

1

PAE2867	-----	-----MSLEL	KEKESELPHD	EQIINDKWRS	KY----	TPIDA	YFKFHRQTV
MT_ACS1	-----	---MSKDTSVL	LEEKRVFKPH	YTVVEEAHIK	NW-----	---	EAELEK GK
Memar_1389	-----	---MAEDFNVK	LEEAKYYTPE	ASCREQSWV-	-----	GD	YNQAYREFLA
Memar_0039	-----	---MAETFAVK	LEE-KCYYPD	PRCKATAWT-	-----	PD	YTRAYSEFLR
MT_ACS2	-----	---MRGQLDAL	LREERIFNPD	EELAENS NIR	AWMDARGIGG		YDELLERASS
Mthe_1413	-----	---MVYEKAEV	SSQENVYRPA	SDLVENS NV	QWMKRRKGRS		EKELRAWCSE
P27095	---	MLKLAGK	EDKCLKTTFV	QDETRIFNPP	KELVEKSIVM	QWMKKKGFKT	EKEMRAWCSS
Mthe_1196	-----	---M	AEETAKTAVL	LEERRLFHPP	KELVENS NV	QWMKKKGFKT	EKEMREWC SK
Mthe_1195	MTVEECHKMA	EEKAATAVL	LEETRVFHP	KELAENS NV	QWMKKKGFKT		EREMRAWT GQ
Mthe_1194	-----	MAETAKTAVL	QEETRIFNTP	QWIEYSNSY	QWMKKKGFKT		EKEMREWCAQ
AF2389	-----	-MTMPGLES	LREERVFYPP	KELAENS NIK	QFMDKHG IKD		EDELKRKRAE
FOSS5188.y2	-----						
Dvul_0401	-----	-MSQERIESM	MDEKRHFAPP	ADSRGRAHV-	-----	SGEAA	REALVRRAAE
TTC0884	-----	---MDRLESV	LKEERVFYPS	EEFRKQAH I-	-----	KSEEE	YQRLYEESVR
STH881	-----	-MDSKQFAAL	LNERDRFPPP	AEFARANV-	-----	A	DESLYEQAAR
Dhaf_0467	-----	-MEEKNLEAL	LEEDRQFIPS	EEFHKNALI-	-----	Q	SADIYEQQQ-
CHY_0659	-----	---MSDNFEAL	LQESRIFEP	AEFKEKAKV-	-----	A	DLSLYEWAER

61

PAE2867	N--	LESFWES	VAKEL--EWF	KPWDKVL D-A	SNPPFYKWFV	GGRLNLSYLA	VDRHVKTWRK
MT_ACS1	D--	HENYWAE	KAERL--EWF	RKWRDVL D-E	SNRPFYRWFV	NGKINMTYNA	VDRWLDTDKR
Memar_1389	D--	PDGFWD C	VAKEL--EWF	QPWDRV K--E	WNYPYARWFL	NGKLNIT HNC	LDRHVYNQRR
Memar_0039	D--	PEGFWDR	VAREL--DWF	EPWDRVR --E	WNYPYAKWFI	NGKLNIT TSC	LDRHVAGGRK
MT_ACS2	D--	PEFWWDE	MASEL--EWF	RPYTRVL --E	WEPHARWFT	GGKFNIT YSA	LDRHVMGL-R
Mthe_1413	N--	YVEFWDE	MAKTY-ADWF	VPYEKVL --E	WNPPHARW FV	GGKCNVAHNA	LDRHARSWRR
P27095	DEHYLEFWDE	MAKTY-VDWH	KPYTKVMD-D	SEMPYFHWFT	GGEINIT YNA		VDRHAKGAKK
Mthe_1196	N--	YVEFWDE	MAKTY-ADWF	EPYKQVL --D	WKPPYAKW FV	GGKVN MAYNA	VDRHAKSWRR
Mthe_1195	H--	YIEFWDE	MAKTY-ADWF	EPYAQIL --E	WKPPYARW FV	GGKCNVA YNA	VDRHAKGAKK
Mthe_1194	N--	YLDFWDE	MAQTY-ADWF	KPYTQIL --E	WNPPYAKW FL	GGKCNVAHNA	VDRHAKSWRR
AF2389	N--	PEFWWSE	MAKEVGIEWF	SEPEKVL --E	WDPPYAKW FV	GAMYNIVHDA	LDKQAE-LRK
FOSS5188.y2	-----						
Dvul_0401	D--	PEGFWGE	RAAQL-IDWF	RPWDTVL DAD	MNEPRIEWFK	GGRLNVAHNC	LDRHVAGNRR
TTC0884	D--	PEGFWGR	VASEL--HWF	EPWRKVL --E	GDLPHPKW FV	GGKTNLSYNA	LDRHVKTWRR
STH881	D--	REGFWAA	QAERL--HWF	RRWDKVL --E	WNPPFAQW FL	GGKLNVA YNC	LDRHLQTHTR
Dhaf_0467	---	SLGFWEE	QAKQL--SWL	SPWEKTL --E	WNPPFAQW FV	GGKLNASANC	LDRHLQDWHR
CHY_0659	D--	FLGFWAD	AAKDI--EWF	LPFEKVL D-D	SDAPFYRWYT	GGKLNVS YNC	VDRHTKSFRR

