HAITAO ZHANG AND GEORGE T. JAVOR*

Department of Biochemistry, Loma Linda University School of Medicine, Loma Linda, California 92354

Received 19 June 2000/Accepted 1 August 2000

The open reading frame at 86.7 min on the *Escherichia coli* chromosome, "yigC," complemented a *ubiD* mutant strain, AN66, indicating that yigC is the *ubiD* gene. The gene product, a 497-amino-acid-residue protein, showed extensive homology to the UPF 00096 family of proteins in the Swiss-Prot database.

The biosynthetic steps of the ubiquinone (coenzyme Q_8 [CoQ₈]) pathway in *Escherichia coli* have been known for many years (11, 16). However, the identities of some ubiquinone biosynthetic genes remained hidden, even after the complete nucleotide sequence of E. coli was published. This was due partly to a lack of sequence data on ubiquinone biosynthetic genes in other organisms and partly to the multiplicity of open reading frames at the suspected locations. Thus, at 86.7 min on the E. coli chromosome, where the ubiD gene was mapped (3), there were no fewer than seven unidentified, hypothetical genes. A systematic search of these some years ago failed to locate the ubiD gene (R. Meganathan, personal communication). This was likely due to the fact that the some of the early gene assignments were incorrect. Recently the identities of the ubiE (8) and ubiF (7) genes were published, leaving ubiD the last of the known ubi genes to be identified.

We prepared a PUC 18 plasmid library which contained chromosomal fragments of *E. coli* K-12 strain AN256, the isogenic *ubiD*⁺ strain of the *ubiD* mutant strain AN66. Strain AN66 (*thr-1 leuB6 ubiD410*) (3) was obtained from the *E. coli* Genetic Stock Center, New Haven, Conn., and strain AN256 (*thr-1 leuB6*) (9) was from C. F. Clarke's laboratory. Chromosomal fragments, 3 to 12 kDa in size, were obtained by partial digestion with *Sau3*AI restriction enzyme. Competent AN66 cells were electroporated in the presence of this plasmid library, and ampicillin-resistant transformants that could grow on plates which contained minimal medium (6), ampicillin (100 µg/ml), leucine and threonine (20 µg/ml each), and succinate (3 mg/ml) were isolated.

In our hands, AN66 cells spontaneously acquired the capacity to utilize succinate, at a frequency of 0.01%. Thus, when competent cultures of AN66 were transformed with the plasmid pBR322, approximately 300 transformed colonies which grew on succinate were obtained. These transformed revertant colonies (which grew faster on succinate plates than the *ubiD*⁺ strain AN256) seriously interfered with the identification of *ubiD* gene-harboring transformants.

Therefore, several cycles of transformations were carried

out to enrich the transformant population with *ubiD* genecontaining plasmids. This was done by recovering all transformed colonies, growing them together in Luria-Bertani medium with ampicillin, and extracting their plasmids. Competent cells were transformed by this preparation, and the procedure was repeated again. After two cycles, besides the 300 or so transformed revertant colonies, a strong haze was also seen on the succinate-containing selection plate.

Cells from this haze were cultured, and their plasmids were extracted. This plasmid preparation produced 1.3×10^5 transformed colonies that were able to grow on succinate as the sole carbon source. One of these was isolated and named AN66p522. Colony sizes of transformed cells on succinate plates were comparable to those of the *ubiD*⁺ strain AN 256.

The ubiquinone contents of strains AN256, AN66, and AN66p522 were determined by a method described earlier (17). Cell preparation included the growth of a 500-ml culture of strain AN256 in Luria-Bertani medium plus glucose (0.3%, wt/vol) and identical volumes of AN66 and AN66p522 in brain heart infusion broth plus glucose (0.3%, wt/vol). At an A_{600} of 0.9 to 1.0, the cells were harvested, washed with distilled water, and lyophilized. We found 0.29 nmol of CoQ₈ per mg of dried cells of strain AN256, 0.05 nmol of CoQ₈ per mg of dried cells of strain AN66, and 0.73 nmol of CoQ₈ per mg of dried cells of strain AN66p522. Thus the chromosomal fragment on plasmid p522 fully complemented the ubiquinone deficiency of AN66.

Sequencing of this chromosomal fragment showed that it was 2,595 bp long. It started near the end of the open reading frame $yigW_2$, 146 bases downstream from the end of the rfaH gene (a regulatory gene of lipopolysaccharide, sex factor, and hemolysin genes, oriented in the opposite direction from $yigW_2$) and ended 252 bases into the *fre* gene. The only other open reading frame located between rfaH and *fre* was yigC. It is our contention that this 1,491-base segment, immediately upstream from the *fre* gene at 86.7 min on the *E. coli* chromosome, is the *ubiD* gene (Fig. 1).

Until recently, the *fre* gene, coding for NAD(P)H flavin oxidoreductase, was designated *ubiB* (5). However, the true



FIG. 1. Composition of the 2,595-bp-long E. coli chromosomal fragment in p522, showing the location of the ubiD gene.

^{*} Corresponding author. Mailing address: Department of Biochemistry, Loma Linda University School of Medicine, Loma Linda, CA 92354. Phone: (909) 796-7311, ext. 48663. Fax: (909) 824-4887. E-mail: gjavor@som.llu.edu.

