```
Mb fold
.........|..||..|................|..||..||...|.......|..|..................|..||..||..|..||...............|..|...||.........|.........|..|..|||.||.....................||..||..|...|......

Plant NsHbs

Marchantia gi|15809396|AAK07743
----------------------------DSATHSVTFFAKIFEIAPGAKAL--FSFLKDS------DIPFDKNPKLKAHALTVFRLTGDSACMLGEKGAIDALHPKLKELGKKHVG------YGV-----IAAHFDVVKAALLSTIESLLPED----WPAEKKVATCGAWSQAYDELAGVMINEMQAIRDAEKA
Euryale Hb1 gi|33590381|AAQ22728
-FSEEQEALVVKSWGVMK---------KDAGQLGVKFFAKIFEIAPSAKRM--FSFLRDS------DLPLEQNPKLKPHALSVFAMTCESAVQLRKAGKVTVRESSLKDLGATHFK------YGV-----VDEHFDVRFAFSRSVFSSIR---------DLFQPEMKIAWGEAYDKLAA---------------
Arabidopsis Hb1 gi|15226675|NP_179204
VFTEEQEALVVKSWSVMK---------KNSAELGLKLFIKIFEIAPTTKKM--FSFLRDS------PIPAEQNPKLKPHAMSVFVMCCESAVQLRKTGKVTVRETTLKRLGASHSK------YGV-----VDEHFEVAKYALLETIKEAVP--------EMWSPEMKVAWGQAYDHLVAAIKAEMNLSN-----
Lotus gi|11230993|BAB18108
GFTAQQEALVGSSYETFK---------KNLPTNSVLFYTVILEIAPTAKDM--FSFL--------KESGPKHSPQLQAHAEKVFALTRDAATQLVAKGEV---TLADASLGAVHVQ------KAV-----TDPHFVVVKEALLQTVKEAVG-------ADEWSDDLSTAWEGAYDGLATAIKKAMG--------
Trema gi|81428|S00838
VFTEEQEALVVKSWAVMK---------KNSAELGLKFFLKIFEIAPSAKNL--FSYLKDS------PIPLEQNPKLKPHAMTVFVMTCESAVQLRKAGKVTVRESNLKRLGAIHFK------NGV-----VNEHFEVTRFALLETIKEAVP--------EMWSPEMKNAWGEAYDQLVAAIKSEMKPSST----
Cichorium  gi|3297847|emb|CAA07547.
GFSEKQEALVKESWEVMK---------QDIPALSLYLYAMILEIAPEARGL--FSFLKDT------NVIPQNNPKLKSHAVKVFKMVCESAIQLREKGEVVVSGSTLKYLGSVHLE------KGI-----VDPQFEVVKEALIRTVAKAMG--------EKCSEEMKSAWSEAYDELADAIKTEMKKEAAETQT
LycopersiconGlb1 gi|12963875|AAK07676
SFSEEQEALVVKSWGSMK---------KDAGEWGLKFFLKIFEIAPSAKKM--FSFLKDS------NVPLDQNPKLKIHAKSVLVMTCEAAVQLRKAGKVVVRDSTLKKIGATHFK------YGV-----VDEHFEVTKYALLETIKEASQ--------EMWSVEMKNAWGEAYDQLVSAIKTEMK--------
Oryza Hb1 gi|17366135|O04986|HBL1_ORYSA
SFSEEQEALVLKSWAILK---------KDSANIALRFFLKIFEVAPSASQM--FSFLRNS------DVPLEKNPKLKTHAMSVFVMTCEAAAQLRKAGKVTVRDTTLKRLGATHLK------YGV-----GDAHFEVVKFALLDTIKEEVP-------ADMWSPAMKSAWSEAYDHLVAAIDQEMKPAE-----
Ceratodon gi|10764841|AAG22831
VYSKDDEVLVKKSWDILK---------KDAQRNGINFFRKVFEIAPGAKSM--YSFKRDS------TIPFEENPKVKNHARYVFMMTGDAAVQLGEKGAYQVLDSKLQKLAATHIN------AGV-----TDDQFEIVKEAILYAIEMGVP--------ELWSPELKSAWGDAYDMLAAAVKAEMHAQRAAAQ-
Physcomitrella gi|7658243|AAF66104
TYSKENEQLVKQSWEILK---------KDAQRNGINFFRKVFEIAPGAKAM--YSFLRDS------TIPFEENPKVKNHARYVFMMTGDAAVQLGEKGAYQVLESKLQKLAATHVN------AGV-----TDDQFEIVKEAILYAIEMGVP--------DLWSPELKSAWGDAYDMLAEQVKAEMHAQRSAATS


Plant SHbs

Canavalia gi|1170183|P42511
AFSEKQESLVKSSWEAFK---------QNVPHHSAVFYTLILEKAPAAQNM--FSFL---------SNGVDPNPKLKAHAEKVFKMTVDSAVQLRAKGEVV---LADPTLGSVHVQ------KGV-----LDPHFLVVKEALLKTFKEAV--------GDKWNDELGNAWEVAYDELAAAIKKAMGSA------
Lupin I gi|126232|P02239
VLTDVQVALVKSSFEEFN---------ANIPKNTHRFFTLVLEIAPGAKDL--FSFLKGS------SEVPQNNPDLQAHAGKVFKLTYEAAIQLQVNGAVA-SDATLKSLGSVHVS------KGV-----VDAHFPVVKEAILKTIKEVV--------GDKWSEELNTAWTIAYDELAIIIKKEMKDAA-----
Sesb II gi|225726|prf||1312241A
GFTDKQEALVNASYEAFK---------KNLPGHSVLFYSFILEKEPAAKGL--FSFLKDS------DGVPQNNPSLQAHAEKVFGLVHDAAAQLRATGVVV---LADASLGSVHVQ------KGV-----TDPHFVVVKEALLKTLKEAA--------GATWSDEVSIAWEVAYDGLAAAIKKAMS--------
Parand I  gi|862963|AAB86653
*VFTEEQEALVVKAWAVMK---------KNSAELGLQFFLKIFEIAPSAKNL--FSYLKDS------PVPLEQNPKLKPHATTVFVMTCESAVQLRKAGKVTVKESDLKRIGAIHFK------TGV-----VNEHFEVTRFALLETIKEAV--------PEMWSPEMKNAWGVAYDQLVAAIKFEMKPSST----*
Casuar I gi|454920|CAA54775
ALTEKQEALLKQSWEVLK---------QNIPAHSLRLFALIIEAAPESKYV--FSFLKDS------NEIPENNPKLKAHAAVIFKTICESATELRQKGHAVWDNNTLKRLGSIHLK------NKI-----TDPHFEVMKGALLGTIKEAI--------KENWSDEMGQAWTEAYNQLVATIKAEMKE-------

Bacterial GCSs

Agrobacterium tumefaciens gi|15888368|NP_354049    499  (18)
Agrtum
LGLGHGERQNLSDMKGVITG--------SLDASLDRFYTKVRAV-PETAKF--FSSEAHI-----------HHAKSMQLKHWSRIASGTFNEDYTN---------AVTAIGRTHARLGLEPRWYI-------GGYALMLDGIVKAVIESMWP--RGLLAKGGSDRVKDALSATIKAALLDMDYSISVYL-----
Azotobacter vinelandii gi|23104391|ZP_00090857  472  (23)
Azovin
LLLGQFPAPVVAQIRELATT--------HQSELPGYFYEQMLQDE-QAML---FLTHE---------------QVKSRLHGTLRQWIVSVFSMSDDDAALQALIAQQKQIGEIHARIKIPIHLVL-------RGARHLRERLFVLLRESDI-------PRQRKLFGQRLISETVDLAMEIMSRAFSDAY-----
Bacillus subtilis gi:37927254|NP388919   433aa  (40)
Bacsub
VRLGDAELYVLEQLQPLIQE--------NIVNIVDAFYKNLDHESSLMDI---INDHSS----------------VDRLKQTLKRHIQEMFAGVID----DEFIEKRNRIASIHLRIGL-----L-----PKWYMGAFQELLLSMIDIYEAS----------ITNQQELLKAIKATTKILNLEQQLVLE-----
Chromobacterium violaceum gi|34495694|NP_899909    375  (25)
Chrvio1
FALAPRDEMLLRAAGNLVES--------HLEELVTRFYELQTST-PEIALL--IGDA----------------DTLQRLRSAQRRYVVDLFSGIY----DLEYVNIRLRIGLVHKRIGVEPK--L------YLAAVDSLKFLLAEKLTELIPD-AEVRLHTLQALDKLMMFDVALVFETYIRSLVAEIE-----
Desulfotalea psychrophila gi|51244053|YP_063937    363  (24)
Despsy