121

PAE2867	NKLAIEWEGE	PVDENGYPTD	RRKLTYYDLY	REVN RVAYML	KQNFVKKGD	KITLYLPMVP
MT_ACS1	NQVAILYVNE	RG-----D	ERKLTYYELY	REVSRTANAL	KS-LGIKKGD	AVALYLPMCP
Memar_1389	NKVAIMWRGE	TEE-----E	ERILTYRQLF	QAVCRFANGL	KR-LGVGKGD	RVC IYMPVVP
Memar_0039	DKPAIVWHGE	SG-----E	KRILSYEALH	REVMRFANGL	SS-LGVEKGD	RVC IYMP LVP
MT_ACS2	NRVAYIWEGE	DG-----S	IRKLTYYDLY	REVNRIANAL	KD-MGVSRGD	RVSIYLP MIP
Mthe_1413	NKVAYYFVGE	PVG-----D	TRAITYYQLY	RDVNKL ANGL	KS-LGVKKGD	RVGIYLP MIP
P27095	DKVAYIWIPE	PTDQ-----P	VQKIT YGDLY	KEVNKFANGL	KS-LGLKKGD	RVSIYMP MIP
Mthe_1196	NKVAYIGVGE	PLG-----D	VRKFTYGDLY	REVNKFANGL	KS-LGVEKGD	RVSIYMP MIP
Mthe_1195	DKVAYIFVPE	PTDQ-----Q	VRKIT YLDLY	KAVNKFANGL	KS-LGVKKGD	RVSIYMP MIP
Mthe_1194	NKVAYYFVGE	PVG-----D	TKTIT YYQLY	QAVNKM ANGL	KS-LGVKKGD	RVSIYLP MIP
AF2389	NKVAYIWEGE	NG-----D	VRKIT YGELY	REVNKL ANAL	KE-LGVKKGD	RVAIYLP MIP
FOSS5188.y2	-----			---NKLSNAI	KS-MGIKKGD	RIMIYLP MIP
Dvul_0401	NKAAIIWQGE	PEE-----D	VRVLT YQMLY	DEVRRFAAVL	RR-MGVHKG D	RVSLYMP MIP
TTC0884	NKAAIVWEGE	PG-----E	ERVLT YHDLW	REVQRFANVL	KR-LGVKKGD	RVTIYLP MIP
STH881	TKAAIWEGE	PG-----D	ERVLT YLDLH	REVCRFANVL	RQ-MGVGKGD	RVTIYLP MIP
Dhaf_0467	NKAAIFEGE	NG-----D	SQVLT YQDLH	REVSKFANVL	KA-NGVNKG D	RVTIYLP MIP
CHY_0659	NKAAIFEGE	PG-----D	SKILT YQELY	REVNKFANVL	KK-LGVQKGD	RVTIYMP MIP

181

PAE2867	ELPITMLAAW	RIGAITSVVF	SGFSADALAE	RINDSQSRIV	ITADGFWRRG	RVVRLKEVVD
MT_ACS1	ELVVSM LACA	KIGAVHSV IY	SGLSV GALVE	RLNDARAK I I	ITADGT YRRG	GVIK LKPIVD
Memar_1389	EQIVAM LACA	RIGAIHTV VF	GGFGVSAL NQ	RITGTDAKV V	VTADV TYRRG	KAIP LKNIVE
Memar_0039	EQVVAM LACA	RIGAVHSV VF	GGFGPDAL AM	RINDAKAK VL	VTADVG YRRG	KTVP LRELAS
MT_ACS2	ELPIAM LACA	RIGAVHSV VF	SGFWAKAF RE	RAADAGAK VA	ITADAF YRRG	KVIK LKETLD
Mthe_1413	ELPVAM LACA	KIGAIHV VFV	SGFSAGAL RE	RINDAGAR VL	ITCDGS YRRG	KPIPI KAQAD
P27095	QLPIAM LACA	KLGVSHI VFV	SGFSSKGL MD	RAAHCGS RAI	ITVDGF YRRG	KPVPL KPNAD
Mthe_1196	ELPIAM LACA	KIGAIHSV VF	SGFSSKAY AD	RVIDAESK IS	ITVDGF WRRG	KIVEL KKQAD
Mthe_1195	ETPIAM LACA	KIGAIHSV VF	SGFSAGGL QS	RVLDAEAK VV	VTTDGF YRRG	KPLPL KPNVD
Mthe_1194	ELPITM LACA	KIGAIHSV VF	SGFSAGGL QS	RVTDAEAK VV	VTSDGF YRRG	KPLPL KPNVD
AF2389	ELPIAM LACA	KIGAIHTD VF	SGFSPMAL RD	RINDAEAK LL	ITADGF YRRG	SVIHL KEDAD
FOSS5188.y2	ELPISML ACA	KIGAIHSV VF	SGVSSLAL RD	RINDSEAK IL	ITCDGF YRRG	KTIPL KQSD
Dvul_0401	ELAVAM LACA	RIGAVHSI VF	AGFSAVSL QN	RIHDCEAK VV	VTADAVL RAG	RRIP LKNVD
TTC0884	EAAIAM LACT	RIGAVHSV VF	GGFSAGAL AD	RIKD AEAKVL	ITADGG FRRG	GIVPL KQNAD
STH881	EAAVAM LACT	RIGAIHSV VF	GGFSAESL RD	RINDSR SKVV	VTADGG WRRG	NIIRM KQIVN
Dhaf_0467	EAVISML ACA	RIGAPHSV VF	GGFSSEAL RD	RVIDAQAK AV	ITSDGS FRRG	NTIPL KDN TD
CHY_0659	EAVIAM LACT	RIGAPHSV VF	GGFSSQAL KD	RIDDAK AKLL	ITADGG YRRG	SIVEL KKNAD

