Supplementary Data

Functional diversity of HemN-like Proteins

Jinduo Cheng#, Wan-Qiu Liu#, Xiaoyu Zhu, Qi Zhang*
Department of Chemistry, Fudan University, Shanghai, 200433, China
#These authors contributed equally to this work.
*Email: qizhang@sioc.ac.cn
**Supplementary Data.** Protein sequences used for phylogenetic analysis in Figure 9 in the main text, showing the accession numbers, the sources, and the identities and coverages in BlastP analysis using *E. coli* HemN as a query.

| Accession No. | Source                  | Ident. | Cover.  |
|---------------|-------------------------|--------|---------|
| HemN          |                         |        |         |
| AAN83245      | *Escherichia coli*      | 100%   | 100%    |
| WP_028069898  | *Sphingobacterium thalpophilum* | 97%     | 38.33%  |
| WP_006365438  | *Chlorobium ferrooxidans* | 94%    | 39.00%  |
| WP_069810592  | *Chlorobaculum limnaeum* | 94%    | 37.87%  |
| WP_026853902  | *Geothrix fermentans*   | 97%    | 42.67%  |
| WP_005034428  | *Holophaga foetida*     | 97%    | 41.91%  |
| AMW06315      | *Gemmatimonas phototrophica* | 100%   | 38.15%  |
| BAH40890      | *Gemmatimonas aurantiaca* | 96%    | 39.46%  |
| AFS53101      | *Leptospirillum ferrarilium* | 97%    | 45.68%  |
| CUW40578      | *Magnetospirillum sp.*  | 97%    | 37.56%  |
| WP_010883524  | *Chlamydia pneumonias*  | 97%    | 37.64%  |
| WP_021756895  | *Chlamydia pecorum*     | 98%    | 38.77%  |
| CAH07682      | *Bacteroides fragilis*  | 94%    | 39.46%  |
| ACB73891      | *Opitutus terrea*       | 98%    | 43.17%  |
| ACM93156      | *Nautilia profundicola* | 100%   | 43.57%  |
| KIN92132      | *Thauera sp. SWB20*     | 98%    | 48.79%  |
| CUI82205      | *Achromobacter sp. ATCC13047* | 100% | 92.56%  |
| WP_025305675  | *Thermocrinis ruber*    | 97%    | 47.23%  |
| WP_012674071  | *Sulfurihydrogenibium azorense* | 99% | 48.15%  |
| EHC13895      | *Fischerella sp.*       | 99%    | 49.67%  |
| WP_036489441  | *Myxosarcina sp.*      | 96%    | 46.67%  |
| AFZ05950      | *Oscillatoria nigro-viridis* | 96% | 50.22%  |
| P74132        | *Synechocystis*         | 98%    | 49.02%  |
| ChuW          |                         |        |         |
| AIU70349      | *Thermococcus eurythermalis* | 61% | 27.87%  |
| WP_051366200  | *Pseudoalterometobula elfii* | 54% | 26.80%  |
| CDC62198      | *Clostridium environmental samples* | 40% | 30.81%  |
| WP_020612921  | *Sediminispirochaeta bajacaliforniensis* | 53% | 28.68%  |
| EEF14474      | *Campylobacter rectus*  | 40%    | 29.63%  |
| WP_013010134  | *Dentirovibrio acetiphilus* | 56% | 27.41%  |
| WP_009353669  | *Veillonella sp.*       | 53%    | 26.88%  |
| WP_051212161  | *Veillonella montpelleriensis* | 84% | 25.18%  |
| EKX99441      | *Selenomonas sp.*       | 85%    | 27.61%  |
| WP_019177805  | *Methanomassiliicoccus luminysensis* | 64% | 26.64%  |
| WP_013506625  | *Desulfurispirillum indicum* | 87% | 24.08%  |
| EGW43871      | *Bilophila sp.*         | 85%    | 23.40%  |
| ChuW          |                         |        |         |
| EIO71851      | *Escherichia coli*      | 66%    | 24.84%  |
| Accession | Organism Name                  | Coverage | Identity  |
|-----------|--------------------------------|----------|-----------|
| WP_004344082 | Thauera linaloolentis            | 53%      | 28.40%    |
| WP_008619339 | Magnetospirillum caucaseum       | 36%      | 29.34%    |
| HutW      | Vibrio alginolyticus            | 36%      | 34.52%    |

