Supplementary Information for

Bacillithiol is an antioxidant thiol produced in Bacilli

Gerald L. Newton, Mamta Rawat, James J. La Clair, Vishnu Karthik Jothivasan, Tanya Budiarto, Chris J. Hamilton, Al Claiborne, John D. Helmann, and Robert C. Fahey

1Department of Chemistry and Biochemistry, University of California, San Diego, California 92093-0314
2Department of Biology, California State University-Fresno, Fresno, California 93740
3School of Chemical Sciences and Pharmacy, University of East Anglia, Norwich NR4 7TJ, England
4Center for Structural Biology, Wake Forest University School of Medicine, Winston-Salem, North Carolina 27157
5Department of Microbiology, Cornell University, Ithaca, New York 14853

Corresponding author: rcfah@ucsd.edu
List of Supplementary Figures

Figure S1. Demonstration that U8mB (11) and U9mB (11) are rapidly interconverting anomers.

Figure S2. Demonstration by chiral HPLC that CySmB purified from hydrolysis of 8 is L-CySmB and not D-CySmB.

Figure S3. NMR spectra of 8: (a) 500 MHz proton NMR spectrum of prep 2 at pH 6.3; (b) 125 MHz carbon NMR spectrum of prep 2 at pH 6.3; (c) gradient COSY NMR correlation of prep 2 at pH 6.3; (d) expansion of gradient COSY NMR correlation of prep 2 at pH 6.3; (e) $^1$H-$^1$H rotating-frame nuclear overhauser effect spectrum (ROESY) of prep 2 at pH 6.3; (f) expansion of ROESY of prep 2 at pH 6.3; (g) expansion of ROESY demonstrating malate to glucosamine glycosidic bond; (h) $^1$H-$^{13}$C heteronuclear single quantum correlation (HSQC) spectrum of prep 2 at pH 6.3; (i) $^1$H-$^{13}$C HMBC spectrum of prep 2 at pH 6.3; (j) expansion of $^1$H-$^{13}$C HMBC spectrum of prep 2 at pH 6.3; (k) 500 MHz proton NMR of prep 1 at pH~3; (l) gradient COSY spectrum of prep 1 at pH~3; (m) $^1$H-$^1$H TOCSY spectrum of prep 1 at pH~3.

Figure S4. Proton NMR spectra recorded at 500 MHz and pH 6.3 for 11 (a) and 12 (b).
Fig. S1. Demonstration that U8mB (11) and U9mB (11) are rapidly interconverting anomers. HPLC analysis of: a, sample of 4 h hydrolysis of 8; b, purified U8mB; c, purified U9mB; d, purified 8.

Fig. S2. Demonstration by chiral HPLC that CySmB purified from hydrolysis of 8 is L-CySmB and not D-CySmB: a, 100 pmol L-CySmB standard; b, 100 pmol D-CySmB standard; c, CySmB from 8h hydrolysis of 8; all samples purified by reversed-phase HPLC. The small peaks eluting at 2-10 min are present in a solvent control injection.
$^1$H-NMR
D$_2$O+NH$_4$HCO$_2$

pH 6.3
500 MHz
$^{13}$C-NMR
D$_2$O+NH$_4$HCO$_2$

pH 6.3

125 MHz
$^1$H–$^1$H gCOSY
D$_2$O+NH$_4$HCO$_2$

pH 6.3

500 MHz
$^1$H–$^1$H gCOSY
$D_2O + NH_4HCO_2$
$pH 6.3$
500 MHz
$^1$H–$^1$H ROESY
D$_2$O + NH$_4$HCO$_2$

pH 6.3
500 MHz
$^1$H–$^1$H ROESY
D$_2$O + NH$_4$HCO$_2$

pH 6.3

500 MHz
$^1$H-$^1$H ROESY
D$_2$O+NH$_4$HCO$_2$

pH 6.3

500 MHz
$^1$H-$^{13}$C HSQC
D$_2$O + NH$_4$HCO$_2$

pH 6.3

500 MHz
$^1$H-$^{13}$C HMBC
D$_2$O + NH$_4$HCO$_2$

pH 6.3
500 MHz
$^1$H-$^{13}$C HMBC
D$_2$O + NH$_4$HCO$_2$

pH 6.3

500 MHz
$^1$H-NMR
D$_2$O+TFA
pH 3.0
500 MHz
$^{1}H,^{1}H$-gCOSY
$\text{D}_2\text{O}+\text{TFA}$
pH 3.0
500 MHz
$^1$H,$^1$H-TOCSY
D$_2$O+TFA
pH 3.0
500 MHz
L-CysMeB-D-GlcN (11)

$^1$H-NMR

D$_2$O + NH$_4$HCO$_2$

500 MHz

pH 6.3
D-CySmB-D-GlcN (12)
$^1$H-NMR
D$_2$O + NH$_4$HCO$_2$
500 MHz
pH 6.3
Table S1. NMR data for 8 recorded in D$_2$O with 10 mM ammonium formate pH 6.3.

