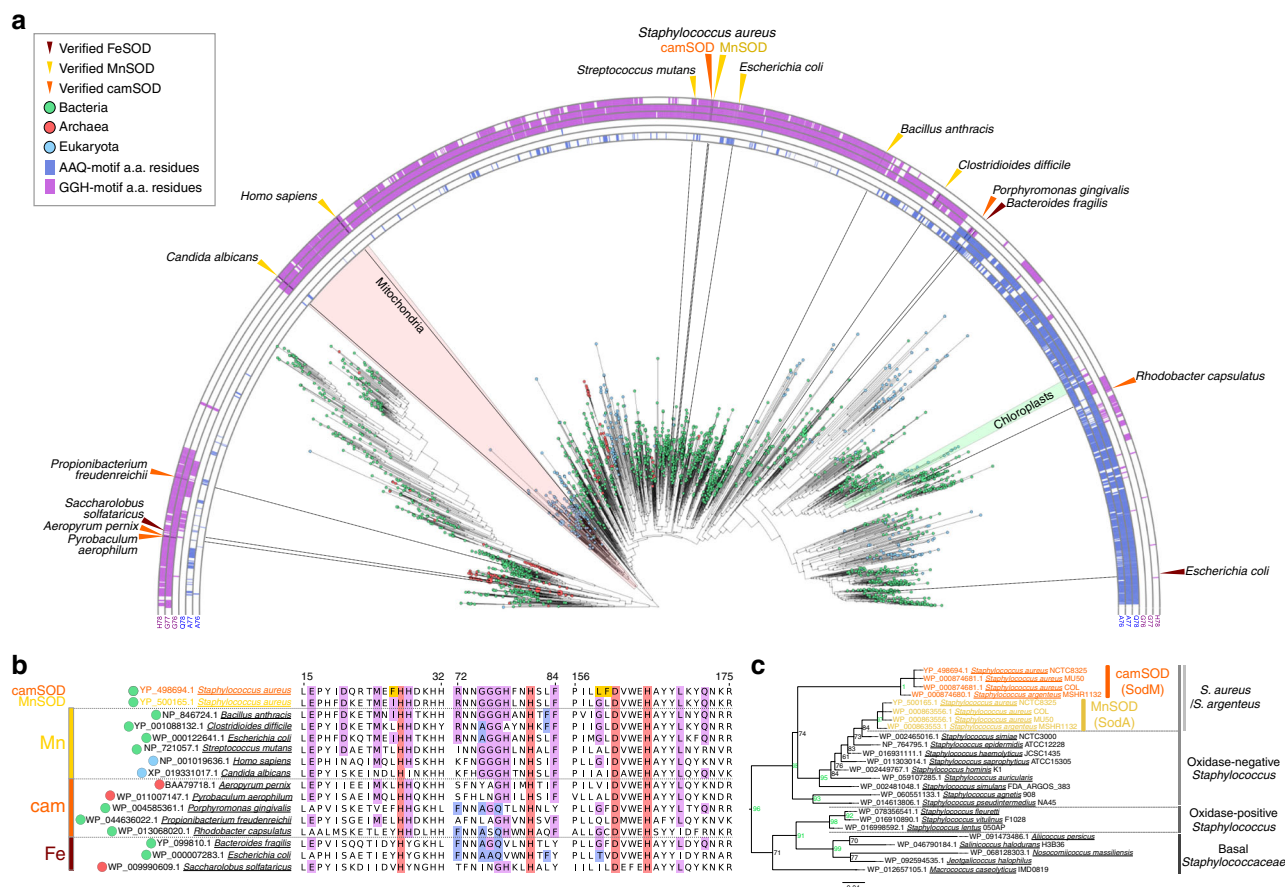


Fig. 4 *S. aureus* camSOD evolved from a manganese-specific predecessor. **a** Amino acid correlation (defined in the text as the AAQ group, blue, and the GGH group, magenta) and unrooted maximum likelihood phylogenetic tree based on alignment of amino acid (a.a.) sequences of 2691 Mn/Fe SOD homologs from bacteria (green circles), eukaryotes (sky blue circles), and archaea (red circles). Known metal-specificities of characterized enzymes (MnSOD = yellow, FeSOD = brown; camSOD = orange) are annotated with triangles. **b** Fragment of the alignment of selected SOD sequences of known metal specificity, with the metal-binding residues (red) and three residues in *S. aureus* camSOD (yellow) that were targeted for mutagenesis shown. Groups 1 and 2 of coevolving residues, identified using amino acid correlation analysis, were mapped onto the alignment, illustrating the GGH or AAQ motifs and associated correlated amino acids. The correlated residues, biological origin and metal specificity are indicated using the color scheme described in **a**, and sequence numbering is based on the *S. aureus* SODs. **c** Sub-tree extracted from the SOD tree in panel **a** containing all identified staphylococcal SOD sequences. Support values correspond to maximum likelihood bootstrap values from 100 rapid bootstrap replicates, with values >90 shown in green. The scale bar indicates number of substitutions per site. *Macroccoccus* and the oxidase-positive staphylococci form an out-group to a strongly supported grouping of oxidase-negative staphylococci, including *S. aureus*, consistent with published Staphylococcaceae phylogenies^{19,40,41}. Within the oxidase-negative staphylococci, all MnSOD (yellow) grouped together while camSOD homologs (orange) formed an out-group to the MnSODs.

Nonetheless, confirmed cambialistic SODs were present in both subgroups^{15,32} and a GGH-containing enzyme has been shown to be an FeSOD³³ (Fig. 4a, b), implying that the three metal specificities can and do interchange evolutionarily.

Inspection of the staphylococcal SOD protein similarity network and the tree (Supplementary Fig. 11) demonstrated that MnSOD is present in all staphylococcal genomes, whereas camSOD was identified only in the genomes of *S. aureus*, and of *S. argenteus* and *S. schweitzeri*, both recognized as early-branching lineages of the *S. aureus* tree^{34,35}. As MnSOD is the closest homolog to camSOD, the most parsimonious explanations are that the *sodM* gene that encodes camSOD arose from a duplication of *sodA*, encoding MnSOD, or that *sodM* was acquired by horizontal gene transfer from a close relative. Regardless, either model predicts that the newly acquired *sodM* would have originally encoded a manganese-specific enzyme. Increased branch length indicates an increased rate of evolutionary change for the ancestral SodM, while SodA remained under purifying selection²² (Fig. 4c). The camSOD sequence is highly conserved across all *S. aureus* (99% identity), and this

homology extends to the second SOD found in *S. argenteus* (97.5%) and *S. schweitzeri* (99%) (Supplementary Fig. 2). Notably Leu159 and Phe160 in camSOD are strictly conserved between *S. aureus*, *S. schweitzeri* and *S. argenteus*. Thus, we conclude that the acquired *sodM* gene originally encoded a MnSOD, which subsequently evolved into the modern camSOD during the emergence of *S. aureus*.

