Molecular Recognition by Ornithine and Aspartate Transcarbamylases

NORMA M. ALLEWELL,*,† DASHUANG SHI,† HIROKI MORIZONO.‡ AND MENDEL TUCHMAN[‡]

Department of Biochemistry, University of Minnesota, St. Paul, Minnesota 55108, and Department of Pediatrics, Medical School, University of Minnesota, Minneapolis, Minnesota 55455

Received October 27, 1998

The ability of molecules to recognize and interact selectively with other molecules underlies a wide variety of important research areas in both chemistry and biochemistry. Examples include chemical and biochemical catalysis, rational drug design, the development of sensors for environmental pollutants and clinical tests, taste and smell, and insect and weed control. Among biological molecules, proteins have evolved to have the greatest versatility and selectivity in molecular recognition (although RNA may prove a close second), and much of what we understand about the principles that underlie biological recognition has come from studying proteins. The ability of proteins to recognize and discriminate among molecules, both large and small, depends on their complex chemistry and surface properties and their ability to undergo changes in their three-dimensional structure as a result of changes in their environment or molecular interactions that alter both their chemistry and their surface properties. Since the underlying principles are

Norma Allewell received her B.Sc. (Hon.) degree in biochemistry from McMaster University (Hamilton, ON, Canada) in 1965 and her Ph.D. in molecular biophysics from Yale University in 1969. Until January 1999, she was Professor of Biochemistry and Vice Provost for Research and Graduate/Professional Education at the University of Minnesota. She is now Associate Vice President for Sponsored Programs and Technology Transfer at Harvard University.

Dashuang Shi was born in Fujian Province, People's Republic of China in 1964. He received his B.S. (1983) and M.S. (1987) degrees in chemistry from Xiamen University and his Ph.D. in 1997 from the University of Sydney, where his advisor was Professor Hans Freeman. He is currently a postdoctoral associate with Professor Allewell

Hiroki Morizono received his S.B. degree in life sciences from MIT in 1987, and his Ph.D. degree in biochemistry from the University of Minnesota in 1997 under the direction of Professor Allewell. He is currently a postdoctoral associate in the laboratory of Dr. Tuchman.

Mendel Tuchman received his M.D. (magna cum laude) from the Sackler School of Medicine, Tel Aviv University, in 1979, and joined the faculty in the Department of Pediatrics at the University of Minnesota in 1985. He is currently Professor of Pediatrics and Laboratory Medicine and Pathology and Director of the Biochemical Genetics and Metabolism Laboratories in the Medical School of the University of Minnesota. In August 1999, he became Director of the Clinical Research Center and Chairman of Metabolism at the Childrens' National Medical Center in Washington, DC.

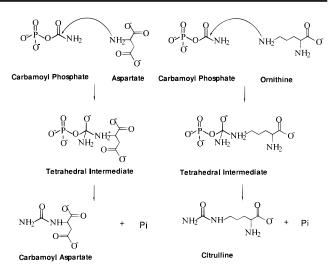


FIGURE 1. Schematic drawings of reactions catalyzed by ATCase and OTCase.

similar for most proteins, these principles can be illustrated with almost any protein or protein family. Here we focus on the transcarbamylases, a family of enzymes that catalyze transfer of the carbamoyl group of carbamoyl phosphate (CP) to an amino group.

The most widely studied members of this family are the aspartate transcarbamylases (ATCases) and ornithine transcarbamylases (OTCases), enzymes that transfer the carbamoyl group of CP to the α -amino group of Laspartate (L-Asp) and the δ -amino group of L-ornithine (L-Orn), respectively (Figure 1). ATCases catalyze the first committed step in the biosynthesis of pyrimidines, one of the components of nucleic acids, while OTCases function in the urea cycle, which eliminates excess ammonia, and in the biosynthesis of the amino acid arginine. This laboratory was largely responsible for defining the thermodynamics of protein-protein and protein-ligand interactions in Escherichia coli ATCase by reaction and differential scanning microcalorimetry, 1-4 electrostatic modeling,^{5,6} hydrogen exchange,^{7,8} solvent perturbation,^{9,10} and related methods¹¹⁻¹⁴ and has recently determined high-resolution crystal structures of both E. coli and human ornithine transcarbamylase, 15,16 as well as examining the enzymology of the human enzyme. 17,18 This work has provided important new insights into the molecular basis of ornithine transcarbamylase deficiency (OTCD), a relatively common clinical condition in which ammonia levels in the blood are elevated as a result of mutations in the OTCase gene.19