LFFTDEDVQVLVSLKEVVSE--------NIEEIVEAFYERVIPFD-EMDG---IGDA----------------ETLRRLKNYQRSYIKTLFAGQY----DEEYVHSRLRVGVVHKRIG------V-----EPKFYISAIHTLLTILREIIT---KNSSNDCQACVRNLRAVEKIIMFDLSLTFDTYINSL----
Escherischia coli gi|15801648|NP_287665    460aa  (13)
Escco1
GLVEQADPPIRAKAAEIAVA--------HAHYLSIEFYRIVRID-PHA-----FLSN---------------EQVERQLKSAMERWIINVLSAQVDD--VERLIQIQHTVAEVHARIGIPVE--I-----VEMGFRVLKKILYPVIFHAKC----------EKARKLQVYHFSINSIDIAMEVMTRAFT-----
Exiguobacterium gi|45532973|ZP_00183969  430aa  (37)
Exigu01
IDLTVEDLQIVRLIRHDVAE--------WMPQMVDAFYAELVRV-PELK----IERH----------------STLERLKGTLAKHILQMFDGQV----TPDYINVRKRIADRHVRIGLQNKWYI-------AAFQKVLNVLSRQIEDSTLPESEQVRIVRSAGKLFNLEQQIVLTMYEEMVDAERQTV-----
Geobacter sulfurreducens gi|39998400|NP_954351    300aa  (11)
Geosul
YRFTDEDAELLGSLFPLAET--------NKERLADQFYDYLLGI-PETAE---FLK----------------EDLVLQKLKQTHQDWFVSLFAGSY----DNRYIHNLQKIGHAHVRVGLNAHY-V------NVAMNVVRQFTLSIIQHLQSPSD----RRQRREAVEKILDINLDIMSASYREEEMRKF-----
Haloarcula marismortui gi|55229828|AAV45247     497aa  (44)
HalmarGCS
TGFDGSDVDNLTAMADETNA--------RAEAVVDDFYDHLQSFD-ETVEI--FGRS----------------TKSVDQLKNTQTQYLRDLVAGTY----DKQYFENRARIGKIHDMLDLGPKIYL-------GAYSIYFEHFLRTIVEDLQSG---------DAARDEALEEMQSRALSVFKLLNLDQQ-----
Magnetococcus sp.MC-1 gi|48832082|ZP_00289126507  507aa  (20)
Magnet1
VGFTEKDASILKKLRPVAAK--------HATAVVNTFYTRLSGF-AHLEKI--IGG----------------AGSSVERLKRTQEEYLVQLFDGEY----GRDYFVRRWRIGQIHNKIGLEPDWYL-------GGYSLYRQLLLPILLDKITTID-------DVQRAMAAIDKILTLDSELAIGSYIDAV-----
```

```
Magnetospirillum magnetotacticum gi|23014982|ZP_00054774  443aa  (11)
Magspi1
LQIDEDTKRALREFREVLSR--------HIDGVLDTFYRHVSNN-PATAKM--FAN----------------PDRMAHARSMQKKHWMESVFLGQ---FDDRYFAQVTEIGKVHQRIGLDPKWYT-------AGYCFVLNMVIGVAVEGSLYSD-------ELTQVLAAVNKAAFLDMDLATSVYIETNT----
Moorella thermoacetica gi|49235973|ZP_00330036  245aa  (13)
Moothe
LTLSEAELGLMEAEKEIFIK--------EADAVVKTFYDHLLQY-PYLENM--IKQH----------------STLERLTQTQKAYFISLTAPRI----DGEYISGRLRIGKKHQKIGLYPKWYL-------GAYRIYLSEIRRVIWHYHADDP---------DLCLRLLEAFTKRIIFDMQLAIENYI-----
Novosphingobium aromaticivorans gi|48849793|ZP_00304036  481aa  (17)
Novaro
FNIDHKDFERFPHIAKVLEN--------YAPPALDKLYDQIATT-PETASF--FGS----------------RQAMRHARDKQIEHWAGMFSGRA----DRSYFESAERIGNVHARIGLEPGWYI-------GGYAMVLEQVINAMFSGIGGIL---GAKRTARSVGSLVKMALLDMEVALSTYFRAEE-----
Sinorhizobium meliloti gi|15964389|NP_384742    533aa  (23)
Sinmel
AGLDEDACALLRDHRQALSP--------RIELALRALSHRLQAS-PDAARH--FDSD-------------RQIDRLHDLQSSHWNVLTDARFDGLRMGLDPRWQIASHAIVLEHLILGA-----I-------EDAWPKSILSLGKARRRELRDLV-------AALVRAAFVDTEIAVSLRFNALRQQHQRQ---
Vibrio vulnificus gi|27366532|NP_762059    306aa  (20)
Vibvul
HDLTEADLALIRKFGQIMVP--------KLDEYVKHFYDWLRNT-PEYEQY--F--------------------GDAQKLQRVQDSQVRYWKTFFDARI--DSAYLERRDVGEIHARVGLP----L-------PTYFAGMNISMVIFTKRMYDTVH----GIAYSSLVTAFTKLLHLDTTIVVDTYSRLI-----

Bacterial Pgbs

Aeropyrum pernix gi|14600601|NP_147118    195aa  (30)
Aerper
VMLGEKDVMYLKKACDVLK--------DQVDEILDLWYGWVASNE-HLIYY--FSNP------------DTGEPIKEYLERVRARFGAWILDTTC-RDYNREWLDYQYEVGLRHHRSKKGVTDGV--RTVPHIPLRYLIAFIYPITATIKPFLAKKGGSPEDIEGMYNAWFKSVVLQVAIWSHPYTK-------
Chloroflexus aurantiacus gi|22974473|ZP00020712   227aa  (63)
Chlaur
VLWTAEDEQYLRMAGEVLG--------DQVDAILDLWYGFVASH-AHLVYY--FTGS------------D-GQPIADYLSRVRQRFGQWILDTCR-RPYDQDWLNYQMEIALRHYRTKKNQTDGV--QSVPMIPLRYMIAFIYPITATIREFLARKGHSAAEVDRMHQAWFKSIVLQVTLWSYPYTR-------
Rubrobacter xylanophilus gi|46106825|ZP_00200180 196aa    (32)
Rubxyl
VLFTGEDERYLRMAGEVLE--------GRLDELLDVWYGFVADH-SHLVYY--FSSP------------E-GEPIQEYLERVRERFKRWVLDACR-RPYDQEWLDYQQEIALRHTREKKNRTDGV--EAPEEVSLRYMISFIYPITATVRPFLEEGGRPAEDVEKMHQAWFKAVVLHVTLWSQPYAR-------
Methanosarcina acetivorans  gi|20091705|NP_617780   195aa  (31)
Metace
VMFTAEDEEYIQKAGEVLE--------DQVEEILDTWYGFVGSH-PHLLYY--FTSP------------D-GTPNEKYLAAVRKRFSRWILDTCN-RSYDQAWLDYQYEIGLRHHRTKKNQTDNV--ESVPNIGYRYLVAFIYPITATMKPFLARKGHTPEEVEKMYQAWFKATTLQVALWSYPYVK-------

2/2Hb1s

Haloarcula marismortui  gi|55379358|YP_137208  120aa
Halmar
-------MASQSIFERIGG-------RDAVDAVVSDFYDRVRDD-PLLEPY--FEET----------------DMDQLRSHQTQFISTVAGGP--------VDYDGDDMQTAHEGMG------I-----TEDAFASVATHLEAALRANGVP----------DDDVEAILTEVAAMKDDIVEA-----------
Gemmata obscuriglobus
Gemobs
------------LWDRLGG-------ESVVKAVVHDFALAALED-PKVN----FTRG------------GKYKPSAAALEKSIVDMISDVSGGP--------FKYSGKSMKEAHKGMA------I-----TDAEFDALAVHLVASLKKF------------------KVADPEIAELVKVVA-----------
Rhodopirellula baltica gi|32475444|NP_868438 127aa
Rhobal
-----MSQSESDLLDQLGG-------MDGVRRVVDEMYVRVLAD-PELTHF--FDGV----------------PIEKLARMQTEFIASITSGD--------IQYTGADLTKVHAGRG------I-----TGVHFSKFCGHLTDALEANNVSSH----------AIDQVLGKLAMYSDKVTGSSNVDG------
Microbulbifer degradans  gi|4886128|ZP_00315184
Micdeg1
--CSSLQGQNQTLYKQLGG-------QDGVSALTRQLLINMAADN-RIAPR--FRGV----------------NIGKFKTGLDTYLCSITDGG--------CVYGGDSIKTIHSGYN------Y-----TATEFNALVENLMQAMETLQIPT----------ATQNKLLAKLAPSYQDVVYQ-----------