J. BACTERIOL.

6244 NOTES

	1									100
1.	MDAM	KYNDLRDFLT	LLEQQGELKR	ITLPVDPHLE	ITEIADRTLR	AGGPALLF	ENPKG	YSMPVLC	NLFGTPKRVA	MGMGQEDVSA
2.	1	M SFRDLPEFLE	K FLEKNGQLKF	TTARUDBDIE	ITEISRRVLA	AQGGPALLI	ENVIKADGSE	STIPVV	NEYASIKRIC	U IGENLKSPAL
3.	MGY	KYRDLHDFIK	DIFFECTIVE	TKEPLSPILE	TTEVTDEVCK	MPGGGKALLE	ENPKG	YRTRVLT	NEYGSEKRIK	KALGYEN
4.5	MGI	AYEDLREFIG	RLEDKGELAR	VKHEVSPILE	MSEVADRTVK	AGGKALLF	ERPKG	YDIPVFM	NAFGTERRMK	LALEVER
6.			MLRR	VSVEASPILE	IPEILRRIMY	R-GSGYAVLF	EKVKG	HEGFRIAG	NIFCSLDVVR	QALGVER
7.		RSLVD	YLRSQHELID	IHVPVDPHLE	IAEIHRRVVE	REGPALLE	НQ	VKGSPFPVLT	NLFGTRRRVD	LLFPDL
8.	1	M AMSSRRLVS	LIRSQNDLID	IFAPVDPYLE	LPEIHRRVIE	NQGPALLF	HN	VQGASFPVLT	NLFGTQKRVD	QIFSKV
9.		MSFLRRHIS	LFRSQKQLID	VFAPVSPNLE	LAEIHRRVIE	DQGPALLF	HN	VIGSSFPVLT	NLFGTKHRVD	QLFSQA
10.	м	MRDFLK	DEKKHDEEKI	ADIPLEVDES.	LCASABAASN	LODKSDALLE	NNT YC	VHNARTAM	NUTCONDNHA	DDDKIP
11.	M	AYDDLRSFLD	TEEKEGOLLE	TTDEVLEEPD	LAAAANATGR	IGENAPALHE	DNVKG	FTDARIAM	NVHGSWANHA	LALGLPK-~~
12.	MRKLN	PALEFRDFIO	VEKDEDDEIE	ITEEIDPNLE	VGAIMRKAYE	SHLPAPLE	KNLKG	ASKDLFSILG	CPAGLRSKEK	GDHG
14.		MEDASLKS	FLEKVGYRV-	VDRTLSREYE	VARLIAETQ-	GC-GPPLL	ARIEG	VRQPVAV	NVVDTREKLY	KALGV
15.		MRE	IINKLNP-II	IDKADKK-FG	VSRILKKYD-	GKPVYI	KDVNG	FEVVG	NL-CSRETLS	KIFNV
16.		MVMKMLRE	IVESFEDLVV	IDKPVKKELE	LTKFLLKYK-	DKPVLF	KDVEG	WEVAG	NLWSSRERIA	KFLNT
17.		MRNF	LDKIGEEALV	VEDEVSTSFE	AASILREHP-	RDLVIL	KNLKE	SDIPVIS	GLCNTREKIA	LSENC
18.	MEDDESCOD	VIADLOBIID	RUEALCREVE	VRSEVDERHD	LAGTAARFE-	GGPOAVIE	EKVAG	HAYPVFV	GLYWSBELLG	ALFDO
Consensus	MERDESGSER		.lel.	iev.p.le		a.lf	q	pv	nl.gs	
00110011000							2	•	5	
	101									200
1.	LREVGKLLAF	LKEPEPPKGF	RDLFDKLPQF	KQVLNMPTKR	LRGAPCQQ	KIVSGDDVDL	-NRIPIMTCW	PEDAAPLI-T	WGLTVTRGPH	-KERQNL
2.	LRELGALLAF	LKQPQIPESF	KETLSILPLA	KRIFSMSPKT	IAKGACHE	VVIVKPNINI	LPIQKCW	PEDISPLI-T	WGIVVTKGPT	~KDRVDHYNL
3.	LEDLGKKLAL	LQQPKPPKKI	SQAIDFGKVL	FUVLKAKPGR	NFFPPCQE	EVINGD-INT	-EDUBTIKCM	PGDAGKII-T	FGOVITKDEE	-SGIBNV
4. 5	LEEIGERLLS	ALEFR-PSSF	MDALKGVGML	KDFMSFIPK-	KTGKAPC	KEVVAESL	-DKFPILKCW	PKDAGRFI-T	FPVVITKDPE	-TGEMNA
6.	LEVIGERLFE	PLKGPPPLGI	GGKLRSLGEV	LSLGKYMPKA	VGRAGFTA	NVLEGREASF	-NLIPAFKVW	PKDGGRYL-T	YALVHVRDPV	-RGVMNM
7.	SSDLFEQIIH	LLSSPPSF	SSLWKHRSLF	KRGISALGMR	KRHLR-PS	PFLYQDAPNL	-SQLPMLTSW	PEDGGPFL-T	LPLVYTQSPE	-NGVPNL
8.	PKGLIPQVIH	LLSSPPKL	SQLWKHRNLL	LRGLS-LGLR	KARFL-KF	PHKKMASVDL	-HQLPMLTSW	PEDGGAFL-T	LPLVYTESPS	-SKIPNL
9.	PDNLIARVAH	LISSTPKL	SSLWKSRDLL	KRISS-LGLK	KARFR-RF	PFVSMSSVNL	-DHLPLLTSW	PEDGGAFL-T	LPLVYTESPT	-LTTPNL
10.	IESLQQRMQA	FLHFNAPKNF	TEGLKVLKDL	WDLRHIFPKK	TTRPK-DL	11KQDKEVND N-ETTEDINE	-LDBPVLKTW	EKUGGAFI-T	MGQVITQSLD	-HQKKNL
11.		DIPV	KEOVEEFARR	WDAFPVAPER	BEEAPWRE	NTOEGEDVDE	FSVLPLFRLN	DGDGGFYL-D	KAAVVSRDPE	DRDDFGKONV
13.	BIAH	HLGLDPKTTI	KEIIDYLLEC	KEKEPLPPIT	VPVSSAPCKT	HILSEEKIHL	-QSLPTPYLH	VSDGGKYLQT	YGMWILQTPD	KKWTNW
14.		TGD	SEAYAKIVDS	TM-RPGRLEY	VDKPPLDE	MPEGF	-EGLPAARFY	EGEAGLYL-S	SGIVIACY	EGVCNA
15.		K-K	EDFIFFMLDA	ME-KEKEGKL	KINNKLKE	KYIVEIPENI	-KNWPIPIYY	EKDAGAYI-T	SGVVVVYDKD	YGYNL
16.		D-N	KGLLELLYEA	ME-KPKPFSV	VEKAEFLK	NREKVNL	-LELPIPKYY	PKDGGPYL-T	SAMVIAKK	EFVNV
17.		R-V	HEITHRIVEA	ME-NPTPISS	VGGLDGYR	SGRADL	-SELPILRHY	RRDGGPY1-T	AGVIFARDPD	TGVRNA
18.		E-V	SELTQKIIEA ALDOUVAAST	SU-NPIKVDK	LVVADGPV	LEVTEAEVOL	-STUPIDINI	LEDGGPYE-D	AGVVFARDPE	TGVRNA
Ly. Consensus			ADFQRVAASI	K5WQ5AFVDF	L VVADGEV		lPi	DgG.v. t	.a.vdp.	N.
Consensus								5 1	·	
	201	MINT TAIDING CH	DCCM DYOFN	67 NUNCED PD	11 11 11 11 11 11 11 11 11 11 11 11 11	mticăvă	געפֿס זיייטרע	PACIT P	CTET TO T	300
1.	201 GIYROQLIGK	NKLIMRWLSH	RGGALDYQEW	CAAHPGERFP -KTKKKELFP	VSVALGADPA	TILGAVT	PVPDTLSEYA PTPENVSEYN	FAGLLR	GTKTEVVKCI NKKIELVOCK	300 SND-LEVPAS TID-LKVPAH
1. 2. 3.	201 GIYRQQLIGK GIYRMQVVSE GVYRLOLOSK	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV	RGGALDYQEW RGGAEHHKRW RGGARHIRK-	CÄAHPGERFP -KTKKKELFP -ÄAEOGKKEE	VSVALGADPA AAVVIGANPV VATALGVDPL	TILGAVT ITLAAVT IIMAAAT	PVPDTLSEYA PIPENVSEYN PIPVDLSEWL	FAGLLR FAGLLG FAGLYG	GTKTEVVKCI NKKIELVQCK GSGVALAKCK	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAD
1. 2. 3. 4.	201 GIYRQQLIGK GIYRMQVVSE GVYRLQLQSK GLYRLQVLDK	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KDGNHHYWK-	CAAHPGERFP -KTKKKELFP -AAEOGKKLE -AKRLGKKLE	VSVALGADPA AAVVIGANPV VATALGVDPL VATAIGGEPP	TILGAVT ITLAAVT IIMAAAT LPYVASA	PVPDTLSEYA PIPENVSEYN PIPVDLSEWL PLPPEVDEYL	FAGLLR FAGLLG FAGLYG FAGIIM	GTKTEVVKCI NKKIELVQCK GSGVALAKCK ERPVELVKGL	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAD TVD-LEVPAN
1. 2. 3. 4. 5.	201 GIYRQQLIGK GIYRMQVVSE GVYRLQLQSK GLYRLQVLDK GMYRMQVFDG	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KDGNHHYWK- KHGAEHFRKM	CAAHPGERFP -KTKKKELFP -AAEOGKKE -AKRLGKKE -AKRLGKKE	VSVALGADPA AAVVIGANPV VATALGVDPL VATAIGGEPP VAVAIGVDPA	TILGAVT ITLAAVT IIMAAAT LPYVASA TLYAATA	PVPDTLSEYA PIPENVSEYN PIPVDLSEWL PLPPEVDEYL PLPSGISEFM	FAGLLR FAGLLG FAGLIM FAGFIR	GTKTEVVKCI NKKIELVQCK GSGVALAKCK ERPVELVKGL KERLKVTECE	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAD TVD-LEYPAN TVD-LLVPAN
1. 2. 3. 4. 5. 6.	201 GIYRQQLIGK GIYRMQVVSE GVYRLQLQSK GLYRLQVLDK GMYRMQVFDG GVYRVMIAGD	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH KEGVVHWQIH	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KDGNHHYWK- KHGAEHFRKM KRGMQAQQDS	CAAHPGERFP -KTKKKELFP -AAEQGKKLE -AKRLGKKLE -AEKGGGKIE -VEKGERRIP	VSVALGADPA AAVVIGANPV VATALGVDPL VATAIGGEPP VAVAIGVDPA AALVIGSDPG	TILGAVT ITLAAVT IIMAAAT LPYVASA TLYAATA TLITGAM	PVPDTLSEYA PIPENVSEYN PIPVDLSEWL PLPPEVDEYL PLPSGISEFM PVPYPIDKHL	FAGLLR FAGLLG FAGLYG FAGIIM FAGFIR FAGVR	GTKTEVVKCI NKKIELVQCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAD TVD-LEVPAN TVD-LLVPAN PNG-IHVPAN
1. 2. 3. 4. 5. 6. 7.	201 GIYRQQLIGK GYRLQLQSK GLYRLQLQSK GMYRMQVFDG GVYRVMIAGD GMYRMQRFDK	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH KEGVVHWQIH ETLGLHFQIQ DWCL WEQIO	RGGALDYQEW RGGAEHHKRW KGGARHLRK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE-	CAAHPGERFF -KTKKKELFP -AAEOGKKLE -AKRLGKKLE -AEKGGGKIE -VEKGERRIP -AEQKKQNLF	VSVALGADPA AAVVIGANPV VATALGVDPL VATALGCPPP VAVAIGVDPA AALVIGSDPG VTVFLSCNPF	TILGAVT IILGAVT IIMAAT LPYVASA TLYAATA TLLTGAM LILSAIA	PVPDTLSEYA PIPENVSEYN PIPVDLSEWL PLPPEVDEYL PLPSGISEFM PVPYPIDKHL PLPENVPELI	FAGLLR FAGLYG FAGIIM FAGFIR FAGVVR FCSFLQ	GTKTEVVKCI NKKIELVQCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEKND CSKLHV-KND	300 SND-LEVPAS TID-LEVPAH TVD-LEVPAN TVD-LEVPAN PNG-IHVPAN PQSGHPLLOP
1. 2. 3. 4. 5. 6. 7. 8.	201 GIYRQQLIGK GIYRMQVVSE GVYRLQLQSK GLYRLQVLDK GMYRMQVFDG GVYRVMIAGD GMYRMQRFDK GMYRMQRFDK	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH KEGVVHWQIH ETLGLHFQIQ DTLGLHFQIQ DTLGLHFQIQ	RGGALDYQEW RGGAEHHKRW RGGARHIRK- KHGAEHFRKM KRGMQAQDDS KGGGAHFFE- KGGGMHFYE-	CAAHPGERFF -KTKKKELFP -AAEQGKKLE -AKRLGKKLE -AEKGGGKIE -VEKGERRIP -AEQKKQNLP -AEQKNONLP	VSVALGADEA AAVVIGANEV VATALGVDEL VATALGCEEP VAVALGVDEA AALVIGSDEG VTVFLSGNEF VTVFLSGNEF VTVFLSGNEF	TILGAVT IILAAVT IIMAAT LPYVASA TLYAATA TLITGAM LILSAIA LILSAIA LITSAIA	PVPDTLSEYA PIPENVSEYN PIPVDLSEWL PLPPEVDEYL PLPSGISEFM PVPYPIDKHL PLPENVSELL PLPENVSELL	FAGLLR FAGLLG FAGLIM FAGIIR FAGVR FCSFLQ LCTFLQ	GTKTEVVKCI NKKIELVQCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEKH GSKLHY-KND GAKLLY-KKT	300 SND-LEVPAS TID-LEVPAH TVD-LEVPAD TVD-LLVPAN PNG-INVPAN PQSGHPLLCD PDTPHPLLYD NDHPHPLLYD
1. 2. 3. 4. 5. 6. 7. 8. 9.	201 GIYRQQLIGK GIYRQQVJSE GVYRLQLQSK GMYRMQVFDG GYYRVMIAGD GMYRMQRFDR GMYRMQRFDR GMYRRVQRFDR GMYRLQVYDK	NKLIMRWLSH NKLIMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH ETUGLHFQIQ DTUGLHFQIQ NTMGLHFQIQ NHUGLHWOIH	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KHGAEHFRKK- KHGAEHFRKK- KGGGAHFYE- KGGGMHFYE- KGGGMHFYE- KDSOLFFHE-	CAAHPGERFP -KTKKKELFP -AAEOGKKLE -AKRLGKKLE -AEKGGKIE -VEKGERRIP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP	VSVALGADPA AAVVIGANPV VATALGVDPL VATALGCPP VAVAIGVDPG VAVAIGVDPG VTVFLSGNPF VSVFLSGNPF VSVFLSGDPF VSTAIGGDLL	TILGAVT ITLAAVT IIMAAAT LPYVASA TLYAATA TLITGAM LILSAIA LTLSAIA LTLSAIA LTLSAIA YTWCATA	PVPDTLSEYA PIPENVSEYN PIPVDLSEWL PLPSGISEML PUPSGISEML PLPENVPELL PLPENISELL PLPENVSELL PLPENVSELL	FAGLLR FAGLYG FAGIIM FAGFIR FAGYVR FCSFLQ LCTFLQ FATFLQ LYGFMR	GTKTEVVKCI NKKIELVQCK GSGVALAKCK ERPVELVKGL GEGLPVYR-L NKKLSFVEKH GSKLHY-KNT GAKLLY-KKT GKKARVMPCL	300 SND-LEVPAS TID-LEVPAH TVD-LEVPAD TVD-LEVPAN PNG-INVPAN PQSGHPLLCD PDTPHPLLYD NDHPHPLLYD SNS-LSVPSD
1. 2. 3. 5. 6. 7. 8. 9. 10. 11.	201 GIYRQQLIGK GIYRMQVVSE GVYRLQLQSK GLYRLQVDK GYRVQLDK GMYRMQRFDG GMYRMQRFDR GMYRVQRFDQ GMYRLQVVDK GIYRMQVKGK	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH KEGVVHWQIH ETLGLHFQIQ DTLGLHFQIQ NTMGLHFQIQ NHLGLHWQIH DRLGIQPVPQ	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KDGNHHYWK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KGGGMHFYE- KGGGMHFYE- KDSQLFFHE- HDIAIHERQ-	CAAHPGERFP -KTKKKELFP -AAEOGKKLE -AKRLGKKLE -AEKGGKLIP -AEQKKONLP -AEQKKONLP -AEQKKONLP -YAKAKVKMP -YAKAKVKMP	VSVALGADPA AAVVIGANPV VATALGVDPL VATALGVDPL VAVAIGOEPP VAVAIGSDPG VTVFLSGNPF VTVFLSGNPF VSVALGGDLL VSTAIGGDLL VTTALGCEPV	TILGAVT ITLAAVT IIMAAAT LPYVÄSA TLYGATA LILSAIA LILSAIA LTLSAIA ITSAIA ITTAAST	PVPDTLSEYA PIPENVSEYN PIPULSEWL PLPEVDEYL PUSSESEM PVPYPIDKHL PLPENVSELL PLPENVSELL PLPYGIVELM PLYYGIVELM PLYYGIVELM	FAGLLR FAGLYG FAGIIM FAGFIR FAGVVR FCSFLQ LCTFLQ FATFLQ LYGFMR MAGAIQ	GTKTEVVKCI NKKIELVQCK GSGVALAKCK ERPVELVKGL GEGLPVYR-L NKKLSFVEKH GSKLHY-KNT GKKARVMPCL GEPYRIVKSK	300 SND-LEVPAS TID-LEVPAH TVD-LEVPAD TVD-LEVPAN PNG-IHVPAN PQSGHPLLCD PDTPHPLLYD NDHPHPLLYD SNS-LSVPSD LSD-LOVPWG
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12.	201 GIYRQQLIGK GIYRMQVVSE GVYRLQLQSK GLYRLQVLDK GMYRMQVFDG GMYRMQVFDG GMYRMQRFDR GMYRVQRFDQ GMYRLQVIGT	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH ETLGLHFQIQ DTLGLHFQIQ NTMGLHFQIQ NHLGLHWQIH DRLGLWQIH DRLGIQPVPQ NRLAFHPA-M	RGGALDYQEW RGGAEHHKRW RGGARHIRK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KGGGMHFYE- KGGGMHFYE- KDSQLFFHE- HDIAIHIRQ- HDIAIHIRQ-	CAAHPGERFE -KTKKKELFP -AAEQGKKLE -AKRLGKKLE -EKGGGKIE -VEKGERRIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEERGINIP -AEERGINIP -AEEKGEDIP	VSVALGADPA AAVVIGANPV VALAIGGEPP VAVAIGVDPA AALVIGSDRG VTVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSTALGCPV IAITLGNDPV	TILGAVT ITLAAVT IPVVASA TLYAATA TLIGAM LILSAIA LTLSAIA YTWGATA YTWGATA MAIVAGM	PVPDTLSEYA PIPENVSEYN PIPPEVDESUN PLPPEVDEYL PLPEVDEYL PLPENVPELL PLPENISELL PLPENISELL PLPENISELL PLPENISEL PLYDQSEYE PMAYDQSEYE	FAGLLR FAGLYG FAGIIM FAGFIR FAGFIR FCSFLQ LCTFLQ LCTFLQ LYGFMR MAGALR	GTKTEVVKCI NKKIELVOCK GSGVALAKCK KERLKVTECE GEGLPVYR-L NKKLSFVEKH GSKLHY-KND GAKLLY-KKT GKKARVMPCL GEPYRIVKSK GAPAPIATAP	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LEVPAN PVD-LLVPAN POSGHPLLCD PDTPHPLLYD NDHPHPLLYD SNS-LSVPSO LSD-LDVPWG LTG-FDVFWG
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13.	201 GIYRQQLIGK GIYRQQVSE GVYRLQUSK GLYRLQVLDK GMYRMQVFDG GWYRVMIAGD GMYRMQRFDR GMYRVQRFDR GMYRLQVFDK GIYRLQVIGT STARGMVVDD	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH KTGGHHQIH DTLGLHFQIQ DTLGLHFQIQ NTMGLHFQIQ NHLGLHWQIH DRLGIQPVPQ NRLAFHPA-M KHITGLVIKP	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KGGGMHFYE- KGSQHFYE- HDIAIHIRQ- HDIAIHIRQ- HDVAQHIRK- QHIRQIADSW	CAAHPGERFE -KTKKKELFP -AAEOGKKLE -AKRLGKKLE -VEKGERRIP -AEQKKONLP -AEQKKONLP -AEQKKONLP -AEQKKONLP -AEERGINLP -AEERGINLP -AEERGEDLP AAIGKANEIP	VSVALGADPA AAVVIGANPV VAIAIGGPP VAVAIGSDPG AALVIGSDPG VTVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSTALGCEPV IATLGKDPV FALCFSVPPA	TILGAVT ITLAAVT ITLAAAT TLYAATA TLLTGAM LILSAIA LTLSAIA LTLSAIA ITTAAST ITTAAST ITTAAST AILVSSM	PVPDTLSEYA PIPENVSEYN PIPVDLSEWL PLPPEVDEYL PLPSGISEFM PVPYPIDKHL PLPENISELL PLPENISELL PLPSGIYELM PLLYDQSEYE PILYDQSEYE PIPEGVSESD	FAGLLR FAGLYG FAGIIM FAGFIR FAGVVR FCSFLQ LCTFLQ LYGFMR MAGAIQ MAGALR YVGAIR	GTKTEVVKCI NKKIELVOCK GSGVALAKCK KERLKVTECE GEGLPVYR-L NKKLSFVEK GSKLHY-KND GAKLLY-KKT GKKARVMPCL GEPYRIVKSK GAPAPIATAP GESVPVVKCE	300 SND-LEVPAS TID-LEVPAH TVD-LEVPAN TVD-LEVPAN PNG-IHVPAN PQSGHELLCD PDTPHELLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG TND-LMVPAT
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14.	201 GIYRQQLIGK GIYRQQVSE GVYRLQLQSK GLYRLQVLDK GMYRMQVFDG GWYRVMIAGD GMYRMQRFDR GMYRVQRFDR GMYRLQVYDK GIYRLQVYDK GIYRLQVIGT STARGMVVDD STARGMVVDD	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH KTGGHWQIH DTLGLHFQIQ NTMGLHFQIQ NHLGLHWQIH DRLGIQVVPQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP	RGGALDYQEW RGGAEHHRRW RGGARHIRK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KGGGMHFYE- KDSQLFFHE- HDIAIHIRQ- HDVAQHIRK- QHIRQIADSW RHLWHLYRK-	CAAHPGERFE -KTKKKELFE -AERGKKLE -AKRLGKKLE -VEKGERRIP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEERGINLP -AEERGINLP -AEERGEDLP AAIGKANEIP -ARERGEDLP	VSVALGADPA AAVVIGANPV VALAIGOPP VALAIGSDPG VAVAIGVDPA AALVIGSDPG VTVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSTAIGGDLU VTALGCPV FALCFSVPPA ATVVGLHPA	TILGAVT ITLAAVT ITLAAAT TLYAATA TLLTGAM LILSAIA LILSAIA YTWCATA ITTAAST ITTAAST AILVSSM AILVSSM VLLAGAT VLLAGAT	PVPDTLSEYA PIPENVSEYN PIPEVDSEWL PLPEVDEYL PLPEVDEYL PLPENVSELL PLPENVSELL PLPENVSELL PLYDQSEYE PIPEGVSESD SPPLGVEELG	FAGLLR FAGLYG FAGIIM FAGFIR FAGVVR FCSFLQ LCTFLQ LYGFMR MAGAIQ MAGALR YVGAIL LAACML LAACML	GTKTEVVKCI NKKIELVQCK GSGVALAKCK KERLKVTECE GEGLPVYR-L NKKLSFVEKH GAKLLY-KNT GKKARVMPCL GEPYRIVKSK GAPAPIATAP GESVPVVKCE GGSMKVYR-S	300 SND-LEVPAS TID-LEVPAH TVD-LEVPAN TVD-LEVPAN PNG-IHVPAN PQSGHELLCD PDTPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG TND-LMVPAT PVHGNPVP-L PN-GLIVD-F
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15.	201 GIYRQQLIGK GIYRQQVSE GVYRLQLQSK GLYRLQVLDK GMYRLQVFDG GVYRVMIAGD GMYRMQRFDR GMYRVQRFNQ GMYRLQVIDK GIYRLQVIGT SIARGMVVDD SIARGMVVDD SIARGMVVDD SIARGMVVDD	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH ETIGLHFQIQ DTIGLHFQIQ NTMGLHFQIQ NHIGLHWQIH DRIGIQPVPQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP YLVIRMV-EQ PDAVIBL-VP	RGGALDYQEW RGGAEHHKRW RGGARHIRK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KGGGMHFYE- KGSQHFYE- KDSQLFFHE- HDIAIHTRY- HDVAQHIRK- QHIRQIADSW RHLWHLYRK- RHLHFLYNK-	CAAHPGERFP -KTKKKELFP -AAEOGKKLE -AKRLGKKLE -VEKGERRIP -AEOKKQNLP -AEOKKQNLP -AEOKKQNLP -AEOKKQNLP -AEEKGNLP -AEEKGEDLP AAIGKANEIP -ARERGEDLP -ARERGEDLP -ALKEKGYLD -SVEHGEFE	VSVALGADPA AAVVIGANPV VATALGVDPL VATALGVDPA AALVIGSDPG VTVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSTALGCPV VATALGCPV IAITLGNDV FALCFGVPPA ATVVVGLHPA VATVIGNPH	TILGAVT ITLAAVT IIMAAT LIVAATA TLITGAM LILSAIA LILSAIA YTWCATA YTWCATA ITTAAST MAIVAGM VLLAAAT VLLAAAT VLLAAAT	PVPDTLSEYA PIPENVSEYN PIPVDLSEWL PLPPEVDEYL PLSGISEFM PVPYIDKHL PLPENISELL PLPYGIYELM PLLYGIYELM PLLYGUSEYE PIPEGVSESD SPPLGVFELG SADITFDELK SVAYGVELF	FAGLLR FAGLG FAGLG FAGLR FAGFLQ LCTFLQ FATFLQ LYGFMR MAGALR YVGALR YVGALL LAAGML FAAALL TAAALL	GTKTEVVKCI NKKIELVQCK GSGVALAKCK KERLKVTECE GEGLPVYR-L NKKLSFVEKH GSKLHY-KNT GKLHY-KNT GARALY-KKT GAPAPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GRPLEVIN-I	300 SND-LEVPAS TID-LEVPAH TVD-LEVPAN TVD-LEVPAN PNG-IHVPAN PQSGHELLCD PDTPHPLLYD NDHPHELLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG PVHGNPVP-L DN-GLLVP-E DG-IPTV-D
