BLOSUM62 Internal Repeat Finder Database

>Nine-Haem Cytochrome C (Electron Transport) pdb 19HC Structure (London) 7 pp 119 (1999) [pdb-sequences-non-duplicates]

GEKPNPKGAAMKPVVFNHLIHEKKI
ADCETCHHTGDP
VSCSTCHTVEGKAEGDYITLDRAMHATDIAARAKGNTP
TSCVSCHQSETKER
RECAGCHAITTPKDDE
AWCATCHDITPSMTPSEMQKGIAGTLLPGDNEALAAETVLAEATVAPVSPMLAPYKVVIDALADKYEPSDFTHRRHLTSL
MESIKDDKLAQAFHDKPE
ILCATCHHRSPLSLTP
PKCGSCHTKEIDAADPGRPNLMAAYH
LECMGCHKGMAVARPRD
TDCTTCHKAAA

>Nonaheme Cytochrome C (Electron Transport) pdb 1DUW Biochemistry 40 pp 1308 (2001) [pdb-sequences-non-duplicates]

SAKPNPKGAAMKPAVFNHLAHEKKIA
NCETCHHTGDPV
ACSTCHTTEGKAEGNFVTLDRAMHATNIAKRAKGNTPV
SCVSCHEQQTKERR
ECAGCHAIVTPKRDQA
WCATCHNVTSSMTPEQMQQGIKGKLPPDQNEALAAETVLNHKPVQPLTAMQGPYKVSIDALADKYEPSNFTHRRHMASLM
ERIKGDKLAEAFHNKPET
LCATCHHRSPLSATPP
KCGSCHTKEIDPANPNRP
NLKAAYHL
QCMGCHQGMNVGRPKNT
DCTTCHKARP

>Pokeweed Lectin-C pdb 1ULK [pdb-sequences-non-duplicates]

SQWGYCGTTEEYCGKGCQSQ
CDYNRCGKEFGGKECHDELCC
SQYGWCG
NSDGHCGEGCQSQ
CSYWRCGKDFGGRLCTEDMCC
SQYGWCGLTDDHCEDGCQSQCDLPT

>Tachylectin 2 pdb 1TL2 EMBO J 18 pp 2313 (1999) [pdb-sequences-non-duplicates]

YPQNKN
DNWLARATLIGKGGWSNFKFLFLSPGGELYGVLNDKIYKGTPPTHDN
DNWMGRAKKIGNGGWNQFQFLFFDPNGYLYAVSKDKLYKASPPQSDT
DNWIARATEVGSGGWSGFKFLFFHPNGYLYAVHGQQFYKALPPVSNQ
DNWLARATKIGQGGWDTFKFLFFSSVGTLFGVQGGKFYEDYPPSYAY
DNWLARAKLIGNGGWDDFRFLFF

>Agglutinin (Lectin) pdb 7WGA J Mol Biol 215 pp 635 (1990) [pdb-sequences-non-duplicates]

Y
GYCGMGG
DYCGKGCQNGACWTSKRCGSQAGGATCTNNQCCSQY
GYCGFGA
EYCGAGCQGGPCRADIKCGSQAGGKLCPNNLCCSQW
GFCGLGS
EFCGGGCQSGACSTDKPCGKDAGGRVCTNNYCCSKW
GSCGIGP
GYCGAGCQSGGCDG

>Wheat Germ Agglutinin pdb 2WGC [pdb-sequences-non-duplicates]

Y
GYCGMGG
DYCGKGCQNGACWTSKRCGSQAGGATCPNNHCCSQY
GHCGFGA
EYCGAGCQGGPCRADIKCGSQSGGKLCPNNLCCSQW
GFCGLGS
EFCGGGCQSGACSTDKPCGKDAGGRVCTNNYCCSKW
GSCGIGP
GYCGAGCQSGGCDA

>Internalin H (Cell Adhesion) pdb 1H6U J Mol Biol 312 pp 783 (2001) [pdb-sequences-non-duplicates]

NAIKIAAGKSNVTDTVTQADLDG
ITTLSAFGTGVTTIEGVQYLNNLIGLELKDNQ
ITDLAPLKN
LTKITELELSGNP
LKNVSAIAGLQSIKTLDLTSTQ
ITDVTPLAGLSNLQVLYLDLNQ
ITNISPLAG
LTNLQYLSIGNAQ
VSDLTPLANLSKLTTLKADDNK
ISDISPLASLPNLIEVHLKNNQ
ISDVSPLANTSNLFIVTLTNQTITNQPVFYNNNLVVPNVVKGPSGAPIAPATISDNGTYASPNLTWNLTSFINNVSYTFN
QSVTFKNTTVPFSGTVTQPLTE

>Tropomyosin (Contractile Protein) pdb 1C1G Proteins: Struct, Funct, Genet 38 pp 49 (2000) [pdb-sequences-non-duplicates]

D
RADEAEADKKAAEDRSKQLEDELVSLQKKLKATEDELD
KYSEALKDAQE
KLELAEKKATDAEADVASLNRRIQLFEE
ELDRAQERLATALQ
KLEEAEKAADESERGMKVIESRAQKDEEKMEIQEI
QLKEAKHIAEDADR
KYEEVARKLVIIES
DLERAEERAELSEGKCA
ELEEEIKTVTNNLKSLEAQAEKYSQKED

KYEEEIKVLSD
KLKEAET
RAEFAERSVTKLEKSIDDLEDELYAQKLKYKAISEELDHALNDMTSI

>Sheep Prion Protein Globular Domain pdb 1UW3 J Mol Biol 336 pp 1175 (2004) [pdb-sequences-non-duplicates]

P
GQGSPGGNRYPPQGGG
GWGQPHGG
GWGQPHGG
GWGQPHGG
GWGQPHGG
GWGQPHGGGGW
GQGGTHG
QWNKPSKPKTNMKHVAGAAAAGAVVGGLGGYMLGSAMSRPLIHFGSDYEDRYYRENMHRYPNQVYYRPVDQYSNQNNFVH
DCVNITVKEHTVTTTTKGENFTETDIKMMERVVEQMCITQYQRESQAYYQRGA

>Human Prion Protein pdb 1QLX PNAS USA 97 pp 145 (2000) [pdb-sequences-non-duplicates]

GQGSPGGNRYPPQGGG
GWGQPHGG
GWGQPHGG
GWGQPHGG
GWGQPHGG
GWGQGGGTHS
QWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVH
DCVNITIKQHTVTTTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGS

>Human Gankyrin pdb 1QYM Proteins: Struct Funct 55 pp 214 (2004) [pdb-sequences-non-duplicates]

LEELKESILADKSLATRTDQDSRT
ALHWACSAGHTEIVEFLLQLGVPVNDKDDAGWS
PLHIAASAGRDEIVKALLGKGAQVNAVNQNGCT
PLHYAASKNRHEIAVMLLEGGANPDA
KDHYEAT
AMHRAAAKGNLKMIHI
LLYYKASTNIQDTEGNT
PLHLACDEERVEEAKLLVSQGASIYIENKEEKTPLQVAKGGLGLILKRMVEG

>IKBB/NF-KB Nuclear Factor pdb 1K3Z [pdb-sequences-non-duplicates]

QHEPFLDFLLGFSA
GHEYLDLQNDL
GQTALHLAAILGEASTVEKLYAAGAGVLVAERG
GHTALHLACRVRAHTCACVLLQPRPSHPRDASDTYLTQSQDCTPDT
SHAPAAVDSQPNPENEEEPRDEDWRLQLEAENYD
GHTPLHVAVIHKDAEMVRLLRDAGADLNKPEPTC
GRTPLHLAVEAQAASVLELLLKAGADPTARMYG
GRTPLGSALLRPNPILARLLRAHGAPEPEDGGDKLSPCEEEGEDEDSDNRDEGDEYDD

>Troposmysin pdb 2TMA Nature 278 pp 413 (1979) [pdb-sequences-non-duplicates]

D
RAEQAEADKKAAEDRSKQLEDELVSLQKKLKGTEDELD
KYSEALKDAQE
KLELAEKKATDAEADVASLNRRIQLVEE
ELDRAQERLATALQ
KLEEAEKAADESERGMKVIESRAQKDEEKMEIQEI
QLKEAKHIAEDADR
KYEEVARKLVIIES
DLERAEERAELSEGKCA
ELEEEIKTVTNNLKSLEAQAEKYSQKED
KYEEEIKVLSD
KLKEAET
RAEFAERSVTKLEKSIDDLEDELYAQKLKYKAISEELDHALNDMTSI

>Human Interferon Gamma (Immune System) pdb 1EKU J Mol Biol 299 pp 169 (2000) [pdb-sequences-non-duplicates]

HSDVADNGTLFLGILKNW
KEESDRKIMQSQIVSFYFKLFKNF
KDDQSIQKSVETI
KEDMNVKFFNSNKKKRDDFEKLTNYS
VTDLNVQRKAIDELIQVMAEFSTEEQQEGPYV
KEAENLKKYFNAGHSDVADNGTLFLGILKNW
KEESDRKIMQSQIVSFYFKLFKNF
KDDQSIQKSVETI
KEDMNVKFFNSNKKKRDDFEKLTNYS
VTDLNVQRKAIHELIQVMAELSPAAKTGKRKRSQMLFRG

>Alkaline V8 Protease pdb 1WCZ [pdb-sequences-non-duplicates]

APVTYIQVEAPTGTFIASGVVVGKDTLLTNKHVVDATHGDPHALKAFPSAINQDNYPNGGFTAEQITKYSGEGDLAIVKF
SPNEQNKHIGEVVKPATMSNNAETQVNQNITVTGYPGDKPVATMWESKGKITYLKGEAMQYDLSTTGGNSGSPVFNEKNE
VIGIHWGGVPNEFNGAVFINENVRNFLKQNIEDINFANDDQ
PNNPDNPDN
PNNPDNPNNPDE
PNNPDNPNN
PDNPDNGD
NNNSDNPDAA

>Ribosomal Subunit pdb 1FFK Science 289 pp 905 (2000) [pdb-sequences-non-duplicates]

DITVEGDNGSVTRRLWYPDI
DVSVDGD
TVVIESDEDNAKTMSTIGTFQSHIENMFHGVTEGWEYGMEVFYSHFPM
QVNVEGDEVVIENFLGEKAPR
RTTIHGDT
DVEIDGE
ELTVSGPDIEAVGQTAADIEQLTRINDKDVRVFQDGVYITRKPNRGDA

>Rack1 Yeast 80S Ribosome pdb 1TRJ [pdb-sequences-non-duplicates]

VTSLATSAGQPNLLLSASR
DKTLISWKLTGDDQKFGVPVRSFKGHSHIVQDCTLTADGAYALSASW
DKTLRLWDVATGETYQRFVGHKSDVMSVDIDKKASMIISGSR
DKTIKVWTIKGQCLATLLGHNDWVSQVRVVPNEKADDDSVTIISAGN
DKMVKAWNLNQFQIEADFIGHNSNINTLTASPDGTLIASAGK
DGEIMLWNLAAKKAMYTLSAQDEVFSLAFSPNRYWLAAATATGIKVFSLDPQYLVDDLRPEFAGYSKAAEPHAVSLAWSA
DGQTLFAGYT
DNVIRVWQ

>50S Ribosomal Protein L6P pdb 1W2B To Be Published [pdb-sequences-non-duplicates]

ITVEGDNGSVTRRLWYPDID
VSVDGDT
VVIESDEDNAKTMSTIGTFQSHIENMFHGVTEGWEYGMEVFYSHFPMQ
VNVEGDEVVIENFLGEKAPRR
TTIHGDTD
VEIDGEE
LTVSGPDIEAVGQTAADIEQLTRINDKDVRVFQDGVYITRKPNRGDA

>Autolysin Choline Binding Domain pdb 1HCX Nature Structural Biology 8 pp 1020 (2001) [pdb-sequences-non-duplicates]

IN
GTWYYFDSSGYMLADRWRKHTD
GNWYWFDNSGEMA
TGWKKIA
DKWYYFNEEGAMK
TGWVKYK
DTWYYLDAKEGAMVSNAFIQSADG
TGWYYLKPDGTLADRPEFTVEPDGLITVK

>Ubiquitin Ligase Complex pdb 1FQV Nature 408 pp 381 (2000) [pdb-sequences-non-duplicates]

IFSCLCLPELLKVSGVCKRWYRLASDESLWQTLDEFR
VQHMDLSNSVIEVSTLHGILSQCSK
LQNLSLEGLRLSDPIVNTLAKNSN
LVRLNLSGCSGFSEFALQTLLSSCSR
LDELNLSWCFDFTEKHVQVAVAHVSET
ITQLNLSGYRKN
LQKSDLSTLVRRCPN
LVHLDLSDSVMLKNDCFQEFFQLNY
LQHLSLSRCYDIIPETLLELGEIPTLKTLQVFGIVPDGTLQLLKEA
LPHLQINCSHFTTIAR

>Human Angiostatin pdb 1KIO [pdb-sequences-non-duplicates]

KNGITCQKWSSTSPHRPRFSPATHPSEGLEEN
YCRNPDNDPQGPWC
YTTDPEKRYDYCDILECEEECMHCSGENYDGKISKTMSGLECQAWDSQSPHAHGYIPSKFPNKNLKKN
YCRNPDRELRPWCFTTDPNKRWELCDIPRCTTPPPSSGPTYQCLKGTGENYRGNVAVTVSGHTCQHWSAQTPHT
HERTPENF
PCKNLDEN
YCRNPDGKRAPWCHTTNSQVRWE
YCKIPSC

>Human Kelch domain of KEAP1 pdb 1U6D J Biol Chem 279 pp 54750 (2004) [pdb-sequences-non-duplicates]

APKVGRL
IYTAGGYFRQSLSYLEAYNPSDGTWLRLADLQVPRSGLAGCVVGGL
LYAVGGRNNSPDGNTDSSALDCYNPMTNQWSPCAPMSVPRNRIGVGVIDGH
IYAVGGSHGCIHHNSVERYEPERDEWHLVAPMLTRRIGVGVAVLNRL
LYAVGGFDGTNRLNSAECYYPERNEWRMITAMNTIRSGAGVCVLHNC
IYAAGGYDGQDQLNSVERYDVETETWTFVAPMKHRRSALGITVHQGR
IYVLGGYDGHTFLDSVECYDPDTDTWSEVTRMTSGRSGVGVAVT

>Collagen-Binding Domain of Adhesin pdb 1P9H [pdb-sequences-non-duplicates]

TAVQISPNADPALGLEYPVRPPVPGAGGLNASAKGIH
SIAIGATAEAAKGA
AVAVGAG
SIATGVN
SVAIGPL
SKALGDSAVTYGAASTAQKD
GVAIGARASTSDT
GVAVGFNSKADAKN
SVAIGHSS
HVAANHGY
SIAIGDRSKTDREN
SVSIGHESLNRQLTHLAAGTKDTDAVNVAQLKKEIEKTQENTNKRSAELLANANAYAD

>Nerve Growth Factor Neurotrophin Receptor pdb 1SG1 To Be Published [pdb-sequences-non-duplicates]

NLGEGVAQPCGANQTV
CEPCLDNVTFSDVVSATEP
CKPCTECLGLQSM
SAPCVEADDAVCRCAYGYYQDEETGH
CEACSVCEVGSGLVFSCQDKQNTV
CEECPEGTYSDEANHVDP
CLPCTVCEDTERQLRE
CTPWADAECE

>Hisactophilin (Actin Binding) pdb 1HCE Nature 359 pp 855 (1992) [pdb-sequences-non-duplicates]

A
VKTHHGHHDHHTHFH
VENHGGKVA

LKTHCGKYLSIGDHKQV
YLSHHLHGDHSLF
HLEHHGGKVS
IKGHHHHYI
SADHHGHVS
TKEHHDHDTTFEEIII

>LDL Receptor pdb 1IJQ [pdb-sequences-non-duplicates]

SEYTSLIPNLRNVVALDTEVAS
NRIYWSDLSQRMICSTQLDRAHGVSSYDTVISRDIQAPDGLAVDWIH
SNIYWTDSVLGTVSVADTKGVKRKTLFRENGSKPRAIVVDPVH
GFMYWTDWGTPAKIKKGGLNGVDIYSLVT
ENIQWPNGITLDLLS
GRLYWVDSKLHSISSIDVNGGNRKTILEDEKRLAHPFSLAVFE
DKVFWTDIINEAIFSA
NRLTGSDVNLLAENLLSPEDMVLFHNLTQPRGVNWCERTTLSNGGCQYLCLPAPQINPHSPKFTCACPDGMLLARDMRSC
LT

>Beta-Lactamase Inhibitor pdb 1JTD [pdb-sequences-non-duplicates]

VPAEAQSGVDAIAGGYFHGLALKG
GKVLGWGANLNGQLTMPAATQSGVDAIAAGNYHSLALKD
GEVIAWGGNEDGQTTVPAEARSGVDAIAAGAWASYALKD
GKVIAWGDDSDGQTTVPAEAQSGVTALDGGVYTALAVKN
GGVIAWGDNYFGQTTVPAEAQSGVDDVAGGIFHSLALKD
GKVIAWGDNRYKQTTVPTEALSGVSAIASGEWYSLALKN
GKVIAWGSSRTAPSSVQSGVSSIEAGPNAAYALKG

>Designed TRP repeat pdb 1NAO Structure 11 pp 497 (2003) [pdb-sequences-non-duplicates]

YKQGD
YDEAIEY
YQKALELDPNNAEAWYNLGNAYYKQGD
YDEAIEY
YQKALELDPNNAEAWYNLGNAYYKQGD
YDEAIEY
YQKALELDPNNAEAKQDLGNAKQKQG

>Polygalacturonase Inhibiting Protein Leucine Rich Repeat pdb 1OGQ [pdb-sequences-non-duplicates]

NPTT
LSSWLPTTDCCNRTWLGVLCDTDTQTYRVNNLD
LSGLNLPKPYPIPSSLANLPYLNFLYIGGINN
LVGPIPPAIAKLTQLHYLYITHTN
VSGAIPDFLSQIKTLVTLDFSYNA
LSGTLPPSISSLPNLVGITFDGNR
ISGAIPDSYGSFSKLFTSMTISRNR
LTGKIPPTFANLNLAFVDLSRNMLEGDASVLFGSDKNTQKIHLAKNSLAFDLGKVGLSKNLNGLDLRNNR
IYGTLPQGLTQLKFLHSLNVSFNN
LCGEIPQGGNLQRFDVSAYANNKCLCGSPLPACT

>Dictyostelium Filamin pdb 1WLH [pdb-sequences-non-duplicates]

CFQPSKFKIHAVDPDGVHR
TDGGDGFVVTIEGPAPVDPVM
VDNGDGTYDVEFEPKEAGDYVINLTLD
GDNVNGFPKTVTVKPAPSAEHSYAEGEGLVKVFDNAPAEFTIFAVDTKGVAR
TDGGDPFEVA
INGPDGLVVDAKV
TDNNDGTYGVVYDAPVEGNYNVNVTLRGNPIKNMPIDVKCIEGANGEDSSFGSFTFTVAAKNKKGEVK
TYGGDKFEVSITGPAEEITLDAIDNQDGTYTAAYSLVGNGRFSTGVKLNGKHIEGSPFKQVLGNPGKKNPEVKSFTTTRT
AN

>Mycobacterium Tuberculosis PKND pdb 1RWI J Mol Biol 339 pp 459 (2004) [pdb-sequences-non-duplicates]

IDFRLSPSGVAVDSAGNVYVTSEGM
YGRVVKLATGSTGTTVLPFNGLYQPQGLAVDGAGTVYVTDF
NNRVVTLAAGSNNQTVLPFDGLNYPEGLAVDTQGAVYVADRG
NNRVVKLAAGSKTQTVLPFTGLNDPDGVAVDNSGNVYVTDTD
NNRVVKLEAES
NNQVVLPFTDITAPWGIAVDEAGTVYVTEHN
TNQVVKLLAGSTTSTVLPFTGLNTPLAVAVDSDRTVYVADRG
NDRVVKLTSLEHHHHHH

>De Novo Protein [pdb-sequences-non-duplicates]

AQELAKKGG
GEELLKQALQQ
AQQLLQQ
AQELAKKGGG
GEELLKQALQQ
AQQLLQQ
AQELAKKGG
GEELLKQALQQ
AQQLLQQ
AQELAKK

>gi|30250173|ref|NP_842243.1| hypothetical protein [Nitrosomonas europaea ATCC 19718]

MIKILLLLTSVLVAMPVAAVDVAPRISDREIIESLAELKAGQKALEEKM
DLRFNAMQEQI
DQRFTAI
DQRFTAMQEQM
DQRFTAV
DQRFTAV
DQHFTAMQKQI
DQRFIAV
DQRFEAI
DRRLDFIQQLMLVTIAGIFGLIGFIIWDRYSTLRPM
DMRLQRLEEDLERDLELQSPEGSKLTRLIHALRELAKE

DKKVEAILRSFSLL

>gi|30250382|ref|NP_842452.1| hypothetical protein [Nitrosomonas europaea ATCC 19718]

MKTTLIKVIAASVTALFLSMQVYASGHTAHVD
EAVKHAE
EAVAHGKEGHTD
QLLEHAK
ESLTHAKAASEAGGNTHVG
HGIKHLE
DAIKHGEEGHVG
VATKHAQ
EAIEHLRASEHKSH

>gi|53804596|ref|YP_113555.1| hypothetical protein MCA1082 [Methylococcus capsulatus str. Bath]

MNVLNRGIIKGLLTAAFATMLG
LSAAQAAEDHKAEA
LKHAEAAVAAGNAAGV
AEHAEAA
KTHAAAANAEKKSPHYDAAITSLNSAIEHGK
MGHADIA
KKSAEEA
VTHLKAAP

>gi|13507876|ref|NP_109825.1| hypothetical protein [Mycoplasma pneumoniae]

MEKKPWEEDISIEEFKKSLNKDKITNLIIKRRWNKGKSTYHLSFNGDFEVVTKKPSTKYVTHKQLDQKLKEFKQDLMVEL
HDTFATKADLRDSEARINQKLEALVQVVLLHGEQINKLTQIVEKQGEQIRELQVEQKAQRQEFNA
RMDRLENL
LVESIESTNK
RFDSIEG
RLDSMDS
RLDSMEN
RLDSIEG
RLDSVEG
RLDSVEG
RLDSMEN
RLDSMET
RLDKIDPPK

>gi|39935203|ref|NP_947479.1| hypothetical protein RPA2134 [Rhodopseudomonas palustris CGA009]

MTKTALVLGTVATLAVASVAAPAPAEAR
NGIGPFIGGLAAGALIAGVASNAYA
YGPGYGY
YGGGPYYGGGPY
YGYAPRYVG
YYGGPAYYGGGY
GWGGPRYYRPRPVY

YGYRSYYRP
YRGGYYGAPVIRAGYGPG
WGGHRHWHHHRHW

>gi|22127541|ref|NP_670964.1| hypothetical [Yersinia pestis KIM]

M
VLILLPV
VLTLLPV
VLILLPT
VPILLPV
VEILHPV
VLVLLPV
MLTELPV
VEVLLPV
VLTLQPV
VEVLLPV
VLVLTPV
VLTLLPV
LLTLLPVKEVLSPV
MLLLEPV

>gi|15836656|ref|NP_297344.1| hypothetical protein [Xylella fastidiosa 9a5c]

MY
SDITTHTL
THSHTHTL
THSHTHTL
THSHTHTL
THSHTHTL
THSHTHTL
THSHTHTL
THSHTHTL
THSHTHTL
THSHTHTL
THSHTHTL
THSHTHTL
THSHTREHDFQPFQINSPITETI

>gi|15837228|ref|NP_297916.1| hypothetical protein [Xylella fastidiosa 9a5c]

MWSNVRASSHSISSMLHITY
PHPSALV
THTSHLI
PHTSYLI
PHTSHLT
PHTSHLT
PHTSHLT
PHTSHLT
PHTSHLT
PHTSHLT
PHTSHLT
PHTSHLT
PHTSHLT

PHTSHLT
PHTSHLT
PHTSHLT
PHTSHLT
PHTSHLT
PHTSHLT
PHTSHLT
PHTLHLT
PHTSHLT
PHTSHLT
PHTSHLT
PHTSHLT
PHTSHLT
PHTSHQGL

>gi|20515634|gb|AAM23913.1| conserved hypothetical protein [Thermoanaerobacter tengcongensis]

MTNEEFMNFIV
EKFNSFE
KRFDNFERRLDEFE
ERFSAFE
KRLEKIE
KRLDGFE
ERFEDVE
KRFENVE
RSFGNLE
KKVEGVE
KRLQDVE
KRVENVE
IRLESVENLVKATFEQTASNSEAITEIKIKLD
KKLESYDKTITELKQKSEEAELDIQLLKKLVVNL

>gi|56551689|ref|YP_162528.1| hypothetical protein ZMO0793 [Zymomonas mobilis subsp. mobilis ZM4]

MSYGLIGNWVMRFRLKGLASLLILFSSVSFAR
GYYSECKISG
GSKFSCGGWFTGKAVVRDR
GRYSECSISG
GSKFSCGGWFTGKAVVRDR
GRYSECSISG
GSKFSCGGWFTGKAVVKDGGDYLKCNLSNGWGNCQGWYTGQAVVKDR
GDYHLCKIED
GEKISCDNAYSGKAVLKEN
GQFHLCTIEN
GYKHSCSRDGYSGEAVSDDE

>gi|21221402|ref|NP_627181.1| hypothetical protein SCE59.16c [Streptomyces coelicolor A3(2)]

MDRTRARAVMRGAGRARYGRTASPGPRADRQPAPPTTPRRVPADRPSADRPPTGEAPVTPGGRWPVGWGVPWGAVACGAP
GRVGGRVSSGPRFREA
TGVLGAR

CSVLGAR
CSVLGGR
WSVLGAR
CSVPGGRSGGVCRGAPWRPGRRAVSVVACRRACGSGRLRV
SSVLGAR
CSVLGAR
CSVLGAR
CPVPGGR

>gi|53802353|ref|YP_112958.1| hypothetical protein MCA0428 [Methylococcus capsulatus str. Bath]

MEVMDEEG
DSCMLCHSFAPNAMTDESKLKPLTTIANEPLK
AVCHDCHVTEQRGP
WRCNLCHDDKTKIWPEDHRFGYVDHHGEAARRDE
AACKECHLDL
AFCTNCHFRRDTMGTGYHPLGYRTLHGLGVRVDT
LDCGRCHNQ
IYCETCHARTRFQ

>gi|33865518|ref|NP_897077.1| putative secreted calcium-binding protein [Synechococcus sp. WH 8102]

MIIGSVEIT
GGSGSNVVYADDSAQIIV
LGGGNDTI
GSAGGDDLLI
GGRGQDLIT
GGADNDTLKGGPQTDLLF
GGQGRDHLN
AGRGDDTLK
GGRGQDTLR
GSSGDDLLKGQLYHDSLH
GGSGDDTLF
GGKGQDTLT
GAKGSDTFK
LSKGNDTIRDFSITEGDLINAPNNLNLQLIQRGNHLLLKDSEHNIKTTLLNINHDDLITYQPDLI

>gi|15837283|ref|NP_297971.1| hypothetical protein [Xylella fastidiosa 9a5c]

MTQKNALMISAATPATSTPSIHLSRSISPFISAK
RVSTCAKRWSM
VFSTCAKRWSM
VFSTCAKRWSM
VFSTCAK
RWSTCAKPSFISASRRAKALPFLDSDSARLCKF
ISNTSAKASASASACAAGTPA
VFSRFANLSVSNATDVTHAPVLTPCDGECNTMISGASA

>gi|15898330|ref|NP_342935.1| Conserved hypothetical protein [Sulfolobus solfataricus]

MADVFIQMLSDAVTEYAKSLVEKSKQGKKLSEQDYLFLMMYLNSKRTDDLRQD
MNKKFDDLKSY
VDKRFDDTKAEIISVKEDIKGLRDY
TERKFDESRAY
TDKKFEE
TNKKVEDLRHY
VDNKFEETRVY
TDKKFEE
TNKKVEDLRNY
VDTKFQELRTY
TDKKFEESRAY
TDVQFVKFKEY
VDKQFEH
VDTRFNDLRSD
MDSKFNELKNDLNLIIKLLVGKEKEEEKS

>gi|20515124|gb|AAM23448.1| predicted divalent heavy-metal cations transporter [Thermoanaerobacter tengcongensis]

MSNFEILMI
GAFAGVAGTGLGGVFAYFFKNPNPKFFSGIMGITA
GLMLSVVMFDLLPHSFEIS
GMPVGVI
GVIIGAFLISLFDSHIEKTKFNKGFIRE
GLLLGIAIALHNFPE
GLAVGSGFMVSSSLGIDIALVIALHDFPEGVAVAVPLSAGGVSPCKVLLYTFLTGLPTAL
GTFIGILSGGISNTFI
GLNLALAGGAMLYVTCGEIIPEARDIYRGKISIL
GMILGVIGGIIITRYL

>gi|5104897|dbj|BAA80211.1| 141aa long hypothetical protein [Aeropyrum pernix]

MEGWPYTMPSLSWWIQRPTTPAPRTSSTTCFI
ALLKASSVSIPNLWAIVFIILIMLTRACSTRRTRSESLLNSSS
SLWKRCS
SLRWRSS
SFWNRCS
SLSWRSS
SLWKRSSRWSCRCS
SFWKRSSRSRNLSS
SLWNLCSSLR

>gi|15964815|ref|NP_385168.1| HYPOTHETICAL PROTEIN [Sinorhizobium meliloti]

MGTGAVDSSPDGIILGAGGGLIALDIQWYAPTGAQGRNSLAGGTPAQCRELYEAEEVPGIGFVDGHALIVARD
YRFAASGPGV
DGFAASGSGV
DGFAASGPRV
DGFAASSLRV
DGFAASGPGV
DGFAASGPRV
DSLTASGPRV
DGFAASGPGV
DGFAASGPRI

DGFAASGPRV
DGFAASCPRV
DGFAASGSRV
DGLTAFGLWI
DSLLASGVRAGIVFG

>gi|22297568|ref|NP_680815.1| ORF_ID:tll0024~hypothetical protein [Thermosynechococcus elongatus BP-1]

MTNPQHLQLLKQGVTAWNQWREQNSDIRPDLGQA
DLTGANLELY
NLTNANLDQA
NLAGADLRNALLKDA
YMAGANLYMSDLIEA
DLQGACLQRA
TLVGADLYKSNIAMA
DLRQANLVGALLRRVSLVEATLAYA
NLTRANLFQANLYLA
DLKGAILQEAILESASLIEANFESAVLIAAKMAKANARRA
NFKGANLYKAEMDDM
DLRNAILDKA

>gi|33862645|ref|NP_894205.1| Hemolysin-type calcium-binding protein [Prochlorococcus marinus str. MIT 9313]

MIAKVDGNLSGEYDGLDEGDSIGRFKSDNGGTLYLFITDKGGD
GNDIELYTKSFSGVLPESLRDPRMIVSD
ADDSLAGAS
ADEVIFGGR
GDDVLLGGA
GDDQVAGGN
GDDWLYGGF
GDDTLKGDRGADTYWLSR
GVDVIIVFSFAENDRISV
GDGVLCSLRQA
GDDLSIRADGIHTTLLDVDKDEFLAADVIDYI

>gi|15602077|ref|NP_245149.1| unknown [Pasteurella multocida]

MKKLATLAALAGALTMAATAPAQAEMKSAGKAE
QCVEGKCLKTKAAEG
KCGEGKCGAHDHKAKAAEG
KCGEGKCGAHDHKA
KAAEGKCGAHDHKAKAAEG
KCGEGKCGSK

>gi|22299550|ref|NP_682797.1| ORF_ID:tlr2007~hypothetical protein [Thermosynechococcus elongatus BP-1]

MEAKLEQRINT
IEAKLQQQINA
VEAKLEQRINA

VEVKLQQQINA
VEAKLEQRINA
VEAKLEQQINA
VEAKLGQRIDT
LESTLEDISKRQAGTDARLWGFLVTLLVATVGLLTKVLFFDLPRP

>gi|37525717|ref|NP_929061.1| hypothetical protein [Photorhabdus luminescens subsp. laumondii TTO1]

MIRPQVAIFVTV
HCRTGSLEKGSGPKGKPQLV
HCRTGSLEKSNSSQYRYPYV
HCRTGSLESKQTKYRSAALV
HCRTGSLEK
TCSLGPLIKCV
HCRTGSLEMHPSTLSPLTIV
HCRTGSLER

>gi|12724189|gb|AAK05317.1|AE006354_2 UNKNOWN PROTEIN [Lactococcus lactis subsp. lactis]

MLIIQ
SQLLTEL
SAMLTVLLST
TKMLTIQ
SQLLTEL
SAALTVLLST
TKMLTIQ
SQLLTEL
SAALTVLLST
TKMLTIQ
SHLLTEL
SAALTVLLST
TKMLTIQ
SQLLTEL
SAALTVLLST
TKMLTIQSHFTDNFSGLLLSFL

>gi|16329677|ref|NP_440405.1| hypothetical protein [Synechocystis sp. PCC 6803]

MKTMDAQEILRCYAQGQR
DFSHINLVRVCLSNA
SLIGAQLIFV
DLGGANLTRAQLDSA
TLKNANLALANMTEVCLIYADLSNA
DLSGANLVGADLTNA
DLSGAKLGGA
DLRKANLSEASLRGA
DLRGVNLIEANLTNTDFSEA
DLTGAYISDGAVINVVNLS

>gi|16330623|ref|NP_441351.1| HglK [Synechocystis sp. PCC 6803]

MAGLAWAKDYHRQWVKMNYIMSNGLSPTPINAPGWLNSGVASLLVNGHVPQGLELTNQNLAGLVFPHGDLSQVALIGCDL
RFA
NLEGADLTDANLIAASLHKS
NLRRANLCRATLNRC
NLSEADLTESDANEALFCQAVFTEVEAHGLRLYRAKVSQA
QLMGAHLHQAYAPEADFSAVAAIAV
DLRWANLRKT
NFRGADLRYG
NFRGANLTQA
DFTGANLKGA
NLRGANLVGT
NLQRADLSDVTW

>gi|15827217|ref|NP_301480.1| hypothetical protein [Mycobacterium leprae]

MKVFDRFNTAILVSAGLCGAALAFSPSAAAA
PLSMPTGGPTCIEQMAGLGAVPAA
VPAVLPGPVAGAVPS
VPVVPSGGF
PPVVPPGGF
PPVVPPGGF
LPVVPPGGF
PPVVPPGGF
PPVVPPGVAPAVGPLLAEDSAALGSGILAGKGVPTASPQAALSGLVVLPGPPASVEA
PPAAAPVVAAATPLLCCGEAAPPVR

>gi|13507839|ref|NP_109788.1| hypothetical protein [Mycoplasma pneumoniae]

MKEKIPFYNEKEFHDMVKKTKKGTFSGWYIIDKDNKSVEFSGSFNRQFKLNKPVIPVNTEYVTRKEFNEYKVSNDQRLTK
IETTLA
AQGEQINKLTQTVE
KQGEQINQLVQVVL
LHGEQINKLTQTVE
KQGEQIKELQVEQK
AQGEQIK
AQGKQIK
AQGKTLKSILQALGGINKRLDKIDPPK

>gi|22299871|ref|NP_683118.1| ORF_ID:tlr2328~hypothetical protein
[Thermosynechococcus elongatus BP-1]

MAELPSKAQSINPVHQAMTKAHTPEDVWRLLGELLEAQKE
TERRFQE
TERRFQE
TERRFQE
TDRRFQE
TDRRFQE
TDRRFQQ
TERLLREEARQLNQQLGKLGNRLGEFVEWQVRPAVLRLFQSRGIAVSQLYSDVILQDGNESLEVDLLVVNTDEAVLVEVK
TKLGQGDVDEHLERIAKFRRLAHQYRGTKLLGAVAGMIVPPEVARYAYRQGLFVLAQSGDGVTIL
NDPQFQPRAW

>gi|13357658|ref|NP_077932.1| unique hypothetical [Ureaplasma urealyticum]

MAKYIKTGVSYINLDNARTINILPEDIDSYLELGGDEAYKTSDLGSELYINYADFESTNILFDLKKEELQAKIDAFLISN
DTVLDLSEVFLDIHFSEDDDY
EEDCCCE
DECCSDE
ENEACCNSEVKV
EEECCGGA
KDDCCGGHEH
EHEVCCDSETKTSET
QEECCGGT
KDDCCGGHEHEHHHHHSHQH

>gi|26554205|ref|NP_758139.1| conserved hypothetical protein [Mycoplasma penetrans]

MTKNNKENKNNKVSKPFTGRYSEKEFASIIKGCKFVDVKNFNVSFDSYEDDKTGDIIETSVFEGNMKIYKTKRHPKPIEV
PPPSDDDKKPTDLSNEIKFIKDTLGIILSDMKEMKHDISVLKDDMVEVKSDIKILKEKVTVLENDVSILK
KDVSEHS
KILQDHS
KILQDHS
KILQDHS
KILQDLQFN

>gi|52784354|ref|YP_090183.1| GsiB [Bacillus licheniformis ATCC 14580]

MANNNDDNKMSREEAGRKGGEKTSQ
NHDKEFYQEIGQKGGEATSQ
NHDKEFYQEIGEKGGEATSR
NHDKEFYQEIGEKGGEATSR
NHDKEFYQEIGEKGGEATSK
NHDKEFYQEIGKKGGEATSR
NHDKEFYQENGEKGGRQRRSG

>gi|13507877|ref|NP_109826.1| hypothetical protein [Mycoplasma pneumoniae]

MKEKIPFYNEKEFNEMMKKTKKGTFSGWYIINPENKSVEFSGNFNRQFKLNKPVIPVNTEYVTRKEFNEYKDSND
QRLTKIE
NKVDKLE
VKVDKLE
EKVDKLE
AKVDKLE
EKVDKLE
AKVDKLE
EKVDKLEAKVDSGFEMLAKILAAIN
KRLDSIE
GRLDKIEPPK

>gi|34496687|ref|NP_900902.1| hypothetical protein CV1232 [Chromobacterium violaceum ATCC 12472]

MGFNPRARVGRDGILGAQYLEAWLF
QSTRPRWARLDQMAVIGAAAAF
QSTRPRWARLSASTSCMLYTQF

QSTRPRWARPPVLARGTNSTMF
QSTRPRWARRGQGGSPGRSVHVSIHAPVLGATRNRVQGITKLSVSIHAPALGATCCASVMSWGRQF
QSTRPRWARRHGIQARRRRVEF
QSTRPRWARRTVPSSIKE

>gi|16080660|ref|NP_391488.1| morphogenetic protein [Bacillus subtilis]

MGHYSHSDIEEAVKSAKKEGLKDYLYQEPHGKK
RSHKKSHRTHKKS
RSHKKSYCSHKKS
RSHKKSFCSHKKS
RSHKKSYCSHKKS
RSHKKSYRSHKKS
RSYKKSYRSYKKS
RSYKKSCRSYKKS
RSYKKSY
CSHKKKS
RSYKKSC
RTHKKSY
RSHKKYYKKPHHHCDDYKRHDD
YDSKKEYWKDGNCWVVKKKYK

>gi|28211188|ref|NP_782132.1| hypothetical protein [Clostridium tetani E88]

MRCIMKDYMMCKCMKDCYMDSMCNDDKMDRPLCNMCGLTPLCPPNYIMPTAKDDKKDMDCKDMTDIRPMYMEQNTCD
MYDDMDYCHCMHDC
MHHCMNPGMDHC
MHHDMNPGMDHC
MHHGMNPGMDHC
MHHGMNPGMDNC
MHYDMNPG
MHHDMHCCMPPY
MHSEMMCHPMMMCHPMMW
GHGGMHPMMG
GYPGMNPMMCGCDGMHPMMYSRELDFEEFDLDED

>gi|15606618|ref|NP_213998.1| putative protein [Aquifex aeolicus]

MMKFLLILIFLASFSFSLTPEEEKQLLKDIAEIKTTLKAFIREA
DERFEQV
DKRFEDV
NRRFEDF
NRRLNDLREDM
NKRFELV
DKRFIELREDM
NKRFELV
DQRFEQLYTFLWIITGIFTTLTASVIAFAWWDRRTIIRKTK
EETFEEMEKELKPEKFRKIMNALREKAKT
DKELEAILKKYGLL

>gi|15898542|ref|NP_343147.1| Hypothetical protein [Sulfolobus solfataricus]

MADDILYRTLGELLSNYVKTLLEKSGQGKRLSEQEYLILMNYYNSKN

MDYVFTELRNIRNELRGI
MEIIDRKSKEA
IDYADKSRKEA
MEYAEKNKKEV
IDYVDKTKKEM
LDYVDKR
FEAIDKR
FEAIDKRFDDLNNRIDDLRELILTLAKKT

>gi|52786096|ref|YP_091925.1| hypothetical protein BLi02352 [Bacillus licheniformis ATCC 14580]

MYY
YEPAKKDYWDYD
YDPYKKD
HDYCKKDFYY
HEPCKKNYYY
HEPCKKDFYY
HEPCKKD
YDYCKKDYDY
HEDYKKDYDY
HKPCKKD
YDYCKKDFYY
HEPCKKDYDY
HKPCKKDCY
YEPRKKCYDSSYKKKNDSWIY

>gi|33865139|ref|NP_896698.1| conserved hypothetical [Synechococcus sp. WH 8102]

MNRSLIAFQALIASSAVLCQTAEASATYSSPDQLNNSIEARIDSV
RNGDWSSLLKTIDNDGELV
AKSKWGNG
GGHKFSNSY
GSGKWKNGK
GGNKWSN
SRNTWGNGG
YHGGWRN
GGGAWGN
GGGAWGN
GGGGFVNW

>gi|15805145|ref|NP_293831.1| hypothetical protein [Deinococcus radiodurans]

MG
DEAGNAQDRAE
AAAARARAESHGADYRREMDSANDDLLSGLGQAKHKTD
QAAEKVRADTQ
EAAQNAR
EKAQDVRANVH
ESAQDFRAGAQ
EQAQTLRADAR
DAAAQAR
DHAQNLRTDVH
EKAQDLREGAH

ETAQNVKADVH
NAAEDVRADVREIDRRSKNQD

>gi|26988133|ref|NP_743558.1| hypothetical protein [Pseudomonas putida KT2440]

MDYLALFFALFNQIIRRRIKPIETNNSQTPGATTMKNKLILTLALSVFAAGAFA
EDGFDRTDAHTFAAVQTQSPRYA
EDGFDRTGGQRFA
EDGFDRTGGQRFA
EDGFDRTGGQRFA
EDGFDRTGGQRFA
EDGFDRTGGQRFA
EDGFDRTGGQRFA
EDGFDRTNAHRIS

>gi|20515642|gb|AAM23921.1| Acetyltransferases (the isoleucine patch superfamily) [Thermoanaerobacter tengcongensis]

MNYISEKAK
IGQNVKIGYFTVIEDNVV
IGDNCVIGNNVTIYKGSIIGNNVR
IDDNVVIGKQPMRAATSIFKDKQEKPPCK
IGDDCIIGTSAVIYAGCE
IGKKCLIADLATVREDVV
IGDMTIVGRGVAIENYCK
IGSRCKIETNAYITAYSELEDEVFIAPCVATSNDNSAGRDPDRFSKMKGVTAKRKSR
IGVNATILPGKV
IGEDAFVGAGSVVTKDVEDGKIVVGNPGRVKK

>gi|15900014|ref|NP_344618.1| choline binding protein I [Streptococcus pneumoniae TIGR4]

MGMAAFKNPNNQYKAITIAQTLGDDASSEELAGRYGSAVQCTEVTASNLSTVKTKATVVEKPLKDFRASTSDQS
GWVESNGKWYFYESGDVKT
GWVKTDG
KWYYLNDLGVMQT
GFVKFSG
SWYYLSNSGAMFT
GWGTDGS
RWFYFDGSGAMKT
GWYKENGTWYYLDEAGIMKT
GWFKVGP
HWYYAYGSGALAVSTTTP
DGYRVNGNGEWVN

>gi|33864800|ref|NP_896359.1| conserved hypothetical protein [Synechococcus sp. WH 8102]