Acquisition of camSOD coincided with virulence factor expansion. To understand the wider context of camSOD emergence within the *S. aureus* lineage, we performed whole genome comparisons and a protein similarity network analysis (Supplementary Data 3) of genomes sampled across the staphylococcal phylogenetic tree (Supplementary Data 4). The *sodA* gene encoding MnSOD was located within a genomic region rich in essential genes^{36–38} (Supplementary Data 5) that are highly conserved across the staphylococci (Fig. 5a, b). Conversely, the *sodM* gene encoding camSOD is located within a highly variable genomic region close to the chromosomal origin that is enriched in virulence factors³⁹ (Supplementary Data 5), consistent with

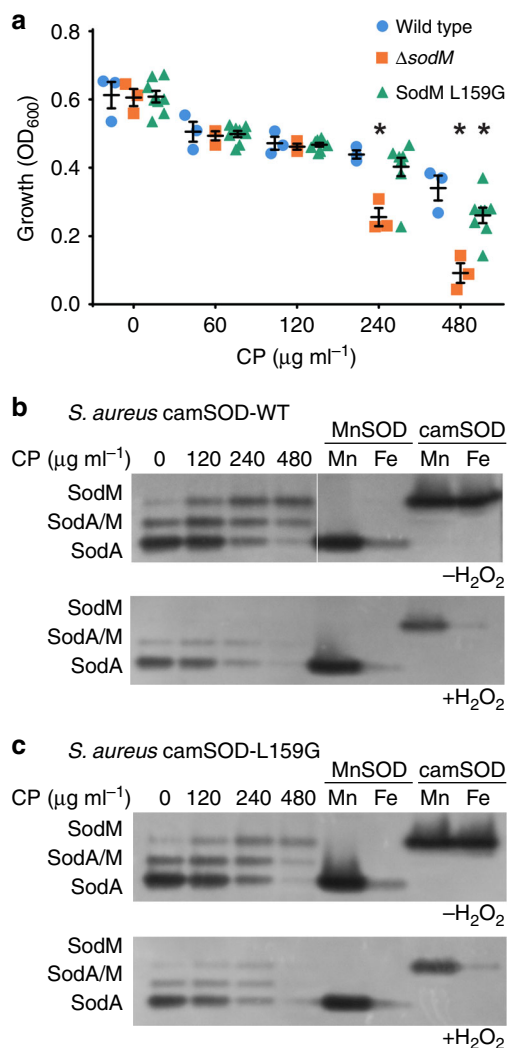


Fig. 6 Iron-dependent camSOD activity enables *S. aureus* to resist host stresses. **a** The Δ sodM mutant strain (lacking the gene encoding camSOD—orange squares) of *S. aureus* shows greatly diminished growth in the presence of paraquat (PQ) relative to wild type (WT—blue circles) under manganese-depleted conditions imposed by the presence of the human protein complex calprotectin (CP)^{8,16,17}. *S. aureus* cells expressing the Leu159Gly variant of camSOD from the native sodM locus (green triangles) also show reduced growth relative to the wild type. Note that the asterisk represents $p < 0.05$ via two-way ANOVA with Tukey's post-test performed in Graphpad Prism when compared to wild-type bacteria grown in the same concentration of calprotectin. Multiple independent clones of the Leu159Gly variant ($n = 8$) were assayed at the same time as those of the wild type ($n = 3$) and the Δ sodM mutant strain ($n = 3$). Each data point represents an independent biological replicate. Black bars represent the mean, with error bars showing standard error of the mean (SEM). The diminished ability of the camSOD variant to use iron (Table 1) renders *S. aureus* less capable of overcoming the host immune response. **b, c** In-gel SOD activity assays using extracts (5.4 μg) prepared from **b** wild type *S. aureus* and **c** the strain expressing the Leu159Gly camSOD variant, cultured under identical conditions to those used in **a**, demonstrate that both forms of the SodM enzyme are detectable at similar levels (upper panels). Treatment of the extracts (left) or recombinant enzymes (right) with H₂O₂ (lower panels), a specific inhibitor of the iron-loaded form⁸, demonstrated that the camSOD was loaded with iron in wild-type cells. The MnSOD enzyme is labeled as SodA, camSOD as SodM. This is a representative gel from a triplicate analysis of independent biological replicates ($n = 3$). Note that molecular weight markers were not used in these native gels, with SOD bands identified in the cell extracts by comparison of their mobility with the purified recombinant proteins. Source data for all panels are provided as a source data file.

ability of camSOD to perform iron-dependent catalysis enables *S. aureus* to resist oxidative stress during physiological manganese deprivation.

Collectively, our data are consistent with a model in which the evolution of camSOD was driven by interaction of the ancestor of *S. aureus* with the host immune system (Fig. 7). Acquisition of a redundant gene encoding a second manganese-specific SOD was followed by a period of rapid evolution of the protein sequence, which conferred the ability to utilize iron for catalysis. This period coincided with the expansion of the organism's repertoire of virulence-associated proteins, enabling *S. aureus* to survive immune assault and establish infection better than less pathogenic relatives⁴⁰. We thus propose that the selection pressure that drove the neofunctionalization of camSOD from a manganese-specific precursor after acquisition involved the combined immune-mediated stresses of the oxidative burst and manganese starvation^{7,8,16,17}.

Discussion

Metalloproteins are ubiquitous in biology, thus understanding the principles that govern metal cofactor utilization has implications for diverse aspects of biology, medicine, and biotechnology. However, the molecular factors that dictate the metal cofactor utilized by metalloenzymes, and how iterative changes to the protein sequence over time result in a change in metal specificity, remain poorly understood. Our inability to accurately predict the metal utilized by the Mn/Fe SODs, a widely distributed family of metalloenzymes, exemplifies these deficiencies. In this study, we developed the pair of *S. aureus* SODs as a powerful model system with which to interrogate how metal specificity can change over evolutionary time, enabling future studies to determine how the architecture of a metalloenzyme imposes cofactor specificity. Exploiting this platform demonstrated that evolution can effect substantial changes in metal specificity by making modest changes in the chemical nature of secondary coordination sphere residues that regulate metal reactivity.

Evolutionary distance between SODs with differing metal specificity has limited our ability to elucidate which residues dictate metal utilization^{6,12,26,29}. Leveraging the extensive structural and sequence homology between the staphylococcal SODs identified two key residues that dictate metal-specificity. Remarkably, the chemical differences between the sidechains of these residues are minor, with Gly (MnSOD) and Leu (camSOD) at position 159, and Leu (MnSOD) and Phe (camSOD) at position 160. While both residues are close to the active site, neither sidechain makes direct contact with the metal or its coordinating ligands, nor do they affect the overall architecture of the active site. Despite this, HFEPR and UV/visible spectroscopy revealed that interconversion of these residues profoundly alters the electronic structure and redox properties of the manganese relative to the wild-type enzymes²⁴. These data confirmed that MnSOD is more susceptible to auto-oxidation than camSOD, and crucially, this difference was inverted in the variants that also exhibited inverted metal specificities. These results support a model in which the SOD architecture selectively tunes the cofactor's reduction potential^{6,28,44}, and identify positions 159 and 160 as playing a critical role in this electronic modulation. Broadly, these results highlight the important contribution of secondary coordination sphere residues in determining metal specificity. Further, they illustrate that even minor secondary sphere changes can have profound impacts on metal-specificity, which is relevant to the study of natural metalloproteins, and the de novo design of synthetic metalloenzymes^{45–49}.

The residue equivalent to Gly159 in *S. aureus* MnSOD was reported as being generally Gly in both manganese-dependent

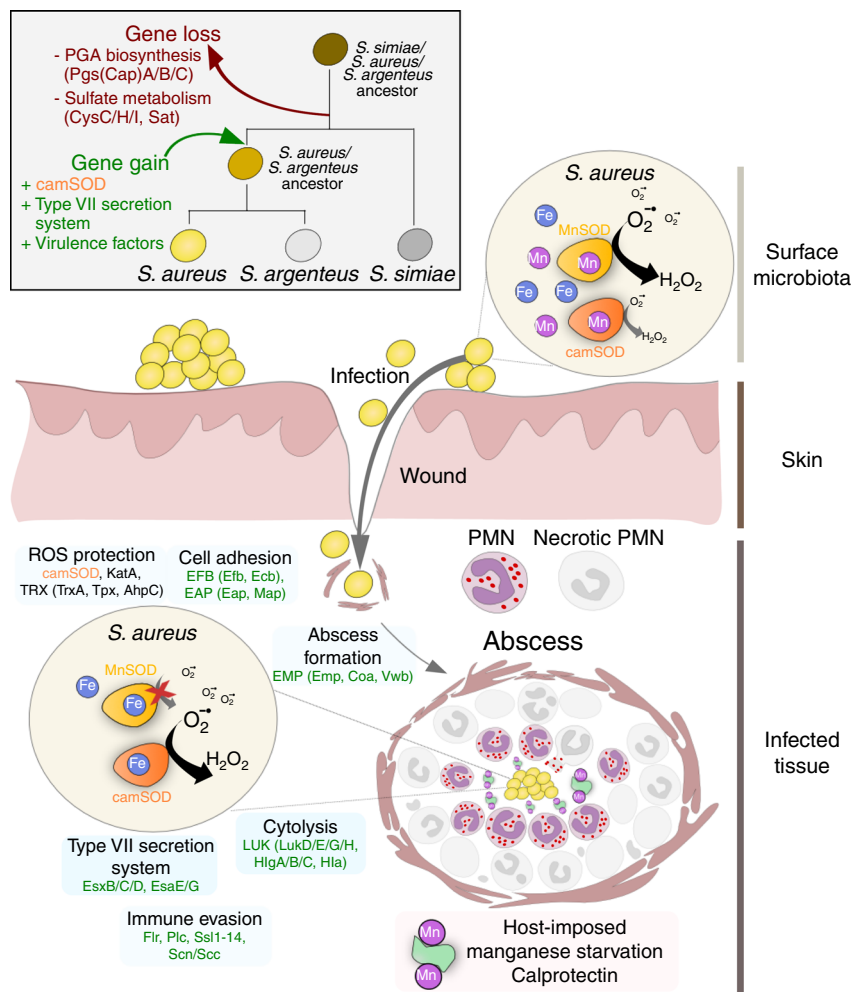


Fig. 7 Evolution of camSOD was driven by altered metal availability in the host. Schematic model depicting the evolution of *S. aureus* camSOD (orange) under selection pressure caused by exposure of the *S. aureus* ancestor to host-imposed manganese-starvation mediated by calprotectin. Acquisition and evolution of camSOD enabled the ancestor of *S. aureus* to maintain its antioxidant defense under these conditions using its iron-dependent catalysis, and coincided with its acquisition of numerous genes (green) involved in infection. Collectively, these acquired systems resulted in increased survival of *S. aureus* during interaction with host immune mechanisms, creating pathogenic *S. aureus*.

and cambialistic SODs, but Thr in iron-dependent SODs^{12,29,50}. Introduction of Thr in this position in manganese-specific or cambialistic SODs increased the cambialism ratio^{12,29}. This was previously explained by modification of a hydrogen-bonding network involving the mechanistically important metal-coordinated solvent²⁹. Our data confirm that a Gly residue in this position strongly favors manganese-specificity, but show that a small aliphatic sidechain at this position confers cambialism when combined with an aromatic sidechain on the neighboring residue at position 160. Importantly, our structural analyses of the staphylococcal SODs and their variants did not identify changes to the hydrogen-bonding network or substrate access channel, and ¹H electron nuclear double resonance (ENDOR) spectroscopy demonstrated the positions of the solvent protons are unchanged between MnSOD and camSOD²⁴, and likely also in the mutants. These observations cast doubt on the hydrogen-bonding network explanation in our system, and highlight the potential of the staphylococcal SODs as a model system with which to uncover the mechanisms by which redox tuning is controlled.

Our data suggest that camSOD evolved from a MnSOD predecessor in the last common ancestor of *S. aureus*/*S. argenteus*/*S. schweitzeri*. The acquisition and subsequent altered metal

utilization of camSOD coincided with the acquisition of numerous other factors that contribute to infection^{8,39,41,42,51}. We exploited our ability to produce manganese-specific camSOD variants to determine whether its iron-dependent catalytic activity provides a direct benefit to *S. aureus* during exposure to stresses encountered within the host. A *S. aureus* strain expressing the Leu159Gly variant camSOD showed reduced growth under manganese-restricted and oxidative stress-exposed conditions relative to wild type. As this variant possesses reduced iron-dependent activity but increased catalysis with manganese, this result clearly demonstrates that *S. aureus* growth under these conditions requires iron-dependent camSOD activity. We previously showed that camSOD is more important than MnSOD to the establishment of infection of wild-type mice⁸. Collectively, these data support a model in which the manganese-specific camSOD antecedent, acquired by the last common ancestor of the extant *S. aureus* clade⁴¹, underwent neofunctionalization through selection for increased iron-dependent activity, a pressure likely imposed by the manganese starvation it experiences during infection.

This study presents experimental evidence of how a change in metal specificity has naturally evolved in a metalloenzyme. We provide evidence that the specificity switch was selected for by

nutrient-starvation imposed by the immune system, implying that metal availability can shape enzyme metal utilization^{18,52}. Our data also demonstrate a critical role for secondary coordination sphere residues in mediating this specificity switch, with wider implications for the study of natural and synthetic metalloenzymes.

Methods

Bacterial strains, culture conditions, and plasmids. *Escherichia coli* strains DH5 α and BL21(DE3) were used for cloning and for recombinant protein expression, respectively, cultured at 37 °C in lysogeny broth (LB) or M9 medium supplemented with 200 μ M FeSO₄, with 1.5% (w/v) agar for solid medium. *S. aureus* strains were cultured at 37 °C in tryptic soy broth (TSB), on tryptic soy agar plates (TSA), or in Tris minimal medium⁵³. Where needed, tetracycline (10 μ g ml⁻¹), erythromycin (10 μ g ml⁻¹), kanamycin (50 μ g ml⁻¹), ampicillin (100 μ g ml⁻¹), or chloramphenicol (10 μ g ml⁻¹) were added to the media. All strains were stored at -80 °C in media mixed with 25% (v/v) glycerol. The construction of the *S. aureus* SH1000 Δ sodA (*sodA::Tn917-LTV1*), Δ sodM (*sodM::tet*), Δ sodA Δ sodM and Newman Δ sodA (*sodA::tet*) and Δ sodM (*sodM::erm*) mutant strains have been previously described^{8,10}. For pull-down analysis, the C-terminally StrepII-tagged variant of camSOD (including its Shine-Dalgarno sequence) was synthesized (IDT) and sub-cloned into the pS10t shuttle vector⁵⁴ using *Bam*HI/*Pst*I restriction cloning, and the resulting pS10t-Strep-sodM construct was introduced into *S. aureus* SH1000 using methods described previously⁵³. Firstly, electrocompetent *S. aureus* RN4220 cells were transformed with pS10t-StrepII-sodM. Positive transformants, selected on TSB-agar plates supplemented with chloramphenicol, were used to generate ϕ 11 phage lysate. Finally, phage transduction was used to introduce pS10t-StrepII-sodM into the *S. aureus* SH1000 background. The chromosomally integrated L159G mutation was introduced in *S. aureus* Newman via homologous recombination using the integration vector⁵⁵ pKOR1 and standard molecular approaches. This mutant was confirmed to be hemolytic by plating on blood agar prior to use.