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17.	201 GIYRQQLIGK GIYRQQVSE GVYRLQLQSK GYRRQVFDG GVYRVMIAGD GMYRMQRFDK GMYRMQRFDK GMYRMQRFDK GTYRMQVFGK GTYRLQVIGT SIARGMVVDD SIHRLLILGR SIHRLLVKDD SFHRMMVLDE SIHRMMVIGD	NKLIMRWLSH NKLIMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH ETLGLHFQIQ DTLGLHFQIQ NHLGLHWQIH DRLGLPVPQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP YLVIRMV-EQ ERAVIRL-VP DRLAVRI-VP	RGGALDYQEW RGGAEHHKRW RGGARHIRK- KHGAEHFRKM KRGMQAQDDS KGGGAHFFE- KGGGMHFYE- KDSQLFFHE- HDIAIHTRY- HDVAQHIRK- QHIRQIADSW RHLWHLYRK- RHLYSMWKD- RHLYSMWKD-	CAAHPGERFP -KTKKKELFP -AAEOGKKLE -AKRLGKKLE -AEKGGKIE -VEKGERRIP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AECKKQNLP -AEEKGEDLP -AEEKGEDLP -ALEKGYLD -SVEHGEELE -AEERGEDLE	VSVALGADPA AAVVIGANPV VATALGVDPL VATALGVDPL VAVAIGVDPA AALVIGSDPG VTVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSTALGCPV IAITLGNDPV FALCFGVPPA ATVVVGLHPA VAIVIGNPA VAIVIGNPA IAITLGNDPA	TILGAVT ITLAAVT IIMAAT TLYAATA TLITGAM LILSAIA LILSAIA YTWCATA YTWCATA YTWCATA YTWCATA VLTAAT VLLAAT VLLAAT LLLAGAT LLAGAT LLAGAT	PVPDTLSEYA PIPENVSEYN PIPVDLSEWL PLPPEVDEYL PLPSGISEFM PVPYIDKHL PLPENISELL PLPYGIYELM PLPYGIYELM PLLYGQSEYE PIAYDQSEYE PIPEGVSESD SPPLGVFELG SADITFDELK SVAYGVSELE	FAGLLR FAGLG FAGIIG FAGIIR FAGFIQ LCTFLQ LTGFMR MAGAIQ LYGFMR MAGAIR YVGAIL FAAALL IASAISLKAF VANTFH	GTKTEVVKCI NKKIELVQCK GSGVALAKCK KERLKVTECE GEGLPVYR-L NKKLSFVEKH GSKLHY-KND GAKLLY-KKT GKKARVMPCL GEPYRIVKSK GAPAPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GRPLEVIN-L EGELELVR-C	300 SND-LEVPAS TID-LEVPAN TVD-LEVPAN TVD-LEVPAN PNG-IHVPAN PQSGHPLLCD PDTPHPLLYD NDHPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG TND-LMVPAT PVHGNPVP-L DN-GLLVP-E DG-IPTPV-D EGVDMEVP-P
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18.	201 GIYRQQLIGK GIYRMQVVSE GVYRLQLQSK GLYRLQVLDK GMYRMQVFDG GWYRMQRFDR GMYRMQRFDR GMYRLQVIGFDR GTYRLQVIGT SIARGMVVDD SIHRLLIGR SIHRILVKDD SFHRMMVIGD SIHRMLVLDD	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH FTGGLHFQIQ DTLGLHFQIQ NHLGLHFQIQ NHLGLHWQIH PRIGIQPVPQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP ZLVIRMV-EQ ERAVIRL-VP CRLAVRI-VP	RGGALDYQEW RGGAEHHKRW KGGARHIRK- KHGAEHFRKM KRGMQAQODS KGGGAHFFE- KGGGMHFYE- KGSGMHFYE- HDIAIHIRQ- HDIAIHIRQ- HDIAIHIRQ- HDVAQHIRK- QHIRQIADSW RHLWHLYRK- RHLWIYRK- RHLHFLYNK- RHLYTYIQK-	CAAHPGERFE -KTKKKELFP -AAEQGKKLE -AKRLGKKLE -VEKGERRIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEEKGEDLP -AEEKGEDLP -AEEKGEDLP -ARERGEDLP -ALKEKGYLD -SVEHGEELE -AEERGEDLE -AEERGEDLE	VSVALGADPA AAVVIGANPV VÄLAIGGEPP VÄVAIGVDPA AALVIGSDPG VTVFLSGNPF VSVALGSDPF VSVALGRDPV IAITLGNDPV FÅLCFGVPPA VAIVIGVHPA VRIVLGNPVH IAITLGNDPA IAILGMDPA	TILGAVT ITLAAVT LTVAATA TLYAATA TLISAIA LTLSAIA LTLSAIA LTLSAIA LTLSAIA LTLSAIA VTWCATA NIVAGM AILVSSM VLLAAST TLLAATT TLLAATT TLLAATT	PVPDTLSEYA PIPENVSEYN PIPEVDSEYN PLPPEVDEYL PLPSGISEFM PVPYPIDKHL PLPENISELL PLPENISELL PLPYQGYELM PLLYQQSEYE PMAYDQSEYE PMAYDQSEYE SIPEGVESD SADITFDELK SVAYGVSELE SIPIDADEME SIPIDADEME	FAGLLR FAGLG FAGLG FAGIR FAGYR FAGYQ LCTFLQ LTFLQ MAGALR YVGAIR YVGAIR YVGAIL FAAALL FAAALL FAAALL VANTFK	GTKTEVVKCI NKKIELVQCK GSGVALAKCK KERLKVTECE GEGLPVYR-L SKKLSFVEKH GSKLHY-KND GAKLLY-KKT GAPAPIATAP GESYPVVKCE GGSMKVYR-S GGEIGVFE-L GRPLEVIN-L EGELELVR-C NGELTLIK-C	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LEVPAN TVD-LLVPAN POSGHPLLCD PDTPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG TND-LMVPAT PVHGNPVP-L DN-GLLVP-E GG-DIPTPV-D EGVDMEVP-P G-DLEVP-Q
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19.	201 GIYRQQLIGK GIYRQQVSE GVYRLQLQSK GLYRLQVLDK GMYRMQVFDG GMYRMQVFDG GMYRMQRFDR GMYRVQRFDR GMYRLQVIGT SIARGMVVDD SIHRLLIGR SIHRLLVKDD SIHRMVLDE SIHRMVLDE SIFRMLVLDD SIGRFQVIGK	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH ETLGLHFQIQ DTLGLHFQIQ NHLGLHFQIQ NHLGLHFQIQ NHLGLHPQIH PRLGIQPVPQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP JLVIRMV-EQ ERAVIRL-VP DRLAVINIDAG	RGGALDYQEW RGGAEHHKRW KGGARHIRK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KGSGMHIYE- KGSQHFYE- KDSQLFFHE- HDIAIHIRQ- HDVAQHIRK- QHIRQIADSW RHLWHLYRK- RHLFLYNK- RHLYSMWKD- RHLYTYQK- RHLYTYQK- RHLGLYLDK-	CAAHPGERFE -KTKKKELFP -AAEQGKKLE -AKRLGKKLE -VEKGERRIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEERGINIP -AEERGEDIP AAIGKANEIP -ARERGEDIP -ARERGEDIP -ALEKGYLD -SVEHGEELE -AQKLGKDLE -AQARGEPLA	VSVALGADPA AAVVIGANPV VATAIGGEPP VATAIGGEPP VATVIGSDPG VTVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSTAIGGDLI IATTIGNDPX TATUGNPYA ATVVVGLHPA VTIVLGNPVH IATAIGMDPA IATAIGMDPA FTLNVGVGPG	TILGAVT ITLAAVT ITLAAT TLYAATA TLYAATA TLISAIA LTISAIA LTISAIA YTWCATA AILVSAM AILVSSM AILVSSM VLIAAST LLLACAT TLLACAT TLLACAT TLLACTT VLIAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	PVPDTLSEYA PIPENVSEYN PIPPUDLSEWL PLPPEVDEYL PLPENVPELL PLPENVPELL PLPENVSELL PLPENVSELL PLPENSELL PLPENSEL PLSYDGSEYE PLSYDGSEYE SADITPDELK SVAYGVSELE SIPIDADEME ARVETDELG	FAGLLR FAGLYG FAGIIM FAGFIR FAGFIR FAGFIQ LCTFLQ FATFLQ LYGFMR YUGAIR YUGAIL LAAGML IASAISLKAF VANTFH IASAFH	GTKTEVVKCI NKKIELVOCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEKH GSKLHY-KND GAKLLY-KKT GEPYRIVKSK GAPAPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GRPLEVIN-L EGELELVR-C GAPLELVAGT	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LEVPAN TVD-LLVPAN PQSHPLLCD PDTPHPLLYD NDHPHPLLYD SNS-LSVPSD LTG-FDVPWG TND-LMVPAT PVHGPVP-L DN-GLLVP-E DG-IPTEV-D EGVDMEVP-P G-DLEVP-Q VGPVEMVA-H
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus	201 GIYRQQLIGK GIYRQQVSE GVYRLQLQSK GLYRLQVLDK GMYRMQVFDG GWYRVMIAGD GMYRWQRFDR GMYRVQRFDR GMYRVQRFNQ GMYRLQVIGT SIARGMVVDD SIHRLLJCR SIHRLLVKDD SFHRMMVLDE SIHRMMVLDE SIFRMMVLDD SICRFQVIGK giyR.qv	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH KEGVVHWQIH DTLGLHFQIQ NTMGLHFQIQ NHLGLHWQIH DRIGIQPVPQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP DRLVIRV-EQ ERAVIRL-VP DRLAVRI-VP DRLVINIDAG 1	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KGGGMHLYE- KDSQLFFHE- HDIAIHIRQ- HDVAQHLRK- QHIRQIADSW RHLWHLYRK- RHLFLYNK- RHLYTYLQK- RHLYTYLQK- RHLYTYLQK- rHLQYLDK- rl.k.	CAAHPGERFE -KTKKKELFP -AEACQKKKE -AKRLGKKLE -VEKGERRIP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEERGINLP -AEERGEDLP AAIGKANEIP -ARERGEDLP -ALEKGYLD -SVEHGEELE -AQKLGKDLE -AAARGEPLA .ag.lp	VSVALGADPA AAVVIGANPV VATAIGGEPP VATAIGGEPP AALVIGSDPG VTVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSTAIGGDLI VTTALGCPV FALCFGVPPA ATVVGLHPA VRIVLGNPVH IATAIGMDPA IATAIGMDPA IATAIGMDPA FILNKGVGPG vai.g.pa	TILGAVT ITLAAVT ITLAAAT TLYAATA TLIGAM LILSAIA LTLSAIA ITTAAST MAIVAGM AILVSSM VLLAAST LLLAGAT TLLAST LLLAGAT ILLAST LLLAGAT LLLAGAT LLLAGAT LLLAST LLLAST LLLAST LLLAST LLLAST LLLAST LLLAST LLAST LLAST	PVPDTLSEYA PIPENVSEYN PIPEVDLSEWL PLPEVDLSEWL PLPEVDEYL PLPENVSELL PLPENVSELL PLPENVSELL PLYDQSEYE PLLYDQSEYE PLLYDQSEYE SADITPDELK SVAYGVSELE SIPIDADEME SIPIDADEME SIPIDADEME P.A.E.	FAGLLR FAGLYG FAGIIM FAGFIR FAGYVR FCSFLQ LCTFLQ LTGFMR MAGAIQ MAGALR YVGAIL LAAGML IASAISLKAF VANTFH VANAFK IASAFH .ag	GTKTEVVKCI NKKIELVOCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L GSKLHY-KND GAKLLY-KKT GKKARVMPCL GEPYRIVKSK GGPAPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GRPLEVIN-L EGELELVR-C NGELTLIK-C GAPLELVAGT g	300 SND-LEVPAS TID-LEVPAN TVD-LEVPAN TVD-LEVPAN PVD-LLVPAN PQSGHELLCD PDTPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG TND-LMVPAT PVHGPVP-L DN-GLLVP-E DG-IPTPV-D EGVDMEVP-P G-DLEVP-Q VGPVEMVA-H l.Vp
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus	201 GIYRQQLIGK GIYRQQVSE GVYRLQLQSK GLYRLQVLDK GMYRMQVFDG GWYRVQRFDR GMYRVQRFDR GMYRVQRFDR GMYRLQVYDK GIYRLQVYDK SIARGMVVDD SIARGMVVDD SIARGMVVDD SFHRLLIGR SFHRLVKDD SFHRMMVLDE SFHRMMVLDE SIARMVIGD SIORFQVIGK GIYR.qv	NKLIMRWLSH NKLIMRWLKL TTMTVHWLSV DKLAVHWQIH KTEGMHWQIH ETIGLHFQIQ DTIGLHFQIQ NTMGLHFQIQ NHIGLHWQIH DRIGIQPVPQ NRIAFHPA-M KHITGLVIKP ERAAIRI-VP YLVIRMV-EQ ERAVIRL-VP DRIAVRI-VP KRIVIRI-VP DRIAVRI-VP DRIVINIDAG 1	RGGALDYQEW RGGAEHHRRW RGGARHIRK- KDGNHHYWK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFYE- KGGGMHFYE- KDSQLFFHE- HDYAQHIRK- QHIRQIADSW RHLWHLYRK- RHLHFLYNK- RHLHFLYNK- RHLYTYIQK- RHLYTYIQK- RHLYTYIQK- RHLYTYIQK- RHLSYLDK- Il.k.	CAAHPGERPP -XTKKKELFP -AEACGKKLE -AKRLGKKLE -VEKGERRIP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEERGINLP -AEERGEDLP -ARERGEDLP -ALKEKGYLD -SVEHGEELE -AEERGEDLE -AAARGEPLA .ag.lp	VSVALGADPA AAVVIGANPV VATAIGGEPP VAVAIGVDPA AALVIGSDPG VTVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSVFLGGDPV FALCFGVPPA ATVVGLHPA VAIVIGVHPA VAIVIGVHPA VAIVIGVHPA TATAIGMDPA IATAIGMDPA IATAIGMDPA IATAIGMDPA IATAIGMDPA TTLNVGVGG VaiG.pa	TILGAVT ITLAAVT ITLAAAT TLYAATA TLIGAM LIISAIA LTISAIA YTWCATA ITTAAST AILVSSM VILGAAT VLLAAST LLIAGST LLIAGST LLIAGST LLIAGST HFAAAPAE .laa.t	PVPDTLSEYA PIPENVSEYN PIPEVDSEWL PLPPEVDSWL PLPSGISEFM PVPYPIDKHL PLPENVSELL PLPENVSELL PLPSGISEFM PLLYGIVELM PLLYGUSELE SIPEGVSESD SADITFDELK SIPIDADEME SIPIDADEME SIPIDADEME P.PE.	FAGLLR FAGLYG FAGIIM FAGFIR FAGFUR FCSFLQ LCTFLQ LYGFMR MAGAIQ MAGAIQ MAGAI LYGAIL LAAGML FAAALL FAAALL FAAALL FAAALK VANTFH VANAFK IASAFH .ag	GTKTEVVKCI NKKIELVQCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEKH GAKLLY-KKT GKKARVMPCL GEPYRIVKSK GGPAPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GGPLEVIN-L EGELELVR-C NGELTLIK-C GAPLELVAGT G	300 SND-LEVPAS TID-LEVPAN TVD-LEVPAN TVD-LEVPAN PNG-IHVPAN PQSGHPLLCD PDTPHPLLYD NDHPHPLLYD SNS-LSVPSD LSD-LDVPWG IND-LMVPAT PVHGNPVP-L DN-GLLVP-E DG-IPTPV-D EGVDMEVP-P G-DLEVP-Q VGPVEMVA-H l.VP
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus	201 GIYRQQLIGK GIYRQQVSE GYRLQLQSK GYRLQVLDK GYRLQVLDK GYRLQVFDG GYRVMQRFDR GMYRVQRFDR GMYRVQRFDR GMYRLQVYDK GIYRLQVIGT SIARGMVVDD SIARGMVVDD SIARGMVVDD SIARGMVVDD SIARGMVVDD SIARGMVVDD SIARGMVVDD SIARGMVVDG SIARGMVVDG SIARMVIGD SIARMVIG SIARVVIG	NKLIMRWLSH NKLIMRWLKL TTMTVHWLSV DKLAVHWQIH KTEGMHQIH KEGVVHWQIH ETIGLHFQIQ DTIGLHFQIQ NTMGLHFQIQ NHIGLHWQIH DRLGIQVVPQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP YLVIRMV-EQ ERAVIR-VP KRLVIRI-VP KRLVIRI-VP CRLVINIDAG 1	RGGALDYQEW RGGAEHHKRW RGGARHIRK- KDGNHHYWK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KDSQLFFHE- HDIAIHIRQ- HDVAQHIRK- QHIRQIADSW RHLWHLYRK- RHLSYMWKD- RHLYTYLQK- RHLYTYLQK- RHLYTYLQK- RHLGYLDR- rl.k.	CAAHPGERFF -KTKKKELFF -AAEQGKKLE -AKRLGKKLE -VEKGERRIP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEERGEDLP -AEERGEDLP -AIGKANEIP -AEERGEDLP -ALKEKGYLD -SVEHGEELE -AAERGEDLE -AAERGEDLA -AEERGEDLE -AAERGEDLA -AEERGEDLE -AAERGEDLA -AEERGEDLE -AAERGEDLA -AEERGEDLE -AAERGEDLA -AAERGEDLA	VSVALGADPA AAVVIGANPV VATAIGOPP VATAIGOPP AALVIGSDPG VTVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSTAIGGDLV TAICGOPPA ATVVGLHPA VAIVIGVHPA VAIVIGVHPA VAIVIGVHPA IATAIGMDPA IATAIGMDPA IATAIGMDPA IATAIGMDPA VAIVIGVPA VAIVIGVPA VAIVIGVPA VAIVIGVPA	TILGAVT ITLAAVT IIMAAAT LIVAATA TLLTGAM LILSAIA UTISAIA YTWCATA YTWCATA ITTAAST VILAAST VLLAGST VLLAGST TLLATTT ILLASTT VHFAAAAPAE laa.t	PVPDTLSEYA PIPENVSEYN PIPEVDSEWL PLPEVDSEWL PLPEVDSEUL PLPENVELL PLPENVELL PLPENVELL PLPYGIYELM PLLYGIYELM PLYGIYELM SPRICYELG SADITFDELK SVAYGVSELE SIPIDADEME SIPIDADEME SIPIDADEME P.PE.	FAGLLR FAGLYG FAGIIM FAGFIR FAGVVR FCSFLQ LCTFLQ FATFLQ FATFLQ HAGAIQ LYGFMR MAGAIQ LYGAIL LAAGML FAAALL FAAALL IASAISLKAF VANTFH VANAFH .ag	GTKTEVVKCI NKKIELVOCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEKH GAKLLY-KKT GKKARVMPCL GEYRIVKSK GAPAPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GGPLEVIN-L EGELELVR-C NGELTLIK-C GAPLELVAGT 9 CKOFPEIVDF	300 SND-LEVPAS TID-LEVPAN TVD-LEVPAN TVD-LEVPAN PNG-IHVPAN PNG-IHVPAN PDTPHPLLYD NDHPHFLLYD NDHPHFLLYD SNS-LSVPSD LSD-LDVPWG TND-LMVPAT PVHGNPVP-L DN-GLLVP-E DG-IPTPV-D EGVDMEVP-P G-DLEVP-Q VGPVEMVA-H l.vp
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus 1. 2.	201 GIYRQQLIGK GIYRMQVVSE GVYRLQLQSK GLYRLQVDK GWYRLQVDK GWYRWQRFDQ GMYRWQRFDR GMYRUQVFDG GMYRLQVIGT SIARGMVVDD SIHRLLUKDD SIHRLLVKDD SIHRLLVKDD SIHRMLVLDD SIGRFQVIGK giyR.qv 301 AEIVLEGYIE SEIVLEGYIE	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV KTTGMHWQIH KEGVVHWQIH ETIGLHFQIQ DTIGLHFQIQ NHLGLHFQIQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP YLVIRMV-EQ ERAVIRL-VP DRLAVRI-VP CRLVIRI-VP DRLAVRI-VP CRLVIRI-VP DRLAVRI-VP CRLVIRI-VP DRLAVRI-VP CRLVIRI-VP DRLAVRI-VP CRLVIRI-VP DRLAVRI-VP CRLVIRI-VP DRLAVRI-VP CRLVIRI-VP DRLAVRI-VP CRLVIRI-VP DRLVINIDAG	RGGALDYQEW RGGAEHHKRW RGGARHIRK- KDGNHHYWK- KHGAEHFRKM KRGMQAQODS KGGGAHFYE- KGGGMHFYE- KGSGMHFYE- KDSQLFFHE- HDIAIHIRQ- HDIAIHIRQ- HDIAIHIRQ- HDIAIHIRQ- HDIAIHIRQ- HDIAIHIRQ- HILWIANK- RHLWIYK- RHLHFLYNK- RHLHFLYNK- RHLFLYRK- RHLFLYEK- RHLGYYEQK- RHLGLYEDK- r1.k.	CAAHPGERFE -KTKKKELFP -AAEAGGKKIE -AKRLGKKLE -AEKGGKIE -VEKGERRIP -AEQKKQNIP -AEQKKQNIP -AEQKKQNIP -AEQKKQNIP -AEEKGEDLE -AEEKGEDLE -ALKEKGYLD -SVEHGEELE -AEERGEDLE -AEERGEDLE -AEERGEDLE -AEERGEDLE -AEERGEDLE -AERGEDLA -AERGEDLA -AERGEDLA -AERGEDLA -AERGEDLA -AERGEDLA -AERGEDLE -AERGED	VSVALGADPA AAVVIGANPV VÄLAIGCEPP VÄLAIGCEPP VÄVAIGVDPA AÄLVIGSDPG VTVFLSGNPF VSVALGSDPF VSVALGEPV IAITLGNDPV FÅLCFGVPPA VAIVIGVHPA VAIVIGNPA VAIVIGNPA VAIVIGVHPA	TILGAVT ITLAAVT ITLAATA TLYAATA TLYAATA LILSAIA LILSAIA LTLSAIA ITTAAST MAIVAGM AILVSSM VLLAAST VLLAGAT VLLAGAT ILLASTT ILLASTT VHFAAAPAE laa.t	PVPDTLSEYA PIPENVSEYN PIPEVDSEYN PLPPEVDSYL PLPSGISEFM PVPYPIDKHL PLPSGISEFM PLPSNVSELL PLPSNVSELL PLPSNVSELL PLYDQSEYE PMAYDQSEYE SIPICAUCSEYE SADITFDELK SVAYGVSELE SIPICAUCSEYE SIPICAUCSEYE SIPICAUCSEYE SIPICAUCSEYE SIPICAUCSEYE SIPICAUCSEYE P.PE. PDE-PAVLGV PDE-PAVLGV	FAGLLR FAGLG FAGLG FAGIR FAGFIR FAGFIQ LCTFLQ LTFLQ MAGALR YVGAIL IASAISLKAF FAAALL IASAISLKAF FAAALK IASAFH .ag. ALNEVFVPIL ALNEIFIPIL	GTKTEVVKCI NKKIELVQCK GSGVALAKCK KERLKVTECE GEGLPVYR-L SKLHY-KND GAKLHY-KND GAKLHY-KNT GAPAPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GRPLEVIN-L GAPLEVN-C NGELTLIK-C GAPLELVAGT G QKQFPEIVDF HQQFPEIVDF	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LEVPAN PVD-LLVPAN POSGHPLLCD PDTPHPLLYD NDHPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG LTG-FDVPWG TND-LMVPAT PVHGNPVP-L DN-GLLVP-E GC-DLEVP-Q VGPVEMVA-H l.vp 400 YLPPEGCSYR