E. coli ATCase, the most thoroughly studied transcarbamylase, is a multisubunit protein containing two trimeric catalytic polypeptide chains and three dimeric regulatory chains which undergoes major changes in its threedimensional structure when substrates bind. It has been

^{*} To whom correspondence should be addressed. Current address: Office of Sponsored Research, Harvard University, 4th Floor, Holyoke Center, 1350 Massachusetts Ave., Cambridge MA 02138. E-mail: norma allewell@harvard.edu.

Department of Biochemistry.

[‡] Department of Pediatrics, Medical School.

widely used for more than 30 years as a model system for understanding how protein-protein interactions mediate signal transduction.²⁰⁻²⁵ Although structural and mechanistic studies of OTCase have lagged behind ATCase, several OTCase crystal structures have recently been determined,^{26–28} including two high-resolution structures from this laboratory, 15,16 generating increased interest and new research opportunities. The new crystal structures indicate that, while OTCases have considerable structural diversity, their fundamental building block is a trimer that is very similar to the catalytic trimers of ATCases, and that they share the conformational flexibility of ATCases. OTCases are important clinically, since OTCD is the most common cause of inherited hyperammonemia (elevated levels of blood ammonia), which in turn produces neurological symptoms, sometimes sufficiently severe to cause death. More than 130 mutations that give rise to OTCD in humans have been identified, 19 many in this laboratory, and this laboratory is actively engaged in developing approaches to gene therapy.

The structural and functional similarities and differences within the transcarbamylase family provide an opportunity to ask a number of questions about molecular recognition. What differences in the L-Asp binding site of ATCase and the L-Orn binding site of OTCase result in ATCase binding L-Asp and OTCase binding L-Orn? Conversely, how similar are the binding sites for CP? Are there similarities between the catalytic mechanisms of the OTCases and ATCases? What are the similarities and differences between subunit interfaces, within the OTCase and ATCase families and between families? What determines whether the basic functional unit, a trimer, forms larger aggregates, or associates with other proteins? What is the relationship between the state of aggregation of the protein and its regulation? This review will focus on these questions.

Sequence Comparisons

More than 30 OTCase genes and more than 15 ATCase genes have been sequenced. These sequences are aligned in Figures 2 and 3, so that similarities and differences can be readily visualized. Similar sequences are likely to give rise to similar three-dimensional structure. Despite their similar folds, E. coli OTCase and the catalytic subunit of E. coli ATCase have only 32% identical sequence. Homology between OTCase and ATCase is greatest in the N-terminal half of the sequence, the region that binds their common substrate, CP, and much weaker in their C-terminal halves, which bind L-Orn in OTCase and L-Asp in ATCase. Although human OTCase has an even lower level of sequence homology with the catalytic subunit of E. coli ATCase (27%) than with that of E. coli OTCase, the homology model of human OTCase that we built using the catalytic subunit of E. coli ATCase as a model successfully predicted the effects on enzymatic function of many naturally occurring mutations found in patients with OTCD (Figure 5) and was in good agreement with the experimentally determined crystal structure.²⁹ Most mutations that produce neonatal OTCD are found at the active site, in the interior of the protein where they would interfere with folding, or between subunits where they would interfere with assembly of the trimer, whereas mutations that produce late onset symptoms are more likely to be found in loops on the exterior of the protein.