Expression and purification of recombinant SODs. The cloning of the *S. aureus* *sodA* and *sodM*, their expression in *E. coli* BL21(DE3) cells, the purification of the recombinant enzymes, the in vitro swapping of bound iron for manganese by unfolding/refolding, and their quantitation, were all as previously described⁸. Expression constructs of *S. aureus* *sodA* and *sodM* were prepared by *Nde*I/*Bam*HI restriction cloning into pET29a vector. Protein was expressed in *E. coli* BL21 cells, in M9 minimal medium supplemented with 200 μ M ammonium iron sulfate or in standard LB broth, for 4 h post induction with 1 mM IPTG. Cell extracts were prepared by sonication of washed cell pellets in 20 mM Tris pH 7.5, 150 mM NaCl, 1x cComplete EDTA-free protease inhibitor (Roche), followed by centrifugation at 19,000 \times g. Cleared cell lysate was subjected to chromatographic separation using an AKTA purification system (GE Healthcare). Recombinant SOD proteins were purified using anion exchange chromatography (Hi Trap Q HP column, GE Healthcare) in 20 mM Tris pH 7.5 buffer with 0–1 M NaCl gradient elution, and subsequently by size exclusion in 20 mM Tris pH 7.5, 150 mM NaCl buffer on a Superdex 200 (GE Healthcare) column. Metal cofactor exchange was achieved by unfolding SOD proteins in buffer containing 2.5 M guanidine hydrochloride in the presence of 5 mM EDTA and 20 mM 8-hydroxyquinoline to remove bound metal ions, followed by refolding through dialysis against 20 mM Tris, 100 mM NaCl, and 1–10 mM metal of interest. The constructs for recombinant expression of variants of MnSOD and camSOD were mutagenized using the QuikChange site-directed mutagenesis approach (Stratagene), using the pET29a-sodA and pET29a-sodM vectors⁸ and the primer pairs listed in Supplementary Table 7.

CD spectroscopy. Circular dichroism data were collected using a JASCO J-810 spectropolarimeter equipped with a PTC-4235 Peltier temperature controller. Spectral measurements were performed within the far-UV range (185–260 nm) at a protein concentration of 0.8 μ M in 10 mM sodium phosphate buffer, pH 7.0 at 20 °C, in a 1 mm path length quartz cuvette (Hellma), using a scan rate of 50 nm min⁻¹, 2 s response time, 0.2 nm data pitch and a 2 nm bandwidth. For each sample, five accumulated scans were averaged, corrected for the buffer reference, and then converted into mean residue ellipticity (MRE, θ). Deconvolution was performed using the CDSSTR program with reference dataset 4 (250–190 nm) on the DichroWeb server^{56,57}. The thermal denaturation of protein was performed by following the change in ellipticity at a fixed wavelength of 222 nm as a function of temperature increasing from 20 to 90 °C. Protein melting temperature (T_m) values were calculated from the first derivative of the thermal denaturation curves, converted to mean residue ellipticity (MRE₂₂₂, θ_{222}).

UV/visible spectroscopy. UV/visible spectroscopy was performed on a Perkin-Elmer λ 35 spectrophotometer, collecting data at room temperature between 200 and 800 nm using 10 mm path length quartz cuvettes. Measurements of resting spectra were performed in 100 mM potassium phosphate buffer, pH 7.5, 100 mM KCl, 1 mM EDTA, with samples prepared at ~800 μ M protein. Spectra were converted to extinction coefficient, ϵ , using the protein concentration corrected for

manganese-loading, as detected by inductively coupled plasma mass spectrometry (ICP-MS). For reductive titrations, each sample was fully oxidized by incubation with 1 mole equivalent of potassium permanganate for 10 min, followed by extensive buffer exchange using centrifugal filtration devices (Amicon, 10 kDa cutoff) to remove the oxidant in the same buffer. Samples were titrated by addition of small volumes of fresh solutions of concentrated sodium dithionite, mixed vigorously by pipette and incubated for 20 min prior to acquisition of each spectrum. Spectra were corrected using the manganese content as detected by ICP-MS, and normalized to allow comparison between titrations.

Electron paramagnetic resonance (EPR) spectroscopy. The 94 GHz EPR spectra were obtained at 5–6 K using a Bruker EleXsys II 680 EPR spectrometer equipped with a “power upgrade 2” and an Oxford Instruments CF935 flow cryostat. The spectra were obtained by measuring the integrated amplitude of a standard two-pulse Hahn echo ($t_p(\pi) = 24$ ns and interpulse time of 400 ns) as a function of the magnetic-field. The zero-field D and E values were determined using a modified version of the method described in our previous work²⁴, because the shapes of the low-field edges of the 94 GHz spectra were different (Supplementary Note 1). To remove any arbitrariness in locating the magnetic-field of the low-field edge, the spectra were numerically differentiated and the lowest-field inflection points were taken as the edge positions. This had the effect of taking into account the distribution in the D values, which produced the different edge shapes, but also overestimated their values by 100 to 200 MHz in comparison to the previously reported values. The same approach was used for the high-field edges (E) as well.

ICP-MS elemental analysis. Metal analysis was mostly performed by ICP-MS using a Thermo x-series instrument and Thermo’s PlasmaLab software in Newcastle, operating in collision cell mode as previously described⁸. For Supplementary Fig. 9, ICP-MS analysis was performed by Durham University Bio-ICP-MS Facility. Briefly, protein samples were diluted in 2.5 % (w/v) solution of HNO₃ (Merck) containing 20 μ g L⁻¹ Co and Pt or In and Pt (Newcastle), or 10 μ g L⁻¹ Ga and Bi (Durham) as internal standards. Matrix-matched elemental standards (containing analyte metal concentrations of 0–500 μ g L⁻¹) were analyzed alongside each set of samples, each in triplicate, to enable metal concentrations to be determined from standard curves.