**Class C RSM**

| Accession | Organism Name                  | Coverage | Identity  |
|-----------|--------------------------------|----------|-----------|
| AIC83437 | Thauera linaloolentis            | 53%      | 28.40%    |
| WP_008619339 | Magnetospirillum caucaseum       | 36%      | 29.34%    |
| Zbam-orf26 | Streptomyces flavoviridis       | 64%      | 28.43%    |
| Blm-orf8  | Magnetospirillum caucaseum       | 36%      | 29.52%    |
| Tlm-orf1  | Streptomyces flavoviridis       | 64%      | 28.38%    |
| EGG43722 | Streptomyces griseoaurantiacu   | 54%      | 32.44%    |
| NosN      | Nocardia                        | 64%      | 28.71%    |
| PbtM3     | Planobispora rosea              | 43%      | 32.06%    |
| TpdU      | Nonomuraea                      | 42%      | 31.19%    |
| TbtI      | Thermobispora bispora          | 65%      | 28.99%    |
| TpdL      | Nonomuraea sp. Bp3714-39       | 42%      | 29.26%    |
| Tpdx2     | Nonomuraea pusilla              | 82%      | 26.21%    |
| Jaw5      | Streptomyces roseoverticillatus | 90%      | 25.52%    |
| SEF54573  | Streptomyces yanglinensis       | 90%      | 26.82%    |
| GAD86125  | Nocardia asteroides             | 91%      | 27.56%    |
| WP_026404953 | Planobispora rosea            | 91%      | 25.47%    |
| WP_007286255 | Streptomyces sp. TP-A0356    | 87%      | 28.05%    |
| WP_030684647 | Streptomyces sp. NRRL B-1347  | 85%      | 28.00%    |
| C10P      | Streptomyces sp. NRRL B-1347    | 85%      | 28.00%    |
| YtkT      | Streptomyces sp. TP-A0356       | 87%      | 28.05%    |
| WP_055490826 | Streptomyces sp. TP-A0356    | 84%      | 26.39%    |

**HemZ**

| Accession | Organism Name                  | Coverage | Identity  |
|-----------|--------------------------------|----------|-----------|
| WP_025117272 | Lysinibacillus fusiformis    | 40%      | 28.64%    |
| Q796V8    | Bacillus subtilis              | 50%      | 23.95%    |
| CON04352  | Streptococcus pneumoniae       | 50%      | 23.95%    |
| WP_010233508 | Clostridium arbusi           | 40%      | 29.17%    |
| WP_007286255 | Intestinibacter bartlettii    | 68%      | 27.19%    |
| WP_039679463 | Terrisporobacter othiienisis | 42%      | 31.68%    |