MIQLGWLALLLALPLQAVEHRPDPRERLLDQRACR
GCNLRRVELQEAHLI
GADLRDADLE
GADLRGANLE
GADLSGARLV

MADLRGATLTNADLT
GTDLRGADLRSAVVINAFAPDVRTEGIRFA
GAKLTGSDLIIGGGD

>gi|30249712|ref|NP_841782.1| hypothetical protein [Nitrosomonas europaea ATCC 19718]

MNIRSTGLSILIGALFAIPATTATAVSPVTDTSIRQGTASYLILADA
HQHQGHQ
GGSAGQGHSGG
SGHAGHG
QGGQGEG
KGRHGGGGGH
SGHGGHGKGD
MEHHGHPPSYAHSVAMQAEALGLSDEQLGKIVRFHLKEDKQAHERIKQKMMESMKAFRKAVGEPATDDETLRKLGQAHID
SFNEMVKYHIDERKAVRSILTPEQIGKLKAVKSDHDH

>gi|30250162|ref|NP_842232.1| hypothetical protein [Nitrosomonas europaea ATCC 19718]

MALTKEDIAEIKELIGEVITERHPEIMNNNVRYELEIRERILRVEEGLKHQRELMQE
GFNRMDK
RFEQIDK
RFESLIAEMNT
RFAQVDK
RFEQVDK
RFEQIDK
RFETMTARMDRFMIWSFATTLTVGGIVVAAIKYLP

>gi|30250370|ref|NP_842440.1| hypothetical protein [Nitrosomonas europaea ATCC 19718]

MLHYTRQSGKLMFKVILLSFIVLIVIPF
DSFAQTRCRP
DSLQNTTCR
DSNGNTMRSRT
DSLGNTTYR
DSDGNTIHSRT
DSLGNTTYR
DNRGNTVRCHK
DALGNTICR

>gi|13507890|ref|NP_109839.1| hypothetical protein [Mycoplasma pneumoniae]

MEFNGNLNHMEKRKSGYVTQKQFNEYKDSNDQ
RLIKIETTLAAQGEQISQLVQIVFLQGKQIKELQAEQKAQRVEFNA
RMDRFESLVLKSLESIGNTLTDFGK
RFDSMET
RLDSMDG
RLDSMET
RLDKIDPPK

>gi|13507906|ref|NP_109855.1| ribosomal protein L23 [Mycoplasma pneumoniae]

MDVTNVLLKPVLTEKVYFNQMGETKKYVFVVNPKASKTRVKLAFELVYGIKPLKVNTLIRKPTTIRGGSRFPGLSKLEKL
AVITLPKGIAISVTGEAPEKTDKPADKTTLKE
STVKEIKDTKNSPEAVVKTAVEALQIKPTAAPVTTAPLQTVAV
KVAKEVKEV
KVEKPVKVEKPTKPA
KVAKEAKTT
KVAKETKAEKSVQTT
KVAKETKTE
KSAKTTKTTATKTTKTK
TTKKEVKK

>gi|13508026|ref|NP_109975.1| hypothetical protein [Mycoplasma pneumoniae]

MGFSDNLNHMEKRKSGYVTQKQFSEFKDANNQRLIKIETTLA
TQGEQLNQLIKVVI
LQGEQIKELQVEQK
AQGEQIK
AKGEQIK
AQSEQIK
TQGETLKLILQALGGINKRLDKVDPPK

>gi|13508083|ref|NP_110032.1| hypothetical protein [Mycoplasma pneumoniae]

MPKIISKRKFNILRKTKEEIFSQIVYTKKRNGLHKAAKEYYFKDKDFTIIEDQQD
RPDKPEELDTPDIP
KPPKPPK
GPDQPEEPGQPG
GPDDPNSGNK
KMPKPDEFVTHRQLQEFKKDLLVELHEI
FPTKPELKRVEEKVDVLFELQKAQGEQIRIQGEQIKELKVEQKAQGETLQLILQTLQKMNDRLDKIEG
KMDKMES
RMDKMET
RLDKLESK

>gi|13508107|ref|NP_110056.1| hypothetical protein [Mycoplasma pneumoniae]

MKEKIPFYNEKEFHYMMKKTKKGTFSGWYIINPENNSVEFSGSFNRQFKLNKPVIPVNTEYVTRKEFNEYKDSN
DQRLTKI
ENKVDKL
EIKVDKL
EKKVDKL
EVKVDKLVETVNAQGEDLNNFKVEVRGTLQSQGETLQLILQTLQGM
SKRLDSV
EGRLDSM
DGRLDSM
ETRLDKIDPLK

>gi|13508204|ref|NP_110153.1| conserved hypothetical protein [Mycoplasma pneumoniae]

MAFL

FISNEWKYQQLIFKINCWHFCYYPKNYLGGSILSKRLFIPPSACRID
LSVFPWA
FICSPWNFCSTWSS
LICSPCFSTVWVSL
LICSPWRSTTWTNW
LICSPCFSTVWVNL
LICSPWAAKVVSIFVSRWLFEFLYSLNSLRVTYSVFTGITGLLSLNCLLNLPENSTLLFSGLMMYQPEKVPFLVFFTMSW
NSFSL

>gi|13508240|ref|NP_110189.1| hypothetical protein [Mycoplasma pneumoniae]

MKEKIPFYNEKEFHDMVKKTKKGTFSGWYIIDKDNNSVEFSGKFNRQFKLNKPVIPVNTEYVTRKEFNEYKNSND
QRLTKIETTLAAQG
EQINKLTQTVEKQGEQINQLVQVVLLHG
EQINKLTQTVEKQGEQIKELQIEQKAQGETLKLILQTLQKMS
DRLDKME
VKMDKME
VKMDKME
VKMDKME
KRIDKLESK

>gi|13508263|ref|NP_110212.1| hypothetical protein [Mycoplasma pneumoniae]

M
KEKISEKEYKALIRKIGKEHFDGEKEEYGDGTVGVWTYELRKYKLKPPVKVKYVTQEQFQEYKDSNNQRLIKIETTLAAQ
GEQIRIQVEQIKELQIEQKAQGETLKLILQT
LQKMSDR
LDKMEVK
MDKMEEK
MDKMEGK
IDKIENR
MDKMEVK
MDKMEKR
IDKLESK

>gi|16077507|ref|NP_388321.1| general stress protein [Bacillus subtilis]

MADNNKMSREEAGRKGGETTSK
NHDKEFYQEIGQKGGEATSK
NHDKEFYQEIGEKGGEATSK
NHDKEFYQEIGEKGGEATSE
NHDKEFYQEIGRKGGEATSK
NHDKEFYQEIGSKGGNARNND

>gi|16078124|ref|NP_388941.1| yhjQ [Bacillus subtilis]

MEQYS
EACIEAC
IDCMKAC
NHCFTKCLEESVQHHLSGCIRLD
RECADICALAVKAMQTDSPFMKEIC
ALCADIC
EACGTECGKHDHDHC

QACAKAC
FTCAEQCRSMAA

>gi|16078274|ref|NP_389091.1| spore coat protein (inner) [Bacillus subtilis]

MNVHTPNLSIRNMVKGIKKAREVFLLDYPLNEQSFEQITPYD
ERQPYYYPRPRPP
FYPPYYY
PRPYYPFYPFY
PRPPYYYPR
PRPPYYPWYGYGGGYGGGYGGGYGY

>gi|16078830|ref|NP_389650.1| similar to spore coat protein [Bacillus subtilis]

MGYYKKYKEEYYTWK
KTYYKKY
YDNDKKHYDCDKY
YDHDKKH
YDYDKKY
DDHDKKYYDDHD
YHYEKKY
YDDDDHY
YDFVESYKKHH

>gi|16078833|ref|NP_389653.1| spore coat protein (outer) [Bacillus subtilis]

MKNRLFILICFCVICLFLSFGQPFFPSMILTVQAAKSTRRIIKRSKNKGGVYMGYYK
KYKEEYYTVKK
TYYKKYY
EYDKKDYDC
DYDKKYD
DYDKKYY
DHDKKDYDYVVEY
KKHKKHY

>gi|16079784|ref|NP_390608.1| yrrS [Bacillus subtilis]

MSNNQSRYENRDKRRKANLVLNILIAIVSILIVVVAANLFINSPSSK
DVSKDSETAQKQESPASG
KTEKKSDEDIK
DSKKDTS
DSEKDSEKSSDS
DSKKDDS
SSKKDDSDSDSSSDSAG
DGLKDAKVTEGGSSS
DVEKTYENPDWKAVGTEQTGEHAATYDSSSQDWAEMLKAISYATGVSTDNMTVLWLGNNGSPQDAKGKILDKATGNKYQV
TITWVDEKGWKPTKVEKLK

>gi|16079802|ref|NP_390626.1| similar to hypothetical proteins [Bacillus subtilis]

MQEGDYEKAA
EAFTKAIEENKEDAIPYINFANLLSSV
NELERAL
AFYDKALELDSSAATAYYGAGNVYVVK
EMYKEAK
DMFEKALRAGMENGDLFYMLGTVLVKLEQPKLALPYLQRAVELNENDTEARFQFGMCLANE
GMLDEALSQFAAVTEQDPGHADAFYNAGVTYAYK
ENREKAL
EMLDKAIDIQPDHMLALHAKKLIDPS

>gi|16079941|ref|NP_390767.1| yscB [Bacillus subtilis]

MNKLIQLALFFTLMLTGCSNSSTSSESKVETTVKTTAAFPQK
ELEKELKKLKPVSLDMKFESPLATELGKRKAKEEA
EKQRQIAAEKKL
EKEREAKR
KKQQEEKAERQRLA
EAERQAEL
ERQRQAAI
QKEQKANA
EKKRQSQAQRQQTEAPSSNSQDPPSSSSQTDKTIQQPASELPDDDGYGYEERKKWHDDQVEWGIKQGYIDPEDAP

>gi|16080364|ref|NP_391191.1| yvqG [Bacillus subtilis]

MVIES
DSKDVRIIAEERSDISAGISGDSGKLFVTENKRKLELTVKEKEFQFLNGFNRSTLIVRLPYD
YKGDLAVRT
SSGDVSVAGNDHLALSGLNAVS
ASGNTSVTDVRIQDLKVKA
SSGDVSISNTVSKTAGIDL
ASGDANLVHVSGSLDVKM
TSGDFNAVLKK
VTGPVSVTL
TSGDANVSLPQNGSFAVNALS
ASGDVSSPYSFADKGHKEQHHITGTQGSGRHPIDIKT
DSGDLAIR

>gi|16080769|ref|NP_391597.1| RNA polymerase (delta subunit) [Bacillus subtilis]

MGIKQYS
QEELKEMALVEIAHELFEEHKKPVPFQELLNEIASLLGVKKEELGDRIAQFYTDLNIDGRFLALSDQTWGLRSWYP
YDQLDEETQPTVKAKKKKAKKAV
EEDLDLDE
FEEIDED
DLDLDEV
EEELDLEADDFD
EEDLDED
DDDLEIE
EDIIDED
DEDYDDEEEEIK

>gi|16081030|ref|NP_391858.1| yxcE [Bacillus subtilis]

MKLKYVKALVAVTVALGVLLPSTISHAKSF
SGRSSSSY
SSRSSSSSYSGSYKS
SPKSSYS
SGSSSSSKKSKT
SDDSSSSISLKKKP
SEKASSSSSKKSSGTFSGATSKVTGKTY
SGKTSKAYVGGRYVSVNHYYHAGFAPSGWFGYYSGFTMGMFMISMMHPWGYTYHPVGGPGYVSYGASPIAWIVDIIALII
ILIIVIALIRAFKAPKTYRRRF

>gi|16081125|ref|NP_391953.1| similar to hypothetical proteins [Bacillus subtilis]

MKYMY
YGIGILI
LSLGISLTIQ
SGLGTSPFDALL
VGLSKKIGLTVGSWEVLISVILLICNAILTRKKPIL
LGLITAFITGIGIDLWLFVVKNTIYLNSLLSKLIC
FGIGLVL
IGLGTAIYLQTKFASTPIDHLTLIIRDLSKRTILFSRTLVYALFLVLAIIFKGP
IGIGTLLTVCLGGMLLHVFMPIVERHFLSKNRRGY

>gi|21240782|ref|NP_640364.1| TonB protein [Xanthomonas axonopodis pv. citri str. 306]

MTEQLVIHRHDYDAGNQGLSWARIIGIAFVIALHLTALMMLLIPAV
APKAPAEKERTTMVTLVDA
PPPPPPP
PPPPPPEDK
PPPPVKNLSPPKPS
PVPPPPEAPVVDV
PEPRPSDIV
TPPSPPSPPAPPSDIGASVDISSKNMN
PPKYPPAAFRAGVQGEVILIVDVDASGNVTNVSVEKSSRNRDLDRAAMDAARKWKFNASTVNGQKAAGRVRVPVNFALN

>gi|21240799|ref|NP_640381.1| conserved hypothetical protein [Xanthomonas axonopodis pv. citri str. 306]

MKPQILLLALSLACTSAWADEDV
SKVNGKISADAGKTYGAL
ETVNGSIEIGAGAQTKNV
ETVNGGIRIGDNARTGGV
ETVNGSITLGQK
VTVSGGL
ETVNGSILAERG
SQISGGV
ETVNGSIGLVGTELGKGI
ETVNGDITVGVG
SHVRGGIEVSKPSFGFSFQPSRKPRVVIGPDAVVDGPLHFEREVNLYVHRSAKIGAVSGATARSFDSDVAPKD

>gi|21241001|ref|NP_640583.1| hypothetical protein [Xanthomonas axonopodis pv. citri str. 306]

MLRQTVLAIALACTIGAAAA
QVAAPPPPPAPNG
PIGGPPAPPT
PVAATPVSVSGTVERFLLNPNGDVDGLWLRDGT
QVGFPPHLSSEVKAAVRAGDSVVVKGFRLGNLPVLQVSSIRSGRSGR
EVVDRPPT
FGAAPPPPPTPGQLTPLQSDGTIQRLVYGPAGEINGALVSDG
TVVRMPPHLALQFADLLRVGAPLSASGLGVATPDGRAIEATQLGRDRSSLRALFTAPRPD
GPLPPPPPVAPVPPAGR

>gi|21241513|ref|NP_641095.1| conserved hypothetical protein [Xanthomonas axonopodis pv. citri str. 306]

MNRKRIVTVVLSSILALGAFAPAAFAGDW
DRDHDRRGHDRHHGPD
HRRDWRD
DRRDDRR
EWRQDRRYYSNGYREGYRDARYQ
DRRYDRNRVYVYDRPAYYRAGPPPRPYWARGQRYHGPVYVVNDY
DRYDLRRPPYGYR
WQRDDRGNLLLVAIATGIIADLVLNN

>gi|21241947|ref|NP_641529.1| conserved hypothetical protein [Xanthomonas axonopodis pv. citri str. 306]

MPQTKTRRAAAKDKREGKSASTQAGEYVKEEIDQVRDGAHGV
KSTKQAIAIGLSKARRDGVDVAPRSSAS
KATKKKAAQDKA
AATSKKPVSPTR
SAGAKKALKTASKGTTAK
KTVGKKALSKQTT
SAAKKRTASDRSTAAK
KAAKTKGPAVR
KAAAKKA
AATRKRAE

>gi|21242100|ref|NP_641682.1| outer membrane protein [Xanthomonas axonopodis pv. citri str. 306]

MTINKLLIAMALGLALTACSKPE
QAQDAAA
SANEAAADAQT
AADQAAASGSQSA
DAAQAAANTAAASADTAA
DAAQAASG
AATDAAADQAA
DAAKTAE
DTAKAAK
DTAEEAKK

>gi|21242193|ref|NP_641775.1| hypothetical protein [Xanthomonas axonopodis pv. citri str. 306]

MFTHAARGGAAGIAVASGCWGAAKGGRPGGAVTVVHPCGA
RRRGWDCGRLRLLGRGK
RWTAWRCSKRQFTRPARGGAAGTAVASGCWGAAKGGRLGGAVTVVHPCGA
RRRGWDCGRLRLLGRGK
RWTAWRCSNGCSPLRRAAARLALRSPQAAEARQKVDGLEVQQTAIHSSGA
RRRGWHCGRLRLLGRGK
RWTAWRCSNGCSPMRRAAARLGLRSPQAAGAPSSSQEAMCNIRITTKRNQVI

>gi|21242536|ref|NP_642118.1| hypothetical protein [Xanthomonas axonopodis pv. citri str. 306]

MRSDLILTDA
ARHPSTSRV
GAHPGATGHPGNAPRAQVRSYRPCDSSENATLHGSASPV
GAHPGATGHTGNAARAQVRSYRPCDSSENATLHGSASPV
GAHPGATGHTGNAARAQVRSYRPCDSSENATLHGSASPV
GAHPGATGHTGNAARAQMRSYRDATAARMPRCTCRL

>gi|21243092|ref|NP_642674.1| conserved hypothetical protein [Xanthomonas axonopodis pv. citri str. 306]

MGAELDGLDVTATATLPTGRV
AAGLAGASVLVGVFAEDFETGTGFLDDGVAGFAGFAGTALGTGA
GAGFSALTGTDLAGAGLVAGFGTDLDATGL
AAGLAAAGAAFF
AAGLAALL
ATGFTALGAADAVGLEA
GAGLAAAFFAGAFF
AGAFLAAGCLVAGF
AAGFAAADV
AAFFATG
LADFLAASTAF
FADFLAATKRSSLVPPGPGKAGLYNVPRGPPATTY

>gi|21243785|ref|NP_643367.1| histone H1 [Xanthomonas axonopodis pv. citri str. 306]

MAT
AKAAKKKPT
AKKAVKKTA
AKKAVKKVAKKAT
AKKAVKKTA
AKKAVKKTAKKAT
AKKAVKKT
AKKATKKTA
AKKAVKKT
AKKATKKTA
AKKTAKKAVKKTAKKAV
AKKATKKSAT
KKTATKKPA
VKKVAKKKPAVRKKKAAPVALPATPAPLI

>gi|15828993|ref|NP_326353.1| LIPOPROTEIN B [Mycoplasma pulmonis]

MNKKYFKKYSSILIFSMSILAPMTLASCSYNLAKEKDKDQKES
TNLSEPNK
SNTSKTNTFQDKKDSTNKIDSQESSKTQS
QNTSESN
QNTKVDSSKTNNLATNQ
NNPSKSNVNI
QETNETKQE
QNINPNNAVISEKQAEND
KNASSLNSKQINNTLKNQDKTKQENDQFKQESKDSSNFTSPKPITHDPINKVISSQSTTRLEMPKND
QSNSESEDNKKSPESPKWWEKLNQGDAF

>gi|15829203|ref|NP_326563.1| unknown; predicted coding region [Mycoplasma pulmonis]

M
NNNWKPNRGSNPYINQNPQTP
QNPQNPQALQNPQNQQNLQNQNPYSNA
QYGYAPQQGQMPVNTPMPQVGQVY
NNAYDPQVGPQPMNTQ
QFSYDPNIQYNTSQF
NNPYANQRNPQFVPSQTQ
QFSYDPNAQYNSAQF
NNPYQGQYQQPKVVNEGA
NPFYQPEFKTQQFEYQAPRQANFNPNISTLERKIKALDKMLSEKLISISEYVERKNALLKTELEKI

>gi|28210376|ref|NP_781320.1| hypothetical protein [Clostridium tetani E88]

MNKSTDSNLKQENTD
TNNSSTKKQENAD
KNNNSTKKQENIN
KNNNSTKTP
KDTNKNKKVS
QNTNSTKISNKNILTKDFSQVSSMEIIYNDTSKKYSIDKDKDFIKKVYDSILNTKTTLNKYAQESDPIFTINLVYKSGEK
NIIKSTETGEFIYRVLDNKSQWSGG
RNNNLFQIIKNNK

>gi|28211437|ref|NP_782381.1| putative phosphatidylinositol 4-phosphate 5-kinase [Clostridium tetani E88]

MINNHIEDKKHTHESEAHLHGGN
YVGEKKDGKAHGEGTYLYKDGSK
YVGQWEDDKMHGEGVHTWACGEK
YIGQWENDDKHGYGIYTWPDGES
YVGQWKNGVKHGEGIYTFSNGEK
YMGEWKDDKIHGEGMHICTDGEK
YIGTWEDDEMVSSKDMQF

>gi|28211654|ref|NP_782598.1| putative grpE protein [Clostridium tetani E88]

MKWRHDYMA
EKKVEDLEMDLN

EEELEESEVNE
DKEFEELDKSEE
ENEVEELISTL
KDENEKL
NNEMEALKDRLLRTTGEYDNYRKRTDREKEGLYASACEDVLKEILPVLDNLERAILA
KGDIEDLKKGVDMTLKQFKDSFKNLGVEEISTENGFDPNLHDAVMHVEDSQYG
EKEVVEVFLKGYKKGDKIIRHSMVKVAN

>gi|28211875|ref|NP_782819.1| conserved protein [Clostridium tetani E88]

MNLSERVCTVNIVVVRINGVEYNLKGNEKEEYLHKIANYVDKTIKTILSNNTKLSTSSAAILTALNTVDELIKS
NEEVEKLLEEINLV
KEYNKRLEKEIEELTDKLQAL
YEENQQV
KEEKEKI
RDEKKELLKNSNSNDTNDENLKEKDLEIKSL
KIEKEDLYNKYNET
KEENESLKVERKEHKFQIQSAKYKIIDLQNRLIETQVELA
KEKNQNNPLINNRIKKKVKS

>gi|38234121|ref|NP_939888.1| Conserved hypothetical protein [Corynebacterium diphtheriae]

MYRVFECFDELIDILEKSTGVPMTSNCMVPRHHMLALIDEARNALPNEI
DDAQDVI
DQQDEILRGAQERARTTIDEAN
EEARRTVTQSR
EEADALVADAE
EHAERVVRQAQ
DEADRLVDGAR
READDTVNRAQ
AEAERLIASGNDQYQRSVDDGL
AEQQRLVSEAEVVRRAN
EEAHRVVDAAHADSNRLRKECDTFVENKLADLEESLSSTLRTISRDRAALRRGAGASGGGNHSQS

>gi|38234825|ref|NP_940592.1| single-strand DNA binding protein [Corynebacterium diphtheriae]

MAQGDVNITVVGNLVADPELRFTPSGAAVANFRIASTPRRYDSQTNQWVDGEALFLQCNIWRQAAENVAESLTKGMRVIV
TGRLRQRSYETREGEKRSVMELEVDEVGPSLRYAAASVNRNPREGGN
GGGNFGG
SQGGFGGNTGGN
SGGGFGA
PQGGFGGGPQNSQSAAPANDPWSSAP
QAGGFGGAEQDPPF

>gi|15606000|ref|NP_213377.1| UDP-3-O-(3-hydroxymyristoyl) glucosamine N acyltransferase [Aquifex aeolicus]

MGSFIGDFVVIGK
NVKIGRNVKIYPFTYVGD
NTVIGDNTVIFSGVHIYR
NTVIGRNVRIHSGAVIGADGFGYHITQEGIKKIPHIGGVIIED

NVEIGANTTIDRALIE
NTLIGKNTKIDNLVMVAH
NCKVGENNILVSQVGLSG
SVKTGKNVILAGQVGVAD
HVEIGDNVIVTAKSGVANNLAPNKTYGANLPAIEWSRWKRIYVYLLRLPELFKKITT

>gi|15606457|ref|NP_213837.1| putative protein [Aquifex aeolicus]

MTKEEAKLLVKIQEIDLETERVNHRL
KKIEEEV
KKLKEELEALLKEKETLLKRKEELENLK
KQLQEEV
KEAEEKL
KVTEEKLMKVTRDVEYKALLREK
SKLEDKILKKSYEI
DQIEEELEKIT
KEIEEKV
PRIERQV
KEIEEELKD
LELEESIAH
RKIHEYVQKR
EEVKREIPEHLLKFYEENKKHFEGLVIVPTEDEACAGCGIKIPSVLLS
KMIKEDSIEQCPSCGRFVYYKL

>gi|15606545|ref|NP_213925.1| putative protein [Aquifex aeolicus]

MMGLLKRLLRVVRAEAHSVVEKLEDPIKI
TEEGLRELRKQLKEAMEALAQVKASQIRF
EREARNEKERAKALVKKAESLLLQAQEGKIS
PEEAEKLAAELIQKAELH
EKNARRL
EEEAKRQREMAVKLQAKIDELKVQIAKYEAELKTLKARLATARAVKKVNKQIAKVDPSDTIAMLERMKEKVE
EEETLAQAY
EELAKQEVALDV
ESEAEKKLK
EVESKQKLEELKKRLGLS

>gi|15606638|ref|NP_214018.1| putative protein [Aquifex aeolicus]

MMLKSLTMENVKVVTG
EIEKLRE
RIEKVKETLDLIPK
EIEELER
ELERVRQEIAKKEDELIAVAREIRHKEHEFTEVKQKIAYHRKYLERADSPR
EYERLLQ
ERQKLIERAYKLSE
EIYELRR
KYEALRE
EEEKLHQKEDEIEE
KIHKLKKEYRALLNELKGLIEELNRKAREIIEKYGL

>gi|15606717|ref|NP_214097.1| putative protein [Aquifex aeolicus]

MSGEDFKPLWEDFLKKENQKKEKDEEDC
SKVIKELL
EEREKLLETIRSY
EEKLDSFE
EEKQKMYFEIKNR
EEKIKEL
EQRLNELSEKGSISLSLSKSI
EEYLSNLKEEY
REKVEEFLKTFLLEFSYYVPQVKVLKEDLRNIIDELIKFKTNLKLYINPEDLRYL
NDELKSLKENLKAEGINLQVLEDKDLEKGSFRIKGEHFTVERDPKEFAKIVFEKVFGDVFKRA

>gi|15606921|ref|NP_214302.1| putative protein [Aquifex aeolicus]

MVIMLIFASSTGTILPSKKAKSMEKDFSIFQGNKLNLKYKRQ
REEVMQKKHRLIFLATVLAGLILFYFGVDTWMK
QKQVQQNQPPPIVIKPVAPVKPKTQESNQTT
KKEVKQEEQKKEEPKKMV
QKQETQE
KREVKKSE
KNEVKQTQE
KKDVKVA
KKVPKTEKKAANLRTYKFQVGAFRYRENAYKMAKIVRSKGFDAQVVKVGSLYRVYAYVKAKNYWEA
KREIKKHFKDAIFVRK

>gi|15606925|ref|NP_214306.1| putative protein [Aquifex aeolicus]

MKRIIALLFFFLIAFGYSLSPEEEKQLIRDIAEIKATLKTF
MEQTDKR
FQDLNQRINEL
REDMNKR
FEQVDKR
FEQVDKR
FEQINNELNRLIQIMVGIFAGQIALVAAVIGFAWWDRRTIIRKSKEET
FEEMEKELRPEKFKKLLNALREKAKTDKELEAILKKYGLL

>gi|16129005|ref|NP_415560.1| curlin major subunit, coiled surface structures; cryptic [Escherichia coli K12]

MKLLKVAAIAAIVFSGSALAGVVP
QYGGGGN
HGGGGNNSGPNSELNIY
QYGGGNSALALQTDARNSDLTIT
QHGGGNGADVGQGSDDSSIDLT
QRGFGNSATLD
QWNGKNSEMTVK
QFGGGNGAAVDQTASNSSVNVT
QVGFGNNATAHQY

>gi|33592345|ref|NP_879989.1| putative membrane protein [Bordetella pertussis]

MVTDAVLNHAGREWGWVALR
GVVAVLFGLMAVLMPGITL
SALVLVW

GAFALAD
GIFALVAGWRIRDQDKPLWPLILVGLTGIAA
GIATFAWP
GLTALVLLYIIAFWAVIGGVFQIAAAIRFRKDIENEWLHGLS
GALSIVF
GALLLFQPGA
GALALVWVIGVYAVLFGVLLLALAFRLKNHVATPA

>gi|33592352|ref|NP_879996.1| putative membrane protein [Bordetella pertussis]

MNPHLPPDDDDRDLRALYRELPRHEPDIMLDASIREQ
AQRAAEADRRARRPRATWHPAWAVAACVVAVSALFVFTDLEQYALQPAAEVELQA
EQEAAQATRDAV
PEPPAPD
PALAAPPPLGALAPEAARGAAAESS
AQDAAPARRGYM
PQRAAPAVRAEKQQQENLAASA
RQRDDPAYAQRVERIRALVREGK
PQEAAAEVLA
WRRAAPGLALPPD
LEQLAPADGGTPGQ

>gi|33592511|ref|NP_880155.1| conserved hypothetical protein [Bordetella pertussis]

MVLAAVIVLPMILDSEPVPVDSDIPIRVPDRNSAFQPTVND
PQAAAPQ
PEQPAPQAPAGT
PEAGAPQ
PPALASA
PPQTPPPGQTPA
QPQTPPAAQ
PPATPPQQAATPPAAPK
PEAKPEPKPA
KPAAPARSDDGARAMALLEGRSP
PPAQSKPAANGNFVLQVASYTTQADAQARRGKLHQAGVTNAFVEQASINGKQQYRLRVGPFPSREAAQAAQARLRTLGYD
NGFIAAQ

>gi|33592564|ref|NP_880208.1| conserved hypothetical protein [Bordetella pertussis]

MSQDPAGFIGRWSRRKSEARQQA
ETAADPAAGVAQASLADTPFAGAPAVAD
RPEETPE
RADAAPVQAPGPAADAETDLPLPTLDDVRALTLDSDFSRFVAGGV
SPDVKNAALKKLFTDPHYNVMDGLDVYIDDYTQLEGLPQETVRRMASAKVFKLFEED
EPASAQGEPADGACGPGQQAPGAMA
EPDAAPAVPAS
APDAIQSALPR
DDAAGPADGP
LPDAEPRRD

>gi|33593522|ref|NP_881166.1| integral membrane protein [Bordetella pertussis]

MELSSAAFWIALLQIIWVNILLSGDNAVVIALAARSLPPDQQKKAIV
VGSAAVIVMRIALTL
VAAKLLLLPWLKL
IGAALLVY
IGVSLLLPEGEDDGAEKGHGSLLVAIRTIMIADLVMSLDNVVAVAAAAMGDTTLLV
LGLAISIPLVI
FGSALLLKVIERFPIIVW
VGAALLGF
IAGELLVGDPALEESVARIDTMLGVTQHDFALMVGA
IGAILVLALGKIFQARQRA

>gi|33593917|ref|NP_881561.1| histone protein [Bordetella pertussis]

MATAKKAAKK
AVKKPAAKK
AAAKKATPAKKAAVKKVAVKKV
AAKKPAVKKV
AAKKPAAKKVAKKAVAKKAVAKKAVAKKAVAKKAVAKK
AVAKKAV
AKKAPAKKAVAKKAVAKKAVAKKAVAKK
AVAKKAV
AKKAPAKKA
APKKPATPPSTAAAPGAKTALNPAASWPFPTGGRP

>gi|15805649|ref|NP_294345.1| hypothetical protein [Deinococcus radiodurans]

MSPRALPGLVACLFAGSLSAACAQSGD
RPLFVQPPQMQGAPA
TPAASTPMWT
PPAQAAPAQPAPV
RPAAAAPVRP
LPAQAAPVSPASVPLTPGV
RPGASSRPGVTPRALPASGP
RPAATAT
QAQTAAPLKLQLMPVWTGQFMV
RPRELAPFNPLGLPQGAVSLR

>gi|15805940|ref|NP_294639.1| hypothetical protein [Deinococcus radiodurans]

MRFERGSQRRSDIFARQDAGGLACLEQFLNAGVLVVGLFSGATSGGFDRRRLGRSGLGAARLGLGRSFCGCLFSGFGTLS
GGSFLGRRLFGSSFLG
AAGLSSVSFSSAALVL
AAGLARAFL
GAGAASAVSV
AASAAASGVTAFLGAAFF
AAGLAAGLGAGL
GAGGAAS
VAWAAGA
AAGAASAVTLRRVRGLSGVMILL
SAGASSATGLALGRTRVLTGSAIEDS

>gi|15806033|ref|NP_294734.1| hypothetical protein [Deinococcus radiodurans]

MTASLTFSVSAALGVTLTGRLCVAAHQGQGAAAQAEAVI
GGRGGGR
SHRDRGQHRHFHRGPFLLVAPE
AGYGQNRPTRRARWRAGVVRCRSCSRSSPSGSADFNPVGNAADLHRHGPFLAVTEGQRVQAEGQRGSWLGR
GGRGRGRRR
GGRRACRGQR
AGRRRERGVAGAQWAAQGEGEQQRA
GGQGRAQGAQRHGVLLGVISEKRVKAR

>gi|15806771|ref|NP_295493.1| hypothetical protein [Deinococcus radiodurans]

M
TVAAAPSAQADRKALWTFLRVLLGLALLAGLAYFLWTPGVWTNKLFVWVLLTILADECAGWFGYLGLALGALVFLAPGAT
PEQWLVVAPLIAAALFALLLVKHSGGVLVLPFAGLIYAGAVLGADKLGHSLDTPIKLLGSEEFRRVALLPMIIGLAFSFV
RQLVNLLLRGHSRRR
ATRPLPTTPVP
VPAADPT
TTAPDPA
TPTPTPEVEEVE
VVAADPAVTVV
TTSASPATPPAPSKPAKSKTKRK

>gi|15807335|ref|NP_296065.1| hypothetical protein [Deinococcus radiodurans]

MGFPLELTTHQPDKEITMKKSILALTILLGSVAYAQDTG
TTTTDTS
TTTTDTTG
TGTTDTTGTGT
TDTTGTD
TTGTDTS
GTTTDTS
TTTDTTSTDTTDTNVQNDAV
TTSTEADGVPGNEKEPAGFPWGLLGLLGLAGLMNRGRPQPTPVVH
TTTTEPRRDT
TVVTGTTTTNNDPNRR

>gi|15807575|ref|NP_296312.1| hypothetical protein [Deinococcus radiodurans]

MASSPRLPEPSPLAGRIPSMSKLDGRIGGLTAGYDLHADWD
GSRLTGRI
GGQIQGKDLRLEVQ
SGDVDGRV
GGVVAGFDV
DGDVSGQRVEVRL
GGRIDGDSLSLDIR
GDEVRGRF
SGRIAGKDVTLTRR
GEQLSGRI
GGQVDGKDVTLTLGDVPLELAALAAVCAYKALEDVQQGTAPSSYGGAS

>gi|15807958|ref|NP_285621.1| hypothetical protein [Deinococcus radiodurans]

MSRL
LLFLGLA
LLLGGAAARRWLAPVPLRWLSTGLA
LLFVGAGLSVYQPLRELGLLAPADVLDYLTAVSAGRAALVLLLAATLFLVAEVQKLPA
LILAGLGGWAL
WGMAGIGHGGVHGDAVRGLHALHAGAMS
IWLGGLGALLRVPAPERLAVARRFSPVAAACVGL
LVLTGLGMSLEHIHSLRVLQSTPYGQTLLLKLLAFVGVLGAAFTVRRSLASTGPGRRLSLEVLALVAVL
LLTARLGGLPLVH

>gi|26986997|ref|NP_742422.1| conserved hypothetical protein [Pseudomonas putida KT2440]

MKKYSSILLLSFSLLSGVAMA
GGTTEAGI
GGALGGVLGSVV
GNSIGGST
GGAIGAGLGGAA
GGALGADKRQRGEAAI
GGALGAAG
GNVVGRSVGGTTGSYI
GAAAGGGA
GGALGNYLGKEADKDDYRDGRRYRRGYDDDRRHWDRGHHYGHRKHKRHWRD

>gi|26987928|ref|NP_743353.1| hypothetical protein [Pseudomonas putida KT2440]

MAPVPTLK
MAVAVAVA
VAVAVATVAVE
VAAAVATVAVE
AAAAVATVAVE
AAVAVVTVVVE
AAVAVATVVVE
AAVVVATVVVE
VVAAVAMVAEAEV
VATAAAVTVVAMP
ATAAVATTVIQTLGAAATMLKRVTIMAITWVGNPGMIMVIMSA
VSLAMITAIMLVVNLVTITATMWVANPVMTTVFMPAARPGMTTAEADWR

>gi|26988896|ref|NP_744321.1| conserved hypothetical protein [Pseudomonas putida KT2440]

MKKPLVAVFACALALGMTAFLPADISPISSAYAKGGG
GGGGGGG
GGGGHGGG
NGAGHGG
GMGGHSGSGKG
LGSDHAGKATRDHG
VSGNHYGSSRN
SDNGHGTTTSGVAHSKDTRGLAKSTAISGTTPGDHNSKGLSNAVGSSSKNDR

>gi|26989392|ref|NP_744817.1| pentapeptide repeat family protein [Pseudomonas putida KT2440]

MNYLPLTLLLALSPALAVNADDAGTDTPLTINGCVIAEASQCPG
ANLRGANLAN
QDLRKMNLAG
ADLRDADLRHAQLDLANLEK
ARLQGANLTR
ASLQQSNLRL
ADLSNSRLVAIQGWGLFAQGAQFEK
ADLRAAYLQFARLSG
AKLQQANLQAADLEMAWLSK
ADLQEANLGD
ANLQEAKFGQSNLAH
ADLRGARQHY
GNFQEANMEGCKGCPETWDK

>gi|26990113|ref|NP_745538.1| curli fiber surface-exposed nucleator CsgB, putative [Pseudomonas putida KT2440]

MKRLYSLLLCLALAGQACWAEDLMDNADLAPGNDLGEPVIIRLLPAGAGQA
AVIEQQGLGNR
AALDQNGQALL
GRIVQSGGAQE
AYILQEGSDLM
AIISQQGNGNS
ASIRQSGSSNN
AAIEQIGNDNS
ASIVQSGSGLN
SSVTQAGNGQHVQITQYR

>gi|26990474|ref|NP_745899.1| hypothetical protein [Pseudomonas putida KT2440]

MNRLIPGIVLLVGALAISGPAAAHGGYGGHGHGG
WYGPGPLLGAAVVGAVVGATVYGGHDRTVYVERQPV
YYAPPPV
YVQPPPQPLYYQP
YYAPPPPPPGYRT
YYGPPPV
YYGPPRW

>gi|26990635|ref|NP_746060.1| hypothetical protein [Pseudomonas putida KT2440]

MRLTLPSLALGLLLCQGAFAGDGTAA
IGGGLGGVLGNV
VGQQLGGKTGAA
IGAGVGGAAGSA
VGARKGNRTEAA
IGGGLGSAGGSL
LGGAVGGKTGST
VGAGLGGA
AGGALGNHLGDNNRGSKKHHRHRH

>gi|26991538|ref|NP_746963.1| conserved hypothetical protein [Pseudomonas putida KT2440]

MKRHLLTLTLSLLAANAFALPASEQHLTSEARSSAAEIAQPLNTV
AEGGSDRLIERSGRV
AEGGSDRLIERSGRV
AEGGSDRLIERSGRV
AEGGSDRLIERSGRV
AEGGSDRLIERSGRV
AEGGSDRLIEQAGRV
AEGGSDRLIEQAGRV
AEGGSDRLVELSRVS

>gi|26991984|ref|NP_747409.1| ferric siderophore transport system, periplasmic binding protein TonB [Pseudomonas putida KT2440]

MMKTRSNMARYGGSLAIVLGVHVVAVLLTLNWSVPQAIELPPAAMMVEL
APLPEPAP
PPPPKAAPQ
PPAPVEE
PPLPKLAE
APKPKIAIPK
PPKPKAKPQ
PPKPEKK
PEPPKEAPPTEQTVDAPPSN
SPPQKSAAPAPSIASNSNALPTWQSDLLRHLAKYKRYPEDARRRGLQGMNRLRFVVDAEGKVVSYSMAGGSGSAALDRAT
LEMIRRAGTVP
KPPPELLNNGTIEVVAPFVYSLDRR

>gi|18976708|ref|NP_578065.1| hypothetical protein [Pyrococcus furiosus DSM 3638]

MPLDFLSSLFKRKKEEEVEGEQEEI
NVEELEERVENREENEQLNQIMERINEIENDIPRIKI
SIDNIKK
QIQELRE
EIERLDKTIKDVMMLYEVVSQ
EINPFKEQLSQETSLSS
EIQELKK
QIEEIRM
ELAQVKNDIKVLAGYGVDIDSILYEVLAEV

>gi|18977446|ref|NP_578803.1| hypothetical protein [Pyrococcus furiosus DSM 3638]

MKEVFKDVKKV
HVRIVNGKLKIEKSEDDNVYIEYYILEGEPRVEIRKEDSELIIETKGKEKRFLNLFAWKSDGKVKMNIKVPEGVTL
EVSSVNAGISASQVSIE
SINSVNGPIILKNSHVK
SISTVNGMIAGSIMAEDL
TVNSVNGYVKLDILDMEGDG
VINTVNGGVHLRISDVCDVNVIAVKRGRSIPLINQNSPYTL
TITSVNGKVNVERF

>gi|18978092|ref|NP_579449.1| hypothetical protein [Pyrococcus furiosus DSM 3638]

MRRGLSLLGFIIGAVIGIIIGALILAGIFLLALGSTISFLDFPNCHCTCETINNAGPIKEIGAYNATSLII
EDVIGTVKIKPSNEPQIKVLSNLPINVSVSEDSLVITCSECREKSFKNGEIILMGNLSSIKT
SDILGRLIVEVPTKRIEV
GDVLGEVNIQAYTEVFES
EDILGRAEIKATREVQI
EDVIGEVKITIPKNSTAKISY
SKIIGKIADLSEDSNKIVKVDI
EDILGELTVENEK

>gi|18978306|ref|NP_579663.1| hypothetical protein [Pyrococcus furiosus DSM 3638]

MEAWRGRSQAVVDNLAPRVVDMLVPEG
FKPTQEEQLSSYDVEKKELATVYMITLVSGSGEKEEEVEE
ETPTQTET
QTPTETRTET
KTPTETTT
TTPTETK
ETPTQTTT
TQPARTETQGGICGPGLIVLLAALGVLRRRS

>gi|15900867|ref|NP_345471.1| conserved hypothetical protein [Streptococcus pneumoniae TIGR4]

MSIEMTVSEIAEVLGLSRQAINNRV
KELPEEDTDKNDKGVTVVTRSGLIKLEEIYKKTIFEDEPVSEDVKQRELMEILVDEKNAEILRLY
EQLKAKD
RQLSEKD
EQMRIKD
RQIAEKD
KQLDQQQQLTLQAMKDQENLKLELDQAKEEVQSTKKGFFARLFGG

>gi|15901340|ref|NP_345944.1| cell wall surface anchor family protein [Streptococcus pneumoniae TIGR4]

MVPKTATSTETKTITRIIHYVDKVTNQNVKEDVVQPVTLSRTKTENKVTGVVTYGEWTTGNWDEVISGKIDKYKD
PDIPTVESQEVTSDSSDKEITVRYDRLST
PEKPIPQPNPEH
PSVPTPNPEL
PNQETPT
PDKPTPE
PGTPKTET
PVNPDPEVPTYETGKREELPNTGTEANATLASAGIMTLLAGLGLGFFKKKEDEK

>gi|21282368|ref|NP_645456.1| conserved hypothetical protein [Staphylococcus aureus subsp. aureus MW2]

MKKLIISIMAVMLFLTGCGKSQEKATLEKDIDNLQKENK
ELKDKKEKLQQEKEKLAD
KQKDLEKEVKDLKPSKE

DNKDDKKDED
KNKDKDKD
KEASQDKQSKDQT
KSSDKDNHKKP
TSADKDQ
KANDKHQS

>gi|21282461|ref|NP_645549.1| conserved hypothetical protein [Staphylococcus aureus subsp. aureus MW2]

MSNQNYDYNKNEDGSKKKMSTTAKVVSIATVLLLLGGLVFAIFAYVDHSNKAKERMLNEQKQEQKEKRQKENAEKERKKK
QQEEKEQNELDS
QANQYQQ
LPQQNQYQYVPPQQQAPTKQRPAKEENDDKASKDESKD
KDDNASQD
KPDDNQK
KTDDNKQPAQP
KPQPQQPTP
KPNNNQQ
NNQSNQQA
KPQAPQQNSQSTTN
KQNNANDK

>gi|21283781|ref|NP_646869.1| probable DNA-directed RNA polymerase delta subunit [Staphylococcus aureus subsp. aureus MW2]

MKIQDYTKQMVDEKSFIDMAYTLLNDKGETMNLYDIIDEFRALGDYEYEEIENRVVQFYTDLNTDGRFLNVGENLWGLRD
WYSVDDIEEKIAPTIQKFDILDAD
DEEDQNLKLLGE
DEMDDDDDIPAQTDDQ
EELNDPE
DEQVEEEINHSD
IVIEEDE
DELDEDEEVFE
DEEDFND