X-ray crystallography. Preparations of the iron-loaded forms of recombinant MnSOD and camSOD, and the manganese-loaded forms of both double mutant variants and of the triple mutant of camSOD (15 mg ml⁻¹), were each subjected to crystallization screening and optimization as previously described²⁴. Each protein preparation was confirmed to contain exclusively the target metal ion by ICP-MS prior to crystallization. Screening trays were set up using a Mosquito liquid handling robot (TTP Labtech) with commercially available matrix screens: PACT, JCSG+, Structure (Molecular Dimensions) and Index (Hampton Research) in 96-well MRC crystallization plates (Molecular Dimensions) using the sitting drop of vapor-diffusion method, incubated at 20 °C. Optimization of initial conditions for salt, polyethylene glycol (PEG) precipitant concentration and pH was performed using the hanging drop of vapor-diffusion method, incubated at 20 °C, in 24-well Linbro plates (Molecular Dimensions). Iron-loaded MnSOD was crystallized in 200 mM MgCl₂, 100 mM Tris, pH 8.5 and 30% (w/v) PEG 4000, whereas iron-loaded camSOD was crystallized in 100 mM PCTP, pH 7.0 and 25% (w/v) PEG 1500. The manganese-loaded forms of the double mutant variant of MnSOD (Gly159Leu-Leu160Phe) was crystallized in 200 mM MgCl₂, 100 mM Bis-Tris, pH 5.5 and 25% (w/v) PEG 3350, whereas the manganese-loaded form of the double mutant variant of camSOD (Leu159Gly-Phe160Leu) was crystallized in 30 mM MgCl₂, 30 mM CaCl₂, 39.1 mM Bicine, 60.9 mM Trizma, pH 8.5, 20% (w/v) PEG 550 MME and 10% (w/v) PEG 20,000. The manganese-loaded form of the triple mutant variant of camSOD (Ile19Phe-Leu159Gly-Phe160Leu) was crystallized in 100 mM potassium thiocyanate and 30% (w/v) PEG 2000 MME. All but the double mutant variant of camSOD were cryo-protected with 20% (w/v) PEG 400. X-ray diffraction data collection (Diamond Light Source, Didcot, UK) and processing, and model building and validation were all as previously described²⁴, except that data processing also used xia2 and phasing also used Phaser⁵⁸. The search model for iron-loaded wild-type MnSOD and manganese-loaded MnSOD Gly159Leu-Leu160Phe was PDB model 5N56. The search model for iron-loaded wild-type camSOD and the manganese-loaded forms of camSOD Leu159Gly-Phe160Leu and SodM Ile19Phe-Leu159Gly-Phe160Leu was PDB model 5N57. The PDB codes for the structures of the iron-loaded forms of MnSOD and camSOD and the manganese-loaded forms of the camSOD triple mutant variants are 6EX3, 6EX4, and 6EX5, respectively. The data collection and refinement statistics are summarized in Supplementary Table 3. All crystallographic images were generated using PyMOL Molecular Graphics System, Version 1.8 (Schrödinger, LLC). To illustrate structural regions where the sequence is conserved, ConSurf was used⁵⁹ to analyze 500 HMMER homologs of the SOD family, sampled from the UniREF90 database. Conservation scores were calculated using the Bayesian method.

Pull-down assay. The StrepII-tagged camSOD was constitutively expressed in a *S. aureus* SH1000 Δ sodM strain¹⁰ and purified using affinity chromatography. Cell

extracts were prepared by freeze-grinding under liquid nitrogen in 20 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA. Thawed, ground material was centrifuged (4000 × g, 20 min, 4 °C) to remove cell debris, and the resulting supernatant was filtered through a 0.45 µm membrane filter. The soluble protein extract was applied to a StrepTrap column (GE Healthcare), equilibrated and subsequently washed in 20 mM Tris pH 8.0, 150 mM NaCl, 1 mM EDTA. Elution (1 ml) used 2.5 mM desthiobiotin (PanReac AppliChem) in this buffer. Fractions were analyzed by SDS-PAGE and ICP-MS.

SOD activity assays. SOD activity was assessed qualitatively in-gel, or quantitatively using a SOD Assay Kit (Sigma-Aldrich) per the manufacturer's instructions, as previously described⁸. *S. aureus* cell extracts were prepared by mechanical disruption in 50 mM phosphate buffer at pH 7.8 with 0.1 mM EDTA. Protein content of lysates was quantified by bicinchoninic acid assay (Sigma). Aliquots of crude extracts (5.4 µg) or purified recombinant proteins (0.3 µg) were resolved on 10% (w/v) native polyacrylamide gels. For staining, gels were incubated in 50 mM potassium phosphate buffer pH 7.8, 1 mM EDTA, 0.25 mM nitroblue tetrazolium chloride (Sigma), 50 µM riboflavin and then exposed to light to detect SOD activity. Incubation of replica gels in 20 mM H₂O₂ prior to activity staining was performed to assess Fe-dependent SOD activity. Note that, as the native gels used for assays were stained with a SOD-specific stain, no molecular weight markers were resolved on these gels.

Calprotectin growth assays. Assays of *S. aureus* growth in the presence of calprotectin and paraquat were performed as previously described⁸. Overnight *S. aureus* cultures were set up in Chelex-treated RPMI medium containing 1% casamino acids (NRPMI), supplemented with 1 mM MgCl₂, 100 µM CaCl₂, and 1 µM FeCl₂. These were further sub-cultured by 1:100 dilution into 100 µl of fresh culture medium consisting of 38% (v/v) TSB and 62% (v/v) CP buffer (3 mM CaCl₂, 20 mM Tris base, 100 mM NaCl, 10 mM β-mercaptoethanol, pH 7.5) supplemented with 1 µM MnCl₂ and 1 µM ZnSO₄. Growth assays were performed in a 96-well round-bottom plate at 37 °C with shaking at 180 rpm for 8 h. Where indicated, 0.1 mM PQ was added to the media.

SOD phylogenetic analysis. SOD sequences were identified using BLASTP search⁶⁰ (*E*-value threshold = 0.005), with *S. aureus* MnSOD sequence as query, against a database of protein sequences downloaded from the NCBI server encoded in 2577 bacterial (including Candidate Phyla Radiation^{61,62}), 289 archaeal (including ASGAR Archaea^{63,64}), and 148 eukaryotic genomes. All 2691 identified SOD homologs were aligned using MAFFT⁶⁵ and trimmed with trimAl⁶⁶ (gappout mode) resulting in a trimmed alignment of 208 positions (Supplementary Data 1). The best fitting phylogenetic model (WAG + G) was selected with PROTTEST⁶⁷, and used to generate a SOD phylogeny in RAXML⁶⁸ with 100 bootstrap replicates. The tree was visualized and annotated with the coevolving amino acid residues (Fig. 4 and Supplementary Fig. 9) using GraPhlAn⁶⁹ (Supplementary Data 1 and 2).

Amino acid correlation analysis. Groups of coevolving amino acids were identified (Supplementary Data 2) in the alignment of the 2,691 SOD sequences sampled across the tree of life (Supplementary Data 1), using amino acid correlation algorithm implemented in PFFstats⁷⁰ following the user guide. PFFstat correlation analyses parameters were as follows: Minimum Score: 60 (18.75% of the maximum positive correlation score of 320); MRSA fraction: 0.2; minimum delta: 0.3. The amino acid correlation network was visualized in Cytoscape⁷¹ (Supplementary Fig. 9).

Protein network analyses. Protein similarity networks (Supplementary Data 3) were generated out of 71,757 protein sequences encoded in 29 analyzed Staphylococcal genomes (downloaded from the NCBI server) with BLAST⁶⁰, and quick edge file optimization implemented in EGN⁷² using the following thresholds: BLAST (*a* = 12, *e* = 1e⁻⁰⁵), simple link parameter (*E*-value threshold = 1e⁻⁰⁵, hit identity threshold = 20%, minimum hit coverage = 20%, without the best reciprocal hit condition enforcement), and with the quick edge file creation (*E*-value threshold = 1e⁻⁰⁵, hit identity threshold = 20%, minimum hit coverage = 20%, minimal hit length = 30 amino acids). Protein sequences were annotated⁷³ using eggNOG-mapper v1. The generated networks were analyzed using igraph⁷⁴ and dnet⁷⁵ packages in R⁷⁶, and visualized in Cytoscape⁷⁷. Analyzed species were classified into two categories: camSOD-positive (*S. argenteus* MSHR1132, *S. aureus* COL, *S. aureus* MU50, *S. aureus* Newman, *S. aureus* USA300LV CA12, *S. aureus* USA300 FPR3757, *S. aureus* NCTC8325); camSOD-negative (*S. capitis* AYP1020, *S. carnosus* SK13, *S. chromogenes* MU970, *S. cohnii* SNUDS2, *S. epidermidis* ATCC12228, *S. equorum* C2014, *S. haemolyticus* JCSC1435, *S. hominis* K1, *S. hyicus*, *S. intermedius* NCTC11048, *S. lentus* 050AP, *S. massiliensis* CCUG55927, *S. microti*, *S. pseudintermedius* NA45, *S. saprophyticus* ATCC15305, *S. schleiferi* 2317 03, *S. sciuri* FDA ARGOS285, *S. simiae* NCTC3000, *S. simulans* FDA ARGOS383, *S. vitulinus* F1028, *S. warneri* SG1, *S. xylosum* HKUOPL8). The Identified in all networks were defined as those networks containing only a single homolog in each of the analyzed genomes; not identified in *S. aureus* networks were defined as