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus 1. 2. 3.	201 GIYROQLIGK GIYROQVSE GVYRLQUQSK GLYRLQVLDK GMYRMQVFDG GWYRWQRFDR GMYRMQRFDR GMYRLQVDFFDK GMYRLQVIGT SIARGMVVDD SIHRLLIGR SIHRLLVKDD SFHRMMVLDD SIHRMLVLDD SIGRFQVIGK giyR.qv 301 AEIVLEGYIE SEIVLEGYIE SEIVLEGYIE	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH ETLGLHFQIQ DTIGLHFQIQ NTMGLHFQIQ NHLGLHWQIH DRIGIQPVPQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP YLVIRMV-EQ ERAVIRL-VP DRIAVRI-VP DRIAVRI-VP DRIAVRI-VP DRIVINIDAG 1	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KGSGMHIYE- KGSGMHIYE- KDSQLFFHE- HDIAIHIRQ- HDIAIHIRQ- HDIAHIRQ- HDVAQHLRK- QHIRQIADSW RHLWHLYRK- RHLFLYNK- RHLFLYNK- RHLYYYLQK- RHLGLYLDK- rl.k. YGDHTGY-YN FGDHTGY-YN	CAAHPGERFE -KTKKKELFP -AAEQGKKLE -AKRLGKKLE -EKGGGKIE -VEKGERRIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEEKGENIP	VSVALGADPA AAVVIGANPV VALAIGGEPP VAVAIGVDPA AALVIGSDPG VTVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSTALGCPV IAITLGNDPV FALCFGVPPA ATVVVGLHPA VAVVIGNPVH IAIAIGMDPA FTLNVGVGPG vai.G.pa VTHITQREDA VTAITMKKP FOCLTHRKNP	TILGAVT ITLAAVT ITLAATA TLIGAA TLIGAA LILSAIA LTLSAIA LTLSAIA TLGATA YTWGATA NIVAGM AILVSSM VLLAGAT TLLATT TLLATT VLLAGAT TLLATTT VLLAGAT TLLATTT VLFAAAAPAE laa.t	PVPDTLSEYA PIPENVSEYN PIPEVDLSWN PLPPEVDLSWL PLPENVDELL PLPENVSELL PLPENVSELL PLPENVSELL PLPENSEL PLPENSEL PLSVOESYE SIPLOVESYE SIPLOVESYE SIPLOVESYE SIPLOVESYE SIPLOVESUE AAPVETDELK SIPLOVESUE AAPVETDELG P.PE	FAGLLR FAGLG FAGLG FAGIR FAGFIR FAGFIQ LCTFLQ LTFLQ HAGALR YVGAIR YVGAIL LAAGML FAALL IASAISLKAF VANTFH VANAFH .ag ALNEVFVPIL ALNEIFIPIL ALNEIFIPIL ALNEIFIPIL	GTKTEVVKCI NKKIELVOCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEKH GAKLLY-KKT GAKLLY-KKT GAKLLY-KKT GAPAPIATAP GESYRVKSK GAPAPIATAP GESYRVKSK GGEIGVFE-L GGSMKVYR-S GGEIGVFE-L GGELELVR-C NGELTLIK-C GAPLELVAGT 9 QKQFPEIVDF HQQFPEIVDF RQQVSEITDF	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LEVPAN TVD-LLVPAN PQSGHPLLCD PDTPHPLLYD NDHPHPLLYD SNS-LSVPSG LTG-FDVPWG LTG-FDVPWG TND-LMVPAT PVHGNPVP-L DN-GLLVP-E GC-DLEVP-Q VGPVEMVA-H 1.Vp 400 YLPPEGCSYR WFPPEGCSYR FLPMEALSYK
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus 1. 2. 3. 4.	201 21YRQQLIGK GIYRQQVSE GVYRLQLQSK GLYRLQVLDK GMYRMQVFDG GWYRVQRFDQ GMYRVQRFDQ GMYRVQRFDQ GMYRLQVIGT SIARGMVVDD SIHRLLUG SIHRLLVKDD SIHRLLVKDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRMVLDD SIFRWVLDD SIFFWVDD SIFRWVLDD SIFFWVDD SIFFW	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH KTGGHHQIQ DTIGLHFQIQ NTMGLHFQIQ NHLGLHVQIH DRLGIQPVPQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP DRLVINIDAG QGETAP-EGP PEGEMLP-DGP PEEPLVDEGP	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KGSGMHIYE- KGSGMHIYE- KDSQLFFHE- HDIAIHIRQ- HDVAQHLRK- QHIRQIADSW RHLWHLYRK- RHLFLYNK- RHLYTYLQK- RHLYTYLQK- RHLYTYLQK- RHLGLYLDK- rl.k. YGDHTGY-YM FGDHTGY-YM FGDHTGY-YM	CAAHPGERFE -KTKKKELFP -AAEQGKKLE -AKRLGKKLE -EKGGGKIE -VEKGERRIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEERGINIP -AEERGEDIP AAIGKANEIP -ARERGEDIP -ARERGEDIP -ALEKGYLD -SVEHGEELE -AGKLGKDLE -AAARGEPLA .ag1p EVDS-FPVFT DVEE-FPIFT GVED-SPLVR POK-YPOMH	VSVALGADPA AAVVIGANPV VATAIGGEPP VATAIGGEPP VATAIGGEPP VIVELSGNPF VSVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSTAIGGDLI IATIGMDPA ATVVVGLHPA ATVVVGLHPA VTVLGNPVH IATAIGMDPA FTLNVGVGPG VALVIGVHPA VTVLGNPVH IATAIGMDPA FTLNVGVGPG VALVIGNPVH VAL	TILGAVT ITLAAVT ITLAAAT TLYAATA TLYAATA TLISAIA LTLSAIA LTLSAIA ITTSAIA ITTSAIA ITTSAIA NIVAGTA AILVSSM VILAGAT TLLACAT	PVPDTLSEYA PIPENVSEYN PIPEVDLSEWL PLPEVDLSEWL PLPEVDEYL PLPENVPELL PLPENVSELL PLPENVSELL PLPENSELL PLPENSEL PLEVDQSEYE PLLYDQSEYE PLLYDQSEYE SIPIDADEME SIPIDADEME DAAVETDELK SIPIDADEME P.DE PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV	FAGLLR FAGLG FAGLG FAGIR FAGFIR FAGFIQ LCTFLQ LTFLQ FATFLQ LYGFMR YVGAIL LAAGML IASAISLKAF VANAFH IASAISLKAF VANAFH .ag ALNEVFVPIL ALNRIYTPIL ACNRIYTPIL ACNRIYTPIL	GTKTEVVKCI NKKIELVOCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEKH GAKLLY-KKT GAKLLY-KKT GAKLLY-KKT GAPAPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GRPLEVIN-L EGELELVR-C GAPLELVAGT 9 QKQFPEIVDF HQQFPEIVDF FNLPEVVDY	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LEVPAN TVD-LLVPAN PNG-IHVPAN PDTPHPLLYD NDHPHPLLYD SNS-LSVPSD LTG-FDVPWG TND-LMVPAT PVHGPVP-L DN-GLLVP-E DG-IPTPV-D EGVDMEVP-P GO-DLEVP-Q VGPVEMVA-H 1.Vp 400 YLPPEGCSYR WFPPEGCSYR WFPPEGCSYR
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus 1. 2. 3. 4. 5.	201 21YRQQLIGK GIYRQQVSE GVYRLQLQSK GLYRLQVLDK GMYRMQVFDG GWYRVMIAGD GMYRMQRFDR GMYRVQRFNQ GMYRLQVYDK SIRGUVGT SIARGMVVDD SIHRLLIGR SIHRLLVKDD SIHRLLVKDD SIHRMVLDE SIFRMVLDD SICRFQVIGK giyR.qv 301 AEIVLEGYIE SEFVLEGYIE SEFVLEGYIE SEFVLEGYIE AEIALEGYVD	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTEGMHWQIH KEGVVHWQIH ETLGLHFQIQ NTMGLHFQIQ NTMGLHFQIQ NTMGLHFQIQ NTMGLHFQIQ NTMGLHFQIA NTMGLHFQIA NTMGLHFQIA NTMGLHFQIA NTMGLHFQIA NTMGLHFQIA NTMGLHFQIA SCH CONTAP-EGP CONTAP-EGP CONTAP-EGP DEEPLVDEGP VDEMRV-EGP	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KDGNHHYWK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFYE- KGSGMHIYE- KDSQLFFHE- HDIAIHIRQ- HDVAQHIRK- QHIRQIADSW RHLWHLYRK- RHLYTYLQK- RHLYTYLQK- RHLYTYLQK- RHLYTYLQK- RHLYTYLQK- RHLGYLDK- r1.k. YGDHTGY-YM FGDHTGY-YM FGDHTGY-YM	CAAHPGERFE -KTKKKELFP -AAEQGKKLE -AKRLGKKLE -VEKGERRIP -AEQKKONLP -AEQKKONLP -AEQKKONLP -AEQKKONLP -AEERGINLP -AEERGINLP -AEERGEDLP AIGKANEIP -AEERGEDLP -ALEKGYLD -SVEHGEELE -AQKLGKDLE -AAARGEPLA .ag.lp EVDS-FPVFT DVEE-FPIFT GVED-SPLVR PVDK-YPQMH PPEP-YPVFT	VSVALGADPA AAVVIGANPV VAIAIGGEPP VAIAIGGEPP VAVAIGVDPA AALVIGSDPG VTVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSTALGCEPV FALCFGVPPA ATVVGLHPA VAIVIGNPA VAIVIGNPA VAIVIGNPA VAIVIGNPA VAIVIGNPA VAIVIGNPA VAIVIGNPA VAIVIGNPA VAIVIGNPA VAIVIGNPA VAIVIGNPA VAIVIGNPA VAIVIGNPA VAIVIGNPA VAIVIGNPA VAIVIGNPA VAIVIGNPA VIAIVMENPA	TILGAVT ITLAAVT ITLAAAT TLLTGAM TLLTGAM LTLSAIA LTLSAIA ITTAAST ITTAAST MAIVAGM AILVSSM VLLAGAT TLLATTT ILLAST	PVPDTLSEYA PIPENVSEYN PIPEVDLSEWL PLPPEVDLSEWL PLPEVDLSEWL PLPENTSELL PLPENTSELL PLPENTSELL PLPENTSELL PLPENTSELL PLPENTSEL SIPECTSEL SIPECTSEL SIPECTSEL SIPECTSEL SIPECTSEL SIPECTSEL PLLYDQSEYE SIPECTSEL SIPECTSEL SIPECTSEL PLLYDQSEYE PLLYDQSEYE SIPECTSEL SIPECTSEL SIPECTSEL PLLYDNEMD PLLYDNEMD PLLYDNEMD PLLYDNEMD PLE - PAVLGV PDE - PAVLGV	FAGLLR FAGLYG FAGIIM FAGFIR FAGYVR FCSFLQ LYGFMR MAGAIQ MAGAIQ MAGAIQ MAGAI IASAISLKAF VANTFH IASAISLKAF VANTFK IASAF	GTKTEVVKCI NKKIELVOCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEK GAKLLY-KNT GKKARVMPCL GEPYRIVKSK GAPAPIATAP GESVPVVKCE GGEJGVFE-L GRPLEVIN-L EGELELVR-C NGELTLIK-C NGELTLIK-C NGELTLIK-C QKQFPEIVDF RQQVSEITDF RQQVSEITDF KFNLPEVVDY RMMHPEIVDI	300 SND-LEVPAS TID-LEVPAN TVD-LEVPAN TVD-LEVPAN TVD-LEVPAN PQSGHPLLCD PDTPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG TND-LMVPAT PVHGNPVP-L DG-IPTPV-D EGVDMEVP-P G-DLEVP-Q VGPVEMVA-H l.VP 400 YLPPEGCSYR WFPPEGCSYR WFPPEGCSYR FLPMEALSYK HLPAEGCFHN NLPVEGAFHN
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus 1. 2. 3. 4. 5. 6. 7. 7. 8. 9. 10. 11. 12. 13. 14. 15. 14. 15. 14. 15. 14. 15. 14. 15. 14. 15. 14. 14. 15. 14. 14. 14. 15. 14. 14. 14. 14. 15. 14. 14. 14. 14. 14. 14. 14. 14. 14. 14	201 GIYRQQLIGK GIYRQQVSE GVYRLQLQSK GLYRLQVLDK GMYRMQVFDG GWYRVQRFDC GMYRVQRFDR GMYRVQRFDR GMYRLQVJDK SINGUST SIARGMVVDD SIARGMVDD SIARGMVD	NKLIMRWLSH NKLIMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH KTGGHWQIH ETLGLHFQIQ NTMGLHFQIQ NTMGLHFQIQ NHLGLHWQIH RRIGLPVPQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP DRLVIRI-VP DRLAVRI-VP DRLVIRI-VP DRLVIRI-VP DRLVIRI-VP DRLVIRI-VP DRLVINIDAG LABYLP-EGP PEEPLVDEGP VDEMRV-EGP LEDLRE-EGP	RGGALDYQEW RGGAEHHRRW RGGARHLRK- KDGNHHYWK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFYE- KOSQLFFHE- HDIAIHIRQ- HDVAQHIRK- QHIRQIADSW RHLWHLYRK- RHLYTYIQK- RHLYTYIQK- RHLYTYIQK- RHLYTYIQK- RHLYTYIQK- RHLGYLDK- I1.k. YGDHTGY-YN FGDHTGY-YT YGDHTGY-YT YGDHFGY-YT	CAAHPGERFE -KTKKKELFF -AEACGKKLE -AKRLGKKLE -AEKGGKIE -VEKGERRIP -AEQKKQNLF -AEQKKQNLF -AEQKKQNLF -AEERGINLF -AEERGINLF -AEERGEDLF AAIGKANEIP -AEERGEDLF -AKEKGYLD -SVEHGEELE -AQKLGKDLE -AAARGEPLA .aglp EVDS-FPVFT DVEE-FPIFT GVED-SPLVR FVDK-YPQMH PPEP-YPVFH KPSRLFPTFR	VSVALGADPA AAVVIGANPV VALAIGOPP VALAIGGEPP VALAIGSDPG VTVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSVFLSGNPF ATAUGEPV FALCFGVPPA ATUVGLHPA VALVIGHPA	TILGAVT ITLAAVT ITLAAT TLYAATA TLLTGAM LILSAIA LILSAIA LILSAIA LITSAIA ITTAAST NILGATA ILLAATT ILLAAT LLLAGAT TLLATT ILLAST LLAGAT TLLATT ILLAST LLAGAT TLLATT ILLAST TLLATT ILLAST LLAGAT TLLATT ILLAST TLLATT ILLAST TLLATT ILLAST TLLATT ILLAST TLLATT TLLATT ILLAST TLLATT ILLAST TLLATT ILLAST ILLAST TLLATT ILLAST TLLATT	PVPDTLSEYA PIPENVSEYN PIPEVDSEWL PLPPEVDSYL PLPSTSES PLPSTSEL PLPENVSELL PLPENVSELL PLPENVSELL PLYDQSEYE SIPEGVSES SADITPDELK SIPIDADEME SIPIDADEME SIPIDADEME SIPIDADEME PLYDQSEYE PPE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-DAVLGK PME-DAWLGK PLE-DVVIGK	FAGLLR FAGLLG FAGLG FAGLR FAGFIR FAGFUR CSFLQ LCTFLQ LTGFMR YVGAIL LAAGML FAAALL FAAALL FAAALL FAAALL FAAALL AAAGML AAAAGML AAAAAAGML AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	GTKTEVVKCI NKKIELVQCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEKH GAKLLY-KNT GKKARVMPCL GEPYRIVKSK GGPIGVFL GGSLGVFL GGEIGVFL GGEIGVFL GGELELVR-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGELTLIK-C NGEVDI GALLEVODI CTLIFEVODI	300 SND-LEVPAS TID-LEVPAN TVD-LEVPAN TVD-LEVPAN TVD-LEVPAN PNG-IHVPAN PQSGHELLCD PDTPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG TND-LMVPAT PVHGNPVP-L GG-IPTPV-D EGVDMEVP-P G-DLEVP-Q GPVEMVA-H l.VP 400 YLPPEGCSYR WFPPEGCSYR FLPMEALSYK FLPMEALSYK FLPMEALSYK FLPMEALSYK FLPMEALSYK FLPMEALSYK FLPMEALSYK FLPMEALSYK FLPMEALSYK FLPMEALSYK
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus Consensus	201 GIYRQQLIGK GIYRMQVVSE GVYRLQLQSK GLYRLQVDK GYRLQVDK GWYRVQRFDQ GMYRWQRFDR GMYRWQRFDR GMYRLQVIGT SIARGMVVDD GIYRMQVKGK GIYRLQVIGT SIARGMVVDD SIHRLLIGR SIHRLLVKDD SIHRLLUKDD SIHRMLVLDD SIGRFQVIGK GIYR, qv 301 AEIVLEGYIE SEIVLEGYIE SEIVLEGYIE AEILIEGYVR AEIVLEGYID SEFILTGEAV	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTGMHWQIH KEGVVHWQIH ETUGLHFQIQ DTLGLHFQIQ NTMGLHFQIQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP YLVIRMV-EQ ERAVIRL-VP CRLAVRI-VP CRLAVRI-VP DRLAVRI-VP DRLAVRI-VP DRLAVRI-VP DRLAVRI-VP DRLAVRI-VP DRLAVRI-VP DRLAVRI-VP DRLAVRI-VP DRLAVRI-VP DRLAVRI-VP DRLAVRI-VP CREVP-EGP PEEPLVDEGP VDEMKV-EGP LEDLRE-EGP	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KDGNHHYWK- KHGAEHFRKM KRGMQAQODS KGGGAHFYE- KGGGMHFYE- KGSGMHFYE- KDSQLFFHE- HDIAIHIRQ- HDVAQHIRK- QHIRQIADSW RHLWHLYRK- RHLWIYRK- RHLWIYRK- RHLHFLYNK- RHLFLYNK- RHLFLYNK- RHLFLYRK- RHLFLYRK- RHLFLYRK- T1.k. YGDHTGY-YN FGDHTGY-YN FGDHTGY-YN FGDHTGY-YN FGDHTGY-YN FGDHTGY-YS	CAAHPGERFE -KTKKKELFP -AAEQGKKLE -AKRLGKKLE -AKRLGKKLE -AEKGGGKIE -VEKGERRIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEEKGEDLE -AEEKGEDLE -AEEKGEDLE -ALKEKGYLD -SVEHGEELE -AQKLGKDLE -AEERGEDLE -AEERGEDLE -AEERGEDLE -AEERGEDLE -AERGE	VSVALGADEA AAVVIGANPV VÄLAIGCEPP VÄLAIGCEPP VÄLAIGCEPP VÄLAIGCEPP VSIAIGVDPA AALVIGSDPG VTVFLSGNPF VSVALGRVPH IAITLGNDPV FÅLCFGVPPA VAIVIGVHPA	TILGAVT ITLAAVT ITLAAAT TLYAATA TLYAATA LILSAIA LILSAIA LILSAIA ITTAAST MIVAGM AILVSSM VILGAAT VLLAGST ULLAGST ILLASTTI LLASTT VHFAAAPAE laa.t IYHSTYTGRP IYLSTYTGRP	PVPDTLSEYA PIPENVSEYN PIPEVDSEYN PLPPEVDSYL PLPSGISEFM PVPYPICKHL PLPSGISEFM PLPSNVSELL PLPSNVSELL PLPSNVSELL PLPSNVSEL PLYDQSEYE PMAYDQSEYE SIPIDAUGSEYE SIPIDAUGSEYE SIPIDAUGSEYE SIPIDAUGSEYE SIPIDAUGSEYE SIPIDAUGSEYE P.PE PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PLE-DVVLGN PQE-DFFLGN	FAGLLR FAGLG FAGLG FAGIR FAGFIR FAGFIQ LCTFLQ LCTFLQ MAGALQ MAGALR YVGAIL FAAALL LAAGML FAAALL LAAGML FAAALL LAAGML FAAALL LAAGML FAAALL LAAGM	GTKTEVVKCI NKKIELVQCK GSGVALAKCK KERLKVTECE GEGLPVYR-L SKLHY-KND GAKLHY-KND GAKLHY-KNT GAPAPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GRPLEVIN-L GRPLEVIN-L GAPLELVAGT G. QKQFPEIVDF HQQFPEIVDF RQQVSEITDF KFNLPEVVDI RMMHPEIVDI QTLLPEVVDI	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LEVPAN TVD-LLYPAN POSGHPLLCD PDTPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG LTG-FDVPWG TND-LMVPAT PVHGNPVP-L DN-GLLVP-E GC-DLEVP-Q VGPVEMVA-H 1.VP 400 YLPPEGCSYR WFPPEGCSYR FLPMEALSYK HLPAEGCFHN NLPVEGAFHN DLPPHGVPCG KSYGEAGFHA
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus 1. 2. 3. 4. 5. 6. 7. 8. 9. 9. 9. 10. 11. 12. 13. 14. 15. 15. 16. 17. 18. 19. 19. 19. 19. 19. 19. 19. 19. 19. 19	201 21 201 201 201 201 201 201 201 201 2	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH ETUGLHFQIQ DTIGLHFQIQ NHLGLHFQIQ NHLGLHFQIQ NHLGLHFQIQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP YLVIRMV-EQ ERAVIRL-VP DRLAVRI-VP KRLVIRI-VP DRLAVRI-VP DRLAVRI-VP DRLVINIDAG 1 QGETAP-EGP PEEPLVDEGP PEEPLVDEGP PEEPLVDEGP CMERRP-EGP CNERRP-EGP	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KLGNHYWK- KHGAEHFRMM KRGMQAQDS KGGGAHFFE- KGSGMHIYE- KDSQLFFHE- HDIAIHRQ- HDIAIHRQ- HDIAIHRQ- HDIAIHRQ- RHLWHLYRK- RHLWHLYRK- RHLWYYLQK- RHLWYYLQK- RHLGLYLDR- I1.k. YGDHTGY-YM FGDHTGY-YM FGDHTGY-YM FGDHTGY-YT FGDHTGY-YD FGDHFGY-YS FGDHFGY-YS	CAAHPGERFE -KTKKKELFP -AAEAGGKKLE -AKRLGKKLE -AKRLGKKLE -VEKGERRIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEEKGEDIP -AEEKGEDIP -AEEKGEDIP -AKEKGYLD -SVEHGEELE -AEERGEDIP -ACKLGKDLE -AAARGEPLA .aglp EVDS-FPVFT GVED-SPLVR PVDK-YPOMH PPEP-YPVFH KPSRLFPTFR LTHD-FPIFK LQHD-FPAFK LQHD-FPAFK	VSVALGADPA AAVVIGANPV VALAIGGEPP VAVAIGVDPA AALVIGSDPG VTVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSIAIGODLI IAITIGNDPA TAIVVGCHPA VAIVIGVHPA VAIVIGNA VAIVIGNA VAIVIGVHPA VAIVIGN	TILGAVT ITLAAVT ITLAAAT TLIGGAM LILGAIA LILGAIA LTLSAIA LTLSAIA LTLSAIA LTLSAIA VTWCATA VINAGM AILVSSM VLLAAT VLLAGT VLLAGT LLLATT ILLATT ILLASTT VHFAAAPAE laa.t IYHSTYTGRP IYLSTYTGRP	PVPDTLSEYA PIPENVSEYN PIPEVDSYL PLPPEVDEYL PLPPEVDEYL PLPPEVDEYL PLPPEVDEYL PLPPEVDEYL PLPPISELL PLPPISELL PLPNSELL PLPNSELL PLPNSELE SIPEGVESG SADITPDELK SVAYGVSELE SIPIDADEME DAPVETDELG P.DE PDE-PAVLGV PDE-PAVLGV PME-DAWLGK PME-DAWLGK PME-DAWLGK PGE-DFFLGN YQE-DFYLGN	FAGLLR FAGLLG FAGLIG FAGIIR FAGFIR FAGFIQ LCTFLQ LTFLQ FATFLQ MAGALR YUGAIR YUGAIL LAAGML FAALL IASAISLKAF VANTFK IASAISLKAF VANTFK IASAFK IASAFK IASAFK IASAFK IASAFK IASAFK IASAFK IASAF	GTKTEVVKCI NKKIELVQCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEKH GAKLLY-KKT GAKLLY-KKT GEPYRIVKSK GAPAPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GGPLEVIN-L GGELELVR-C NGELTLIK-C GAPLELVAGT 9 QKQFPEIVDF RQQVSEITDF KFNLPEVVDF RMMHPEIVDI PLLMPGVRL PLVMPGVRL	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LEVPAN TVD-LLVPAN PQSGHPLLCD PDTPHPLLYD NDHPHPLLYD NDHPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG LTG-FDVPWG LD-LMVPAT PVHGNPVP-L DN-GLLVP-E GC-DLEVP-Q VGPVEMVA-H 1.VP 400 YLPPEGCSYR FLPMEALSYK HLPAEGCFHN NLPVEGGFHN NLPVEGAFHA KSYGEAGFHA
