

Supplementary Information

The Thaumarchaeon *N. gargensis* carries functional *bioABD* genes and has a promiscuous *E. coli* Δ *bioH*-complementing esterase EstN1

Jennifer Chow¹, Dominik Danso¹, Manuel Ferrer² and Wolfgang R. Streit¹ *

¹ Microbiology and Biotechnology, University of Hamburg, 22609 Hamburg, Germany

² Institute of Catalysis, Consejo Superior de Investigaciones Científicas, 28049 Madrid, Spain

*Corresponding author.

Table S1: List of primers used in this study. Restriction sites are underlined.

Target and primer designation	Forward primer sequence (5'→3')	Reverse primer sequence (5'→3')
Ngar_c21820_for / _rev	GGGAATCCAGCACGAAGATTAG	CATCAAGACGAGAAATCACGAATC
Ngar_c24650_for / _rev	ACAGCGCTTCATATCATACTACC	ATAGCTGTCTGCCACTGACAATC
Ngar_c14400 (<i>estN1</i>)_for / _rev	TGTTCAAGACTACGGTGCAGAG	TTATAGTAGCGCCAGCGGATTC
Ngar_c30910 (<i>estN2</i>)_for / _rev	CTCTACGGGATAGAAGATATTAAC	TTGCCATTATAGCAAACACTAC
Ngar_c32780_for / _rev	TAAAGAACGCGATGAAGAATATG	TTCTTGGCACTCATGATATCTC
Ngar_c35080_for / _rev	AACAAAGAAGCCAAAGGAAGTC	CAAGAGCGTGATAATAGTTACCG
<i>estN1_NdeI_for / _HindIII_rev</i>	<u>CATATG</u> TATGTCAACGGCTATGC	<u>AAGCTT</u> CTGGTCACTTCTGACCAAG
<i>estN2_NdeI_for / _HindIII_rev</i>	<u>CATATG</u> CTTTTCTATCCTCGCCTTC	<u>AAGCTT</u> CCCAACAGGCTTTACATCCTTTGTC
<i>Ecoli_bioA_FRT_for / _rev</i>	ATGACAACGGACGATCTTGCCTTTGACCAAC GCCATATCTGGCACCCATAAATTAACCTCAC TAAAGGGCG	TTATTGGCAAAAAAATGTTTCATCCTGTACC GCGCGTTAACCGCTGCGGTAATACGACTC ACTATAGGGCTC
<i>Ecoli_bioB_FRT_for / _rev</i>	ATGGCTCACCGCCACGCTGGACATTGTCGC AAGTCACAGAATTATTTGAAATTAACCTCAC TAAAGGGCG	CCGCGTTGTAATATTCGTCGGTGTCCGGGG TCATCAGCGCTGTCAAGATAATACGACT CACTATAGGGCTC
<i>Ecoli_bioC_FRT_for / _rev</i>	CAAGCCATTGCAGCGGCATTTGGTCGGGCAG CCGCACACTATGAGCAACAAATTAACCTCAC TAAAGGGCG	AATCACTCCCAAAAAAGATGATACGTCAG AGGATATCGCCCTGCTGTTTAATACGACT CACTATAGGGCTC
<i>Ecoli_bioD_FRT_for / _rev</i>	GTTATTTTGTACCGGAACGGATACCGAAGT GGGAAAACGTGCGCCAGTAATTAACCTCA CTAAAGGGCG	CTACAACAAGGCAAGGTTTATGTACTTTCC GGTTGCCGATTTTCTGGATTAATACGACT CACTATAGGGCTC
<i>Ecoli_bioF_FRT_for / _rev</i>	CAGGAGAAAATCAACGCGGCGCTCGATGCG CGGCGTGCTGCCGATGCCCTAATTAACCTCA CTAAAGGGCG	GCAGACGGTCGATATCCTGCATTTTCATGCG CAGCGTTAGCGTTAAGCGCTAATACGACT CACTATAGGGCTC
<i>Ecoli_bioH_FRT_for / _rev</i>	ATGAATAACATCTGGTGGCAGACCAAAGGTC AGGGGAATGTTTCATCTTGTAAATTAACCTCAC TAAAGGGCG	CTACACCCTGCTTCAACGCCACCAGCAG GTGACAAAACCTCGGCCGATTAAATACGACT CACTATAGGGCTC
<i>Ecoli_bioA_NdeI_for / _HindIII_rev</i>	<u>CATATG</u> ACAACGGACGATCTTGCC	<u>AAGCTT</u> TTGGCAAAAAAATGTTTCATCC
<i>Ecoli_bioB_NdeI_for / _HindIII_rev</i>	<u>CATATG</u> GCTCACCGCCAC	<u>AAGCTT</u> TAATGCTGCCGCGTTGTAATATTC
<i>Ecoli_bioC_NdeI_for / _HindIII_rev</i>	<u>CATATG</u> GCAACGGTTAATAAACAAG	<u>AAGCTT</u> CTCAGAGCAATCACTCCC
<i>Ecoli_bioD_NdeI_for / _HindIII_rev</i>	<u>CATATG</u> AGTAAACGTTATTTTGTACC	<u>AAGCTT</u> CAACAAGGCAAGGTTTATG
<i>Ecoli_bioF_NdeI_for / _HindIII_rev</i>	<u>CATATG</u> AGCTGGCAGGAGAAAATCAAC	<u>AAGCTT</u> GTTGCCATGCAGCACCTCC
<i>Ecoli_bioH_NdeI_for / _HindIII_rev</i>	<u>CATATG</u> AATAACATCTGGTGGCAGAC	<u>AAGCTT</u> CACCCTCTGCTTCAAC
<i>Ngar_bioA_NdeI_for / _EcoRI_rev</i>	<u>CATATG</u> AACCTGAGAGATGC	<u>GAATTC</u> GATAGCTTCTCCACTTTTC
<i>Ngar_bioB_NdeI_for / _HindIII_rev</i>	<u>CATATG</u> ATGGCCGATTGAGC	<u>AAGCTT</u> TGCTTTGAGGCAATCTC
<i>Ngar_bioD_NdeI_for / _HindIII_rev</i>	<u>CATATG</u> CGCGGCATATTCGTAACAG	<u>AAGCTT</u> CACGGACAGTAGCAGGCTATC
<i>Ngar_bioF_NdeI_for / _EcoRI_rev</i>	<u>CATATG</u> GCCTCAAAGCATAAGAAG	<u>GAATTC</u> CATGATGCCTGTTTTTCTACC
<i>Ngar_bioC1_NdeI_for / _HindIII_rev</i> (Ngar_c05970)	<u>CATATG</u> TCTGCAGAGGATTCGACTG	<u>AAGCTT</u> TTGACGCCGCCGATAATC
<i>Ngar_bioC2_NdeI_for / _HindIII_rev</i> (Ngar_c21300)	<u>CATATG</u> CGCTGCCAGTTTACGAAATAC	<u>AAGCTT</u> TGGCTTTTACGAGATACTATGG
<i>Ngar_bioC3_NdeI_for / _HindIII_rev</i> (Ngar_c35810)	<u>CATATG</u> GAAAAGAGGCGTCAACAG	<u>AAGCTT</u> CCTTACGGCCCTGACAAC
<i>Ngar_bioC4_NdeI_for / _HindIII_rev</i> (Ngar_04400)	<u>CATATG</u> TCCGGTCTTGAAGCC	<u>AAGCTT</u> GTAAGCGGCGATTATCACCG