Mycobacterium avium  (HbN) 131aa
Mycavi1
ILARFRKAEPASIYDRIGG-------HEALEVVVEDFYVRVLADE-QLSGF--FTGT----------------NMNRLKGKQVEFFAAALGGP--------HPYTGAPMKQVHQGRG------I-----TMHHFGLVAGHLADALTAAGVPS----------ETVSEILGAIAPLAPEIATGEA---------
Myxococcus xantus (116aa
Myxxan1
-----------SVYEQLGG-------EPAMAAAVEVFYRKVLADD-HISHF--FEDV----------------DMERQAAKQKAFLTMVTGGP--------VHYSGKDMRAGHAPLVKRG---L-----NDSHFDAVAGHLKATLEEL--------------GVAAPLVARVMTIAESARADV----------
Nostoc punctiforme  gi|23126739|ZP_00108627 159aa
Nospun2
----AHNSAGSSLYKRLGG-------YNAIAAVIDDSAKNIFAD-PLIGKY--FIGL--------------STNSKQRLRELLIAQFCQAAGGP--------CIYTGRTMKLSHSGIGRG----L-----TNGEFYAFVNDIALALDKNGVKQPE----------KNQVLGFANSLRDQIVEKP----------
Synechocystis sp.PCC6803 gi|16330583|ref|NP_441311 124aa
Syncys
---------MSTLYEKLGG-------TTAVDLAVDKFYERVLQDD-RIKHF--FADV----------------DMAKQRAHQKAFLTYAFGGT--------DKYDGRYMREAHKELVENHG--L------GEHFDAVAEDLLATLKEMGV--------------PEDLIAEVAAVAGAPAHKRDVLNQ-----
Xanthomonas axonopodis  gi|21243436|NP_643018 131aa
Xanaxo
TSACASTQSRQTLYDELGG-------QAGIEALVETMLSRIADD-PRIVDK--FARV----------------NIVMLNERLVQKFCHVADGP--------CPDTAKSMQQAHAHLA------I-----REGDFNALVEDLNWAMDQRKIP----------RRTQNRLLARLAAMHGEIVNH-----------

2/2Hb2s

Agrobacterium tumefaciens  gi|15889651|NP_355332
Agrtum1
SKEATMSSETVTLYEAIGG-------DATVRALTRRFYELMDTL-PEAARCRAIHPA----------------DLSGSEAKFYDYLTGYLGGP----PVYVEKHGHPMLRRRHFVAP------I-----GPAERDEWLLCFRRAMDETIE----------NAKLREIIWAPVERLAFHMQNQEADNP-----
Bdellovibrio bacteriovorus gi|39576266|NP_969438 137aa
Bdebac1
------MSEEKKPYELLGG-------EEVLRQLCKRFYEIMDTI-PEAKGIRDMH----------------PGNLRGSEEKLFMFLSGWLGGP----GLFVEKYGHPRLRMRHFPFK------I-----GKSERDQWMMCMVQAFDELNIAEP----------LRSELLHSLLRLADHMRNVEEPESSE----
Bordetella pertussis  gi|33593344|NP_880988
Borper1
QPISEPLDTSKTVFDMLGG-------EPGVRELVDRFYDLMDME-SD------FKALR-----------DAHGPSLEQARDKLFWFLCGYFGGP----H-YIERFGHPRLRARHLPFS------I-----GEIERDQWVACMGRAMQDQQL--------------PAALVDRLLQAFYGTADWMRNRAG-----
Bradyrhizobium japonicum  gi|27376186|NP_767715 131aa
Brajap1
--MMTDSDVAISMFERIGG-------SATIDLLVDRFYERMDTL-PEAKV---IRAM--------------HAADLGLIRDVLKRYLTEWTGGP----RLYSVEKGHPRLRQRHIGFA------I-----GDAERDAWLTCMRGALEETVAD----------AAARQDLDKALSGLADWMRNR-----------
Burkholderia cepacia gi|46323938|ZP_00224300  142aa
```

```
Burcep1
AHSLHAAHAPPSPFDRLGG-------AAALARIVDAFYRQMDTR-PDAAGIRAMHGA-----------------DLGPVKTVLVTYLCEWLGGP----RHYSAQRGHPRLRMRHRAFA------I-----GMAERDAWLACMRAALDECGVEP----------ALRDELMQALFKLADWLRNTGR---------
Chloroflexus aurantiacus gi|22973316|ZP_00020053 140aa
Chlaur
HYGEVYIMSEPTIYEQIGG-------EATFRRIVDIFYARVEAD-PRLRHL--FP------------------ADLEPGKEHQRLFLMQYFGGP-----TYSERRGHPRLRMRHAPFP------I-----GPRERDAWLEHMLAALNEAGVPE-----------PARSVMENYFRHAAQAMMNRLGED------
Corynebacterium efficiens  gi|25028901|NP_738955
Coreff
------MTSTENFYDSVGG-------EETFNLIVHRFYEQVRTDD-ILGPM--YPQ-----------------DDWEGAEARLRMFLVQYWGGP-----DYSEQRGHPRLRMRHAHFP------I-----GVTAAERWLELMTRALDSLEEGTL-------TDDQRAAIWDHMVRAADMLINSNPDPHQR---
Deinococcus radiodurans  gi|15806025|NP_294726
Deirad
-MTTPLDLARGSLYDRIGP--------DTLALLVHRFYAHVARN-PDLAPI--FPA------------------DLSETARKQLAFLTGFTGGP-----LYHELYGHPRLRARHLPFP------I-----TPGRARAWLACMNAALRETPGLAEA---------DAHELYAALARVAVHMVNTEEAPDSD---
Exiguobacterium  gi|45532753|ZP_00183753  126aa
Exiguo
-----------MLYEDLGQ-------KQAIEQLVARFYERVYADH-LLRPL--FP------------------SDRRQVEQAQTRFLIQLTGGP----KQYESYDERMNLAMIHRLLP------I-----TEVHAIRWIELMTMTIEETISDSEA----------ATRLIERLRIGALNVLRICDAHQNR---
Kineococcus radiotolerans  gi|46365419|ZP_00227907.1  149aa
Kinrad
RKPAGPDGGTRSFFDEVGG-------HETFVRLVDAFYEGVAADE-VLRPM--YP-----------------EADLGPARQRMLMFLEQYWGGP----ATYSEQRGHPRLRMRHASFK------V-----NPDARDRWLTHMRAAVVSLGLAP----------AQEGVLWDYLDRAAHSMLNTFED-------
Leptospira interrogans  gi|24216675|NP_714156  145aa
Lepint
GGPPGPIPGLQLVFGSVGE--------NSLRKLVSDFYDQIPSS-SIS-FM--FP------------------ENLEDSKIKSADFLIQVTGGP-----LYSQNYGPPKMRARHLPFP------I-----DEKARRVWLSCYRKVLDDWDAEV----------SAKEVLWIFFKDFSTWMVNLESKTEE-----
Methylobacillus flagellatus  gi|46120937|ZP_00201861  137aa
Metfla
DKIGEPGSSKATLFELLGG-------AEQVRAIVERFYDIMDTD-PRAAGIRAMHAP-----------------DLTSAREKLFMFLTGWTGGP----QLYMERYGHPRLRMRHMSFP------I-----DDSARDQWMYCMVKAMHDTGVDE-----------AIIEKMGAALYGVADFMRNREG--------
Mycobacterium avium gi|41408389|NP_961225 (HbO) 129aa
Mycavi2
-----------SFYDAVGG-------AETFQAIVSRFYAQVPEDE-ILREL--YPL-----------------DDLEGAEERLRMFLEQYWGGP-----TYSDRRGHPRLRMRHVPFR------I-----TPLARDAWLRCMHTAVASIDSKTLD-------DEHRRELLDYLEMAAHSLVNSP----------
Oceanobacillus iheyensis   gi|23098673|NP_692139 128aa
Oceihe2