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus 1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 19. 10. 19. 10. 10. 11. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 19. 19. 19. 19. 19. 19. 19. 19	201 21 201 201 201 201 201 201 201 201 2	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH ETLGLHFQIQ NTMGLHFQIQ NTMGLHFQIQ NHLGLHWQIH DRLGIQPVPQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP YLVIRMV-EQ ERAVIRL-VP DRLVINIDAG 1 CGETAP-EGP PGEMLP-DGP PEEPLVDEGP VDEMRV-EGP ZAGERRP-EGP CNERRP-EGP AGERRP-EGP	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KGSGMHIYE- KGSGMHIYE- KDSQLFFHE- HDIAIHIRQ- HDVAQHLRK- QHIRQIADSW RHLWHLYRK- RHLYTYLQK- RHLYTYLQK- RHLYTYLQK- RHLGLYLDK- rl.k. YGDHTGY-YM FGDHTGY-YM FGDHTGY-YM FGDHTGY-YD FGDHTGY-YS FGDHTGY-YS FGDHFGY-YS FGDHFGY-YS	CAAHPGERFE -KTKKKELFP -AAEQGKKLE -AKRLGKKLE -EKGGGKIE -VEKGERRIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEERGINE -AEERGEDIP -ARERGEDIP -ARERGEDIP -ARERGEDIP -ALEKGYID -SVEHGEELE -AEERGEDIE -AQKLGKDIE -AAARGEPIA .ag1p EVDS-FPVFT DVEE-FPIFTF UVES-FPVFT DVEE-FPIFTF LTHD-FPIFK LQHD-FPEFH FIEP-YPVLE	VSVALGADEA AAVVIGANPV VALAIGGEPP VALAIGGEPP VAVAIGVDEA AALVIGSDEG VTVFLSGNPF VSVELSGNPF VSVELSGNPF VSVELSGNPF VSTAIGCDLU IAITLGNDPV IAITLGNDPV IAITLGNDPA TRIVIGNEVH IAIAIGMDFA FTLNVGVGFG VALVIGVHEA VTLIGMEVH IAIAIGMDFA FTLNVGVGFG VALVIGNEVH IAIAIGMDFA FTLNVGVGFG VALVIGNEVH IAIAIGMDFA FTLNVGVGFG VALVIGNEVH IAIAIGMDFA CRIVIGNEDA VTHITQREDA VTHITQREDA VTHITQREDA CRIVINKKDA CRIVINKKDA CRIVINKKDA VKTISYKKDS	TILGAVT ITLAAVT ITLAATA TLIGGAM LILSAIA LTLSAIA LTLSAIA TLTSAIA YTWGATA NIVAGM AILVSSM VLLAGAT LLLAATT ULLAGAT LLLAGAT LLLAGAT ULLAGAT ILLASTT VHFAAAAPAE laa.t IYHSTYTGRP IYLSTYTGRP IYLSTYTGRP IYLSTYTGRP IYLSTYTGRP IYLSTYTGRP IYLSTYTGRP IYLSTYTGRP IYLSTYTGRP IYLATVVGRP IYPATVVGRP IYPATVVGRP IYLATVVGRP	PVPDTLSEYA PIPENVSEYN PIPENVSEYN PIPEVDLSEWL PLPEVDEYL PLPENVDELL PLPENVSELL PLPENVSELL PLPENVSELL PLPENSEL PLPENSEL SPBLCQSEYE SIPECVSESD SIPECVSESD SIPECVSESD SIPECVSESD SIPECVSESD SIPECVSESD PLLYQSEYE PLLYQSEYE PLLYQSEYE PLLYQSEYE PLLYQSEYE PLLYQSEYE PLS-DVLGW PME-DAWLGW	FAGLLR FAGLLG FAGLG FAGIR FAGFIR FAGFIR FATFLQ LCTFLQ FATFLQ MAGALR YVGAIL FAALL FAALL FAALL IASAISLKAF VANAFK VANAFK VANAFK IASAFH .ag ALNEVFVPIL ALNEIFIPIL ALNEIFIPIL ATERIFLPIL ATERIFLPIL RLOEYLSPLF KLOEYLSPLF	GTKTEVVKCI NKKIELVOCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEKH GAKLLY-KKT GAKLLY-KKT GAKLLY-KKT GEYRIVKSK GAPAPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GRPLEVIN-L GGELELVR-C NGELTLIK-C GAPLELVAGT 9 QKQFPEIVDF HQQFPEIVDF KFNLPEVVDY RMMHPEIVDI QTLLPEVVDI PLIMPGVQLL PMVMPGVRL QMNAPNLIEY	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LEVPAN TVD-LLVPAN PQSGHPLLCD PDTPHPLLYD NDHPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG TND-LMVPAT PVHGNPVP-L DN-GLLVP-E GC-DLEVP-Q VGPVEMVA-H 1.VP 400 YLPPEGCSYR FLPMEALSYK HIPAEGCFHN NLPVEGAFHN NLPVEGAFHA KSYGEAGFHA KSYGEAGFHA KSYGEAGFHA
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus 1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 19. 10. 11. 19. 19. 10. 10. 11. 11. 15. 16. 17. 18. 19. 19. 19. 19. 19. 19. 19. 19. 19. 19	201 21YRQQLIGK GIYRQQVSE GVYRLQLQSK GLYRLQVLDK GMYRMQVFDG GWYRVQRFDG GMYRWQRFDR GMYRVQRFDR GIYRMQVKGK GIYRLQVIGT SIARGMVVDD SIHRLLVKDD SIHRLLVKDD SIHRMLVLDD SIHRMLVLDD SICRFQVIGK GIYR.qv 301 AEIVLEGYIE SEIVLEGYIE SEIVLEGYIE AEIVLEGYID AEIVLEGYID AEITLGESVD AEITLGESVD AEITLGESVD	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH ETLGLHFQIQ DTLGLHFQIQ NTMGLHFQIQ NTMGLHFQIQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP URIXTHV-EQ ERAVIRL-VP DRLVIRI-VP DRLVINIDAG .1 QGBTAP-EGP DRLVINIDAG .1 QGBTAP-EGP CEMLP-DGP PEEPLVDEGP VDEMRV-EGP LEDLRE-EGP AGERRP-EGP CEELEL-EGP AGEREY-EGP	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KGSGMHLYE- KGSGMHLYE- KDSQLFFHE- HDIAIHIRQ- HDVAQHLRK- QHIRQIADSW RHLWHLYRK- RHLFLYNK- RHLFLYNK- RHLYTYIQK- RHLGYTYQZ- RHLGLYLDK- rl.k. YGDHTGY-YM FGDHTGY-YM FGDHTGY-YT FGDHTGY-YS FGDHFGY-YS FGDHFGY-YS FGDHFGY-YT FGDFFGY-YS	CAAHPGERFE -KTKKKELFP -AEACGKKLE -AKRLGKKLE -VEKGERRIP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEQKKQNLP -AEERGINLP -AEERGEDLP AIGKANEIP -AEERGEDLP -AERGEDLP -ALEKGYLD -SVEHGEELE -AQKLGKDLE -AAARGEPLA .ag.lp EVDS-FPVFT DVEE-FPIFT UVEE-FPIFT KPSRLFPTFR LTHD-FPIFK LQHD-FPAFK LQHD-FPAFK LQHD-FPIFK	VSVALGADĚA AAVVIGANPV VATAIGGEPP VATAIGGEPP VATAIGSDPG VTVFLSGNÞF VSVFLSGNÞF VSVFLSGNÞF VSVFLSGNÞF VSTALGCEPV IATLGKDPV FALCFGVÞPA ATVVGLHPA VALVIGNPVH IATAIGMDPA FILNUGVGPG VALVIGNPVH IATAIGMDPA FTLNUGVGPG VALVIGNPVH IATAIGMDPA VTNITUREDA VTNITUREDA VTNITUREDA VTNITUREDA VTNITUREDA VTNITUREDA VTNITUREDA VTNITURENA VTNITUREDA VTNITUREDA VTNITUREDA VTNITUREDA VTNITUREDA VTNITUREDA VTNITUREDA VTNITUREDA VTNITUREDA VTNITUREDA VTNITUREDA	TILGAVT ITLAAVT ITLAAAT TLITGAM TLITGAM LILSAIA LTLSAIA TTTAAST MIVAGM AILVSSM VLIAGST LLLAGAT TLLAGAT TLLAGAT TLLAGAT LLLAGAT LLLAGAT ILLAGAT ILLAGAT ILLAGAT TLLASTT VLFAAAAPAE laa.t IYHSTYTGRP IYLSTYTGRP IYLSTYTGRP IYLSTYTGRP IYLSTYTGRP IYLSTYTGRP IYLSTYTGRP IYLSTYTGRP IYLSTYTGRP IYLSTYTGRP IYPATVVGRP IYPATVVGRP IYPATVVGRP IYLATVVGRP IYLATVVGRP IYLATVVGRP	PVPDTLSEYA PIPENVSEYN PIPENVSEYN PIPEVDLSEWL PLPEVDLSEWL PLPEVDEYLL PLPENVSELL PLPENVSELL PLPENVSELL PLPENVSELL PLYDQSEYE PALYQSEYE SADITPDELK SVAYGVSELE SIPIDADEME AAPVETDELG P.PE PDE-PAVLGV PDE-PAVLGV PDE-DAVLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK	FAGLLR FAGLLG FAGLG FAGIR FAGFIR FAGFIR CSFLQ LCTFLQ LTGFMR YUGAI LYGFMR YUGAIL LAAGML LAAGML IASAISLKAF VANAFH .AGAI ALNEVFVPIL ALNEIFIPIL ALNEIFIPIL ALNEIFIPIL ALNEIFIPIL ALNEIFIPIL ATERIFLPII FAERIFLPAI KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF LTERLFLPLL	GTKTEVVKCI NKKIELVOCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEH GSKLHY-KND GAKLLY-KKT GKKARVMPCL GEPYRIVKSK GGPJATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GRPLEVIN-L EGELELVR-C GAPLELVAGT 9 QKQFPEIVDF KFNLPEVVDY RMMHPEIVDI QTLLPEVVDY RMMHPEIVDI PLIMPGVQLL PLVMPGVRLL CMNAPNLIEY KEAYPNEIVA	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LEVPAN TVD-LLVPAN PQSGHPLLCD DDTPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG TND-LMVPAT PVHGPVP-L DN-GLLVP-E DG-IPTPV-D EGVDMEVP-P G-DLEVP-Q VGPVEMVA-H 1.VP 400 YLPPEGCSYR WFPPEGCSYR WFPPEGCSYR HLPAEGCFHN NLPVEGAFHN DLPPHGVFQG KSYGEAGFHA KSYGEAGFHA KSYGEAGFHA
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus 1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 17. 18. 19. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 10. 11. 12. 13. 14. 15. 16. 17. 17. 18. 19. 10. 11. 12. 13. 14. 15. 16. 17. 17. 18. 19. 10. 11. 12. 13. 14. 15. 16. 17. 17. 18. 19. 10. 17. 17. 18. 19. 19. 10. 17. 18. 19. 19. 10. 11. 12. 13. 14. 15. 16. 17. 17. 17. 18. 19. 17. 17. 18. 19. 17. 17. 17. 18. 19. 17. 18. 19. 19. 19. 19. 19. 19. 19. 19. 19. 19	201 201 GIYRQQVIGK GIYRLQVSE GVYRLQVSE GVYRLQVDK GYRLQVIDK GMYRMQVFDG GWYRVQRFDR GMYRVQRFDR GMYRVQRFNQ GMYRLQVIGT SIRGUVGKGK GIYRLQVIGT SIRRILVKDD SFHRLLIGK SFHRMVLDE SFHRMVLDE SIRRILVKDD SFRFQVIGK GIYR.QV 301 AEIVLEGYIE SEFVLEGYIE AEIVLEGYIE AEIVLEGYIE AEIVLEGYIE AEIVLEGYIE AEIVLEGYIE AEIVLEGYIE AEIVLEGYIE AEIVLEGEI AEFVLEGEI SEFVLEGEI SEFVLEGEI SEFVLEGEI	NKLIMRWLSH NKLIMRWLKL TTMTVHWLSV DKLAVHWQIH KTGMHWQIH ETIGLHFQIQ DTIGLHFQIQ NTMGLHFQIQ N	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KDGNHHYWK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KGSGMHIYE- KDSQLFFHE- HDIAIHIRQ- HDVAQHIRK- QHIRQIADSW RHLWHLYRK- RHLYTYLQK- RHLYTYLQK- RHLYTYLQK- RHLYTYLQK- RHLYTYLQK- RHLGYLDK- I1.k. YGDHTGY-YM FGDHTGY-YM FGDHTGY-YT FGDHTGY-YS FGDHFGY-YS FGDHFGY-YS FGDHFGY-YS FGDHFGY-YS FGDHFGY-YS	CAAHPGERFE -KTKKKELFP -AEACGKKLE -AKRLGKKLE -AEKGGKIE -VEKGERRIP -AEQKKONLP -AEQKKONLP -AEQKKONLP -AEQKKONLP -AEERGINLP -AEERGINLP -AEERGEDLE -AERGEDLE -ACKLGKDLE -AAARGEPLA .ag.lp EVDS-FPVFT GVED-SPLVR PVDK-YPOMH PPEP-YPVFH KPSRLFPTFF LQHD-FPIFK LQHD-FPIFK LQHD-FPIFK GGRR-MPVIR	VSVALGADPA AAVVIGANPV VAIAIGSPP VAIAIGSPP AALVIGSDPG VTVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSTALGCEPV FALCFGVPPA ATVVGLHPA ATVVGLHPA VAIVIGNPVH IAIAIGMDPA IAIAIGMDPA IAIAIGMDPA VAIVIGNPVH IAIAIGMDPA VAIVIGNPVH IAIAIGMDPA VAIVIGNPVH IAIAIGMDPA VTIVLGNPVH IAIAIGMDPA VTIVLGNPVH IAIAIGMDPA VTIVLGNPVH IAIAIGMDPA VTIVLGNPVH VIIIGNPVH VIIG	TILGAVT ITLAAVT ITLAAAT TLLTGAM TLLTGAM LILSAIA LTLSAIA ITTAAST ITTAAST ITTAAST ILLAAT LLLAGAT LLLAGAT LLLAGAT LLLAGAT LLLAGAT LLLASTT ILLASTT ILLASTT ILLASTT ILLASTT ILLASTT ILLASTT ILLAST	PVPDTLSEYA PIPENVSEYN PIPEVDSEWL PLPEVDSEUL PLPEVDEYL PLPENVSELL PLPENVSELL PLPENVSELL PLPENVSELL PLPENVSEL PLYDQSEYE PLLYDQSEYE SIPICAUSEYE SIPICAUSEYE SIPICAUSEYE PLLYDQSEYE PLLYDDY PLLYDDY PLLYDDY PLLYDDY PLLYDDY PLLYDDY PLLYDDY PLLYDDY PLLYDY P	FAGLLR FAGLLG FAGLG FAGIR FAGFIR FAGFIR CSFLQ LYGFMR MAGAIQ MAGALR YVGAIL LAAGML IASAISLKAF VANTFH JASAISLKAF VANTFK IASAFK IASAFK JASAFK JASAFK JASAFK JASAFK JASAF	GTKTEVVKCI NKKIELVOCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEK GAKLLY-KNT GKKARVMPCL GEPYRIVKSK GAPAPIATAP GESVPVVKCE GGEIGVFE-L GRPLEVIN-L EGELELVR-C NGELTLIK-C N	300 SND-LEVPAS TID-LEVPAN TVD-LEVPAN TVD-LEVPAN TVD-LEVPAN PQSGHPLLCD PDTPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG TND-LMVPAT PVHGNPVP-L DG-IPTPV-D EGVDMEVP-P G-DLEVP-Q VGPVEMVA-H l.VP 400 YLPPEGCSYR WFPPEGCSYR WFPPEGCSYR WFPPEGCSYR SYGEAGFHA KSYGEAGFHA KSYGEAGFHA KSYGEAGFHA KSYGESGFHA KSYGESGFHA
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus 1. Consensus 1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 10. 11. 12. 13. 14. 15. 16. 17. 17. 18. 19. 19. 10. 11. 12. 13. 14. 15. 16. 17. 17. 18. 19. 19. 10. 11. 12. 17. 18. 19. 19. 10. 11. 12. 17. 18. 19. 19. 19. 10. 17. 17. 18. 19. 19. 19. 10. 11. 12. 17. 17. 18. 19. 19. 19. 19. 19. 19. 10. 17. 18. 19. 19. 19. 19. 19. 19. 19. 19. 19. 19	201 21 201 201 201 201 201 201 201 201 2	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV KTAGHHWQIH KEGVVHWQIH KEGVVHWQIH ETIGLHFQIQ DTIGLHFQIQ NTMGLHFQIQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP YLVIRMV-EQ ERAVIRL-VP KRLVIRI-VP DRLAVIRI-VP CRLVIRI-VP DRLAVIRI-VP CRLVIRI-VP CRLVIRI-VP CRLVIRI-VP CRLVIRI-VP CRLVIRI-VP CRLVIRI-VP CRLVIRI-VP CRLVIRI-VP CRLVIRI-VP CRLVIRI-VP CRLVIRI-VP CRLVIRI-EGP CEMLP-EGP AGERRP-EGP AGERRP-EGP AGEREY-EGP SKRRRI-EGP	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KDGNHHYWK- KHGAEHFRKM KRGMQAQODS KGGGAHFYE- KGGGMHFYE- KGGGMHFYE- KDSQLFFHE- HDIAIHIRQ- HDVAQHIRK- QHIRQIADSW RHLWHLYRK- RHLHFLYNK- RHLHFLYNK- RHLHFLYNK- RHLHFLYNK- RHLHFLYNK- RHLHFLYNK- RHLYTYFQK- RHLGLYLQK- RHLGLYLQK- RHLGLYDK- T1.k. YGDHTGY-YN FGDHTGY-YN FGDHTGY-YN FGDHTGY-YS FGDHFGY-YS FGDHFGY-YS FGBHFGY-YS FGEFTGH-YS FGEFTGH-YS FGEFTGH-YS	CAAHPGERFE -KTKKKELFP -AAEQGKKLE -AKRLGKKLE -AEKGGGKIE -VEKGERRIP -AEQKKQNIP -AEQKKQNIP -AEQKKQNIP -AEQKKQNIP -AEERGINIP -AEERGINIP -AEERGINIP -AEERGEDLE -AEERGEDLE -ACKLGKDLE -AEERGEDLE -AAARGEPLA .ag.lp EVDS-FPVFT GVED-SPLVF VDK-YPQMH PPEP-YPVFH KPSRLFPTFR LTHD-FPIFK LQHD-FPEFH FIEP-YPVE GGRS-MPIIK GGRR-MPVIR SQGHPCPLYT	VSVALGADEA AAVVIGANPV VÄLAIGCEPP VÄLAIGCEPP VÄLAIGCEPP VÄLAIGCEPP VSVAIGVDPA AALVIGSDPG VTVFLSGNPF VSVELSGNPF VSVALGEPV IÄILGOPV IÄILGOPV IÄILGNDPA VAIVIGVHPA VAIVIG	TILGAVT ITLAAVT ITLAAAT TLYAATA TLYAATA TLYAATA LILSAIA LILSAIA LITSAIA TTYAATA YTWCATA VILGATA VILGAT VLLAGAT VLLAGAT VLLAGAT VLLAGAT ILLASTT VHFAAAPAE ILLASTT ILLASTT VHFAAAPAE ILLASTT ILLASTT ILLASTT VHFAAAPAE ILLASTT ILLASTT VHFAAAPAE ILLASTT ILLASTT ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT ILLASTT VHFAAAPAE ILLASTT ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT VHFAAAPAE ILLASTT ILLASTT VHFAAAPAE ILLASTT ILLASTT VHFAAAPAE ILLASTT-	PVPDTLSEYA PIPENVSEYN PIPENVSEYN PLPPEVDSYL PLPSGISEFM PVPYPICKHL PLPSGISEFM PUPENVSELL PLPSGISEFM PLPSNVSELL PLPSGISEFM PLPSNVSELL PLPYGSEYE PMAYDQSEYE PMAYDQSEYE SIPIDAUGSEYE SIPIDAUGSEYE SIPIDAUGSEYE SIPIDAUGSEYE PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-DAVLGV PDE-DAVLGV PDE-DAVLGV PQE-DFYLGN YQE-DFYLGN YQE-DFYLGN YQE-DFYLGN YQE-DFYLGN YDE-DXMIGI WNECDYLVGP	FAGLLR FAGLLG FAGLG FAGIR FAGFIR FCSFLQ LCTFLQ LTFLQ MAGALR YVGAIL IASGM FAAALL ALNEFFILL ALNEFFILL FAERIFLPIL ATERIFLPIL FAERIFL	GTKTEVVKCI NKKIELVQCK GSGVALAKCK KERLKVTECE GEGLPVYR-L SKLHY-KND GAKLHY-KND GAKLHY-KND GAKLHY-KNPCL GEVTIVKSK GAPAPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GRPLEVIN-L GRPLEVIN-L GAPLEVN-C NGELTLIK-C GAPLELVAGT G. QKQFPEIVDF HQQFPEIVDF KFNLPEVVDI QTLLPEVVDI QTLLPEVVDI PLIMPGVRQL PLVMPGVRL QMNAPNLIEY RAPFP-EVQA IESGLPILDA	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LEVPAN TVD-LEVPAN POSGHPLLCD PDTPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG LTG-FDVPWG LTG-FDVPWG TND-LMVPAT PVHGNPVP-L DN-GLLVP-E GC-DLEVP-Q VGPVEMVA-H 1.VP 400 YLPPEGCSYR WFPPEGCSYR FLPMEALSYK HLPAEGFHN DLPPHGVPCG KSYGEAGFHA KSYGESGFHA KSYGESGFHA YMPENGVFNG -VNAMYTHGL FMPYEAQALW
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus 1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 10. 11. 12. 13. 14. 14. 15. 16. 17. 18. 19. 19. 19. 10. 11. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 19. 19. 19. 19. 19. 19. 19. 19	201 GIYROQLIGK GIYROQVSE GVYRLQLQSK GLYRLQVDK GYRVQVFDG GVYRVMIAGD GMYRMQRFDR GMYRNQRFDR GMYRLQVIGT SIARGMVVDD SIHRLLIGR SIHRLVKDD SHRMLVLDD SHRMLVLDD SHRMLVLDD SHRMLVLDY SHRMLVLDY SHRLVLEGYIE SIVLEGYIE SELVLEGYIE SELVLEGYIE SELVLEGYIE SELVLEGYIE SELVLEGYIE SELVLEGYIE SELVLEGYIE SELVLEGYIE SELVLEGYIE SELVLEGYIE SELVLEGYIE SELVLEGYIE SEVVLEGEI SENVLEGYIE SENVLEGYIE SENVLEGVIE SENVEGTES	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV KTTGMHWQIH KTTGMHWQIH ETIGLHFQIQ NTMGLHFQIQ NHLGLHFQIQ NHLGLHVQIH PRIGIQPVPQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP YLVIRMV-EQ ERAVIRL-VP KRLVIRI-VP DRLAVRI-VP KRLVIRI-VP DRLVINIDAG 1 QGETAP-EGP PGEMLP-DGP PFEPLVDEGP PGEMLP-DGP PFEPLVDEGP CMERRP-EGP CMERRP-EGP CAERRP-EGP CAERRP-EGP CAERRP-EGP CAERRP-EGP SAKRAI-EGP LTDTHL-EGP	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KLGNHHYWK- KHGAEHFRMM KRGMQAQDS KGGGAHFFE- KGSGMHIYE- KGSGMHIYE- KDSQLFFHE- HDIAIHRQ- HDIAIHRQ- HDIAIHRQ- HDIAIHRQ- RHLWHLYRK- RHLWHLYRK- RHLWHLYRK- RHLWYYLQK- RHLWYYLQK- RHLGYLDK- I1.k. YGDHTGY-YN FGDHTGY-YN FGDHTGY-YN FGDHTGY-YT FGDHFGY-YS FGDHFGY-YS FGDHFGY-YS FGCFTGH-YS FGEFTGH-YS FGEFTGH-YS	CAAHPGERFE -KTKKKELFP -AAEAGGKKLE -AKRLGKKLE -AEKGGKLE -VEKGERRIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEEKGEDIP	VSVALGADEA AAVVIGANPV VALAIGGEPP VAVAIGVDEA AALVIGSDEG VTVFLSGNEF VSVELSGNEF VSVELSGNEF VSVELSGNEF VSVELSGNEF VSIVIGUERA VAIVIGUERA	TILGAVT ITLAAVT ITLAAAT TLYAATA TLYAATA TLYAATA LILSAIA LILSAIA LTLSAIA LTLSAIA LTLSAIA VICAGTA VICAGTA VICAGTA VLCAGT LLLAATT VLCAGT LLLAATT VLCAGT VLCAGT LLLAATT VLCAGT	PVPDTLSEYA PIPENVSEYN PIPEVDSEW PLPPEVDEYL PLPSGISEFM PVPYDISELL PLPSGISEFM PLPPUSELL PLPSGISEFM PLPYDSELL PLPYDSELL PLPYDSELE PLPYDSEYE SIPEOVESG SADITPDELK SVAYGVSELE SIPIDADEME SIPIDADEME P.DE PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-DKYLGW PME-DAWLGK PME-DAWLGK PME-DAWLGK PME-DAWLGK PLE-DVVISK FQE-DFFLGN YQE-DFFLGN PLE-DVVISK FQE-DFFLGN PLE-DVVISK FQE-DFFLGN SLEHVNIMGF	FAGLLR FAGLLG FAGLG FAGIR FAGFIR FAGFIQ LCTFLQ FATFLQ MAGALR YUGAIL LAAGML FAALL IASAISLKAF VANTF	GTKTEVVKCI NKKIELVQCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEKH GAKLLY-KND GAKLLY-KND GAKLLY-KKT GEVTIVKX GAPAPIATAP GESVPVVKCE GGSJKVYR-S GGEIGVFE-L GGSHKVYR-S GGEIGVFE-L GGSLELVR-C NGELTLIK-C GAPLELVAGT 9 QKQFPEIVDF RQQVSEITDF KFNLPEVVDT PLUMPGVRL PMVMFGVRL PLVMFGVRL QMNAPNLIEY KEAYPNEIVA RAEFP-EVQA IESGLPILDA RAALPRVRAV	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LLVPAN PVG-IHVPAN PQSGHPLLCD PDTPHPLLYD NDHPHPLLYD NDHPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG LTG-FDVPWG USD-LDVPWG LG-FDVPWG G-DLEVP-Q VGPVEMVA-H 1.VP 400 YLPPEGCSYR FLPMEALSYK HLPAEGCFHN NLPVEGGFHN NLPVEGCFHA KSYGEAGFHA KSYGEAGFHA KSYGEAGFHA KSYGEAGFHA SYMENGVFHGL FMPYEAQALW RLTPASGGWL