>gi|21284018|ref|NP_647106.1| ORFID:MW2289~hypothetical protein, similar to TpgX protein [Staphylococcus aureus subsp. aureus MW2]

MKRLVTGLLALSLFLAACG
QDSDQQKDGNKEKDDKAKTEQQDKKT
NDSSKDK
KDNKDDSKDVN
KDNKDNSA
NDNQQQSNSNAT
NNDQNQTNNNQA
NNNQKSSYVAPYYGQNAAPVARQIYPF
NGNKTQALQQLPNFQTALNAANNEANKFGSNNKVYNDYSIEEHNGNYKYVFSFKDPNANGKYSIVTVDYTGQAMVTDPNY
QQ

>gi|21284086|ref|NP_647174.1| hypothetical protein [Staphylococcus aureus subsp. aureus MW2]

MKRTDKYRDSYQ
YDNQNQN
HRRQSEDASYRQQYAKGDPEEHPERYYNGRDYRREQILEEENEKSRRSKKWLYIIIAILLIIVAIFVTRALLNNDSDKVS
NDPKVSQN
YKKQVENQDGQ
INQQVDN
AKENIKNNQKTDDIIKN
LQNQIDN
LKQQEQNKADSKLTQF
YQDQINKLTEANNALKNNASQGKIESMLNDINTKFDSIKSKLESLFKDENGGAN

>gi|21284348|ref|NP_647436.1| conserved hypothetical protein [Staphylococcus aureus subsp. aureus MW2]

MAKGNQGIKWSSL
IMGVLLLMLAVVIFTFPIENFYAITWLIGLFVLINGVIQIIYRRKAKALVGGNQNWIL
FMGIVDI
LFGLIVIFNVGASSAFFIYMFAFWFIFSSIAGLFTFTGSGSLK
LISVIFN
LLGIVFGVILLFNP
LMGIVFIST
MIAIAFV
FVGVIYVVDALA

>gi|12045012|ref|NP_072822.1| ribosomal protein L29 (rpL29) [Mycoplasma genitalium]

MTIAKELKQKSNEELVKLVIKLKGELLEYRFKLAHGELDKPHLIAKVRKLLAVVLTILTERKLNWQVEKDKYKLLSRKTN
ELIVNSWK
QKLSTKPESKQET
KKAEVKPKVESKPESKQET
KKAEVKPLKQET
KKVEVKPKVEPKPLKQET
KKVEARIETK
TKVESKPLKQEV
KKVEAKKSVSKPQKPVKAKMIKTKEKKQ

>gi|33864553|ref|NP_896112.1| RNA-binding region RNP-1 (RNA recognition motif) [Synechococcus sp. WH 8102]

MSIFVGNLPFRAEQEDVIELFAQFGEVANCALPLERDTGRKRGFAFVEMADEAVEDAAIEGLQGAELMGRPLRINKAEPR
GS
APRGGGG
GYRGGGG
GYGGGGGYGGGGGRDGGG
GYGGGGG
GYRGGGG
GYGGGGGGGRDGGG
GYGGGGG
GYRGGGDAGDRPSGARGWEDRSYGARDSG
GEGGGGHDDGRSRRRRGSS
GGGGGGDDYS
GYGGAEG

>gi|33865398|ref|NP_896957.1| conserved hypothetical protein [Synechococcus sp. WH 8102]

MATEGGHCGSKPKRIAIGVAP
LGTISIGIVP
MGVICIGVVP
MGVVSIGVVA
MGVFNAAIVG
MGLVAVGVNT
MGVITAGPMS
MGLIQIRSTTNPRYLAYPSREEAEAQAKVLGCEGMHRMGDRYWMPCNEHPK

>gi|33866008|ref|NP_897567.1| conserved hypothetical protein [Synechococcus sp. WH 8102]

MSDSYGEPQQQGA
QGDGSRGDGY
RGEGGRG
GGRGGRGP
GNREGGGFRIRLSENEMRSARALQEAFNLRSTVAVLGFALRTLGQMLEDGQLDELIEQQRNQAPRGRREG
GGRDGGGRDGGGRGRRSDDDRQ
GGRGSRPDPFARPAKPQPEPEPTPEPESTPEEATPSDAPAEDTPAEASTDLPNEEESKAATPEA

>gi|33866530|ref|NP_898089.1| conserved hypothetical protein [Synechococcus sp. WH 8102]

MSTTGSAPLNLREWTAPEKLLPNGSPLVPVDAR
GADWHGIELG
ELDLRGAKLCRCDLR
GTDLSHCQLD
GADLRLARYDSQTVVPV
GFDLSNSGAVGP
GAKLNGVYLNSSDLR
GMDLRGSMLLGSYLS
GADLSGAILD
GVSLAGSDLRSSTMRGAMCRGTRFGTCELDMADLR
GANLEGAVLETVESIR
GADFSLCTGLGAQIDALLSRSVAELDCWNPLTRSTTRDSLESLRAARSD

>gi|15893289|ref|NP_361003.1| similarity to O-linked GlcNAc transferase [Rickettsia conorii]

MG
KYDLALE
AYNKAIEVDPSHPY
AYNNKAEILRKLE
KYELALE
AYNKAIEIDPNYSDAYFNKSKLFDELG
NYEVALE
AYDRALTYRHDDDLDYSDIYII

>gi|26553464|ref|NP_757398.1| hypothetical protein [Mycoplasma penetrans]

MKENKNIKNNKNKPFKGKYSEKEFIEMIKGCKFIDVDNFFISLSSYEDQKTGDIIETSVFEGNMKASKTKKYQKPKDPKD
PIWEVLGKILDKLEV
IESDIRI
LKEDVTV
LKEDVAV
LKEDVAV
LKEDVAV
LKEDMSKIKRCPTITRELAQLN

>gi|26553688|ref|NP_757622.1| predicted integral membrane protein [Mycoplasma penetrans]

MNDNF
NNNGNGPALD
PTQGPNPYN
NNGGYNN
NNNPYNNNGY
NNNGYNPYM
NNQAYNPQENVNLFTKIKKHSKNMLIFYFTAIVLAIIGAILISVGAAKIALSLSNNYSGEYYDLSSFINGEWIVGVLFIV
LTGIFYLVCFVIGIILVIKVNDAQGKHPEFQDITMHTVLLIIGIVILPICAIIDCFILYSKSKKALASLNAR

>gi|26553691|ref|NP_757625.1| hypothetical protein [Mycoplasma penetrans]

MRNKYRNKRDDYEYDFEDHEEDNLSN
HNNDYNREN
RNNDYSN
NQGSYNNQ
NNNGNGY
NNNGYNN
NLGGYNNQNNNY
GNNGYNNPNNNYY
GNNGYNNPNNNYY
GNNGYNNGPYM
MNNHYNQSRRIL

>gi|26553796|ref|NP_757730.1| hypothetical protein [Mycoplasma penetrans]

MQNNLKKS
IKTTNAESTNKNLYYDSFNGNMKTSVFINRIRDSQTVNLKDYSEKVIDYYIDQETNERVKTVKFKAKSIEGNIMIRENNP
SPDILEEILKQLKT
ISSDVSE
LKTDVAQ
LKTDMVE
VKADIVE
LKTDVAQ
LKTDMVE
VKADIVE
LKTDVSQ
LKKDVSRIMKCPTIVRELAEID

>gi|26553797|ref|NP_757731.1| hypothetical protein [Mycoplasma penetrans]

MCINILFNRIFILKIYYDSFNGNMKTSVFINRIRDSETVNLKDYSEKVIDSYIDQETNERVKTVKFKAKSIEGNIMIREN
NPSPDILEEILKQLKTIS
SDVSELK
TDVAQLK
TDMVEVK
ADVTQLK
KDVSRIMKCPTIV
SELSEID

>gi|26553798|ref|NP_757732.1| hypothetical protein [Mycoplasma penetrans]

MKFKAKSIEGNIMIRENNPSPDILEE
ILKQLKTISG
DVSELKT
DVAQLKT
DMVEVKA
DILELKT
DVSQLKK
DVSRIMKCPTVVRELAEID

>gi|26553926|ref|NP_757860.1| hypothetical protein [Mycoplasma penetrans]

MFITILNLKFSIVIFIINTWNINK
ICILLDDKNNVRIITSKAKCYMEFTKIKKEKIIQEVIPPPKDIDGSIEGMLRKILDKMDVIELD
LKTLKED
VTILKSD
VTQLKTD
MVEVKSD
VAQLKTD
MVEVKSD
IAILKADMIDVKSGKYCKNHNLYENA

>gi|26553982|ref|NP_757916.1| hypothetical protein [Mycoplasma penetrans]

MLLKKLLLIGGGVVVGSTALASAIAIPISQSQLNSSSAEVPTDPNGNNQ
PNDQNGNNS
QNGENNNN
QNPDNNG
NGGENINPAP
DANEDNNLIKDVKLPEGGVVFRNFGDGFIQKEESFAYQVYSKGISS
ENDLNNKLNEWFN
QNKDKDNFNGTHFFSVKDISFLSHFSNESYNYEVITYVNNSAKIIKTVDSSNN
QDSQNSDSSILKNKVYSLEVTLKPASGHAWTDGTSQSKTITVYGVYSGVSNKSDK

>gi|26553990|ref|NP_757924.1| hypothetical protein [Mycoplasma penetrans]

MLLKKLLLIGGGVVVGSTALASAIAIPISQSQLNSSSTEVPTDP
NGNNQPND
QGGNNNQ
NGDNNNNPQVP
DNDNNGNSGENVNPEVDNNLIKDVKLPEVEVFFRDFKSTSLDINFETIGDKVFNKSINS

NEELNNHLESWFNQNK
NNENFNDTYFVQSNLLGFHVPEYGYNYKIASFISGSAKINTEIL
NDENKNPQTFNKVYSFEIDLTPNFGHSWIDGTSGSKKITVYGIYSVS

>gi|26554019|ref|NP_757953.1| hypothetical protein [Mycoplasma penetrans]

MEIIKTKKEKLINEVIPPPKDKDSSIEVILGKILDKM
DVIELDLKTLKE
DVAQLKT
DMVQVKA
DIVEMKK
DIVELKT
DVAQLKT
DMVEVKS
DVAVLKA
DMIDVKSGKYCKNHNLYENA

>gi|26554020|ref|NP_757954.1| hypothetical protein [Mycoplasma penetrans]

MDQNNKNNNKPFKGQFSEKEFISVIKGCKIVNVDHFITDSNVFVDDKNNVRIITTRAKCYMEIIKTKKEKLINEVIPPPK
DKDSSIEVILGKILDKMEVIELDLKT
LKEDVAQ
LKTDMVE
VKSDVAQ
LKTDMVE
VKSDVAL
LKQDVAI
LKADMIDVKSGKYCKNHNLYENA

>gi|26554079|ref|NP_758013.1| conserved hypothetical protein [Mycoplasma penetrans]

MKILISPSKTQNKN
IDSFSNNEENILFTNKTNELNDLL
IKSINNEKFNS
WLELKNNSLIEETIKD
IKEFKNNKAYKAIEFYDGLQFKNILYSQLSETQKTKLNESLIIISGFYGIVFPNSYIKPYRLMLGSKIN
IPNYKNLYDFW
FKEFNQKLEELNPDKLIINLASGEYSKLIDNSIFNVINVD
FKLFKSNKYVSLSTFSKQCRGYFINQFLNINLDINKIKDLD
ILGFKFNN
ELSYKNNYIFTKAY

>gi|26554093|ref|NP_758027.1| hypothetical protein [Mycoplasma penetrans]

MNDKFNNNQKINSDSGSFTGPLKEKDFVRKIKGCKIVNVRDFEESFNVYENESENEITETTYIKGYVEMVVSRKYQKPII
LGPTDGRGGGTICGGDAHDEINKKLDYLIDEMK
EMKQDIK
ILKEDVA
VLKEDVA
VLKEDVA
VLKEDVT
ILKEKVS

VLEADMADVKAGKYCKSHNSPDFV

>gi|26554191|ref|NP_758125.1| conserved hypothetical protein [Mycoplasma penetrans]

MIDYYIDQETNERVETVKFKAKTIEGNIMIRENNPSPDILEEILKQLKT
ISSDVAQ
LKTDMVQ
VKADIVE
LKTDVAQ
LKTDMVE
VKADITQ
LKKDVSRIMKCPTIVRELAQID

>gi|26554204|ref|NP_758138.1| conserved hypothetical protein [Mycoplasma penetrans]

MTKNNKENKNNKVSKPFTGRYSE
KEFTSIIKGCKFVDVKNFNVSFDSYEDEKTGDIIETSVFEGNMKIYKTKRHRKPIEVPPPGDDDKKPTDLSNEIKFI
KDTLGIILSDMKEM
KKDINYL
KDDVSIL
KNDVSVL
KDDMVEV
KSDIKIL
KDDVSKIKQCPTIAKELAQL

>gi|26554253|ref|NP_758187.1| hypothetical protein [Mycoplasma penetrans]

MV
EPNEFDQFEFDQSYDQNLF
YPQEYVDTVQYY
QPIESIHDQQFE
QPVEYVQDQQFE
QPVEYVQDQQFQ
QPVEYVQTQFDQNIQQTPQN
ENAEYIDTSSWTEQDFINYYAQVDFYNKYPNAFVYQGKIYNSDLDQYVPSKIGDLPNEPNLQIVKAPEFEEKSIKRNDFL
KNN

>gi|26554281|ref|NP_758215.1| hypothetical protein [Mycoplasma penetrans]

MQCHLIILSKTRNKLGGTLMNKNKDML
KLVVDLAKKVTNVEFEDLQLNEDKKAGSVFVKF
ELEQDMSEKITLALVYKELKEMRT
EFRTDIS
ELKTDVA
QLKVDVA
ELKTDVA
QLKVDVAAIKACPTIQKELKLAQ

>gi|26554282|ref|NP_758216.1| hypothetical protein [Mycoplasma penetrans]

MVKTHML
KLVVDLAKKVTNVEFEDLQLNEDKKAGSVFVKF
ELEQDMSEKITLALVYKELKEMRT
EFRTDIS
ELKTDVA
QLKVDVA
ELKTDVA
QLKVDVT
ELKTDVA
QLKVDVAAIKACPTIQKELKIAQ

>gi|26554283|ref|NP_758217.1| hypothetical protein [Mycoplasma penetrans]

MVKTHML
KLVVDLAKKVTNVEFEDLQLNEDKKAGSVFVKF
ELEQDMSEKITLALVYKELKEMRT
EFRTDIS
ELKTDVA
QLKVDVA
ELKTDVA
QLKVDVAAIKACPTIQKELKIAQ

>gi|26554284|ref|NP_758218.1| hypothetical protein [Mycoplasma penetrans]

MVKTLYAKISSRFSKKSYKCWIWRSSINEDKKAGSVFVKF
ELEQDMSEKITLALVYKELKEMRT
EFRTDIS
ELKADVA
QLKVDVT
QLKTDMV
EVKADIAAIKACPTIQKELKLAQ

>gi|26554301|ref|NP_758235.1| conserved hypothetical protein [Mycoplasma penetrans]

MNLYYDSFSG
NMKTSVFISRIRDSETVNLKDYSEKVIDSYIDQETNERVKTVKFKAKSIEGNIMIRENNPSPDILEEILK
QLKTISSDVA
QLKTDMV
EVKADVA
QLKTDMV
EVKADIV
ELKTDVS
QLEKDVSRIMKCPTIIRELAEID

>gi|26554302|ref|NP_758236.1| conserved hypothetical protein [Mycoplasma penetrans]

MVVKQKGIY
MKNDLKKS
IKTNAEKTNKNLFYDSFSGNMKTSVFISRIRDSETVNLKDYSEKVIDSYIDQETNERVKTVKFKAKSIEGNIMIRENNPS
PDILEEILKQLKT
ISSDVSQ

LKTDMVQ
VKADIVE
LKTDVAQ
LKTDMVE
VKADIVE
LKTDVSQ
LKKDVSRIMKCPTIVRELAEID

>gi|26554420|ref|NP_758354.1| hypothetical protein [Mycoplasma penetrans]

MPEAKTTRSSVSKAPLVSVAANS
EANPTVNPSLNTNGLQQV
KVSPTLTG
RPSAQSV
RPNPTMVARPTAQPA
RPNPTMPARPTAQPV
RPNPAVNG
RPSPAVL
RPTPTRPTVISQ
RPNATYSRPAGSYSGK

>gi|26554434|ref|NP_758368.1| hypothetical protein [Mycoplasma penetrans]

MKSKNNDMLKLIV
DLAKKVTNVEFDNLQLDEDKNTGSVFVRFELEQNMSEKITLALVYK
ELKEMRT
DISELKT
DVSHLKT
DMVEVKA
DIAELKT
DVAAIKACPTIQKELLLTK

>gi|26554458|ref|NP_758392.1| ribosomal protein L29 [Mycoplasma penetrans]

MLMIKELRTKTDNELGNLISKLKIQLLEMRFKIANGEVEGLHKFKEIKKTIAMAMTVLSERNVKISFSTFSTQLIKQKNG
KQEIVSLSIVDERKDAVKSQPKAKPVEA
NKPAVSESKKEEPKKVETNNKPV
AKPVVKK
EQPVQSKPA
AKPTAKPVAKPV
AKTTASN
SKPATSKPV
SKPTGKPV
AKPAASN
SKPVASKPMAKPA
AKPTGKPVAKPA
AKPAASN
SKPTTSKP
ASKTTAKAPVKGKK

>gi|15827879|ref|NP_302142.1| conserved hypothetical protein [Mycobacterium leprae]

MAVYRVFEALDELGAIVEEARGVPMTAGCVVPRGDVLELIDDIKDAIPGELDDAQDVLDARDAMLND
AKAHADSMVSSATTESESLLSH
ARAEADRILSD
AKSQVDRMASE
ARQHSERMLGD
AREESIRIATVAKREYEASLNR
AQSECDRLIENGNISYEKAIQEGIKEQQRLVSQNEVVQA
ANAESTRLIDT
AHAEADRLRGECDIYVDNKLAEFEEFLNGTLRSVGRGRHQLRTAAGTHDYVTR

>gi|15827894|ref|NP_302157.1| probable histone-like protein [Mycobacterium leprae]

MNKAELIDVLTQKLGSDRRQATAAVENVVDTIVRAVHKGDSVTITGFGVFEQRRRAARVARNPRTGETVKVKP
TSVPAFRPGAQFKAVVAGAQRLP
LEGPAVKRGVATSAAKKAAI
KKAPVKKALAKKAA
TKAPAKKA
VKAPAKKITTA
VKVPAKKATKVVKKVA
AKAPVRKAT
TRALAKKAAV
KKAPAKKVTAAKRGRK

>gi|15828288|ref|NP_302551.1| proline rich antigenic protein [Mycobacterium leprae]

MT
DQPPPSGSNPT
PAPPPPGSSGGYEPSFAPSELGSAYPP
PTAPPVGGSY
PPPPPPGGSY
PPPPPPGGSYP
PPPPSTGA
YAPPPPGPAIRSLPKEAYTFWVTRVLAYVIDNIPATVLLGIGMLIQTLTKQEACVTDITQYNVNQYCATQPTGIGMLAFW
FAWLMATAYLVWNYGYRQGATGSSIGKTVMKFKVISEATGQPIGFGMSVVRQLAHFVDAVICCIGFLFPLWDSKRQTLAD
KIMTTVCLPI

>gi|15677322|ref|NP_274477.1| hypothetical protein [Neisseria meningitidis MC58]

MKKLLIAAMMAAALAACSQ
EAKQEVKEAVQAVESDVKDTAA
SAAESAA
SAVEEAK
DQVKDAAADAKASAEEAVTEAK
EAVTEAAK
DTLNKAA
DATQEAA
DKMKDAAK

>gi|15677503|ref|NP_274659.1| conserved hypothetical protein [Neisseria meningitidis MC58]

MKKLSRIVFSTVLLGFSAALPAQTYSVYFNQNGKLTATMSSAAYIRQYSVVAGIAHAQDF
YYPSMKKYSEPYIVASTQIKSFVPTLQNGMLIL
WHFNGQKKMAGGFSKGKPDGEWVN
WYPNGKKSAVMPYKNGLSEGTGYR
YYRNGGKESEIQFKQNKANGVWKQ
WYADGSIKTEMVMVNDEPAKILT
WDESGRLLSELSIRHHQRNGVVLE
WYEDGSKKSEAVYQDDKLVRKTQWDKDGYLIEP

>gi|15677912|ref|NP_275080.1| hypothetical protein [Neisseria meningitidis MC58]

MPSEPPSDGIAR
HPKSTIKMAKKPNKPFRLTPKLLIRAVLLICIAAIGALAIGIVSTFNPNGDKTLQAEPQHTD
SPRETEFWLPNGVVGQDAAQPEHHHAASSEPA
QPDGTDESGSGLPSPAAPKKNRVKP
QPADTAQTDRQPDDAGTQAEN
TLKETPVLPTNVPRPE
PRKETPEKQA
QPKETPKENHTKPD
TPKNTPPKPHKEILDNLF

>gi|32490818|ref|NP_871072.1| dnaQ [Wigglesworthia glossinidia endosymbiont of Glossina brevipalpis]

MKINTERQIVLDTE
TTGMNKNGPHYYGHRIIEIGA
IEMINRRLTGRCFHTYLKPDRLVEIEAFKIHGISDEFLFFQPTFEEIMEKF
INFIKGSELIIHNSVFD
IGFINNE
IQLCNKNLNNINYYCSVIDTLKLARNIFPGKRNNLDALSDRYGIDTTKRILHGALLDAEILSNVYLLMTGGQIP
INFSKKKN
NNFINKKIKNN
FNNFKKNKKYNKIIKPIVIYANEEEIHEHKKFMKS
INFINKNN

>gi|32491283|ref|NP_871537.1| yfiO [Wigglesworthia glossinidia endosymbiont of Glossina brevipalpis]

MDIKFKNIIVIITFLLAIESCSLKN
TNFIQYKIKNEENNLKNNELYK
NRYKKYKRYLYFINKY
NNYYYYKEAKNNII
NLYYTYKDYEIENNSILEEIIDQFIESNPRCPYIDFLIYIQGLINMDLDKNT
INFFIYKKKFIE
NTKYAYLALNNFKKLIYKY
PNSDFYKSS
KNYIFHLNERIAFHELQIIKFYFKKHAYSASNFRILEMINKFPNTKSSK
ESLIYFKY

>gi|16329525|ref|NP_440253.1| unknown protein [Synechocystis sp. PCC 6803]

MASFLALLPRSLTTFLYAVAALLRFYGNIDTTPIPRIPLTIFGWSFLAFTLGTAALLVNLGLEWNTGNR

SRNREIETRERETR
RDNLADE
ERNRASE
EREKADR
ERDRADQ
ERNRADQ
ERKRAASRARIQNRGFVLQTRYQLAPSPEARATLIDFLSFLQEYGE

>gi|16329754|ref|NP_440482.1| unknown protein [Synechocystis sp. PCC 6803]

MKNKLGIILVCNCLFSLFINRGVGAEETINIGDCSFNGHPLYGKI
QLVGSFPDLTV
QVVSSFPDLKV
QLVESFPNQCGQW
QIVTSFPDTKV
QIVESFPDVKI
QYVDSFPGLP

>gi|16330057|ref|NP_440785.1| unknown protein [Synechocystis sp. PCC 6803]

MATTADEVWALLGELIESQKE
TDRRMQE
TDRRMKL
TDLQMQE
TDRRMQE
TDRLIKENGRFIKELGKQIGGLGKKFGSFTEGLALPS
MERILQERFGMEVISPSVRASKGGQHLEIDVLAYANSDINTAYVVEVKSHPREESITQLKNLLQKFRTFFPEHKDKKLYG
ILAAVDWSDDLKQQALREGLYVASIHDEVFELAETPPDFEPKLW

>gi|16330154|ref|NP_440882.1| unknown protein [Synechocystis sp. PCC 6803]

MTMNPLPLQKEAEKRSPRLNSPFKSLMRI
HWWMALAYLIVFCTGTLMVNLERGQFLRSELYDFHKSIGILTMAL
LTWRIFTLLQ
VWWRKYTKRPPKLSL
AWWRTFTLHLSLYVFMGVIPISGFLLSNSFKANNVRFFSITLPDIFPQNEAMVEIGRNA
HFWLAYTFLIFIFIHML
TYWKVIR
ANWRRWVNFIGKFQQQT

>gi|16330278|ref|NP_441006.1| unknown protein [Synechocystis sp. PCC 6803]

MAINNRSSWTAFVIALTAIGTVAEPGNALATNSQNALSTRLNNISSALQQRANQMSPEETPINRANLQAGFANGG
GGGGFGNARR
GGWGDGA
GGGGFANV
GRGGWADGS
GGGGFANVNNPWGNGWG
DGGGFVNRG
GGGGFVNRW

>gi|16330460|ref|NP_441188.1| unknown protein [Synechocystis sp. PCC 6803]

MTNFPPLLRE
AKRDNRSIHFSATIPDFQTFVAIVTTLLVFLTGVALKYDTIYFRNYEISIPQWAGLSFAGLTFSFVVLDQLLRGFVRLEE
IQQTT
EARNRRA
EEEQRRSEDRARTVAREAKQDREREEELA
RAENERA
EARDRRA
EAREREARRVGQ
EARRTRIENQYRLALIRYQLEPNSPHRRQLENILALLDEYGDTL

>gi|16331083|ref|NP_441811.1| unknown protein [Synechocystis sp. PCC 6803]

MLMNFLVICQKFFTPNLFPWKAIARVQREKPQSLGRWQFVVRTGILVATFILALGSLASPSLALDYNRG
NLVGADFSHQ
DLRGSIFDHA
NLRGADFTGA
NLQGARFFSANMDGAILEGA
DARGVDFESARLTHA
NLRNARLEGSFGTNTKFGEV
DIEGADLTDIILRPDTEDYLCGLAKGTNPVTGRETKETLFCP

>gi|16331309|ref|NP_442037.1| unknown protein [Synechocystis sp. PCC 6803]

MDNKNENQENSQIDDTDIAAGAIAVG
TVAGAGLSA
TIGGMGL
TVAGAGIGIGMA
PVAAAGAVVGSAAFGAKKAIEGQDPMAVGAIATG
AVVGAGTSA
ALGGMGLALTG
TAIGIGMAPI
AAAGAVV
GLAGYGLYKMFGGE

>gi|16331795|ref|NP_442523.1| hypothetical protein [Synechocystis sp. PCC 6803]

MNNKKLSLTLPLVGIL
ALTACNQPKEEIDPLSQLREKGECLDC
NLAGADLREF
NLENARLNRSDLSGA
NLSGVNLRRALLDRA
NLTGANLSETDLTEA
ALTEANLAGA
DLSGANLERSFLRDV
DLTGANLKGANLAWA
NLTAANLTDVDLEEAEFWETTMPDGSRR

>gi|32265904|ref|NP_859936.1| conserved hypothetical protein [Helicobacter hepaticus ATCC 51449]

MKE
YYESGQLSRET

LFENGKTNGVVKS
YYKSGQLEEEV
TYENDKVNGVRKE
YYESGKLKYEAPQKNNRLNGVVKS
YYKSGQLEYESPFENDKANGVRKE
YYESGKLKKEILYEYDRTITRKE
YDESGKLIKSQQNK

>gi|12723308|gb|AAK04529.1|AE006280_5 UNKNOWN PROTEIN [Lactococcus lactis subsp. lactis]

MKKIIYSIGLISLLNFGTIAYGET
ATGTSASSGDATAQSSTGTS
STSNSAS
SASSSSQNTSSSEGTN
STSQNQT
QTSTTAQNQTDNVGQNQTPPVVENRDSS
TDKNSGSQK
ETKNSTTIINETVNNNAKGNNILLMT
LTSLSASVTIIGGIIAFVRHLRGRGKKKNPVSKAQSSKNGK
TPSNSMTAEPSVSSKKTPNEEKVLTETQQINMQADELLKNHKKQQKEYRRRRR

>gi|12723537|gb|AAK04732.1|AE006296_6 DNA-directed RNA polymerase delta chain (EC 2.7.7.6) [Lactococcus lactis subsp. lactis]

MKITALEGQKISELSMIEVAHALLEQNGKEMQFSEIIRAIQDY
LEKSDDEIKASISRFYTEINTDGSFIPLGNNIWALRSWYAIDEIDEEVIALDE
IEDEEEEKPAKKRKKVNAFGIEDEID
PEDEEGT
KETTEEDMSYDTQA
EDEDKDDVAAYDAELAE
VELDNVDEEVDID
VEDDEDD
SDDTDED

>gi|12724400|gb|AAK05510.1|AE006372_9 prophage pi3 protein 39 [Lactococcus lactis subsp. lactis]

MNKKLITTAVVAAGIFGSATFGAYAANAWAGHQNMVAVQQNISILKQRLLDRN
EQLKQANNSSQQYA
DQLNQLN
NRINQLK
DQINQDNSNLQNQAASYQ
NQLNVLNQQKEEVA
RQLNQANQDKANMA
QQISDLNSKLTAAQQKT
DELSQAVTDAQQTKDLSDDAVNATK

>gi|37524852|ref|NP_928196.1| hypothetical protein [Photorhabdus luminescens subsp. laumondii TTO1]

MRELTIQEFDL
VAGAGAGETAGEI

AGAAAGGAAGRV
AGSVAGGYAGA
AAGAAIGSVIAPGPGT
AAGAVIGVVVGRF
AGGGAGA
AGGAAAGRY
IGGEIGGRID
SGGNSNGAGTNYCNSSGSNYCH

>gi|37525718|ref|NP_929062.1| hypothetical protein [Photorhabdus luminescens subsp. laumondii TTO1]

MPYRQLRNFSVSITAVVPRA
LPYRQLRKMKLIVSLTIFRA
LPYRQLRNTMLGYAALSRSA
LPYRQLRNLGPHAPQRGMGA
LPYRQLRKVCTLRKATKWCA
LPYRQLRKNPRSLSVIASGA
LPYRQLRSAITALGC

>gi|37525731|ref|NP_929075.1| hypothetical protein [Photorhabdus luminescens subsp. laumondii TTO1]

MTMDMSS
LPYRQLRNRRFERDSSCNCS
LPYRQLRKYRIVSKTYPLCS
LPYRQLRNISGTNPPLIKSS
LPYRQLRNVEQMPTRVVSSS
LPYRQLRKVSMCWWVSRLSS
LPYRQLRNHVICTNLLTTGS
LPYRQLRKLAQITPTRKLCS
LPYRQLRNPIFK

>gi|37525732|ref|NP_929076.1| hypothetical protein [Photorhabdus luminescens subsp. laumondii TTO1]

MPYRQLRKLIAGEIAAVSCS
LPYRQLRNVVFHVVIYCNCS
LPYRQLRKSGVRKSRTQISS
LPYRQLRKPMCSADSLTSSS
LPYRQLRKISEQRNIVDSGS
LPHRQLRKLRKLSLMLRARS
LPYRQQRTRLSALKKQPIFFSICINCYAFVHILHY

>gi|37525942|ref|NP_929286.1| hypothetical protein [Photorhabdus luminescens subsp. laumondii TTO1]

MPVRQCTAHSPFVCGRSFSK
LPVRQCTGHAIASGSFQISK
LPVRQCTSPALLPFVGIISK

LPVRQCTNPHYIEECKRISK
LPVRQCTFKE

>gi|37526068|ref|NP_929412.1| hypothetical protein [Photorhabdus luminescens subsp. laumondii TTO1]

MAKIYEFSANDIATIESHKHLPNTMKYGNGNGGGGGDMLEARVAKLESDVEHIKATMSDVKADLKSVISDVGTMKTDVAV
IMQKLDVINNTLSKKPSEEQ
LNARLTKSENR
FDSKFTEVEGK
FDARLTKSENR
FDSKFTEVEGK
FDARLTKSENR
FDSKFTEVEGK
FDARLTKSENR
FDSKLTE
VESKQTESVLIETLESSRIDNTPFTFLGNAVANAVDSVSDEWASLPFASKGATLSHVISPLAI

>gi|37526530|ref|NP_929874.1| hypothetical protein [Photorhabdus luminescens subsp. laumondii TTO1]

MIDQNNSNNIVTPNVTISDVTVSSIQIAPEIARDNVHVIIVDDKGAPRDFLLDNSD
GGYQGSGSLIALILAHLNSKKVTIRTTGYEYNKHGYINGVLIKEDKG
GGGTGGGTGGGTGGGT
GGGTGGGTGGGTGGGTGGGA
GGGTGGGAKGSAGGGSGG

>gi|37526690|ref|NP_930034.1| hypothetical protein [Photorhabdus luminescens subsp. laumondii TTO1]

MSG
SDDKKIEFNTSDVLNALLHRGQNMQHLATQESVDNLRRETEVNFA
SLDKCFE
SIDKRFE
SVEKRFE
LIDRRFE
SIDKRFD
VMDKRFE
SIDKRFD
KLEGKFDRLQWFIVAAALALLFKEYIFKAAAVIQ

>gi|33862287|ref|NP_893847.1| RNA-binding region RNP-1 (RNA recognition motif) [Prochlorococcus marinus str. MIT 9313]

MSIFVGNLPFRAEQEDVIELFAPFGEVANCALPLERDTGRKRGFAFVEMADESAEPAAIEALQGAEMMGRPLRINKAEPR
GSAPRRG
GGYGGGGGGGG
GGYGGGGG
GGYGGGG
GGYGGGG
GGYGGGGYGG
GGYGGGG
YGGGGGGGDRGSGARGWED

RSYGGASPETSSYDEGRTRRRRGGAAPEGEGGDSY
SGYGGAEG

>gi|33863131|ref|NP_894691.1| conserved hypothetical protein [Prochlorococcus marinus str. MIT 9313]

MNTADKTGDDKASAHCGSKPKKLA
IGIAPLGTIS
IGIVPMGVVC
IGVVPMGVVS
IGVVAMGVIN
ASVVGMGLLS
VGVSTMGVWTAGPLS
MGLVSLGGSEIKNSHQHHLPVSGKSSSQTDPRLMAYPTQLEAEVKARELGC
KGVHQMGTHWMPCDQHSQH

>gi|33863420|ref|NP_894980.1| Type I antifreeze protein [Prochlorococcus marinus str. MIT 9313]

MAYPESQVVMGGLVHIPIIIGVFWALNNLTTGGS
KAKKAAEA
QAKQAAEEA
AAKAAAEA
AAKQAAEQA
AAKAAAEAA
AAKKAAE
AAASAAPAATAATPVSGEAETSQASNNDTQATPAPDQEVL

>gi|33863420|ref|NP_894980.1| Type I antifreeze protein [Prochlorococcus marinus str. MIT 9313]

MAYPESQVVMGGLVHIPIIIGVFWALNNLTTGGS
KAKKAAEA
QAKQAAEEA
AAKAAAEA
AAKQAAEQA
AAKAAAEAA
AAKKAAE
AAASAAPAATAATPVSGEAETSQASNNDTQATPAPDQEVL

>gi|33863785|ref|NP_895345.1| conserved hypothetical [Prochlorococcus marinus str. MIT 9313]

MTPRSLLGIYVLLASSAVVVQSAEAAVYNKPNLSNPIEAKIQTVRTGSWKALLKENQLSSNERNVARAWG
NGGGRAW
GNGGGRAVRRA
WGNGGGW
GNGGGGW
MNGGGGWMNG
GYGGRGFANW

>gi|33863962|ref|NP_895522.1| TPR repeat [Prochlorococcus marinus str. MIT 9313]

MTSMPRHTAAIAVGVATAAVTAVGLSFFRVGKPLLIVSTAVAGVAGVAVSARREGEENDVNAFIKSGNAKY
GIGDYQG
SIEDFSKALIIDPQY
ALAYYNRGNAKDEIEDYKG
AIDDYSKALEIDANH
AIAYNNRGVANRKSGDNRA
AIADYNKALEIDSKL
AVAYLNRGISKENDGNLTG
ACADWREASALGDEIAAKWVKESCR

>gi|33864143|ref|NP_895703.1| hypothetical [Prochlorococcus marinus str. MIT 9313]

MGTLENK
GGTINNQGVLFCTYAPHFPEASGSVDGGKIYLYD
SGTLNNNAWLCIKARSRLDIDA
SSTLNNLSFLQIIGDEG
GGTLDNN
GGTVNNKCGDTPPGEDDDFGGLKVY
AGKLNINNGGKLYSYGSSTIKTGSE

>gi|33864143|ref|NP_895703.1| hypothetical [Prochlorococcus marinus str. MIT 9313]

MGTLENK
GGTINNQGVLFCTYAPHFPEASGSVDGGKIYLYD
SGTLNNNAWLCIKARSRLDIDA
SSTLNNLSFLQIIGDEG
GGTLDNN
GGTVNNKCGDTPPGEDDDFGGLKVY
AGKLNINNGGKLYSYGSSTIKTGSE

>gi|54020446|ref|YP_115819.1| 30s ribosomal protein S6 [Mycoplasma hyopneumoniae 232]

MPKYEIMTILDPKAEMAIIDNLLKTVFGDNSTEKLRKLETTNLAYPIRKSKIAQYFLVDLNAPTNLIEEFVRRANITREI
WRYLIVNLDSEKGLNKKPKIR
ERNRKYTPRRDRFEKPNFRGNPKSRFDQQDQQA
TKNQQNF
QQNQQNQTSQYR
ENSRQNQDDFQQV
SSNQQNF
GQNQQNQSGYHR
ENNRQNQENIHQN
SKNHQNQTSQTQ
RNRRQYQPI
KNPKFNQKEKEN
YNNKKPQSSN

>gi|47458865|ref|YP_015727.1| hypothetical protein MMOB0300 [Mycoplasma mobile 163K]

MSF
GVVSFLCILKQLHKQSLQTVRF
GVVSFLCILKRKNVYYLDTDSF
GVVSFLCILKLFFWFISWAECF
GVVSFLCILKQMKQTTKNSQRF
GVVSFLCILKQEREYNSLSFSF
GVVSFLCILKLNTESIGAWVCF
GVVSFLCILKQSIK

>gi|47458932|ref|YP_015794.1| hypothetical protein MMOB0970 [Mycoplasma mobile 163K]

MQEKSFKTLKNWNFLKSLY
KKDKIKGEVLKKKILEIIRKSEIQ
KLEEIFNDGNSYKYIIEGVFIKKEATKSNQN
ILNAIENLSKSFDTKIENLSSGIKSINT
RLDKIEN
RLENVEN
RLEKVEN
RLENVEKDVVLIKSLPTIQKELKARNI

>gi|47459072|ref|YP_015934.1| 50S ribosomal protein l23 [Mycoplasma mobile 163K]

MNHNEIIKYPLLTEKSYKTMAQNVYVFAVDRRARKIEIKDAIEFIFEVKVEKINLFNVPQKEKKVGRFKGLTNSYKKAYV
YLREGKINIFPEEIEKEQKVEKQLI
KEKSTSELKLEEKIAAKIAAKEQSEIKDEIVKSVP
KKVTSIK
KEITSKEVTPI
KKTTSVK
KETSPKEITSV
KKTTTKEKDKTLVVAKKATIDTKVKSTTT
KKTTTKKV

>gi|47459272|ref|YP_016134.1| expressed protein [Mycoplasma mobile 163K]

MMKINNEIKDLR
EIDVYDENGKKLKGIVLFTFEEAGDTFICYTIKEEVFFSKLNEDNSLSNL
EEDEYKIVEQIWEDFI
QSDKFKELAEKYNFEGIGGPLVDGGDDEDDD
DEDEDDED
DEDEDDEDEDDDEEYYGDYDE
DEDEYEE
DYEEFSKDIKDSYKNLIEKRNIIN

>gi|42560611|ref|NP_975062.1| Hypothetical transmembrane protein [Mycoplasma mycoides subsp. mycoides SC str. PG1]

MCECQIIMKKLFNILGINFLLVLSIITAIYLKNNNNTSIILKNEST
NETNKKR
LEELEKE

LLETKDLLDKEYTFSSKL
LEENKKLQGILNDYLKIKAE
LEEFKELLDNEYTLNNEL
LKEKEKFKIKISDLEKEIENLQNQL
IQSNKNVTQLENI
IQEIKKTRTNK
ESENKKINKQLEEKNLN

>gi|42560833|ref|NP_975284.1| Prolipoprotein [Mycoplasma mycoides subsp. mycoides SC str. PG1]

MKKLLYTLSTLVVGVFSTVSLVSCTKKNNNEPEQLSKLE
QLKNEIS
SLETNVN
QYDEQINKTELEKQQLLEEIQKIEQEII
NLKTEENNKTKELSLLT
NKKNSINQQID
NLNSQIISIDKTSKKEKEKINLLTEEIQKIKN
NLTNATNQKNINLEQIS
NLENQVNQLKTKTSQLEKE
LLEKKLNKEQLLLNKK
ETESKVNKFKAALDNSNKQKEDLIN

>gi|42560936|ref|NP_975387.1| ribulose-phosphate 3-epimerase [Mycoplasma mycoides subsp. mycoides SC str. PG1]

MKIVPSIYCANLIDLKTEITNLTKAQINQIHFDVMDGVFVKNYALSNKLVDEIKQQFPNLIIEVHIMGIDLLDKISLFKN
ADYLTFHYNA
INDLNQAKLW
INEIKKYNLKPGIALDLNNE
INDIKKI
LNEIEVVTFMSIKPGFTGQKFEESSWTK
LNLIKELKLKYPHLKFQIDGGVR
WENIKKLIDCQLDWIVVGSLLFN
ENDYLKV
VSRIKQLETN

>gi|42561100|ref|NP_975551.1| Prolipoprotein [Mycoplasma mycoides subsp. mycoides SC str. PG1]

MKKLLTIFGSISLISMTSITSIACKNINKSKKE
DPSKPDR
TPKTPEN
DSTKPEENQPKKP
DDKKPNNNDNSSNDSN
SNSNPNNKP
DSTKPSE
NPKKPNEQH
QGDDPNNHQPQADQLEKSEQPKRTKEQNFELIKKYGAEVLALLESLSEDKTEQ
LYKKPENSELLKLVTNIGQFYQKLGSSTTIDDFEKEFKEKKFLNKLYEEWDKTVTEYEKQKDSILEILKQ

>gi|42561280|ref|NP_975731.1| hypothetical protein MSC_0755 [Mycoplasma mycoides subsp. mycoides SC str. PG1]

MFYSAKSFN
QNISSWNISNVVNMSAMFFEAEKFN
QPIGNWNVSKVTDMDFLFDGAITFN
QNINDWKTDSLTDLTYTFRSAISFN
QPLNNWNVSKVKSMDGMFTEATSFN
QDINLWNT
NNVESMNGMFSSPDKFN
QNLSSWDVKKVGNAQNFTNALKSVMTKDKLPKFNQKYSDDKYFMVVKK

>gi|42561388|ref|NP_975839.1| Hypothetical transmembrane protein [Mycoplasma mycoides subsp. mycoides SC str. PG1]

MIFIKLYKNILTNKEKAMRDYNSSNNQ
NDRYYQNDNKNNNY
HNQYYDNRYD
QNQYSDYD
DDRYYDQDYR
YDQNYQDQNYEQQEYNLQQVEYNNSNSDQQVKFIPDKKIKKIVRFNSIKSFISIFILVGLIGYFTVFL
INHYYQFLYDPNN
IKQYYKTLQGWLFTMNGFKVWHLSVIIAILSVCCIIYIVVASTLFSN
YNKYLKDMQHRTEEYQAQKLCIPYPKKPEQGIPPLLIKKMYEKQI
KKPYYAN

>gi|21219450|ref|NP_625229.1| putative secreted proline-rich protein [Streptomyces coelicolor A3(2)]

MHGPEPTRSAAAGRVLALVTLTFTLVGIPVGAARADACAYASTGPDGTDAVAVAGSASWPTF
PPCPKPTPTPTP
PPDPEPT
PPPPEPTP
PPAPKPPPPK
PPAPEPEPPP
PPAPRPA
PPRPAPEPSP
TPTPAPT
PPPVRPA
APAPAPSVTP
TPTPRPSLSPVHYPRHRAARAPQRSSRGTTSPLVFVLLITVPAVVAVAALRAR