241

PAE2867	AALEKATGVE	SVIVL PRLGL	KDV----PMT	E--GRDYWWN	KLMQGIPPNA	YIEPEPVESE
MT_ACS1	EAILQCPTIE	TTVVVKH-TD	IDI----EMS	DISGREMLFD	KLIEGEGDRC	--DAEEMDAE
Memar_1389	EAVVNAPSVE	RIVILRRDDD	KPV----ELH	P--EMEV DYY	DLMEGVEREC	--PAEPMDAE
Memar_0039	EALAHAPGVE	RVVVLRR-ET	PAV----ELD	P--DREVDYA	DLMARAAPDC	--PAEPM DSE
MT_ACS2	MVADEIPSLE	RVIVADR-MG	EDV----SMV	D--GRDIYWS	EAVEGMDDEC	--PCQELDPE
Mthe_1413	EALQDAPSVE	RQIVYRR-TG	QSI----EWK	D--GFDIWWH	ELVKNQ PDEC	--ETLQMDSE
P27095	EAAGGAPSVE	KIIVYKR-AG	VDV----SMK	E--GRDVWWH	DLVKGQSEEC	--EPVWVDPE
Mthe_1196	EAIQDAPT VK	HQIVYKR-TG	QDI----PWN	K--DRDIWWH	DLVKDQPAEC	--ETEQLDPE
Mthe_1195	EAVQNAPSVE	KVVVVKR-AG	LDV----PMK	E--GRDIWYH	ELVANQ PDEC	--ETEKM DSE
Mthe_1194	EAVQNAPSVE	KVVVVKR-VG	LDV----PMK	E--GRDIWYH	DLVKDQPAEC	--YTEELDPE
AF2389	KALEDAPSVE	KVIVARR-IG	LDV----PMK	E--GRDYWWD	DVTRNQPKEC	--ETEVM DAN
FOSS5188.y2	EALKDAQSIE	NVIVFKR-TG	EEV----PWN	A--NIDNWWH	AVTEKESDFC	--ETEIM DSE
Dvul_0401	EALRQCPSVE	KVVVVNR-GG	LEV----TME	E--GRDLWWH	EVMADRTLDV	DRPCEEMDAE
TTC0884	EALKDATSVE	HVVVVRRTG	EEV----PWT	P--GRDHWWH	ELMEAAPDRC	--DPESM EAE
STH881	EAVVDCASVE	KVIVVKR-IG	HESLVEHGWH	P--GRDHWYH	TLMRNAPVTC	--PVEEMDAE
Dhaf_0467	IALDGVDCVE	HVFVIQR-TK	QTV----QMK	E--ERDLWYH	EEMAKASPVC	--PAEPM DAE
CHY_0659	AALEGETTIE	KVVVVKR-TG	QEV----PMT	E--GRDYWYH	ELMADAALYC	--EPEQCDAE

301

**

PAE2867	HPSFILY TSG	TTGKPKGIVH	DTGGWAVH VY	ATMKWVFDIR	DDDIWFCTAD	IGWVTGHSYV
MT_ACS1	DPLFILY TSG	STGKPKGVLH	TTGGYMGVVA	STLEMTFDIH	NGDLWWCTAD	IGWITGHSYV
Memar_1389	DPLFILY TSG	TTGSPKGI VH	TCCGYMVGTY	YTTRHVFDVK	DNDIYWCTAD	PGWITGHSY G
Memar_0039	DPLFILY TSG	STGAPKGI VH	TCCGYAVGTY	YTTRHVFDVK	ENDVYWCTAD	TGWITGHSY I
MT_ACS2	DPLFILY TSG	TTGKPKGVVH	THGGYAVGVS	STYRFVFDVK	DQDIWWCLAD	IGWITGHSY I
Mthe_1413	DPLFILY TAG	AGGKPRGVVH	AHGGFCVGPA	YTTSWVFDIK	DTDVYWSTAD	IGWITGH TYI
P27095	HRLYILY TSG	TTGKPKGIEH	ATGGNAV GPA	QTLHWVFDLK	DDDVWWCTAD	IGWVTGHSY I
Mthe_1196	HRLFILY TSG	TTGKPKGIEH	AHGGYCVGVP	QTLHWVFDLK	EDDVWWCTAD	IGWITGHSY I
Mthe_1195	DRLFILY TSG	TTGKPKGIEH	VHGGYCVTPA	QTTHWVFDLK	DDDVWWCTAD	VGWITGHSY V
Mthe_1194	DRLFILY TSG	TTGKPKGIEH	AHGGFCVGPA	YTTAWALDVH	EEDVYWCTAD	CGWITGHSY V
AF2389	DILFILY TSG	TTGKPKGVMH	AHGGYAVGTA	ATLKFVLDLK	ENDVYWCTAD	IGWITGDSY I
FOSS5188.y2	DRLFILY TSG	TTGKPKGIEH	VHGGYSVGIA	NTL SFVFDLK	DDDVYWCAAD	IGWITGHSY I
Dvul_0401	DMLFILY TSG	STGKPKGVVH	TTGGYLT YAA	HTTQWVFDVQ	DDDVYWCTAD	IGWITGHSY I
TTC0884	EPLFILY TSG	STGKPKGVLH	TTGGYMTYVY	YTTKL VFDLK	DEDVYWCTAD	VGWITGHSY V
STH881	DPLFILY TSG	STGKPKGVLH	TTGGYLTQVA	ATTKYVFDLK	DEDVYWCSAD	IGWITGHSY V
Dhaf_0467	DMLFILY TSG	TTGKPKGVVH	TTGGYMGVVS	TTHR WVFDLK	EEDVYWCTAD	VGWITGHSY I
CHY_0659	DMLFILY SSG	TTGKPKGIQH	TTGGYLVGVH	TTFKYIFDYR	EEDIYWCTAD	IGWITGHSY I