| position | $\delta_H$ (J in Hz) | $\delta_C$ | gCOSY | ROESY | HMBC$^c$ |
|----------|----------------------|------------|-------|-------|---------|
| bimane   |                      |            |       |       |         |
| 1        |                     | 162.6      |       |       |         |
| 2        |                     | 114.0      |       |       |         |
| 3        |                     | 147.5      |       |       |         |
| 4        |                     | 150.0      |       |       |         |
| 5        |                     | 111.6      |       |       |         |
| 6        |                     | 163.3      |       |       |         |
| 7        | 1.80 s              | 6.2        | $3''^e$, 8 | 1, 2, 3 |
| 8a       | 3.91 m$^b$          | 24.9       | $7, 9, 3''^d$ |
| 8b       | 3.88 m$^b$          |           | $7, 9, 3''^d$ |
| 9        | 2.38 s              | 5.6        | 10     | 8, 10  | 4, 5$^e$ |
| 10       | 1.71 s              |             | 9      | 9      | 4, 5, 6 |
| L-cysteine |                 |            |       |       |         |
| $1''$    |                     | 168.2      |       |       |         |
| $2''$    | 4.09 m$^b$          | 52.4       | $3a'', 3b''$ | $3''^d$ | $1''$ |
| $3a''$   | 3.08 dd (5.4, 12.5) | 32.5       | $2'', 3b''$ | $7, 8, 2'', 3b''$ | $1''$ |
| $3b''$   | 3.12 dd (4.5, 12.4) |           | $2'', 3a''$ | $7, 8, 2'', 3a''$ | $1''$ |
| D-glucosamine |               |            |       |       |         |
| $1'$     |                     | 99.0       | $2'$  | 2', 2'' | 3', 2''$ |
| $2'$     |                     | 54.3       | $1'$  | 1', 4'  | 1''$   |
| $3'$     |                     | 59.2       | $4'$  | $5'$    |        |
| $4'$     | 3.49 dd (9.0, 10.0) | 69.8       | $3', 5'$ | 2', $6a'^e$, $6b'^e$ | $5'^e$ |
| $5'$     | 3.88 m$^b$          | 72.2       | $4', 6a', 6b'$ | $3', 6a'^e$, $6b'^e$ |        |
| $6a'$    | 3.69 dd (2.3, 12.7) | 59.9       | $5', 6b'$ | $4', 6b'$ |        |
| $6b'$    | 3.74 dd (3.7, 12.7) |           | $5', 6a'$ | $4', 6a'$ |        |
| L-malate |                      |            |       |       |         |
| $1''$    |                     | 179.4      |       |       |         |
| $2''$    | 4.09 dd (2.5, 11.6) | 80.0       | $3''^d$ | $1', 3a'', 3b''$ | $1''$ |
| $3a''$   | 2.55 dd (2.4, 15.2) | 41.5       | $2'', 3b''$ | $2'', 3b''$ | $4''^e$ |
| $3b''$   | 2.41 dd (11.7, 15.1) |         | $2'', 3a''$ | $2'', 3a''$ | $4''^e$ |
| $4''$    |                     | 179.1      |       |       |         |

$^a$ $^{13}$C shift were assigned according to their HSQC correlations. $^b$ A multiplet is reported due to overlapping peaks. $^c$ HMBC data was collected with the optimal $J_{1,3}$=8.0 Hz and are given by correlations from $^1$H to 13C. $^d$ Could not distinguish between the two hydrogens on this carbon. $^e$ Denotes that a weak correlation was observed.
Table S2. Thiol content of selected bacteria.

| Genus/species/Strain                  | BSH | Cys | CoA | GSH | MSH |
|--------------------------------------|-----|-----|-----|-----|-----|
| Deinococcus radiodurans              | 0.6 | 0.15| 0.8 | <0.01| <0.01|
| Pseudomonas aeruginosa               | <0.01| 0.7 | 0.1 | 1.1  | <0.01|
| Bacillus anthracis Sterne            | 0.2 | 0.2 | 0.9 | <0.03| <0.03|
| Bacillus cereus                      | 0.3 | 0.3 | 0.1 | <0.1 | <0.01|
| Bacillus megaterium                  | 0.3 | 0.5 | 1.2 | <0.1 | <0.01|
| Bacillus pumilis                     | 0.7 | 0.3 | 0.3 | <0.1 | <0.01|
| Bacillus subtilis 6051               | 0.3 | 0.4 | 0.5 | <0.1 | <0.01|
| Bacillus subtilis CU1065             | 0.6 | 0.6 | 0.07| <0.01| <0.01|
| Bacillus subtilis JH642              | 0.6 | 0.6 | 0.6 | <0.01| <0.01|
| Geobacillus stearothermophilus       | 0.07| 0.3 | 0.2 | <0.01| <0.01|
| Staphylococcus aureus                | 0.7 | 0.5 | 0.4 | <0.01| <0.01|
| Staphylococcus epidermidis           | <0.01| 0.07| 0.4 | <0.01| <0.01|
| Staphylococcus saprophyticus         | 0.2 | 0.3 | 0.3 | <0.01| <0.01|
| Streptococcus agalactiae             | 0.2 | 0.2 | 0.09| ≤0.1 | <0.01|
| Streptococcus mutans                 | <0.01| <0.01| <0.01| 0.4  | <0.01|
| Streptococcus pyogenes               | <0.01| 0.08| 0.04| 0.8  | <0.02|
| Enterococcus faecalis                | <0.01| 0.4 | <0.01| 2    | <0.02|
| Rubrobacter xylanophylis             | <0.05| 0.3 | nd  | 3    | <0.01|
| Arthrobacter aurescens               | 0.09| 0.3 | nd  | <0.01| 0.8  |
| Arthrobacter histodinolovorans       | 0.08| 0.3 | nd  | <0.01| 0.6  |
| Kocuria rhizophila                   | <0.01| 0.13| 0.5 | <0.01| 4.9  |
| Mycobacterium smegmatis              | <0.01| 0.15| 3   | <0.02| 13   |

*aMean of triplicate determinations; error ≤ 25%. The values given are equivalent to ~3 times the intracellular thiol concentration in mM.

*bIncludes dephospho-CoA

*cNot determined
Table S3. The intracellular thiol concentration and redox ratio of exponentially growing bacteria.

| Strain                  | Thiol   | Thiol Content | Intracellulara | Redox Ratiob |
|-------------------------|---------|---------------|----------------|--------------|
|                         |         | µmole/gm dry wt or nmoles/10^9 cells* | Thiol (mM)     |              |
| E. coli e               | GSH     | 19            | 6.3            | 320          |
| M. smegmatis d          | MSH     | 40*           | 6.6            | 400          |
| M. bovis BCG d          | MSH     | 17-25*        | 3.5            | 100          |
| M. tuberculosis e        | MSH     | 20*           | 3.3            | 360 ± 100    |
| B. subtilis JH642 f     | BSH     | 0.6           | 0.2            | 400 ± 140    |
| B. subtilis JH642 f     | Cys     | 0.6           | 0.2            | 120 ± 80     |
| B. subtilis JH642 f     | CoA     | 0.6           | 0.2            | ND           |
| D. radiodurans f        | BSH     | 0.6           | 0.2            | 100 ± 20     |
| D. radiodurans f        | Cys     | 0.15          | 0.005          | 46 ± 6       |
| D. radiodurans f        | CoA     | 0.8           | 0.3            | ND           |

*a*The intracellular thiol concentration was estimated from the thiol content in µmole/gm dry weight using 3 µl H₂O per mg dry weight. We assumed that 1 x 10^9 cells contain 2 mg of dry weight for conversion to intracellular concentration.