----MEQMEHRTIYEAIGG-------YRTMNNIVHAFYTRVEVH-PDLKPI--F------------------PDDLTETIRRQMMFLTQFFGGP-----LYEEERGHPMLRRRHLPFP------I-----TPTRRDAWLTCMAAALEEAEIKEP-------YRTAMMERLTLTANHMMNTPE------------
Photobacterium profumdum  gi|46913965|CAG20747 160aa
Phopro
QKTTPVYGMGDATLLAAGG-------EEGVKKLVDAFYDYMDTL-PEAEVVRAMHKP-----------------DLTDSREKLTTFLTGWMGGP----SRYAEKYGSMNIPGAHRHLS------I-----GIAEKEAWLLCMQKALDDQGYDD--------VFKRYVMVQLSFPAEMCRNRS------------
Rubrivivax gelatinosus  gi|47573541|ZP_00243579  136aa
Rubgel
--TPPESTAPATPYEWVGG-------EHAVRALVDRFYDLMDLE-PG------FHDLR-----------RLHPSTLDGSRDKLFWFLCGWLGGP----DLYVERFGHPRLRARHLPYS------I-----GVRERDQWMACMEQAMHECELDP-----------LLVSRLKASFANTADWMRNRGT--------
Staphylococcus aureus gi|15923993|NP_371527  121aa TrHb2
Staaur
---------MTTTPYDIIG-------KEALYDMIDYFYTLVEKDE-RLNHL--FP------------------GDFAETSRKQKQFLTQFLGGP------YTEEHGHPMLRKRHMDFT------I-----TEFERDAWLENMQTAINRAAFPQG-------VGDYLFERLRLTANHMVNS----------------
Thermobifida fusca  gi|48835434|ZP_00292434  163aa
Thefusc
--MTSQHDDQMTFYEAVGG-------EETFTRLARRFYEGVAAD-PVLRPM--YP-----------------EEDLGPAEERLRLFLMQYWGGP------YSERRGHPRLRMRHFPYR------I-----GAEERDRWLTHMRAAVDDLALPA-------HLEQQLWEYLVYAAYAMVNVPEDAQPP-------
Thiobacillus denitrificans  gi|52006448|ZP_00333828  126aa
Thiden
---------MQTHYDRLGG-------ADTIRRLVDRFYDLMDED-PDY-----YG-IR-----------KLHPTDLTESRNKLAWFLSGWTGGP----PEYTDRFGHPFLRRRHLPFV------I-----GESERDQWMGCMIRAMQDVGLDA-------ALQQELTAALFQTADFMRNQPQ------------

2/2Hb3s

Acidithiobacillus ferrooxidans TIGR_243159|contig:10428:a 132aa
Acifer3
---MEQVSPPKRRLEPLCEK----ITLPVIRAVVDDFYNRIQHH-PTLAEP--FSIV----------------QDWELHKDRLVHYWWTVSGGL-------PYKDYRYALGDKHAPLG------I-----THSLVDDWLALFHETMLDHMDAD--------MARRWHGMAAGIGESLRLMFAPR----------
Agrobacterium tumefaciens  gi|15887600|NP_353281
Agrtum2
VHNAEIQDRAEKAMAEIGV------DAAFIDLLVETFYGRVLEHQTLGPV---FD-----------------AGRWPEHMARMKQFWTAIAFKN---------GSYGGKPVQAHLG--------V--QGMSAELFPQWLALFSVTLDDIAPSRKAHD---WFMETAERIARSLTLSLFYNPAMDDP--------
Bdellovibrio bacteriovorus  gi|42523942|NP_969322  128aa
Bdebac2
-------MTTPTSRKPLES-------REDVMVLVDSFYAKVRAD-GYIGPI--FTDV--------------AKVDWEEHLPKLYNFWSDLLLGE---------DSYRGRPFPPHTKLN------L-----ERGHFEQWLRLFVETVDEHFVGLKANE--------AKERAYRIARNFMINLQLLSLE-------
Bordetella pertussis gi|33593122|ref|NP_880766
Borper2
----------MGAPDLCTE--------EDITQLVHQFYAEVRRD-AELGPI--FN---------------THVDDWDVHLAKLVDFWSSILRGT-----RGTPMPRHVALPGLHAG--------L---------FERWLRLFRATAAAQPNQAMAEQ--------ACLMAGRIAQSLWYGYQLHRSPDQA----
Bradyrhizobium japonicum (gi|27381333|NP_772862 154aa
Brajap2
AAAARRNLLTQDAIERTGI------TEEMIAELVTRFYGRVRED-ALLGPV--FA-V---------------VQNWDEHLAKLGDFWSSVVLMS-----SPMRAHLPLGLVGDHFDRW------L-------DLFEQTARDVCPPAAAALFIDKA-----------RRIADSFEMASATVAGRIASPRHVLRS-
Burkholderia cepacia gi|46315472|ZP_00216054  186aa
Burcep3
--MTALPASPDSAPARPRD---TEPTEENIRDLVYAFYDRVRAD-PLLGPV--FDAK--------------LEGRWDDHLPKMVTFWSSLVLGT-----K----GYRGNVQQAHQPLDG-----I-----EPAHFSRWLSLFLKTVETRYAPPAA-----------IRFMEPALRIAQSLQLSRFGWDYQIPP-
Campylobacter jejuni gi|15791829|NP_281652
Camjej
-----------MKFETINQ--------ESIAKLMEIFYEKVRKDK-DLGPI--FNNAI-----------GTSDEEWKEHKAKIGNFWAGMLLGE---------GDYNGQPLKKHLD--------L--PPFPQEFFEIWLKLFEESLNIVYNEE----------MKNVILQRAQMIASHFQNMLYKYGGH---
Caulobacter crescentus  gi|16127286|NP_421850
Caucre
-----MSLSPRSQARADA----APFTDEEIDRLVETFYARIRQHH-RLGPI--FEAA-------------IGPDGWPEHLAKLKDFWSSVLNTS---------GRYKGQPVAAHQGVEG-----V-----TEGLFTPWLGLFHMTCAELFDAP----------RAAIIGQKAERIAATLKLGLFHRPGAVA-
Desulfitobacterium hafniense  gi|23119438|ZP_00102524
Deshaf
-------MFKRGAMRSSHD-----LDEAALAALVRRFYAAVREDD-RLGPV--FA---------------AAIDDWDAHERRLIDFWTSLMLGS------GRYRGNPAQMHLRHAGR-------I-----QPELFAHWLALWRRCSEHTLPPAQ----------AAAVQARAGRIAEHLQRLLAQA------
Mycobacterium avium (Actinobacteria) 148aa (HbP)
Mycavi3
```

```
-------------MPDLRD-------RADVEALLRRFYGRALDDE-VLAEP--FARL--------------RATGLDDHVPTMCDFWETVLFRA---------GRYRGSALQAHRDIHRRAP--L-----SDRHFRRWLTLWHITVDEMYRGPAAD-------RAKIQAARIAWAMHRRLTGADSPELL-----
Novosphingobium aromaticivorans  gi|48849793|ZP_00303907 127aa
Novaro
------------MNDLSITE-------GDIDRLVPEFYARVRKDE-VLGPI--FNGA---------------IEDWPDHLERLKAFWSSIMLTS---------GRYKGQPLIAHVKHEAH----M-----TPENFARWLELWRQTTGDILGPD------------KAELFRQKAARIAESLTMGVQFYRERQRA

Chlamydomonas reinhardtii 2/2Hbs

Chlamydomonas reinhardtii 160981
Chlrei1
--MAADTAPADSLYSRMGG-------EAAVEKAVDVFYERIVAD-PQLAPF--FANV-----------------DMKKQRRKQVAFMTYVFGGS--------GAYEGRDLGASHRRL-------IREQGMNHHHFDLVAAHLDSTLQELGVAQ----------ELKAEAMAIVASARPLIFGTGEAGAAN----
Chlamydomonas reinhardtii 160982
Chlrei2
AAGVVDAAAPPSIFDRLGG-------EAAVEKAVDVFYERIVAD-PQLAPF--FEGV-----------------DMRTQRRKQQAFMTYAFGGA--------TGYTGRDLAAAHRRL-------IRDKGLKEEHFDMVAGHLAGTLQLLGIGA----------DLVDEVIALVATTKPVIFERV----------

Ciliate 2/2Hbs

Paramecium caudatum gi|121272 121aa
Parcau