361

PAE2867	VLGPLLMGAT	EVIYEGAPDY	P---QPDRWW	SIERYGVTI	FYTSPTAIRM	FMRYGEEWPR
MT_ACS1	VYGPLLLGTT	TLLYEGAPDY	P---DPGVWW	SIVEKYGVTK	FYTAPT AIRH	LMRFGDKHPK
Memar_1389	VYGPLLN GAT	CLIAEATPDY	P---DPGTYW	DLIEEYGVTI	FYTAPT AIRM	FMRVGEEWPD
Memar_0039	VYGPLEV GTT	VVLAEGTPDY	P---DPGAYW	RLVQDLGVTI	FYTAPT AIRM	FMRVGEEWPA
MT_ACS2	VYAPLIEGAT	SVIYEGAPDH	P---DPGRIW	SMVERYGVSI	LYTAPT TVRL	FVKYGD KWPE
Mthe_1413	VYGPLCLGAT	SVMYEGSPDY	P---DFGRWF	QIIEDYGVSV	IYTAPT AIRM	FMKEGEEWPR
P27095	VYAPLILGMT	SLMYEGAADY	P---DFGRWW	KNIQDHKVTV	LYTAPT AVR M	FMKQGA EWPD
Mthe_1196	VYGPLSLGAT	SIYEGSPDY	P---DFGRWW	SIIEEFGVNV	LYTAPT AIRM	FMRAGEQWPA
Mthe_1195	VYGPLCLGAT	SILYEGAPDY	P---DFGRWF	SIIQEHKVSV	FYTAPT AIRM	FMKAGEQWPQ
Mthe_1194	VYGPLCLGAT	SILYEGAPDY	P---DIGRWW	SIIEEYGVSV	FYTAPT AIRM	FMKAGDQWPK
AF2389	VYAPLILGAT	SVIYCGVPDY	P---KPDRWW	EIEEYGVTV	FYTAPT AIRM	FMRLGEEWVE
FOSS5188.y2	VYAPLILGVT	SVMYEGSPDY	PRLIDYGKW-	--IEKYKVS	FYTSPTAIRM	FMKFGESHIQ
Dvul_0401	VYGPLALGAT	SLMFEVPSW	P---SPDRFW	RIVEKFRVNI	FYTAPT VVRA	LMREGTDWTE
TTC0884	VYGPLLN GAT	TVMYEGAPNW	P---EPDRFW	RIVDKYGVTV	FYTAPT AIRS	FMKWGE GWPG
STH881	IYGPLANGAT	TLMYEGAPDY	P---DRGRIW	EIIQKYRVNI	FYTAPT LIRS	FMRWGE GWPS
Dhaf_0467	VYGPLANGAT	VLMYEGSPDY	P---NRDRFW	EIVEKYKVTI	LYTAPT AIRT	FMKWGPQYPQ
CHY_0659	VYGPLSNGAT	VVLYEGAPDW	P---QKDRFW	EIEEYRVNI	LYTAPT AIRT	FMRWGE KWPK

421

^ ^ *

*

PAE2867	KHDLSTLRII	HSVGE PINPE	AWRWAYRVLG	NEKVAFGSTW	WMTETGGIVI	SHAPGLYLVP
MT_ACS1	RYNLES LKIL	GTVGE PINPE	AWMWYRNIG	REKCP IIDTW	WQTETGMHLI	APLP---VTP
Memar_1389	RYNLSSLRIL	GSVGE PLNPE	AFEWFYRTIG	KDRCP IVDTW	WQTETGMHMV	TTMI---GEP
Memar_0039	KYDLSTLRVL	GSVGE PLNPE	AFEWYRAIG	GGRCPIVDTW	WQTETGMHMV	TTMI---GEA
MT_ACS2	KYDLRTLRL	GSVGNP INPE	AWMWYRTVG	GGRCPIMDTW	WQTETGMHII	TPLP---VTP
Mthe_1413	KYDLRSVRLM	GSVGEAMNPD	AFLWWRKHVG	NDWAPIMDTW	FQSETGCHVI	APLP---ITP
P27095	KYDLSSLRLL	GSVGE PINPE	AWMWYREHIG	RGELQIMDTW	WQTETGTFLN	SPLP---ITP
Mthe_1196	KYNMKCLRL	GTVGE PINPE	AWVYRKNIG	RDELQIMDTW	WQTETGTFIG	SPLP---ITP
Mthe_1195	KYDLSSLRLL	GSVGE PINPE	AWIWYRKYFG	SDRCPIMDTW	WQTETGCFV	SPLP---ITP
Mthe_1194	KYNLKSIRIL	ASVGE PLNPE	AYVWFRNIG	GGQAPI IDTW	WQTETGCHVI	APLP---MTP
AF2389	KHDLSSLRIL	GSVGE PINPE	AWYWYKXIG	KERCPIMDTW	WQTETGHFIV	TPLP---VTP
FOSS5188.y2	KHNLDSIRLL	GSVGE PINPE	AW-----	-----	-----	-----
Dvul_0401	KHDLSSLRVL	GSVGE PINPE	AWMWYHHTIG	KGRLP IVDTW	WQTETGGIMI	SGLPY--ATT
TTC0884	KHRLDSLRL	GTVGE PINPE	AWLWYHVI	KGRCP IVDTW	WQTETGGIMI	TTPLG--AHA
STH881	KYDLSSLRLL	GTVGE PINPE	AWMWYHKHIG	GERCP IVDTW	WQTETGAIMA	TPLPG--VVH
Dhaf_0467	SRDLSSLRLL	GSVGE PINPE	AWMWYKYIG	GERCP IVDTW	WQTETGMIMM	TPLPG--ITS
CHY_0659	GRDLSSLRLL	GTVGE PINPE	AWIWYHEHIG	GGRCPIVDTW	WQTETGMIMI	TPLPG--VIP