>gi|21219785|ref|NP_625564.1| hypothetical protein 2SCG18.24 [Streptomyces coelicolor A3(2)]

MGNSKNVQDIFKDLVGDMTDSWKEAFDDLLNRDDDSGSSRGSGSSRDSGNAVSTLAGLPPNALAAL
AGAVNPATAL
AGVTNPLAALGGI
GAAGNPLAGLA
AGAAGPLGNIA
GAAQNPLAALTGA
AGGGNPLAAAL
GGAGNPLAAL
GGAANPLAAV
GGAAGALGGLGQLA

GGLGHAA
SGAGDLMALPAQLTQLLQVLPKLLEAVQGLQNLTNLPAQAGQQGLAALQGLTGQQ
GPGGKPSQGGQSSTSG

>gi|21219871|ref|NP_625650.1| conserved hypothetical protein [Streptomyces coelicolor A3(2)]

MRNHGYER
GHGGHGP
GHGEGGPFGRGGFDGRRAAFG
PFGPGGP
GGGPGGPFGP
GFGHGGPWG
GRGRGGPRGRARRGDVRASILALLKDRPMHGYEMIQEIAERSGGAWKPSPGSVYPTLQLLEDEGLIA
SESEGGKKLFALTEAGRTAAEEGPEAPWEEASRGVDWEALNDIRQAGFGLMEAFGQVWKTGSKEQREKALAVIGDARKKL
YLILADED

>gi|21220555|ref|NP_626334.1| putative DNA-binding protein [Streptomyces coelicolor A3(2)]

MVAKKKTTAAKQS
AAEESAAEQDPARKK
AAEKTAAKKS
PARKTAAKKS
PAEKTAAKKSTAKKSTAKKVGAAE
AAEQTGATTV
VAKKTPGTA
TAAKTAVP
KARGTAAVPGDLAVRPGEEPWTPQEVEEARGELQSEADRLRTEIDTSERSLQGMMRDSGDGAGDDEADTGSKNITREHEL
ALAATAREVLSQTERALDRLDAGTYGLCENCGNPIGKARMQAFPRATLCVECKQKQERRY

>gi|21220659|ref|NP_626438.1| hypothetical protein [Streptomyces coelicolor A3(2)]

MPWAGVVRGVGHVTRR
GLGTRGGRGREGDAARRCGVPLRPPVPPQRHDCPQRRRRGGVGPQP
GRGRAVSRVRREA
GRGARAGLSRARREA
GRGARSGL
SRARRGG
GCGAWAGRAAYGEEGACR
GVSARSGWRVNAPSLFK

>gi|21220867|ref|NP_626646.1| putative membrane protein [Streptomyces coelicolor A3(2)]

MDYCHPCRRHLNGALACPGCGTPCH
ELPASEGHLTAGEKNPPRPARDAAADAETPEEALDEGGDAG
ADDAAPGGRSSRRDRKAAAHRRRRRRTLLVAAGFVLAAGGLSLAELGMDAP
HSPEKPAAAGGETTDGEATREVGEASAVPVGDAPAATTTTT
SGTPSPGASASTTVSA
SASESPDAAESPTA
TGPDSPSA

SSPSSPGPTGPAN
PGPSSPGSDDDPPPT
TTPTPPEPPDPTPTETCDRFLWWCT

>gi|21220961|ref|NP_626740.1| putative membrane protein [Streptomyces coelicolor A3(2)]

MRDAAAVLRPRLPLRPPPEEEGRLEDALVPDWLLARVGDAVFF
AEVFFAAVFCAGDFC
AAADFAAVDP
VAVDLAALDPVVVDL
AAVDLAAAFSAPVVRRV
AAADLAGVAVPAPAAFVPLAL
APAVFAAAVLPVAAF
AAAVFAVAAFAVAALLVVFLADVRAGLLPTAPRAPAAS
ASPPFAAPPSPPRPFCPFCPRGCVSQATAAPVTATGHSTSPATPSTPAPV

>gi|21221446|ref|NP_627225.1| hypothetical protein SCE33.05c [Streptomyces coelicolor A3(2)]

M
PPGPVDGT
PPGPPDTWRSLPPGLRSGAPPVPPGGV
ADELYDGV
APGPYDGE
ADGPYDGE
ADGPYDGE
APGAEEGV
APGPPWGGGGEVGGDAGAVGRGWVRVMTLRMGVPVRRPG

>gi|21221714|ref|NP_627493.1| conserved hypothetical protein [Streptomyces coelicolor A3(2)]

MPTTVNDLLRTYPADLGGVDREAMAR
CIEECLR
CAQACTA
CADACLSEPTVADLTKCIRTDMD
CADVCTATAAVLSRHTGYDANVTRAVLQA
CATVCAA
CGDECARHAGMHEHCRV
CAEACRS
CEQACQELLAGLG

>gi|21221720|ref|NP_627499.1| putative integral membrane protein [Streptomyces coelicolor A3(2)]

MSSTRVPPGARGSYPAPGLALTPGPVRNQALTVIAVPFVLIVVLAVVLATTGPDD
GAPATDGLY
GSGALPDW
GSGTTDDGTTDDGTS
GDGTDPDGTY
GNGTYEDGTY
GDGTYDNDPY

GDGTADETPTDDTTTDTATDDTATEDTTANGSGPEATVTAYFDAINNRDFAAAWELGGKNLDQDYDSFVA
GFGTTQRDDVT
VTGTSGDDVSVTVVAWETD
GTQTTFEGTYTVSGGVITSADMREGN

>gi|21221783|ref|NP_627562.1| hypothetical proline-rich protein [Streptomyces coelicolor A3(2)]

MQLGRLDPAGSASSA
PTPVTAPPPPPPAL
PDPADPPPDSPP
PDPEPFPALPPPE
PDPEPEPLESPL
PDSSRPPGAWLIAG
PDPEGPGVIDGAGFLPLA
PSSRPLP
PSPPPETTIHVINSANRAKRDSSTALRRQ

>gi|21221808|ref|NP_627587.1| hypothetical protein [Streptomyces coelicolor A3(2)]

M
PTPYGPRGGMAFGAQELRVLRRALALALNPSPVSAQEAQDCLHLADSLDEAMREGARLRAFLLADLARYRAALPGTATGY
LALLDEALGAGHRPDADDLAALGALRGTPAAAALLARCA
PETPTAPTASTVPLPAARARTRLHAVPGGRQSG
KEPPQKP
AAKPSPR
PAPAPAT
PKRPVPT
PGEVFPPRK
PAPPAPPAPDSRHHLAAG

>gi|21221970|ref|NP_627749.1| membrane spanning protein [Streptomyces coelicolor A3(2)]

M
RAATGGGRIGGGRAV
RGRTGGGRSA
RGRTGEGRAA
RGRTGGGRAV
KDQTGRGRAV
RAATGGGQAM
GDATGEGGAV
RDGAGGSRAMNASDRGSATVWSVAAIAVLCFVFGIVLALGQAVVARH
RAAGGADLAALAAADDWAQGATAACERAGRVARAQEVRLVRCVLTGQISDV
TAASGSGPFTAEVRARAGPAPLETPAPAPPLPSARSLPPTPPAAPAPPPPPPAAAP

>gi|21222147|ref|NP_627926.1| putative secreted protein [Streptomyces coelicolor A3(2)]

MLRPGADFLGPPPGRYEAIRRAASRRRLLRAAAGAAATAAAAALIALPLNLAAPAP
PARPTPPLAPA
PAAPTPEA

TPSPAPPSESTRAPDTG
PALPTGESA
PPGATPS
PSEPSPSTPSSEP
SAEPSDP
PVRQAPTNAPG
TVVPTPSAHGPAEQGGPAVTDGTAATTG

>gi|21222314|ref|NP_628093.1| putative single-strand DNA-binding protein [Streptomyces coelicolor A3(2)]

MAGETVITVVGNLVDDPELRFTPSGAAVAKFRVASTPRTFDRQTNEWKDGESLFLTCSVWRQAAENVAESLQRGMRVIVQ
GRLKQRSYEDREGVKRTVYELDVDEVGASLRSATAKVTKTSGQGRGGQ
GGYGGGG
GGQGGGGWGGGPG
GGQQGGGAPADDPWAT
GGAPAGGQQGG
GGQGGGGWG
GGSGGGGGYSDEPPF

>gi|21222497|ref|NP_628276.1| putative lipoprotein [Streptomyces coelicolor A3(2)]

MRPLYVPVRMAATALAVVAAAGCMSVGDDGSGGRAKPSQSAGERGGEVPEGGPAVSGGNAGYGAAAADDGHGKGKKGAGG
DRAGGDGKGKGGEGGKGKDAS
ASAPPAAHP
TGAPPPNGGNHGPEP
TRTPPAEPTR
TAEPEPTP
TEEPPASPTPDPT
VAEPSSS
AHEPPGTQLEQ
REPAPAAGTPA

>gi|21222644|ref|NP_628423.1| hypothetical protein SCD8A.22 [Streptomyces coelicolor A3(2)]

MYA
GGGPSGGGAPAPASYGGGGSLYG
GGGSSGPGRDGGGSSRVRLSTTVAAVAEARAAVVCPARRERRKENRWTGGMSPDTAP
GGAVSGVVSAGPACV
ATGISGRAADG
RGGVSGATGSGAPVEAPGDTGLRRGTALAGARWTTGGAAPPPPGSAGGV
RCGAAGAASVAAAA
KGGTADVARAA
KGGVADV
AGGVRAV
KGGVADVAAAVVVAV
KGGVVGAKAWAVARCTGGGT

>gi|21222983|ref|NP_628762.1| NADH dehydrogenase subunit NuoB2 [Streptomyces coelicolor A3(2)]

MN
PGEPADPAP
VPLPEPKRLGTLARLAPEPMKVVLNWGRRYSLWVFNFGLACCAIEFIAASMARHDFIRLGVI
PFAPGPRQADLMVVSGTVTDKMAPAVKRLY
EQMPEPKYVISFGACSNCGGPYWDSYSVTKGVDQIIPVDVYVPGC
PPRPEALLQGILKLQEKIARESLGERYGVSPRPSTAALQSGLVR
PPEPEPEP
EPEPEPEPGQGPGQGPGSGEGTDTGSDAGRGR

>gi|21223120|ref|NP_628899.1| hypothetical protein [Streptomyces coelicolor A3(2)]

M
ALEPDPP
PPSPHPPSPP
PPSPHPSS
PPSSSSP
SSSPSPPVPAPPVPGLGSLPRAEFGFPGPLRDRLVAAVLDGSKTSTTGLVADYEHEGEPLPEAGRREVVVDSRERPVAVI
EVTEVRVVPLAEVDLAHAVDEGEGDTSVAGWRAGHERFWHGAEMRAALGDPGFTVDDATPVVLERFRIVTLLPAPGG

>gi|21223303|ref|NP_629082.1| putative membrane protein [Streptomyces coelicolor A3(2)]

MTTTGTPTEPRYAWGRRGRHRRPRPRKVLLAAGGLALAAGALSVVRLL
PGPGGDTGGVRAGAEPRVESGGRDPS
PDPATAPAAGTVGEPSGAGPSATSAMGGLSGS
PGPGLIPGLAPAPSATGPAV
PLPTAPPVRTP
GAPETPK
PGEGAGERERERGDDTGERA
PAPGTPQPSRSDRPV
PHPVPPPP
PGPGTPPPSDPA
PEPADPGLCVPVIGLCVDLLGD

>gi|21223507|ref|NP_629286.1| putative ATP-binding protein [Streptomyces coelicolor A3(2)]

MSLPLTRRIARAALLVA
AGAAAGVGA
AGSASAAPELPATPNLGGLTALDGANVGNT
AASAAESATET
AGDTGGKAVKKAVPT
AGKTGGSVVKKATPAAQNA
AGDAAGS
AGDIVGDTAATATKGGLPTDALGKGGLPVQGLPLG

>gi|21223911|ref|NP_629690.1| histone-like DNA binding protein [Streptomyces coelicolor A3(2)]

MNKAQLVEAIADKLGGRQQAADAVDAVLDALVRAVVAGDRVSVTGFGSFEKVDRPARYARNPQTGERV
RVKKTSVPRFRAGQGFKDLVSGSKKLPKNDIAVKKAPKGSLSGPPPTISKAAGKKA
AAKKATG

AAKKTTG
AAKKTSA
AAKKTTAKKTTGAAKT
TAKKTTAKKS
AAKTTTA
AAKKTAAKKA
PAKKATAKKAPAKKS
TARKTTAKKATARKK

>gi|21224001|ref|NP_629780.1| putative membrane protein [Streptomyces coelicolor A3(2)]

MAVLSIISRTASRRGVRALGVVAASAALALSASGNALACNISEFSAEAKCDGDNGVITVTDVDPAGIPATVTVYLKDGDK
QVGQQTVKGSKDGTTVSFPVDWQPNTTYRVHVKADRYVDEDIQPDVTTP
STPCKTEDT
PTPTPSE
SSSTPSDETES
PAPEPSTSA
PAPSQSESTA
PAAVPNDNA
PSPAAGDSNLAETGANSNTGMIAGIAAALVVVGGGAVFFGLRRRGANSNS

>gi|21224014|ref|NP_629793.1| putative integral membrane protein [Streptomyces coelicolor A3(2)]

MTDAIRPEGEPSGGHDPWAPPESGPSLDKGPAADAGQPQQPPRQSQPPASDAPPPWQPPPGVHDQATVTSMPGAGFTGPE
SAVPPPPLAPGGAGVPGGYGYPGYGQGYGWPGMRTPPQNG
LGTASMV
LGILACALFCL
YGVVSLV
LGVLAVVFGIKGRRKAESGLADNHGQAQAGFV
MGIVGIV
LGVAVIVLIA
IGITAAINEDSDGYDPYYGSARPVPVAVTAES

>gi|21224157|ref|NP_629936.1| hypothetical protein [Streptomyces coelicolor A3(2)]

M
PLIPEEPQIHEK
AQGPRATPASGRTAPTPRPVPGPR
PAAPPRPGRPGPVRPA
PPAQRATRDTAATPGPSGPAA
PAAPEAT
SAAPQIQLI
PATPDGAL
DAAEEAVDLLLDSGRTPGDVLVITTGELHPWAAHELSFGDAAYWAQHDAGDDVFYTDATVADRAASRAVVVVAL
NGGPEAAAVTALTAAHSRAGALLIVCGDPQRINSVLGAGV

>gi|21224797|ref|NP_630576.1| putative membrane protein [Streptomyces coelicolor A3(2)]

MGRPGAHPRSGHG

SPAGGYPQPMDRMDHLEHLDRQLVDELAQVARETVRDELREQTRRQRRKATLYAASGTAALYAGAAVALAVGLALALALP
GWAAALITAALLAVVAYLLRGAARP
HPSRPGP
APGTAGHDHVAGHDH
VAGGGAPAA
APGGLGVPYPPMPPV
APGGVGGATG
APGAGTP
APGGTGPAAPRQDDLDPENKRRRG

>gi|21225515|ref|NP_631294.1| hypothetical protein SC7A12.05. [Streptomyces coelicolor A3(2)]

MTSTTTSQQELFRFLEDRFA
CAQACTE
CARACALRASLVDPDGTENQELVRRRGIM
CAEVCDATCRVLSEQNQVDEAAIRAQLEW
CRQVCLESAHVFDDHPGAEE
SAQACRD
CARACGEFLATLR

>gi|21225718|ref|NP_631497.1| putative membrane protein. [Streptomyces coelicolor A3(2)]

MRSARMLLTTAAASAVFVLGAPGAYAAAGGDWDHDDSSYSKEHDNGSK
HDEPRGGMHTGGGALAAVTEGDWN
HDDTSSGKEHDNGSK
HDEPRGGMHTGGGALAAVNEGDWN
HDDASSGKEHDNGSK
HDEPRGGMHTGGGALAAVNEGDWDGGGRESKEGGSQTYQQDE
GESSRGGGE
HEKPRGGMHTGGGGLATPTTTAGGLAVLAVAATGLYAVRRKKSAHGVA

>gi|29375722|ref|NP_814876.1| DNA-directed RNA polymerase, delta subunit, putative [Enterococcus faecalis V583]

MYLLEINVFEGLNKKELSMIEVAHAILEQHADVMDFSDLVNQIQNYLGKSDSEIRDQLAQFYTDLNIDGSFISLGDNRWG
LRSWYPI
DSIDEEVTHGL
EEDEEDAPRRRKRKKVNAFIT
NPNDEDVIDYN
DDDPEDAELT
NDDEEDILYDDE
EDEDEEIKAYNSDLQEIGADS
DDDEEDI
PQIEEDLTIID
DDDVEDEDDFDDEYSEE

>gi|29376109|ref|NP_815263.1| LysM domain protein [Enterococcus faecalis V583]

MSKKDKKKNQAREPWEQSIYEPDQNGGGSRL
AKRQQQRGNSLFLTVLVILLLLIIAIPIGTFLWMMQDKKPNESASKNSQPSSSLVQSSSKEKKKESTSKSVESSEPASSQ
PAENTTPSSSDAA

AQQQQDQQ
AQQQQQQQQ
EQQQQQE
AQNQQQQQ
AQNQQPVYDTVQSGEGARQVAERNGMTLEQILALNPGIDTSVFYPGQPLRIK

>gi|29376293|ref|NP_815447.1| general stress protein, putative [Enterococcus faecalis V583]

MAKKGGFFLGAVIGGTAAAVAALLLAPKSGKELRDDLSNQTDDL
KNKAQDYTDYAVQKGTELTEIAKQKAGVL
SDQASDLAGSV
KEKTKDSLDKAQGVSGDMLDNFKKQTGDLSDQF
KKAADDA
QDHAEDLGEIA
EDAAEDIYIDVKDSAAAAKETVSAGVDEA
KETTKDVPEKA
AEAKEDV
KDAAKDVKKEFKG

>gi|29376880|ref|NP_816034.1| conserved hypothetical protein [Enterococcus faecalis V583]

MKLVKNVLL
ILIGLVI
IGIGASTLRVAGLGVDPFTAMN
IGLSSMFGMQ
LGIFQIIVNAFILFYVYLKNKKA
IGLGTILNMLLVGILIQQISTWLQPYANSLGGSTIGKLLS
LFVGVLIFTFGCALYMATDQGESPYDAIAPTICQQTNKPYASVRKAQDLLVLLIGFIAGGP
IGIGTIICA
FGAGPLISFWNKLLQPLVK

>gi|29377114|ref|NP_816268.1| hypothetical protein [Enterococcus faecalis V583]

MWQFLTKNNEDEEYEKMPDFENEQSMKA
QPTVTEEAIFDDEVVYAKEQSAKVKQVQQQV
PTELSEESFSF
VEKLEEENQLYQQS
IKKLTEQLLMQTNEVEAL
QKQVVEKDVQLKHVKETLSDKETTITSL
QKQLSEEKMQ
QRQTSEENLDTAVTLS
QKEIGEVLLEAKRQAKETISQANQQVATVHEEMEQRLATLTR
MKQVAREYQAYCEQMQTIKNESTGTYQQIEQLLAEIKR

>gi|15964011|ref|NP_384364.1| CONSERVED HYPOTHETICAL PROTEIN [Sinorhizobium meliloti]

MTKFRNFVVSCALAAASAVPPVTSVQAMPRPALDIAIKAPAGVENATHRGR
YGWRGGYSRR
GYYGGSGFYPGWSPYYRRSGIS
FYFGPSYRYR

PYYGNRYYYGP
RFYGPGYY
SYPGPAYYGYRSYRP
RYYGGGYVRSGGNSHVSWCLSRYRSYNPASNRY
LTYGGVYRVCYSPYR

>gi|15964498|ref|NP_384851.1| HYPOTHETICAL TRANSMEMBRANE PROTEIN [Sinorhizobium meliloti]

MNQHEELREKP
PLDPATGQPT
PRDPEVPVTRSRRSSWPLLVIILAGLILALIAWLPAELTRDAENE
PVAPGQTET
TPSDATP
PAAQPAPTDQAAPTNEAA
PTDQATPPAGGTTA
PAAPETQT
PAAPETPAA
PAAPATPAQPQSAPAQ

>gi|15964759|ref|NP_385112.1| HYPOTHETICAL PROTEIN [Sinorhizobium meliloti]

MFLRAAACALLLMPVAASAAEGFATANVNMRSGPSTYYPAVTVIPVGESVEIHGCLSESPWCDVSFYGGRGWVAGRYVQA
AYRSSRVYVEPEYYRPLGIPTVVFEFDSYWDRNYRGRDFY
RDRDRWRHD
RDWDRERE
RDRPEWERRQV
RERRDWEHDRE
RERREWEPERG
SDDDRWRRIEGAEREQSRRERLYDGQRTIVECDAGDFRCEE

>gi|15964892|ref|NP_385245.1| CONSERVED HYPOTHETICAL PROTEIN [Sinorhizobium meliloti]

MAPFDPKMQE
CIENCLA
CYRVCLSTAMTH
CLEAGGEHTKPPHFRLMMA
CAEICRDAAHFMLLGTPHHKHLCGE
CAEICAE
CADDCERIGDMGD
CVDACRR
CAESCRQMAA

>gi|15965655|ref|NP_386008.1| HYPOTHETICAL GLYCINE RICH TRANSMEMBRANE PROTEIN [Sinorhizobium meliloti]

MKTLDVFLTSTGLALMVGFSTSTTIPAHAQNLTGPMLVAEQCSPLTDANYEACCMAQNRADILTAGQIEQCPPLSTAAIG
NSFSTTSTDSGPRNLN
GDDGGGG
NPGGGGGGGGGGGGGGGGGGGGGGGGGGDPGGGG
GQDTGKG
NQDHGKG

NQDNGKG
SQDNGKG
SQDNGKG
SQDNGKG
SQDNGNQDPQR

>gi|15965659|ref|NP_386012.1| PUTATIVE GLYCINE-RICH CELL WALL STRUCTURAL TRANSMEMBRANE PROTEIN [Sinorhizobium meliloti]

MRRTGKTIIGSLLGLAYVVATPFASLASFDDGLRQSTALAGPHGDVLVVADAGSAPGGSQETG
SGGGQDTG
SSGGQDS
GSGGGEG
SGGGQDTG
SGGGQDTG
SGGGQDTG
SGGGQDS
GSGGGEG
SGGGQDSG
SGGGQDTG
SGGGQDTG
SGGGQDS
GSGGGEG
SGGGQDSG
SGGGQDTG
SGGGQDSG
SGGGQDS
GSGGGEG
SGGGQDTG
SGGGQDS
GSGGGEG
SGGGQDSG
SGGGQDT
GSGGGETPASTSSD

>gi|15965826|ref|NP_386179.1| PUTATIVE GLYCINE-RICH PROTEIN [Sinorhizobium meliloti]

MKRINRSIFRTFFILSTATIGAVIVTTDADAVQRQRMTARQCSPLSEDNFAACCVAINRTKILTPAQIKMCPPLTTASIG
AGTGRHDRSDSTASGPMRG
GGKGGGSKGGGGVGGGGGGGGGGGGGGGGGGG
GGGGGGD
NGGGGQD
PGKGDQD
NGKGNQD
GGKGHQD
NGKGGQD
GGKGHQD
NDKGNQD
GGKGHQD
NDKGNQD
GGKGHQD
NNKGDQDG
RGKGGKDK
GGKAKGDK

>gi|16760144|ref|NP_455761.1| TonB protein [Salmonella enterica subsp. enterica serovar Typhi]

MTLDLPRRFPWSTLLSVGIHGAVVAGLLYTSVHQVI
ELPAPAQPITVTMVSPADLEPPQA
VQPPPEPVVEPE
PEPEPIP
EPPKEAPVVI
EKPKPKPK
PKPKPKPEPVKKVE
EQPKREV
KPAAPRPASPFENSAPMRPTSGTASATSKPAVSVPTGPRALSRNQPQYPARAQALRIEGRVKVKFDVTAAGRVENVQILS
AQPANMFEREVKNAMRKWRYEAGKPGSGLVVNIIFRLNGAAQIE

>gi|16760632|ref|NP_456249.1| conserved hypothetical protein [Salmonella enterica subsp. enterica serovar Typhi]

MQQEH
RECIEQC
YECAAAC
DICASSCLREDNVEMMKHCIELD
MQCAAICRLAAQFMALESEYSQKLC
RLCADIC
KACAEECARHDHDHC
QNCARAC
SQCADACLKMAA

>gi|16762879|ref|NP_458496.1| conserved hypothetical protein [Salmonella enterica subsp. enterica serovar Typhi]

MAKSALFSVRKNE
PCPQCGAELVIRSGKHGPFLGCS
RYPECDYVRPLKSQADGHIVKILEGQ
LCPECGAVLVLRQGRFGMFIGCS
QYPQCEHTVVIDKPDETAI
ACPACQQGHLVQRRSRYGKIFHSCD
RYPECQFVINFTPVAG
ECPECHYSLLIEKKTAQGVKRFCASKQCGKPVPVE

>gi|15618299|ref|NP_224584.1| Histone-like Protein 2 [Chlamydophila pneumoniae CWL029]

MIGAQKKQSG
KKTASRAVRKPA
KKVAAKRTVKKATV
RKTAVKKPAV
RKTAAKKTVA
KKTTAKRTV
RKTVAKKPAV
KKVAAKRVV
KKTVAKKTTAKRAV
RKTVAKKPVA

RKTTVAKGSPKKAAACALACHKNHKHTSSCKRVC
SSTATRKHGSKSRVRTAHGWRHQLIKMMSR

>gi|16799813|ref|NP_470081.1| lin0738 [Listeria innocua]

MNPRKPEF
ILTLIAG
ITGIIAGITGIAGGG
ILAALSGSEELSGAATMSTTDTEALAAAGGFAVILS
AIALVIGIALIVFAVLIKRNAKVFG
ILTLVAGVGGFFLVS
LLWIVPG
VLAVIAGIMCLARKVQTF

>gi|16800079|ref|NP_470347.1| similar to tetrahydrodipicolinate succinylase [Listeria innocua]

MEQMDAHQIISFIQNSKKATPVKVYLKGDLEKIDFPSDVKTFITGNAGTI
FGEWAVVEPLLEANKANIEDYVIENDRRNSAIPLLDMKNINAR
IEPGAVIRDQVT
IGDNAVIMMGASINIGSV
IGDGTMIDMNVVLGGRATVGKNCH
IGAGSVLAGVVEPPSAQPVIVEDNVV
VGANVVVLEGVR
IGEGAVVAAGAIVTKD
VAPGTVVAGIPARELKKLDAKTASKTEIMQELRQL

>gi|16800653|ref|NP_470921.1| similar to B. subtilis late competence protein ComC (type IV prepilin peptidase) [Listeria innocua]

MVYFLLIIYSSIFISFIQVAGECFPIKKSFLFRFSECNHCQKTLSFSQMIPIYSFLFLKGKSKCCQTPIP
LNYFLLELTT
PIYILLLHYQFSFSYRFVIGC
IIYYFLAFFFVTDILYLHVPNSV
MILFFLVLLTCYISFQQPAVS
LIYSSLISNCFY
LLFFILFRKGIGLGDIKILIILSSFLGFEMGY
IVFFLSIVLGAIILLSALVLKKVEKNKQVPFVPYI
FISFIIVSILMK

>gi|16801121|ref|NP_471389.1| similar to unknown proteins [Listeria innocua]

MAKTRKSKREYEEPADFDMNSEGQDEFNNEPPMLSRSDSRKGKKKTIFRYPLLNLLIVFFLLIPIVIVIVFVTVQNMGST
NEVKTETESTVNVSDNTQT
KAEKEKAK
KAAEEKAAAEKAAEEK
KAAEEKAAADK
KSQEDEAA
KAAAAKKEQ
EAAEEKAAAEKAAAD
KAAKEKAAA
EAAQKKAN

ESTQKKAGGSHTVQAGDTLYSIARSSYGQAGAAAGVEKIKQANGLGSNNVPVGTVLTIPQ

>gi|16801545|ref|NP_471813.1| similar to proteins involved in resistance to cholate and to NA(+) and in pH homeostasis [Listeria innocua]

MNVIIEIIISIMILIGGLLSILAAIGVIRLPDVYTRTHAAGISNTFGVSLLLFATVGYFFHSGEGFNARVLLAVLFIFLT
TPVASHLINRAAYDTGVPLAIRIRDQLRSVKKDDIKKKKSLIIRQEQIEKARQE
REELEER
MEWERRE
EKIDERED
QEEQERE
REEQTIEEQ
SDDSEHEIIE
QDESETD
SDEDESEK

>gi|15602174|ref|NP_245246.1| unknown [Pasteurella multocida]

MTE
VHHRVDGLEINKQARDNPQL
VHHRVDGLEKLGGLENLRFN
VHHRVDGLESLRKRCFARLV
VHHRVDGLEKPIQLTMEEAN
VHHRVDGLEKKPILCTSIKR
VHHRVDGLEIGDSYS

>gi|15602307|ref|NP_245379.1| unknown [Pasteurella multocida]

MKKSVLAALVLGASLSVTGCFDKESKAGQ
EVEKAKASVV
ETKEAVVNAAN
EVKNSAVEAAK
EAKETVATKVE
EVKETTAAKVE
EVKAATAAKVE
EMKETTAAKVE
EMKEASAAKME
EVKAATATKME
EMKAAVADAKN
DVAEKAA
EVKDAAAAKAE
EAKAAVAEKSA
EMKDAAAA
KVEDAKVASSEAKEMVTEKAT
EMKDAMKA
EVNKAADAVSEKTE
EVKEVIVPKAQ

>gi|15603706|ref|NP_246780.1| unknown [Pasteurella multocida]

M
VNCRIGSLEILHKQLGHQSI
VNCRIGSLEKDSRAGRKSYS

VNCRIGSLESLCIAFFNSTY
VNCRIGSLEKNQNLHLFHLD
VNCRIGSLERIY

>gi|15899654|ref|NP_344259.1| Conserved hypothetical protein [Sulfolobus solfataricus]

MS
LEEEIKKVLLNNPSILVDVLTAKPEIVYQVLARLTPWQSLATKQ
DVERLDK
EIEEVKKTMATKQ
DVERLDK
EIEEVKKTMATKQ
DVERLDK
EIEEVKKTMATKEDLKRLETIITGLGARWGILNEDSFRQGISEILSNSGWKISREVLYDRDGYVYDEPSDVEYDIVVKDS
SVILVEITSSIRRG
DLPTIKKKKEFYEKVKNIKVNLVYVITPFIHDRYPERIKAMAKDMGIEIIYPSS

>gi|20515959|gb|AAM24210.1| Molecular chaperone GrpE (heat shock protein) [Thermoanaerobacter tengcongensis]

M
EELEKDKIERN
EEMSEEVKGEGPPSELEQS
EEVVEEKI
ETEVEQKKEPSL
EEIVEEL
RKKLEEKEKEAKEYLDIAQRIKAEFDNYRKRTEKEKAEMISYGQ
EQVIIELLPVIDNFERALANEGDYNSLREGLELIYRQFKKVLDKFEVREIEAEGQMFDPYKHHALA
QEEVEGKQPNEIIEVFQKGYYLKDKVIRPSLVKVAK

>gi|20516101|gb|AAM24340.1| hypothetical protein [Thermoanaerobacter tengcongensis]

MGNNNFESRAKDSFKGLIGFFKSIG
LIIFTLG
VWSWIIG
AFIWVLG
AFGWFLL
AYVWRKNLFIPQAVYSTAF
AFVLSLGWAIIICAI
MFVWTKYHYYRYYRKNKRKLKMPVFEAPILAWKQLSLEALDIAKIMEDKQIPQETNKSLFGNLTSVKEKDYLAPFTPMIM
KKNFCSPKGNIIVSEGEEITPEVIKRIAEEGLYWEFIHEISEYIPAEEESK

>gi|20516284|gb|AAM24506.1| hypothetical protein [Thermoanaerobacter tengcongensis]

MEYLNERVSYLRGLADGLGIDDSTKEGKLILAILDALEDFADAIEELEASQSELNDYVGE
LDEDLSEI
EDRLYGD
EDEDYDYEDEEDEDYEYD
YDEDEEED
EDEDYEEYVEVECPNCHMLMSVEEELL

EDEDAELVCPHCNATIKVKDLKI
VDDDDED

>gi|20516667|gb|AAM24859.1| hypothetical protein [Thermoanaerobacter tengcongensis]

MILTSMAGCKTATKK
PAPARYT
PTPTRPA
PAPSKAT
PTPGYTTPAP
TAPTAPT
PAPVRKPTTESMRASRVATSVARI
PEVNKATVVISGTTALVGVDMKAKVQGTHEKDVKKKIEKAVKDTDKSITRVYVTADPDLYKRIDNIARGISEGRPVSEFA
KQISEIIKRITPGM

>gi|20517446|gb|AAM25572.1| hypothetical protein [Thermoanaerobacter tengcongensis]

M
FQSLLGRLKTSSNIKTANKSPV
FQSLLGRLKTNMAIGKFVAGEA
FQSLLGRLKTRFDILWEVRRRK
FQSLLGRLKTAKEGECMNRTKM
FQSLLGRLKTIVFG

>gi|20517662|gb|AAM25769.1| hypothetical protein [Thermoanaerobacter tengcongensis]

M
FLAYLRGIETCKHQLNQLLSREK
FLAYLRGIETNSALNLLRIDAS
FLAYLRGIETFVCAVLRAGYPR
FLAYLRGIETGRLCHAAPTSFM
FLAYLRGIETIV

>gi|22297715|ref|NP_680962.1| ORF_ID:tlr0171~hypothetical protein [Thermosynechococcus elongatus BP-1]

MKHSFTIMAKFNAEELLQAYAAGER
NFRGSDLMGLSLFDT
NLSDADFSESTLENAYLPYCQLNRIQATAA
NLRHSELGDA
QLYQANLAAAEENAYLPYCQLNRIQATAA
NLRHSELGDA
QLYQANLAAADLEGASLVRA
NLRTANLERA
NLRDANLQGADLRYA
NLRQANLQGA
NLQQANLERADLTEAQVQRCNFFRATGVDLSAAVCDRTTIYPDGYFYP

>gi|22297733|ref|NP_680980.1| ORF_ID:tlr0189~hypothetical protein
[Thermosynechococcus elongatus BP-1]

MYSTIQKLGSLPQFCIGTKWCQLDKLQLSTPVECPVITHPHYPNSAVVHRWVLSSCLMGLGCVTLQLAPATAQIAVY
EGKRVTGDVTLTLP
DGSTYKGEL
LNGRFDGQGILTMA
NGNRYEGEF
RNGRYHGQGVLTYA
DGGRYEGGFADGIFSGKGILQLA
NGQRYEGTF
LNGQYHGEGVLTFP
DGTRYEGQF
LAGKYHGTGTLSFG
NGTSYTGQF
RNGLFEGEGVLTLPDGSRIIATWHNGRREPR

>gi|22298048|ref|NP_681295.1| ORF_ID:tlr0505~hypothetical protein
[Thermosynechococcus elongatus BP-1]

MVMALGNIAIAVKTMQRSTEPNFLVKLRAGVFSLALFSLVVPAVANAENPAHVRQLLETRQCPNCDLRG
ANLMDANLFK
ANLAGANLEGAYLAG
ANLFGANLARANLSRASLYLAILVN
ANLTGAQLVGARLTLSNLVN
ANLSRANLEN
AVMKGANMIDVNLTA
ANLKGADLDR
ANLSGANLSQTIFDRPANP

>gi|22298427|ref|NP_681674.1| ORF_ID:tlr0884~hypothetical protein
[Thermosynechococcus elongatus BP-1]

MKRQVFVAIAMMSGAIASPTAFAMPASMQQLLQTNACPGCDLRN
VDLRGYNLAG
ANLRGANLQG
ANLQDTNLMV
ANLVGANLSHANLTAAYLERANLEE
ANLSGANLSRAILRHSHARRANFSN
SNLTGADVRYGDFRR
AHLNGATFEQ
ANMQWARLDNATVGQTNFRDAYRPRMPYRLNLP

>gi|22298510|ref|NP_681757.1| ORF_ID:tlr0967~hypothetical protein
[Thermosynechococcus elongatus BP-1]

MCAPKRGRSQQFITTVIVCWRMASFSFRIEVRLSDAERRPTMRQQLFSTLTVLVALVGLSRAVVAQPMAIATLTPPELQE
LQRLVPRSA
EDYYNLGLMHQGRGDLAAAIDAFSQAIKLNPSA
DYYFARG
LAYYDQGDMQRAIADFTTALGDDPQFI
AAYYNRGMAYLALQNYDAAIADFDAALARDPQFV
AAYYSRG
MAHFDRGNVELAQRDYERARALNPAMT

AQYYDRAPRPLTGGP

>gi|22125451|ref|NP_668874.1| hypothetical [Yersinia pestis KIM]

MLYVCFPRRTCLPFIYKIKAVASASLAVF
ASVPAVPLSLAVPLSLAV
PWFPAVPLFPAVPLFPAVPLFPAVPLFLAVPLSLAVPLSLAVPLFPAV
PWFPAVPLSLAVPLFLAV
PWFPAVPWFPAV
PWFPAVTLDLVATLDLVAMSDPVAMSDPVALLDPVAMSLVAPLLATRVTVAVRQ

>gi|22125498|ref|NP_668921.1| putative lipoprotein [Yersinia pestis KIM]

MRVASQFQNRLVGTIILVALGVIVLPGLLDGKKKHYEDEFAA
IPLVPKPGDSQEIDVVPPLN
QPLPITPPEGAAQAVGAELNDDSVSQAANDAGNLAV
EPTIVAPPPIQS
KPVEVKPTVT
KPVDVKPTET
KPIETKPAEK
KPLEVKPKPEIK
PPVEEKAPVGQAYVVQLGALKNAAKVNEIVAMLRLSGHRAFTVPATPVQGEITRLYVGPDASKQKLQSALPELNALSGLN
GQVKPYTTSR

>gi|22127542|ref|NP_670965.1| hypothetical [Yersinia pestis KIM]

MLCPVMLVLMPVMLW
LLPVMLL
LVPVKEE
LHPVIEL
LTPVSEL
LAPVKEV
PTPVNDV
LTPVKEVPIKVKPT
LTPVTLV
LMPVTDK
LVPVVET
LMPVAET
LMPVTEKLKSL

>gi|22127845|ref|NP_671268.1| FKBP-type peptidyl-prolyl cis-trans isomerase (rotamase) [Yersinia pestis KIM]

MKVTKDLVVSLAYQVRTEDGVLVDESPVSAPLD
YLHGHGSLIAGLENAL
EGHEAGDSFDVRVNADEGYGSYDENLVQRVPKDVFMGVDELEVGMRFLADTDQGPVPVEITAVEDEHVVVDGNHMLAGQD
LNFHVEVVAVREATEEELQ
HGHVHGEHD
HHHEHGDGCCGG
HGHDDHE
HEHEHGKGGCGKSGGCGCH

>gi|5104270|dbj|BAA79585.1| 212aa long hypothetical protein [Aeropyrum pernix]

M
DGEPPLHQR
GHPGPGSLRHRHR
HSSPPPQQPAGGGEGQPGGHHNP
HRGSPPSALNPH
HHPPPPHPRQGNAAGLQR
PGPKPRH
RRPPKPAPNYRGLHSRRHQNPRAGRDPTGVLDPRPRNNHS
LHTPGDHG
HTPEAPGHREPGCRREVHDNPNTRPLTLRSTRQTPPNNTRHTPGEDASTPPGTGCQAKGSPSGSRESRGLQHSAKSVDRP
G

>gi|5104510|dbj|BAA79825.1| 128aa long hypothetical protein [Aeropyrum pernix]

MAA
SRQDSPCTSTRWARPRGSCTCSSTT
YPYPAPCGTSSG
ASSPQPPCKRCPPSSPTGPGGRGR
PRRSRPGRWR
RRRGKLCPGSPRSSGPGGPLSGPGVGS
ARRPRPCP
GRSCRPC
SRRGTPATGCSC

>gi|5104531|dbj|BAA79846.1| 248aa long hypothetical protein [Aeropyrum pernix]

MLQQHSL
PHVLAPR
PPPLPGRPEELEGRLYTPLRAPLREPLDEGRVGHCGRLQAQQR
EHPLGSRLRATLNHSHVPRRV
PRLLPVRNVAVPAEAPGVHQFVDDHVLEPLAQPPLNGQHSIVPAGP
VFPVPQNPRFIHLLLV
PNPDPRGHRLAEVKLHRPPSPLNLAQDALGLLTIRLQEIPGLEPPAHLRGVNAGYVVPVGR
PHPLPQRPSVPGRQPDVHKGAARTP
PSPLPERLPGTEPLLPSQAKPV

>gi|5104682|dbj|BAA79997.1| 143aa long hypothetical protein [Aeropyrum pernix]

MLRRPLLSCGMGRVVG
GLCSRPP
RQPPPPP
GSPSSPGSPWRTRPA
RVSSPPRRPCPRGLRALRGGRT
GSRCPPPRLGPSPRAGGRGA
GASSPPPRGRLSSGRGLWGGACPHTPSPCTPGARR
GGGEPWPWPRAPSCWSPARSS

>gi|5104687|dbj|BAA80002.1| 229aa long hypothetical protein [Aeropyrum pernix]

MAQGHRGE
RLGPGGGGEDEQVQGQLHPPKAGPRLVGRRRHQVGRGPRGRRLGARRRQLRAQRGRQRSIPPKPVAGLRQGELREACEE

GGEVGGQVVRKQAQLHNSKGYQADGGGQLQDRP
RGGVVQAAGELQVVPQEVGRRAQGGPPPEVHRGPDPTHSPLRPPHR
RGGVGGHG
QGGLRLHSILAGARRVEDKP
RGGGGGGDC
AGGAGGR
QGGPQPHRRGRHPSRHSCG

>gi|5104807|dbj|BAA80122.1| 220aa long hypothetical protein [Aeropyrum pernix]

MTRRILLAN
APHAPPP
LPHIPPPQPYLLSP
PQHLPAP
LPHGPQVAPGVEAVELPRQGGVG
ALHSPPEEVEVHRPVLNAGAELRQPVVY
QHHTPSPDYVVRGGGVPVYDARPVEEAEVVASLLHKPP
GPEAPPVYDVLEEGALHILHHNQPEAGEDGEEGWGADPRGPGCRYNPG
LPPSPQPDKPLGEPGVHVGLPNPSLIHGL
EPHKLPQPDLGLGPPLQ

>gi|5104893|dbj|BAA80207.1| 235aa long hypothetical protein [Aeropyrum pernix]

MAPAIPPYRGRRIPCSPPGGVGCCGVDAED
RPLPRRPHDTRRKSIHGRQELWQEANQPAYGRRVGAQ
RTRPGPR
HPRPGAPR
NTRPGRPPSL
LPRPGAPPGSRAFEEGLQAPGARGGDAKVHGRLPPPNNTPHAALPPPPGHYACEAPPALHRV
HRRPHNPP
RPPPTTPGSAGQAIQELP
AHPPGTPLGGSNRQGRGPRCIV
YNRLGSPSRGRRPVQARGRGGDE
GPRRAPPLLD

>gi|5104918|dbj|BAA80232.1| 138aa long hypothetical protein [Aeropyrum pernix]

MDPELERELRRWL
SIPFRDMLDSIMLFTSFLWVHFLFPLGICPGRLGRGVGGGPGL
SIPFRDMLAGISGPNNYLVLPPPL
SIPFRDMLTDLGFGYIAYAIL
SIPFRDMPRQLWGGSNGGGKL
SIPFRDMRRCPAHSRV

>gi|5105230|dbj|BAA80544.1| 200aa long hypothetical protein [Aeropyrum pernix]

MELTSIPQDIGTARMLGIAGSA
LVLLGILIPF
LSLIGFVLVLVSLYMLSKQMGEESIFRNAVISVV
LSIAGFIVFAFLTGASFFAMAGALGSGELTALTAAGIG
SLLAGFVILYIAMVASAVFFRRALEALAGRTGVDMFRHAGL
AYLVGSITL
IVLVGGL

ILLIAWV
LLLVGFIQLRPEQGMETQAAVQEGPGESIGQT

>gi|5105252|dbj|BAA80565.1| 138aa long hypothetical protein [Aeropyrum pernix]