-----------SLFEQLGG-------QAAVQAVTAQFYANIQADATVATF---FN-----------------GIDMPNQTNKTAAFLCAALGGP--------NAWTGRNLKEVHANMG------V-----SNAQFTTVIGHLRSALTGAGVAAALV----------EQTVAVAETVRGDVVTV-----------
Tetrahymena pyriformis gi|121274 121aa
Tetpyr
------MNKPQTIYEKLGG-------ENAMKAAVPLFYKKVLADE-RVKHF--FK-----------------NTDMDHQTKQQTDFLTMLLGGP--------NHYKGKNMTEAHKGMN------L-----QNLHFDAIIENLAATLKELGVT--------------DAVINEAAKVIEHTRKDMLGK-------

Plant 2/2Hbs

Arabidopsis GLB3 gi|18418064 175aa
Arabid
DQAEAFAIDESNLFDKLGL--------QTFINLSTNFYTRVYDDE-EEWFQSIFS-----------------NSNKEDAIQNQYEFFVQRMGGP----PLYSQRKGHPALIGRHRPFP------V-----THQAAERWLEHMQNALDDSVD-----------QDSKIKMMKFFRHTAFFLVAGNELKNQN--
Glycine  gi|44888958  172aa
Glycin
SIADAFAIDEANLFHKLGL--------QTLVNLSTNFYNRVYDDE-QEWFRSIFA-----------------NSDKQNAIQNQIEFLVQRMGGP----TLYSQRRGHPALIARHRPFP------L-----THEAANRWLHHMGQALDSTSA---------IDDDSKIKLINFFRHTAYFLVAGVEHKDQN--
Oryza  gi|50725383 172aa
Oryza
AAGDAFAIDDGNVFEALGG------TTQPFVDLSTNFYTRVYEDE-EEWFRQIFA-----------------GSKKEDAIRNQYEFLVQRMGGP----QLFSQRRGHPALIARHRPFP------V-----THQAAERWLHHMQQAVDTTDS---------IDAATKTKMMYFFRHTAYFLVAGNEMTRQG--

Thalassiosira pseudonana

Thalassiosira pseudonana  Scaff_18
Thapse2
QSLLATTSLTPTLLERIGG-------SSGFITLSTLFYNRVFNDKANPW----FLGI-------------FATSTKDEAIDNQFRFLVQTFGGE---ELYRQKKGKYTRLVGRHANYQ------I-----QIAAAERWVVHMEGAIDEHEALVDDEE-----ARECLKLYFRYTAFYIVVASEFMRDDQ-----

Bacterial FHbs

Azotobacter vinelandii gi|23104720|ZP_00091182  393aa
Azovin
MLTEQQRAIVKATVPLLET.........GGEALTSHFYRIMLAEHPEVRPL..FNQA..............HQASGEQPRALANAVLMYARHI...DRLEALGPLVAQIVNKHVSLQ......V.....LPEHYPIVGACLLRAIREVLGAETA.......TDAVLDAWGAAYGMLAELLIGAEEAAYAGAAV
Bacillus anthracis gi|30261545|NP_843922  403aa
Bacant
MLSEKTIEIVKSTVPLLQE---------KGVEITTRFYEILFSEHPELLNI--FNHT---------------NQKKGRQQQALANAVYAAATYI---DNLEAIIPVVKQIGHKHRSLG------I-----KAEHYPIVGTCLLRAIKEVAGA----------PDEVLNAWGEAYGVIADAFISIEAEMYEEAAH
Burkholderia cepacia gi|46316131|ZP_00216711  394aa
Burcep1
------MARVKATAPVLAV---------HGATITKHFYQRMFARHPELKNL--FNQT---------------HQKTGSQPETLAKAVYAYAANI---DNLGALGGAVTKIAHKHASLN------I-----RPEHYPIVGENLLASIVEVLGDAV--------DAETLEAWRVAYGQLAQILIGVEADLYAGAAW
Burkholderia fungorum gi|22983511|ZP_00028694 414aa
Burfun
MLSAEHRAIVKATVPLLES.........GGEALTTHFYRILIGEHPEVRPM.-FNLA..............NQASGAQPRALANSVLMYARHI...DQLEQLGGLVSQIINKHVALN......I.....LPEHYPIVGQCLLRAIREVLGAEIA.......TDAVIAAWAAAYRQLADLLIGLEEKIYAEREA
Chromobacterium violaceum gi|34498943|NP_903158  404aa
Chrvio1
MLSEQTRSLVKATVPVLQQ---------HGVALTSHFYRRMFEHNPELKNV--FNQG---------------HQHSGQQQQALAMAVLAYARHI---DDPSPLLPVLTRVAHKHVSLG------I-----RAEHYPIVGKHLLASIRELLGEAA--------GDDLIQAWAEAYGLLADTLIGIENGMYGDATS
Corynebacterium efficiens gi|25029214|NP_739268  404aa
Coreff
RLSEAHAEMIKATLPLVGS---------KINDITPVFYNRMFTAHPELLRST-FNRA---------------NQKQGDQQKALAASIATFATML-VDPEAPDPVDMLARIGHKHVSLG------I-----TADQYNIVHEHLFGAIVEVLGADVV-------TPPVAEAWDAVYWLMANVLIDLESDLYASNDL
Deinococcus radiodurans gi|15807907|NP_285566  403aa
Deirad
MLTPEQKAIVKATVPALEA---------HGETITRTFYASMFAAHPELLNI--FNPA---------------NQQTGKQARSLAASVLAYAAHI---DHPEALGGMVGRIAHKHVSLE------V-----LPEHYPIVGQYLLGAIAGVLGDAA--------KPEILDAWAAAYGELADLMIGIEKGMYDAGAG
Escherischia coli gi|16130477|NP_417047  396aa
Esccol
MLDAQTIATVKATIPLLVE---------TGPKLTAHFYDRMFTHNPELKEI--FNMS---------------NQRNGDQREALFNAIAAYASNI---ENLPALLPAVEKIAQKHTSFQ------I-----KPEQYNIVGEHLLATLDEMFSP----------GQEVLDAWGKAYGVLANVFINREAEIYNENAS
Exiguobacterium gi|46114160|ZP_00184406.2  384aa
Exiguo
MLTPSAITIIKSTVPVLAE---------HGHSITRVFYQRLFENHPEMKHI--FNQS---------------NQKNDRQSQALATAVYAAAAYI---DQLETLKPTLLPVLHKHRSLQ------I-----KPYMYDIVGTELIGAIQDVLKDAA--------TPDIIDAWTAGYGEIAKLFISLEAELYQQDEQ
Magnetospirillum magnetotacticum gi|23006958|ZP_00049038  403aa
Magmag
PLSARTIAVVKATIPALET---------HGLTITRRMYERLFQNT-DIRDL--FNQSH--------------HGETGSQPKALALAVLAYARNI---DNLGVLTGAVERI-QKHVALN------I-----LPEHYSFVADALIGAIQDVLGEAA--------TPEIVEAWGEAYWFLAELLIGREATIYRDLAS
Mycobacterium tuberculosis gi|15607526|NP_214899 390aa
Myctub
--MGLEDRDALRVLQNAFKL--------DDPELVRRFYAHWFALDASVRDL--F-----------------PPDMGAQRAAFGQALHWVYGEL-VAQRAEEPVAFLAQLGRDHRKYG------V-----LPTQYDTLRRALYTTLRDYLG------------HPSRGAWTDAVDEAAGQSLNLIIGVMSGAAD
```

```
Oceanobacillus iheyensis gi|23097746|NP_691212 406aa
Oceihe
LLDKKTTEIIKATVPVLKE---------HGEAITKHFYKILLENNPELKNV--FNQT---------------NQRKGAQSKALANTVYAAAANI---EKLEEILPHVKQIAHKHVSLN------I-----KPEQYPIVGKYLLIAIKEVLGDAA--------TDEIIEAWEKAYFVIADIFISVEKEMYNEKKN
Photorhabdis luminescens gi|37527172|NP_930516.1 396aa
Pholum
MLDNQTIATVKSTIPLLSA.........TGPKLTAHFYERMFKHNPELKNI..FNMS...............HQLNGDQREALFNAICAYAANI.....DNLKVLLPAVEKIAHKHASLN....I.....QPEHYQIVGTHLLATLNEMFQPGN..........EILDAWGKAYGVLADIFINREEQIYHSGEL
Rhodopirellula baltica gi|32477410|NP_870404 417aa
Pirell