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus 1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 19. 10. 11. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 19. 19. 19. 19. 19. 19. 19. 19	201 201 21YRQQLIGK GIYRMQVYSE GVYRLQLQSK GLYRLQVLDK GMYRMQVFDG GWYRVQRFDR GMYRMQVFDK GMYRMQVFDK GMYRLQVJGT SIARGMVVDG SIHRLLIGE SIHRLVKDD SFHRMMVLDD SIHRLLVKDD SIHRLVLGYIE SIHRMVLDD SIHRMLVG SIRGFQVIGK GIYR.qV 301 AEIVLEGYIE SEIVLEGYIE SEIVLEGYIE SEFILIGEGY SEFILIGEGY SEFILIGESY SEFILIGESY SEFILIGESY SEFILIGESY SEFILIGESY SEFILIGESY SEVILEGYIE SEVVFEGTLS GAAMVADVWI AEIILEGVIE SEVVFEGTLS GAAMVADVWI	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV KTTGMHWQIH KTTGMHWQIH ETLGLHFQIQ NTMGLHFQIQ NTMGLHFQIQ NHLGLHWQIH DRLGIQPVPQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP YLVIRMV-EQ ERAVIRL-VP DRLAVRI-VP KRLVIRV-EQ DRLAVRI-VP DRLAVRI-VP DRLAVRI-VP DRLAVRI-VP DRLAVRI-VP DRLVINIDAG .1 CGETAP-EGP PEEPLVDEGP PEEPLVDEGP VDEMRV-EGP LAEYLP-EGP AGERRP-EGP CNERRP-EGP CNERRP-EGP CKLEL-EGP AGEREY-EGP SRKRRI-EGP LTDTHL-EGP	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KGSGMHIYE- KGSGMHIYE- KDSQLFFHE- HDIAIHIRQ- HDVAQHLRK- QHIRQIADSW RHLWHLYRK- RHLFLYNK- RHLYTYLQK- RHLHYTYLQK- RHLGLYLDK- I1.k. YGDHTGY-YN FGDHTGY-YN FGDHTGY-YT FGDHTGY-YT FGDHTGY-YT FGDHFGY-SS FGDHFGY-SS FGDHFGY-SS FGDHFGY-YT FGEFTGH-SS FGEFTGH-SS FGEFTGH-SS FGEFTGH-SS FGEFTGH-SY	CAAHPGERFE -KTKKKELFP -AAEQGKKLE -AKRLGKKLE -EKGGGKIE -VEKGERRIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AECKGINIP -AEERGINIP -AEERGEDIP -AEERGEDIP -ARERGEDIP -ARERGEDIP -ALKEKGYLD -SVEHGEELE -AEERGEDIE -ACKLGKDIE -AAARGEPIA .ag1p EVDS-FPVFT DVEE-FPIFT DVEE-FPIFT UVEE-FPIFT LTHD-FPIFK LQHD-FPEFH PIEP-YPVLE GGRR-MPVIR SQGHPCPLYT RVRR-QPVIR IVRR-QPIK	VSVALGADEA AAVVIGANPV VALAIGGEPP VALAIGGEPP VAVAIGVDEA AALVIGSDEG VTVFLSGNPF VSVELSGNPF VSVELSGNPF VSVELSGNPF VSVELGEPV IAITLGNDPV FALCFGVPEA ATVVVGLHEA ATVVVGLHEA ATVVVGLHEA TRIMCKEP FALCFGVPEA VALVIGNPVH IAIAIGMDEA FTLNVGVGFG vai.G.pa VTHITQREDA VTHITQREDA VTAITMKKP FOCLTHEKNP VTAITMKKP ITHITHEENP CNCLYHKKDA CRKIYHEKDA VKIISYKKDS IKRVYHENNP VERVSYRHEP VERVSYRHEP VERVSYRHEP VERVSYRHEP VKAMSYRDNA LEAAYI-KEG IEKLY-RKEK	TILGAVT ITLAAVT ITLAAAT TLIGAAA TLIGAAA LILSAIA LTLSAIA LTLSAIA YTWGATA NIVAGM AILVSSM VILAAAT VLLAGAT TLLATTT TLLATTT TLLATTT TLLATTT UHFAAAAPAE laa.t IYHSTYTGRP IYLSTYTGRP IYLSTYTGRP IYLATVVGKP IYPATVVGKP IYPATVVGKP IYPATVVGKP IYPATVVGKP IYPATVVGKP IYPATVVGKP IYPATVVGKP IYPATVVGKP IYPATVVGKP IYPATVVGKP IFEHLYLGMP VFESLLIGMP OFIFHALLP-G PIFHALLP-G	PVPDTLSEYA PIPENVSEYN PIPENVSEYN PIPENVSEYN PLPPEVDEYL PLPENVSELL PLPENVSELL PLPENVSELL PLPENVSELL PLPENVSEL SIPEGVSESD SPEGVSESD SPEGVSESD SIPEGVSESD SIPEGVSESD SIPEGVSESD SIPEGVSESD SIPEGVSESD PLLYOGSFYE PLLYOGSFYE PLLYOGSFYE PLLYOGSFYE PLLYOGSFYE PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-DKILGP PDE-DKILGP PME-DAWLGK FQE-DFYLGN YQE-DFYLGN YQE-DFYLGN PLE-DVVIGK FQE-DFYLGN SLEHVNLMGF GIEHKTLMGM	FAGLLR FAGLLG FAGLIG FAGIIR FAGFIR FAGFIQ LCTFLQ LTFLQ MAGALR YVGAIL IAAGML FAALL IAAGML FAALL IAAGML FAALL IAAGML FAALL IAAGML IAAGML FAALL IAAGML FAALL IAAGML IAAGML FAALL IAAGML FAALL IAAGM	GTKTEVVKCI NKKIELVOCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEKH GAKLLY-KKT GAKLLY-KKT GAKLLY-KKT GAVAPPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GRPLEVIN-L GGELELVR-C NGELTLIK-C GAPLELVAGT 9 QKQFPEIVDF HQQFPEIVDF ROQVSEITDF KFNLPEVVDY RMMHPEIVDI QTLLPEVVDI PLIMPGVQL PLVMPGVRL QMNAPNLIEY KEAYPNEIVA RAEFP-EVQA IESGLPILDA RRALPRVRAV	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LEVPAN TVD-LLVPAN PQSGHPLLCD PDTPHPLLYD NDHPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVFWG TND-LMVPAT PVHGNPVP-L DN-GLLVP-E GC-DLEVP-Q VGPVEMVA-H 1.VP 400 YLPPEGCSYR WFPPEGCSYR HIPAEGCFHN NLPVEGAFHN NLPVEGAFHN NLPVEGAFHN SYGEAGFHA KSYGEAGFHA KSYGEAGFHA KSYGEAGFHA KSYGEAGFHA KSYGEAGFHA VMPENGVFHN -VNAMYTHGL FMPYEAQALW RLTPASGCWL RLTEGGCCWL
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus 1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 10. 11. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 10. 11. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 10. 11. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 10. 11. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 10. 11. 11. 12. 15. 16. 17. 18. 19. 19. 10. 11. 11. 12. 15. 16. 17. 18. 19. 19. 19. 10. 11. 11. 12. 18. 19. 19. 10. 11. 11. 12. 17. 18. 19. 19. 10. 11. 11. 12. 15. 16. 17. 18. 19. 19. 10. 11. 12. 18. 19. 19. 10. 11. 12. 18. 19. 19. 10. 11. 12. 19. 19. 10. 11. 11. 12. 13. 14. 15. 16. 17. 18. 19. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 10. 11. 11. 12. 15. 16. 17. 17. 18. 19. 10. 11. 11. 12. 11. 11. 12. 11. 11. 12. 12	201 201 21YRQQLIGK GIYRQQVSE GVYRLQLQSK GLYRLQVLDK GYYRLQVLDK GYYRVQRFDG GWYRVQRFDR GMYRVQRFDR GMYRLQVIGT SIARGMVVDG SIHRLLUKDD SIHRLLUG SIHRLLVKDD SIHRLLVKDD SIGRFQVIGK GIYR, qv 301 AELVLEGYIE SEIVLEGYIE AEIVLEGYIE AEIVLEGYIE AEIVLEGYIE AEIVLEGYIE AEIVLEGYIE AEIVLEGYIE AEIVLEGYIE SEFFLIGEGI AEIVLEGEI SEFFLIGESP CDIVIEGFVD AEVVLEGEI SEVVFEGTLS GAAMVADVWI AEIILEGKIL SEFVFKAKIT AEIILEGKIL	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH KTGMHWQIH ETLGLHFQIQ NTMGLHFQIQ N	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KGSGMHLYE- KGSGMHLYE- KDSQLFFHE- HDIAIHIRQ- HDVAQHLRK- QHIRQIADSW RHLWHLYRK- RHLYTYLQK- RHLYTYLQK- RHLYTYLQK- RHLYTYLQK- RHLYTYLQK- T1.k. YGDHTGY-YM FGDHTGY-YM FGDHTGY-YM FGDHTGY-YM FGDHTGY-YS FGDHFGY-YS FGDHFGY-YS FGDHFGY-YS FGDHFGY-YS FGEFTGH-YS FGEFTGH-YS FGEFTGH-YS FGEFTGH-YS FGEFTGH-YD FVDLTGT-YD FVDLTGT-YD	CAAHPGERFE -KTKKKELFP -AAEQGKKEE -AKRLGKKLE -EKGGGKIE -VEKGERRIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AECKGEDIP AAIGKANEIP -AEERGEDIP -ARERGEDIP -ARERGEDIP -ALEKGYLD -SVEHGEELE -AEERGEDIP -ACKLGKDE	VSVALGADEA AAVVIGANPV VATAIGGEPP VATAIGGEPP VATAIGGEP VTVLISGNEF VSVELSGNEF VSVELSGNEF VSVELSGNEF VSVELSGNEF VSVELSGNEF VSTVUGGE VATVIGKEP ATVVVGLHPA ATVVVGLHPA ATVVVGLHPA TATLGMDPA TATLGMDPA TATLGMDPA TATVGVGEG VALVIGNEPVH IATAIGMDPA TTLNVGVGEG VALVIGNEPVH IATAIGMDPA TTLNVGVGEG VALVIGNEPVH VALVGVER VALVIGNEPVH VALVGVER VALVG	TILGAVT ITLAAVT ITLAAAT TLIGAA TLIGAA LILSAIA LTLSAIA TLISAIA ITLSAIA ITLSAIA YTWGATA AILVSSM VILGAST TLLAGAT TL	PVPDTLSEYA PIPENVSEYN PIPENVSEYN PIPEVDLSEWL PLPEVDLSEWL PLPEVDSELL PLPENVSELL PLPENVSELL PLPENVSELL PLPENSELL PLFUGSEYE PARTOSEYE SPECVSESD SPELGYFEIG SADITPDELK SVAYGVSELE SIPIDADEME P.L.YOQSEYE PARTOSEYE PARTOSEYE PARTOSEYE PLC-PARTOSEY PDE-PARTOSEY PDE-PARTOSEY PDE-PARTOSEY PDE-DARTOSEY PDE-DARTOSEY PLE-DKYLGW PCE-DFYLGN YQE-DFYLGN YQE-DFYLGN YQE-DFYLGN WHECDYLVGP SLEHVNLMGF GIEHKTLMGG GFEHELLQGL	FAGLLR FAGLLG FAGLIG FAGIIR FAGFIR FAGFIQ LCTFLQ LTFLQ FATFLQ LYGFMR YVGAIL LAAGML LAAGML IASAISLKAF VANAFH .AAGALH .AGAL	GTKTEVVKCI NKKIELVOCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEKH GSKLHY-KND GAKLLY-KKT GKKARVMPCL GEPYRIVKSK GAPAPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GRPLEVIN-L EGELELVR-C GAPLEVNF- GAPLEVIN-L GAPLEVAGT 9 QKQFPEIVDF HQQFPEIVDF KFNLPEVVDF KFNLPEVVDY RMMHPEIVDI PLIMPGVQDL PLVMFGVRL QMNAPNLIEY KEAYPNEIVA RAEFP-EVQA RESGLPILDA RRALPRVRAV RNTVPTVKNI	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LEVPAN TVD-LLVPAN PQSGHPLLCD DDTPHPLLYD NDHPHPLLYD SNS-LSVPSD LTG-FDVPWG TND-LMVPAT PVHGPVP-L DN-GLVP-E DG-IPTPV-D EGVDMEVP-P GC-DLEVP-Q VGPVEMVA-H 1.VP 400 YLPPEGCSYR WFPPEGCSYR WFPPEGCSYR WFPPEGCSYR NLPVEGAFHN NLPVEGAFHN NLPVEGAFHN SYGEAGFHA KSYGEAGFHA CNAMYTHGL -VNAMYTHGL VLTEGGCCWL
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus 1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 10. 17. 18. 19. 10. 17. 18. 19. 10. 17. 18. 19. 10. 17. 18. 19. 10. 17. 18. 19. 10. 17. 18. 19. 10. 17. 18. 19. 19. 19. 10. 19. 19. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 19. 19. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 10. 17. 18. 19. 19. 19. 19. 19. 19. 19. 19. 19. 19	201 201 21YRQQLIGK GIYRQQVSE GVYRLQLQSK GLYRLQVLDK GYRLQVLDK GYYRVQFDG GYYRVMIAGD GMYRMQVFDG GYYRVQRFDR GIYRMQVKGK GIYRLQVIGT SIARGMVVDD SIHRLLIK SIHRLLVKDD SIHRLLVKDD SIHRMVLDD SIHRMVLDD SIHRMVLDD SIFRCVIGK GIYR.qV 301 AEIVLEGYIE SEFVLEGYIE AEIVLEGYIE AEIVLEGYID AEILIGGYD AEIVLEGYID AEILIGESVD AEIVLEGYID SEFILIGESVD AEIVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII SEVVLEGEII	NKLIMRWLSH NKLIMRWLKL TTMTVHWLSV DKLAVHWQIH KTTGMHWQIH KTGMHWQIH KTGMHWQIH DTIGLHFQIQ NTMGLHFQIQ NTMGLHFQIQ NTMGLHFQIQ NTMGLHFQIQ NTMGLHFQIQ NTMGLHFQIQ NTMGLHFQIQ NTMGLHFQIQ NTMGLHFQIQ RIATHQI RIATH	RGGALDYQEW RGGAEHHKRW RGGARHLRK- KDGNHHYWK- KHGAEHFRKM KRGMQAQQDS KGGGAHFFE- KGGGMHLYE- KDSQLFFHE- HDIAIHIRQ- HDVAQHLRK- QHIRQIADSW RHLWHLYRK- RHLFLYNK- RHLFLYK- RHLFLYK- RHLFLYK- RHLFLYK- RHLFYYFQ- RHLGLYLDK- r1.k. YGDHTGY-YM FCDHTGY-YM FCDHTGY-YT FCDHTGY-YT FCDHTGY-YS FCDHTGY-YS FCDHFGY-S FGCHFGY-S FGCHFGY-S FGCHFGY-YS FGCHFGY-YS FGCHFGY-YS FGCHFGY-YS FGCHFGY-YS FGCHFGY-YS FGCHFGY-YS FGCHFGY-YS FGCHFGY-YS FGCHFGY-YD FGCHFGY-YD FGCHFGY-YD FGCHFGY-YD FGCHFGY-YD FGCHFGY-YD FGCHFGY-YD FGCHFGY-YD FGCHFGY-YD FGCHFGY-YD FGCHFGY-YD FGCHFGY-YD FGCHFGY-YD FGCHTGT-YD FVDLTGT-YD FVDLTDT-YD	CAAHPGERFE -KTKKKELFP -AEACGKKLE -AKRLGKKLE -AEKGGGKIE -VEKGERRIP -AEQKKONLP -AEQKKONLP -AEQKKONLP -AEQKKONLP -AEERGIN -AEERGEDIP AIGKANEIP -AEERGEDIP -AERGEDIP -AEERGEDIP -AEERGEDIP -ALKEKGYLD -VHGEEFE -AQKLGKDIE -AAARGEPLA .ag.lp VDS-FPVFT VUS-FPVFT VUS-FPVFT VUS-FPVFT VUS-FPVFT VUS-FPVFT VUS-FPVFT VUS-FPVFT VUS-FPVFT VUS-FPVFT VUS-FPVFT SVED-SPLVR PVDK-YP0H FPS-YPVFH CGRS-MPIIK GGRR-MPVIR SOGHPCPLYT VVRC-QPVVR IVRK-QPIVI IIRD-QPIIN	VSVALGADPA AAVVIGANPV VALAIGGEPP VALAIGGEPP VAVAIGVDPA AALVIGSDPG VTVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSVFLSGNPF VSVFLSGNPF ATVVGLHPA ATVVGLHPA ATVVGLHPA VALVIGNPVH IATAIGNDPA FALCFGVPPA ATVVGLHPA VALVIGNPVH IATAIGNDPA FTLNVGVGPG VALVIGNPVH IATAIGNDPA FTLNVGVGPG VALVIGNPVH IATAIGNPA CHAIYARNA CHAIY	TILGAVT ITLAAVT ITLAAAT TLIGAA TLIGAA LILSAIA LTLSAIA ITTAAST MIVAGTA IITSAIA ILTSAIA ILTSAIA ILTSAIA ILTSAIA ILTSAIA ILLSAT ILLAGAT TLLAAT ILLAGAT ILLAGAT ILLAGAT ILLASTT	PVPDTLSEYA PIPENVSEYN PIPENVSEYN PIPEVDLSEWL PLPEVDLSEWL PLPEVDSELL PLPENVSELL PLPENVSELL PLPENVSELL PLYDQSEYE PLLYDQSEYE SADITPDELK SVAYGVSELE SIPIDADEME AAPVETDELG P.PE PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-DAWLGK PLE-DAWLGK PLE-DAWLGK PLE-DAVLGW TQE-DFYLGN YQE-DFYLGN YQE-DFYLGN YQE-DFYLGN YQE-DFYLGN YQE-DFYLGN TDETHTLIGS SLEWNUMGF GIEHKTLMGM GYEHKLLQGL	FAGLLR FAGLLG FAGLIG FAGIIR FAGFIR FAGFIR CSFLQ LCTFLQ LTFLQ LTGFMR YVGAI LAAGML LAAGML IASAISLKAF VANAFH .AGAI IASAISLKAF VANAFH .AGAI LAAGML IASAISLKAF VANAFH .AGAI ALNEVFVPIL ALNRIYTPIL ALNRIYTPIL ALNRIYTPIL ALNRIYTPIL ATERIFLPIL FAERIFLPAI KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEYLSPLF KLQEFLYAV PCEPRILKGV PCEPRIYRAV PQEPRIYRAV	GTKTEVVKCI NKKIELVOCK GSGVALAKCK ERPVELVKGL KERLKVTECE GEGLPVYR-L NKKLSFVEK GARALY-KNT GAKLLY-KKT GKKARVMPCL GEPYRIVKSK GAPAPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GRPLEVIN-L EGELELVR-C NGELTLIK-C GAPLELVAGT 9 QKQFPEIVDF KFNLPEVVDI QTLLPEVVDI PLIMPGVQDL PLIMPGVQDL PLIMPGVQDL PLIMPGVQDL PLIMPGVQL PLIMPGVQL PLIMPGVQL PLIMPGVQL PLIMPGVQL PLIMPGVQL PLIMPGVQL NTVPTVKNI KRVPFVKNI	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LEVPAN TVD-LLVPAN PQSGHPLLCD DDTPHPLLYD NDHPHPLLYD SNS-LSVPSD LSD-LDVPWG TND-LMVPAG TND-LMVPAG TND-LMVPAG DG-IPTPV-D EGVDMEVP-P G-DLEVP-C UGPVEMVA-H 1.VP 400 YLPPEGCSYR WFPPEGCSYR WFPPEGCSYR WFPPEGCSYR NLPVEGAFHN NLPVEGAFHN NLPVEGAFHN SYGESGFHA KSYGESGFHA KSYGESGFHA KSYGESGFHA KSYGESGFHA KSYGESGFHA KSYGESGFHA KSYGESGFHA KSYGESGFHA KSYGESGCWL VLTEGGCCWL VLTEGGCCWL
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus 1. Consensus 1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 10. 17. 18. 19. 10. 17. 18. 19. 10. 17. 18. 19. 10. 17. 18. 19. 19. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 19. 10. 17. 18. 19. 19. 10. 11. 11. 12. 13. 14. 15. 16. 17. 17. 18. 19. 10. 17. 18. 19. 19. 10. 17. 18. 19. 17. 18. 19. 19. 19. 10. 17. 18. 19. 19. 19. 19. 19. 19. 19. 19. 19. 19	201 201 21YRQQLIGK GYYRQVSE GVYRLQLQSK GVYRLQVDK GVYRVNIAGD GWYRWQRFDR GMYRWQRFDR GMYRWQRFDR GMYRWQRFDR GYYRLQVIGT SIARGMVVDD SIHRLLIGR SHRILVKDD SFHRILIGR SHRILVKDD SFHRMWVIGD SHRMLVLDD SIHRMLVLDD SGHYLGYIGK GIYR, QV 301 AEIILEGYIE SEIVLEGYIE SEIVLEGYIE SEIVLEGYID SEFILTGEAI AEIILEGYI AEIILEGYI AEIILEGFI SEWFEGTLS SGAMVADVWI AEIILEGKI SEFVLEGKIE SEWFEGKLS SAMVALECER	NKLIMRWLSH NKLLMRWLKL TTMTVHWLSV DKLAVHWQIH KTGMHWQIH KEGVVHWQIH ETIGLHFQIQ DTIGLHFQIQ NTMGLHFQIQ NRLAFHPA-M KHITGLVIKP ERAAIRI-VP YLVIRMV-EQ ERAVIRL-VP KRLVIRI-VP DRLAVIRI-VP CRLVIRI-VP DRLAVIRI-VP CRLVIRI-VP CRCR-SCR CRCR CR	RGGALDYQEW RGGAEHHKRW RGGARHIRK- KDGNHHYWK- KHGAEHFRKM KRGMQAQODS KGGGAHFYE- KGGGMHFYE- KGGGMHFYE- KGGGMHFYE- KDSQLFHE- HDIAIHIRQ- HDVAQHIRK- QHIRQIADSW RHLWHLYRK- RHLHFLYNK- RHLHFLYNK- RHLHFLYNK- RHLHFLYNK- RHLHFLYNK- RHLYTYFQK- RHLGLYLQK- RHLGLYLQK- RHLGLYLQK- RHLGLYLQK- RHLGLYLQK- TGDHTGY-YN FGDHTGY-YN FGDHTGY-YN FGDHTGY-YD FGDHFGY-YS FGBHFGY-YS FGBHFGY-YS FGBHFGY-YS FGEHFGY-YS FGEHFGY-YS FGEHFGY-YS FGEHFGY-YS FGEFTGH-YS FGEFTGH-YS FGEFTGH-YS FGEFTGH-YS FGEFTGH-YS FGEFTGH-YS FGEFTGT-YD FVDITGT-YD FVDITGT-YD FVDITGT-YD FVDITGT-YD	CAAHPGERFE -KTKKKELFP -AAEQGKKLE -AKRLGKKLE -AKRLGKKLE -VEKGERRIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEQKKONIP -AEERGINIP -AEERGINIP -AEERGINIP -AEERGEDLE -ALKEKGYLD -SVEHGEEGLE -AQKLGKDLE -AAERGEDLE -AQKLGKDLE -AAERGEDLE -AQKLGKDLE -AAERGEDLE -AQKLGKDLE -AAERGEDLE -AQKLGKDLE -AEERGEDLE -AQKLGKDLE -AERGEDLE -AQKLGKDLE -AERGEDLE -AQKLGKDLE -AERGEDLE -AQKLGKDLE -AERGEDLE -AQKLGKDLE -AERGEDLE -AQKLGKDLE -AQKLGKDLE -AQK	VSVALGADEA AAVVIGANPV VÄLAIGCEPP VÄLAIGCEPP VÄLAIGCEPP VÄLAIGCEPP VSVAIGVDPA AÄLVIGSDPG VTVFLSGNPF VSVELSGNPF	TILGAVT ITLAAVT ITLAAAT TLYAATA TLYAATA TLYAATA LILSAIA LILSAIA LITSAIA ITTAAST MIVAGM AILVSSM VILGAAT VLLGAST VLLGAST ILLASTT VHFAAAPAE ISAA ILLAST VHFAAAPAE ILLAST ILLAST VHFAAAPAE ILLAST ILLAST VHFAAAPAE ILLAST	PVPDTLSEYA PIPENVSEYN PIPENVSEYN PIPENVSEYN PLPPEVDSYL PLPSGISEFM PVPYPICKLL PLPSGISEFM PLPSNVSELL PLPSNVSELL PLPSNVSEL PLPSNVSEL PLYDQSEYE PMAYDQSEYE SIPIDAUGSEYE SIPIDAUGSEYE SIPIDAUGSEYE SIPIDAUGSEYE SIPIDAUGSEYE PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-PAVLGV PDE-DAVLGV PDE-DAVLGV PDE-DAVLGV PDE-DAVLGV PDE-DAVLGV PDE-DAVLGV PDE-DAVLGV PDE-DAVLGV TQE-DFFLGN YQE-DFFLGN YQE-DFFLGN YQE-DFFLGN SLEHVNLMGG GFEHKLLQGL GFEHKLLQGL GFEHKLLQGL	FAGLLR FAGLLG FAGLG FAGIR FAGFIR FCSFLQ LCTFLQ LTFLQ MAGALR YVGAIL IAAGM FAAALL ALNEIFIPLL ALNEIFIPLL ALNEIFIPLL ALNEIFLP	GTKTEVVKCI NKKIELVQCK GSGVALAKCK KERLKVTECE GEGLPVYR-L SKLHY-KND GAKLY-KND GAKLY-KNT GKKARVMPCL GEVTIVKSK GAPAPIATAP GESVPVVKCE GGSMKVYR-S GGEIGVFE-L GRPLEVIN-L GRPLEVIN-L GRPLEVIN-C GAPLELVAGT G. QKQFPEIVDF HQQFPEIVDF KFNLPEVVDT RMMHPEIVDI QTLLPEVVDI PLIMPGVRQL PLVMPGVRQL PLVMPGVRQL PLVMPGVRQL PLVMPGVRL QMNAPNLIEY RAEFP-EVQA IESGLPILDA RRALPRVRAV KNTVPTVKNI KNVPTVENV RNAVPTVENV	300 SND-LEVPAS TID-LKVPAH TVD-LEVPAN TVD-LEVPAN TVD-LEVPAN POSGHPLLCD PDTPHPLLYD NDHPHPLLYD SNS-LSVPSD LSD-LDVPWG LTG-FDVPWG LTG-FDVPWG TND-LMVPAT PVHGNPVP-L DN-GLLVP-E GC-DLEVP-Q VGPVEMVA-H 1.VD 400 YLPPEGCSYR WFPPEGCSYR FLPMEALSYK HLPAEGFHN DLPVHGVFUG KSYGEAGFHA KSYGEAGFHA KSYGESGFHA YMPENGVFIN -VNAMYTHGL FMPYEAQALW RLTPASGCWL VLTEGGCCWL YLTEGGCCWL