MILSH
PKPLVEPG
PNGLHIPSEHGCWEAQLMRGGRLN
PHPRHKPGNGCRVH
AVPQHEPLYLSPDARSYVHRQ
PHPGSPG
QYPSHYPGSLLQAPGPEDLAQLGVGKHGRVCHHNA
AHARRRPPAHA
SHPLQPPCPSLLV

>gi|5105838|dbj|BAA81150.1| 228aa long hypothetical protein [Aeropyrum pernix]

MVYRLVVKYVLEKGPAFTILKVELNE
GESIWVESGSYMLNKGDIEVKTTMGGGIVRGLLRAVAG
GESLFFNIITARS
KSEVWIAPPVTGDIAAVELG
GEEVYIQDSSYLAHIGDVEVGVGWRGFKGLIAE
GELVWVKAKGRGTVFINSFGALEHIRVGP
GEKITIDNGHFVAIKGGKWNIRKFGGWKTFFLG
GEGVVVDVEG
PADLWVQTRNLPAFAGILGKFIPSKS

>gi|5105885|dbj|BAA81197.1| 119aa long hypothetical protein [Aeropyrum pernix]

MDRKARASLTVA
SPLATAS
SVSATFSITSISSGFVMSKTY
ATSSSAS
SASSTASSA
PASSVTSSRLPV
SSSATASSSLSTSSDNRSRTLKVDRL
QAYSTLSTPVLRSRGTTTSTSPLQP

>gi|5105931|dbj|BAA81243.1| 190aa long hypothetical protein [Aeropyrum pernix]

MDTGLGLRLPRLQEQPLRQHSSLRDGD
GRRQEGP
GRRRGGGCGDGVQGYPLPRRLRLHTARRA
PRPPGGGGKGDREEDRCS
GEDRGGVQVALHTPDPE
GPRRGKQVLRGPRERGCEAEGDS
GRQAGHAGDSEGGTGLR
DSEAGGGQAPR
GVAWGGGEGLEGGRWVY
EAAQGGGAAGREACHREEGGPGLL

>gi|11497623|ref|NP_068843.1| hypothetical protein [Archaeoglobus fulgidus DSM 4304]

MKAASYGVP
FQSHFGLISTIRGNDGLPRVVVF
FQSHFGLISTVKIVEVLEKDEALSIPFWSDFNCHLQMDQRIRLAF
FQSHFGLISTGTLQTSGTNANFLSIP
FWSDFNLELLKRIKKTKML
FQSHFGLISTRNMVWGRSELQRAI
FQSHFGLISTQSSQLTMPVKMQTFNPILV

>gi|11498652|ref|NP_069880.1| hypothetical protein [Archaeoglobus fulgidus DSM 4304]

MIEMLIGIAFAGFILAIPYVYERVGGKLSLGKLSLGMPKLRSKKEV
EEKIKEV
DRELEEVVNAGVSINPKPAEEKVVPLDDANVEPDDNLLEEMETASEIKVEAPEPD
EEKLPDIPDLPELNSDLDMDFEDLGQEIPLDAD
EQEEEEE
EEEVEEVEFDEEDDLISSLAKEVETKE
EEEIDLLRDLKGQKFSAEEL
EAELQEMIQRLKVLSGGSS

>gi|11498824|ref|NP_070053.1| hypothetical protein [Archaeoglobus fulgidus DSM 4304]

MIPLVALLVLLTLQASPGDIVSVNVTDTATLTVGDSCMYFVDNLQPSINATPGEYEVKIGINCTPGLKEVYADGSVLAQI
DVNETTIDYQAYAASLEKENLALQKE
VESLKEKLKISQEQ
IETLKSQ
LEDLQNKAKMLGI
QNELQKQQ
IEELQKKLERA
KTELQKKKSD
LDELEEK
IRELNRQSSIYRLATFFMVSLFVGSFVALVFVARKE

>gi|11499465|ref|NP_070706.1| hypothetical protein [Archaeoglobus fulgidus DSM 4304]

MFTAFNLH
FQEPPFLTRQPQLVHQLWYQ
FFQSPFSGASLSYFSEISIALINQI
FFQSPFSGASLSYTYYTDEFTKELISFNLH
FQEPPFLTSSSSSLITTPLI
IFQSPFSGASLSYLSVIEDMESGER
DFQSPFSGASLSYSRHFNEDAVRED
DFQSPFSGASLSYAISSKKIGGKEY
IFQSPFSGASLSYRIYSHRCHQNSLLSISIFRSLPFLLVMKAVKINGVD
FFQSPFSGASLSYFPIQKTK

>gi|11499466|ref|NP_070707.1| hypothetical protein [Archaeoglobus fulgidus DSM 4304]

MVLFYEVLTIS
IFRRLPFLQRMLNKSQAKSKTCFQSPFSGASLSYTSVAVHLCPHPCSLSIS
IFRSLPFLHNRKVQAAGPLRGLSIS
IFRSLPFLPDYKRNPSPVQRHLSIS
IFRSLPFLRMGHMTQMAGQATFQSPFSGASLSYPSPLPLPELLFCIFQSPFSGASLSYAEIIPHYYHPYFALSIS
IFRSLPFLQAA

>gi|11499515|ref|NP_070756.1| tungsten formylmethanofuran dehydrogenase, subunit C (fwdC) [Archaeoglobus fulgidus DSM 4304]

MMEMLIIKPKTEFVVTVEAELTPDLAEMDVDEVKKFKVYY
GKDLVELGELFEVTKDGDD
KKLLLDGDFSRVKWIGCRMTD
GEIVVKGSVGANCGAYMAG
GKIVIEGNADDWLGAEMKD
GEIIVKGNAANLVGCAYYGDATGMSG
GKILIEGNAGNYIGEKMAG
GEIEIKGNAGDFIGTEMKG
GTIIIHGSCGFVGGDMKA
GEIKIGGEFELMPTFVRQEDAWV
GDVNVKGEGVVKRL

>gi|23464668|gnl|REF_nestle|BL0040 narrowly conserved hypothetical protein [Bifidobacterium longum NCC2705]

MPFPIERGLTRMTDNQ
QQPYGSDS
VPQPAQPAQPA
QPSYGQDAAQQ
PQGYGAPQ
PAAYAQPA
AGSYGQQ
QPQYGQP
AGAYAQPQYG
QPQYGQPAQPAAGYGQ
QPGYGQAYQQPNAYA
QPQYTAAPYAAAPVGQSSKSKLAAGLLGIFLGGLGVHNFYLGNTGKAVAQLLLTLVGWILLGLGPIAAAIWGFVEGIIIL
CSKPGSPWHKDANGYELSD

>gi|23465013|gnl|REF_nestle|BL0415 single-strand binding protein [Bifidobacterium longum NCC2705]

MAGETVITIVGNLTADPELRTIGSGVTVASFTIASTPRTWNRQT
NQFEDGQALFMRCSAWRDLADHCAQSLKKGMRVIATGRLQQRSYQAQDGSQRTVVEMQLDEIGPSLRYATAQVTRQSSGQ
GGFQGRS
GGFQGGQQPNQG
YGNQGGQ
GGYAGGQGGYN
GGYQGGGAGYSQQSAAPANPSAPAVDPWSNAGASDAGSFSTFGASSDFGGDDSDPEF

>gi|23465210|gnl|REF_nestle|BL0627 hypothetical protein [Bifidobacterium longum NCC2705]

MAYCGHCGNRVDRPARFCPACGASVD
SAPVASV
PVPAASS
AVPTASIPAPAV
PAPSVPPSIP
AAPVAPSVY
AAPAAPQ
PAPTAPLPALPI
PAPAAYQATDATAVLAPPRKSSKKLIAVLCALLATILVAALAFIVPNFTAVKAMLG
FQPTNPLVVTFDAINSLSTLRSTKYELRTHSSTDDSLDARMSGMYSLGQNHQ

>gi|15679902|ref|NP_275211.1| TPR-repeat-containing protein [Methanothermobacter thermautotrophicus]

RIGWLGEGSVLIYLKK
YKKALKC
FEKAIELNPKNYRAWGTKGITLHNLKI
YEEALKC
YDKVLQLNPQDDKAWNNKGLVFNELGR
YDESLEC
YEKALQINPKLAEAWNNKGVVLSELGR
YEEALEC
YEKALEIDPEDDKTWNNKGLVLEELGK
YKDALEC
FQKALEINPEFADAWKWKGIILEDLKK
PEESLKC
YKKALKLNPPKQNTMVHARKNTTKTRQTQRSTKML

>gi|15678835|ref|NP_275952.1| unknown [Methanothermobacter thermautotrophicus]

MKGGESVSDTK
IPGLISL
ILGILVILFPVFSVFTLSV
LTGVAVLFVAIWLFLLGSGTWKVNRGAGVLYILLG
VLGVIVAATIAGNVALFSFLAAFWIY
ITGIILI
IAGITALFSREHRAGRAAAISVV
VLGFIYIILGIFVSNPVFLAF
LIGLSLV
IDGIGLLMY

>gi|30062775|ref|NP_836946.1| membrane protein, energy transducer [Shigella flexneri 2a str. 2457T]

MTSITLDLPRRFPWPTLLSVCIHGAVVAGLLYTSVHQVIEL
PAPAQPISVTMVA
PADLEPPQAV
QPPPEPVVEPEPEPE
PIPEPPKEAPVVIEKPKPKPKPK
PKPVKKVQEQPKRDVKPVE
SRPASPFEN
TAPARPTSSTAT
AATSKPVTSVASGPRALSRNQPQY

PARAQALRIEGQVKVKFDVTPDGRVDNVQILSAKPANMFEREVKNAMRRWRYEPGKPGSGIVVNILFKINGTTEIQ

>gi|28572598|ref|NP_789378.1| putative integral membrane protein [Tropheryma whipplei TW08/27]

MIGVARKLFQNPSGGGISRISMPVCLTGVSALLLALVFITPASHAYAGRAAAHVFATRNYNCLGTFLFAYRQRGSSAPKP
VCREFYQRK
PSSKPPVPHKPVPAK
PSSKPPAQQRPPVPHNNPRPLPAK
PTSQPTSKPQRPSVPVVHPVPPAA
PTTSPTTS
PTTSPTTS
PTTSPTTVTAPSEPA
PTTDTTKQLSTSNGRARLPWWVMLVTGLSGLFLGSGLVALPFLL

>gi|28572782|ref|NP_789562.1| hypothetical protein [Tropheryma whipplei TW08/27]

MLPPSAPVTIHH
SAQPSGKHSSSTQAPSR
ATKPSAP
SAQPPASE
ATQPAAPAASQVTQ
SAQPVVPSATAKP
AAKPSAPTQTTQPAQAILHWVASIGIKRLTHLHLPQAQQFLQPYCSLDTANSKPKWENMTFSVP
AAVPSATAKPAA

>gi|15836614|ref|NP_297302.1| TonB protein [Xylella fastidiosa 9a5c]

MSEELIVHRYEQDDVSDALNWPRILGIAFVIALHLAALMLLLIPAASL
KAPPEKQRLAAVTIV
DSPPPPP
PPPPPPQEKK
PAPPIKTLA
PPRPAPVPQPPQPPVVDV
PQPRPSDVVV
PPAPPAPPTPPSDISASVDISSKKMN
PPKYPPTALRAGIQGEVVLIIDVDADGNVTNVSVEKSSRNRDLDRAAIDAASKWKFNPSVINGKKSAGRVRVPVNFSL

>gi|15836623|ref|NP_297311.1| hypothetical protein [Xylella fastidiosa 9a5c]

MQQRRRYTGIYPKHLNG
RHNIISN
QQSAISN
QQSAISN
QQSAISN
QQSAISNMYPGFFITCICADKMMTT

>gi|15836785|ref|NP_297473.1| hypothetical protein [Xylella fastidiosa 9a5c]

MNQMSGAGVIGNATRADFFLR
ATGFFLATA

FVDFLAIT
FEGFLAIVF
FIAFFVTFFNTGFGVF
FTAFFAVTV
FTTFLEF
FTAAFFTVAAFFT
ATAFFTTAVFFTIAFFNAGFLAL
FIAFSAAIFFIAG
FLVFFTA
FTAFFMVIFLAATFFVAM

>gi|15837520|ref|NP_298208.1| acyl-(ACP)-UDP-N-acetylglucosamine [Xylella fastidiosa 9a5c]

MKNTFNYVFSNGQPATAALDWINPDGSKGGIVAANANINPSVVIDRTSVV
DANVTIGAGTVI
GGKTKIGRNSVIGTKVTITCNADI
GNNVCIGKESKINNKVRIEDHVVI
GESVSIGYNTHL
GQSVNIGYNVHL
GQSVSIGHKVHL
GESVSVDDNVHI
GESVSIGDYVHL
GESVSIAKLACIARHASISHRACI
GESVRVVEFARIAPGAIVSQ

>gi|15837693|ref|NP_298381.1| hypothetical protein [Xylella fastidiosa 9a5c]

MPRLEKQSRNQLSRTTSTCSQCLCFSQKIANATNPVGNQAEKSVNTFLRFICRRFDTMNRCSNYASYRCPD
HVPSRSHLRISASPH
HAGTQAHR
HTGTQAHR
HTGTQAHR
HTGTQAHGSCSN

>gi|15839117|ref|NP_299805.1| phage-related protein [Xylella fastidiosa 9a5c]

MTSVAFDTLKFANRLKTAGVPAAHAEAEAEALAEVLEINLQGLAESESKNGKALARLEADM
KEGFAQV
DQRFAQVEKTM
DQRFAQVEKTM
DQRFAQV
DTRFAEIKGEMLLLKWMLGVLVAGVAALIIKAFF

>gi|56478039|ref|YP_159628.1| hypothetical protein ebA4595 [Azoarcus sp. EbN1]

MGVALIATFAAGAAFADKPEWAGGGKPDKQQN
ERRGGDDAR
ERRSDDR
PRRADEH
ERRGDNDYR
ERRGDDRVS

RHRDSDRPRVDIRFGDRERVVVREYYSERVRTGSCPPGLAKKGNGCMPPGQAKKWSVGRPLPRDVVFYDLPPAIVVELGV
PPRGHRYVRVAADILLIAVGTGMVVDAIEDLGRM

>gi|56479412|ref|YP_161001.1| hypothetical protein ebB244 [Azoarcus sp. EbN1]

MSQQTTLSPEMLA
CIELCRR
CQMVCLGMATGHCLEKGGRHVEPEHLRTMIV
CAEICQAAANVMATNSTLHRQVCAV
CADICDA
CISSCRDLDEMEE
CIFACER
CKNSCEAMTSV

>gi|56479537|ref|YP_161126.1| hypothetical protein ebA7231 [Azoarcus sp. EbN1]

MPSARRQSRTSSQPDGTTRPVARSSTSTSSATAPNRRSATRAPESSSARSSPCAASPCAARPNTKCAWEKTDEDRHRDR
PRPLPHD
PAAPHHR
PRPLPHD
PAAPHHR
PRPLPHD
PAAPHHRRL
PLERGHRRGDRRQPARGGRHTAHVAAALHRRLIDA
PVRRPRRRDPADRHARGR
VRLRPARRAGRPSAAKARPGARLRRRNRRATRRRRPPLRR
PRGRPHRA
APERPHHH
PRGNPLCATPC

>gi|52784598|ref|YP_090427.1| hypothetical protein BLi00799 [Bacillus licheniformis ATCC 14580]

MIHETAK
IGKNVVLGEHAVIEENVV
IGDNVTIGHHAIIKKDTHIGSGVK
IGDLAVLGKAASSNKKMARQPKQAGAPLRIEDDAI
VGASAVIYRDVLLEQGVF
VGDMASIRENVAIGRESI
IGRNAMVENNTR
IGSKATIQTGCYITADMTIEDEVFIGPCCSTSNDKYMGKGNYPYQGPTIKRGAK
IGNNATLLPAVV
VGEGAVIGAGAVITKDVPAGKTA
VGNPGRLMK

>gi|52784871|ref|YP_090700.1| YnzH [Bacillus licheniformis ATCC 14580]

MYFYHDHCK
KDYDYDY
HKKDHDC
KKYYYYYYYYYPCF
KKYDYDYDY
KKFDHFDHCK

KKKHYDYW

>gi|52784941|ref|YP_090770.1| hypothetical protein BLi01169 [Bacillus licheniformis ATCC 14580]

MSGVHDHDIKKAVDQLKSEGRDHYLDREPEEYAGSGKERRPHNIWDLWW
GIKPDRSKKHHKSKKHDDYEKPDK
SKKPDKSKKPDD
CKKPDKSKKPDD
CKKPDKSKKPDD
CKKPDKSKKHHD
CKKRDKCG
RKRPDHHR
CKKTHHHKRGGYETITKWSDGNLTEVIYRKK

>gi|52786218|ref|YP_092047.1| hypothetical protein BLi02479 [Bacillus licheniformis ATCC 14580]

MKSKLLLTAAASVAGLSLASAALMTQTDGVK
VAERNVTSAPMQ
IAERNVTSAPVQ
VAERNVTSAPVQ
IAERNVTSAPVQ
VAERNVTSVKTSPIH
IGERNVT

>gi|52786698|ref|YP_092527.1| hypothetical protein BLi02966 [Bacillus licheniformis ATCC 14580]

MTKQ
QQNQQQKQKQQKQKHEQLLGANITDTELAQEF
GNIQQQL
AQLEQTVNQMGSQQQGQSGQQS
SQLSQGINQLK
QQIQQFSSQSQ
QQMQQTD
QQIQQSIQQAV
QQLNQALQGIQTNQAM
NQMSQAL
QQAQQSFNQQGQQGQQGQQN
QQNQQQN
QNFQQPLQ

>gi|52786852|ref|YP_092681.1| hypothetical protein BLi03127 [Bacillus licheniformis ATCC 14580]

MNTELKYMLQSVLKEELSPINQRLD
GIDKRLD
KIDARFV
EIDKRFD
EVDARFV
EVDKRFN
AIDKRFK

EIDGRLN
KVENRLN
AMDKRLNRLETDIDELKRGQERLPKTIIENIGQFTENIVEHADDKAAALNDRVFSVETAIQRIYRLSS

>gi|52787642|ref|YP_093471.1| RpoE [Bacillus licheniformis ATCC 14580]

MSLKQYSEEQLKEMALVEIAYEIFSEHKKPITFQELTDQVASLLGMGKEELEDRIAQFYTDLNIDGRFLALSDQTWGLRS
WYPY
DQLDEETQPTVKAKKKKAKKAVE
EDLDLDEF
EEIDEDDILLDDVE
EDLDIDADEF
DEIDEADD
DELDDLE
DEILDDD
EDFDEEEDEE

>gi|16752654|ref|NP_444919.1| Hc2 nucleoprotein [Chlamydophila pneumoniae AR39]

MIGAQKKQSG
KKTASRAVRKPA
KKVAAKRTVKKATV
RKTAVKKPAV
RKTAAKKTVA
KKTTAKRTV
RKTVAKKPAV
KKVAAKRVV
KKTVAKKTTAKRAV
RKTVAKKPVA
RKTTVAKGSPKKAAACALACHKNHKHTSSCKRVC
SSTATRKHGSKSRVRTAHGWRHQLIKMMSR

>gi|15835917|ref|NP_300441.1| histone-like protein 2 [Chlamydophila pneumoniae
J138]

MIGAQKKQSG
KKTASRAVRKPA
KKVAAKRTVKKATV
RKTAVKKPAV
RKTAAKKTVA
KKTTAKRTV
RKTVAKKPAV
KKVAAKRVV
KKTVAKKTTAKRAV
RKTVAKKPVA
RKTTVAKGSPKKAAACALACHKNHKHTSSCKRVC
SSTATRKHGSKSRVRTAHGWRHQLIKMMSR

>gi|34495655|ref|NP_899870.1| hypothetical protein CV0200 [Chromobacterium
violaceum ATCC 12472]

MKTPWIKPVLASLAFGG
WWLLLTLCRD

RWVMVTLAAALL
LWEALPSGLRGRALLLAGVGCLLDLALAASGLLRFAGVGALPLWMMV
WWLSFAC
WWLWLLQSWRPSPRLLMMLGGLAAPLARYAAFKSQALQPGAA
WWLLLPLLALM
WAALLPQTLCLIRHAPLADRH

>gi|34495946|ref|NP_900161.1| cobalamin (5'-phosphate) synthase [Chromobacterium violaceum ATCC 12472]

MRGLILAVQFLTRLPTPQLADFKPEW
LAASARWFPAVGVIVGALLLAALQLGSLADPWLAA
LLALLAWVMVTGGLHLDGLADLADALGAAHRSCERFFEVLKDPHLGSFGVLSL
LLAVIAKLV
LLMLIARQAGTAWG
LLLVPAWARGFAVIWSATLPAIAPGSGERFAWRRDWAGMGLNLAA
LLALSAWLAPALLLA
PLAGLAWRTFLQRRLGGMTGDCLGAGVELCEIA
MLGLLLWRVPAPFGWLH

>gi|34496340|ref|NP_900555.1| probable alginate regulatory protein [Chromobacterium violaceum ATCC 12472]

MKKLTLLALSAVFGLSSAAGFA
AQVAPASAPAAKAAPA
AKKAEKKTPAKIHH
AKKGEKVAASAPQKAQ
AAKAEASAP
AKKAAHKHAKKHA
AKKAEAS
AQKAQAAKSDASAP
AKKAHKSHAKKH
VVKKEAS
AQKAQAAKSDASAP
AKKAAHKHAKKH
AAKKEAS
AQKAQAAKSDASAP
AKKAHKHHAKK
AHKAAASAAK

>gi|34496766|ref|NP_900981.1| hypothetical protein CV1311 [Chromobacterium violaceum ATCC 12472]

MASVNQF
AYVPGTN
TYKFAYGGSIPNMPVVNMPADTNWLRWAMLNDGEYYRMYFFKGSTANT
LYQAAFNPATG
AYEFGYNSIKELQITGAPADADASSLAMLYDGSTKTYRLYLRRLGAPTV
LYQFGFNRETS
RYEYGYNSIPTLNVVKAPGDTDWHRWSMLFDGSAYRLYAFKVGSTDT
FYQFAYNRQTN
QYEYGFDSIPVLSLVDIPANSNLTSMAMLFGQGDYRLYFQTL

>gi|34497204|ref|NP_901419.1| hypothetical protein CV1749 [Chromobacterium violaceum ATCC 12472]

MPDRQLRKPEPPLSVMLTRS
LPDRQLRNAQFNPAAAVECS
LPDRQLRNRPGEFGVDHSSS
LPDRQLRNPARWANPTTTGS
LPDRQLRKHPDTHIDVTPSS
LPDRQLRKGQPALALARKRS
LPDRQLRNT

>gi|34497906|ref|NP_902121.1| hypothetical protein CV2451 [Chromobacterium violaceum ATCC 12472]

MRPPAGEVRPDRGPAAAPAGTAPS
RRGCRTGLRTGPASPASAPASPSSPNAADPTCRLWRWCRLPGRRI
RSICRRRPSQPPGPRNRSP
CRTARSR
RWRCRRRAGA
ARALPRRRSGAGSRPAAPGCWPRTDC
SRARRRSRRS
AAGCRRARPDSCRTACPR
SGRCRRNSCSGTPDASHGGRDGWTASPAPGPA
SRSCRRRHYGCTAAPRN
EKTMRRRTSPAAGP
ARPAARRKE
TRSARNTRNCVCPATG

>gi|34498626|ref|NP_902841.1| hypothetical protein CV3171 [Chromobacterium violaceum ATCC 12472]

MTAFRRVRSCAAPFPPGGSPCGCCAESVARDGPTSAGGPAWTATRI
PGRRRRGSGCRAAAARRPAGDAAAALL
PAAGGRRRAARGL
SGRRSRPSGRCWQI
PSARARRCRHPQ
PGRRACRSVALPASI
PGRRGKRCAS
PGKPVRR
PAWRGSRRPPAPACRRAASGNRPAVARFPAPRPRPGCPPAGWPRSAPARP
SCRRGTRPDGGRRNGPGPAPWSVGRP

>gi|34498820|ref|NP_903035.1| hypothetical protein CV3365 [Chromobacterium violaceum ATCC 12472]

MSVSSLGNSSISLQYTQVDIQESFSS
DAGQGGGAVQGGGHRHRGGGGELTQSVLQALQQSGLLQQASAG
SSGDSGGPT
GGGDAGAVG
AAGNTGQAVHAFIHSLFQALRQEGA
SSASGSGSASDND
GDVDSGG
QDQDSGAAVGGSAYGGLQNNLQQLLQDLANGTQNDAASGIQSAFQSLQQ

ALGLSGGQSGQSSSLQTFLQNLAKDLQQNQPGAHAGVGALLSTSA

>gi|34499419|ref|NP_903634.1| hypothetical protein CV3964 [Chromobacterium violaceum ATCC 12472]

MENLGPVRRRHHRLVGAAVAGVAAELAAHQPVGHPDAAPGAADAAERRRQAVELSEGKTRG
GRGAAAG
GTGGGRAQRQ
ARGGARAARRGIARAPARGRLPLAVDRG
GPGQGRAGLA
GRDGRRRPLAGAGQ
GRGGGRRGDGRAVPAPVGDVVRPV
RRGAGRGLAAAVPGAD
APGAGRRATDPGQGT
EAAGGRRDPGRRRPCRRRLPLPARQRTACRPRAAGAAAGSGAAGP
GGAGGRR
AAGGARRR

>gi|53802323|ref|YP_112971.1| TonB domain protein [Methylococcus capsulatus str. Bath]

MAAVPLVQTDSKSSGAESLYRPGSAILVVVSLAHLAVWWSYRQLKLLE
PEERPSQPIEVSLEMIA
PAPKPAEPAAQ
PAPAPTAPPKPVTP
PKPKPVAKPKPLTKTL
PRPLPVQPERTLGDAA
PRSEPAAPPPEASSAAGGEGVPGGRSGSSV
PASTPASFNANYLHNPLPEYPAFARKQRWQGRVTMKVHVL
PSGIPAEVELQASSGHDILDESALETVRRWRFV
PATKGGKAVASWVVVPLEFSLTH

>gi|53805164|ref|YP_113181.1| TonB domain protein [Methylococcus capsulatus str. Bath]

MPQQAIEVSLEMIAPA
PKPAEPAAQ
PAPAPTAP
PKPVTPPRPKPVAK
PKPVAKTLPR
PLPVKPAEEASTQSEVSSAPP
PESASAARPDGGAASAPA
VPPSTPANFSANYLHN
PVPEYPAFARKQRWQGKVMLKVHVLPTGLPAEVELQASSGHDILDESALETVRRWRFVPATKGGKAVASWVVVPLEFSLTH

>gi|53803732|ref|YP_114654.1| TonB-dependent receptor, putative [Methylococcus capsulatus str. Bath]

MNIKLLRRYLPFAVGTVLVLLVALGVYLARDLFDKPAQTKKQIQQITVV
QPPPPPP
PPPPQEK
PPEPEVKEEKIE

EPEPEPE
PEPEQAE
EPPPGEDLGVDAEGGAGSDGFGLVGKKGGHGLIGGGGGNAIIWYGQQLQRTVSSELQRKLKDRARNSRYQAVLHIWISPQ
GAISRVELGKSSGIAEIDEALKAALATIHGQL
EPPPERMPQPVKIRVQSEG

>gi|53803594|ref|YP_114725.1| hypothetical protein MCA2309 [Methylococcus capsulatus str. Bath]

MNHRYSGAILGLAGVLFGIVVHARGFGGA
VQQSSRAAQQE
RQQSAEQ
MQQSRQQNLQG
MQESAQQNRNA
NQQSRQQYSAEKTYQRQLTGSQVHGQRARAYDD
VQENIQQTEKRRGAAKLLGGALRAKQ
MQQQQAQPQQ

>gi|53803104|ref|YP_115156.1| hypothetical protein MCA2759 [Methylococcus capsulatus str. Bath]

MGRPGENLDSDVRQRSRAR
RKPGSRRTAGGQPLIRTGTHHTAFSAVPSLRKSARLIRTPLADRLR
HRPTARKARRGRDKYRDSARNRAARRTPGPGRSGTPYRCWSR
RRSRRRRRACLRD
RRPQRRKDSS
RAPCRSRGTRPQHLHG
RRPAHQHDAELRRRSPPHP
LRPRRKHHQQPDLRYLHP
RRPRRRGIG

>gi|53802838|ref|YP_115429.1| hypothetical protein MCA3042 [Methylococcus capsulatus str. Bath]

MKKPFSTHILSAALLL
AAPLAGAADYPPDFKPSVIYRDPSLT
GKPAAPEA
APPAQPQ
AAPAAPAR
PAPAAPAKTEAAPEPQ
AAPSSPAKPAPDSGSDYYLFGGVVAALIGFVLWSSRRPATAHPAAT
AAAAAPAAAPAPAATGVAKYLQAQGLATGPETGVAKYLKALPEPVRTPETGVARYLKNLPLPEVAAAAETGVARYIKNLP
KPAVVATGETGVTKYLKSLNG

>gi|53802832|ref|YP_115435.1| hypothetical protein MCA3048 [Methylococcus capsulatus str. Bath]

MKKPFSTHILSAALLL
AAPLAGAADYPPDFKPSVIYRDPSLT
GKPAAPEA
APPAQPQ
AAPAAPAR
PAPAAPAKTEAAPEPQ

AAPSSPAKPAPDSGSDYYLFGGVVAALIGFVLWSSRRPATAHPAAT
AAAAAPAAAPAPAATGVAKYLQAQGLATGPETGVAKYLKALPEPVRTPETGVARYLKNLPLPEVAAAAETGVARYIKNLP
KPAVVATGETGVTKYLKSLNG

>gi|39933314|ref|NP_945590.1| hypothetical protein RPA0237 [Rhodopseudomonas palustris CGA009]

MKILKLSGLAAIAGLLLIAAPGERAQAAPLSTPGIAASVQRDAIPQTTEVQY
RHHRHMH
RHHHRSH
RHHHVRRHYGH
RHHHHHHY
RHHHHRHWR

>gi|39933316|ref|NP_945592.1| hypothetical protein RPA0239 [Rhodopseudomonas palustris CGA009]

MADLRVIRMPAAQSSRPSPGPLRLVRRIGVALMAATALTISAITPRPALALSPVTPATKSISSAAATQPIQVQHHFGHGG
GP
RVGGFHGGGFRGGGFRGGPAFHGGFRGGPAVGP
RFYGGGRHFYAGPRVYGAP
RFYGRPRYFGA
RHYGFHRRYYRP
RVYGYYGPAYYGSAYYYP
AYYGYYRPRVCRVVW
TYYGPRRVCHV
RHWRHHWH
RHWRHHH
RHHHHHRHWRVY

>gi|39933685|ref|NP_945961.1| hypothetical protein RPA0608 [Rhodopseudomonas palustris CGA009]

MRNGQNNKRLRNRNNGGNNNNNRRGQNPMTRVFESNGPDIKIRGTASHIAEKYVQLARDARSSGDPVAAENYYQHAEHYF
RLIAAAQEQF
RQNQPQPRTDNDTLDNDDDSEGDFSNFGAEPGLV
PVQQPQPYQ
QREQPQPYQ
PREQPQP
REHRPQPQFV
PREQPQPAAEGVDRLPSFITG
PQPQLGPNAFEGNGGGGERFP
PRRRRRPRGEPAAAQAVAPAPSGEDAGAGE

>gi|39934294|ref|NP_946570.1| hypothetical protein RPA1219 [Rhodopseudomonas palustris CGA009]

MERRNMEYLSALGPLLILMLFGGGLVIAAYLTFDDMRSRGGSFDATSFITSIRSTVDRVRALPPTTLIMIGAVVVAALFV
YLVSASG
YYARPAAVE
TSAAPAAVE
TAATPAPE
AAAPPPAAK

AAPAPAP
ATKAPAADKQA
AAAAPAAIAAVRNVTSTSCASDGCPVTCAANETLTSAYCVNGGAAKLADQLTVKDGSVT
AKCSPSASAIKVACARK

>gi|39934566|ref|NP_946842.1| hypothetical protein RPA1494 [Rhodopseudomonas palustris CGA009]

MTHIITRLYDSSSRLSALESDLKSNSFKYAVVTAGRSSAEETAAALLKAGVSRAEAATLADYLKKGGVVVAAQAEFGYGA
KLTKILDGYGPNPISLKSSSSSRAADQSDAAPFSNLIHAPVLLETSG
PYKSYSGDLLLVD
SDKTFSGTPLVISSDG
PYQSFSGTPLLLDSSG
PYKSFSGTPLLLDTSG
PYKSYSGEPLLINNST
PFSSWLGLPVLL

>gi|39934567|ref|NP_946843.1| hypothetical protein RPA1495 [Rhodopseudomonas palustris CGA009]

MSE
EYKGHSGHPLILKQEG
EYKGYSGEPLILKQEG
EYKGYSGTPLILEQKG
EYQSFSGTPLILKQEG
EYRGFSGAPLILKQDG
EYKSFSGYPLLLNI

>gi|39934881|ref|NP_947157.1| hypothetical protein RPA1812 [Rhodopseudomonas palustris CGA009]

MDIKASGRSALIIAAGLTVGLWC
VAPATPSFAAP
QAAASDS
GSAAKPAPKKKAV
KHHAKPAKAK
QAQAKPSDTK
PSDANPSEAKQSEAKSVVKSPVEEGRTA
DAQTEPKPLADASGALPASIANA
NAQATPAETA
QAQPTPSADASRAAAPTPAAGATKPAVIAADQVNDLDRAATDAPQL
QAQMQPQQPAAAETAATPQPQA
ETQAKATAHDDDAWDKASLIGKIFIAAGGLLTLASAARMFMA

>gi|39936159|ref|NP_948435.1| hypothetical protein RPA3096 [Rhodopseudomonas palustris CGA009]

MRLSHRLIPLAVLVALSSALAGCSSMSSFDPTDMFDFLDQKKKLAGERKPVFPEGVPGLEQGVPRNLYKDEVEQQQRQQA
AV
EAAPAPV
EDRPKPAARSKSRHS
SAARRPA
AAAPEPAAEPAAEEEGGAAAV

PPAPKPAKQSRRRVTTPPADDQ
SSAPAPAQSQ
SSGTAPSFPAPLPSGSFQR

>gi|39936243|ref|NP_948519.1| hypothetical protein RPA3180 [Rhodopseudomonas palustris CGA009]

MDDDKPIIEQVGEVITEATDA
AKTVVKKARKA
AKKVAKKVKAG
AKKAAKA
TKKAGKKA
AKKAAKKA
AKKAPKKAKKA
AKKAAKKTVKK
AKKAAPKK
AKKAKKSKR

>gi|39936655|ref|NP_948931.1| Collagen triple helix repeat:Antifreeze protein, type I [Rhodopseudomonas palustris CGA009]

MAEAESTRRGRPGPQ
GPRGRPGEPGRP
GPQGHPGRPGPEGPRGKQ
GPVGKPGPQGKA
GPQGKPGIAGKP
GPDGKPGPI
GPQGKPGPQ
GPRGEQGLRGEQGPRGEQ
GPQGPRGEQ
GPRGEPGPAGALPSIEQVMPWLHLIFDAYEDYKAQRAREARELEERLAAEALEQAAREAAEREVAAAIEAANAEAEIVLD
DETHAEGGKKKKKRKHKD

>gi|39936714|ref|NP_948990.1| hypothetical protein RPA3652 [Rhodopseudomonas palustris CGA009]

M
ARKYSKKASAKV
GRALKKRKAGTLKSGGS
GKKVKSRKQAIAIGLSEARAE
GAKAPKKKI
GKKASKTS
AKKASKSGRKTS
TKKTTKKA
TKSSSKRTGAKAAS
ERKSSKSSSARKSS
GRKSSKRKS
GKKTSSD

>gi|39936932|ref|NP_949208.1| Nuclear protein SET [Rhodopseudomonas palustris CGA009]

MPLLSNDHYRVGRSKTGLGLFAL

KPIKKGAKIIRYFGPLLDAHNKKHDGIENKYLFELNKRWTIDGSVRKNVARYINHACKPNAESDVNPRKKRVIIRAIKNI
EPGEEINYDYGTDYFKAYLKPIGCKCVSCEK
KRKKKAAEARAEKAKLKL
KAERKAQAEAE
KAARKAA
KAAGKAKPEKPSKPE
KVAKSKAPKAVKAKTTAKS
KGSSKAVKTKVVTAKAGKT
KAAQPAKT
TSSRKTAASKSKAA

>gi|39937370|ref|NP_949646.1| hypothetical protein RPA4310 [Rhodopseudomonas palustris CGA009]

MMVRNDIAPAPVLPSQQRENVMRSKTGLMIAAALLAASSA
SAQTGGTGG
GSGAGGTAGG
AASPGGISAP
SSVPQNSGGAGI
SSAPGGSS
SSATGTGMSGTTTGSSLNSGTNQPGVPN
TSTPGSIGTGASP
SGLPGDS
TSSPGHPGQVGK

>gi|39937407|ref|NP_949683.1| hypothetical protein RPA4347 [Rhodopseudomonas palustris CGA009]

MAEADLDAIIRQLAKQQTKALT
AAVKKRRDALL
ARAAKAK
DAAGKAQGKALATAAYEQGL
AAAKRLQTAADNAADAYARAMR
KAAEDAATAA
AAASKPKADDKP
KATAKSKTAKS
KPAKAAGDKAETKPAKTGAKK
KAAKKAKR

>gi|39937516|ref|NP_949792.1| hypothetical protein RPA4456 [Rhodopseudomonas palustris CGA009]

MNESVNDDSGRNQAPNQDKAAAPPRKSADKRAPLVDNLIDLVVDGAVEITKVAAKRAVRRLVSGPV
KATVTKAAARAGLAEFTKPLAKPSPS
KGKTTAQAGAASGRKKPAKTAA
KSAASKSGAS
KAKTSKASAS
KAKAAKAKTATSK
TAKASSA
KSRTTKSAASKS
KAKSAKTAKTKKR

>gi|39937765|ref|NP_950041.1| hypothetical protein RPA4707 [Rhodopseudomonas palustris CGA009]

MPISNRRLRRRLVPRRVGIFTAAMLATALLGIPALA
QNYAPPPAD
RPPPPPPARTDNPGLVDEIQKLLKTPSSLLPSTSLLPSFGSDSKD
APPPPPP
TPPPPPQQSPAGPPASSQT
APPPPPPPIGG
RPPSPPP
APPPPPAPAASSAPIVPSMVSGRQVCPAAPGSGADCQAAADILCRSKGYGGGKSLGVDATEKCSAKVLIPG
RPREPGDCRTENFVTRAWCQ

>gi|56551029|ref|YP_161868.1| hypothetical protein ZMO0133 [Zymomonas mobilis subsp. mobilis ZM4]

MTRDGLKSPL
QKAAERGDAKAQYALGNAYSKGQDVSKSDEQAVSWY
QKSASQGYAPAQAALGYAYSSGLGVTHDDQQAVSFF
QKAANQGLAEAQYSLAIAYYTGRGVTQNYEQASFWF
QRSANQGFVPAQFYLGVMYRNGAGIPEDDDRALFWF
HKAADKGYADAQYNLGLIYHEGKVVKKDEKQATFWY
QQAANQGLVEAEFNLGIAYLKGQGVQKDKDKATFWL
EKAADKGDSHAQDVLEMMNK

>gi|56551031|ref|YP_161870.1| hypothetical protein ZMO0135 [Zymomonas mobilis subsp. mobilis ZM4]

MQADKERYDREMISSLEVRAKAGNSLAEEVLGEAYAHGRGRPQD
DEKAVYW
YQKAADKGMKESQYNLGDAYLHGRGVGVD
YEKAAFW
YRKAADQNFIQAQYNLGLLYVKGQGLPKS
DEHAAFWWQKAAEQGEAKAQFNLGVFYHNGRAVPKN
DARAIFWMEQAAHQGLVEAQMLLAMAYASGQGAPKD
KEKAVYW
YQKAADQGNAAAQKALQALSY

>gi|56551032|ref|YP_161871.1| hypothetical protein ZMO0136 [Zymomonas mobilis subsp. mobilis ZM4]

MKIILWIVALSFLAVRPVLAGNDRTVRLLSDGYSIVEIQKNANAGDSQAQFSLGYAYAFGKNLPQD
YEKAAFW
YRKAAEQGFAPAQCQIAVAYVSGAGVSQN
YKRAAFWFDKSARQGDSSAQYYLGLLYTQGIGVPQN
DETAVFWLQKAAHQKYRAAEYDLGNAYYDGKGVLRN
GEKAMFWWQKSADQGFAAAQFR
LGRAYYWGDVVAQDQKVAQAWIKRAADQG
YVAAIVWLQNAQH

>gi|56551890|ref|YP_162729.1| hypothetical protein ZMO0994 [Zymomonas mobilis subsp. mobilis ZM4]

MHSTMSSNRLTKPVLGFLLGATAAFGLSPVIAGAAYAAEQGVF
QKAGNSIQNTA
DNAGKAVS
DTAEDAHDGAKNVTNKARHSAKRSWNKTKSTAKKTT
DKSGDAL
DKSWDKTKSSAETAT
DNAGHTIARSA
DKAGDAIENTT
DKAGTGIKKGANTVGKAFSGAWKDVTGGNKSKGK

>gi|56552263|ref|YP_163102.1| hypothetical protein ZMO1367 [Zymomonas mobilis subsp. mobilis ZM4]

MSGCQDVR
MSGCQDVR
MSGCQDVR
MSGCQDVR
MSGCQDVR
MSGCQDNSPKMIWASCFLMIRNLTFKNKNNLLLRGIGLE

>gi|56552388|ref|YP_163227.1| hypothetical protein ZMO1492 [Zymomonas mobilis subsp. mobilis ZM4]

ML
PVCPSARL
PVCPSARL
PVCPSARL
PVCPSARL
PVCPSARL
PVCPSARLKNNYFMGQVR

>gi|56552483|ref|YP_163322.1| hypothetical protein ZMO1587 [Zymomonas mobilis subsp. mobilis ZM4]

MGKNLFWAGIGLLALTLPVAVMAYPR
GYHYGHYGRSGYIRNGYSSGYVRGYDHS
GYIRHYGWS
GYVPRYGSM
GYNHRYN
PYNHNPM
RYNHSYNHLDRQLLYPSNRW
DGAYRYNGYER

>gi|56552537|ref|YP_163376.1| hypothetical protein ZMO1641 [Zymomonas mobilis subsp. mobilis ZM4]

MRPITLKIAIASLAFLSLSACESPFRANVTRFQMMPPPAGQSFVIQASNPQDKGGLEFEHYADIVAHQLWQQGYQPAAPG
APSTLIVKLDYGIDNGQQKVASTSY
GYGGWGG
PWGGWGGPWGGWG
GYGGWGGW
GYGGWGG

GWGGGTDVYSYTLYTTFFKMEIDRSNDGQRVFEGRARAQVTTDNMTQLVPDLIQAMFVNFPGHSGEDVEVKLTNNKKKHK

>gi|56552816|ref|YP_163655.1| hypothetical protein ZMO1920 [Zymomonas mobilis subsp. mobilis ZM4]

MLKNILRLLVAAIAASSLTATVVTPALAHH
HGGPGGP
GGWGGGPGP
GGWGGGP
GGWDGPRHHHHRG
GGWGPGAAFGMGALGLMTGAIIANGSYNRGYYDGG
GGYYGGYS
GYYGGGYANCNTRMQWDPYIGDYVSMRACY

>gi|49476747|ref|YP_034567.1| hypothetical protein BT9727_0212 [Bacillus thuringiensis serovar konkukian str. 97-27]

MLQNFYYTNTLDRNTINAVNTTAPLADPTKSPFVTCNTCSFIFFILPSFVFCYKTFIILIRLIEI
RLMLLIL
PLHLLILRS
HLLLLVILALSFCILLSFSFYKTSIILIRLIRI
RLMLLIL
LLHLLILRS
HLLLLVILALSYYVSSFLISYLYVISKKCLFHSFYVDVTLNLIT

>gi|49480527|ref|YP_035770.1| hypothetical protein BT9727_1436 [Bacillus thuringiensis serovar konkukian str. 97-27]

MFQQSNVYQQS
NPYAQQNMYQYNTDTYLRYNMYPF
EPHYGNQNYYQPFE
VSFMNQPQQPQ
QPYMNQQP
QPYMNSQYYMPPP
SPYSNQQAMFYPPK
QPYPTQNKQKQQQQPSQFSSFVSQFKNSDGNYDVNKMMNTAGQMMNAMNQVTGIVKQVGGFFGK