*b*The redox ratios are expressed as thiol/disulfide. The disulfide is estimated as half the thiol content released by DTT from an N-ethylmaleimide blocked sample (see Supplemental Methods). Redox ratios cannot be determined (ND) for CoA with DTT as it cleaves thioesters of CoA giving a high disulfide analysis.

*References* 1, 2, 3, and 4.

*This study.*
Table S4. Candidate Bacillithiol disulfide reductase genes from *B. subtilis* strain 168.

| Gene/synonym       | E value | Annotated function                  | Functional validation |
|--------------------|---------|-------------------------------------|-----------------------|
| Homologs of *M. tuberculosis* mycothiol disulfide reductase (Mtr, Rv2855) blastp results<sup>a</sup> |
| BSU14610/pdhD      | 1e-44   | Dihydrolipoamide dehydrogenase      | Dihydrolipoamide dehydrogenase<sup>4</sup> |
| BSU24060/yqiV      | 9e-39   | Dihydrolipoamide dehydrogenase      | none                  |
| BSU08090/yfjH      | 3e-37   | Dihydrolipoamide dehydrogenase      | none                  |
| BSU03300/nasD      | 8e-9    | Nitrate reductase                   | none                  |
| BSU03320/nasB      | 6e-8    | Nitrate reductase                   | Nitrate reductase<sup>5</sup> |
| Homologs of *E. coli* Glutathione reductase (GOR) blastp results<sup>a</sup> |
| BSU14610/pdhD      | 5e-51   | Dihydrolipoamide dehydrogenase      | Dihydrolipoamide dehydrogenase<sup>4</sup> |
| BSU08090/yfjH      | 2e-42   | Dihydrolipoamide dehydrogenase      | none                  |
| BSU24060/yqiV      | 5e-34   | Dihydrolipoamide dehydrogenase      | none                  |
| BSU03320/nasB      | 1e-8    | Nitrate reductase                   | Nitrate reductase<sup>5</sup> |
| BSU40100/ahpF      | 2e-7    | Alkyl hydroperoxide reductase       | Alkyl hydroperoxide reductase<sup>6</sup> |

<sup>a</sup>BLASTP 2.2.19+<sup>7</sup> search of NCBI microbial genomes.
Table S5. NMR data for 11 and 12 recorded in D$_2$O with ammonium formate pH 6.3.

| position | $\alpha$-11 $\delta_H^a$ | $\beta$-11 $\delta_H^a$ | $\alpha$-12 $\delta_H^a$ | $\beta$-12 $\delta_H^a$ |
|----------|-------------------------|-------------------------|-------------------------|-------------------------|
| bimane   |                         |                         |                         |                         |
| 1        |                         |                         |                         |                         |
| 2        |                         |                         |                         |                         |
| 3        |                         |                         |                         |                         |
| 4        |                         |                         |                         |                         |
| 5        |                         |                         |                         |                         |
| 6        |                         |                         |                         |                         |
| 7        | 1.87 s                  | 1.88 s$^c$              | 1.86 s$^c$              |                         |
| 8        | 3.97-3.91 m             |                         | 4.04-3.93 m             |                         |
| 9        | 2.46 s                  |                         | 2.45 s                  |                         |
| 10       | 1.78 s                  |                         | 1.79 s                  |                         |
| L-cysteine |                    |                         |                         |                         |
| 1"       |                         |                         |                         |                         |
| 2"       | 4.20 t (6.0)$^c$        | 4.24 t (6.0)$^c$        | 4.26 t (6.5)$^c$        | 4.24 t (6.5)$^c$        |
| 3"       | 3.26-3.16 m             |                         | 3.21-3.12 m             |                         |
| D-glucosamine |          |                         |                         |                         |
| 1'       | 4.79 d (8.5)            | 5.25 d (3.5)            | 4.74 d (8.0)            | 5.19 d (3.5)            |
| 2'       | 3.78 m                  | 3.90 m                  | 3.76 m                  | 3.95 m                  |
| 3'       | 3.86-3.70 m$^b$         |                         | 3.89-3.76 m$^b$         |                         |
| 4'       | 3.71-3.40 m$^b$         |                         | 3.65-3.52 m$^b$         |                         |
| 5'       | 3.52-3.48 m$^b$         |                         | 3.69-3.58 m$^b$         |                         |
| 6'       | 3.89-3.75 m$^b$         |                         | 3.81-3.77 m$^b$         |                         |

$^a$ $J$ values are provided in Hz. $^b$ Assignments were made based on gCOSY data and are noted by a tentative region. The complexity of the anomeric mixture complicated these assignments. $^c$ Multiple spin systems were observed and were tentatively assigned to each anomer based on their integration.
SUPPLEMENTAL METHODS

**Bacterial strains and culture conditions.** All bacterial strains were started from a single colony in 5 ml trypticase soy broth media (BD Biosciences). This starter culture was used to inoculate 100 ml of broth and the cultures were incubated at appropriate temperatures under shaking conditions until the cultures reached mid-log phase growth.