FIG. 2. Homology between the *ubiD* gene product and 18 members of the UPF 00096 family of proteins, as arranged by the Multaline program. The identity of each protein is shown in Table 1. Lowercase letters designate consensus amino acids in 50 to 90% of the sequences, and boldface capital letters indicate 90 to 100% homology.

	401									500
1	LAVVTTKK	OVAGHAK	RVMMGVWSFL	ROFMYTKE-V	IVCDDDVNAR	DWNDVIWAIT	TRMDPARDTV	LVENTPIDYL	DFASP	VSGLGSKMGL
2	VVVVSIKK	SYPGHAK	RIMLGIWSYL	ROFMYNKF-I	IVVDDDIDVR	NWOEVIWAIA	TRSDPKRDTS	FIDNSPIDYL	DFASP	KSGLGSKMGI
2.	AATTSIDK	AYPGOAK	BAALAFWSAL	POFTYTKF-V	IVVDKSINIR	DPROVVWAIS	SKVDPVRDVF	ILPETPFDSL	DFASE	KIGLGGRMGI
4	FCEVSIKK		KVAYALLG-L	GLMSLEKH-I	VVFDDWINVO	DIGEVLWAWG	NNVDPORDVL	ILK-GPIDVL	DHATN	EVGFGGKMII
	LATVSTKK		KVMYATWG-T	GMT-SLTKT-V	VVVDDDVNVH	DMREVVWAVT	SRFDPARDVV	ILPPSPTDSL	DHSAY	IPNLAGKLGI
5.	MAEVSIRK		KALLALMG-L	GOLSLTKI-I	VVVDHDINVH	DVNOVIWAVS	SHVDPORDVL	VVPHSHTDEL	DPATP	TPMYGSKLGI
0. 7	LAAATUKE		RSALBILG-E	GOUST KE-L	WITDOSVDLE	NFPSLLECVL	ERMNFDRDLL	ILSETANDIL	DYTGS	GFNKGSKGIF
	LTCAUVER		ATSLRILG-E	GOUSLTKE-L	MITEHHVDLD	NEPKLLETIL	SRIVPERDLI	TESETSNDTL	DYTGP	KLNKGSKAIF
o. o	TTANVVKE		TTALRILG-E	GODSLTKF-L	MUTDOEVPLD	RESVULETIL	ERLOPDRDLI	TESETANDTL	DYTGP	SLNKGSKGIF
10	TTINETUT		OVMHAFWG-V	COMSEVENAT	FUNEDAPNIE	DTNATTEYLL	ENESKENALT	SOGVCDAL	DHASP	EYAMGGKLGI
10.	LIDAKINI		AUCMENT TTP	HGEGYCKM-W	TVUDEDVDPF	NLPOVMWALS	TEMHREHDAV	TTPDLSVLPL	DPGSN	PSGITHKMIL
11.	MUTTORNAL-		AVCMEAMTTE	HCLCYVAO-V	TLUDEDVDPF	NLPOVMWAMS	AKVNPKDDVV	VIPNLSVLEL	APAAO	PAGISSKMIT
12.	MVIISIAK	ONIETTOFE	CVENCELVER	TRUCEIVHEI	TLVADETDIE	NEKEVIWAYV	TRHTPVADOM	AFDDVTSFPL	APEVSOSSES	KTMKGGKCVT
13.	TITYADTKOT	QALKITESE	TATMATTAT	DOTEN	VVVDSDVDVD	DPMOVERATA	TRECADEDLY	TIPRARGSTL	DPSA	ADGLTAKMGL
14.	HAVIAVER	DEGDGR	NATINAFACI	DST VUV	TWODDDTNTE	DINDUFYATA	TRVOGDKDIV	TISCAKOSSE	DPSSDL	KNKLTAKVGV
15.	HAVVQIEK	OUEGDGR	NATLAAFASH		VUUDEDUNIV	DDPFVFWATA	TREOPDRDLV	TIPNARGSSL	DPSG	KDGLTAKWGT
16.	HAVVSIIK	QREGDGK	NALDAARAGA	DOT VUV	VWWDEDTDVI	DEFETEVATA	TRUCCORDET	TVPGARGSSL	DPAA-I.	PDGTTTKVGV
17.	HAAVSIKK	OTEGDGK	NUTWARDAR	POLKU	VVVVDTDVDVE	DPODIEVATA	TRVKGDRDLM	TVPNVRGSST	DPVA-E	SDGTTTKIGL
18.	HAAISINK	ULGDGK	ONTIMENT	PSERHR	TWINEDUDTE	NCODUENAM	TRUNCODICODIC	VIENAEGHGE	NPTF	PNYLGTKVGF
19.	HCVVKIAQ	KRAGWAK	QATLATEAAL	Prince	TWO dyp	d v wai	tr n d	in L	do a	ασΚα
Consensus	.a.vsi.k	.y.g.ak	d	•••		uv.war.	crpu	1.9	up .u	
	501									600
1	501			7 D 3 7 11						
			TEEDDDVVAH			LATENNGKSA				
	DATNKWPGET	-QREWGRP	IKKDPDVVAH	IDAIW	DE	LAIFNNGKSA				
2.	DATNKWPGET	-QREWGRP -NRKWGKK	IKKDPDVVAH IEMSQEVIDK	IDAIW IDSMW	DE DG	LAIFNNGKSA LNI YGLGDINLTE	VNPNLEGYDV			
2.	DATNKWPGET DATDKMYPET DATTKIPPET	-QREWGRP -NRKWGKK -DHEWGEV	IKKDPDVVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKKE	IDSMW VSQRW	DE DG AE	LAIFNNGKSA LNI YGLGDINLTE LGIE	VNPNLFGYDV			
2. 3. 4.	DATNKWPGET DATDKMYPET DATTKIPPET DATTKWKEEG	-QREWGRP -NRKWGKK -DHEWGEV YTREWPEV	IKKDPDVVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKKR VEMDAETKEK	IDAIW VSQRW IDEIW VDAIW	DG AE DR	LAIFNNGKSA LNI YGLGDINLTE LGIE TRNMVL	VNPNLFGYDV			
2. 3. 4. 5.	DATNKWPGET DATDKMYPET DATTKIPPET DATTKWKEEG DATKKWRDEG	-QREWGRP -NRKWGKK -DHEWGEV YTREWPEV YEREWPDV CCKOWPFF	IKKDPDVVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKKR VEMDAETKRK VAPDPETVEL	IDAIW IDSMW VSQRW IDEIW VDAIW VEGRW	DE DG DR DR 	LAIFNNGKSA LNI YGLGDINLTE LGIE IRNMVL YGLD	VNPNLFGYDV			
2. 3. 4. 5. 6.	DATNKWPGET DATDKMYPET DATTKIPPET DATTKWKEEG DATKKWRDEG DATKKUPEEY	-QREWGRP -NRKWGKK -DHEWGEV YTREWPEV YEREWPDV GGKQWPEE D2P2VCPS	IKKDPDVVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKKR VEMDAETKRK VAPDPETVRL LEGISOIGVE	IDAIW IDSMW VSQRW IDEIW VDAIW VEGRW	DE DG DR 	LAIFNNGKSA LNI YGLGDINLTE LGIE IRNMVL YGLD LOOLDIPALL	VNPNLFGYDV	LUILVED	LSSALSSTKE	FIWRTFTRSS
2. 3. 4. 5. 6. 7.	DATNKWPGET DATDKMYPET DATTKIPPET DATTKWKEEG DATKKWRDEG DATKKLPEEY LGVGAPIRSL	-QREWGRP -NRKWGKK -DHEWGEV YTREWPEV YEREWPDV GGKQWPEE PRRYRGPS	IKKDPDVVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKKR VEMDAETKRK VAPDPETVRL LPGISQIGVF	IDAIW IDSMW VSQRW IDEIW VDAIW CRGCL CRGCL	DE DG DR NE VLETS VLETS	LAIFNNGKSA LNI YGLGDINLTE LGIE IRNMVL YGLD LQQLDIPALL LQQVNIDALL	VNPNLFGYDV KEPHLADWP- NHPDLSSWP-	LVILVED	LSSALSSTKE	FIWRTFTRSS
2. 3. 4. 5. 6. 7. 8.	DATNKWPGET DATDKMYPET DATTKIPPET DATTKWKEEG DATKKWRDEG DATRKLPEEY LGVGAPIRSL MGIGPAIRDL	-QREWGRP -NRKWGKK -DHEWGEV YTREWPEV YEREWPDV GGKQWPEE PRRYRGPS PHKYRGKS	IKKDPDVVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKKR VEMDAETKRK VAPDPETVRL LPGISQIGVF LPNITNMGTF IHCVODIABE	IDAIW IDSMW VSQRW IDEIW VDAIW CRGCL CPGCL CPGCL	DE DG DR NE 	LAIFNNGKSA LNI YGLGDINLTE LGIE IRNMVL YGLD LQQLDIPALL LQQVNIDALL LEDBCIKSLI	VNPNLFGYDV KEPHLADWP- NHPDLSSWP- HHPDLKSWP-	LVILVED LVVLTEN	LSSALSSTKE LNETLASSKD LBETLOSEKD	FIWRTFTRSS FLWKTFTRLA FLWRTFTRCA
2. 3. 4. 5. 6. 7. 8. 9.	DATNKWPGET DATTKIPPET DATTKIPET DATTKWKEEG DATKKWRDEG DATKKWRDEG DATKKLPEY LGVGAPIRSL MGIGPAIRDL MGIGKAIRDL	-QREWGRP -NRKWGKK -DHEWGEV YTREWPEV YEREWPDV GGKQWPEE PRRYRGPS PHKYRGKS PHGYQGGK	IKKDPDVVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKKR VEMDAETKRK VAPDPETVRL LPGISQIGVF LPNITNMGTF IHGVQDIAPF	IDAIW IDSIW VSQRW VDAIW VEGRW CRGCL CRGCL CRGCL	DG DR DR DR 	LAIFNNGKSA LNI YGLGDINLTE LGIE IRNMVL YGLD LQULDIPALL LQQVNIDALL LEDRCIKSLL VEKKNKSVIE	VNPNLFGYDV KEPHLADWP- NHPDLSSWP- HHPDLKSWP- LANNLGPFF	LVILVED LVVLTEN LIILADN H.BIVIFVEH	LSSALSSTKE LNETLASSKD LRETLQSEKD ASNDLNNPYM	FIWRTFTRSS FLWKTFTRLA FLWRTFTRCA LLWRTVNNID
2. 3. 4. 5. 6. 7. 8. 9. 10.	DATNKWPGET DATCKMYPET DATTKWREG DATKWREG DATKKWRDEG DATKKLPEY LGVGAPIRSL MGIGPAIRDL MGIGKAIRDL DATSKSNTPY	-QREWGRW -NRKWGKK DHEWGEV YTREWPEV YEREWPDV GGKQWPEE PRRYRGPS PHKYRGKS PHGYQGGK PTLINDSA	IKKDPDVVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKKR VAPDPETVRL LPGISQIGVF IHGVQDIAPF ILALLQDKMQ	IDAIW IDSMW VSQRW IDEIW VDAIW VEGRW CRGCL CRGCL CRGCL NIVLLKQYIP	DG DR DR NL QE VLETS VLETS HTRNPICVIS	LAIFNNGKSA LNI YGLGDINLTE LGIE IRMMVL YGLD LQQLDIPALL LQQVNIDALL LEDRCIKSLL VEKKDKSVIE	VNPNLFGYDV KEPHLADWP- NHPDLSSWP- HHPDLKSWP- LAKNLLGFEE	LVILVED LVVLTEN LIILADN HLRIVIFVEH	LSSALSSTKE LNETLASSKD LRETIQSEKD ASNDLNNPYM	FIWRTFTRSS FLWKTFTRLA FLWRTFTRCA LLWRIVNNID
2. 3. 4. 5. 6. 7. 8. 9. 10. 11.	DATNKWPGET DATCKNPPET DATTKNPET DATTKWKEEG DATKKWRDEG DATKKLPEEY LGVGAPIRSL MGIGPAIRDL MGIGKAIRDL DATSKSNTPY DATTPVAPET	-QREWGRW -NRKWGKK -DHEWGEV YTREWPEV YEREWPDV GGKQWPEE PRRYRGPS PHKYRGKS PHGYQGGK PTLLND-SA RGHYSQPLDS	IKKDPDVVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKKR VEMDAETKRK VAPDPETVRL LPGISQIGVF LENITNMGTF IHGVQDIAPF LLALLQDKMQ PLTTKEWEQA	IDAIW IDSMW VSQRW IDEIW VDAIW CRGCL CRGCL CRGCL NIVLLKQYYP LMDLMNK LODLINDRY	DG DR DR DR 	LAIFNNGKSA LNI YGLGDINLTE LGIE IRNMVL YGLD LQQLDIPALL LQQVNIDALL LQQVNIDALL LEDRCIKSLL VEKKDKSVIE	VNPNLFGYDV KEPHLADWP- NHPDLSSWP- HHPDLKSWP- LAKNLLGFEE	LVILVED LVVLTEN LIILADN HLRIVIFVEH	LSSALSSTKE LMETLASSKD LRETIQSEKD ASNDLNNPYM	FIWRTFTRSS FIWKTFTRLA FIWRTFTRCA LLWRIVNNID
2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12.	DATDKWPGET DATDKWYPET DATTKIPPET DATTKWKEEG DATKKWRDEG DATKKURDEG DATKKLPEEY LGVGAPIRSL MGIGPAIRDL MGIGKAIRDL DATSKSNTPY DATTPVAPET DATTPVAPET	-QREWGRW -NRKWGKK DHEWG-EV YTREWPEV GGKQWPDV GGKQWPEE PRRYRGPS PHKYRGKS PHGYQGGK PTLLNDSA RGHYSQPLDS RGNFSTPAKD	IKKDPDVVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKR VEMDAETKRK VAPDPETVRL LPGISQIGVF LPNITNMGTF IHGVQDIAPF ILALLQDKMQ PLTKEWEQK LPETAEWAAR	IDAIW IDEW VDAIW VECRW CRGCL CRGCL CRGCL RIVLLKQYYP LMDLMNK LQRLIAARV	DG AE NE 	LAIFNNGKSA LNI YGLGDINLTE LGIE IRNMVL YGLD LQQLDIPALL LQQVNIDALL LEDRCIKSLL VEKKDKSVIE	VNPNLFGYDV KEPHLADWP- NHPDLSSWP- HHPDLKSWP- LAKNLLGFEE	LVILVED LVVLTEN HLRIVIFVEH	LSSALSSTKE LNETLASSKD LRETIQSEKD ASNDLNNPYM	FIWRTFTRSS FLWKTFTRLA FLWRTFTRCA LLWRIVNNID
2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13.	DATNKWPGET DATCKWYPET DATTKIVPET DATTKWKEG DATKKWREG DATKKUPEY LGVGAPIRSL MGIGPAIRDL MGIGKAIRDL DATSKSNTPY DATTPVAPET DATTPVAPEV NCIFRQQYER	-QEEWGRP -NRKWGKV DHEWG-EV YTREWPEV YEREWPDV GGKQWPEV PRRYRGPS PHKYRGKS PHGYGG-GK PTLLND-SA RGHYSQPLDS RGHYSQPLDS SFDITCNFE	IKKOPDUVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKKR VEMDAETKRK UPGISQIGVF LPNITNMGTF IHGVQDIAPF LLALLQDKMQ PLTTKEMEQK LPETAEWAAR KGYPKGLUDK	IDANW IDSMW VSQRW VDAIW VEGRW CRGCL CPGCL RGCL NIVLLKQYIP IMDLMNK LQRLIAARW VNENWKRYGY	DG DR DR QE 	LAIFNNGKSA LNI YGLGDINLTE LGIE IRNMVL YGLD LQQLDIPALL LQQUNIDALL LEDRCIKSLL VEKKDKSVIE	VNPNLFGYDV KEPHLADWP- NHPDLSSWP- HHPDLKSWP- LAKNLLGFEE	LVILVED LVVLTEN JIILADN HLRIVIFVEH	LSSALSSTKE LNETLASSKD LRETIQSEKD ASNDLNNPYM	FIWRTFTRSS FLWKTFTRLA FLWRTFTRCA LLWRIVNNID
2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14.	DATNKWPGET DATDKWYPET DATTKIWPET DATTKWREG DATKKWREG DATKKUPEEY LGVGAPIRSL MGIGKAIRDL DATSKSNTPY DATTPVAPET DATTPVAPET DATTPVAPEN NCIFRQQYER DATKPLDAGM	-QEEWGRF -NRKWGKK -DHEWGEV YTREWPDV YEREWPDV GGKOWPEE PRRYRGFS PHGYQGGK PTLLND-SA RGHYSQPLDS RGNFSTPAKD SFDVITCMFE GYERGRIEGF	IKKOPDVVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKRK VEMDAETKRK UPMIDAETKRK LPGVDIAPF LLALLQDKMQ PLTTKEMECK LPETAEWAAR KGYPKGLVDK KWGSRRCHQP	IDANW USQRW VDAIW CPGCL CPGCL NIVLLKQYYP LMDLMNK LQRLIAARV VNENWKRYGY GD	DG DG DE DE QE VLETS VLETS HTRNPICVIS K	LAIFNNGKSA LNI YGLGDINLTE LGIE IRNMVL YGLD LQQLDIPALL LQQUNIDALL LEDRCIKSLL VEKKDKSVIE	VNPNLFGYDV KEPHLADWP- NHPDLSSWP- HHPDLKSWP- LAKNLLGFEE	LVILVED LVVLTEN LILADN HLRIVIFVEH	LSSALSSTKE LNETLASSKD LRETIQSEKD ASNDLNNPYM	FIWRTFTRSS FLWKTFTRLA FLWRTFTRCA LLWRIVNNID
2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14.	DATNKWPGET DATDKWYPET DATTKINPET DATKKWRDEG DATKKURPEY LGVGAPIRSL MGIGGAIRDL DATSKIRPEY DATTPVAPET DATTPVAPET DATTPVAPET DATTPVAPET DATTPLAPEN NCIFRQOYER DATMSLIKGR	-QEEWGRF -NRKWGEV -NRKWGEV YTREWPDV YEREWPDV GGKQWPEE PRRYRGPS PHGYQGGK PHGYQGGK PHGYQGGK RGHYSQPLDS RGHYSTPAKD SFDYITCNPE GYERGRIFGF EHFERAKIPD	IKKOPDVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKKR VEMDAETKRK UPGISQIGVF LPHITNMGTF IHGVQDIAPF IHGVQDIAPF ILALLQDKMQ PLTTKEWEQK LPETAEWAAR KGYPKGLVDK KWGSRRCHQP K	IDANW VSQRW VDAIW CRGCL CRGCL CRGCL CRGCL NIVLLKQYYP LMDLMNK LQRLIAARV VNENWKRYGY GD	DG DG DR 	LAIFNNGKSA LNI VGLGDINLTE LGIE IRNMVL YGLD LQQLDIPALL LQQVDIDALL LEDRCIKSLL VEKKDKSVIE	VNPNLFGYDV KEPHLADWP- NHPDLSSWP- HHPDLKSWP- LAKNLLGFEE	LVILVED LVVLTEN LIILADN HLRIVIFVEH	LSSALSSTKE LNETLASSKD LRETIQSEKD ASNDLNNPYM	FIWRTFTRSS FLWKTFTRLA FLWRTFTRCA LLWRIVNNID
2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15.	DATNKWPGET DATDKWYPET DATTKIPPET DATTKWKEEG DATKKWRDEG DATKKLPEEY LGVGAPIRSL MGIGPAIRDL DATSKSNTPY DATTPVAPET DATTPVAPET DATTPVAPET DATTPLDAGM DATKELDAGM DATKELDAGM	-QEEWGRP -NRKWGKV DHEWG-EV YTREWPEV YTREWPEV GGKOWPEE PRRYRGPS PHKYRGKS PHGYQGGK RGHYSQPLDS RGHYSQPLDS RGHYSTPAKD SFDYITCNFE EHFERAKIPD E-FERASLDF	IKKOPDUVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKKR VEMDAETKRK UPGISQIGVF LPRITNMGTF IHGVQDIAPF IHGVQDIAPF IHGVQDIAPK ULAILQDKMQ PLTTKEWEQK LPETAEWAAR KGYPKGLVDK KWGSRRCHQP K	IDANW USQRW IDEIW VDAIW CRGCL CRGCL CRGCL CRGCL CRGCL CRGCL GD	DG AE DG NE VLETS VLETS HTRNPICVIS K	LAIFNNGKSA LNI YGLGDINLTE LGIE IRNMVL YGLD LQQLDIPALL LQQVNIDALL LEDRCIKSLL VEKKDKSVIE	VNPNLFGYDV KEPHLADWP- NHPDLSSWP- HHPDLKSWP- LAKNLLGFEE	LVILVED LVVLTEN LIILADN HLRIVIFVEH	LSSALSSTKE LNETLASSKD LRETIQSEKD ASNDLNNPYM	FIWRTFTRSS FLWKTFTRLA FLWRTFTRCA LLWRIVNNID
2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17.	DATNKWPGET DATDKWYPET DATTKIWYPET DATTKWREG DATKKWREG DATKKUPEY LGVGAPIRSL MGIGFAIRDL MGIGKAIRDL DATSKSNTPY DATTPVAPET DATTPVAPET NCIFRQOYER DATKPLDAGM DATKPLDAGM DATKPLDAGM	-QEEWGRF -NRKWGKK -DHEWGEV YTREWPDV GGKQWPEE PRKYRGFS PHKYRGFS PHGYQGGK PTLLNDSA RGHYSQPLDS SFDYITCNFE GYERGRIFGF EHFERAKIPD EHFERAKIPG E-FEKASLDF	IKKOPDVVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKRK VAPDPETVRL LPGISQIGVF LPNITNMGTF LLALLQDKMQ PLTTKEMEQK KGYPKGLVDK KWGSRRCHQP K	IDANW VSQRW VEQRW VEGRW CPGCL CPGCL NIVLLKQYYP LMDLMNK UQRLIAARV VNENWKRYGY GD	DG DG DE GE VLETS VLETS HTRNPICVIS K	LAIFNNGKSA LNI YGLGDINLTE LGIE IRMWUL YGLD LQQLDIPALL LQQUNIDALL LEDRCIKSLL VEKKDKSVIE	VNPNLFGYDV KEPHLADWP- NHPDLSSWP- HHPDLKSWP- LAKNLLGFEE	LVILVED LVVLTEN DILLADN HLRIVIFVEH	LSSALSSTKE LNETLASSKD LRETIQSEKD ASNDLNNPYM	FIWRTFTRSS FLWKTFTRLA FLWRTFTRCA LLWRIVNNID
2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 16. 17. 18. 16. 17. 18. 19. 10. 11. 12. 13. 14. 15. 16. 16. 17. 18. 18. 19. 10. 11. 12. 13. 14. 15. 16. 16. 17. 18. 18. 19. 10. 11. 12. 13. 14. 15. 16. 17. 18. 18. 19. 19. 19. 19. 19. 19. 19. 19	DATNKWPGET DATDKWYPET DATTKIWYPET DATTKWREG DATKKWREG DATKKUPEEY LGVGAPIRSL MGIGKAIRDL DATSKSNTPY DATTPVAPET DATTPVAPET DATTPVAPET DATKPLDAGM DATMSLIKGR DATKPLDAKK DATAPL-ASA DATKSL-KTL	-QEEWGRF -NRKWGKK -NRKWGEV YTREWPDV YEREWPDV YEREWPDV GGKQWPEE PRYRGFS PHGYQG-GK PHGYQG-GK PHGYQG-GK FTLIND-SA RGHYSQPLDS RGNFSTPAKD SFDYITCMFE GYERGRIPGF EHFERAKIPD EHFERAKIPD EKFQRVSRSE DKFERVSFCE	IKKOPDVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKKR VEMDAETKRK VAPDPETVRL LPGISQIGVF LPNITNMGTF HGVQDIAPF LLALLQDKMQ PLTTKEWEQK LPETAEWAAR KGYPKGLVDK K	IDANW USQRW VDAIW VGERW CRGCL CRGCL NIVLLKQYYP LMDLMNK LQRLIAARV VNENWKRYGY GD	DG DG DR 	LAIFNNGKSA LNI YGLGDINLTE LGIE IRNMVL YGLD LQOLDIPALL LQOUDIDALL LEDRCIKSLL VEKKDKSVIE	VNPNLFGYDV KEPHLADWP- NHPDLSSWP- HHPDLSSWP- LAKNLLGFEE	LVILVED LVVLTEN LIILADN HLRIVIFVEH	LSSALSSTKE LNETLASSKD LRETIQSEKD ASNDLNNPYM	FIWRTFTRSS FLWKTFTRLA FLWRTFTRCA LLWRIVNNID
2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19.	DATNKWPGET DATDKWYPET DATTKIPPET DATTKWKEG DATKKWRDEG DATKKUPEY LGVGAPIRSL MGIGGAIRDL DATSKIRDL DATSKINTPY DATTPVAPET DATTPVAPET DATTPLDAGM DATMSLIKGR DATMSLIKGR DATMSLIKGR DATMSLIKGR DATMSLIKGR DATMSLIKGR	-QEEWGRP -NRKWGKV -NRKWGEV YTREWPEV YTREWPEV PRYRGPS PHKYRGKS PHGYGQGK PHLINDSA RGHYSQPLDS RGNFSTPAKD SFDYITCNFE EHFERAKIPD E-FERASLDF EKFQRVSFGE AFDRAKTKAM	IKKOPDVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKKR VEMDAETKRK VEMDAETKRK LPGISQIGVF LPNITNMGFF IHGVQDIAPF IHGVQDIAPF IHGVQDIAPK UALLQDKMQ PLTTKEWEQK LPETAEWAAR KGYPKGLVDK K TLDGLDIVGA	IDANW VSQRW VDAIW CRGCL CRGCL CRGCL CRGCL RIVLLKQYYP LMDLMNK LQRLIAARV VNENWKRYGY GD	K	LAIFNNGKSA LNI YGLGDINLTE LGIE IRNMVL YGLD LQQLDIPALL LQQVNIDALL LEDRCIKSLL VEKKDKSVIE	VNPNLFGYDV KEPHLADWP- NHPDLSSWP- HHPDLKSWP- LAKNLLGFEE	LVILVED LVVLTEN LIILADN HLRIVIFVEH	LSSALSSTKE LNETLASSKD LRETIQSEKD ASNDLNNPYM	FIWRTFTRSS FLWRTFTRLA FLWRTFTRCA LLWRIVNNID
2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. Consensus	DATNKWPGET DATDKWYPET DATTKIWPET DATTKWREG DATKKWRDEG DATKKUPEY LGVGAPIRSL MGIGFAIRDL MGIGKAIRDL DATSKSNTPY DATTPVAPET DATTPVAPET DATTPVAPUN NCIFRQQYER DATKPLDAGM DATKSLKGR DAKPLDKKK DATKSL-KTL DCTRFFPHTP dat	-QREWGRE -NREWGEV YTREWP-EV YTREWP-DV GGKQWP-EE PRKYRG-PS PHKYRG-FS PHKYRG-S GYTLND-SA RGHYSQPLDS SFDYITCNFE GYERGRIFGF EHFERAKIPD EHFERAKIPD EHFERAKIPD EHFERAKIPA	IKKOPDVVAH IEMSQEVIDK LESDPAMAEQ IEMSPEVKRK VAPDPETVRL LPGIQIGVF LPNITNMGTF LLALLQDKMQ PLTTKEWEQK KWGSRRCHQP K TLDGLDIVGA	IDANW VSQRW VEQRW VEGRW CPGCL CPGCL NIVLLKQYYP LMDLMNK LQRLIAARV VNENWKRYGY GD KR	DG DG DR GE VLETS VLETS HTRNPICVIS K	LAIFNNGKSA LNI YGLGDINLTE LGTE IRMW/L YGLD LQQLDIPALL LQQUNIDALL LEDRCIKSLL VEKKDKSVIE	VNPNLFGYDV KEPHLADWP- NHPDLSSWP- HHPDLKSWP- LAKNLLGFEE	LVILVED LVVLTEN DIILADN HLRIVIFVEH	LSSALSSTKE LNETLASSKD LRETIQSEKD ASNDLNNPYM	FIWRTFTRSS FLWKTFTRLA FLWRTFTRCA LLWRIVNNID