481

PAE2867	MKPGTNGPPL	PGFEVDVDE	NGNPA--PPGV	KGYLVIKKPW	PGMLHGIWGD	PERYIKTYWS
MT_ACS1	LKPGSVTKPL	PGIEADVDE	NGDPV--PLGK	GGFLVIRKPW	PAMFRTL FND	EQRYIDVYWK
Memar_1389	MRPGFAGKPI	PGVVADVDM	AGNPC--PPGT	GGLLVVREPW	PSMRTI WND	DERYCK--YWN
Memar_0039	MKPGFAGKPI	PGAVVDVDR	TGKPV--PPGT	GGFLVIREPW	PAMLR TVHGN	DERYCA--YWN
MT_ACS2	LKPGSAGKPF	PTVIADVDD	EGRSL--RGS	GGHLVIKTPW	PAMFRTL FRE	PERYVDAYWS
Mthe_1413	LKPGSPAFL	PGYNVDLLDV	NGRAV--GPGE	SGNIVLTAPW	PTMLRGLIYGE	PEKLREIYYD
P27095	LKPGSCTFPL	PGYDISILDE	EGNEV--PLGS	GGNIVALKPY	PSMLRAFWD	KERFMKEYWQ
Mthe_1196	LKPGSCTFAL	PGYSMDVWDE	AGKPV--PPGE	GGNIVILEPY	PSMLRDFYKD	PDYRFR TYWE
Mthe_1195	LKPGSPTFPL	PGFNTDIYDE	DANPV--PPGQ	GGNIVNPTPW	PSMLRAF YKD	PERYMKEYWQ
Mthe_1194	EKPGSVAFL	PGFNTDIYDE	DGNSV--PLGY	GGNIVQKTPW	PSMLRAFFRD	PERYMKEYWQ
AF2389	LKPGSATKPF	PGIEADVDFD	SGNSLYGKNV	GGYLV IKKPW	PGMLRGVWRN	PERYFKTYWE
FOSS5188.y2	-----	-----	-----	-----	-----	-----
Dvul_0401	LKPGSATQPL	PGVDAIVRP	DGSPA--GPNE	GGHLVIRKPW	PGMLRGIFGS	PERYRSTYFE
TTC0884	MKPGHAGKPF	FGVVEILDG	EHRPVENPDE	GGHLCITRPW	PSMLRTVWGD	PERFLQ QYFS
STH881	TKPGSTTRPF	PGIEMAVVNE	EGQEV--PPGA	GGYLVVKSPW	PSMLRTVWGD	PDYRVSTYFG
Dhaf_0467	LKPGSCTVPF	PGVRIEVVDS	AGHPV--PKGG	GGYLAIKEPW	PAMLRNIYGD	PARFERTYFG
CHY_0659	TKPGSATKPF	PGVEADVND	KGEPV--PPGQ	GGYLV LKKPW	PAMLR TLYGD	PERYKNTYWS