*Staphylococcus aureus* (ATCC 25923), *Staphylococcus epidermidis* (ATCC 12228), *Staphylococcus saprophyticus* (ATCC 15305), *Streptococcus agalactiae* (ATCC 27956), *Streptococcus mutans* (ATCC 35668), *Streptococcus pyogenes* (ATCC 19615), *Enterococcus faecalis* (ATCC 29212), *Pseudomonas aeruginosa* (ATCC 10145), and *Deinococcus radiodurans* (ATCC 13939) were cultured at 37 °C in trypticase soy broth (BD Biosciences). Environmental isolates *Arthrobacter aurescens* (AW100 from A. Wright), *Arthrobacter histodinolovorans* (AW101 from A. Wright), and *Kocuria rhizophila* (ATCC 9341) were cultured in trypticase soy broth at 30 °C. *Bacillus cereus* (ATCC 14579), *Bacillus megaterium* (QMB1551, from A. Claiborne), *Bacillus pumilus* (ATCC 945), *Bacillus subtilis* (ATCC 6051), *Bacillus subtilis* (CU1065 from J. Helmann) and *Bacillus subtilis* (JH642 from J. Helmann) were cultured in trypticase soy broth at 37 °C. *Bacillus anthracis* Sterne (from D. Guiney) was cultured in brain heart infusion medium (Becton, Dickinson and Co.) at 37 °C. *Mycobacterium smegmatis* mc²155 (ATCC 700084) was cultured in 7H9 Middlebrook broth (Difco) with 0.4% glucose and 0.05% Tween 80 at 37 °C. The thermophiles, *Rubrobacter xylanophilus* (DSM 9941 from J. Batista) and *Geobacillus stearothermophilus* (ATCC 12890), were cultured in trypticase soy broth at 55 °C. Before harvesting the cultures for thiol analysis, the purity of the cultures was confirmed by restreaking the liquid cultures on trypticase soy agar plates to check for colony morphology and by Gram stain. The cultures were then divided into two aliquots and harvested by centrifugation for 10 min at 5,000 x g at 4 °C. The cells were either processed immediately or frozen and stored at -80 °C until thiol analysis. For the *Bacillus* strains, the starter colonies were heat-inactivated by heating at 65 °C for 15 min to ensure that only spores were used to start the culture. Growth curves of *D. radiodurans* (ATCC 13939) and *B. subtilis* (JH642) were performed in triplicate by diluting a starter culture in 100 ml trypticase soy broth to a 0.05 optical density at 600 nm. Cultures were harvested at appropriate time intervals and the harvested samples were restreaked and Gram stained to confirm purity.

**Analysis of biological thiols.** All chemicals were of reagent grade or higher from Fisher Biosciences, except as noted. Analytical reversed phase HPLC chromatography was performed on a Waters model 600E liquid chromatograph with a Waters Wisp 712 autoinjector, a Waters 486 tunable absorbance detector, a Waters (Micromass, ZMD) quadrupole mass detector and a St. Johns 2001A Fluoro-Tec filter fluorometer. Thiols were determined by labeling with monobromobimane (mBBr, Molecular Probes-Invitrogen) during extraction in warm (60 °C) 50% acetonitrile and analysis by high performance liquid chromatography (HPLC) with fluorescence detection as previously described, except as noted below. HPLC A buffer was adjusted from pH 3.6 to pH 4.0 and this reduced the retention of the unknown thiol-bimane derivative detected in *B. anthracis* from 12 min (U12) to 10 min; the unknown was redesignated 8. This change increases the resolution of 8 (10 min) from CySmB (9, 12 min).
Method A
Column: 4.5 x 250 mm Beckman Ultrasphere C18 (5 µm) ion pairing (#235335).
Temperature: 23 °C
Flow rate: 1.2 ml/min.
Buffer A = aqueous 0.25% v/v acetic acid adjusted to pH 4.0 with NaOH.
Buffer B = HPLC grade methanol
Binary linear gradient:
| Time (min) | %B |
|------------|----|
| 0          | 10 |
| 5          | 10 |
| 15         | 20 |
| 30         | 30 |
| 32         | 100|
| 34         | 10 |
| 45         |    |
| 47         | 100|
| 50         |    |
| 60         |    |
|            | reinject. |

The retention times (min) for the bimane standards were: Cys-GlcN-Ins\(^8\) (8.3), bacillithiol (10), cysteine (11.5), glutathione (18), mycothiol (21) and N-acetylcysteine (24).

A tetrabutylammonium ion pairing HPLC chromatography analysis (Method B) was used to confirm method A thiol analysis and provide analysis of coenzyme A, 4’-phosphopantetheine, and \(\text{H}_2\text{S}\).

Method B
Column: 3.9 x 150 mm Waters Symmetry C8 (5 µm).
Temperature: 23 °C
Flow rate: 1.0 ml/min.
Buffer A = aqueous 10 mM tetrabutylammonium phosphate (Fluka 86842) adjusted to pH 3.4 with acetic acid.
Buffer B = 10% water, 90% HPLC grade methanol containing 10 mM tetrabutylammonium phosphate adjusted to pH 3.4 with acetic acid.
Binary linear gradient:
| Time (min) | %B |
|------------|----|
| 0          | 10 |
| 15         | 25 |
| 30         | 50 |
| 40         | 75 |
| 45         | 100|
| 47         | 100|
| 50         |    |
| 60         |    |
|            | reinject. |

The retention times (min) for the bimane derivatives were: Cys-GlcN-Ins\(^8\) (1.3, unretained), cysteine (7.5), mycothiol (15), bacillithiol (19), glutathione (26), N-acetylcysteine (33.2), \(\text{H}_2\text{S}\) (34.4), 4’-phosphopantetheine (37), dephosphoCoA (40.5) and CoA (43). Authentic thiol-bimane standards were prepared as previously described\(^{11}\) and responses were found to be linear in the
range of 1-500 pmole standard. All unknown thiol-bimane derivatives were calculated on the basis of cysteine-bimane as standard.

**Determination of BSH (13) Redox Status.** The redox status of 13 in *B. subtilis* JH642 and *D. radiodurans* was estimated for exponential phase cells from the soluble thiol and disulfide content. One liter cultures of *B. subtilis* and *D. radiodurans* were grown in trypticase soy broth as described above and harvested at OD$_{600}$ = 1.4 and 0.52, respectively. The cell pellets from the 1 liter culture were frozen at -70 °C. Each pellet was split into 6 equal portions of ~50 mg residual dry weight each. Three cell pellets were extracted with 1 ml of pre-warmed (60 °C) 50% aqueous-acetonitrile containing 2 mM mBBr and 20 mM 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES) pH 8.0. These thiol samples were incubated at 60 °C for 15 min, cooled on ice and acidified with 5 µl of 5M methanesulfonic acid. The cell debris was removed by centrifugation as described above. Three cell pellets were extracted with 1 ml of pre-warmed (60 °C) 50% aqueous-acetonitrile containing 5 mM N-ethylmaleimide (NEM) and 20 mM HEPES pH 8.0. These disulfide samples were incubated at 60 °C for 15 min and cooled on ice. The cell debris was removed by centrifugation. A 0.8 ml aliquot of the supernatant was mixed with 2.5 mM dithiothreitol (DTT) to react with residual NEM and reduced in volume to 0.2 ml on a speedvac. The 0.2 ml supernatant was reduced with 2 mM DTT for 20 min at 23 °C and mBBr was added to a final concentration of 6 mM and reacted in the dark for an additional 10 min. The labeling reaction was quenched with 5 µl 5 M methanesulfonic acid. The redox ratio was expressed as thiol/disulfide or RSH/RSSR. The redox ratio was not provided for coenzyme A (CoA) as DTT cleaves thioesters such as acetylCoA leading to an overestimation of the disulfide form of CoA and an underestimate of the CoA redox ratio$^{12}$.