networks not containing protein homologs from any of the camSOD-positive genomes, and containing homologs from at least 11 out of 22 camSOD-negative species; *S. aureus*-specific networks were defined as networks containing protein homologs from 7 out of 7 genomes from camSOD-positive species, and none from the camSOD-negative species. To be defined in the group labeled as Enriched in *S. aureus*, each camSOD-positive species had to contain more protein homologs than any of the camSOD-negative species within the network. Heatmaps of the identified protein networks (Fig. 5 and Supplementary Fig. 12) were generated using gplots package (<https://CRAN.R-project.org/package=gplots>) within R.

Staphylococcus species tree. To generate a species tree of the analyzed Staphylococci (Supplementary Data 1), single orthologues from 104 Identified in all networks (Supplementary Data 4), were aligned using MAFFT⁶⁵, trimmed with trimAl⁶⁶, and concatenated with module Nexus of the Biopython⁷⁸ package to a final alignment of 24,091 amino acid sites. The best fitting phylogenetic model was selected using ModelFinder⁷⁹ implemented on the CIPRES Science Gateway V. 3.3 web server⁸⁰, according to the BIC criteria. Phylogenetic tree was generated under LG+F+R3 model with 1000 ultrafast bootstrap replicates using IQ-TREE⁸¹ implemented on the CIPRES Science Gateway V. 3.3 web server⁸⁰.

Circular genome comparison. Circular genome comparison was generated with BRIG⁸², using *S. aureus* NCTC8325 genome⁸³ as a reference, which was compared to 29 analyzed Staphylococcus genomes using BLASTN⁶⁰ (Fig. 5). *S. aureus* NCTC8325 virulence factors (PATRIC⁸⁴, and VFDB⁸⁵ databases), and essential genes (PATRIC⁸⁴ and AureoWiki³⁸ databases) annotations (Supplementary Data 5) were downloaded from the respective databases on 29th March, 2019.

Protein network analyses of metal acquisition systems. Networks containing homologs of proteins known to be involved in manganese and iron import were isolated using igraph⁷⁴ and dnet⁷⁵ packages in R⁷⁶, and visualized in Cytoscape⁷⁷ (Supplementary Fig. 12). Sequence of the proteins present in the networks were aligned using MAFFT⁶⁵, and the alignments were trimmed with trimAl⁶⁶ and used to generate phylogenies under LG + G model in FastTree⁸⁶. Homologs of the proteins of interest were identified based on the network and tree topologies, together with manual inspection of protein sequence alignments, and used to generate the heatmap with gplots package.

Other bioinformatic methods. Multiple sequence alignments were visualized and annotated using Jalview⁸⁷. Phylogenetic trees were visualized and annotated using FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>), Archaeopteryx⁸⁸, and BRIG⁸². Linear genome fragments were aligned and visualized using EasyFig⁸⁹.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Structural data that support the findings of this study have been deposited in the Protein Data Bank with the accession codes PDB 6EX3, PDB 6EX4, PDB 6EX5, PDB 6QV8, and PDB 6QV9 (see Supplementary Table 2). Source data for Table 1, Fig. 6 and Supplementary Fig. 1b–d, f are provided with the paper. There are no restrictions on any data within the manuscript. Biological materials will be provided on request.

Received: 8 August 2019; Accepted: 29 April 2020;

Published online: 01 June 2020

References

- Waldron, K. J., Rutherford, J. C., Ford, D. & Robinson, N. J. Metalloproteins and metal sensing. *Nature* **460**, 823–830 (2009).
- Andreini, C., Bertini, I., Cavallaro, G., Holliday, G. L. & Thornton, J. M. Metal ions in biological catalysis: from enzyme databases to general principles. *J. Biol. Inorg. Chem.* **13**, 1205–1218 (2008).
- Anjem, A., Varghese, S. & Imlay, J. A. Manganese import is a key element of the OxyR response to hydrogen peroxide in *Escherichia coli*. *Mol. Microbiol.* **72**, 844–858 (2009).
- Gu, M. & Imlay, J. A. Superoxide poisons mononuclear iron enzymes by causing mismetallation. *Mol. Microbiol.* **89**, 123–134 (2013).
- Sobota, J. M., Gu, M. & Imlay, J. A. Intracellular hydrogen peroxide and superoxide poison 3-deoxy-D-arabinoheptulosonate 7-phosphate synthase, the first committed enzyme in the aromatic biosynthetic pathway of *Escherichia coli*. *J. Bacteriol.* **196**, 1980–1991 (2014).
- Sheng, Y. et al. Superoxide dismutases and superoxide reductases. *Chem. Rev.* **114**, 3854–3918 (2014).