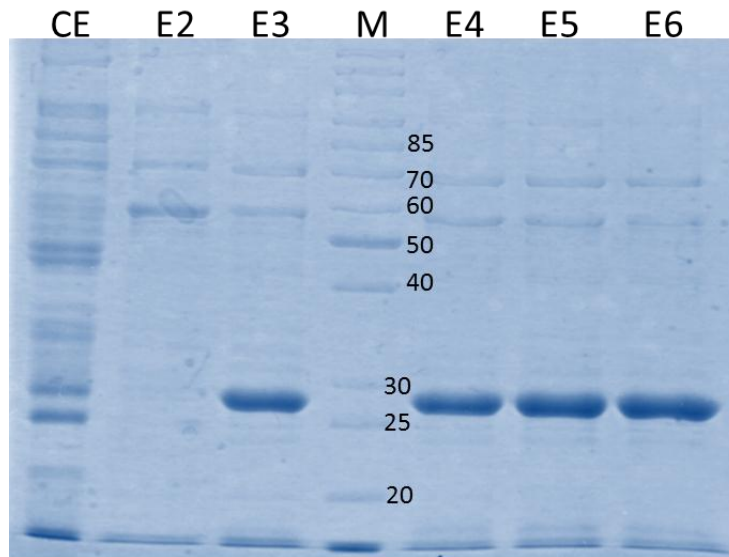


Fig. S1: SDS-PAGE with purified EstN1 visible as strong bands between 25 and 30 kDa. EstN1 was expressed from a pET21a vector in *E. coli* Rosetta-gami 2 (DE3). It was His₆-tag purified via Ni-NTA-agarose. The full-length 12% acrylamide gel shows cell extract (CE) as well as five different elution fractions (E2-E6). M; protein marker (unstained PageRuler #26614, Thermo Scientific) with protein sizes in kDa.

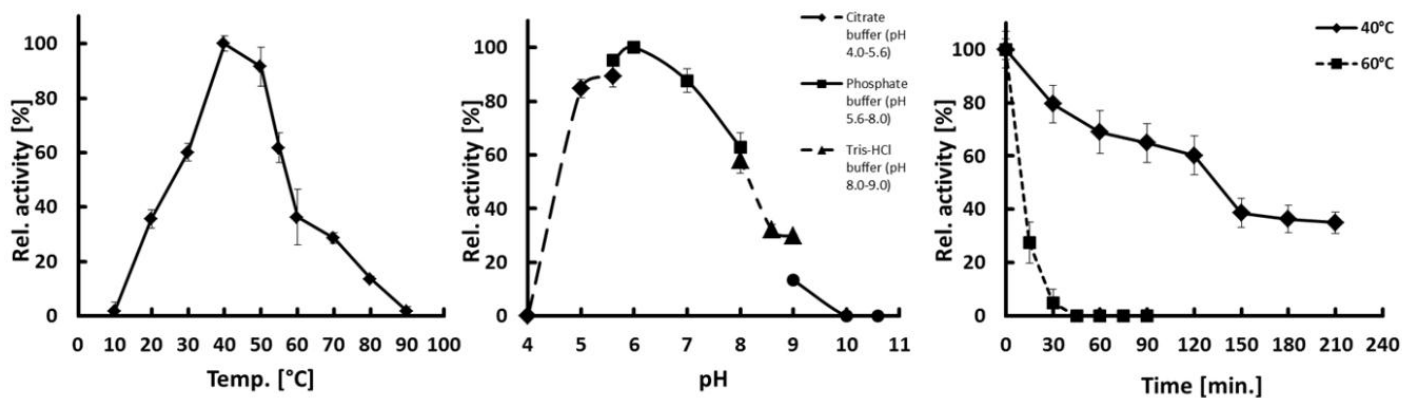


Fig. S2: Biochemical properties of EstN1 tested with *para*-nitrophenol hexanoate as ester substrate measured at E405 nm. EstN1 shows its highest activity at 40°C (left) and is active at a slightly acidic to neutral pH (middle). After 2 hours at 40°C, it still has 60% activity, while incubation at 60°C leads to activity loss after 30 minutes (right).

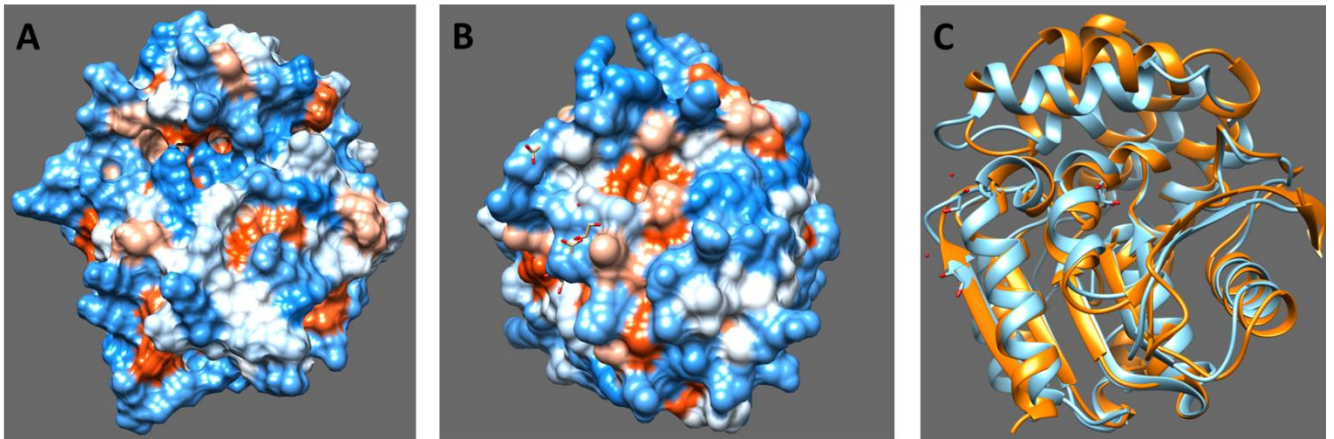


Fig. S3: Surface hydrophobicity of the homology modelled enzyme structure of *N. gargensis* EstN1 (A) and the crystal structure of *E. coli* BioH (B; PDB 1M33). Hydrophobic regions are shown in red. Superimposition of EstN1 (orange) and BioH (blue) is shown in C.

```

N.gar_EstN1_      1 mtyvngyatrylehgpp--dgktlillhgigasawrsrviptlskyfrvitpdivgfgysdkptveytm
N. everglade     1 mtkvnghttryldygsrvkagkdlvllhglgasserwllvaptlskyfrviavdvvgfgysdkptveytm
N. viennensi     1 mtkvnglatryldygsrvkagkdlvllhglgasserwllvaptlskyfrviavdvvgfgysdkptveytm
E.coli_BioH      1 ---mnniwwqtkgqg-----nvhlvllhgwglnaevwrcideelsshftlhlvdlpgfgrsrg----fga

N.gar_EstN1_     69 dffldffftgfldnldvskaiivvgssfgghlatefairhnrkvdklvlv--spagmmrtstptldgyimaa
N. everglade     71 dffidffdgflqnlgiiekphlvvgssfgghlaaeyairnkrridkmalv--spagamrtstpildqyimaa
N. viennensi     71 dfftdffdkflqnlgiiekphivvgssfgghlaaeyairnrkidklala--spagvmrtstpildqyimaa
E.coli_BioH      59 lsladmaeavlqqap-dkaiwlgwslgglvasqialthpervqalvtvasspcfsardewpgikpdvlag

N.gar_EstN1_    137 lyptyenayrafremahdpdavteeivmdfvnrnr-----lpnakyafmstllgmryapklqgrlgkiis
N. everglade    139 lyptfenalkafsdmahdpsivteemvvdvkrmn-----lpnakyafmstllgmryslplrglssvia
N. viennensi    139 lyptyenalkafsdmahdpsivdegvtvdfvrrmn-----lpnakyafmstllgmryspplrglssvia
E.coli_BioH    128 fqqqlsddfqrtrverflalqtmgtetarqdaralkktvialpmpevdvlnnggleilktdlrqplqnvsm

N.gar_EstN1_    202 ptllvwgdsdrmpvqyakeynei-pdselvvikncghtpyvekpmtfnkllkflvrdsq---
N. everglade    204 ptlimwgdedrmipvqyakdfrev-pnselvvikdcghtpyvekpmtfnriilkflagkeelvp
N. viennensi    204 ptlviwgdedrmipvqyakefrev-pnnelvvikdcghtpyiepktfnriivkflagkeelvp
E.coli_BioH    198 pfllrlygyldglvprkvvpmldklwphsesyifakaahapfishpaefchllvalkqr-----

```

Fig. S4: Global alignment of the amino acid sequences of *N. gargensis* EstN1 and its counterparts from *N. evergladensis* and *N. viennensis* together with *E. coli* BioH. The alignment is based on BLOSUM 62 scoring matrix. While the two archaeal proteins share a 71% identity with EstN1, BioH has only 22% identity to its functional analogue EstN1. The common GX SXGG-motif around the active site serine is framed. Asterisks mark the catalytic triad composed of Ser, His and Asp. Circles mark amino acids that EstN1 and BioH have in common compared to the enzymes from *N. evergladensis* and *N. viennensis*.

E. coli Rosetta-gami 2 (DE3) Δ *bioH*

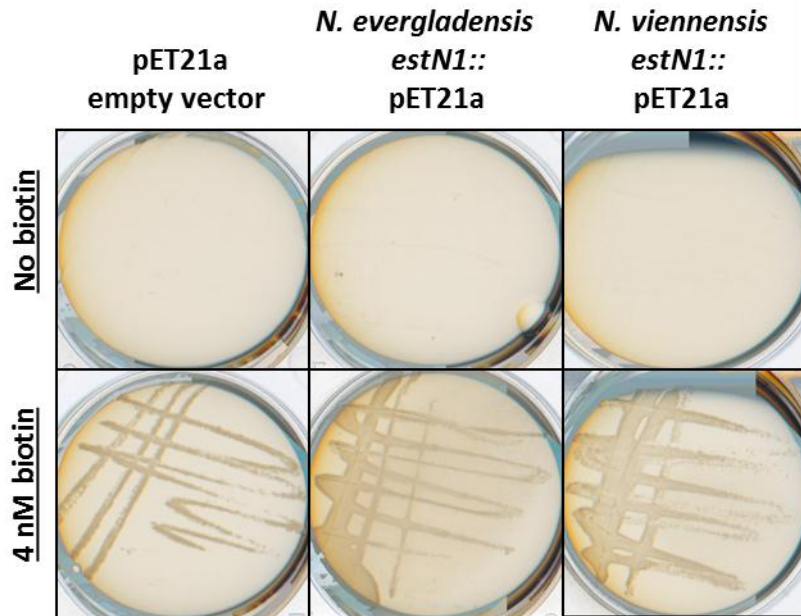


Fig. S5: Complementation of auxotrophic *E. coli* Δ *bioH* insertion mutants with possible *estN1* homologues from *N. evergladensis* and *N. viennensis*. Both enzymes were not able to complement for growth on biotin-free M9 medium.



Fig. S6: Hidden-Markov-Model (HMM) logo based on 49 sequences found by Uniprot database searches. Positions of the conserved amino acids are given above the letters. The model was visualized with skylign.org. The conserved GX SXGG motif is framed.