**Isolation of 8.** Two independent samples of 13 were isolated from *D. radiodurans* cells as the bimane derivative 8. Five liters of *D. radiodurans* were cultured to early stationary phase (50 h) in trypticase soy broth at 37 °C with shaking (225 rpm). The cells were pelleted by centrifugation (5000 × g) at 4 °C. The cell pellet (20 g) was extracted for 30 min with 200 ml of 50% aqueous acetonitrile (60 °C) containing 20 mM HEPES pH 8.0 and 0.5 mM mBBr. The extract was adjusted to pH ~3 with trifluoroacetic acid and cooled on ice. The cell debris was pelleted by centrifugation (10000 × g) and the 50% acetonitrile extract was retained. The pellet was suspended in 50% aqueous acetonitrile and the washed cell debris was pelleted by centrifugation (10000 × g). The monobromobimane extract and the 50% acetonitrile pellet wash were combined and concentrated to 50 ml under reduced pressure. The extract was clarified by centrifugation and the supernatant purified by solid phase extraction on a Sep Pak C18 cartridge (5 g, Waters Corp.). The Sep Pak cartridge was eluted in 3 volumes of aqueous 0.1% TFA containing 0%, 10%, 20% and 50% methanol. Compound 8 eluted in the 10% methanol fraction. This fraction was concentrated in a speedvac to ~5 ml for purification by preparative HPLC.

The purification of crude 8 was achieved by two consecutive preparative HPLC separations. The sample was fractionated on a reversed phase C18 column (Vydac, 22 x 250 mm, #218TP1022) using a linear gradient (0-20% B over 40 min) of aqueous 0.25% (v/v) acetic acid adjusted to pH 4.0 with NaOH (A buffer) and HPLC grade methanol (B buffer). The separation was monitored at 385 nm absorbance for bimane derivatives. The fractions containing 8 were identified by analytical HPLC as described above. Residual sodium acetate buffer was removed by a second
preparative HPLC separation over the same column and gradient, except that buffer A was 0.1% aqueous trifluoroacetic acid. The purified 8 was evaporated to dryness in a speedvac and exchanged with 99.8% D₂O (Aldrich) twice prior to suspension in 0.5 ml 99.96% D₂O (Aldrich) for NMR, pH ~3.

The second 20 g *D. radiodurans* cell pellet was processed as above for 8, except that the final preparative HPLC used aqueous 10 mM ammonium formate (pH 6.3) for the A buffer. The amount of residual ammonium formate was decreased by repeated solubilization in water and evaporation in the SpeedVac (5 times) and with 99.8% D₂O (2 times). The sample was suspended in 0.5 ml of 99.96% D₂O for NMR.

**Hydrolysis and composition of 8.** After recording the NMR spectra for the pH 3 sample of 8, an aliquot was used for acid hydrolysis. In a screw top 1.5 ml polypropylene microcentrifuge tube 100 µl of 2.7 mM 8 in D₂O was mixed with 500 µl of 6 N HCl and incubated at 65 °C. Samples (80 µl) were taken at intervals and mixed with 240 µl of cold (4 °C) 1 M NaOH to quench the hydrolysis and the samples were stored at -70°C for analysis.

For analysis of bimane labeled thiols the quenched samples were diluted 5-fold into 10 mM aqueous methanesulfonic acid and analyzed by analytical HPLC as described above. For analysis of glucosamine and other amines the AccQ-Fluor (Cat. # WAT052880, Waters Corp.) fluorescent labeling kit was used with slight modification of the manufacturer’s instructions. A 20 µl aliquot of quenched sample was neutralized with 5 µl 1 M NaOH and clarified by centrifugation at 13000 x g for 5 min. A 7.5 µl aliquot was mixed with 39.4 µl of borate buffer (Waters Corp.) and reacted with 15.6 µl of AccQ-Fluor reagent for 15 min at 60 °C. The labeled sample was diluted with 188 µl water for HPLC analysis. The amine content was analyzed by HPLC and fluorescent detection of AccQ-labeled amines as previously described.

For analysis of bimane labeled thiols the quenched samples were diluted 5-fold into 10 mM aqueous methanesulfonic acid and analyzed by analytical HPLC as described above. For analysis of bimane labeled thiols the quenched samples were diluted 5-fold into 10 mM aqueous methanesulfonic acid and analyzed by analytical HPLC as described above. For analysis of glucosamine and other amines the AccQ-Fluor (Cat. # WAT052880, Waters Corp.) fluorescent labeling kit was used with slight modification of the manufacturer’s instructions. A 20 µl aliquot of quenched sample was neutralized with 5 µl 1 M NaOH and clarified by centrifugation at 13000 x g for 5 min. A 7.5 µl aliquot was mixed with 39.4 µl of borate buffer (Waters Corp.) and reacted with 15.6 µl of AccQ-Fluor reagent for 15 min at 60 °C. The labeled sample was diluted with 188 µl water for HPLC analysis. The amine content was analyzed by HPLC and fluorescent detection of AccQ-labeled amines as previously described.

Glucosamine was the only amine compound analyzed that increased with time of hydrolysis (Fig. 1a).