7. Buvelot, H., Posfay-Barbe, K. M., Linder, P., Schrenzel, J. & Krause, K. H. *Staphylococcus aureus*, phagocyte NADPH oxidase and chronic granulomatous disease. *FEMS Microbiol. Rev.* **41**, 139–157 (2017).
8. Garcia, Y. M. et al. A superoxide dismutase capable of functioning with iron or manganese promotes the resistance of *Staphylococcus aureus* to calprotectin and nutritional immunity. *PLoS Pathog.* **13**, e1006125 (2017).
9. Valderas, M. W., Gatson, J. W., Wreyford, N. & Hart, M. E. The superoxide dismutase gene *sodM* is unique to *Staphylococcus aureus*: absence of *sodM* in coagulase-negative staphylococci. *J. Bacteriol.* **184**, 2465–2472 (2002).
10. Karavolos, M. H., Horsburgh, M. J., Ingham, E. & Foster, S. J. Role and regulation of the superoxide dismutases of *Staphylococcus aureus*. *Microbiol. (Read., Engl.)* **149**, 2749–2758 (2003).
11. Gregory, E. M. Characterization of the O₂-induced manganese-containing superoxide dismutase from *Bacteroides fragilis*. *Arch. Biochem. Biophys.* **238**, 83–89 (1985).
12. Yamakura, F., Sugio, S., Hiraoka, B. Y., Ohmori, D. & Yokota, T. Pronounced conversion of the metal-specific activity of superoxide dismutase from *Porphyromonas gingivalis* by the mutation of a single amino acid (Gly155Thr) located apart from the active site. *Biochemistry* **42**, 10790–10799 (2003).
13. Meier, B., Barra, D., Bossa, F., Calabrese, L. & Rotilio, G. Synthesis of either Fe- or Mn-superoxide dismutase with an apparently identical protein moiety by an anaerobic bacterium dependent on the metal supplied. *J. Biol. Chem.* **257**, 13977–13980 (1982).
14. Martin, M. E. et al. A *Streptococcus mutans* superoxide dismutase that is active with either manganese or iron as a cofactor. *J. Biol. Chem.* **261**, 9361–9367 (1986).
15. Whittaker, M. M. & Whittaker, J. W. Recombinant superoxide dismutase from a hyperthermophilic archaeon, *Pyrobaculum aerophilum*. *J. Biol. Inorg. Chem.: JBIC: a Publ. Soc. Biol. Inorg. Chem.* **5**, 402–408 (2000).
16. Corbin, B. D. et al. Metal chelation and inhibition of bacterial growth in tissue abscesses. *Science (N. Y., NY)* **319**, 962–965 (2008).
17. Kehl-Fie, T. E. et al. Nutrient metal sequestration by calprotectin inhibits bacterial superoxide defense, enhancing neutrophil killing of *Staphylococcus aureus*. *Cell Host Microbe* **10**, 158–164 (2011).
18. Dupont, C. L., Butcher, A., Valas, R. E., Bourne, P. E. & Caetano-Anolles, G. History of biological metal utilization inferred through phylogenomic analysis of protein structures. *Proc. Natl Acad. Sci. USA* **107**, 10567–10572 (2010).
19. Kirschvink, J. L. et al. Paleoproterozoic snowball earth: extreme climatic and geochemical global change and its biological consequences. *Proc. Natl Acad. Sci. USA* **97**, 1400–1405 (2000).
20. Ji, H. -F., Chen, L., Jiang, Y. -Y. & Zhang, H. -Y. Evolutionary formation of new protein folds is linked to metallic cofactor recruitment. *BioEssays* **31**, 975–980 (2009).
21. Valasatava, Y., Rosato, A., Furnham, N., Thornton, J. M. & Andreini, C. To what extent do structural changes in catalytic metal sites affect enzyme function? *J. Inorg. Biochem.* **179**, 40–53 (2018).
22. Conant, G. C. & Wolfe, K. H. Turning a hobby into a job: how duplicated genes find new functions. *Nat. Rev. Genet.* **9**, 938 (2008).
23. Gregory, E. M. & Dapper, C. H. Isolation of iron-containing superoxide dismutase from *Bacteroides fragilis*: reconstitution as a Mn-containing enzyme. *Arch. Biochem. Biophys.* **220**, 293–300 (1983).
24. Barwinska-Sendra, A., Basle, A., Waldron, K. J. & Un, S. A charge polarization model for the metal-specific activity of superoxide dismutases. *Phys. Chem. Chem. Phys.: PCCP* **20**, 2363–2372 (2018).
25. Vance, C. K. & Miller, A. F. Novel insights into the basis for *Escherichia coli* superoxide dismutase's metal ion specificity from Mn-substituted FeSOD and its very high E(m). *Biochemistry* **40**, 13079–13087 (2001).
26. Tabares, L. C., Gatjens, J. & Un, S. Understanding the influence of the protein environment on the Mn(II) centers in Superoxide Dismutases using High-Field Electron Paramagnetic Resonance. *Biochim. Biophys. Acta* **1804**, 308–317 (2010).
27. Sjodin, M. et al. Tuning the redox properties of manganese(II) and its implications to the electrochemistry of manganese and iron superoxide dismutases. *Inorg. Chem.* **47**, 2897–2908 (2008).
28. Miller, A. F. Redox tuning over almost 1 V in a structurally conserved active site: lessons from Fe-containing superoxide dismutase. *Acc. Chem. Res.* **41**, 501–510 (2008).
29. Osawa, M. et al. Conversion of the metal-specific activity of *Escherichia coli* Mn-SOD by site-directed mutagenesis of Gly165Thr. *Biochim. Biophys. Acta* **1804**, 1775–1779 (2010).
30. Leveque, V. J., Vance, C. K., Nick, H. S. & Silverman, D. N. Redox properties of human manganese superoxide dismutase and active-site mutants. *Biochemistry* **40**, 10586–10591 (2001).
31. Tu, W. Y. et al. Cellular iron distribution in *Bacillus anthracis*. *J. Bacteriol.* **194**, 932–940 (2012).
32. Tabares, L. C., Bittel, C., Carrillo, N., Bortolotti, A. & Cortez, N. The single superoxide dismutase of *Rhodobacter capsulatus* is a cambialistic, manganese-containing enzyme. *J. Bacteriol.* **185**, 3223–3227 (2003).
33. Yamano, S., Sako, Y., Nomura, N. & Maruyama, T. A cambialistic SOD in a strictly aerobic hyperthermophilic archaeon, *Aeropyrum pernix*. *J. Biochem.* **126**, 218–225 (1999).
34. Tong, S. Y. et al. Novel staphylococcal species that form part of a *Staphylococcus aureus*-related complex: the non-pigmented *Staphylococcus argenteus* sp. nov. and the non-human primate-associated *Staphylococcus schweitzeri* sp. nov. *Int. J. Syst. Evol. Microbiol.* **65**, 15–22 (2015).
35. Zhang, D. -F. et al. Preliminary comparative genomics revealed pathogenic potential and international spread of *Staphylococcus argenteus*. *BMC Genomics* **18**, 808 (2017).
36. Chaudhuri, R. R. et al. Comprehensive identification of essential *Staphylococcus aureus* genes using Transposon-Mediated Differential Hybridisation (TMDH). *BMC Genomics* **10**, 291 (2009).
37. Forsyth, R. A. et al. A genome-wide strategy for the identification of essential genes in *Staphylococcus aureus*. *Mol. Microbiol.* **43**, 1387–1400 (2002).
38. Fuchs, S. et al. AureoWiki The repository of the *Staphylococcus aureus* research and annotation community. *Int. J. Med. Microbiol.: IJMM* **308**, 558–568 (2018).
39. Bae, T. et al. *Staphylococcus aureus* virulence genes identified by bursa aurealis mutagenesis and nematode killing. *Proc. Natl Acad. Sci. USA* **101**, 12312–12317 (2004).
40. Suzuki, H., Lefebvre, T., Bitar, P. P. & Stanhope, M. J. Comparative genomic analysis of the genus *Staphylococcus* including *Staphylococcus aureus* and its newly described sister species *Staphylococcus simiae*. *BMC Genomics* **13**, 38 (2012).
41. Planet, P. J. et al. Architecture of a species: phylogenomics of *Staphylococcus aureus*. *Trends Microbiol.* **25**, 153–166 (2017).
42. Cheng, A. G. et al. Genetic requirements for *Staphylococcus aureus* abscess formation and persistence in host tissues. *FASEB J.: Off. Publ. Federation Am. Societies Exp. Biol.* **23**, 3393–3404 (2009).
43. Conroy, B. S., Grigg, J. C., Kolesnikov, M., Morales, L. D. & Murphy, M. E. P. *Staphylococcus aureus* heme and siderophore-iron acquisition pathways. *Biomaterials* **85495708** (2019).
44. Hunter, T. et al. A single mutation is sufficient to modify the metal selectivity and specificity of a eukaryotic manganese superoxide dismutase to encompass iron. *Chem. A Eur. J.* **24**, 5303–5308 (2018).
45. Hosseinzadeh, P. & Lu, Y. Design and fine-tuning redox potentials of metalloproteins involved in electron transfer in bioenergetics. *Biochim. Biophys. Acta* **1857**, 557–581 (2016).
46. Cook, S. A. & Borovik, A. S. Molecular designs for controlling the local environments around metal ions. *Acc. Chem. Res.* **48**, 2407–2414 (2015).
47. Hosseinzadeh, P. et al. Enhancing Mn(II)-binding and manganese peroxidase activity in a designed cytochrome c peroxidase through fine-tuning secondary-sphere interactions. *Biochemistry* **55**, 1494–1502 (2016).
48. Span, E. A., Suess, D. L. M., Deller, M. C., Britt, R. D. & Marletta, M. A. The role of the secondary coordination sphere in a fungal polysaccharide monooxygenase. *ACS Chem. Biol.* **12**, 1095–1103 (2017).
49. Mann, S. I., Heinisch, T., Ward, T. R. & Borovik, A. S. Coordination chemistry within a protein host: regulation of the secondary coordination sphere. *Chem. Commun. (Camb., Engl.)* **54**, 4413–4416 (2018).
50. Parker, M. W. & Blake, C. C. Iron- and manganese-containing superoxide dismutases can be distinguished by analysis of their primary structures. *FEBS Lett.* **229**, 377–382 (1988).
51. Kim, H. K. et al. IsdA and IsdB antibodies protect mice against *Staphylococcus aureus* abscess formation and lethal challenge. *Vaccine* **28**, 6382–6392 (2010).
52. Lane, T. W. & Morel, F. M. A biological function for cadmium in marine diatoms. *Proc. Natl Acad. Sci. USA* **97**, 4627–4631 (2000).
53. Tarrant, E. et al. Copper stress in *Staphylococcus aureus* leads to adaptive changes in central carbon metabolism. *Metallomics: Integr. Biometal Sci.* **11**, 183–200 (2019).
54. Morrissey, J. A., Cockayne, A., Hill, P. J. & Williams, P. Molecular cloning and analysis of a putative siderophore ABC transporter from *Staphylococcus aureus*. *Infect. Immun.* **68**, 6281–6288 (2000).
55. Bae, T. & Schneewind, O. Allelic replacement in *Staphylococcus aureus* with inducible counter-selection. *Plasmid* **55**, 58–63 (2006).
56. Whitmore, L. & Wallace, B. A. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* **32**, W668–673 (2004).
57. Sreerama, N. & Woody, R. W. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**, 252–260 (2000).
58. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **63**, 32–41 (2007).