MLSEKTIRIVKEITPLVAA---------NAETITRRFYERMFEANPEVKAF--FNQA---------------HQHSGGQQKALAGAICAYFTHI---DNPAVLMPAVELIAQKHVSLG------I-----KPEHYPIVGSNLLAAIGDVMGDAA--------TPEIVEAVSEAYGFLADIFIGREGAIYEEQAS
Pseudomonas aeruginosa gi|15597860|NP_251354 393aa
Pseaer
MLSNAQRALIKATVPLLET---------GGEALITHFYRTMLGEYPEVRPL--FNQA---------------HQASGDQPRALANGVLMYARHI---DQLQELGPLVAKVVNKHVSLQ------V-----LPEHYPIVGTCLLRAIREVLGEQIA-------TDEVLEAWGAAYQQLADLLIEAEESVYAASAQ
Ralstonia metallidurans gi|48771323|ZP_00275665 406aa
Ralmet
MLSAASRPYIDASVPVLRE---------HGLAITTHFYREMFAARPELKRI--FNMG---------------NQANGSQQQSLASAVFAYAANI---DRADVLAPVVERIVHKHVAVG------L-----TPAHYPIVGKYLLEAIAAVLGDAA--------TPPLLAAWDEAYWLLAGELIAAEARLYERHGV
Sinorhizobium meliloti gi|16263102|NP_435895 403aa
Sinmel
MLTQKTKDIVKATAPVLAQ---------HGYAIIQHFYKRMFQAHPELKNI--FNMA---------------HQERGEQQQALARAVYAYAANI---ENPESLSAVLKDIAHKHASLG------V-----RPEQYPIVGEHLLASIKEVLGDAA--------TDEIISAWAQAYGNLADILAGMESELYERSEE
Staphylococcus aureus gi|15923230|NP_370764  381aa
Staaur
MLTEQEKDIIKQTVPLLKE----------GTEITSIFYPKMFKAHPELLNM--FNQT---------------NQKRGMQSSALAQAVMAAAVNI---DNLSVIKPVIMPVAYKHCALQ------V-----YAEHYPIVGENLLKAIQDVTGLEE--------HDPVIQAWAKAYGVIADVFIQIEKEIYDQMMW
Streptomyces avermitilis gi|29832496|NP_827130 400aa
Strave1
MLSEQSAATVRATLPAVGA---------AVGEITARFYDRLFAARPELLRDL-FNRG---------------NQAAGTQRQALAGSIAAFATYL-VEHPDERPDAMLDRIAHKHASLG------I-----APGQYAVVHEHLFAAIAEVLGDAV--------TPEVAAAWDEVYWLMANALIAIERRLYAQHEN
Vibrio cholerae gi|15600953|NP_232583  394aa
Vibcho
MLTQEHINIIKSTIPLLES---------AGPALTQHFYQRMFSHNPELKHI--FNMT---------------HQKTGRQSVALFEAIAAYAKHI---DNLAALTSAVERIAHKHTSFN------I-----QPEHYQIVGHHLLETLRELAPDAF--------TQPVEEAWTAAYFFLAQVFIDREGALYLERKQ

Bacterial SDgbs

Acinetobacter sp.ADP1 gi|50086221|YP_047731 254aa (1-141)
Acinet*
.MTPQQIELVKSTVPVLRE.........HGVTLTSYFYNRMLNNHPELKNV..FNLD...............HQSTGRQPRALAAAVLAYAEHI...EQPEKLAKAVERITTKHVSLN......I.....QPEQYAIVGENLLHSISEVLNVPF........ESELIEAWKQAYLQLADILIEVEKEKYHQLAK
Aquifex aeolicus gi|15605769|NP_213146 139aa
Aquaeo*
MLSEETIRVIKSTVPLLKE---------HGTEITARMYELLFSKYPKTKEL--F------------------AGASEEQPKKLANAIIAYATYI------DRLDNAISTIARSHVRRN------V-----KPEHYPLVKECLLQAIEEVLNP----------GEEVLKAWEEAYDFLAKTLITLEKKLYSQP--
Bradyrhizobium japonicum gi|27377918|NP_769447 142aa
Brajap1*
-MTPEQITLIQQSFAKVAP---------ISETAAVLFYDRLFEVAPSVRAM--F------------------PEDMTEQRKKLMGMLAAVVGGL---SNLDSILPAASALAKRHVA--------Y---GAKAEHYPVVGATLLWTLEKGLGEA--------WTPELATAWTDAYGVLSGYMISEAYGAQAQAAE
Bradyrhizobium japonicum gi|27379994|NP_771523  139aa
Brajap2*
TLMTISPNPIEQSFERAAS.........RCADLTPLVYRRLFDEHPETRAM..F......................RSQGSELVKGSMLALT.I.......EAILDFAGTRSGHFRLIACEV..V.....SHDAYGTPRELFIAFFAVIRDTL.........RDLLGDAWSAEIAQAWDTLLTDIEAYVTVPA.
Campylobacter jejuni gi|15792891|NP_282714  140aa
Camjej*
-MTKEQIQIIKDCVPILQK---------NGEDLTNEFYKIMFNDYPEVKPM--FNME---------------KQISGEQPKALAMAILMAAKNI---ENLENMRSFVDKVAITHVNLG------V-----KEEHYPIVGACLLKAIKNLLNPDE----------ATLKAWEVAYGKIAKFYIDIEKKLYDK---
Chromobacterium violaceum gi|34498391| 148aa
Chrvio2*
PLTQQQIRLVQESFAKVEP---------IADEAAKLFYDKLFEYAPDLRPL--F------------------KKDMAAQRQMLMSTLKLAVKGL---DDLEQLTPVLRKLAQRHVDYG------V-----KPGDYTPVGNALLWTLKQGLGEAS-------WTQELRAAWVDAFRLMATVMKQAAYPEIQPIQD
Clostridium perfringens gi|18310935|NP_562869  144aa
Cloper*
MLDQKTIDIIKSTVPVLKS.........NGLEITKTFYKNMFEQNPEVKPL..FNMN...............KQESEEQPKALAMAILAVAQNI...DNLEAIKPVVNRIGVIHCNAK......V.....QPEHYPIVGKHLLGAIKEVLGDGA........TEDIINAWAKTYGVIAEVFINNEKEMYASR..
Desulfitobacterium hafniense gi|23119439|ZP_00102525  267aa
Deshaf*
ALNEAAAAVVRATVPALAA---------HGHAVTTAMYRRLFRDE-RIAAL--FNHA---------------NQGGDNAQAHALAGAILAYAQHI---DDLPALSGALERIAQKHVGYH------I-----LPEHYPFVAGALLGAIEETLGEAA--------TPQVLAAWGEAYWFLADLLQRREAELRGELAQ
Gloeobacter violaceus gi|37521141|NP_924518 142aa
Glovio*
TAMALQVKLLEQSFEGVKP---------NAHAFAASFYDNLFSDFPQTQAL--FA-----------------HSDMQAQQQKLLASLVLVVENL---RQPQVLSTALQDLGNRHAGYG------I-----VPEHYPMVGTSLLKTFETYLGDAW--------TPEVKQAWVDAYGAITGLMLTGAES-------
Nostoc punctiforme gi|23129187|ref|ZP_00111020.1|   126aa
Nospun*
.....MYEIMFQNHPEVKEQ..FSMA...............AQADGSQPARLATAVYSYANQI.....DNLPALKSMVEKIAHRHVQTH....V.....TPEQYPIVGESLLQAMKDVLGEAA........TEEVMAAWTEAYQALSEVFIHREHDIYVGEDK
Novosphingobium aromaticivorans gi|23110290|ZP_00096448  148aa
Novaro*
IASEHAKTIVKATAPAIEK---------HGVDITTAMYKRLFQND-EVKAM--FDQA---------------AQDSGEQPRRLAGAILAYAKNI---DKLENLGSAVQRMVQRHVDTG------V-----KAEHYPYVAEALLPAIRDVLGAEVA-------TDEVLAAWGEAYWMLADILIAAEKQAYLEAVA
Photobacteriumm profundum gi|46915547|CAG22319.1  133 aa
Photpro*
...MDIHEVFNDSYTRCNR..........NPHFMPLFYERFLEKD.ETFRSM.F................THVDMENQMKMIKASLLIIMLAA....TSEQARATVKKFGKRHGPDG......L...GIKPLDIDLWFECLIDTIKECDPA.........YNQSVEQAWRQCFNEGITIMKAECKS.......