The analysis of L-malate (Sigma-Aldrich) was conducted with a K-LMALR kit (Megazyme, Ireland) based on the stereospecific oxidation of L-malate to oxaloacetate with reduction of NAD⁺ by malate dehydrogenase. Due to the residual NaCl from neutralization of the acid hydrolysis sample (0.17 M) a calibration curve for L-malate (5-50 nmoles) in 0.2 M NaCl was found to be linear and identical to that determined in the absence of NaCl. The enzymatic assay was specific for L-malate and <1% of the 340 nm change was observed for D-malate (Sigma-Aldrich). A 160 µl sample of the quenched hydrolysis was neutralized with 50 µl 1 M NaOH and mixed with 10 µl of 1 M glycylglycine buffer pH 10 (Megazyme) and reduced to dryness in a SpeedVac. The sample was reconstituted with 50 µl of water (~60 µl total) and centrifuged for 5 min at 13000 x g to remove particulate matter. A 50 µl aliquot was diluted into a 1 ml assay solution according to the manufacturers instructions and the change in $A_{340}$ with addition of malate dehydrogenase was recorded. L-Malate was released from 8 in a time dependent manner (Fig 1a) with 83% of the initial 8 recovered as L-malate in the 18 hr hydrolysis sample. The recovery of the total bimane labeled products was 52% at 18.3 hr. The cysteine-bimane derivative (CySmB) is known to hydrolyze to a stable, non-fluorescent cysteine-derivative which elutes near norleucine on amino acid analysis, hence complete recovery of CySmB was not expected.

The analysis of L-malate (Sigma-Aldrich) was conducted with a K-LMALR kit (Megazyme, Ireland) based on the stereospecific oxidation of L-malate to oxaloacetate with reduction of NAD⁺ by malate dehydrogenase. Due to the residual NaCl from neutralization of the acid hydrolysis sample (0.17 M) a calibration curve for L-malate (5-50 nmoles) in 0.2 M NaCl was found to be linear and identical to that determined in the absence of NaCl. The enzymatic assay was specific for L-malate and <1% of the 340 nm change was observed for D-malate (Sigma-Aldrich). A 160 µl sample of the quenched hydrolysis was neutralized with 50 µl 1 M NaOH and mixed with 10 µl of 1 M glycylglycine buffer pH 10 (Megazyme) and reduced to dryness in a SpeedVac. The sample was reconstituted with 50 µl of water (~60 µl total) and centrifuged for 5 min at 13000 x g to remove particulate matter. A 50 µl aliquot was diluted into a 1 ml assay solution according to the manufacturers instructions and the change in $A_{340}$ with addition of malate dehydrogenase was recorded. L-Malate was released from 8 in a time dependent manner (Fig 1a) with 83% of the initial 8 recovered as L-malate in the 18 hr hydrolysis sample. The recovery of the total bimane labeled products was 52% at 18.3 hr. The cysteine-bimane derivative (CySmB) is known to hydrolyze to a stable, non-fluorescent cysteine-derivative which elutes near norleucine on amino acid analysis, hence complete recovery of CySmB was not expected.
The acid hydrolysis of 8 showed two transient species eluting at 8 and 9 minutes, U8mB and U9mB, which were maximally produced at 8 hours and declined thereafter with increasing production of CySmB and glucosamine (Fig. 1a, panel 2). U8mB and U9mB were isolated from the 4-hour hydrolysis sample (Fig. 1b, panel 1) by analytical HPLC using Method A (above), except that buffer A was 0.1% aqueous trifluoroacetic acid (TFA). The acidic pH increased the retention time of U8mB, U9mB, and 8 by 2-3 minutes improving resolution (Fig S1a). The protonated CySmB eluted at 30 minutes in high methanol content eluent at the end of the gradient (not shown). The isolated U8mB peak was concentrated and reinjected on the same protocol and chromatographed as two peaks, coeluting with U8mB and U9mB (Fig. S1b, panel 2). When the U9mB peak was isolated, concentrated and reinjected the same pattern was observed (Fig. S1c). When the initial substrate 8 was similarly isolated and reinjected a single peak coeluting with 8 was observed (Fig. S1d). The intermediates U8 and U9 appear to be rapidly interconverting species consistent with a Cys-bimane conjugate linked to glucosamine present in the α and β anomeric forms. This data also suggests that the anomeric hydroxyl of glucosamine has been exposed by acid hydrolysis of 8.

Chiral chromatography of CySmB from 8 (BSmB) hydrolysis. To determine the absolute stereochemistry of the CySmB moiety from 8 we hydrolyzed a sample of 63 nmol of 8 for 8 hours as described above. The CySmB peak sample was purified by analytical HPLC (see Fig. 1b, panel 1) using fluorescence and mass detection (ESP) with single ion resonance (SIR) at 312 Da. Authentic standards of L-cysteine (Sigma) and D-cysteine (Sigma) were derivatized with monobromobimane as previously described and purified using the same analytical HPLC protocol. The D-CySmB and L-CySmB standards were separated on a Chirobiotic T (4.6 x 250 mm, 5 µm, Supelco #12024AST) teichoplanin containing chiral HPLC column operated at 23 °C and 1 ml per min isocratic 60% methanol-water, essentially as described for amino acid derivatives. The L-CySmB (Fig. S2a) and D-CySmB (Fig. S2b) standards were detected using bimane fluorescence and eluted at 12.8 and 17.9 min, respectively. CySmB purified from the hydrolysis of 8 (BSmB) coelutes with L-CySmB and not with D-CySmB (Fig. S2c).