541

PAE2867	RF-----PG	MFYAGDYAIK	DKDGYIWWLG	RADEVIKVAG	HRLGTYELES	ALISHPAVAE
MT_ACS1	QI-----PGG	VYTAGDMARK	DEDGYFWIQG	RSDDVLNIAG	HRIGTAEVES	VFVAHPAAVE
Memar_1389	TV-----PG	CYTAADLAVK	DKDGYIMVIG	RSDDLIVVAG	HNIGTAEVES	ALVSHEAVAE
Memar_0039	TI-----PG	CYTAGDLAVK	DEDGYIMVLG	RADDLIVVAG	HNIGTAEVES	ALVSHDAVAE
MT_ACS2	TF-----PG	IYLSGDVARI	DEDGYFWIQG	REDDVLNVAG	HRISTAEVES	ALVSHPDVVE
Mthe_1413	YYWSI--KPG	IYLSGDRARR	DADGYWWILG	RIDDVLKVAG	HRISNAEVES	AALSHPNVAD
P27095	FYWDVPGRRG	VYLAGDKAQR	DKDGYFFIQG	RIDDVLVSVAG	HRIANAEVES	ALVAHPKIAE
Mthe_1196	TYWNV--RAG	TYLAGDKGRR	DEDGYFWIQG	RIDDVIKVAG	HRIGNSEVES	AAVSHPKVAE
Mthe_1195	MYWDI--RPG	VYLAGDKATR	DKDGYFFIQG	RIDDVLKVAG	HRISNAEVES	ALVSHPAVAE
Mthe_1194	MYWDI--KPG	TYLAGDKATR	DKDGYWWIQG	RIDDVLKVAG	HRISNAEVES	AAVSHPAVAE
AF2389	KF-----KD	VYLTGDAARV	DEDGYFWIQG	RLDDVLNVAG	HRIGNSEVES	ALVSHPAVSE
FOSS5188.y2	-----	-----	-----	-----	-----	-----
Dvu1_0401	RF-----PG	MYESGDGART	DTDGYFWIMG	RLDDVINVSG	HRMGTAEVES	ALVAHPSVAE
TTC0884	QH-----PG	VYFSGDGAKR	DKDGYMILG	RVDDVLNVAG	HRLGTMEIES	ALVAHPAAVE
STH881	RF-----GHQ	VYFTGDGAKR	DEDGYWIIIG	RVDDVINVSG	HRIGTMEVES	ALVDHPLVAE
Dhaf_0467	NW-----PG	VYFPGDGAKW	DKDGYFWILG	RVDDVINVSG	HRIGTMEVES	ALVDHPSVAE
CHY_0659	KF-----PG	WYFTGDGAKK	DEDGYFWILG	RVDDVINVSG	HRIGTMEVES	ALVEHPLVAE

601

PAE2867	SAVVGVPDAI	KGEVPIAFV	LKQGVAPSDE	LRKELREHVR	RTIGPIAEP	QIFFVTKLPK
MT_ACS1	AAVIGKADPI	KGEVIKAFVI	LKKGHKLNA	LIEELKRHLR	HELGPVAVV	EMVQVDSL
Memar_1389	AAVIGKPDAL	KGNTIKAFVI	LRNGQKPGEK	LRNDLVYHVR	MTLGPVAVV	EIDFVDSL
Memar_0039	AAVIGKPDPL	KGNTIKAFVT	LCVGVSPGDG	LTDDLARHVR	KSLGPVAVV	EIEFMDRL
MT_ACS2	AAVVGKPDIL	RGEETIAAFVT	LRDKVEPT-R	LKGVLRHVR	REIGPIASPS	YIEFVEDL
Mthe_1413	AAVIGRPDKV	KGENIILFV	LKEGINPSEE	LKKDIRNHVR	ATMGPIAMPS	EVYFVSAIP
P27095	AAVVGKPDDEV	KGESIVAFVI	LRVGNPESPE	LAKDAIAFVR	KTLGPVAAPT	EVHFVNDL
Mthe_1196	AAVIGKPDVP	KGEVIVFAI	LREGVQESSE	LKKDIAKHVR	ETLGPVAMPE	AVYFVKDVP
Mthe_1195	AAVIGKPDDEV	KGEVIVAFVI	LREGVKESED	LKKELAKHVR	SVLGPVAYPE	IVYFVKDVP
Mthe_1194	AAVIGKPDDEV	KGEVIVAFVI	LKEGVQESSE	LKKDIAKHVR	SVLGPVAYPE	IVYFVKDVP
AF2389	AAVVGKPEHEV	KGEAIV----	-----	-----	-----	-----
FOSS5188.y2	-----	-----	-----	-----	-----	-----
Dvu1_0401	AAVVGMPHAV	KGEAIVAFVT	LGADAEETEE	LRAELRAWVR	KEIGPIATPD	VLQFAEGL
TTC0884	AAVVGKPDDEV	KGEAIVAFVT	LKEGHTPSDA	LKEELRAHVA	KVIGPIARPD	EIRFTDAL
STH881	AAVIGRSHAV	KGQAITAFVT	LKEGRRGTPD	LVDELKQHV	TKIGALARPE	EIYFAADL
Dhaf_0467	AAVIGKNHEV	KGQALACFVT	LKEGIEITPD	LEDELKKHVA	KKIGALARPD	DIFFTAEL
CHY_0659	AAVIGKSHEV	KGQAIIAAFVT	LKEGVEGTPE	LVQELKQFVA	QKIGALARPD	DIFFTAEL