>gi|49480532|ref|YP_035868.1| hypothetical protein BT9727_1536 [Bacillus thuringiensis serovar konkukian str. 97-27]

MSLFKNRIPKNSVSFSEETYELQFPKPIPV
QVEEEIQVDHE
ELLAQQQSLHTELN
QLRQEQQ
MLERERQ
QLLHDQEQF
QIHVQQQMKEIESARI
QFQKEQQETAYEWTELLWDQSFQLAEKIVNQAVDSRLLDVLPILTGIVQTLPTSFEKLVITVHPETFE
RIQEEKENTKEYWLLQLVEWKYDFSLQFGEFVLEEEKEFFEFKFAPIFA
KLRQKWEEE
KLFEEQNV

>gi|49481252|ref|YP_036192.1| exosporium protein B [Bacillus thuringiensis serovar konkukian str. 97-27]

MKRDIRKAVEEIKSAGMEDFLHQDPSTFDCDDDCTTKIECSDDCNC
PRTRCTRVKHCT
FVTKCTHVKKWT
FVTKCTRVRVQKWT
FVTKVTRRKECV
LVTKRTRRKHCT
FVTKCVRFEKKFYWTKRCYCKKCEFFPNGHGGSCDDSCDDGKKWNDCKDGGHTFPSCKNKKFDHFWYKKRNC

>gi|49477618|ref|YP_036436.1| hypothetical protein BT9727_2109 [Bacillus thuringiensis serovar konkukian str. 97-27]

MENQLKEIFGALIAAIGTITSAIGSTPFYFISSNVREDLNIY
GNTLQAV
GNALEADGQE
GISLEKI
GNEIQSTGNVTVISGLVIDFKDETKIKLVIA
GNWTQALGGLTALADEFEDASDKDESFNII
GNLLQAI
GNSLQAIGGIYELKSIRKERLDSKDELVNDTEGNLDNQVNSELDKKKE
GQSIDTI
GSWIQAVGSVFSLIGQIREESEELEGSDN

>gi|49477699|ref|YP_036553.1| hypothetical protein BT9727_2226 [Bacillus thuringiensis serovar konkukian str. 97-27]

MIAGVLRKIIVA
MIGGNLKISIIVM
TIGGNLKISTIVM
TIGGNLKISIIVM
TIGGNLKISTIVM
TIGGNLKISTIVM
TIGGNLKISTIVM
TIGGNLKISIIVM
TIGGNLKINTIVIMIGSLSEISVVVNVSLFVVIVKRNIVKFFNASLKAKNMS

>gi|49477823|ref|YP_036780.1| hypothetical protein BT9727_2454 [Bacillus thuringiensis serovar konkukian str. 97-27]

MNQFQQELQSLNLNDYQTGNVVYWDQQQSQYPYYYIQDDARR
CGGCGGC
GGRCGGC
GGRCGGC
AGRCGGCIG
CAGCFGCFNCWNWWII

>gi|49479603|ref|YP_036958.1| hypothetical protein BT9727_2634 [Bacillus thuringiensis serovar konkukian str. 97-27]

MRTENLIIN
GYGSSNGGEFHKVQLN

GKGTVNGNVECEKFECN
GYGAVTGDLKSSSARIS
GSGKVDGTVHAKTMRID
GKGTITKDVKANSLKIA
GKGTIGGNVTGEEFKIN
GQATIDGNCEVDTFSSD
GQFEVGGLLSADEININIHGTCRAKEIGGQTIKVRHRLGTFSRLFKSVFGLQLEAELLEGDNIDIDYAHIRTVRGNNVTV
GPNCEIELIEYTGVLTVDKSANVKEIKQV

>gi|49478041|ref|YP_037250.1| conserved hypothetical protein, glycine-rich repeat [Bacillus thuringiensis serovar konkukian str. 97-27]

MNSQRQGYA
GWWQGGH
GGWPGGH
GGWQGGH
GGWQGGH
GGWQGGH
GGWQGGH
GGWQGGH
GGWQGGH
GGWPGGH
GGWQGGH
GGWQGGH
GGWPGGH
GGWQGGH
GGWQGGHGY

>gi|49478075|ref|YP_037307.1| conserved hypothetical protein, glutamic acid-rich region [Bacillus thuringiensis serovar konkukian str. 97-27]

MSKKLLLFALMSLCVIALSACKGN
EEKLKGAPNEQKV
DEDKKVE
EEKRKQEEQQRVE
EEKRKQEEQQRVE
EEKRKQEEQQRAE
EEKRKQEEQQRVE
EEKRKQEEQQRVE
EEKRKQEEQQKVE
EEKRKQEEQQRVA
EEKRKQEEKKAQQQQSTQQ
ERTQKQEKPTQATGGKPTRSQISVGSHVVIQLDNDYSKTVSGVVKDILTHSETHTYGIKVRLQDGQIGRVQSVK

>gi|49478098|ref|YP_037340.1| hypothetical protein BT9727_3017 [Bacillus thuringiensis serovar konkukian str. 97-27]

MKRKISTAVVGLSVFSFGLFGVNIGDGGVVKAE
HGETFAPKLEN
HGDYGGAPAFN
VGDTPAPKSVSIGDGGGAPAYDHGRPPISSYN
IGDTWAPSYTNT
HGEGMAPKQES
HADFFAPTLEA

HGNTGAPALN
VGDTPTPKLEG
HEHFPAPKQETHADLSSKLDS
HADTWAPALN
IGDTGGAPKQ
EGHTHFPAPKFEA
HGNTGGAPANSEHGTGW

>gi|49478182|ref|YP_037447.1| hypothetical protein BT9727_3124 [Bacillus thuringiensis serovar konkukian str. 97-27]

MKMKKVAIASITIGTMLTMAACSPSADKETEKKEQVTAQNEGEVKT
TSTDETS
KTTKDTDKETNTSSN
EKTDKDT
KNTKESSGTESNG
KTSTQNENV
KTKEKETTG
KTNDQTTNG
KTTDQSTNSDGKTTKDPKKSVAVKTSVKEVAALVDSLDKQLAANPKLETVNKLGKQINAKWDVIEKELETSHPAEY
KTIGQSMYPLIVGAEKEKIDITKMKSLTTKTKKDLNQLLTKLS

>gi|49478600|ref|YP_038345.1| hypervariable Bacillus group-specific protein; vrrA gene product [Bacillus thuringiensis serovar konkukian str. 97-27]

MFPKSPNRQMYPNPGQ
QPYTPYPIPQLPPMAQKKKGFLAKLFKKHDPT
EPFMQMV
PPYRQMEG
PPPMMHQ
QPPPQYR
QQYQQQYP
QQYSQQY
QPYMQHHPEQMIPPQMYESNETRGGAATTAASSSGIGSFFSNLISNPTNMINNIEKVSQVVQSVSPVVEQYGPIMRNLPS
IVKILTSGKSTEENPTEDQTEDLTEKVEVATPPPPQKKRKRKKMVIEPVIEKEVREEPVQKIATKPKLYV

>gi|49478826|ref|YP_038428.1| hypothetical protein BT9727_4110 [Bacillus thuringiensis serovar konkukian str. 97-27]

MAGGSRFQ
QKQQKRRQNGVLNIAIAIVLVAVAIVAYQLFVPDTKEQASSSDKKVSQQTTK
ENKSEKAKG
KEETKKNEQEKTEAKKKEEE
EEEKLKA
EEEQKKA
EEEAKANEKVPA
EKTQPKATDAYTKSSWKPVGTEQGATPAMTFKKGTADWNEMNQAISAAIDVPVEQLIIHRIGNNGKNKAYGNVQDKQSGK
RYYVNIDWVENEGWKPVLVQTLN

>gi|49480717|ref|YP_038783.1| hypervariable Bacillus group-specific protein [Bacillus thuringiensis serovar konkukian str. 97-27]

MKGMDNNA

PHGFFGGGSDSYEQMMLMGGAGQQGYGGIPSWMGGSPGGFSTSVAGVQTGFPVPGVGVVAGGIGGFPQVIHG
HHGHHHHQG
HHGHHGHHHHQG
HHGHHGHHHHQG
HHGHHGHHHHQG
HHGQHGHHH
HHGHHGHHQQ
QVHHHGHHHIHPQAILYQT
HQGHHGHHD
HHGHHGHQG
HHGHQGHHGQHGHHGQHSQQ
HQQHQQYQQYQQQSSPWAGGIGAGAAGAAAGAASISGAAG
HAGHVGH

>gi|49478855|ref|YP_039045.1| hypothetical protein BT9727_4736 [Bacillus thuringiensis serovar konkukian str. 97-27]

MYSTYYKTTKKKTVAKLKVTRKLEDILPAFY
IVFISRHHNKKKLPSQQLLFHIPHPVPTGLYQRFFKYIDDFYTFIGDFF
IISTVRH
RISTYRQFSAIY
LYSIPRHYNKKKLPSQQ
LLSHIPHPVPTGLYQRFFKYIGDFYTFIGDFF
IISTIRH
RISTYRHPSTLKTHNIPYPPSNLLPNSTIKKITNKIASIFNATYATLKTT

>gi|49480420|ref|YP_039332.1| DNA-directed RNA polymerase, delta subunit [Bacillus thuringiensis serovar konkukian str. 97-27]

MDFKQYSPEELKECSMIEVVHSVLGDKRQATTFNELVQEIAQVLGLSQEQVNAKIAQFYTDLNIDGRFINLGENRWGLRS
WYPYEQIDEEILPQPKPKKKRKV
EDDGFDDYIE
EDEDFDDAD
VTEDEDDDVEDLDKVLEDEDG
DDDDLDDLDE
DDDDFAE
EELEYDET
EEEEEEEL

>gi|49474943|ref|YP_032984.1| Cytochrome o ubiquinol oxidase subunit IV [Bartonella henselae str. Houston-1]

MSMHNETHNPSTSSYLVGFILAVFFT
LGSFIPVMYGMMDSWAISTKVVYLIGMAIIQIIVQIIFFLHLNSGPDAKWNLSALWFAAICVFVIIGGTWWAISHLNYNM
MGGSGRVVEPEISGGTNPSVSEQLSDIQEIEKHLSGKDISAQKA
PAVQIPVEPV
PGSEMPLGQAPTVQTPIEPI
PGSEIPLEQA
PAVQTPIEPV
PGSEIPLGQA
PAVQTPVELA
PGAEMPME

>gi|49476136|ref|YP_034177.1| hypothetical protein BH14780 [Bartonella henselae str. Houston-1]

MWMYCRKKLKIVFYMAIFIACSASFAQSAAHWVSKYFTHDSE
SSATDKP
SDAVDETA
PGATDKP
SDAVDETA
PGATDKP
SDAMDEAA
PDAADKP
SDAMDEAA
PDAADKP
SDAMDEAA
PDAADKP
SDAVDEAA
PDAADKP
SDAMDEAA
PDAADKP
SDAVDEAA
PDAADKS
SDAVDEAA
PDAADKS
SDAVDEAA
PDAADKP
SDAVDEAA
PDAADKS
SDAVDEAA
PDAADKS
SDAMDEAA
PDAADKASDVGE

>gi|57237395|ref|YP_178408.1| membrane protein, putative [Campylobacter jejuni RM1221]

MTHEQKIAFFKIKR
LSVYIFFCALYNFYYFFSYEGVTSSYEAGNEILYIGGIFLCFFSCIYIV
LAFIFMV
VRIHFLFSSNKILIYFLLAIINLGASFFLLENKILFLGL
LALFFVFYFLFSLT
LALLFKEVK
MSFLFAF
LCMLLAF
LSFIFIYK
LSLLFLFLTLHF
LSFAYAFERYKILKKLD

>gi|57237468|ref|YP_178481.1| conserved hypothetical protein [Campylobacter jejuni RM1221]

MRIFLALFLFFNSLLALSELEEGL
KFYEANKFDKAYEIFKSLCE
KDISKACFSLAFMYESAKGMSKDLNQAY
KFYDKACKLGLASACSNMALLLQNQGYKNEAL
LAFNKACTLGESLSCNNIA

LFYEKEKDGQMAS
SFYKRSCDLKNARACYQLG
SLYDKGELVKASVKSAL
AFYSKSCTLGFGESCYLLGRYNQLEKQDLTKAK
RYFGMACD
QKHQEACAAYKELNSKDIELY

>gi|57237473|ref|YP_178486.1| conserved domain protein [Campylobacter jejuni RM1221]

MQISSSYNSYDYYSSKPSF
DKTNQSDST
QKDEKNLSE
NKDDKEKNGEQTQMVNGVELSQKEVAQVRELQSIDRNVKAHEAAHQAAGGGLAGAASFTYTRGPDNQMYATAGEVPISMQ
KGNTPEETIANARQIAAAAMAPADPSPQDYKVAANAAKMEFEARAEAMKLKAEEAK
EKEEENK
EKDDKNSENSLEKT
DKDSKNNDKNSNF
DKDFKNFVARTY
QQNSQNNDIKFNIAS

>gi|16124535|ref|NP_419099.1| hypothetical protein CC0280 [Caulobacter crescentus CB15]

MTMRSTLMLIAAGAAALSLAACSQEHAA
EVKEGAKAAASDVKDAAHEIANDP
DVKEAGTAIKE
SVKETGA
DLKDAAGEAGQ
ELKEAGK
DLKHDAKEGAAESKK
QLHEATK

>gi|16126129|ref|NP_420693.1| hypothetical protein CC1886 [Caulobacter crescentus CB15]

MTIKPGRRNRGMARRASLIAAVSVLSGLAVAGVAVARQAGTPQA
QTPPAAP
AATPAPPAPVG
ELRPAQPV
QVAPATAN
PNPPASPAPGAATQTKPTT
ATKPAEPAKRARYSVAILQALDKVTTETMRFEVPIGQPIRYKTLIFTVRACETAAADEVAPESAAYVVVDTQPKA
QAGRAAPPGRQIYKGWMYASSPGLNPLQHPVYDAWLIACK
QSIPVAPAAKP

>gi|16126134|ref|NP_420698.1| pentapeptide repeat family protein [Caulobacter crescentus CB15]

MCAN
CDLAGRKLMNAKFTGANFAKAVLIG
ADLRGAMFYGSDFGWADLSR
ADLRGAEMRGANFTRANFTD

ARMSGVESSGANFAQATLVR
VDLSSAELHA
ANMAGANLQKARFANAELIA
ANLSGANARD
ADFSNADINH
ANFQGARFDGARFHN
ADMTGSNLRGGIFNS
ADFRNADLTMVNLSG
ADLSSARGLEQEQLDEACGDSSTRLPSGLSVRTCNGLRSGQHFMLLRAVPKPPAPPAPPKAPKGPK

>gi|16126566|ref|NP_421130.1| TonB protein, putative [Caulobacter crescentus CB15]

MALALNPNGFGIP
GLDNPPPKPAKVLLIGLGVSGALHIALIGYLAYQKWTQPL
PLMSPDTGFIVETVPLT
PKPPPPSP
PTEKPPPAAAAP
KLHIPAPT
PLETPPPLFADPRIIPAGPTTTGPVTTLA
PAATTPPS
PPQPPPKVITRANWQRIPSADEMARYYPESAQRRGLSGQVTLNCMVALNGTVRDCVVATESPADEGFGSAALKISRFFKM
KPQTENGDPVDGATVRIPIRFNAGA

>gi|16127165|ref|NP_421729.1| cytochrome c, membrane-bound [Caulobacter crescentus CB15]

MSDLTFNKIAGGVLLTGLVIFGLREASSIVFSKPEVEKAGYEIAVQE
ESAGGAAVVESLPDWNAVLTPANVA
AGQAVAAKCAACHKFDASNANGTGPGLWDV
VGRKPAAHADFNYSTGMKEFAAKTPVWDYDHLYEFLKAPGKYVSGTNMTFVGLKKQEDRIAMIAYMHSLGSKLPVPPPR
PVAAPAPA
EGAAPAA
EGAAPAA
EGAAPAA
AGAAPAAP
AAAAPAAPKA

>gi|16127252|ref|NP_421816.1| hypothetical protein CC3022 [Caulobacter crescentus CB15]

MSSLKLPADNPTFLAMVGSCQVWSPGLDNAGKRPRSCCLGLPNAHFETRSPPMFAKKKETPAKLALTPLE
PQRAATPQVAQPQ
PQPAPAP
ARPKPASLIAQGLTIRGDVVGDIEVHLDCVVMGDVKVGKLIVGPNARIEGAVVAQAMEIHGMVIGAVTAQSVRLYGTARV
DGDITHEQLAMEPGAEFQGRSLKFRR
PAPTPQP
GQQAPAA
PTPAAPPTPPSASQTPYNHSAH

>gi|56418815|ref|YP_146133.1| hypothetical protein GK0280 [Geobacillus kaustophilus HTA426]

MTMSDALLAQVLDRLNAMDGTLQKLVEGQE
RLEREQIAI
RQEQEEI
RKEQQEI
RKEQQEI
RLEQQRL
RQEQDSL
RAGQKELNGIVAALLH
RQDETDAKLEALAADVHQLHQEFASLKQTVDRLEQKMDSRYEQLAAQIEDIKLDLNFIAHRTVDNERIIYKLKEMLFR

>gi|56418970|ref|YP_146288.1| hypothetical protein GK0435 [Geobacillus kaustophilus HTA426]

MEEKTIIQAV
LQALEQHSDKISSK
MQEMEQRLHHE
MQTMEQRLHHE
MQTMEQRLRSE
MQEMGQQLRNE
MQEMGQQLRSE
MQEMGQQLRAEIHAVETRLEAKIDSLREE
MQQMNRQLNERIDHLEEELTDTQASVEVLTTKVLHHDRKLRQLTKQA

>gi|56419424|ref|YP_146742.1| hypothetical protein GK0889 [Geobacillus kaustophilus HTA426]

MERDTIVKEVVEALKMFSHD
IRTQIDEMGSQ
LRAEIQDTANQ
LRAEMQDMANQ
LRAEMQDTANQ
LRAEMQQ
FRAEVNERFDRLEQKFAG
LRVELTETQETVHFLASKTIQHEKKLHHLAKQLESIR

>gi|56419518|ref|YP_146836.1| hypothetical protein GK0983 [Geobacillus kaustophilus HTA426]

MNHWDPYMNGPFSTSSGHDNEAGWMENN
VPPMAPLSATTPNEATPTQVMPMAY
YPPAAPYGYW
CPPVAPPAPLVSPVHLDSPDSPDSSDMMLPAPPLGLGAGPNMAHASVWPMGGKEN
TPPMTPDDAAIP
YGMTAPYAEN
VPPAGSTAPVAVPGTLSGMAAGATPGF
VPGAHPYAPLY
YPPMAPQP
YAPVMPYGYGYP
AAPLPPYYGGGTPYAPGFPISYGPYR

>gi|56419949|ref|YP_147267.1| hypothetical protein GK1414 [Geobacillus kaustophilus HTA426]

MSEEILLQLFDRLGHIENTLHLLVESHKRL
EAEQQEM
RKEQQEI
RREQQEI
RKEQQEI
RREQQEM
RKEQQEI
RREQQEM
RKEQQEI
RREQQEM
RKEQQEI
RREQQEM
RKEQQEI
RREQQEM
RKEQEEF
RRIQQALLEGQKRL
EEEQAHIKAGQKELYDLMSALLHRQDETDAALDALAMDVHKLHGEVSSIHQRLDALEKKVDSHAQSFTEQLNNIKLDITF
LSNKVIEHEREIYKIKNVIL

>gi|56420021|ref|YP_147339.1| hypothetical protein GK1486 [Geobacillus kaustophilus HTA426]

MGAISNEMIWQAIK
ELAEQLQ
QTNGQVQ
QLGQQVQ
QTNEQVQ
QLGQQVQ
QTNEQVQ
QLGQQVQ
HLNSQVQ
QLNHQVQ
QTHEQVD
QLAKQVQ
QTNEQLN
QLSQQVADMDKRITDVANGQNILVEELFENKKEIKRVKTALNMY

>gi|56420241|ref|YP_147559.1| hypothetical protein GK1706 [Geobacillus kaustophilus HTA426]

MTAGERQGDGENQGEQ
KQQESQASH
QEKKGKK
EDEQHEK
EKEKKKKGEEKAKDDKKTDSAET
KEEKSEKEKKG
KEENHKMK
KHEEDKKKGD
PHEKSKK
KDEEKPKKEDQNHKHD

>gi|56420540|ref|YP_147858.1| hypothetical protein GK2005 [Geobacillus kaustophilus HTA426]

MEE
LLRQLTTLVQGLVDDMSV
LKQDMQTVKQKLGQLEQRVDKLEQRMDKQEEELRW
LKEQMKENTEMIKAIRDNQLEQRSIQDG
MRHEIAS
LRGEMAG
IREQMATKDQLSR
LREQMATKEDISLIHQRLDYQLGRIARVEEEMYLLKKAE

>gi|56420556|ref|YP_147874.1| hypothetical protein GK2021 [Geobacillus kaustophilus HTA426]

MSEPWHLILDKLEIMQQEMAEM
KANMATKQELEDI
KTRMATKEELEHI
KANMATKQELENI
KANMATKQELEDM
KANMATKAELNEI
KADMAKGFAAVHQAIREIDAIVKRLEQNQEQQMQLLLRQERIIDVLCRRSLEHEAAIADLRLAIKS

>gi|56420588|ref|YP_147906.1| hypothetical protein GK2053 [Geobacillus kaustophilus HTA426]

MEGETIVQQILKALDLHSGQLRQQ
MKEMEDRLRVE
IQDTENRLRAE
IQDTENRLRAE
IQDTENRLRAE
IQDTENRLYAK
IRETEDRLHQK
LLDLEKRTDARFERLETKMNYWRLEWSETQETVDFLARKTLQHERKLRSRAGQNEIS

>gi|56420952|ref|YP_148270.1| hypothetical protein GK2417 [Geobacillus kaustophilus HTA426]

MSHELVELLRSVLKEELQP
IHERLDR
LEAGQQR
LEERQQR
LEERQQR
LEERQQR
LEEQQQR
LEEKLDHFIIETRSHFKKIEDELADHRRMFEILSHDWRNANIAIEYLNHKIGQHDLEIYQLKKRLELQETK

>gi|56421243|ref|YP_148561.1| hypothetical protein GK2708 [Geobacillus kaustophilus HTA426]

MDGELRALLNAVLERLDVQGAYI
EQMRMDIVKLQGSVKQLEDKV
DRLQSDV
DAMKKDM
DAMKKDM
DAMKKDM

DAMKKDM
LGVKKDVAALKEGQVRQEKIMERLAIRSIEQEAEI
DELRKEISA

>gi|56421437|ref|YP_148755.1| hypothetical protein GK2902 [Geobacillus kaustophilus HTA426]

M
EALLNQILDRLDQLHSSVGVL
TSEVNEM
KNQLNKI
EARAGSI
EERVNSI
EARVDSI
ESRVNNIETNMATKDELAEL
RSKVDDIEAKMATKDELAEL
RSTVNGL
QSNVNEIQAKMATKDDLVPIRQAVMEIDQIVKRIEVNQERHEHILAILSKRSIEHEASIASLRQAQ

>gi|56421643|ref|YP_148961.1| hypothetical protein GK3108 [Geobacillus kaustophilus HTA426]

MEELLRQILS
KLDNMEKGFLVFQEETRQRFE
QMDGRLTEMNG
RLVGLEQEMGAVKE
RLDRVET
RLGGVET
RLDGVET
RLNGVEI
RLDGVET
RLDGVET
RLDGVET
RLNGVET
RLDGVET
RLNGVET
RLDGVET
RLDGVET
RLDGVENELREVKE
TLHRVETQLVDDVQILLRKIHEKVEEKSFEIYALNRRLHHAEAAIEQLQSKA

>gi|56421856|ref|YP_149174.1| hypothetical protein GK3321 [Geobacillus kaustophilus HTA426]

MSQELKELL
RSVLQEELKPVHDRLERLEGAVGHIRQD
VSQLQQD
VSQLQKE
VTQLQEGQRTLETT
VSQIQKD
VAQLQEGQNALQRD
VAQLQEGQRTLETT
VSQIQKD

IIQLQEG
QNALQRE
IVQLQEGQRTLETT
VSQIQKD
IIQLQEGQQIVEHEMAG
MKQMQERILEELRISKHNQALILNDVTSIRKSVDVT
VSYLQRRSDVMFDKLIDHEKEILNLKSKMPN

>gi|56421863|ref|YP_149181.1| hypothetical protein GK3328 [Geobacillus kaustophilus HTA426]

MKKKALLIRLFSALVALFLAAGCN
MNRDNNNPP
PPNNQND
VNDQDVNPADD
INQNDNVDNDKDNDG
IPDNRDTD
LNDNQNNQLP
PGDDTNNGPGDDERDTVPDREDPIEDPRD
ANDKDNKDE

>gi|56421925|ref|YP_149243.1| DNA-directed RNA polymerase delta subunit [Geobacillus kaustophilus HTA426]

MSLQQQYSPEELQEMSFVELANLILLDKREALPFDQLVREAAALAG
ISEEEMEARLAQYYTDLNIDGRFICVGENVWGLRAWYPFDQTEDETVTIVKPKKKKK
ALDDEYDEYDEL
LEDEDLDYDDLDEYDGEELELDEDEL
LEDEEFDL
DEDVAFDDDI
LGDEEFELDDA
PLDEELD
LEEPDEDE

>gi|54307588|ref|YP_128608.1| hypothetical protein PBPRA0374 [Photobacterium profundum SS9]

MTFPSNISALAIAVSTLVSTVAFADDNFR
QHDAHIHGVVELNIAQDGNDLLLEITAPGSDIVGFEHAPQTQDQTQAIQQAMLTLKKPSSLFEINSDANCELTDSDVNQS
LSK
QNDGHDHS
NHDGHDHS
NHDDHDHS
NHDDHDHS
NHDDHDHS
NHDDHDHS
NHDDHDHSEHANKHGEFTAQYSFSCKNIAQLNEIKLYWFNHFPSTEKISVQAITDKGQQANQLTENNTTFKFQ

>gi|54308452|ref|YP_129472.1| hypothetical cytochrome c-type protein NrfB precursor [Photobacterium profundum SS9]

MGNYKVAMNTMITFFVA
MCIFGFSINVAADSNTQDATTATPAVINQESGQHVVKFIRDRDY

ACTQCHKDEKDVLKGAHGDAIHAQTGRDV
TCVDCHSNVGENHRDGAKVVTKFAPAQSVAGSEKPAADVAWIKQQNN
TCINCHEPKDLREVNWTHDVHALDL
SCASCHTVHPDTDPMKGIDRKSKIK
MCVDCHSDQKKEK

>gi|54308460|ref|YP_129480.1| hypothetical 17 kDa Surface antigen [Photobacterium profundum SS9]

MNKNIQN
NNNNNNN
NNIKIKN
KDVNIKNKDI
NNVNKNN
NVVINNNRHNTVVVPSNRHRTYRNVVVVRPHGHTYYG
NYYNDNDAWKWLAFTAITLKLL
DNLNEQQEREHEAAQVKATTAPINEQINWSDGGASGSVTATREGTTTSGRYCREFQQKITVGGKNENAYGTACQNQDGSW
EIVSTGS

>gi|54310376|ref|YP_131396.1| putative single-strand binding protein [Photobacterium profundum SS9]

MELSMASRGVNKVILVGNLGNDPEIRYMPSGSAVANITIATSESWRDKATGEQREKTEWHRVALFGKLAEVAGEYLRKGS
QVYIEGQLQTRKWQNQQGQDQYTTEVVVQGFNGVMQMLGGRQGG
GQQQQQQQQQGNWG
KPQQPAAAPQQSV
APQQQQQ
APQQPQQ
APQQPQQQYNEPPMDFDDDIPF

>gi|50841759|ref|YP_054986.1| hypothetical protein PPA0270 [Propionibacterium acnes KPA171202]

MSSPRAVPKAPISAWMVTSK
FELGKPRLV
PGPGKPRLV
SGPGKPRLV
SGPGKPRPV
PGPVRWK
PEPNKTRDAGNNEGRHGKQGKGSTLSKKPGREEESDGDHKSRETTTIRVAVNDQNRSQVEKWLGDNDVGPSNAVPLLVKA
RFMNPLDADATNDFAKFLYQSEDTQIDTQTSSETSDLKDDSYDILVAKVGEKDARESSSSTGQMGYLLENGKRVPKISEG
CK

>gi|50843015|ref|YP_056242.1| hypothetical protein PPA1537 [Propionibacterium acnes KPA171202]

MVARAPRLVGEARQAVAEELAGRYHDGASIRSLARESGRSYGLVQKLLREAGVEFRRRGGADPASPETKAERETVE
KSQEIDQQDSVDSNTNGGGTSPSPEDIDALKLAVEAAVA
RAEKADRKAH
KAEKALRKLRRAGAGKSRRKKAKATLNKRRA
KAEKADR
KVRKARR
RLDKAEHSIDNRQF

>gi|50843333|ref|YP_056560.1| hypothetical protein PPA1879 [Propionibacterium acnes KPA171202]

MKRTVLAGAVAACLGMAVFAVTTPDARAESAPIPVPTSKATAHSRLANRPSSGITWSPRHHGIVPPEIH
GGTIYLPDIS
AGPIILPTIK
GGTITLPDIK
GGTITLPDIS
GRPIILPKIE
GGAVSIPPIVFPSLHARGSIQVNGTLPGWANGRGNFLAAGGMGAPA
GRPMSLPQTGA

>gi|50843620|ref|YP_056847.1| hypothetical protein PPA2183 [Propionibacterium acnes KPA171202]

MLPENVAGKSALRLFPRRNVCNDGDVRRPKEVVMSRFNDNEFVP
DDDVDDYDDLDDYDE
LSDDDDYG
DDGDDDYDD
DDYDDDL
EDASEDEIDFVVALYNDDGERVAVPMEPMLANDFDELITQLRRLPGDAGAVGMVSIDHQFYVIVRVRGRNVQVFLSDATQ
ANDWPIARDVADFLGEDI
PDPDDDPETMGDGDMLADAGLGEFDMEAF
ADADEDSDEILSDIAARVHFGPQFERAANASAR

>gi|27467370|ref|NP_764007.1| hypothetical protein SE0452 [Staphylococcus epidermidis ATCC 12228]

MDKKKLIKTLINVLPIVIVPLVTERKRIKE
HPDVQKVTNATSKATSKTTSAISNKASDVKEYVGDKKQEFD
NKRELKKIAKENDPAYIEKKGEKLAKQN
RKEADKMDKKLQKNIEKR
HKEEQKQ
REANQKQRIKDMKKTQKYEEQVGLTPGKID
HEIEKKGEKLEKDNRKDIKKLDKKLQKNIEKR
HKEEQKQREEAEKA
RKKEFKKYENYVADSVVKQ
HKESNHS

>gi|27467472|ref|NP_764109.1| hypothetical protein SE0554 [Staphylococcus epidermidis ATCC 12228]

MFKNEREVNNMSGISKFISAIVVLLLLIGLAFGIYSFVDSSKKGNERLSDK
TTQQEKKKDDKDKKEKKDKKSVEEKKN
NTQQTQQVP
QTQQTQQTQ
QTVQTPQRPTTQTPVR
QNPQTGEKRAKEPKTD
ERKKTRQDSSNSKESPNEKDQS
STQQKPQQNKSNQSSE
NRQSTQNKSN
QTQSTRPQNNGNSGNQNAS

SGQNTQKQNNNNQNSQSSKKD

>gi|27467708|ref|NP_764345.1| hypothetical protein SE0790 [Staphylococcus epidermidis ATCC 12228]

MNFKKTVAIVLTSAVLLAGCTI
DKKEIKKYDDQVQKAMD
QEKTVNQVSK
KINELEEKKQKLFKKVNDKDQST
RKKAAEDIVENVEQ
RQKEFEKEEKALDN
SEKEFKQAKQYLEHVENKA
KKKEVEQLDSAIKEKYKSHDAYAKAYKKALN
KEKELFSYLNEDNATQSEVDG
KSKDLSKAYKEMNNKFNAYSKAIEKVKR
EKQDVDQLK

>gi|27469139|ref|NP_765776.1| hypothetical protein SE2221 [Staphylococcus epidermidis ATCC 12228]

MKKFLATIVVIILICALGFSLYAAINHHTN
SNKNDKKIEHNR
KNKQNDT
SNQNRDTNQHQDQTQPTNNDY
NNDNQSGTEQPAQQPNYHQY
PNNNQQSG
SNKNNSSENNKQKPN
QNKTNQSYHQPAQSTPQQS
SQHNNQSDSQQN
GNSNNNS
NNQNHGT
NDKQNKNR

>gi|27469243|ref|NP_765880.1| hypothetical protein SE2325 [Staphylococcus epidermidis ATCC 12228]

MTKQDLSLSVFTNE
NYKNLRYTSS
SFRNSMYDELEVNKS
RFKNCNFNEGIFKNIEAICNC
KFTTCGFNNCIFEDVHFYKN
QFKDSTFVNTPFDQSVFNST
LFQNAMFDSNLIRSVKWTDIIFKNV
SFKNVEIEGTTFKDV
KFKNCEFKNVIITNSTMSQKLMNELQKQDVTLENIDTSI

>gi|27469260|ref|NP_765897.1| hypothetical protein SE2342 [Staphylococcus epidermidis ATCC 12228]

MKRLLGTLIAATLVLSACS
QNDTKED
ENKKSEN
TTEKKSDDKKDKK

TNEDKKSGEQKK
SQEKKNNKSMQESA
TNEQVQSQQQ
TNNNQQQYSDEYLANNSVSNDDQLSSEEQQKYNEN
TNENKQNSHVDLNSLPATDFRTDWMSESGQKQVNDLANQKDSGQISQYEFNDRVSDVMNNEIDNQQP

>gi|51891313|ref|YP_074004.1| hypothetical protein STH175 [Symbiobacterium thermophilum IAM 14863]

MAAVQEDRLMVLQMLADG
RLTVEEADTLLRSLEETAAHADAAGAGPARRADSAADRHEPGARLRVLL
QQMLDEANAALDRALRSVEERLAELERREPYRQIRQLADSLDRVAA
REAVREARAAA
KEAVREARAAA
KEAVEEA
LEEVEEKV
ADAIETA
KEAVADAIEAA
EEAVADALEAA
EEAVADA
AHAAEEAEENVRFRASGRFGS

>gi|51891316|ref|YP_074007.1| hypothetical protein, glycine-rich [Symbiobacterium thermophilum IAM 14863]

MRWEVCSMPQWLMRAVEPIVAHGLREMRTEGADREHILREVALMSALVGTGMPAVQAIRLVEAMEPQLLGMPRGEAREPE
H
AGYPDMGGWGMGFPGMGFPG
MGYPGMGYPG
MGYPGMGYPG
MGYPGMGG
WGKDGFGKGMMMGPWGKMPGTGP
MGWPQGGQNYMQGTQPQ

>gi|51891365|ref|YP_074056.1| hypothetical protein, proline-rich [Symbiobacterium thermophilum IAM 14863]

MRRPWLRAAALILSTALILSGCGSAD
NPPPSPP
PQTEPKP
TPNPTPPET
PETPSPPADA
PEDPAEPE
PDVPALPGDTEALREALRVQSPITYTIVLAHDAPREMDKTAYLDQMLAEQGYPGKNEILLVLFPADNYNIRFAMGSLVFD
RRITLQQMLELVQSQYLTRSRQGDPAGGLAALINAINERAK

>gi|51891394|ref|YP_074085.1| hypothetical protein STH256 [Symbiobacterium thermophilum IAM 14863]

MPASRREAGEGA
GVGCSGSKGVPPPHPHPLAI
GRRCHGGWPYMPASRREAGEGA
GVGCSGSKGVPPPHPHPLAI

DRRCHGGWPYMPASRREAGEGA
GVGCSGSKGVPPPHPHPLAI
GRRCHGGWPYMPASRREAGEGA
GVGCSGSKGVPPPHPHPLAI
GRRCHGGWPYMPASRREAGEGA
GVGCSGSKGVPPPHPHPLAI
DRRCHGGWPYMPASRREAGEGA
GVGCSGSKGVPPPAGNRQAVP

>gi|51891456|ref|YP_074147.1| hypothetical protein STH318 [Symbiobacterium thermophilum IAM 14863]

MGRTFRGRPRFQ
SSPGLAAGCYPRPSAALWSGKTFQ
SSPGLAAGCYLTRGQGPSW
WRPGFNPHPA
SRPGATLCAPWGAPTAPRSFNPHPA
SRPGATPHA
SRGGRTARRFQ
SSPGLAAGCYRQVPVLLGPRIPVSILTRLRGRVLPSSSRRPSPATGSFNPHPA
SRPGATRTGQAIVMWDDEVSILTRPRGRVLRPIP
SRSHARPRFQ
SSPGLAAGCYD

>gi|51891583|ref|YP_074274.1| hypothetical protein STH445 [Symbiobacterium thermophilum IAM 14863]

MSVHGVVKRGRGSARVLRMGLTMLLIIGLVSGVPAHGSKLADVGAVEGNPVGQ
PDEESPSD
PGGEDPDSVGQ
SDEEFPSDE
GDEDSGSVGQ
PDEESPSG
PGGEDPDSVGQ
PDEEFPSGGGDEDSGFVGQ
PDEESPSDPGGEGSDPGDVDPG
DDEDGPAFPVAPGTEDDGGTDPAQEGLFAALAWCCPGPRWSPTTAAPPP

>gi|51891635|ref|YP_074326.1| hypothetical protein, glycine-rich [Symbiobacterium thermophilum IAM 14863]

MNWD
DMGTMPY
GGTAMPM
GMDGMPY
GKWNMPY
GKGGMPF
GKGGMPF
GKGGMPF
GKGGMPFDKGMPF
GKGGMPF
GKGGFGKMP
GWGGVPGGGFAKQYPP
GMVGIPYGFYPGYGLVQ

>gi|51891922|ref|YP_074613.1| hypothetical protein, glycine-rich [Symbiobacterium thermophilum IAM 14863]

MDRQVAAGTVRLRWRDANVAQSGDLMARLLNLVQEEKRDMESLYAEEPVRAHLLQATGDDGSG
DGSGSGDGSGGGSG
DGSGSGDGSGSG
DGSGADDADAAM
TGGGSGGDEA
DADGSGA
DAGSGGEDGSGDEGVDATMTDDGSGA
DAGSGGEDGSGDEGVDATM
TGGGSGG
DADGADD
DSFGSGM

>gi|51891929|ref|YP_074620.1| hypothetical protein STH791 [Symbiobacterium thermophilum IAM 14863]

MAGSV
AGRIGGRWAGNDLRVVYADGEL
TGRMGGKWVGKDVALSFDGAEP
TGRWGGRYDGQDLRFSIRGTEG
TGRMGGDVVGWDIALVQDGDRI
TGRIGGAVVGADVRLAARDGLL
TGRIGGNFDGKNCTLQVDGVPYLVAAALAAMVFYQLYLDAASTAAST

>gi|51891930|ref|YP_074621.1| hypothetical protein STH792 [Symbiobacterium thermophilum IAM 14863]

MGNDIRVSCE
GGRLSGRF
GGAVIGKDVDLTFD
GAVLRGRW
GGVWNGKDLHFTIRGTEG
RGRMGGSVVGWDIALVQD
GDRIAGRI
GGAVSGADLRLTAG
DGRLTGRI
GGSVFGKDCDLQVEGVPYLVAAALAAMVYYQMHLDSSNNASSSSGGA

>gi|51892003|ref|YP_074694.1| hypothetical protein STH865 [Symbiobacterium thermophilum IAM 14863]

MSKPMMPYGRPGWPVMPQMTPG
MMPGMMPG
MMPGMMPG
MMPGMMPGITPG
MVPGMTPG
TTPGMMPPMTSPQMM
MPPGAPPELVYGGQVMIPQEQSYVENILRMNRGKTATVYATYDNNPEWAARVFRGQVENAARDHVVLSDPQTGMRYVILM
VNIDYLTFDEPLRYPEEIYRPPGTF

>gi|51892328|ref|YP_075019.1| acetyl-CoA carboxylase, biotin carboxyl carrier protein [Symbiobacterium thermophilum IAM 14863]

MDVEEL
TPAAGAGFSPAAP
GPAGIAAAA
GPSPAAS
AAAPAAPPVM
AAPPAAV
APAPAAAPASAAAS
APAPAAQPAASGGA
APAEGKAVC
APLPGMIVSVPVQVGQQVREGDVVAVLEAMKMENDITSPYAGTVKQVLVSKGTSVSMNDPLVVIG

>gi|51892777|ref|YP_075468.1| hypothetical protein, proline- and glutamine-rich [Symbiobacterium thermophilum IAM 14863]

MRRLARSFRSA
GGPQQNPAGS
QAPAQQPGGPQDAPPQGRDQAYLEP
QGAQQPPWLPSLNGETAASMLTALRAAKQQLVQELEANLKQLKAVLEETHRIQKQMEAVLMEARAEIGLQQRAGG
HGQDQQP
QRQQQPQARQPQA
WGQMQQPQQLQQP
QAQGQPPA
QGHLQQPQG
HGEPQPPQAPG
HGRPQQPQAPSAGATGGP
RGAQPPSPGAPGPRRPRPPLSSDPPPPWEPPEPASEPG
WSPWQPPV

>gi|51893024|ref|YP_075715.1| hypothetical protein, glycine-rich [Symbiobacterium thermophilum IAM 14863]

MRGSAGAEGGHAQNLH
PGSGRRFRFCAGA
PGQRGGHAQNLH
PGSGRRFRFCAGA
PGQKGGHAQNLH
PGSGRRFRFCAGA
PGQKGRHAQNLH
PGSGGRFRFCAGA
PGQKGRHAQNLH
PDSGRRFRFCA

>gi|51893105|ref|YP_075796.1| hypothetical protein, glycine-rich [Symbiobacterium thermophilum IAM 14863]

MGDGACCGRNVQ
WREGDGCTFRLEL
GGRGRGCARNEQ

WREGDGCTFRLEL
GGRGRGCVRTVQ
WGEGDGCTFRLEL
GGRGRGCARNVQ
WGEGDGCTFRLEL
GGRGRGCVRNVR
WREGDGCTFRLELRGRGRRYGQNVQWWASAGAHLVGCTFRRGHDPREDR

>gi|51893340|ref|YP_076031.1| hypothetical protein STH2202 [Symbiobacterium thermophilum IAM 14863]

MHGLTTGASRRHRLLPDRETLRDAGIELLSYVGLVGDQGVLGFPGLQWPTAWRREVERAVGFRGFPPEGKYLATQRLLSY
PLWVLVSLVVVATGLLKAMR
YAYPLPAAVVRWSTTLHDWAFIATIVLLAIHVAAVTVVRSNWPLLRSMFTGTVSLEHVRAVHGRWYLELLREMEANT
AAGSAPT
EAASAPA
AAGSIPA
TAGPTPA
SAGPAPA
GAGDVPRPDHTDRGGKEAPR

>gi|51893342|ref|YP_076033.1| hypothetical protein STH2204 [Symbiobacterium thermophilum IAM 14863]

MLAAGESRRMGRPKMLLPWDGGTVLEAVVRSLLAAPVRRVAVVLGHRASEHRAALRQLARESRVLVVENPRYREGMLTSA
QCGVAVLLRKIPWGMAALS
REETRVATALL
PEETRDGTAPL
REQTRNGAVLL
REEPRDGAEPR
REESRNGTTLL
REETRNRAVLL
REETQDGQALL
REETWDGAGLL
REETRDGAALF
REETRDGAGLL
REETQDGQALL
REETRDGAGLRREQA

>gi|51893793|ref|YP_076484.1| hypothetical protein STH2655 [Symbiobacterium thermophilum IAM 14863]

MSL
AYGAGRFVACADGGLSASTDGLTWTPVPLDARYWHIDLSCTGKEFVATGCITSRTCEVILTS
ADGVSRYQRVGLAAFDQV
AYGNGRYVL
SFGNAAWTWSSADLTNWTTVNTLLFLALTFDQGMFVAADADEGTIRTSADGVEWTEVFHPDQPVYISDL
AYGGGRFVAVGGTNSGPRRAVVL
TSGDGRQWSLQVQEEAEGLVRV
VYGAGRFVANSEAALAVSEDGTGCRVQVLRGDEVRARRI