ubiB gene is now shown to be the former open reading frame *yigR*, approximately 6 kb downstream from its previous location (12).

The *yigC* segment and its upstream region were isolated by PCR from the AN256 chromosome (primers used were as follows: in the forward direction, 5'-GATCATCGGTGCCAG GCAATTCACAGCC-3', in the reverse direction, 5'-TCAGG CGCTTTTACCGTTGTTAAAA-3'). It was cloned into a pNoTA/T7 shuttle vector (manufactured by 5 Prime \rightarrow 3 Prime Inc.), and this construct was transformed into AN66 cells. This plasmid, designated p613, complemented the *ubiD* mutant trait of AN66 cells to the same extent as the larger plasmid, p522.

Based on its nucleotide sequence, the product of *ubiD* gene is a 497-amino-acid protein, its molecular mass is 55,603.7 Da, and its theoretical isoelectric point is 5.31. The ubiD gene product is one of two enzymes (3-octaprenyl-4-hydroxybenzoate carboxy-lyase) which catalyze the decarboxylation of 3octaprenyl-4-hydroxy benzoate to 2-octaprenylphenol. Earlier work with this enzyme suggested that it is a membrane-associated protein, although during cell fractionation much activity was found in the cytoplasmic fraction (9). Analysis of its amino acid sequence for transmembrane helices indicated zero (13), one (positions 215 to 235) (K. Hofmann and W. Stoffel, Biol. Chem. Hoppe-Seyler 347:166, abstr. MF C-35), or two (positions 226 to 232 and 334 to 340) (4) such regions, depending on which program was used. This enzyme's molecular mass by gel filtration measurement was reported to be approximately 340,000 Da (9). This suggests that it is a hexameric protein in vivo.

We isolated the *ubiD* gene from strain AN66 by PCR and

sequenced it, for the purpose of locating the site of mutation. The long gene was sequenced in overlapping segments, and the last fragment was sequenced in the reverse direction as well. (The following primers were used: 1, 5'-ATGGACGCCATG AAATATAACGATT-3'; 2, 5'-GCGTGGCGATGGGCATGGGCATG GGGCAGG-3'; 3, 5'-GCATTCCCATTATGACCTGCTGG CCGG-3'; 4, 5'-GGTGCCGATCCCGCCACGATTCTCGG-

TABLE 1. Proteins of UPF0096 family

Organism	Name of protein	Primary accession no.
Escherichia coli	YigC (UbiD)	
Rickettsia prowazekii	Y821_RICPR	Q9ZCD6
Synechocystis sp. (strain PCC6803)	Y936_SYNY3	P72861
Aquifex aeolicus	Y612_AQUAE	067542
Archeoglobus fulgidus	Y209_ARCFU	030030
Aeropyrum pernix	YF71_ARCFU	Q9YBM7
Chlamydia trachomatis	Y085_CHLTR	084087
Chlamydia psitacci	Y66K_CHLPS	034023
Chlamydia pneumoniae	Y328_CHLPN	Q9Z8L0
Helicobacter pylori	Y396_HELPY	025157
Bacillus subtilis	YCLC_BACSU	P94405
Streptomyces sp. (strain D7)	VDCC_STRD7	Q9X697
Saccharomyces cerevisiae	YD39_YEAST	Q03034
Aeropyrum pernix	YK78_AERPE	Q9YA60
Methanococcus janaschii	YB33_METJA	Q58533
Pyrococcus horiochii	Y963_PYRHO	058701
Methanobacterium thermo- autotrophicum	YD94_METTH	P41655
Methanobrevibacter smithii	YPUE_METSM	P22349
Rhodospirillum rubrum	YCOM_RHORU	P72315

3'; and 5, 5-GGGCGTCCGCCAGATGAGCCCGCGGCGG TG-3' [all forward direction] and 5'-TCAGGCGCTTTTACC GTTGTTAAAA-3' [reverse direction].) Comparison of the results with the published nucleotide sequence showed a single deviation. Codon 452, GGG (glycine), appeared in the mutant as AGG (arginine). A homology study (see below) shows that this glycine is a consensus residue.

A comparison (14) between the amino acid sequences of the ubiX and yigC (ubiD) gene products showed no significant similarities. A Blast search of the entire protein database of 525,243 sequences yielded 86 hits of similarities with yigC (ubiD). One of these hits was a 29% sequence identity and 48% similarity with a 4-hydroxybenzoate decarboxylase from *Clostridium hydroxybenzoicum*.

In the Swiss-Prot database yigC is listed as a member of the family UPF (uncharacterized protein family) 00096, with a taxonomic range of archaea, eubacteria, and eukaryota. Table 1 lists these proteins, and Fig. 2 shows regions of homology among them, as assigned by the Multaline program (2). Clearly not all of these proteins can be functional homologues of ubiD. Two proteins are from the same organism (Aeropyrum pernix). The gram-positive organisms Bacillus subtilis and Streptomyces sp. synthesize menaquinones, and the cyanobacterium Synechocystis sp. makes plastoquinone instead of ubiquinone (1). Helicobacter pylori also utilizes menaquinone instead of ubiquinone (10). Archaebacteria also employ a variety of ubiquinone analogues (15). However, due to the extensive homology between these proteins, it is reasonable to expect that they all function in some membrane-associated decarboxylation process

We thank R. Meganathan for communicating unpublished results to us. We are grateful to Catherine F. Clarke for strain AN256 and also for sending us the manuscript of reference 12 before it was published.

REFERENCES

- Collins, M. D., and D. Jones. 1981. Distribution of isoprenoid quinone structural types in bacteria and their taxonomic implications. J. Bacteriol. 45:316–354.
- Corpet, F. 1988. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. 16:10881–10890.
- 3. Cox, G. B., I. G. Young, L. M. McCann, and F. Gibson. 1969. Biosynthesis of

ubiquinone in *Escherichia coli* K-12, location of genes affecting the metabolism of 3-octaprenyl-4-hydroxybenzoic acid and 2-octaprenyl phenol. J. Bacteriol. **99:**450–458.

- Cserzo, M., E. Wallin, I. Simon, G. von Heijne, and A. Elofsson. 1997. Prediction of transmembrane alpha helices in procaryotic membrane proteins: the Dense Alignment Surface method. Protein Eng. 10:673–676.
- Daniels, D. L., G. Plunkett III, V. Burland, and F. R. Blattner. 1992. Analysis of the *Escherichia coli* genome: DNA sequence of the region from 84.5 to 86.5 minutes. Science 257:771–778.
- Davis, B. D., and E. S. Mingioli. 1950. Mutants of *Escherichia coli* requiring methionine or vitamin B₁₂. J. Bacteriol. 60:17–28.
- Kwon, O., A. Kotsakis, and R. Meganathan. 2000. Ubiquinone (coenzyme Q) biosynthesis in *Escherichia coli*: identification of the *ubi*F gene. FEMS Microbiol. Lett. 186:157–161.
- Lee, P. T., A. Y. Hsu, H. T. Ha, and C. F. Clarke. 1997. A C-methyltransferase involved in both ubiquinone and menaquinone biosynthesis: isolation and identification of the *Escherichia coli ubiE* gene. J. Bacteriol. 179:1748– 1754.
- Leppik, R. A., I. G. Young, and F. Gibson. 1976. Membrane-associated reactions in ubiquinone biosynthesis in *Escherichia coli*. Biochim. Biophys. Acta 436:800–810.
- Marcelli, S. W., H. T. Chang, T. Chapman, P. A. Chalk, R. J. Miles, and R. K. Poole. 1996. The respiratory chain of *Helicobacter pylori*: identification of cytochromes and the effects of oxygen on cytochrome and menaquinone levels. FEMS Microbiol. Lett. 138:59–64.
- Meganathan, R. 1996. Biosynthesis of the isoprenoid quinone menaquinone (vitamin K₂) and ubiquinone (coenzyme Q), p. 642–656. *In* F. C. Neidhardt et al. (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press Washington, D.C.
- Poon, W. W., D. E. Davis, H. T. Ha, T. Jonassen, P. N. Rather, and C. F. Clarke. 2000. Identification of *Escherichia coli ubiB*, a gene required for the first monooxygenase step in ubiquinone biosynthesis. J. Bacteriol. 182:5139– 5146.
- Sonnhammer, E. L. L., G. von Heine, and A. Krogh. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences, p. 175– 182. *In J.* Glasgow, T. Littlejohn, F. Major, R. Lathrop, D. Sankoff, and C. Sensen (ed.), Proceedings of the Sixth International Conference on Intelligent systems for molecular biology. AAAI Press, Menlo Park, Calif.
- Tatusova, T. A., and T. L. Madden. 1999. Blast 2 sequences—a new tool for comparing protein and nucleotide sequences. FEMS Microbiol. Lett. 174: 247–250.
- Thurl, S., W. Whitke, I. Buhrow, and W. Schafer. 1986. Quinones from archaebacteria. II. Different types of quinones from sulphur-dependent archaebacteria. Biol. Chem. Hoppe-Seyler 367:191–197.
- Young, I. G., P. Strobant, P. C. G. Macdonald, and F. Gibson. 1973. Pathway for ubiquinone biosynthesis in *Escherichia coli* K-12: gene-enzyme relationships and intermediates. J. Bacteriol. 114:42–52.
- Zeng, H., I. Snavely, P. Zamorano, and G. T. Javor. 1998. Low ubiquinone content in *Escherichia coli* causes thiol hypersensitivity. J. Bacteriol. 180: 3681–3685.