661

PAE2867	TRSGKIMRRL	LKAVATGAP-	-LGDVTTL	ETSVEEAKRA	YEEIKAEMAR	T
MT_ACS1	TRSGKIMRRI	LRAREEGED-	-LGDSTLEE	-----	-----	-
Memar_1389	TRSGKIMRRV	LKAQELGMD-	-PGDISTLEE	-----	-----	-
Memar_0039	TRSGKIMRRV	LKARELGMD-	-PGDISTLEE	-----	-----	-
MT_ACS2	TRSGKIMRRV	IKAIVRGED-	-VGDVSTLSN	PESVSMLEDR	VRTF-----	-
Mthe_1413	DRTGKPVRAV	IKAKALGAA-	-LGDTSVIN	KDAIDAIPAI	-----	-
P27095	TRSGKIMRRV	VKARALGNP-	-VGDISTLMN	PEAVDGIPKI	V-----	-
Mthe_1196	TRSGKIMRRV	IRAKAMGQP-	-VGDISTLAN	PEAVDAIPKI	V-----	-
Mthe_1195	TRSGKIMRRV	IKAKALGNP-	-VGDISALAN	PEAVDAIPLI	K-----	-
Mthe_1194	TRSGKIMRRV	IKAKALGKP-	-VGDISALAN	PESVENIPLI	V-----	-
AF2389	-----	-----	-----	-----	-----	-
FOSS5188.y2	-----	-----	-----	-----	-----	-
Dvu1_0401	TRSGKIMRRI	LRKIAAGATS	EFGDTSTLAD	PGVVS DLIEG	RLQLTGR---	-
TTC0884	TRSGKIMRRL	LRQIAAGEKE	IKGDTSTLED	RSVVERLKEG	A-----	-
STH881	TRSGKIMRRL	LRDIAEGRA-	-LGDTTTLAD	PNVVAQLRQQ	YESQEAG---	-
Dhaf_0467	TRSGKIMRRL	LRDIAEGRA-	-IGDTTTLAD	ASVVNTLKAN	YQE-----	-
CHY_0659	TRSGKIMRRL	LRDIAEGRA-	-LGDTTTLTD	PAVINKIKEQ	YKDEG-----	-

Figure S1

Alignment of ACS sequences used to construct the phylogenetic tree displayed in [Fig. 2](#). The four residues forming the acetate binding pocket ([71](#)) are indicated by asterisks, and two residues diagnostic of a strong preference for acetate as a substrate ([42](#)) are indicated by the symbol ^. Download [Figure S1, DOC file, 0.046 MB](#).

1408.46 TTVSNIAAALAED-GKKVLVVGCDPKADSTRNLVGRKIP-TVLDVFRKK--GPDNMELED
FS40622.45 TTVSNIAAALAED-GKKVLVVGCDPKADSTRNLVGRKIP-TVLDVFRKK--GPDNMELED
2826344 TTVSNIAAALAED-GKKVLVVGCDPKADSTRNLVGRKIP-TVLDVFRKK--GAENMKLED
150014349 TNVSNIAAALAES-GKKVMVIGCDPKADSTRNLMGRKIP-TVLDALRKN--GANNLELED
44732 TNVGNMAAALAED-GKKVLVVGCDPKADSTRMLMHGKIN-TVLDTFRDK--GPEYMKIED
L34 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILRGMMQTTIMDTLREE--GEDEITLEK
L48 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILRGMMQTTIMDTLREE--GEDEITLEK
K10 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMISRGMMQTTIMDTLREE--GEDEITLEE
ctd005.70 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILRGMMQTTIMDTLREE--GEDEITLEK
ctd005.15 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILRGMMQTTIMDTLREE--GEDEITLEK
ctd005.17 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILRGMMQTTIMDTLREE--GEDEITLEK
L37 TTTQNMAAALAHFNDEKVLHSCDPKADSTRMILRGMMQTTIMDTLREE--GEDEITLEK
L9 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILRGTMQTTIMDTLREE--GEDGITLEK
L53 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILRGMMQTTIMDTLREE--GEDEITLEK
1408.9 TTQQNTAAAMAHYFSGKVMIHGCDPKADSTRLLILGIKMQETIMDTLREL--GEDEVTLDK
L63 TTQQNTAAAMAHYFSGKVMIHGCDPKADSTRLLILGIKMQETIMDTLREL--GEDEVTLDK
L8 TTQQNTAAAMAHYLSKVMIHGCDPKADSTRLLILGIKMQETIMDTLREL--GEDEVTLDK
1408.31 TTQQNTAAAMAHYFGRKVMIHGCDPKADSTRLLILGIKMQETIMDTLREL--GEDEVTLDK
1408.44 TTQQNTAAAMAHYFSGKVMIHGCDPKADSTRLLILGIKMQETIMDTLREL--GEDEVTLDK
1772570 TTQQNTAAAMAHYHGKIVIHGCDPKADSTRLLVGGVAQTTIMDTLREL--GEDAVTAEN
19915040 TTQQNTAAAMAHYHGKIVIHGCDPKADSTRLLVGGVAQTTIMDTLREL--GEEAVTVDN
1854552 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
2622682 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
132663323 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
150011527 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
A1.2 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
B1.32 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
1408.30 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
118197434 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
A1.46 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
E1.54 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
44623 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
150014559 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
1537.15 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
1537.52 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
1537.102 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
L74 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
L7 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK
216814 TTTQNMAAALAHFNDDKVLHSCDPKADSTRMILGGKQTTMMDTLREL--GEGACTPDK