>gi|51894247|ref|YP_076938.1| conserved domain protein , histidine- and aspartate-rich [Symbiobacterium thermophilum IAM 14863]

MSRAGCRSYAGGREGGERMSWRD
DHDRRSGLCADLEDDRDCGCGGHGWNG
RHDQDGGDRWDGDHGHG
HHDHRGDHD
HHGHHDH
HHNHHGGHDG
HHHRHRRRQRRLIPIGSVTVDCTPVTFRTPGQAINVVSCPVSWIIPGTRCIATIHTPPVQTPVLPIHRHRKACGCH

>gi|50083581|ref|YP_045091.1| hypothetical protein; putative signal peptide [Acinetobacter sp. ADP1]

MKKTISLIATALVSTVLLAGCQNMSPSDQRIGAAALGG
AVGGGVGN
HVGGGLGAGVGA
AAGAGVGSN
AKGGSSNTTTSS
AVGAGIGSVVGKA
IFGGDSGAAIGG
AIGGGAGAAIQEKR

>gi|50083733|ref|YP_045243.1| conserved hypothetical protein; putative membrane protein [Acinetobacter sp. ADP1]

MKTVGNDLIRNELHHQRKWYLLF
GILMIIF
GVLMLGSLAMATLSVILLF
GGLMMIG
GLIHLVAAFKLFQGGTRWLWALF
GVLYLLAGYYAFKTPITTAIVLTNLL
AIFLLIAGVFRTFNAFILKPIAGWGWTLFS
GILTFVT
GVLILISPDAPFWVLGLFLAVDLLFQGVNYLSLASAIKHLPHSSTTST

>gi|50083986|ref|YP_045496.1| hypothetical protein ACIAD0770 [Acinetobacter sp. ADP1]

MYHLCCKLHFLNYPLAIRCIKEGKFIAIDHDQALQKFNFNRNDNEDSDMTNQNKL
DELKDKAADLG
DELKHKAKEAKLEGEKAI
DDLKHKAQEAKLEGEKKL
DDLKKDAEKEHSEQKAAFDEKV
DDAKTKW
DSAKHDAQDKF
DHLKTEASHKL
EEAKLKA
QELKHDASEKL
DELKAQATAKF
DELKKTATDTLNKFKNNDKG

>gi|50085463|ref|YP_046973.1| host factor I for bacteriophage Q beta replication, a growth-related protein [Acinetobacter sp. ADP1]

MSKGQTLQDPFLNSLRKERIPVSIFLVNGIKLQGHIESFDQYVVLLKNTVSQMVYKHAISTVVPARNPRPAGAPTGAPQG
GGFGGQG
GGFGGQG
GGFGGQG
GGFGGQG
GGFGGQG
TGFGGQG
GGFGGQGGFGGQ
GGFGGQGGFGGQ
GGFGGQGGSGFDNDTKFEDGQEDDSNNR

>gi|50085471|ref|YP_046981.1| putative membrane protein [Acinetobacter sp. ADP1]

MLMSQFDVNLLVLLVLLACGIFSQNAAVT
IAAGVLIVFKLTPLSEFFPFLQQHG
LNVGVII
LTIGVLTP
IASGKLSGETILKSFFSYKSILA
ITVGLLVAWLGGRGVKLMSSQPD
IVAGLLIG
TVAGVALLRGVPVGPL
IAAGILSLVLWN

>gi|50085565|ref|YP_047075.1| fragment of Orfc273-4 from Vibrio metschnikovii [Acinetobacter sp. ADP1]

MNVIPSLFL
NHLCDDEPNILTQSATVPFL
NHLCDDEPHIWGAPNGCTFL
NHLCDDERAEALSVNAGNFL
NHLCDDEQ

>gi|50085568|ref|YP_047078.1| fragment of hypothetical protein PM0309 and/or PM0310 of Pasteurella multocida [Acinetobacter sp. ADP1]

MLGLKH
VRHRIDDLEKKKCSQLHTVT
VRHRIDDLETGGSVPFSGAN
VRHRIDDLEKNHIIRLGKCV
VRHRIDDLENENKPQYCQCS
VRHRIDDLESFRVIRVQLRH
VRHRIDDLEIETYELCVIRV
VRHRIDDLETYQNRWDLCRT
VRHRIDDLERIML

>gi|50085571|ref|YP_047081.1| fragment of Orfc273-4 from Vibrio metschnikovii [Acinetobacter sp. ADP1]

MGAGIFFL
NHLCDDEQPMHYSNIRRAFL
NHLCDDERFASILRSMMIFL
NHLCDDEQMIF
LNVCFDEFL

NHLCDDEQKQILRDAAKNFL
NHLCDDELYSIFLPTKSSFL
NHLCDDELK

>gi|50085573|ref|YP_047083.1| fragment of Orfc273-4 from Vibrio metschnikovii [Acinetobacter sp. ADP1]

MVKGRSAAQQFL
NHLCDDELTVEIIIPISVFL
NHLCDDERGGRSKLPFALFL
NHLCDDEHLGRTATTPNHFL
NHLCDDELT

>gi|50085577|ref|YP_047087.1| fragment of Orfc273-4 from Vibrio metschnikovii [Acinetobacter sp. ADP1]

MVCILCCCCNYFL
NHLCDDEPHSLTVDAPLIFL
NHLCDDERVGEAYDYDSTFL
NHLCDDEHVASLKKFELAFL
NHLCDDERQFCCMHQGMLFL
NHLCDDELLTTHSSVTPVFL
NHLCDDELFGQCDFCRDPFL
NHLCDDELLRVNGTFYGIFL
NHLCDDEQKDWWDKLGADFL
NHLCDDEHVATAMKAITDFL
NHLCDDEPFCASKVGLGFFL
NHLCDDEHKATR

>gi|50085578|ref|YP_047088.1| fragment of hypothetical protein PM0309 and/or PM0310 of Pasteurella multocida [Acinetobacter sp. ADP1]

MRFVTNA
VRHRIDDLEINQDRITINSK
VRHRIDDLEKQALLSYVSLK
VRHRIDDLEKACLDAYNKIA
VRHRIDDLETNVIRISRLQR
VRHRIDDLEKIACA

>gi|50085579|ref|YP_047089.1| fragment of Orfc273-4 from Vibrio metschnikovii [Acinetobacter sp. ADP1]

MCFL
NHLCDDERFTNSLNSISIFL
NHLCDDELTDVAVVPAPSFL
NHLCDDECKSLIGQDQVHFL
NHLCDDEHGGS

>gi|50085640|ref|YP_047150.1| hypothetical protein ACIAD2564 [Acinetobacter sp. ADP1]

MQVKKLLIGLAFIVFASLIGVMIAIWVFKHF

NIYIPLENQAVSI
DLQEPLQAKV
KIQDALDVDVSGRV
NATIPIRE
ELNIPLTQTLTPRVYF
DNQVPIKTVIPVQETLKV
DQNLPIDTKVQVKILGK
DITLPLKG
TIPIKLDVPL
NIQVPLDQSVHLKFDAPVKTVLKE
NLHIPLDATLKTNIPIQG
HLNVPIKTALAASVDVQNTLPIKIKQG
ELKIPLSSITLQKDQASSNNVAQTPPASKTE

>gi|50085803|ref|YP_047313.1| conserved hypothetical protein; putative signal peptide [Acinetobacter sp. ADP1]

MQIIVTIIMRKRMFPITLIIS
LSIGLISFQTYAEPTLSDIKQMAAQGKSTAQ
YHLGMMYLSGEQGVTKDTTQALKWLTLADQNGSVGAK
YSLGLMYMTGTGVSQNQSTAFEWFSKAAKFGHAQAQ
YTVGRMYSEGVGVEKNMPQAFEWIQKAALQGYPPAE
FSLGLMYNDGRGVAQNKQQAIKWYTQAAEHQQSNAQ
YNLGIMYLNGEGTSKNPPLAKKWLQRAANAGDQDAKDVLKGLK

>gi|50086052|ref|YP_047562.1| hypothetical protein; putative signal peptide [Acinetobacter sp. ADP1]

MLFKKIAIAAAIATTLAFVGCAKKTEEAA
ADANAAA
SQATAAA
SEAQAAVDA
SAASAPASAAVAA
TAATDAA
AAATDAA
HAASAAASK

>gi|50086268|ref|YP_047778.1| hypothetical protein; putative signal peptide [Acinetobacter sp. ADP1]

MDLQKDRKMKMMKSAILTASILA
SATVFAG
NASVDVGA
NAQVESGGLLNRIGSGVKNGVTDTAHAVHSTAVKAT
DATVDAGSHVAHKTH
DFAVDAK
DKTVNVAEKVGDKSEQAWQDTKSGTKKAVNGTEKFAAKEKAKAESHWNTKKKQSEAKHHAKAKTGV
NANTQAGIKVGSVNA
NAGVDANVSPRNVGVKTQVGVGQ

>gi|30264447|ref|NP_846824.1| hypothetical protein BA4605 [Bacillus anthracis str. Ames]

MAGGSRFQ
QKQQKRRQNGVLNIAIAIVLVAVAIVAYQLFVPDTKEQASSSDKKVSQQTTK
ENKSEKAKG
KEETKKNEQEKTEAKKK
EEEKLKA
EEEQKKA
EEEAKANEKVPA
EKTQPKATDAYTKSSWKPVGTEQGATPAMTFKKGTADWNEMNQAISAAIDVPVEQLIIHRIGNNGKNKAYGNVQDKQSGK
RYYVNIDWVENEGWKPVLVQTLN

>gi|30264593|ref|NP_846970.1| hypothetical protein BA4764 [Bacillus anthracis str. Ames]

MSSWRNNINGYCSCNNQNGVHVDSCCFSCDGTVPKLGPTGPTGATGA
TGATGVTGVTGV
TGATGITGATGI
TGATGITGVTGP
TGATGITGATGP
TGATGITGA
TGPTGVTGTSITATYAFANN
TSGTAISVLLGGTNVPLPNNQNIGPGITVSGGNTVFTVANAGNYYISYTINITASLLVSSRITINGAPLAGTINSPAVAT
TSFSATIITTLAAGSAISLQLFGLLAVATLSTTTPGAVLTIIRLS

>gi|30264806|ref|NP_847183.1| vrrB protein [Bacillus anthracis str. Ames]

MDNNA
PHGFFGGGSDSYEQMMLMGGAGQQGYGGIPSWMGGSPGGFSTSVAGVQTGMPTSMGGFSTSGAGMPTVTGGFPTSVAGVQ
TGFPVPGVGVVAGGIGGFPQVIHG
HHGHHHHQGHQG
HHGHQGHHHHQG
HHGHHGHHHHQG
HHGHHGHHQQ
QVHHHGHHHIHPQAILYQT
HQGHQGHHD
HHGHHGHQG
HHGQHSQQ
HQQHQQYQQYQQQSSPWAGGIGAGAAGAAAGAASISGAAG
HAGHVGH

>gi|30265363|ref|NP_847740.1| DNA-directed RNA polymerase, delta subunit [Bacillus anthracis str. Ames]

MDFKQYSPEELKECSMIEVVHSVLGDKRQATTFNELVQEIAQVLGLSQEQVNAKIAQFYTDLNIDGRFINLGENRWGLRS
WYP
YEQIDEEILPQPKPKKKRKVEDDG
FDDYIEEDEDFDDADV
TEDEDDD
VEDLDKV
LEDEDGDDDD
LDDLDED
EDDFAEEELE
YDETEEEEEEL

>gi|30261652|ref|NP_844029.1| hypothetical protein BA1580 [Bacillus anthracis str. Ames]

MFQQSNVYQQS
NPYAQQNMYQYNTDTYLRYNMYPF
EPHYGNQNYYQPFE
VSFMNQQQQPQ
QPYMNQQP
QPYMNSQYYMPPP
SPYGNQQAMFYPPK
QPYPTQNKQKQQQQPSQFSSFVSQFKNSDGNYDVNKMMNTAGQMMNAMNQVTGIVKQVGGFFGK

>gi|30262056|ref|NP_844433.1| hypothetical protein BA2029 [Bacillus anthracis str. Ames]

MKRRIGTAVVGLSVLGFG
IFGFTHGNGGGVVQAASEGDGGG
APAYAHGNHGG
APTYTHGEGW
SPSYAHGHTGG
APSYTHGHTG
APVYGYGDYP
IPSFNVGDTPAPRLDADGDHGG
APTYTHGHFPAPKLE
SDGYTGG
APSYNKEDTW
SPSYAHGNTG
APSNTEPGGGI

>gi|30262070|ref|NP_844447.1| hypothetical protein BA2045 [Bacillus anthracis str. Ames]

MKRDIRKAVEEIKSAGMEDFLHQDPSTFDCDDDCTTKIECSDDCNC
PRTRCTRVKHCT
FVTKCTHVKKWT
FVTKCTRVRVQKWT
FVTKVTRRKECV
LVTKRTRRKHCT
FVTKCVRFEKKFYWTKRCYCKKCEFFPHGHGGSCDDSCDHGKDCHDDGHKWNDCKGGHKFPSCKNKKFDHFWYKKRNC

>gi|30262180|ref|NP_844557.1| hypothetical protein BA2162 [Bacillus anthracis str. Ames]

MFYNNQ
PPYPQQPFYPQQQEQYEEQELQQQE
QQYEQNP
YATPQNQE
LQYPQNP
YTAPQTQEQQFQQNSYDTRPS
YEYPQNP
YAAPQNQE
LQYPQNP
YVTPQTQE

QQFQQNPYPTQPQQTQYQQQMYQPNYDARVSPPKPPTFDITQPQILPPGPTLDITQPQILPPGPITEPTQQQIQQVVGTQ
FLPLKKPVLDFVKPWVDYGLNEAKHTSHKHALTEVAAIMFLVGKGFNPTIAHYIVESWEKNEQF

>gi|30262210|ref|NP_844587.1| hypothetical protein BA2195 [Bacillus anthracis str. Ames]

MQVDPN
TPVDPKP
PVVEPKP
PVVDPKP
PVVDPKPPEPTH
PTITVSPKTGDILVEWYAGNGKYYSSVFFNDFYTSQKELDENALNNLTNYYSKKEDWTLVSQLNTPSKSKSVAEHKLALD
NVEQILTSGEYSNERALFYLRNQMNVLNESVTTQEELQRYLSENKKDYIKSVAEKVYQTATEEKREWSNTGDVASLEQAL
SKFYLVTSVFESVDPALV

>gi|30262336|ref|NP_844713.1| hypothetical protein BA2328 [Bacillus anthracis str. Ames]

MENQLKEIFGALIAAIGTITSAIGSTPFYFISSNVREDLNIY
GNTLQAV
GNALEADGQE
GISLEKI
GNEIQSTGNVTVISGLVIDFKDETKIKLVIA
GNWTQALGGLTALADEFEDASDKDESFNII
GNLLQAI
GNSLQAIGGIYELKSIRKERLDSKDELVNDTEGNLDNQVNSELDKKKE
GQSIDTI
GSWIQAVGSVFSLIGQIREESEELEGSDN

>gi|30262659|ref|NP_845036.1| hypothetical protein BA2677 [Bacillus anthracis str. Ames]

MNQFQQELQSLNLNDYQTGNVVYWDPQQSQYPYYYIQDDARR
CGGCGGC
GGRCGGC
GGRCGGC
AGRCGGCIG
CAGCFGCFNCWNWWII

>gi|30262843|ref|NP_845220.1| hypothetical protein BA2879 [Bacillus anthracis str. Ames]

MRTENLIIN
GYGSSNGGEFHKVQLN
GKGTVNGNVECEKFECN
GYGAVTGDLKSSSARIS
GSGKVDGTVHAKTMRID
GKGTITKDVKANSLKIA
GKGTIGGNVTGEEFKIN
GQATIDGNCEVDTFSSD
GQFEVGGLLSADEININIHGTCRAKEIGGQTIKVRHRLGTFSRLFKSVFGLQLEAELLEGDNIDIDYAHIRTVRGNNVTV
GPNCEIELIEYTGVLTVDKSANVKEIKQV

>gi|30263292|ref|NP_845669.1| lipoprotein, putative [Bacillus anthracis str. Ames]

MKMKKVAIASITIGTMLTMAACSPSAD
KETDKKEQVTAQNEGEVKT
TSTDETS
KTTKDTDKETNTSSNEKTDKDA
KNTKESSGTESNG
KTSTQNENV
KTKEKETTG
KTNDQTTSG
KTTDQSTSTNSDGKTTKDPKKSVAVKTSVKEVAALVDSLDKQLAANPKLETVNKLGKQINAKWDVIEKELETSHPAEY
KTIGQSMYPLIVGAEKEKIDITKMKSLTTKTKKDLNQLLTKLS

>gi|30019706|ref|NP_831337.1| Spore coat protein [Bacillus cereus ATCC 14579]

MF
QQPNLYQQ
TNPYAQQNMYQYNTDTYLRYNMYPFEPYYGNQNYYQPF
EVSFMNQAQQPQQQ
QQPYVNQAQQPQQ
QQPYVNQAQQSQA
QQPYMNQPQ
QQPYMNSQYYMPP
PSPYANQQAMFYPP
KQPYPTQGKQKQQQPSQFSSFVSQFKTSDGNYDVNKMMNTAGQMMNAMNQVTGIVKQVGGFFGK

>gi|30019793|ref|NP_831424.1| hypothetical protein BC1646 [Bacillus cereus ATCC 14579]

MSLFKNRIPKNSVSFSEETYELQFPKPASV
QIEEEELQADHA
ELRAQQESLHMEMN
QLRQEQQ
MLERERQ
QLLQDKEQFQMHIHEQMEQMEAARV
QFQKEQQEIAYEWTELLWDQSFHLAEKIVNQAVDSRLLDVLPILTGIVQTLPTSFEKLIITVHPETFE
RIQEEKENTKEYWLLQLVEWKYDFSLQFGEFVLEEEKEFFEFKFAPIFA
KLRQKWEEE
KLFEEQNV

>gi|30020167|ref|NP_831798.1| Spore coat protein G [Bacillus cereus ATCC 14579]

MKRDIRKAVEEIKSAGMEDFLHQDPSTFECDDDKFTHHH
CTTGCKCTTRGKCPRTRCTRVKH
CTFVTKCTHVKK
WTFVTKCTRVRVQK
WTFVTKVTRRKE
CVLVTKRTRRKH
CTFITKCIRFEKKFFWTK
RSFCKKCEFFPHGHGGSCDDSCDHGKDCHDSGHKWNDCKGGHKFPSCKNKKFDHFWYKKRNC

>gi|30020620|ref|NP_832251.1| hypothetical protein BC2492 [Bacillus cereus ATCC 14579]

MI
AGVLRKSIAVMI
GGNLKKSIIVMM
IGVLRKSIVMMI
GGNLKKSVIVMT
IGVLRKNIVMMI
GGNLKKSVIVMI
IGVLNEISVVVNVSLFVAIVKEKIGNIFNVNVKERICLLEP
IGLFRKVDPSPTSLLLLNF

>gi|30020697|ref|NP_832328.1| Collagen triple helix repeat protein [Bacillus cereus ATCC 14579]

MILIEYKNRDIFNTFIFSDVKPIIIPHHP
TGPTGVTGRTGI
TGPTGVTGRTGI
TGPTGVTGPTGVTGKTGV
TGPTGDTGSTGI
TGPTGVTGPAGVGLTNYLYIFDTTNQSIVVGGSVTFNTNGPIIGSAFTHIAGTGNITINTRGTYVAEFQLEAARENQFSF
ALNGTPIPGGRYGTGSPHSINQGTVAFTVT
TVPSTLTLINNTSSAGTITLSNNDGGSLTNVSASISIFQVG

>gi|30020811|ref|NP_832442.1| hypothetical protein BC2687 [Bacillus cereus ATCC 14579]

MEGFIMNQFQQELQSLNLNDYQTGNVVYWDQQQSQYPYYYIQDD
ARRCGGCGGC
GGRCGGC
GGRCGGC
AGRCGGCIRCAGCFSCFNCWNWWII

>gi|30020991|ref|NP_832622.1| Spore coat protein X [Bacillus cereus ATCC 14579]

MCEC
EVCRRKKNRDPLN
REFEPKK
RVPEPKK
EDCEPKK
EDCELKK
EDCDPKK
EDCESLNDYSFSKNANFLEEAIQTDEIDQISEEYIEIVDSADVQVTTTDTKAALSIQAALQAAIVVVVSISIADSEKADK
ITQELFQKSAIKQINRQKTFIKNSRNVTVTTTDTDIAVNIQILLQILLALLVKLNIL

>gi|30020999|ref|NP_832630.1| hypothetical Cytosolic Protein [Bacillus cereus ATCC 14579]

MRTDNLIIN
GYGSSNGGEFHKVQLN
GKGTVNGNVECEQFECN
GYGAVTGDLKSSSARIS

GSGKVDGTVHAETMRID
GKATITQNVKANSLKIA
GKGTIGGNVTGEEFKVN
GQATIDGNCEVDTFSSE
GQFTIGGLLSADEININIHGTCRAKEIGGQTIKVRHRSGTFSRLFKTVFGLQLEAELLEGDNIDIDYAHIRTVRGNNVTV
GANCEIELIEYTGVLTVDKNANVKEIKQV

>gi|30021296|ref|NP_832927.1| hypothetical protein BC3186 [Bacillus cereus ATCC 14579]

MGQV
SRFQEKQQNRRQKAIFNIAFTLVLVTVGIVTYQLFTPSNTSTK
AIAQEKKV
TKIEEKKTEDKET
AKIEEKKTEDKEA
TKLEEKKTKDKEI
AKIEEKKSEDKEA
AQLEGKQKVEKSVKTNGTE
NKEEEKTVSQEKDPQTNPSWKPIGTEQGAKPDMKFKEGTVDWSEMKKAISYAVDVPESQLIFDFIGNNGNNKAYGNVRDK
QSNKKYKVNIDWVENQGWKPASVQVVK

>gi|30021434|ref|NP_833065.1| hypothetical protein BC3326 [Bacillus cereus ATCC 14579]

MKMKKVAIASITIGTMLTMAACSPSAD
KETEKKEQVTAQ
NDGEAKTTSADEASKTT
KDTEKETNTSSN
EKTDKDA
KNTKESSGTESNGKTSTQNENV
KTKEKETTGKTNDQTTSG
KTTDQSTSTSTNNNGKTT
KDPKKSVAVKTNVKEVVALIDSLDKQLAANPKLETVNKLGKQINEKWDVIEKELEASHPTDSKTIGQSMYPLIVGAEKEK
IDINKMKSLTMKTKKDLNQLLTKLS

>gi|30022366|ref|NP_833997.1| VrrA protein [Bacillus cereus ATCC 14579]

MYPNPGQQPYTPYPIPQLPPMAQKKKGFLAKLFKKHDPTE
PFMQMVP
PYRQMEGQ
PMMHQQPPPP
QYHQQYPQQYQQ
QYPQQHPQYQ
PYMQQHPEQMIPPQMYESNETRGGSTTTAATASSGIGSFFSNLISNPTNMINNIEKVSQVVQSVG
PVVEQYG
PIMRSLPSIVKILTSGKSTEENQTENVTEPVEVESPPPAPQKKKRKRKKIVIEPVIEKELPKEPVQQTATKPKLYV

>gi|30022452|ref|NP_834083.1| hypothetical protein BC4371 [Bacillus cereus ATCC 14579]

MAGGSRFQQ
KQQKRRQNGVLNIAIAIVLLLVAIVAYQLFVPSTKEQASSTDKKVSQQTTKNEKAGKSEDK
ESAKKKEKEQEQA

EAKKKAE
EEKLKAE
EEQKKAE
EEAKANEKVSAEKTVSKAKEAYTKPSWKAVGTEQGSSPAMTFKQGSSDWNEMQSAIAAGVEIPKEQLVFHFVGRDKTGKS
KAYGDVQDKQSGKRYYVNIDWVDNEGWKPVLVQTLN

>gi|30023055|ref|NP_834686.1| hypothetical exported repetitive protein [Bacillus cereus ATCC 14579]

MKKKTVGYLAIAGALSFGIIGGVGIPAFAATNTPAV
DQPAKTAKKDL
DDATKQKVKTIMDDTKKQLEELGVKL
PEKEKRKDMFAGL
DEQAKEKAKSILEQEKSGKLT
REQAKEELTKLGVKF
PEKGKHKEMFAGL
DEQAKEKAKSILEQEKSGKLT
HEQAKEELTKLGVKMPEKGKHGDMFC

>gi|30023370|ref|NP_835001.1| DNA-directed RNA polymerase delta chain [Bacillus cereus ATCC 14579]

MDFKQYSPEELKECSMIEVVHSVLGDKKQATTFNELVQEIAQVLGLSQEQVNAKIAQFYTDLNIDGRFINLGENRWGLRS
WYPYEQIDEEILPQPKPKKKRK
VEEDGFDDYIE
EDEDDFDDADVN
EDEDDDVEDLDKVL
EEEDGDD
DDLDDLD
EDDDDFA
EEELEYDET
EEEEEEEL

>gi|53711536|ref|YP_097528.1| hypothetical protein BF0245 [Bacteroides fragilis YCH46]

MKTIFRMLSVLLLTTGLLSSCIQIGEGIQPSKKLITRDYKVKEFNKID
AGTVGNIYYTQSTDGKTDLQIYGPDNIVALIQVAVKDNTLFLSIDK
SKKVRNFKKMKITITSPTLNGIS
FKGVGDVHIENGLTTDNLDIE
SKGVGNVDIQSLTCQKLNVQ
SMGVGDVKLEGTAQIAALH
SKGVGNIEAGNLRANAVEAS
SQGVGDITCNATESIDAA
VRGVGSIKYKGSPTIKSLS
KKGVGTIKNI

>gi|53712712|ref|YP_098704.1| hypothetical protein BF1419 [Bacteroides fragilis YCH46]

MCELGYFISSYRLSPYKFRFFPISNYRIHRSGMIFLFYFCCKPKKNIMKKILFILFTVQFILIPRITGSYGANSIGNNHG
HSVTNQGKKTIK
KDIFGDTVI

EDNHGNRKTIK
KDIFGDTVI
EDNHGNRKTIK
KDIFGDTVIEDNRGNRKTIK
KDIFGDTVIENNCGNRKTIK
KDIFGDTVIEDNRGNRKSIK
KDIFGNTVIENNKGYKKTIK
TDIFGNKIIEDNHGKKQIIK
KDIFGNVIIENY

>gi|53713708|ref|YP_099700.1| conserved hypothetical protein BatC [Bacteroides fragilis YCH46]

MYMSRNKYILFALLLSLSAGAFAQKAERDYIRKGNRLFKDSVFVDAEVNYRKALEANPKSTISMYNLGNTLSQQQKF
KDAMEQYVAATNIEKDKAKLGQIYHNMGVLFQSGKDYQKAVEAYKMSLRNNPKDDETRYNLALAQKLL
KDQQQNQQN
QDQNQDQN
KDDQQKQQDK
KDQNKQNDQN
KDQQQQQPPKSEKNDNEMSKENAEQLLNSVMQDEKGVQDKV
KKQQTLQGRRLEKDW

>gi|17988179|ref|NP_540813.1| Hypothetical Membrane Spanning Protein [Brucella melitensis 16M]

MNAGPCRNGWCRIGNGSGYGWISARYVRFASYAGGYAAQPPATVIVNNGGYDD
WGPALGWGLGIGWASSNWGPY
WGPGWGGYRGGPNYIHGCIGRNCQSG
WGPGRRG
WNPHWGRGPQRRAQ
WGRGWPQHVNRPFWGPGPA
MGPHFGG
FGHGFGGGYRGMHLGNMRSLHAGR

>gi|15616701|ref|NP_239913.1| flagellar hook-length control protein [Buchnera aphidicola str. APS (Acyrthosiphon pisum)]

MHESSNI
KKKNDFDFKNLKNKTLQKKINVFD
TIKNHIKIINNEEKIILPLNIKKNKKIVFDSKNINKINPMI
SKRHSCKILS
NKKEFIKSYIEPFMSSDTKNSIEWKKLISHKILLSIS
NKYNQAKIHLKPESLGSIYVSI
SMKNNIATLKFI
SEHNEVKTFLDSYMPFLYNS
LKKNGIKLKEFKISIS
SKKNQLKNYKNLFYKNIYYTDHIKKTLNWNEKKIDYIKDKTIDVYV

>gi|15834936|ref|NP_296695.1| Hc2 nucleoprotein [Chlamydia muridarum Nigg]

MLGVQKKRST
RKTAARKTVVRKPAA
KKTAAKKASV

RKVAAKKTVA
RKTVAKKAVAARKPAA
KKTAAKKAPV
RKVAAKKTVA
RKTVAKKAVAA
RKTVAKKSVAA
RKTAAKKAPV
RKVAAKKTVA
RKTVAKKAVAA
RKPAAKRTVSTKKTAVAAKAGVCMKKHKHTAACGRVAASGVKVCASSAKRRTHHNRSRTAHSWRQQLMKLVAK

>gi|15893371|ref|NP_346720.1| Ferredoxin [Clostridium acetobutylicum ATCC 824]

MGVATMVTDKY
QKCIDAC
NRCSQAC
YECFKACLNEPDVNARRTCISIL
FECAQMCQMSSALMSMDAQFAIDHC
KLCSVIC
DKCAQECSMFQDPHC
QKCANEC
RTCSNECRMMSGM

>gi|15895985|ref|NP_349334.1| Hypothetical protein, CF-30 family [Clostridium acetobutylicum ATCC 824]

MYYNYNDFDDYDYDHGYHGYYDHDYHDHDYDDYDYDHDHDHDYDHDHDYDHDHHREEEFDPMRYEDDVLDAEAECPFVSD
CPATRSQMT
PPGPGQM
GPGYGQM
GPGYGQM
GPGYGQM
GPGHGQMMPPGPPPTYVPQIQSASNGPVIKAVDSGAIRMCRYKYTYIWLNNRRSFWAYLTFVGPRSIAGYRWNGRRWIYF
GLDLRQIRSFVCR

>gi|29653705|ref|NP_819397.1| ankyrin repeat domain protein [Coxiella burnetii RSA 493]

MHYAARGGSKKAMAYAIRELGIDATSRNTNN
IDALHYAALCGS
VEAMEYMIDIVEINPNSVGGYN
ANVLHYAALSGS
IEAMEYAIKTLGIDKYSTDANG
KSILHYAAQSGA
IEAMRYVIDVLGMNAGDINNSG
VNVLHYAVLSGS
VKAIEYVINTLEIDLGSTTNNG
ANALHWAAKSGSMKAVRFIRKLSHDRNLQLNVCDFRGRSAFWYADHSNNASLIKRVLAMPFEELMLETSIEEIKPSEDCL
VM

>gi|29654665|ref|NP_820357.1| hypothetical protein CBU1368 [Coxiella burnetii RSA 493]

MKGFRRHANTNLHDPVPLIRDFFRRAKDQ
RPHPQPLTARDYATIREILEARRHR
RFNGQSVANLLLAIAYHHTQW
RLLPRSLAAQLWDAIAKNVE
RLNPQGIANTLWTLATMNVRRR
ELEVQGLSDRLLDAVYYNAE
QFNSQDIANTLWALAAMGMRWR
ELEEQGLSDRLLDAVRYDAE
RFNPQGIANTLWALVAMGMTWGELEAQELNDRLLDAVGSNAP
RFNSQDITNTL

>gi|29655316|ref|NP_821008.1| hypothetical protein CBU2035 [Coxiella burnetii RSA 493]

M
PPEEEPPVTA
PPEEEAPPETGL
PPEEEPPVTE
PPEEETPPETGLPPEEEQGPGRGEGRGQDKEDKEDKENQGKGKGKGENKEGDDDGYSFEKADSHELDPNGFSDHEAHGDT
HSEGASVQEYHTGEADVDISGDSSGGVDQDPSG

>gi|50119870|ref|YP_049037.1| TonB-like protein [Erwinia carotovora subsp. atroseptica SCRI1043]

MYPLYRSRHAISWLPLPVFATCLFFASQQPPLKVQQQYDETVMALTLAEPEPIPQPEPL
PEPKPVLQ
PKPEPEPIPVDEPDPIIEAPPV
IPPKPEVK
PKPKPEVKPKAETK
PKPTPAQAKP
TAPRPEVPIKRSAPTTNAAPSAPSVNVAALENSYAQALRAQLEQTKRYPTGRQASLER
PEGRVEVWLEVDRTGRVINSGINSKAPNMLLNRAAQASLQSIKQVRAFPADAFAGQSTKRFLATFDYQAQ

>gi|50122597|ref|YP_051764.1| hypothetical protein ECA3676 [Erwinia carotovora subsp. atroseptica SCRI1043]

MSRIVN
FLSCLYGSEPILGYAPRPYR
FLSCLCGSEQHAARFCPQCA
FLSCLCGSELSSADSSVTGY
FLSCLCGSELHSPRKGRRCD
FLSCLYGSEPLIDVSLSIIC
FLSCLYGSEPEPVLQVVMAT
FLSCLYGSEHA

>gi|15830379|ref|NP_309152.1| EspF-like protein [Escherichia coli O157:H7]

MAEH
IPPAPNWP
APPPPVQNEQSRPLPDVAQRLMQHLAEHG
IQPARNMAEH
IPPAPNWPAP
TPPVQNEQSRPLPDVAQRLMQHLAEHG

IQPARNMAEH
IPPAPNWPAP
TPPVQNEQSRPLPDVAQRLMQHLAEHG
IQPARNMAEH
IPPAPNWPAP
TPPVQNEQSRPLPDVAQRLMQHLAEHGINTSKRS

>gi|15830674|ref|NP_309447.1| CsgA [Escherichia coli O157:H7]

MKLLKVAAIAAIVFSGSALAGVVP
QYGGGGGN
HGGGGNNSGPNSELNIY
QYGGGNSALALQADARNSDLTIT
QHGGGNGADVGQGSDDSSIDLT
QRGFGNSATLDQWNGKDSHMTVK
QFGGGNGAAVDQTASNSTVNVT
QVGFGNNATAHQY

>gi|15831006|ref|NP_309779.1| TonB [Escherichia coli O157:H7]

MTLDLPRRFPWPTLLSVCIHGAVVAGLLYTSVHQVIEL
PAPAQPISVTMVA
PADLEPPQAV
QPPPEPVVEPEPEPE
PIPEPPKEAPVVIEKPKPKPKPK
PKPVKKVQEQQKRDVKPVE
SRPASPFEN
TAPARPTSSTAT
AATSKPVTSVASGPRALSRNQPQY
PARAQALRIEGQVKVKFDVTPDGRVDNVQILSAKPANMFEREVKNAMRRWRYEPGKPGSGIVVNILFKINGTTEIQ

>gi|15833804|ref|NP_312577.1| EspF [Escherichia coli O157:H7]

MLNGISNAASTLGRQLVGIASRVSSAGGTGFSVAPQAVRLTPVKVHSPFSPGSSNVNARTIFNVSSQVTSFT
PSRPAPPPPTSGQASG
ASRPLPPIAQALKEHLAAYEKSKGPEALGFK
PARQAPPPPTSGQASG
ASRPLPPIAQALKEHLAAYEKSKGPEALGFK
PARQAPPPPTSGQASG
ASRPLPPIAQALKEHLAAYEKSKGPEALGFK
PARQAPPPPTGPSGLPPLAQALKDHLAAYEQSKKG

>gi|38704232|ref|NP_944582.1| hypothetical protein ECs5584 [Escherichia coli O157:H7]

MKRHSTLFLFTLLTLTTVPAQADIIDDTIGNIQQAINDASNPD
RGRDYEDSRDDGW
QREVSDD
RRRQYDD
RRRQFED
RRRQLDD
RQHQLNQ
ERRQLEDEERRMEDEYGQ

>gi|56708188|ref|YP_170084.1| hypothetical protein FTT1110 [Francisella tularensis subsp. tularensis Schu 4]

MSKCKTHQNHE
HKHQDDCGHTK
IKHGDHYDYLHDGHLH
HEHNGHYDEHTQEITEKNPDGC
HPIKEDCSS
HVHGPNCGHEA
VPHGDHIDYIVDGRLH
HPHGDHCDDHGPVEVIKNDKK

>gi|39995552|ref|NP_951503.1| hypothetical protein GSU0444 [Geobacter sulfurreducens PCA]

MKCPKCGYNSFEFLDNCKKCSHDLSAFKQTHGIRAVVLPLAGL
TVAAPAV
AVAAPSAANAANAEESFTWE
VPAAPESFKPGDDIFSDLDLGLSPSP
APAEPASAGLDLGDFSFDD
TPAAPAVPAPSTTDADAGTDDFATLLETGDNSEI
TAAAPGAAQVPEMESPWEVPTDSFGGFGD
TPAPAAEPAAAGDFDLESFSWDDKETKTAT
TPAAPKGPEAGVESFPGTDFDSLFGDAGEEEK

>gi|39995553|ref|NP_951504.1| hypothetical protein GSU0445 [Geobacter sulfurreducens PCA]

M
GAAGVVAVFVSLSSQLKLSRSKSPAAAGSA
AGAGVSPNPPKLSVGTSQGLSISGTCAAP
GAAAVISLLSPVSRSVAKSSVPASASVVEGAGTA
GAAGVSSKEKSPRSSPADA
GSAGAGDGESPRSRSEKMSSPGLKDS
GAAGTSQVKLSSAFAAFAAD
GAATATA
GAATVSPAKGSTTARIPWVCLKAERSWLHFLQLSRNSKLL

>gi|39995829|ref|NP_951780.1| hypothetical protein GSU0723 [Geobacter sulfurreducens PCA]

MRSYARGPNVGEVMGSWRVILVCMMLGMAVSPAWGDGGGVS
WERLTPEQREVLQHFAER
WNSFPPERQERLLDSAER
WRAMAPEARRNALERYRH
WHDLTQEQRDLIRKRIKEFHKLAPADQERIKRSYQR
FQNLPPQEREVLKGK
WRSLSPEEKRQFKEK
WRSMTDEQRREFLRDRSTDDSGAR

>gi|39996636|ref|NP_952587.1| hypothetical protein GSU1536 [Geobacter sulfurreducens PCA]

MSLILDALRKMEQERKARRGATENIRPDVLGHRGAQRSDSRKRWLFAAGAAVILLAGLAVGLLLKGGNDATRTAAVQRRD
AGRPAPLTAGYQDEGTESAP
APAPVPQPPVP
QPQPAPALE
APRPVPAPPAAA
VPRPVPA
APRPDEEEEESAPVVRAPS
TPAPVAVAQPADLSVTGIAWQDERRLRRAVVNGTLVAEGNVVAGARIVSIGERKVRFSKDGRTFDVPLSGGH

>gi|39996770|ref|NP_952721.1| lipoprotein, putative [Geobacter sulfurreducens PCA]

MRTNNGFPRGKRIRAVLSVVVVAVALLLGCAAGADEGIK
EGGKKVG
EGFKTIGKETGQAFK
EGGKEIGQGVKQMGK
EAGQEAK
KTGRSVGEWFR
ETGRKTG
EAFREMGRSIKRFFTGE

>gi|39996799|ref|NP_952750.1| TonB-dependent receptor, putative [Geobacter sulfurreducens PCA]

MDQLIQTRRSDAPMLWTIALSIALHGALYSVTSLLPRHVAGVEGMQVVSVDLSGSEIRPVA
PSPPDSKPA
PVPEPAVASDMSLPVENEQ
PPAEEVL
PPPQAVE
PPPPPVVS
APPSPLSLGMSRGFFRGIAEGESLRSDVREYYFTLLETINERWWTVANASGMDLGRSEAMFTIVMKKSGEMVDVQLVKST
GSPAYDRLILQAIQAA
NPLPPLPDSYTSELFLAPVRLVAPRGLLFS

>gi|39997392|ref|NP_953343.1| cytochrome c family protein [Geobacter sulfurreducens PCA]

MPSFLKPLRAAIPAALAVLCSLPPAASALTLTSN
DCGACHVGSNLDRHHRLISEAGM
ECLACHQLHDLPAGGYTILLSR
ECLACHNATVHNGVSHTVTTPA
DCARCHRESLETTHAENGWHGGDPAINAEIF
PCLRCHTSTSTSIIETISAGLAGETP
SCMNCHPYGMKRSRGTSSVRR

>gi|39997736|ref|NP_953687.1| cytochrome c family protein [Geobacter sulfurreducens PCA]

MKRLIALAGLVTLTLAGCGGGSNAGNASS
PVATSAHTPIWVTYHRFPTTESFSNNGAEAL

NQCKVCHGTNLLGAAEGAGA
PACLSCHVVDPVPYP
ERCYSCHGSDPVKPFFSWYSTMRSTRNGMPIDPAFVQRVRNGSLRHLKHDAVPPAEMENP
DACLRCHAPQTEFPDRPLPEIHHKLVFTVRDITGDGIPEL
VDCTSCHVYTIDPELNIPIRAV
RDCVLCHPTIVR

>gi|39997826|ref|NP_953777.1| cytochrome c family protein [Geobacter sulfurreducens PCA]

MKKWFIALLLLTVSAFTVQMALADKMSHKEYATTPI
GECNACHKGEGIAPNHDADWVRGHRVVASRAG
KNCADCHVQ
QFCLDCHQGGGIEADLSTRTFMRDYVPKSHRSNFLSIHPTKALDNP
QTCTRCHDQ
SYCNECHARFPKGSLRIKSHLMLGPNGQKYSFGLGEHAIEARRNL
QSCQTCHPEG
DVCIQCHSSGKTSPHPRNWNSIKNNYKDRAGS
RVCTKCHLPGTY

>gi|39997832|ref|NP_953783.1| cytochrome c family protein [Geobacter sulfurreducens PCA]

MKKWFIALLLLTVSAFTVQMALADKMSHKEYATTPI
GECNACHKGEGIAPNHDADWVRGHRVVASRAG
KNCADCHVQ
QFCLDCHQGGGIEADLSTRTFMRDYVPKSHRSNFLSIHPTKALDNP
QTCTRCHDQ
SYCNECHARFPKGSLRIKSHLMLGPNGQKYSFGLGEHAIEARRNL
QSCQTCHPEG
DVCIQCHSSGKTSPHPRNWNSIKNNYKDRAGS
RVCTKCHLPGTY

>gi|39997901|ref|NP_953852.1| cytochrome c family protein [Geobacter sulfurreducens PCA]

MRRHSNKTNPVAAGILFITLACWAGNAMGWGTSGDSRSTWGGSGSSLNLNLAFTD
CAKCHTSTNNVNRHHDLITKKGKQ
CLACHTMTADNSGQYTVQVQRD
CQACHTSSVHDNVQHNVST
CSRCHGSDVINIHSGWRSYTSTLSV
CYLCHTSTNAKVKATIAKGVSGQTVY
CTDCHGSNPHSWGGTWGR

>gi|37519754|ref|NP_923131.1| hypothetical protein glr0185 [Gloeobacter violaceus PCC 7421]

MKNAQTVGWNLWLQWVLAY
ALGSLAGG
AIARPLGE
AVGSVLGAAVAGL
VVGAALGL
AQAVVLGRYLKAGTPWIRATALGS

ALGSLLGGTTEWLALNRMGLPDLPSNVLGMT
VVGACLGVAQWWAVRKQFPQFGWWILGNTIAL
AIGPIVTLLAVLALGYFGPALTVFLRPLVSAAVSGGVLVGLLREAKIE

>gi|37519939|ref|NP_923316.1| hypothetical protein glr0370 [Gloeobacter violaceus PCC 7421]

MEEEMEQGQPGFF
EGLGEQADDFAGGVS
EGIAGLWGGITQGAEDAWNATTDIVG
EGMANVAGFADGVTDYLQ
EGTVEAG
EFLAEAGADVVGTTVGVASEGSEFVN
EGLAEVGELAQ
DGGAELSAWGEERSAADPEGSPLGEAAELTGGVLQAGGEWTEETGEFGDNLY
EQIGETGDDIVAYDPEFDAPQDDYVDQLDETLN

>gi|37519999|ref|NP_923376.1| hypothetical protein gll0430 [Gloeobacter violaceus PCC 7421]

M
AAPKCPNPQCEYFNRTLPNTAQVCPMCGTPLGNVVRSGSAAPVA
AAPPPPPIVSPYQQ
PAPPPHVY
APAAPPP
AYPQPAP
AAYPQPPVAPAAAPPASQGRPTLKLVHALIGREFAAPGDESYIGRRGGTVKPAPEIDLTGIPNDQVISRPHACILWDAAY
ATYTITDNNSRNGTFVNGQPLTPGVSYQLVDGDTLQLGRENLVQFKVAIA