Pirellula sp.1 gi|32471656|ref|NP_864649 166aa
Pirell*
DMTPEQVTLVKESWEKVKP---------ISEQAAELFYGRLFTLDPSLRSL--F------------------KGDMSEQGKKLMSTITLAVTSL---DRLETILPTVQALGRKHAVEYE-----V-----PDSSYATVGEALIWTLGQGLGDDF--------TEDVKEAWLLTYTTLSGAMLSGKSTAA-----
Pseudomonas aeruginosa gi|15599162|ref|NP_252656   130aa
Pseaer*
...MNAADRVMQSYGRCC..........ASTGFFDDFYRHFLASSPQIRAK..FA................TTDMTAQKHLLRAGIMNLVMYA.....RGMSDSKLRALGASHSRAALD....I.....RPELYDLWLDALLMAVAEHDRDC.........DAETRDAWRDVMGRGIAVIKSYY.........
Rhodopseudomonas palustris gi|3993677|NP_949047  141aa
Rhopal1*
MMTPQAIALVQQSFARLTP---------ISDETATLFYERLFATAPELRPL--F------------------HGDMRRQGKKLMSTLNVVVTGL---TDLTTILPATGRLARLHVAYG------V-----SAAHYTPFGAALLWTLERELGAEW--------TPELAAAWRDAYAMLSETMLAAAYGDTTPS--
Silicibacter gi|52009457|ZP_00336820  143aa
```

```
Silici*
ILRQIEVQLIKVSFNRVFA.........QKAALAEKFYHHLFLELPDAEVM..F..................TRDFSHQTEMFARVLTTGMQSL...GRDREMMVLVDDLLQRHKHLG......L....TLDQMYTAQRALHLAFCEVMQAEL.........TAAEVSAWDNAIGRLCRALAAGIEPPAA....
Thermobifida fusca gi|48835900|ZP_00292898 177aa Thefus*
RPDLRTVSLVQALCADLLD.........QPEVLADRFYHHLFRLLPSCRDL..F..................PEDMGEQHVRMARVLVEVVNHL...DEPEKTWDTLRKLGEYHYVRWG.....L.....GLEEYRCVGHALIEAARDISLEW.........APSVGSAWVTVYEWIVSAMLSGAAEAAQARSG
Vibrio parahaemolyticus gi|28900962|NP_800617  142aa
Vibpar*
...MNVIEVFNDSYERCVL..........SEHFFERFYAKFWSKG-KSFRQK.F................EGVDMQKQVRMLKGSIVFFMMTD....TSSDARKMMEKYGKKHGIGNIG....I.....KPEDIDIWFESLLEAVQECDYEY.........NSDVEKAWRACFKSGLEIMKRECGTPPSKSDT
Vitreoscilla stercoraria gi|70622|pir||GGZLB[70622]  146aa
Vitste*
MLDQQTINIIKATVPVLKE---------HGVTITTTFYKNLFAKHPEVRPL--FDMG---------------RQESLEQPKALAMTVLAAAQNI---ENLPAILPAVKKIAVKHCQAG------V-----AAAHYPIVGQELLGAIKEVLGDAA--------TDDILDAWGKAYGVIADVFIQVEADLYAQAVE

Eukaryote FHbs

Aspergillus nidulans gi|40739893|gb|EAA59083|AN3522.2  410aa
Aspnid1
ALTPQQKQIVKATIPALEQ---------HGVTITTLFYKEMLERHPELNNI--FNSS---------------HQATGVQPAALAHAVWAYASNI---DHPEALATAVSRIGHKHASLG------V-----RSEHYPIVGEHLLRAIKKVLGDAV--------TPPVADAWTAAYQQLADIFISFEDDLYKQATQ
Debaromyces hansenii gi|50428015|XP_462620.1 403aa
Debhan1
ELTPREKELIKASVPILEE---------SGDVLTSKFYNHMLTDFPEVKPF--FNES---------------NQKTMKQPKILAFALLHYAKNI---DDIAPLTAFVNQIVVKHVGLQ------V-----KPEHYPIVGQCLLDTMETLLGPEIA-------TEEFLTAWATAYGNLAQILINAEAEKYKEMEW
Dictyostelium discoideum gi|5821408|BAA83810
Dicdis1
SLSQQSISIIKATVPVLQV---------HGVNITTTFYRNMFKANPQLLNI--FNHS---------------NQREGKQQNALANTVLQAAIHI-----DKLNELNLAPIVHKHVALG------V-----LPEHYPIVGTNLLGAIKEVLQDAA--------TDEILGAWGEAYGVIAQAFIDAEAALYKVTEE
Eremothecium gossypii gi|45185029| NP_982746.1   436 aa
Eregos
ASMVSQPATPAPALDNVNTNTI------ASSLFCAQFYANLLDMDANLEKL--Y--------------------PSIKHQAVAYAGVLTTAISHL---ENLSLMDGYLSDLGKRHNRILG-----I-----EPPHFELMGIAFLKTIQDRFGIYC--------TIELTETWSRLYSYLANSILQFGIDPVLELNH
Giardia lamblia (intestinalis) gi|27981493| 458aa
Gialam
TLSEDTLRAVEATAGLIAA---------QGIEFTRAFYERMLTKN-EELKNI-FNLA---------------HQRTLRQPKALLDSLVAYALNI---AELQGFFSAAERVANKHTSFG------I-----QPAQYQIVGAHLLATIEDRITK----------DKDILAEWAKAYQFLADLFIKREEEIYAATEG
Giberella zeae gi|3551511|EAA73242 415aa
Gibzea1
ALTAAQVAIVKSTAPILKE---------HGKTITTTFYRNMLGAHPELKNY--F--------------------SLRGAQQAALANSVLAYATYI---DDL---SHAVERIAQKHVSLF------I-----KAEHYPIVGTHLIGAIGEVLGSAL--------TTEIKDAWVAAYGQLADIFIQREGQLYDAAGE
Kluyveromyces lactis gi|50304451|XP_452175.1  395aa
Klulac1
MLSVNTKNIVKATVPVLEQ---------HGVTITKIFYNNMLSENPVLLNT--FNRV---------------NQKKGRQPTALAMTVLAAAKNI---DDLSVLLPAVNEIGQKHRALQ------V-----EPAQYDIVGKYLLLAIKEVLGSAA--------TPEIIQAWGEAYKVIADIFISVEKEMYKKAAW
Neurospora crassa gi|28919225|EAA28703  418aa
>Neucra1
TLTDAQISIVKSTAPVLKQ---------HGEAITTVFYNDLITENPSLKNI--FSLT---------------SQATGAQPRALAHAVLAYATYI---DNLSALSEAVARIAHKHVSLQ------V-----EPAQYAIVGQYLIQAIGKVLGDAA--------TPEIVDAWTAAYGVLANVFIGVEGGMYEENEK
Saccharomyces cerevisiae gi|1362285|S57699|FHP_YEAST  399aa
Saccer1
MLAEKTRSIIKATVPVLEQ---------QGTVITRTFYKNMLTEHTELLNI--FNRT---------------NQKVGAQPNALATTVLAAAKNI---DDLSVLMDHVKQIGHKHRALQ------I-----KPEHYPIVGEYLLKAIKEVLGDAA--------TPEIINAWGEAYQAIADIFITVEKKMYEEALW
Schizosaccharomyces pombe gi|19115929|NC003424 427aa
Schpom1
ELNESQKQYIRSSIPILES---------SGVNLTKAFYQKMLGNYPEVLPY--FNKA---------------HQISLSQPRILAFALLNYAKNI---DDLTSLSAFMDQIVVKHVGLQ------I-----KAEHYPIVGHCLLSTMQELLPSDVA-------TPAFLEAWTTAYGNLAKILIDSEKKVYQSQPW
Yarrowia lipolytica gi|50548235|XP_502088  374aa
Yarlip1
