nature research

Corresponding author(s): Emily P. Balskus and Amie K. Boal

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Geneious Pro 467 Version 7.1.6 (Biomatters) was used for pBLAST searches and sequence alignments. MEGA-X was used to construct phylogenetic tree. MODELLER (Max Planck Institute for Developmental Biology) was used for building homology models. ProDRG program was used to build the HMA ligand Data collection software for x-ray cystallography was Blu-Ice Genome sequencing and annotations were performed by Era7 Bioinformatics Structural homologs of SznE were identified using HHPred (HHSuite v. 3) The closest structural homologs' sequences were aligned with SznE using clustalo (version 1.2.4).
Data analysis	Pymol Molecular Graphics system, Version 1.8 (Shrödinger, LLC) was used to view PDB files and for generating structural figures Geneious Pro 467 Version 7.1.6 (Biomatters) and MEGA-X were used to view sequence alignments NMR spectra were visualized and analyzed using MestReNova, version 10.0.0-14411 Data analysis XDS, HKL2000, ShelX, HKL2MAP, ARP/wARP, PHASER, Refmac5, COOT, TLSMD server for crystallographic characterization of SznF Microsoft Excel was used to construct bar graphs, Supplementary tables, and pie charts Adobe Illustrator CC2015 was used to construct figures Extracted ion chromatograms were generated with Agilent ChemStation software Mauve was used for comparative genomics analysis O2 consumption assay was analyzed using NeoFox Viewer version 2.40.

Genome neighborhood analysis of SznF homologs were performed with the software tools available at Integrated Microbial Genomes and Microbiome system (Joint Genomic Institute)
Extinction coefficients for determining protein concentrations were performed by protparam

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The nucleotide sequences for the szn biosynthetic gene cluster and individual genes will be deposited into NCBI (accession numbers: TBD). The structure factor and coordinate of SznF have been deposited in the Protein Data Bank (PDB code for HMA-SznF: 6M9R and PDB code for SeMet-SznF: 6M9S). Additional data that support the conclusions of the paper can be requested from the corresponding authors.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculations were completed, since the report explores biosynthetic functions of two enzymes. For analysis of metabolite production by various mutants, at least triplicates performed on different days were performed.
Data exclusions	No data were excluded from analyses
Replication	The number of times an experiment was repeated is detailed in the Statistics and Reproducibility section of the manuscript. All attempted replications yielded similar results. Measures taken to ensure reproducibility include performing biochemical assays on different days using different enzyme preparations and picking different monoclonal colonies for in vivo knockout analysis.
Randomization	Randomization is irrelevant since our study does not involve group allocation.
Blinding	Blinding is irrelevant since our study does not involve group allocation

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging