Structure and Activity Analyses of *E. coli* K-12 NagD Provide Insight Into the Evolution of Biochemical Function in the HAD Enzyme

Superfamily ^{†,‡}

Lee W. Tremblay[§], *Debra Dunaway-Mariano*^{\parallel^*}, *and Karen N. Allen*^{§*}

[§]Department of Physiology and Biophysics, Boston University School of Medicine,

Boston, Massachusetts 02118-2394

^{II}Department of Chemistry, University of New Mexico, Albuquerque, NM 87131-0001

SUPPORTING INFORMATION

1 Figure of sequence alignment of NagD homologues from various bacterial species. 2 pages total.

| E.coli_K-12 salomnella Yursinia Shweanella Streptomyces Nocardia Leifeonia Arthrobacter Enterococcus_faecalis Streptococcus_pneumoniae Staphylccoccus_aureus Thermotoga | | . MTIKNVICDIDE . MTIKSVICDIDE . MTIKSVICDIDE DRKPIESWITDMDE DEPILSVITDMDE RRDEIECWITDMDE SGBIECWITDMDE MSLDYQGYLIDIDE . MKYKGYLIDIDE | VLMHDNVAVPGAAE VLLHDNTAIKGAM VLLHDNKIIPGSDH VLIHEGVPIPGAD VLVBENKALPGAP VLVBENKALPGAP VLVHENKALPGAP TIYLGKEPIPAGK TIYLGKEPIPAGK TIYKGKDRIPAGE | ELTGILEKGLPLVL PLARIQDAGMPLVI FIQRILEQGNPLVI FIKRLRDSGKPFLV LUQWRDQGTPFLV LUQWRDQGTPFLV LUQWRDQGTPFLV PVERLQEKDLPFLP FVHRLQKRDIPYLF FVHRLQKRDIPYLF | LTY PSOTGCDLA LTY PSOTGCDLT LY PSOTGCDLT LY PVOTGKDLQ LY NSIYTPRDLA LTNNSIYTPRDLA LTNNSIPTPRDLA LTNNSIPTPRDLA VINNTTKSPETVA VINNTTKSPETVA VINNTTKTPESVK |
|--|--|--|--|--|---|
| E.coli_K-12 salomnella Yursinia Shreptomyces Nocardia Leifsonia Arthrobacter Enterococcus faecalis Streptococcus_pneumoniae Staphylococcus_aureus Thermotoga | 60 NRFA.TAGVDVPDSV NRFI.TAGVNVPDSV NRFI.TAGLVPESA NRLS.AAGIDVPESA ARLR.RMGLEVPIES ARLR.SGLDVPESS ARLR.SSGLEIPEEN QRLANEFDIHVPASL EMLAQNFNIDTPLST EKLR.EMHIDAKPEE RKLR.NXGVDVPDDA | FYTSAMATADFLRR FYTSAMATADFLRR FYTSAMATADFLR IWTSALATADFLKH IWTSALATATFLND IWTSALATATFLND IWTSALATADFLRQ VYTATLATIDYMKE VYTATLATIDYMKE VYTSALATADY ISE | CE. GKKAYUV CD. GKKAYVJ CE. GSKAPVJ CRPGGS. AYVJ CRPGGT. AYVV CNPGGS. VYVJ CVRGSDSGNRAYTJ ANRGKK. VPVJ LGLEKT. VYVV GSPGAS. VYMI | GEGALIHELYKAGP GEGALIHELYKAGP GEGALTHELYKAGP GEGALTHELYKAGP GEGALTHELYKAGP GESGLTTALHDIGY GESGLTTALHEIGY GEAGLTTALHEIGY GEAGLTTALHEAGP GEAGLIDLILEAGP GEAGLKEAIKAAGY | TI.TDVNPDFVIV TI.TDINPDFVIV TI.TDINPDFVIV TL.TDDPDFVIV TL.TDDPDYVVL TM.TETNPDYVVL TM.TETNPDYVVL W.DETNPDYVVV W.DETNPDYVVV VE.DKENPAYVVV VIENDEHVDYVVI |
| E.coli_K-12 salomnella Yursinia Shweanella Streptomyces Nocardia Leifsonia Arthrobacter Enterococcus_faecalis Streptococcus_pneumoniae Staphylococcus_aureus Thermotoga | 120 130 GETRSYNWDMMHKAA GETRSYNWDMMHKAA GETRSYNWDMMHKAA GETRSYNWDMIHKAA GETRTYSFEANTKAV GETRTYSFEANTKAV GETRSYSFDAITTAI GETRSYSFDAITTAI GLDEVSYEKPATAT GLDBQVTYEKFATAT GLDEVTYEKLAIAT GFDKTLTYERLKKAC | PFVANGARFIATNP YFVANGARFIATNP GFVARGARFIATNP RLINAGARFICTNP RLIGKGARFIATNP RLIGKGARFIATNP LJIQKGARFIATNP LAIQKGALFIGTNP LAIQKGAHFIGTNP | DTHGRGFYPAC DSHGHGFAPAC DTHGPAYSPAC DTHGPAYSPAC DETGPSTEGPLPAT DATGPSAEGPLPAT DATGPSAEGPLPAT DATGPSAEGPLPAT DATGPSKDGPMPAT DKNIPTERGLLPGA DLNIPTERGLLPGA DVSIPKERGPLPGN | CALCAGIEKISGRK GALCAPIEKISGRK GALCSPIERITGKK GSVAALITKATGKD GAVTAMITKATGRD GAVTAMITKATGMK GSIJALITKATGRE GSVTFVETATQTK GSLITLLEVATRVK GSLITLLEVATRVK | P. FYVGXPSPWII P. FYVGXPSSWII P. FYGXPSSWII P. YYGXPNPLMM P. YYVGXPNPLMM P. YVIGXPNPMMF P. YIVGXPNMMF P. YIGXPNAMMF P. YIGXPNAIIM P. VYIGXPAIIM |
| E.coli_K-12 salomnēlla Yursinia Shweanella Streptomyces Nocardia Leifsonia Arthrobacter Enterococcus_faecalis Streptococcus_pneumoniae Staphylococcus_aureus Thermotoga | RAALNKMQAHSEETV RAALNKMQAHSEETV RAALNKMQAHSESTV RSALNHIDGHSENTV RTGLNAIGAHSETSA RSALRRLGAHSQSTV RSALNRIGAHSENTA RSALNRIGAHSENTA ERAIAHLGVEKEQVI | IVGDNLRTDILAGF IVGDNLRTDILAGF IIGDNMRTDILAGF MIGDRMDTDVLAGM MIGDRMDTDVISGL MIGDRMDTDVISGL MIGDRMDTDIIAGM MVGDNYETDIQSGI MVGDNYLTDIRAGI VVGDNYLTDIRAGI | QAGLETILVLSGVS QAGLETILVTSGVS QAGLETILVTSGVS BAGMQTFLVTSGIS BAGLHTILVTSGIS BAGLHTILVTSGIT QNGIDSLLVTSGFT QNGIDSLLVTSGFT DNGIPTLLVTTGFT | TINDIDSMPFRPSW TLTDIDAMPFRPSY KLEDIDKEPFRPNH RPEQVENFPYRPSO TRAAVEQYPYRPSO HKDDIAAYPFRPDE HKDDIAAYPFRPNQ PKSAVPTLPTPTY KAEEVAGLPIAPTH TLEDVQNKNVPPTY | YPSVABIDVI IYPSVADIDII YPSVADIDII VFACAGDIDVV VIDSVADLVERV. VIDSVADLVRRV. ILPGVDALVAREP ILNSVADLKSQI. VVSLDBWTFEG. VSSLABWDFDEN SFKDLNBAIAELE |
| E.coli_K-12 salomnella Yursinia Shweanella Streptomyces Nocardia Leifsonia Arthrobacter Enterococcus_faecalis Streptococcus_pneumoniae Staphylccoccus_aureus Thermotoga | NPFG. VESEF K | | | | |

Figure S1. The sequence alignment of NagD homologues from various bacterial species. In red-highlight and boxed in blue are shown the stringently conserved residues. In red text and boxed in blue are conserved residues. Accession codes for sequences or PDB: *E. coli* K-12: P15302; *Salomnella enterica*: NC_004631.1; *Yursinia pseudotuberculosis*: NC_006155.1; *Shweanella oneidensis*: NC_004347.1; *Streptomyces*: CAB65573; *Nocardia farcinica*: YP_121565; *Leifsonia xyli*: NC_006087.1; *Arthrobacter*: AAHG1000001.1; *Enterococcus faecalis* NP_814917; *Streptococcus mutans*: AAN59079; *Staphylococcus aureus*: YP_040313; *Thermatoga maritima*: NP_229540.