>gi|37520242|ref|NP_923619.1| hypothetical protein glr0673 [Gloeobacter violaceus PCC 7421]

MSRVSFRVISTPAAVLLVFAISGPVSCPMLAQTKGGPVGKAAVAELMRSRQCV
SCNLAGANLS
NMDLSGVDLSFASLRAVNLR
GSDLTGALLN
GADLRGAKIE
GANLTGVGCF
NCQLGGASLRGATMAGMTLS
NADLTGADLS
NTDFARAGGLARSTIGAILPDGTICRPPATSPGVCPYR

>gi|37520537|ref|NP_923914.1| hypothetical protein glr0968 [Gloeobacter violaceus PCC 7421]

MLNTVRRLIRAAAVGGTACALAFGLTIGAASAADRDNNPP
GRGGGPGTNWENPR
GPAGGPGASPDGRWCLPA
DRDGNPP
GRVGGRGTNWENPP
GPTGGPGSSPDRRGYYAAR
DRDGNPP
GAVGGPGTNWENPR

GPVGGPGASPDRRRC

>gi|37520958|ref|NP_924335.1| hypothetical protein glr1389 [Gloeobacter violaceus PCC 7421]

MVGSALLHAGLFALKLP
EPKPPEIPKAQKMVMVET
APKLPPP
KPKVEPP
KPKPPEPKSFSR
KPVPKKSAPQKAKASRTILSSKAVSADAGTVSENQLAGSRGTGYGTEKGDDFGSIAPLGVDGGSGGTDVV
ETPPPPPPPPPPLVNARPKGAV
QPEYPEIAQQNNWEGRVIVKAYINADGSVGEVQVAKSSGHSELDNAALAAVRNTKFEPARRGEESVGAWVRIPVTFSLQ

>gi|37520960|ref|NP_924337.1| RNA-binding protein [Gloeobacter violaceus PCC 7421]

MTLFVGNLPFSATEQEIVEAFTEYGEVKSVKIPMDRETGRPRGFAFVDLESETAEQAAIDDLDGATWNNREIRVNKAEPR
PDRRA
GGGGGGGNR
SGGGGGGYR
GSGGGGG
GAGGGGGRDRRY

>gi|37520978|ref|NP_924355.1| hypothetical protein gll1409 [Gloeobacter violaceus PCC 7421]

MRERAMSNFFKDFMNVESFIYSTWMNLFGGAFSRLTRPQSGASSVVPGVA
APLRPGMR
PPQRPGAP
APQRPGAPPAQGA
APQRPGAP
APQRPAAGA
PGPRPPA
APQRPGM
PPGMPAPNPNRQLDKSAQGAQLFGMNINN
LPQIPGLNA
PPSGKGNAVQTRAEQEIYLLKARGMNFVSQGVHRSFSQMYTNLRSFFERRA

>gi|37521373|ref|NP_924750.1| hypothetical protein glr1804 [Gloeobacter violaceus PCC 7421]

MPLGSVEILNTFLMALQLF
LLGEDLT
LLGEHLV
LLGEHLV
LLGKHLLLHVLHDRQDLAEDVGHLLNDRAGSSQMLQRLGK
LLVEHHLTGAQFGKEVELVAELLDQSFQNTLDTLDSIGRLAVGVHGLPGLEVHGDTLFNHC

>gi|37521382|ref|NP_924759.1| hypothetical protein gll1813 [Gloeobacter violaceus PCC 7421]

MTKECYKSWRHAWTPATNVLQPPPRDRRARRSLMDENMNTAPPST
EAPLEVEEKPVV
AESVEATSE
DEPVVEQVLDQEATA
EEPHLEETTA
AEPVEEETIV
AEELLEEPTLEAPTAVE
ETAVEEESISEHAAPEEMVGEPAE
AEVTEEEAAVV
EEPVAEDSEVAA
AEAVSETPL
EETTEEEAIVEPTEA
AEDTAEEHVPGEQAVEQTASVVDADTDETRTEQATA

>gi|37521803|ref|NP_925180.1| phycocyanin alpha phycocyanobilin lyase related protein NblB homolog [Gloeobacter violaceus PCC 7421]

METPEEKL
ASPNLGDRLS
AVGALRDIPAEHSVALLVKAAQDDNARVRYAAVS
LLGQKEDRSVLGLLRTLLTSDPEFDVRAA
AAAALGDLHAQEAYEDLAAAFAQDNEWLVRFSILA
ALGELGDRRAVDLLVTALKGDPLIRTVAAG
ALGQLGDRRAVGELVAYIADEDWQLRFRIAQ
ALGELGGEEACSALQTLAADPNDAVSSTARS
ALEQLAQ

>gi|37521812|ref|NP_925189.1| hypothetical protein gll2243 [Gloeobacter violaceus PCC 7421]

MKTHIAGLLALGLLLSAAPGMAQTTTPYGGSSSTTP
SNPRSNPTTDE
SGPRSTEPRTDATPSAGTPAAERPGSS
MDPMTRPGSSMDSKNKMQPTTPSS
GNPRSNPMNDK
SGPTSEPRTDATPSAGTPAAERPNTSMDPMKPDT
STKKSNKTSSGMGKSSDRVDVTRTTDLPE
SQSRSNDAVIADPPSGVSN
TDPNLQPKSTTRPA
GEPKPKP

>gi|37522042|ref|NP_925419.1| hypothetical protein gll2473 [Gloeobacter violaceus PCC 7421]

MVTRERDPKTGKVTFK
GTPGNDNIRVSNRRNTAGEVDGLLVIVLDAHGKECERFSLDREEFKDGLCIR
GGAGDDIIQVDEDVEFDLELY
GEEGNDLII
GGAGNDLID
GGPGNDRLY
GGEGNDTIH
GGDGDDLID
GGPGDDKLY
GGAGNDKIL

GGPGNDLLD
GGPGDDCLV
GGPGDDILI
GGEDNDRLF
GGEGMDQLI
GGTGKDIVHADLSDLPVQRGNDEQDTVHIDRG

>gi|37522337|ref|NP_925714.1| hypothetical protein gll2768 [Gloeobacter violaceus PCC 7421]

MGRKGGPGADFGSGF
AQAVQRAPPEEAGLYRFGSALLHIYNALRQVKKSLEVYCLEQGEYSFLSDNPVWLPDIGLAIGCERGVYQGIEREWLYWY
DQGGQRYLTPEERT
AEAARLA
TEAAQQA
AEAERRA
EQAERRV
AEAQQQILK
EAAARRA
VEAQLQALRHRLQEQGIDPGA

>gi|37522624|ref|NP_926001.1| hypothetical protein gll3055 [Gloeobacter violaceus PCC 7421]

MPMLVNPHVLWHDWMRFSIVCERCHRRLLFSQWSGGGGEEWPQIKKVECRLEARFGVQRAERCELCGRPFAAVRRPVRSM
GAGELLARYRRGQRSFLKVALVGD
DLSTAALDGARMFEA
NLSGSRLVGTGLRSA
VLTGASLVGA
DFAAARLQRA
DLSGAELVGANLSGT
DLQGAKLIGG
DLRGASLYWA
NLEGAQLRGALLDDA
SLRGARLRGATLPDGRLLD

>gi|37522759|ref|NP_926136.1| hypothetical protein gll3190 [Gloeobacter violaceus PCC 7421]

MRSEVGCTQIRAILRFSPMTPSNGQPPSSSDRLDRIERILAAVAASSAEH
DRQIGVLLSRDAVY
EQRMDAF
ERRTEVL
ERQLEAFNWRTEVL
DRRMEAL
DQHTEVL
ERRMEAF
DQRTEAF
DLRLVSL
DQRLATIAQQQ
ERNTEAISRLSEEVAGVNRTTADAAATATNAITGLVDVADNILARIDDIMARIGDIMARIDANSSQIVGLQTENRRILEI
LERRNGGRA

>gi|37522760|ref|NP_926137.1| hypothetical protein gll3191 [Gloeobacter violaceus PCC 7421]

MIEAEKIKQGEVE
DLRGAGLEGI
DLAGVRLERH
NLAGANLTGAHLAGAVLHC
DLWGANLMGANLSES
DLRGANLRGA
NLMSANLQNA
SLGGANLEGTNLMGLNGIGADLRHA
DLRGARLSGANLAGA
DLSGANLAGA
DLRGADLEEANLSSAAVAGA
SFAGANLLCA
DLSGIDTRLA
DFEGACLVGTLLDAV

>gi|37522930|ref|NP_926307.1| hypothetical protein glr3361 [Gloeobacter violaceus PCC 7421]

MSEPWQQWLKALTERLDNQINLMAELRVGQVVAN
QQMAVTQ
QQIGETQ
RQLGETQ
RQVGETQ
RQIGETL
QLVGETQ
RQLAATNTAVLQLTDAFGRILPVIETQQ
RQLQSHQERIEDLYRAQGHILR
RLFGEEEQQP

>gi|37523071|ref|NP_926448.1| hypothetical protein glr3502 [Gloeobacter violaceus PCC 7421]

MLKAQAPELSRHSGAGESAMRRGWIVGMLIGTVLVLVGWTPLMRAFLQDFVLVVDASGSRQTPKNYRGAKLAGQ
DLRGQDLEGA
DLRSADLRGT
NLRGAHLSGANFQGA
NLEGADLSTAILENA
DLRGASLVGA
DLRAADLAGANLSAARLVRA
DLSAANLTHA
DLRNADLHGARLDDTLLQQT
DLRGLDLARLDLEDTDASAARVR

>gi|37523076|ref|NP_926453.1| hypothetical protein gll3507 [Gloeobacter violaceus PCC 7421]

MPVGSAASQITPLRVNQHNCTLRTKEGQGGRATTRQLHKLKGVFMIRRNFATALALACACAAMGTAQAQTTTS
TPGGGTVVE
GERGGKAYV
GPNDGKVLV
APDGRKYVS

GPDGGRLYVDEDGRKYAE
GPNGAKVYVGDDVKVYK
GPQGGTAVE
GPDGGRVYQ
GPRGGEAAV
GPEGRKRYE
GPEGTKVYS
GPRGTVIKDKDGQVIRQPR

>gi|37523136|ref|NP_926513.1| hypothetical protein glr3567 [Gloeobacter violaceus PCC 7421]

MVTNVV
RALRQGAWIAGFVAVGQLASM
AAVSADP
VQLRLQPQRPTAAVNPIDFGANRYILTDEAFAQAVSEYLQGRPDRAVVQFG
EAVRLDPKDGPAFNNRANARTALGDSQGALADYD
RAVELDPGYAFIYFNRGITKSGLGDRQGALADYD
QALEIDPLNVFAYNNRGILRYKLGDRAGSLTDFD
RALRLDPTYAFAYHNRGIVRSKLGDSRGAEADSQKAAELMGKSPRRRSA

>gi|37523542|ref|NP_926919.1| hypothetical protein gll3973 [Gloeobacter violaceus PCC 7421]

MVLSPETIVYPDSDGKPMADNTEQFRWIVYIKEGLEWLFASDPNVFVAGDLLWYPVEGNPRLCQAPDVLVAFGRPKGKRG
SYRQWLEAGVAPQVVFEVLSPGNGVSEMTRKFQFYQHHGVEEYYLYDPEGGTLDGFVREGAALLGIDPMHSWVSPRLGAR
FELGERGDLCLWRPDGTPFESYVEMAA
RAEQERL
RAEQERQ
RAEQERQ
RAEQERQ
RAEQERQ
RAEQAEQQALRLAQRLRELGIDPQA

>gi|37523592|ref|NP_926969.1| hypothetical protein gll4023 [Gloeobacter violaceus PCC 7421]

MKAKLAAAVAVLM
VAGTAAVSAETVTREVPAFSRIE
VRGSTDIDVRAGANQLVVLLEGAPGVIERIVTEVSGETLVISGPRGSSMNIGGRGPRALIRVPQLQSIT
ILGSGDARLSGIRSDTFSVS
IAGSGDVVVAPAEGLSTVDRLQVD
VAGSGDANLAALRARTAAVS
VAGSGDVQVYTTEALDAS
VAGSGDVFYSGNPQRVNRK
VAGSGTIVSR

>gi|37523818|ref|NP_927195.1| hypothetical protein gll4249 [Gloeobacter violaceus PCC 7421]

MDELERSYWVLGLESGASPEEVSQAYKDLVFVWHPDRVPTENTRLHQKAVEKLKEINRAYEHLRSHRPSPRKVHQAHQTQ
PASNFRPVYRPSASATPANAYRPRPAPPPPPPPPPPPPSRPDMS
GADLRGANLKEKDFE

GRNLSNANLS
GADLSDAFLHKINLS
GADLSGANLFRANLLRANLR
GANLQDAHLIGADFS
GADLSGANLR
GAQMSAAGRLMVKLTAAILT
GAVMPDGQVHA

>gi|37523863|ref|NP_927240.1| hypothetical protein glr4294 [Gloeobacter violaceus PCC 7421]

MAELARHQIVLSIINTPAGATLRFRGL
NLAGLDLSGGL
DLSRANFSLA
DLSGVQLQQADLEQARFWKT
DLRMADLRHT
HLAGAMLAEA
RLARANLRQA
DLRGANLQGANLAHA
DLSAADLTDAALQGA
DLAGAILDDCKLERA
DLSGAVLTGASLQHAVLIDANLHGARLTGA
DLQGAHLIDATCTGA
DLSGASLRAA
DLSRAVLVDT
DLSGANLAGADVTHTDLRRAILRNVHWDGPSYSQS

>gi|15789683|ref|NP_279507.1| hypothetical protein VNG0441H [Halobacterium salinarum NRC-1]

MNAER
FPDAVADRTRTALDRLGSSKRLVAITGADLTTPTVLARAA
ASEAAAADTFDAL
AAAADTD
AAAAACDSFAATERDHYDRVTAALDAD
DADTAADPGAIHDYLRERPDDTTRLGAVVGRGLVADRTLAQFVGFF
VNDAAHDRADLFRDL
RRDTSAD
TADAAAALASCCDGDDDWAR
AADAAAD
AVDAAYQEYVAALNGMGVDAKSVC

>gi|15789804|ref|NP_279628.1| hypothetical protein VNG0597H [Halobacterium salinarum NRC-1]

MADDDQPADDAAEDDGEEKSFRERVEEIREERAEERDEGDGEDPRERMADAMG
GGGGPPG
MGGGGGGGGNPFAQM
MGGMMGG
GGGGPPG
MGGGPGGDESGGTSETAERELEMLRKELERTNNKLDRIANALED

>gi|15790293|ref|NP_280117.1| hypothetical protein VNG1235C [Halobacterium salinarum NRC-1]

MEVLVVGAGAVGRWVAEVVA
APAAFADTDPQ
RAVAAADATEN
ARAVAVDTAASFAVVAVAVPMR
AAGAAID
AHAAAAERAIVDFTGVMGEPLA
AMAAAGPPGVERASFHPLFAPAHAPGRIAVTDGDGGPTVDRLRGWLADAGNELVEVPAAVHDDAMETIQGRTHAAIL
AFALAADDVPDALATPVYEELSRLADRVTGGTPGVYADIQSTFGGSQ
EVAAAADRI
AAAAAGDGDAFAEVYEDAS

>gi|15790933|ref|NP_280757.1| hypothetical protein VNG2091H [Halobacterium salinarum NRC-1]

MVGVGLAAIVLLKGMFAAPRPPASVAHVTAHGTGFPSGHAIGSAVLYGGLAALLDVGGRRQRYAAGAAMVVVVSATRLAL
GVHYLVDV
VAGAGVG
VAVVGGVLAVTRRRVGYGFGLAV
ILAAGAVAVAGPTTDA
VAALGGAAGGLVGWRI
VAARGAVGRAVRP
VAGVGAVVA
AGGVAAASV
GVGAGAIPVFGA
HAAAGAVFVGLPAAQNFSR

>gi|15791235|ref|NP_281059.1| hypothetical protein VNG2465C [Halobacterium salinarum NRC-1]

MSLGGG
GQQQLQQ
LSQEIQA
IEEEVEELEADVASLRQEQTEIEEAKEALDVLETGATVQVPLGGDAYVRAEVKDMDEVVVSLGGGYAAEQDSDAAASVLD
EKKATIDGRIDDVQAEIAD
LSEEAEQ
LEQQAQQ
AQQQMMQQQMQAQQQPQDGEQ

>gi|56459133|ref|YP_154414.1| Zn-finger domain associated with topoisomerase type I [Idiomarina loihiensis L2TR]

MSSD
ERCPRCERPLVIKHVGHNSFLGCTGYPDCDYKRGLREQSEIEPEDLG
VPCPECGQELQLKSGRYGLFVGC
SGFPECEFVTEPNVEEEET
IRCPECGKGQLHQKTSRRGTVFYAC
DQYPKCEFTVNQPPVD
ETCPKCRYPLLVKKKTAAGIR
KVCPQKQCDYKSDAL

>gi|56460324|ref|YP_155605.1| Uncharacterized conserved secreted protein with internal repeats [Idiomarina loihiensis L2TR]

MKKVLSALTITALMASPVFADAQHKHGDDKTGS
MSGMMHHEK
MTSMNEHMQEMREAMEKIKSESDPKK
RKKLMEE
HMKSMQK
GMQMMNHGMDKGMNKGEHKKGDKAEQMGKENMMN
KMKMMEHRMDMMQKM
MEQMMEHQSEEMKQGS

>gi|58336515|ref|YP_193100.1| hypothetical protein LBA0168 [Lactobacillus acidophilus NCFM]

MKKLVFWGVIAAMFILGGGGAGAYISHTNAVSHNNSTYNKSNNNNGKPVKSGQSDDKN
SSSNKEDNQ
SDNSSSDKS
SSSSSSDKDAKN
ASSSSDDKDKNSSDDQ
SSSDKSQ
DDQSSEDQ
SSSTSTSHEQ
STQSSTEQEQ
SSSAASTSNE

>gi|58336573|ref|YP_193158.1| RNA polymerase delta subunit [Lactobacillus acidophilus NCFM]

MGLDKFKDKNRDELSMIEVARAILEDNGKRMAFADIVNAVQKYLNKSDEEIRERLPQFYTDMNTDGEFISMGENVWALRS
WFPYESVDEEVNHP
EDEEEDDSRKHHKKVNAFLASATGD
DDIIDYDNDDPE
DDDLDAATD
DSDDDYS
DDDSDYD
EDNDDADDVLPDGIEGQLSQLN
DEDDDEDD

>gi|58338174|ref|YP_194759.1| hypothetical protein LBA1932 [Lactobacillus acidophilus NCFM]

MKKFVALILGCITICGLYGCSKNSSQIINKTFSLPAYNRIE
IDSANDDIKIISGKHYRVRYVGQNKLKPT
IKVKNGVLVIDSPKKSTIINSNIFGSKNIKQTLTIEMPKKELKYLS
VDTSNGDVSVDY
LEVQKGT
IDTSNGNVNLKNLITKKDFE
IDTSNGTVKVGKSNVEGYA
LSTSNGSIVVKGENKSDEFEKNTSAKNVLS
IDTSNGNIYVN

>gi|50954133|ref|YP_061421.1| hypothetical protein Lxx02840 [Leifsonia xyli subsp. xyli str. CTCB07]

MGFLDRLLGREEPRRDAPQSAEGWNADSQSTHGSNAYARNAPQPPARSDDEIAVERYRYLLRTAPPETIEQVHEEAFARL
TADQRRLLFQQLSEGAPA
GERPRGDDPRSLAQAATRSELRQPGTLERS
LGGSGGGVGAAS
GMGFGGMLATSLLGSVAGYVIGSALVSTFLPTPGADQA
ADGTAEDGSG
GDGSGGDGTATDT
GSGDAGVETASADAG
WGGDSGDWGG
GGGGGWD
GGGFGGDFL

>gi|50955749|ref|YP_063037.1| membrane protein [Leifsonia xyli subsp. xyli str. CTCB07]

MNA
VFSTVSSRSQSKSLAVHRARWGMVFVAMLVAVPVSTTPAAFAET
TPSTISVGKSPVAVAVAVAVGPHSKNVYAANQ
DGNTVSVIDAASHQ
VTSTVQVGQKPTGVAVGSHWKTVYVTNG
LDSTVSVIDVPSYR
VTSTIPVGTFPDAVAVAPDGKTAYVTNF
VDSTVSVIDAASLK
VTSTFPVGNSPRGVAVAPDGKTVYVTND
KDNTVSVLSLPATPPTPSVTPAAAERPISALPVVAG

>gi|52425100|ref|YP_088237.1| hypothetical protein MS1045 [Mannheimia succiniciproducens MBEL55E]

MLLFVYLGAKMKKLALAALILGSSLALT
ACDQAKEQASQTTETVTETAKDVKDN
AVEKAGEVKDN
AVEKANE
VKDAAAEKMDAAVDATKEKVAAAKEA
VADKAEEVKNA
VSDKAAEMKEA
VSDKANE
VKDAAEK

>gi|52426254|ref|YP_089391.1| SlpA protein [Mannheimia succiniciproducens MBEL55E]

MKVAKNIVVGIAYQVRTEDGVLVDEAPTNQPLEYLQGHNNLVIGLENALEGKAVGDKFEVRVKPE
EGYGEYNENMVQRVPKDVFVGVDELAVGMRFIADTDMGPLPVVITEVSENDVVVDGNHMLAGQELLFTVEVVSAREATPE
E
IAHGHIHS
GDHDHGHGGCGC
GGHGHEHDHDHSHGGCGC
GGHGHHHD
HGHDHHHGEGCCGGHGKK
EGHGHGNC

GCGGHGH

>gi|45358088|ref|NP_987645.1| hypothetical protein MMP0525 [Methanococcus maripaludis S2]

MNKNLKKTTGAILMVITLVPYCMAMDEQTSEN
SSGIFNNILSWLKSALGTNNQIEPE
TVEFENNSSKNLPM
MPELNNSEMNNS
TPEIINN
TPENITNKTRNTPNN
MPENFNEYVPEQRNMSG
MPENVNNTKY
PPQIPENFENN
TPIEINSNNE

>gi|45358303|ref|NP_987860.1| possible purine NTPase [Methanococcus maripaludis S2]

MRIILDKIRKTIENNNESGKNILDTDKITLELGKLDNKVNGISKELKGHSK
TFKDLEEVLPEYFEDTEKNTK
KIQELGNTLDKIL
KYIEQEK
KIKEQQDL
KIKELEK
RINDLEEINKNWTVVKTWMDNRKTNEKINSEKLQLLET
KFNGIEKYINTERYKKTIKRETDNEQVLSILKNGCSQPKDLVKNFKGGTKALYD
TLKRLEKSSAIIRKKDGKQVFYELKQK

>gi|45358458|ref|NP_988015.1| TPR repeat [Methanococcus maripaludis S2]

MKGLNMRNNYGIQKWHSHGALLCNSEK
YAEALEC
YDKILSYYPDDFLAVFGKGMVFLKLKE
YEKALQCFNSVLNMNPSYIPAIKNKKQVEEKLKKQESENYNKFSKLGVENYKQGN
FEKSLSY
FEKAFEINPYSETLRKNIETTRLKLKEILPIRWNNKGVEYYKVKN
YQKALEC
FEKAVKLNPK
FESALKNKTMVQKLIKN

>gi|45358520|ref|NP_988077.1| hypothetical protein MMP0957 [Methanococcus maripaludis S2]

M
DLFLIFLAGISTALGPCIVTVLPFVLAYTFGISNSRFE
GFLVSFSFMV
GFSIVFSTLGVLSSAF
GIFLNFTLLKYIAGSFAIIF
GILVLFNKNFTFRKE
GVFQKFSKKLSNMNDISLHVKILNSIVLGIVYGFGANICADPILAGILTFVASKADVYY
GFFALLTYSLGY
GIPIIFLSTLGAESKQIVEKIVKKKFITYIS

GIILIIL
GLFLIFSV

>gi|45358603|ref|NP_988160.1| A1A0 ATPase, subunit K [Methanococcus maripaludis S2]

MVFENPLLL
GAIGAGLAVGIAGLG
SGIGAGIT
GASGAGVVAEDPNKFGTAIVFQALPQTQGLYGFLVAILILFVFKSAPEWAMLA
AGIGTGLAGL
SAIGQGIASAAGLGAVAEDDGIFGKAMVFSVLPETQAIYGLLVAILLLVGVFASPGVTTI
AALGAGLAVGFAGL
SGIGQGITAA
GAIGATARDP
DAMGKGLVLAVMPETFAIFGLLIAILIMLGIMF

>gi|15597177|ref|NP_250671.1| hypothetical protein PA1981 [Pseudomonas aeruginosa PAO1]

MKTFLLCAA
LLGAGLAQAADDIDDVRVIAGCRLEPGARCAKAD
LRGADLRNLDLGRIDLAGAD
LSGADLRHARLDLANLEKAN
LRGADLTRASLQQAN
LRGADLSGARAVAIQAWGLFAQGAQ
WQGADLTAAYLEFARLSGGRLHQAT
LRAADLEMTWLSRAD
LKGADLRDANLQEVKLAEAN
LEDADLRGSKVRFGNFQGSNMQGCKDCPAGWQ

>gi|15597386|ref|NP_250880.1| hypothetical protein PA2190 [Pseudomonas aeruginosa PAO1]

MTDKRQGMSTSEA
GQKGGAATSRSHGKEFYQEI
GHKGGQASGGNFANDPQRAAEA
GRKGGQQSGGNFANDPERAAEA
GRKGGQQSGGNFANDREKASEA
GRKGGQH
SHGGGRSS

>gi|15597509|ref|NP_251003.1| hypothetical protein PA2313 [Pseudomonas aeruginosa PAO1]

M
QHIIGMPLQ
LIIIGMPHSIIFAISMQQFLNISMS
MPCIGIILQCMASPCMEHSIMHFIGA
MPIIGMPIP
MPIVGIIAMQAWD
MPIIARLPH
CIIMGMPQFIIIAIISALLRNWPMS

IPIAGIILHSMPSSIVSHCMV
AIIIGMPIMGIIGMAAFMFIGICMAFIMAVTPVGEAGVGANGADSRQRRRTGKMPVQQPSRALHGG

>gi|15597758|ref|NP_251252.1| hypothetical protein PA2562 [Pseudomonas aeruginosa PAO1]

MFNTRSVLALVVSLSLIGGTPAVLAD
PGNGKGNPHGNQ
MDGGKGNKGN
KGNQGGHD
NHGGPGN
KGGGAGQ
KGNGGGNWDSGPRIDVGGVRVILGDNRGYWNPGPALPPGIQKNLARGKPLPPGIAKKLDGRLLGRLPHYDGYEWMQAGTD
LLLVAVATGLIYEVLDNVLD

>gi|15598012|ref|NP_251506.1| hypothetical protein PA2816 [Pseudomonas aeruginosa PAO1]

MLKLFATGLAASFLLLPPAQAAPPAPYGVQPHQQA
VQRAGEQRQRQ
LQEQRQRFDEQR
LQLQQDQ
LQRQQQN
LQRQRQQRQMQDN
LIRQQQLDQQRWRLEQDQRR
LDSERRQ
LENRRRQSQSPAIR

>gi|15598090|ref|NP_251584.1| hypothetical protein PA2894 [Pseudomonas aeruginosa PAO1]

MP
AAPKTAS
KKVAPAAEQVAEPKPP
AKPKPAA
APPKPASRPV
AKDKPAPAKRASTARLDPEV
RKPLPSAKLDLRLPKELVQKMAPPGTEET
HKPKPLLPPMFEEKDDSDFRLNGQLLYRDRELGVSRPLEGAELQFEFRQ

>gi|15598991|ref|NP_252485.1| hypothetical protein PA3796 [Pseudomonas aeruginosa PAO1]

MGLLDSIL
GSALGGSEGQGGARDPKVALLIGLVTMLMSRQGGAAQG
GGDGGGLLGSLGSIL
GGAGGGAAS
GGLGGVLGDLL
GGAAGGAGQAAPGVEAGSL
GGLGGLFQMLQNAGLGDALN
SWIGGGPNQSVSAADISRVFGDGQLQQLADSAGVSQGEAAEHLSSLLPELVNKLTPDGQAPQGDLDIGSLLARFS

>gi|15600262|ref|NP_253756.1| translocation protein TatB [Pseudomonas aeruginosa PAO1]

MFGISFSELLLVGLVALLVLGPERLPGAARTAGLWIGRLKRSFNTIKQEVEREIGADEIRRQLHNEHILSMEREAQKLLA
PLTGQNPPQET
PPPAAES
PAPSVPTP
PPTSTPAV
PPADAAA
PPAVAASTP
PSPPSETPRNP

>gi|15600313|ref|NP_253807.1| hypothetical protein PA5120 [Pseudomonas aeruginosa PAO1]

MRHSLWLLLAAILSLPAQAGTECRDIHD
RDLRRMCNALERGDSGDCGDIDS
RDMRRYCGALLAPGQRYDCDDIRD
GDTRRQCRAIV
RGDRKRCDDIDS
RDMRRQCRAVVSRAPWQCDGIDD
RDMRRICRVILSR

>gi|15600375|ref|NP_253869.1| hypothetical protein PA5182 [Pseudomonas aeruginosa PAO1]

MLGSVVG
DAVGGSTGAAIGSGIG
GAAGGAV
GAGRGNKTE
AAIGGGL
GAAGGNVIGRQIG
GSTGGLIGAAL
GGAGGGALGNHYGDGNRRYDDDDDYRDRRYYRRAGYRDGYYRHDNGHHYGQYKKWKRHKHHRRYYDD

>gi|15600376|ref|NP_253870.1| hypothetical protein PA5183 [Pseudomonas aeruginosa PAO1]

MRKSLLVIAFGLLSAQAALAGDSTK
TAVGGGVGGALGNVVG
GAIGGSTGAAIGAGLG
GAAGGAMTAKDGRKTE
AALGGGL
GAAGGSVIGNKLG
GSTGATIGAGLG
GAAGGAVGNNLGKDNDSGHRGKKHRKHKHR

>gi|28867307|ref|NP_789926.1| tonB protein [Pseudomonas syringae pv. tomato str. DC3000]

MISTRQKLTRYSGSLLLVLAVHAIAIIVAVRWPASQAIELPPAAMMVELAPLPEPAPPPPPKVVQ
PPQPPAP
EEQVPMPE

VAQAPKPEIVVPK
KVEKPKP
KPQPPKPQK
KPEPPQE
KPADEKPVDT
PPSTAKPE
KSAAPAPAP
TPSPPSTALPSWQGDLLSHLGKYKKYPEDARRRGMQGVARLRFVVDADGNVLSYSIANSSGSPALDRATMEMIRRAQP
LPKPPKEILNNGTIEILAPFVYSLDKR

>gi|28867884|ref|NP_790503.1| single-stranded DNA-binding protein [Pseudomonas syringae pv. tomato str. DC3000]

MARGVNKVILVGTCGQDPEVRYLPNGNAVTNLSLATSEQWTDKQSGQKVEKTEWHRVSMFGKVAEIAGEYLRKGSQVYIE
GKLQTREWEKDGIKRYTTEIVVDMQGTMQLLGGRPQGDAQQGQGGGGNY
NQSAPRP
QQSRPQQSAP
QQSAPQQNY
NQQPPQQ
RDSRPAP
QQQAPQPAADFDSFDDDIPF

>gi|28868617|ref|NP_791236.1| exchangeable effector locus protein [Pseudomonas syringae pv. tomato str. DC3000]

MRAYKNLTAKIGGFLLALTIIGTSLPAFAV
NDCDLDNDNSTGATCGG
NDKDLDNDNVTDAAFGG
NDKDMDNDHHTDAAFGG
NDKDLDNDHHTDAAFGG
NDKDLDNDNKTDAAFGG
NDRDLDNDNNTDNYNGTPSAAKK

>gi|28868975|ref|NP_791594.1| membrane protein, putative [Pseudomonas syringae pv. tomato str. DC3000]

MIEVVMFVL
LGAAMGT
LGGLFGI
GGGLVAIPA
LGVLFGLDQQLAQGTALLMVLPNVLLALWRYNQRNRILLRNALMLIIPSFIFAWLTSMLAVRIDPQSMRLGFVAFLVILT
VFNVAQMYWRNTLASGGLRHEKYLWLLGVGSGV
TGGLFGVGGGVVATPI
LTSVFGATQVVAQGLALSLAAPSTGITLVTYALHDHVNWSMGIPLA
VGGLASISWGVKLAHSLPERTLRLMFCVFLVLCAVILALKV

>gi|28869872|ref|NP_792491.1| hypothetical protein PSPTO2684 [Pseudomonas syringae pv. tomato str. DC3000]

MIRQIPKIALFVGALALAGQANADRGWGGPAVVGAIVGAAVVGAAVSSGRDRTVYVE
REPVYVQ
QQPVYYA
PPPVYYQ

PAPVYQVVERYPPP
PPPGYYHRYGPP
PGPGYYYGPPRGRW

>gi|28870968|ref|NP_793587.1| hypothetical protein PSPTO3813 [Pseudomonas syringae pv. tomato str. DC3000]

MALLDNVVKQRMVGALVLIAVAVIFLPMLFTREDETRQVQVEAPSAPHA
PAVSQVKVD
PVPVPEPQVLPQE
PVPGDADMGSNNQPPAMP
IAPAPTA
PATAPASKPAAKPT
PAPAPAATA
PAPAPAKPAAPAGVDANGLSISWSVQLASMSNRANADNLQKTLRTQGYNAYIRTADGVNRVFVGPLIERAEADRLRDQLD
KQQKLEGIVVRFQPERG

>gi|28871434|ref|NP_794053.1| hypothetical protein PSPTO4297 [Pseudomonas syringae pv. tomato str. DC3000]

MKTTKLIFGLAFSVLASSAFALPAVDTVNNSDMSKASLVAEGGTKHTNV
GSIRVSAD
GADHVGANRLAAD
GADRVGANRQAAD
GADHVGANRLAAD
GADRVGANRQAAD
GADRVGANRLG

>gi|28872516|ref|NP_795135.1| hypothetical protein PSPTO5408 [Pseudomonas syringae pv. tomato str. DC3000]

MTPRSRRHPYEIPYMKYSSILLLSLTL
ISGAASAGNLESG
VGGALGGVLGSVVGQQ
IGGSTGSAIGAG
VGGAAGGAVGADRRNRTEAA
IGGGLGAAGGNVLGRS
VGGNTGSL
VGSAVGGGGGAALGNYMGNESRSDDRRYRGDRDDRREYRGHGKRNKHYGHRRHGRD

>gi|32471726|ref|NP_864719.1| hypothetical protein-transmembrane prediction [Rhodopirellula baltica SH 1]

MFELLHSNYRSRASVRMGATVVEMAIVSTVLFTILISGIELTRVTMLRHSADHAAYIGARRGIITGATAENVEEVVQGHM
DAIGIRDATVTVIPEEITEATTQVEVEVGVPLAMNTWISPELFGKKLKGRARLLTERAAMVMSQ
SMPTPPP
PPPPPPE
PEPEPEPN
PEPEPQPQPNPDPEPT
PEPEPEPPPPPPPPLL

>gi|32473144|ref|NP_866138.1| hypothetical protein-transmembrane prediction [Rhodopirellula baltica SH 1]

MNKTFLMTLLSLSLGFATVGCDVEKTR
DGEMPDV
DVDGGQLPGY
DVDAPEVDVTTEES
KIKVPDV
DVDMKEK
TITTPDV
DIDLPADE

>gi|32473416|ref|NP_866410.1| hypothetical protein RB4940 [Rhodopirellula baltica SH 1]

MSPCSLFPALLLLHKPKLACSVASRRTAAANLPQFAAKHQNL
FAASQLQHQ
FAANQHLHQS
FAANQLQLQH
LAANLLQLQ
FAANQHLF
AAQNQLQ
FAATHVQSQSVACLQRSRLVSQPKSAAQILAATESTEFVRFRSELETAFYRHKSECCSRTVLVFLCALGSQLLTWGQDST
QGIRTRSSPLLLQPSVCKIVECTSTDHAIGLPPLSSLFNSTNCHDGISQYSRSRRRPTES

>gi|32473417|ref|NP_866411.1| hypothetical protein RB4941 [Rhodopirellula baltica SH 1]

MLQEEQ
LLRTCPS
LLRSTRTC
LLRASSSTS
LLRTSTCTSR
LLRTSSSSST
LLRTCSSSS
LLRTSTCL
LPRTSSSLLRPMCKAKVWLVCKDQGSFRSQKVLPRSLLQLNQRNSFVSDLNWKQRFTDTRASAVHEQCSFFYAHSAPSFS
PGGKIQPKALELDRRHCCSSPQSAKS

>gi|32473590|ref|NP_866584.1| hypothetical protein-signal peptide prediction [Rhodopirellula baltica SH 1]

MPQTHNCVEGYRVYSLSTYKEFPMNFARTMAAVAV
AVCISSASSASACGLLGCG
NDCCDAA
PVCCEAPA
PVCCEPA
PVCCPPAPVKVSWCVKN
PLTCCPVEVSACV
PACCENVIPE
LVCCKKGFLGRKVMTFKFGCCDTCVDVVFTPLGRVIVRD

>gi|32473993|ref|NP_866987.1| hypothetical protein RB5957 [Rhodopirellula baltica SH 1]

MMLRTSMMFAATVLIGVAILFATRS
AISSDNASAQNGE
ACPCNQC
DVGCDCCVNG
DVNCDNCSCE
ACECDACDTAASDFAVAKMSCCA
GKTCENAANEKAASSDVLTSAVAA
ACVCGQC
DADCNCCLDE
SVDCDNCSCE
VCQCEGCVDAPATGA

>gi|32474601|ref|NP_867595.1| hypothetical protein-transmembrane prediction [Rhodopirellula baltica SH 1]

MS
FEVSKHVHMKRTRLLRLAPSEKPTLAGQVWDVAEIAPSAEMLFTQSI
PLVDKPVDRNSERKEQAVKIRKSHCIALVASSFALVVVAGIRDARDETASMQMAAASTTVSPAEPAT
YFVTRPVREQREKTVH
YTVTKPVYEQRTKTVH
YTQMTPVHETRTKRDDATGKTTQ
YTVCKMVPQPREKTVN
FTVCKMVPEQRTKHIS
YTVTRMVREERILD

>gi|32475337|ref|NP_868331.1| hypothetical protein RB8376 [Rhodopirellula baltica SH 1]

MTILRTGTNEKYAEGFEAAFGGGTPASKKT
AKKATKKA
AKKTSKKAKTQPAQSKKKS
AKKSVKKA
AKKSVKKAVKAKAT
PKKAVKKA
AKKSVKKSVKKT
AKKSVKKS
AKKAKKKS

>gi|32475659|ref|NP_868653.1| hypothetical protein RB8987 [Rhodopirellula baltica SH 1]

MRVAS
CELRVAS
CELRVAS
CEERGAR
SEERGAR
SEERGAR
SEERGARS

>gi|32476072|ref|NP_869066.1| hypothetical protein-signal peptide and transmembrane prediction [Rhodopirellula baltica SH 1]

MSSHLFETRIVPMKSSCAVVLTLV
CLLMSTGCVGPMWNA
CGPCGPGVG
CDSGACGDSCGG
CDSCDGCGELYIDPWINEPADCYDP
CDSCGNYNGQSCGK
CRAVFTGIKSIWGYK
RDGCGTG
CDTCAGGRPLFPVADLGGCDGCGG
CDSCSTG
CDSCSGG
CDSCSGGSYGHSSDGVILHEGETIISDSGHGHTSTVAARTAQNRRNSPARQIYRTRSVNGQQPTLAY

>gi|32476441|ref|NP_869435.1| RNA-binding protein [Rhodopirellula baltica SH 1]

MFREFRGRGKKMGLTGVSGWFTLRSQLSTTFGLVGLFWKRFPVKRARMREASCLVGDGLQGIRLVTNIYVGNLSFKATEE
ELRGAFEQYGEVSAVNIIMDRETGRSRGFAFVEMADAEGAKDAIENLNGHEIDGRSVTVNEARPREPR
SGGGGGGGG
YGGGGGGRG
RGGGGGG
YGGGGGNRGGG
YGGGGGG
YGGGGGGG
YGGRGDR

>gi|32476565|ref|NP_869559.1| hypothetical protein RB10689 [Rhodopirellula baltica SH 1]

MLVLVLVA
RQLVSSSV
RQFVSSSV
RQFVSSSV
RQFVSSSV
RQFVSSSV
RQFVSSSV
RQFVSSSV
RQFVSSSVLFFDYEYEHRSTT

>gi|32476936|ref|NP_869930.1| hypothetical protein RB11348 [Rhodopirellula baltica SH 1]

MTYPSDY
PEMPPYPFVPPF
DPPGSEPPYDPEDDDDGPIILAPPPL
PPPPVWPD
DYPPGTPDPG
PRPPWWPDDYPW
PPPPEEPVVIE
SPPPIHVWS
RLPPDHPYHV
PPGYSAPD

NLPPGHPGEAY
PGMPDYPVVHPGDWGDDWPSYPGILDD

>gi|32477026|ref|NP_870020.1| hypothetical protein RB11522 [Rhodopirellula baltica SH 1]

MTYPSDY
PEMPPYPFVPPF
DPPGSESPYDPEEEDDGPIILTPPPL
PPPPVWPD
DYPPGTPDPG
PRPPWWPDDCPW
PPPPEEPVVIE
SPPPIHVWP
RLPPDHPYHV
PPGYSAPD
NLPPGHPGEAY
PGMPDYPVVHPGDWGDDWPSYPGLYDD

>gi|32477033|ref|NP_870027.1| hypothetical protein RB11539 [Rhodopirellula baltica SH 1]

MWAMHHAVIPATAPSNEPDELGGSPGSSQSLLTNSQPHDQIHVLLQGSLDPLLSSEGSSVESLSLGRSEPSL
LSDALSLGRSEPSLGGSESEVESLLSEGWLLSLPDESPTLGSSLALGRSLSEPSDPDL
LVDSDAESEPEP
LDDVGALVESSLRES
LSESDVLVDR
ESESDPLP
LSDSDALT
LPDPESLPELDP
LPDAEPLWLPLAEA
LSEYDAEALAE

>gi|32477035|ref|NP_870029.1| hypothetical protein RB11542 [Rhodopirellula baltica SH 1]

MLDSLWESLPESLTDCDVESDCDWDSD
SESLDSESLSLLVDEFQLYEGKLSE
SDSLRDRLSDVLD
SDWLDSEF
DDSLVDDALSDVDD
CELLTDGVL
ADSLDCEV
DESLDEPESLD
CDSLDDESL
ADELDPDTLGRLTDALESFDSLV
DESLDWE
LDELFDESLDADDDPAELDDPLMLGRLIESLESLDSLF
DESLDCE
FDELSDESLDADDPSELDEPLILGRLVESLES
LESLLDESLD

>gi|32477418|ref|NP_870412.1| hypothetical protein RB12276 [Rhodopirellula baltica SH 1]

MSTATASMQE
CIQNCQE
CQTICADMLTSHCLVEGGDHVEQTHVKRMLD
CIAACGACIDFMSRNSENHSLYCRA
CAEICKA
CAESCEKVGGMEK
CVECCRT
CEQSCSSMAA

>gi|32477454|ref|NP_870448.1| hypothetical protein-transmembrane prediction [Rhodopirellula baltica SH 1]

MVDEPDWSPDSIGLASFLLR
NFLLGCLLFG
DLLLGCFLDGFLLY
SLLLGCLA
SGFLDALLRSLLLN
GFLLDCLA
SGLLGCLHALFSSAF
DGLLGCFFLGHVLPPRERNCC

>gi|32477816|ref|NP_870810.1| hypothetical protein RB13001 [Rhodopirellula baltica SH 1]

MLR
ASCFVLR
ASCFVLR
ASCFVLR
ASCFVLR
ASCFVLR
ASCFVLR
ASCFVLRARARARARARARARYPTPLFYFTLPPGGSSEARGGPSTENATISKCQFHFFNGIQSELPCTVAPPKQTTTVIA
HTLHLTPTPIPNLQCPLCNLQSIPFGPKPLSQLYPPRCSSIHSCTTQITAAMLSPLQFSG
MSCVVCQYCGALNPGRG