**Archaea**

| Accession | Organism Name                  | Coverage | Identity  |
|-----------|--------------------------------|----------|-----------|
| AFK50646  | Thermogladius cellulolyticus   | 66%      | 25.23%    |
| ESQ25637  | uncultured Acidilobus          | 42%      | 27.41%    |
| AFZ70518  | Caldisiaera lagunensis          | 85%      | 23.17%    |
| CCC80926  | Thermoproteus tenax            | 39%      | 30.81%    |
| Accession | Species                      | Identity | Similarity | E Value |
|-----------|------------------------------|----------|------------|---------|
| ABM80220  | *Hyperthermus butylicus*     | 65%      | 28.48%     |         |
| WP_055408506 | *Pyrodictium delaneyi*     | 44%      | 30.77%     |         |
| AEM37888   | *Pyrolobus fumarii*         | 68%      | 26.35%     |         |
| WP_011998375 | *Ignicoccus hospitalis*     | 65%      | 23.86%     |         |
| WP_014025565 | *Pyrolobus fumarii*         | 68%      | 26.35%     |         |
| MenK       | *Gordonibacter pamelaeae 7-10-1-b* | 49%     | 27.71%     |         |
| WP_011138079 | *Wolinella succinogenes*    | 66%      | 25.16%     |         |
| WP_028766718 | *Shewanella fidelis*       | 64%      | 26.25%     |         |
| EGG57223   | *Parasutterella excrementihominis* YIT 11859 | 66%     | 25.89%     |         |
| MenK2      | *Gordonibacter pamelaeae 7-10-1-b* | 49%     | 27.71%     |         |
| XP_011398951 | *Auxenochlorella protothecoides* | 71%     | 24.93%     |         |
| XP_005649356 | *Coccomyxa subellipsoidea* | 52%      | 28.00%     |         |
| XP_001703480 | *Chlamydomonas reinhardtii* | 50%      | 27.35%     |         |
| XP_002946784 | *Volvox carteri f. nagariensis* | 52%      | 28.00%     |         |
| CAL50029   | *Ostreococcus tauri*        | 67%      | 27.52%     |         |
| WP_002507726 | *Micromonas comoda*        | 65%      | 24.78%     |         |
| EDQ75811   | *Physcomitrella patens*     | 65%      | 23.95%     |         |
| EFJ27703   | *Selaginella moellendorffii* | 65%      | 24.19%     |         |
| XP_011627754 | *Amborella trichopoda*     | 65%      | 25.57%     |         |
| ABI93924   | *Arabidopsis thaliana*      | 67%      | 27.27%     |         |
| XP_002262672 | *Vitis vinifera*           | 82%      | 24.94%     |         |
| XP_012487614 | *Gossypium raimondii*      | 50%      | 27.35%     |         |
| XP_003522422 | *Glycine max*              | 66%      | 23.81%     |         |
| XP_009616975 | *Nicotiana tomentosiformis* | 67%      | 26.96%     |         |
| XP_010250874 | *Nelumbo nucifera*        | 66%      | 27.44%     |         |
| BAF29564   | *Oryza sativa*             | 80%      | 22.11%     |         |
| XP_009415069 | *Musa acuminata malaccensis* | 66%     | 24.60%     |         |
| XP_010930027 | *Elaeis guineensis*       | 65%      | 26.58%     |         |
| Animal_Fungi |                           |          |            |         |
| EEC11362   | *Ixodes scapularis*        | 83%      | 22.02%     |         |
| ESA03321   | *Rhizophagus irregularis*  | 41%      | 26.80%     |         |
| CEP17932   | *Parasitella parasitica*   | 61%      | 25.52%     |         |
| CEG73035   | *Rhizopus microsporus*     | 63%      | 26.82%     |         |
| CDS13916   | *Lichtheimia ramosa*       | 77%      | 23.77%     |         |
| EDV19558   | *Trichoplax adhaerens*     | 59%      | 24.00%     |         |
| XP_012557630 | *Hydra vulgaris*          | 87%      | 24.52%     |         |
| HXP_001636936 | *Nematostella vectensis* | 51%      | 25.21%     |         |
| EFX89896   | *Daphnia pulex*            | 52%      | 22.76%     |         |
| ELT99882   | *Capitella teleta*        | 82%      | 23.00%     |         |
| XP_009064549 | *Lottia gigantea*        | 82%      | 21.03%     |         |
| XP_005090240 | *Aplysia californica*     | 83%      | 22.84%     |         |
| XP_012695163 | *Clupea harengus*       | 83%      | 23.00%     |         |
| Accession   | Species                  | Identity | Similarity |
|-------------|--------------------------|----------|------------|
| NP_001077026 | Danio rerio              | 52%      | 24.28%     |
| XP_010862851 | Esox lucius              | 40%      | 26.20%     |
| XP_007556327 | Poecilia formosa         | 88%      | 20.84%     |
| XP_008327612 | Cynoglossus semilaevis   | 92%      | 21.02%     |
| XP_005999660 | Latimeria chalumnae      | 82%      | 22.25%     |
| OCA21266    | Xenopus tropicalis       | 83%      | 23.80%     |
| XP_007441654 | Python bivittatus       | 85%      | 22.25%     |
| XP_008102717 | Anolis carolinensis      | 84%      | 23.02%     |
| XP_005297689 | Chrysemys picta bellii   | 52%      | 28.22%     |
| XP_009562321 | Cuculus canorus          | 52%      | 31.02%     |
| KFV63667    | Picoides pubescens       | 50%      | 29.66%     |
| XP_005144383 | Melopsittacus undulatus  | 50%      | 29.24%     |
| XP_007505827 | Monodelphis domestica    | 52%      | 26.83%     |
| XP_012405492 | Sarcophilus harrisii     | 52%      | 27.05%     |
| XP_012581510 | Condylobatida cristata   | 53%      | 26.53%     |
| XP_001013399 | Mus musculus             | 52%      | 26.34%     |
| XP_008516193 | Equus przewalskii        | 52%      | 28.22%     |
| XP_011228768 | Aliuroidea melanoleuca   | 52%      | 27.39%     |
| XP_004446612 | Dasytus novemcinctus     | 55%      | 26.15%     |
| XP_012518368 | Propithecus coquereli    | 52%      | 26.72%     |