Synthesis of 2-[(2R)-2-(amino)-1-oxo-3-[[2,5,6-trimethyl-1,7-dioxo-1H,7H-pyrazolo[1,2-a]pyrazol-3-yl)methyl]thio]propyl]amino]-2-deoxy-D-glucose (11, L-CySmB-D-GlcN) and 2-[(2S)-2-(amino)-1-oxo-3-[[2,5,6-trimethyl-1,7-dioxo-1H,7H-pyrazolo[1,2-a]pyrazol-3-yl)methyl]thio]propyl]amino]-2-deoxy-D-glucose (12, D-CySmB-D-GlcN). Both L-CySmB-D-GlcN and D-CySmB-D-GlcN were prepared using the same method as outlined below for L-CySmB-D-GlcN. All solvents and reagents were purchased from Aldrich except where specified. Fmoc-L-Cys-S-trityl-polystyrene (0.35 mmol) was pre-swollen by agitation in N,N-dimethylformamide (DMF) (10 ml) for 10 min and then filtered. A freshly prepared milky suspension of D-glucosamine (250 mg, 1.4 mmol) in DMF (3 ml) was then added and whilst agitating with nitrogen gas a solution of 1-hydroxybenzotriazole hydrate (HOBT, Alfa-Aesar) (189 mg, 1.4 mmol), (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBOP 729 mg, 1.4 mmol, Novabiochem) and 2,6-di-tert-butylypyrididine (314 µl, 1.4 mmol) was added and the resulting mixture was agitated for 4 hours by which time the slurry had turned into a dark yellow/brown solution. The solution was filtered, washed successively with 10 ml portions of DMF (4 × 10 min) before repeating the above coupling procedure two more times. The washed resin was then agitated in 20% (v/v) piperidine in DMF (10 ml) for 30 minutes. The
solution was filtered and the resin was washed successively with 10 ml DMF (4 \times 10 \text{ min}) followed by 10 ml dichloromethane (4 \times 10 \text{ min}, \text{CH}_2\text{Cl}_2). The product was then cleaved from the resin by 3 \times 3 \text{ min} treatments with 3 ml portions of TFA:CH_2Cl_2:triethylsilane (40:60:3) and the resin finally washed with 10 ml CH_2Cl_2. The combined filtrates were concentrated on a rotary evaporator and washed by repeated addition and deactivation of ice cold ether (3 \times 6 \text{ ml}) to give a yellow/brown oil, which was dissolved in water (10 ml) and freeze dried overnight to give L-Cys-d-GlcN as a yellow gel (71 mg). A portion (28 mg) was dissolved in HPLC grade water (4 ml) and mixed with a solution of mBBr (37 mg, 0.136 mmol) in acetonitrile (2 ml). The mixture was added to a solution of NaHCO_3 (20 mg, 0.238 mmol) in HPLC grade water (1 ml) and stirred for 45 min at room temperature in the dark. The reaction was then quenched with 10 ml water and washed with dichloromethane (3 \times 20 ml) and the aqueous layer was freeze dried to give crude yellow-green solid. This was purified by HPLC on a Phenomenex Luna C_{18} 10-\mu m (250 mm x 10 mm) column at 35 °C and a flow rate of 4.8 ml min^{-1} using the following elution conditions: 0-6 min, 10% methanol in buffer A (10 mM ammonium formate adjusted to pH 4.0 with acetic acid) followed by a linear gradient to 20% methanol from 6-18 min. Product containing fractions were pooled, concentrated and freeze dried to give 26 mg as a yellow-green solid. Analytical HPLC showed the compound not to be of the desired purity. So the compound was re-purified (2nd time) as above to afford L-CySmB-D-GlcN as a yellow-green solid (7 mg, 15% yield). The \alpha/\beta anomers of L-CySmB-D-GlcN eluted at 7.9 and 9.0 min under these conditions. D-CySmB-D-GlcN was prepared in a similar manner (3 mg, 7% yield) and its \alpha/\beta anomers eluted at 8.3 and 11.0 min under the same HPLC elution conditions. As observed for U8mB and U9mB, when either of these individual anomer peaks were collected and reanalysed by HPLC they were shown to represent a rapidly equilibrating mixture of \alpha/\beta anomers.

**NMR analysis of 8.** Unless otherwise stated, \textsuperscript{1}H- and \textsuperscript{13}C-NMR were recorded on a Bruker DMX500 or Jeol ECA500 at 500 MHz for \textsuperscript{1}H and 125 MHz for \textsuperscript{13}C spectra. Chemical shifts are referenced to the residual signal for D_2O at \delta 4.80 ppm (\textsuperscript{1}H) and NH_4HCO_3 at \delta 8.453 ppm (\textsuperscript{1}H) and \delta 171.2 ppm (\textsuperscript{13}C). All spectra were collected at 23 °C.

Hydrolysis studies showed that the composition of 8 includes L-cysteine, D-glucosamine, L-malate and the bimane thiol label. A combination of 1D and 2D NMR studies confirmed the composition of 8 and established connectivity of these moieties using \textsuperscript{1}H, \textsuperscript{13}C, gCOSY, TOCSY, NOESY, ROESY, HSQC, and HMBC experiments. As noted in the appended spectra, the composition of 8 was confirmed by \textsuperscript{1}H-\textsuperscript{1}H gradient homonuclear correlated spectroscopy (gCOSY, Figs. S3c and S3d). One bond \textsuperscript{1}H-\textsuperscript{13}C correlations were determined using a heteronuclear single quantum correlation (HSQC, Fig. S3h) experiment. These experiments verified the presence of cysteine, bimane, glucosamine, and malate. The cysteine \alpha carbon proton (2"") had a chemical shift of 3.65 in acidic D_2O (pH−3, Fig. S3k) and a chemical shift of 4.09 ppm in ammonium formate (pH 6.3, Table S1, Fig. S3a), both significantly upfield from that found for the cysteine \alpha-carbon proton (4.5-4.7 ppm) adjacent to an amide linkage found in the bimane derivative of mycothiol\textsuperscript{17,18}. This upfield shift was observed in desacetylmycothiol\textsuperscript{19} for the cysteine \alpha-carbon proton adjacent to a free amine, consistent with a free cysteine \alpha-amino group in 8. Long-range \textsuperscript{1}H-\textsuperscript{13}C correlations were examined using a heteronuclear multiple bond correlation (HMBC, Figs. S3i and S3j) experiment and a 1"" to 2"" correlation established the cysteine-glucosamine amide bond, similar to that found in mycothiol\textsuperscript{17,18}. A weak HMBC correlation between 1' and 2''' (Fig. S3i) suggests a glycosidic linkage between the anomeric
carbon of glucosamine and the 2-hydroxyl of malate. This interaction is more clearly demonstrated in the ROESY spectra.