59. Ashkenazy, H. et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–350 (2016).
60. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
61. Brown, C. T. et al. Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature* **523**, 208–211 (2015).
62. Probst, A. J. et al. Differential depth distribution of microbial function and putative symbionts through sediment-hosted aquifers in the deep terrestrial subsurface. *Nat. Microbiol.* **3**, 328–336 (2018).
63. Spang, A. et al. Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521**, 173–179 (2015).
64. Zaremba-Niedzwiedzka, K. et al. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **541**, 353–358 (2017).
65. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evolution* **30**, 772–780 (2013).
66. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics (Oxf., Engl.)* **25**, 1972–1973 (2009).
67. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics (Oxf., Engl.)* **27**, 1164–1165 (2011).
68. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxf., Engl.)* **30**, 1312–1313 (2014).
69. Asnicar, F., Weingart, G., Tickle, T. L., Huttenhower, C. & Segata, N. Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ* **3**, e1029 (2015).
70. Fonseca-Junior, N. J., Afonso, M. Q. L., Oliveira, L. C. & Bleicher, L. PFstats: a network-based open tool for protein family analysis. *J. Computat. Biol.* **25**, 480–486 (2018).
71. Su, G., Morris, J. H., Demchak, B. & Bader, G. D. Biological network exploration with Cytoscape 3. *Curr. Protoc. Bioinforma.* **47**, 8.13.11–24 (2014).
72. Halary, S., McInerney, J. O., Lopez, P. & Baptiste, E. EGN: a wizard for construction of gene and genome similarity networks. *BMC Evolut. Biol.* **13**, 146 (2013).
73. Huerta-Cepas, J. et al. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* **44**, D286–293 (2016).
74. Csardi, G. & Nepusz, T. The igraph software package for complex network research. *InterJournal. Complex Syst.* **1695**, 1–9 (2006).
75. Fang, H. & Gough, J. The ‘dnet’ approach promotes emerging research on cancer patient survival. *Genome Med.* **6**, 64 (2014).
76. R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing (R Development Core Team, 2008).
77. Smoot, M. E., Ono, K., Ruschinski, J., Wang, P. L. & Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics (Oxf., Engl.)* **27**, 431–432 (2011).
78. Cock, P. J. et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics (Oxf., Engl.)* **25**, 1422–1423 (2009).
79. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermini, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
80. Trifinopoulos, J., Nguyen, L. T., von Haeseler, A. & Minh, B. Q. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* **44**, W232–235 (2016).
81. Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evolution* **32**, 268–274 (2015).
82. Alikhan, N. F., Petty, N. K., Ben Zakour, N. L. & Beatson, S. A. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* **12**, 402 (2011).
83. Gillaspay, A. F. et al. in *Gram-Positive Pathogens*. 2nd edn. (American Society of Microbiology, 2006).
84. Wattam, A. R. et al. Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Res.* **45**, D535–d542 (2017).
85. Chen, L., Zheng, D., Liu, B., Yang, J. & Jin, Q. VFDB 2016: hierarchical and refined dataset for big data analysis-10 years on. *Nucleic Acids Res.* **44**, D694–697 (2016).
86. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2-approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
87. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. *Bioinformatics (Oxf., Engl.)* **25**, 1189–1191 (2009).
88. Han, M. V. & Zmasek, C. M. phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinforma.* **10**, 356 (2009).
89. Sullivan, M. J., Petty, N. K. & Beatson, S. A. Easyfig: a genome comparison visualizer. *Bioinformatics (Oxf., Engl.)* **27**, 1009–1010 (2011).

Acknowledgements

K.J.W. was supported by a Wellcome Trust- and Royal Society-funded Sir Henry Dale fellowship (098375/Z/12/Z). A.B.-S. was supported by a BBSRC Ph.D. studentship and funding from Newcastle University's Faculty of Medical Sciences, except during a revision period when she was funded by a Wellcome Trust Collaborative Award (204877/Z/16/Z). E.S.M. was also supported by a BBSRC Ph.D. studentship. A.B. was funded by Newcastle University's Faculty of Medical Sciences. T.E.K.-F. was supported by operating grants from the National Institutes of Health (R01 AI118880) and the Vallee Foundation. The high-field EPR measurements benefited from the biophysics platform of the Institute for Integrative Biology of the Cell, supported by the French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INBS-05. We thank Diamond Light Source for access to beamline I03, I04, and I24 (mx7864, mx9948, and mx18598). The contents of this work are solely the responsibilities of the authors and do not reflect the official views of any of the funders, who had no role in study design, data collection, analysis, decision to publish, or preparation of the manuscript.

Author contributions

A.B.-S. designed and performed most in vitro experiments, with further contributions from E.S.M., E.T., and P.C. Y.M.G. performed *S. aureus* experiments. A.B. assisted A.B.-S. with X-ray crystallography, and K.M.S. performed all bioinformatic analyses with assistance from C.B. for network analysis. S.U. performed HFEPR experiments and analyzed these data, with input from L.C.T., who also prepared the auto-oxidation calculations. K.J.W. and T.E.K.F. conceived and managed the study, and wrote the manuscript with S.U., with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-16478-0>.

Correspondence and requests for materials should be addressed to T.E.K.-F. or K.J.W.

Peer review information *Nature Communications* thanks Jeffrey Harmer, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020