1408.46	IVFEGFGGVYCVESGGPEPGIGCAGRGVITAVDMLNRLGVFEELKPGVVIYDILGDV
FS40622.45	IVFEGFGGVYCVESGGPEPGIGCAGRGVITAVDMLNRLGVFEELKPDVVIYDILGDV
2826344	IVFEGFGGVYCVESGGPEPGVGCAGRGVITAVDMLNRLGAFEELKPDVVIYDILGDV
150014349	IVFKGFGETYCVESGGPEPGVGCAGRGVITAIIDILNRLEAFETIKPDVVIYDILGDV
44732	IVYEGFNGVYCVESGGPEPGVGCAGRGVITAVDMLDRLGVYDQLKPDVVIYDILGDV
L34	IMQTGYGGIRCVESGGPEPGVGCAGRGIITAINLMEEHDGYAD-DLDFVFFDFLGDV
L48	IMQTGYGGIRCVESGGPEPGVGCAGRGIITAINLMEEHDGYAD-DLDFVFFDVS GDV
K10	IMQTGYGGIRCVESGGPEPGVGCAGRGIITAINLMEEHDGYAD-DLDFVFFDVLGDV
ctd005.70	IMQTGYGGIRCVESGGPEPGVGCAGRGIITAINLMVEHDGYAD-DLDFVFFDVLGGV
ctd005.15	IMRTGYGGIRCVESGGPEPGVGCAGRGIITAINLMEEHDGYAD-DLGFVFFDVLGDV
ctd005.17	IMRTGYGGIRCVESGGPEPGVGCAGRGIITAINLMEEHDGYAD-DLDFVFFDVLGDV
L37	IVQTGYGGIRCVESGGPEPGVGCAGRGIITAINLMEEHDGYAD-DLDFVFFDVLGDV
L9	IMQTGYGGIRCVESGGPEPGVGCAGRGIITAINLMEEHDGYAD-DLDFVFFDVLGDV
L53	IMQTGYGGIRCVESGGPEPGVGCAGRGIITAINLMEGHDGYAD-DLDFVFFDVLGDV
1408.9	VMKIGFGGIKCVESGGPEPGVGCAGRGVITAINLMEEHVAYED-DVDFVFFDVLGDV
L63	VMKIGFGGIKCVESGGPEPGVGCAGRGVITAINLMEEHGAYED-DVDFVFFDVLGDV
L8	AMKIGFGGIKCVESGGPEPGVGCAGRGVITAINLMEEQGAYED-DVDFVFFDVLGDV
1408.31	VMKIGFGGIKCVESGGPEPGVGCAGRGVITAINLMEEHGAYED-DIDFVFFDVLGDV
1408.44	VMKIGFGGVVKCVESGGPEPGVGCAGRGVITAINLMEEHGAYED-DVDFVFFDVLGDV
1772570	VINTGFDGIKCVESGGPEPGVGCAGRGVITAINLMEEHGAYSE-DLDFIHFDVLGDV
19915040	VVNTGFKGIRCVESGGPEPGVGCAGRGVITAINLMEEHGAYSD-DLDFVHFDFVLGDV
1854552	VIETGFGGIKCVESGGPEPGVGCAGRGVITAITLMEHGVYEN-DLDFVFFDVLGDV
2622682	VIETGFGGIKCVESGGPEPGVGCAGRGVITAITLMEHGVYED-DLDFVFFDVLGDV
132663323	VIEVGFGGVVKCVESGGPEPGVGCAGRGVITAITLMEQHGVYED-DLDFVFFDVLGDV
150011527	VIETGFCGVRV CVESGGPEPGVGCAGRGVITAITLMEQHGVYED-NLDFVFFDVLGDV
A1.2	VVKIGFGGIRCVESGGPEPGVGCAGRGVITAINLMEEELGGYPD-DLDFLFFDVLGDV
B1.32	AGKIGFGGIRCVESGGPEPGVGCAGRGVITAINLMEEELGGYPD-DLDFLFFDVLGDV
1408.30	IRKVGFGGILCVESGGPEPGVGCAGRGVITAVNLMIEELGGYPD-DLDFLFFDVLGDV
118197434	IRKVGFGGILCVESGGPEPGVGCAGRGVITAVNLMIEELGGYPD-DLDFLFFDVLGDV
A1.46	VRKVGFGGILCVESGGPEPGVGCAGRGVITAVQMMIDLGGYPD-DLDFLFFDVLGDV
E1.54	VRKVGFCGILCVESGGPEPGVGCAGRGVITAVQKMMELGGYPD-DLDDLIFDVLGDV
44623	VRKIGFKDILCVESGGPEPGVGCAGRGVITAVDMMRELEGYPD-DLDNLFFDVLGDV
150014559	VRKQGFGGIYCVESGGPEPGVGCAGRGVITAVDTMRELEGYPD-DLDNLFFDVLGDV
1537.15	VRMIGFKDIHCVESGGPEPGVGCAGRGVISAQSMIDLEGYPD-DLDYLFFDVLGDV
1537.52	VRMIGFKDIHCVESGGPEPGVGCAGRGVISAQSMIDLEGYPD-DLDYLFFDVLGDV
1537.102	VRMIGFKDIHCVESGGPEPGVGCAGRGVISAQSMIDLEGYPD-DLDYLFFDVLGDV
L74	VRMIGFKGIHCVESGGPEPGVGRAGRGVISAQSMIDLEGYLD-DLDYLFFDVLGDV
L7	VRMIGFKDIHCVESGGPEPGVGCAGRGVISAQSMVDLEGYQM-ILIISSLMFLEM-
216814	VIYKGYGGVDCVEAGGPPAGACGGYVVGETVKLLKELNAFDE--YDVILFFDVLGDV

Figure S2

Alignment of partial dinitrogenase reductase (NifH) sequences that were used to construct the NifH phylogenetic tree in [Fig. 4](#). Sequences from the Lost City and other hydrothermal vent environments are identified by their clone names; all other sequences are identified by their GenBank accession numbers. Download [Figure S2, DOC file, 0.029 MB](#).