MLTDAQIKIIKDSAPFLAE---------NGTDFAAKMYKYMFATYPEVIRF--FNQS---------------DQKNLAQPKILAHALVAYASNI---DNLGVLSDFVEQIVVKHVGLQ------I-----QPEQYPIVAASIIHTLKEMLGEAA--------TPEFIEAWTVAYTQLANILIDAEKKLYEKQAS

Cyanidioschyzon merolae CMR319C  185aa
Cyanmer
EPQQESIAIVQSTWSQAVQ---------QRERLGQVFYDRLFALYPELQPM--F------------------RSDPALQRIRLVDMVDAGVKLL---NSRRDLEQALRDLGKRHVK--------Y---GTQEEQYPIVGENLLHALESILGSKHF-------SEDMRKAWLDVYAYWSSVMLEGAREAQ-----

Thalassiosira pseudonana Scaff_137 158aa
Thapse1
GLSPEDLSLVQKSWAKVVP---------IASVAADLFYTKLFELDPELRPL--F------------------PSDLADQKKKLMAMISVAVDGL---TDLEALVPAVQDLGRRHAKY-------Y---KVTAPMFDTVGAALLDTLEKGLGEGW--------DEEHKEAWTLVYGVL-----------------

Vertebrate globins

Homo sapiens Ngb tr|Q9NPG2  151aa
HsaNgb
-----MERPEPELIRQSWR-----AVSRSPLEHGTVLFARLFALEPDLLPL--FQYN---CRQFSSPEDCLSSPEFLDHIRKVMLVIDAAVTNV---EDLSSLEEYLASLGRKHRAVG------V-----KLSSFSTVGESLLYMLEKCLGPA---------TPATRAAWSQLYGAVVQAMSRGWDGE------
Gallus gallus Ngb  gi|50749024|XP_426451.1  248aa
GgaNgb
MESGMLSRTQQALIRESWR-----RVSGSPVQHGVVLFSRLFDLDPDLLPL--FQYN---CKRFASPQECLAAPEFLDHIRKVMLVIDAAVSHL---EDLPCLEEYLCNLGKKHQAVG------V-----KVESFSTVGESLLYMLEKCLGAA--------FSPDVREAWIELYSAVVKAMQRGWEVLPE-GD-
Danio rerio tr|Q90YJ2 Ngb 159aa
DreNgb
--MEKLSEKDKGLIRDSWE-----SLGKNKVPHGIVLFTRLFELDPALLTL--FSYS----TNCGDAPECLSSPEFLEHVTKVMLVIDAAVSHL---DDLHTLEDFLLNLGRKHQAVG------V-----NTQSFALVGESLLYMLQSSLGPA--------YTTSLRQAWLTMYSIVVSAMTRGWAKNGEHKSN

Homo sapiens Cygb sp|Q8WWM9|CYGB_HUMAN 190aa
HsaCygb
ELSEAERKAVQAMWARLYAN---------CEDVGVAILVRFFVNFPSAKQY--FSQF----KHMEDPLEMERSPQLRKHACRVMGALNTVVENL---HDPDKVSSVLALVGKAHALKH-----KV-----EPVYFKILSGVILEVVAEEFASD--------FPPETQRAWAKLRGLIYSHVTAAYKEVGWVQQV
Gallus gallus Cgb  gi|56961664|NP_001008789.1  179aa
GgaCygb
EISDAEKKVIQETWSRVYAN---------CEDVGVSILIRFFVNFPSAKQY--FSQF----KHMDDTLEMERSLQLRKHAQRVMGAINTVVENL---DDPEKVSSVLALVGKAHALKH-----KV-----EPVYFKKLTGVMLEVIAEAYGND--------FTPEAHGAWTKMRTLIYTHVTAAYKEAGWVSYP
Danio rerio  Cygb gi|15387697|NP_571928  159aa
DreCygb1
SLTEEDVCVIQDTWKPVYAE---------RDNAGVAVLVRFFTNFPSAKQY--FEHF----RELQDPAEMQQNAQLKKHGQRVLNALNTLVENL---RDADKLNTIFNQMGKSHALRH-----KV-----DPVYFKILAGVILEVLVEAFPQC-------FSPAEVQSSWSKLMGILYWQMNRVYAEVGWENSK
```

```
HBA_HUMAN
VLSPADKTNVKAAWGKVGA---------HAGEYGAEALERMFLSFPTTKTY--FPHF----------DLSHGSAQVKGHGKKVADALTNAVAHVD------DMPNALSALSDLHAHKLR-----V-----DPVNFKLLSHCLLVTLAAH--------LPAEFTPAVHASLDKFLASVSTVLTSKYR--------
HBA_SQUAC
VLSAADKTAIKHLTGSLRT---------NAEAWGAESLARMFATTPSTKTY--FSKF---------TDFSANGKRVKAHGGKVLNAVADATDHLD------NVAGHLDPLAVLHGTTLC-----V-----DPHNFPLLTQCILVTLAAH---------LTELKPETHCALDKFLCEVATALGSHYR--------
HBAD_ANSAN
-LTADDKKLLAQLWEKVAG---------HQDEFGNEALQRMFVTYPQTKTY--FPHF----------DLHPGSEQVRSHGKKVAAALGNAVKSLD------NISQALSELSNLHAYNLR-----V-----DPANFKLLSQCFQVVLAVH--------LGKDYTPEMHAAFDKFLSAVAAVLAEKYR--------

HBB_HUMAN
HLTPEEKSAVTALWGKV-----------NVDEVGGEALGRLLVVYPWTQRF--FESF----GDLSTPDAVMGNPKVKAHGKKVLGAFSDGLAHLD------NLKGTFATLSELHCDKLH-----V-----DPENFRLLGNVLVCVLAHH--------FGKEFTPPVQAAYQKVVAGVANALAHKYH--------
HBB_ANSAN
HWSAEEKQLITGLWGKV-----------NVADCGAEALARLLIVYPWTQRF--FSSF----GNLSSPTAILGNPMVRAHGKKVLTSFGDAVKNLD------NIKNTFAQLSELHCDKLH-----V-----DPENFRLLGDILIIVLAAH--------FAKEFTPECQAAWQKLVRVVAHALARKYH--------
HBB_SQUAC
HWTGEEKALVNAVWTKT-----------DHQAVVAKALERLFVVYPWTKTY--FVKF--------NGKFHASDSTVQTHAGKVVSALTVAYNHID------DVKPHFVELSKKHYEELH-----V-----DPENFKLLANCLEVELGHA--------LHKEFTPEVQAAWSKFSNVVVDALSKGYH--------

MYG_HUMAN
GLSDGEWQLVLNVWGKVEA---------DIPGHGQEVLIRLFKGHPETLEK--FDKF----KHLKSEDEMKASEDLKKHGATVLTALGGILKKKG------HHEAEIKPLAQSHATKHK-----I-----PVKYLEFISECIIQVLQSK--------HPGDFGADAQGAMNKALELFRKDMASNYKELGFQG--
MYG_THUAL
----ADFDAVLKCWGPVEA---------DYTTMGGLVLTRLFKEHPETQKL--FPKF-----AGIAQADIAGNAAISAHGATVLKKLGELLKAKG------SHAAILKPLANSHATKHK-----I-----PINNFKLISEVLVKVMHEK----------AGLDAGGQTALRNVMGIIIADLEANYKELGFSG--
MYG_CHICK
GLSDQEWQQVLTIWGKVEA---------DIAGHGHEVLMRLFHDHPETLDR--FDKF----KGLKTPDQMKGSEDLKKHGATVLTQLGKILKQKG------NHESELKPLAQTHATKHK-----I-----PVKYLEFISEVIIKVIAEK--------HAADFGADSQAAMKKALELFRNDMASKYKEFGFQG--

Urochordate Hbs

Ciona intestinalis gi|34364568|emb|CAD68145  156aa
Ciona1
PFTDEELKLLRDSWDEVK--------KLGMKEVGLHIFTGLLNAAPSLRTL--FYTIDLPDEEELTIDVMRENKKVVAHATRIANAISKFIKFLD---QPDELEKLLTSLGESHARRQ------V-----DPESFEYVAPVILSVIGGH--------LKLPSNSPTLQAWVKAYGVLRNGIVSAMEA-------
Ciona intestinalis gi|34364572|emb|CAD68147  153aa
Ciona3
SLTSEQVVLLRSSWQTIG--------KLGMSNVGLAVLHRLFNDVPE--TLPFFHSV-LSPTQQTEIEVLKSNAKVVRHASRVGLSIDKIINLLD---NGEELVKYLLFLGQVHVKRS------I-----PRKYFSAMGPVLLSVISAV--------LEKDLDAPVMQAWATAYGVIEQGIIDGM---------
```