**RSAD1**

HemW and other bacterial proteins

| Accession   | Species                                | Identity | Similarity |
|-------------|----------------------------------------|----------|------------|
| WP_009962340 | Verrucomicrobium spinosum              | 80%      | 24.61%     |
| EAQ77478    | Blastopirellula marina                 | 89%      | 25.00%     |
| KLU03830    | Rhodopirellula islandica               | 80%      | 26.48%     |
| ADB19145    | Pirellula staleyi                      | 89%      | 25.24%     |
| ACD05922    | Akkermansia muciniphila                | 84%      | 24.30%     |
| ALA58796    | Nitrospira moscoviensis                | 83%      | 27.53%     |
| CUW28423    | Streptomyces reticuli                  | 64%      | 27.10%     |
| KOX89034    | Thermus aquaticus                      | 41%      | 34.87%     |
| WP_052195205 | Deinococcus swuensis                  | 52%      | 26.86%     |
| ADI13526    | Truepera radiovicina                   | 36%      | 33.13%     |
| WP_062189862 | Anaerolinea thermolimosa              | 53%      | 31.84%     |
| EFH90847    | Ktedonobacter racemifer               | 82%      | 27.04%     |
| WP_014433607 | Caldilinea aerophila                  | 52%      | 33.33%     |
| ACM05800    | Thermomicrobium roseum                | 87%      | 26.78%     |
| WP_039744988 | Geobacter pickeringii                | 52%      | 29.92%     |
| ACU91423    | Desulfomicrobium baculatum            | 83%      | 28.57%     |
| EFV43164    | Bilophila wadsworthia                 | 65%      | 27.33%     |
| WP_015946007 | Desulfovibrio vulgaris              | 83%      | 26.49%     |
| WP_015898519 | Acidobacterium capsulatum            | 84%      | 25.62%     |
| ACX75702    | Fibrobacter succinogenes              | 53%      | 25.97%     |
| Accession  | Species                          | Identity | Coverage |
|------------|----------------------------------|----------|----------|
| WP_013008727 | *Deferribacter desulfuricans*    | 83%      | 25.13%   |
| EDM26343   | *Lentisphaera araneosa*         | 85%      | 23.54%   |
| ALD66217   | *Spiroplasma cantharicola*      | 40%      | 32.31%   |
| AFG37251   | *Spirochaeta africana*          | 89%      | 23.92%   |
| ACV39949   | *Leptotrichia buccalis*         | 65%      | 27.56%   |
| ADO82048   | *Ilyobacter polytropus*         | 65%      | 26.42%   |
| ERT66913   | *Cetobacterium somerae*         | 65%      | 26.37%   |
| ACK42388   | *Dictyoglomus turgidum*         | 84%      | 24.87%   |
| ACI19107   | *Dictyoglomus thermophilum*     | 84%      | 26.10%   |
| EGG81213   | *Lachnospiraceae bacterium*     | 65%      | 30.79%   |
| AGB18790   | *Thermoanaerobacterium*         | 66%      | 26.95%   |
| WP_036947005 | *Pseudobacteroides cellulosolvens* | 83%    | 25.19%   |
| YP_001255455 | *Clostridium botulinum*       | 66%      | 26.95%   |
| CAB61616   | *Bacillus subtilis*             | 82%      | 25.66%   |
| WP_010735215 | *Enterococcus munditii*       | 88%      | 24.88%   |
| WP_038057790 | *Thermodesulfobacterium*     | 64%      | 23.36%   |
| AII03372   | *Thermodesulfobacterium commune* | 83%   | 23.85%   |
| HemW       | *Enterobacteriaceae*            | 52%      | 32.79%   |