Spatial $^1$H-1H spin couplings were examined by rotating-frame overhauser enhancement spectroscopy (ROESY) and established a 1’ to 2’’ glycosidic bond from the anomeric proton of glucosamine to the 2’’’ proton of malate (Figs. S3e and S3g). The ROESY spectrum was key to determine the linkage between the L-malic acid and glucosamine moieties. As illustrated in Fig. S3g, three cross peaks were observed at 4.13–4.18 ppm as given by correlations between protons at 2’’’ and 3a’’, 2’’ and 3’, and 2’’’ and 1’. The position of these peaks was apparent due to the presence of two maxima for the 2’’’ proton and one for the 2’’ proton with the 2’’’ downfield of the 2’’. The observation of an NOE between the anomeric proton 1’ of the glucosamine and the 2’’ of L-malic acid supports a glycosidic linkage with the C2 hydroxyl of L-malic acid. Consistent with this linkage are the nearly identical $^{13}$C chemical shifts of the carboxylates of malate (1’’’, 4’’’, Table S1). The anomeric carbon–proton coupling constant for glucosamine (1’) of ($J = 3.5$ Hz, Table S1) is consistent with an α-glycosidic linkage between glucosamine and malate similar to mycothiol$^{17,18}$ and not the β-glycosidic linkage ($J = 7.9$ Hz) in the malate containing glucoside from Synadenium pereskiifolium$^{20}$.

**NMR analysis of 11 and 12.** The $^1$H-NMR spectra for 11 and 12 were recorded at 500 MHz in D$_2$O containing NH$_4$HCO$_3$ at pH 6.4 as shown in Fig. S4a and Fig. S4b, respectively. The spectra indicated the presence of both α- and β-anomers (Fig. 1b), which are largely separable on the timescale of the HPLC experiment but rapidly re-equilibrate when collected (Fig. S1b,c). The interconversion of the anomers is slow on the NMR timescale so that the observed spectrum is a composite of that for each α- and β-anomer. Assignment of the NMR spectra for 11 and 12 was possible after collection of gCOSY spectra as noted in Table S5.

**High Resolution Mass Spectral Analyses.** A Thermo Scientific LTQ Orbitrap XL mass spectrometer was used for high resolution electrospray ionization mass spectral analysis. A mass of 611.1648 Da was determined for 8 and corresponded to a formula for the positive ion of C$_{23}$H$_{32}$O$_{12}$N$_4$SNa (theoretical mass 611.1630 Da). The mass for 11 was 495.1521 (predicted composition C$_{19}$H$_{28}$N$_4$O$_8$SNa, theoretical mass 495.1520) and for 12 was 495.1519 (predicted composition C$_{19}$H$_{28}$N$_4$O$_8$SNa, theoretical mass 495.1520).

**Analysis of the bimane derivative of N-acetylbacillithiol.** The possibility that (13) is an intermediate in the biosynthesis of the N-acetylated form, analogous to MSH, was eliminated by preparation of this derivative and showing that it elutes on HPLC at a position where no bimane derivatives are found for either D. radioduran or B. subtilis. Bacillithiol-bimane BSH (0.1 mM) isolated from D. radiodurans was incubated at 23 °C with 10 equivalents of acetic anhydride (Fisher Biosciences) in 10 mM ammonium bicarbonate as described previously for acetylation of GlcN-Ins$^{21}$. The resulting N-acetyl-bacillithiol-bimane eluted 15 min later than bacillithiol-bimane using Method A HPLC (above). A D. radiodurans culture was analyzed at 8 time points from early exponential to late stationary phase culture (24 independent samples) and no N-acetylbacillithiol was found (<0.02 nmoles per mg dry weight). A B. subtilis JH642 culture was analyzed at 11 times from early exponential to late stationary phase (33 independent samples) for N-acetyl-bacillithiol and none was found (<0.01 nmoles per mg dry weight). Thus, N-acetylated
bacillithiol, the bacillithiol derivative analogous to mycothiol in structure, is not found at any significant level during the growth of *D. radiodurans* or *B. subtilis*.

References

1. Smirnova, G.V., Muzyka, N.G., Glukhovchenko, M.N. & Oktyabrsky, O.N. Free Radic. Biol. Med. 28, 1009-1016 (2000).
2. Ung, K.S. & Av-Gay, Y. FEBS Lett. 580, 2712-2716 (2006).
3. Buchmeier, N.A., Newton, G.L. & Fahey, R.C. J. Bacteriol. 188, 6245-6252 (2006).
4. Hoch, J.A. & Coukoulis, H.J. J. Bacteriol. 133, 265-269 (1978).
5. Ogawa, K. et al. J. Bacteriol. 177, 1409-1413 (1995).
6. Antelmann, H., Engelmann, S., Schmid, R. & Hecker, M. J. Bacteriol. 178, 6571-6578 (1996).
7. Altschul, S.F. et al. Nucleic Acids Res. 25, 3389-3402 (1997).
8. Newton, G.L., Ta, P. & Fahey, R.C. J. Bacteriol. 187, 7309-7316 (2005).
9. Koledin, T., Newton, G.L. & Fahey, R.C. Arch. Microbiol. 178, 331-337 (2002).
10. Nicely, N.I. et al. Biochemistry 46, 3234-3245 (2007).
11. Fahey, R.C. & Newton, G.L. Methods Enzymol. 143, 85-96 (1987).
12. Fenton, S.S. & Fahey, R.C. Anal. Biochem. 154, 34-42 (1986).
13. Anderberg, S.J., Newton, G.L. & Fahey, R.C. J. Biol. Chem. 273, 30391-30397 (1998).
14. Newton, G.L., Fahey, R.C., Cohen, G. & Aharonowitz, Y. J. Bacteriol. 175, 2734-2742 (1993).
15. Peter, A. et al. J. Chromatogr. A 871, 105-113 (2000).
16. Stewart, M.J., Jothivasan, V.K., Rowan, A.S., Wagg, J. & Hamilton, C.J. Org. Biomol. Chem. 6, 385-390 (2008).
17. Newton, G.L. et al. Eur. J. Biochem. 230, 821-825 (1995).
18. Spies, H.S. & Steenkamp, D.J. Eur. J. Biochem. 224, 203-213 (1994).
19. Bornemann, C., Jardine, M.A., Spies, H.S.C. & Steenkamp, D.J. Biochem. J. 325, 623-629 (1997).
20. Hermansson, K., Kenne, L., Rukunga, G.M. & Samuelsson, G. Phytochem. 29, 513-515 (1990).
21. Newton, G.L., Av-Gay, Y. & Fahey, R.C. J. Bacteriol. 182, 6958-6963 (2000).