Tertiary Structure

As shown in Figure 4, the overall topology of the subunits of both OTCases and ATCases is α/β , with 14–16 α -helices and 9–10 β -sheets. Each chain consists of two domains, a carbamoyl phosphate binding domain and a second domain that binds either L-Orn or L-Asp. Each domain consists of a central core made up of a β -sheet surrounded by α-helices. Helices H5 and H11 in E. coli OTCase and analogous helices in other transcarbamylases connect the two domains, with a highly conserved hydrophilic cluster holding helices H1, H5, and H11 (H12 in ATCase) together. Active chimeras produced by recombinant DNA methods consisting of the CP domain of E. coli ATCase and the L-Orn domain of *E. coli* OTCase, or vice versa, with the substrate specificity of the second domain, were reported a number of years ago.³⁰ However, more recent efforts to reproduce this result and to exchange secondary structure elements between the two proteins yielded only insoluble proteins, highlighting the importance of specific sidechain packing, even in the context of a shared protein fold.31

The individual chains of both the catalytic subunit of E. coli ATCase and those of E. coli and human OTCases undergo domain closure when substrates and substrate analogues bind. A loop known as the 240s loop in E. coli ATCase swings in toward the active site, enabling a number of adjacent residues to interact with bound substrate. A second loop, known as the 80s loop, undergoes a smaller motion. These motions are illustrated in Figure 6, which superimposes the liganded and unliganded catalytic subunits of E. coli ATCase and liganded and unliganded subunits of E. coli OTCase. In E. coli ATCase, movement of these loops propagates to the subunit interfaces between catalytic subunits and between catalytic and regulatory subunits, resulting in a large change in quaternary structure, the T-to-R transition. This conformational change was one of the earliest models of ligand-induced signal transduction and has served as a paradigm for signal transduction in many other systems, including those involved in intercellular signaling in the nervous system and in regulating cellular proliferation of normal cells and neoplastic growth of cancerous cells.

Carbamoyl Phosphate Binding Site

The binding sites for CP in OTCase and ATCase are very similar (Figure 7). Both active sites contain a STRT motif (Ser55-Thr56-Arg57-Thr58 in *E. coli* OTCase), which binds the phosphonate oxygens of CP, and Arg and His residues (Arg 106 and His 133 in *E. coli* OTCase), which bind the carbonyl oxygen of CP. (Sequence numbers throughout are those of *E. coli* OTCase and ATCase, unless otherwise

OTC_HUMAN	*NKVQ	LKGRDLLTLK	NFTGEEIKYM	LWLSADLKFR	IKQKGEYLPL	LQGKS	LGMIFEKRST	RTRLSTETGL	ALLGGHPCFL	TTQDIHLGVN
OTC_MOUSE	*SQVQ	LKGRDLLTLK	NFTGEEIQYM	LWLSADLKFR	IKQKGEYLPL	LQGKS	LGMIFEKRST	RTRLSTETGF	ALLGGHPSFL	TTQDIHLGVN
OTC_RAT	*SQVQ	LKGRDLLTLK	NFTGEEIQYM	LWLSADLKFR	IKOKGEYLPL	LQGKS	LGMIFEKRST	RTRLSTETGF	ALLGGHPSFL	TTQDIHLGVN
OTC RANCA										TTQDIHLGVN
OTC PACTA										GKDDIQLGVN
				IDRSSEIKQA						
										GKEDIQLGVN
	MOTIFACIE						47	57	67	77
		3	13	23	33	37				
				LQLAAKLKAD						
				LDLSRDLKRA						
				IDRAIECKRL						
				VRNASSHKHS						
OTCA_PSEAE	~~~~~~~	~SVRHFLSFM	DYSPEELIGL	IRRGSELK.D	LRNRGVLYEP	LKSRV	LGMVFEKAST	RTRLSFEAGM	IQLGGQAIFL	SPRDTQLGRG
OTC2_BACSU	~~MHTVTQTS	LYGKDLLTLK	DLSEEDINAL	LAEAGEL	KQNKIQPI	FHGKT	LAMIFEKSST	RTRVSFEAGM	AQLGGSALFL	SQKDLQLGRG
OTC_PYRFU	~~~~MVVS	LAGRDLLCLQ	DYTAEEIWTI	LETAKMFK.I	WQKIGKPHRL	LEGKT	LAMIFQKPST	RTRVSFEVAM	AHLGGHALYL	NAQDLQLRRG
										DSGSTQLGRD
										DPNSSOIGHK
										EPSASOIGHK
				LDLSHDLKAK						
							**	••		
							••	••		•
				AKEASIPIIN						
				AKEASIPIVN						
				AKEATIPIVN						
OTC_RANCA	ESLKDTARVL	SGMTDAVLAR	VYHQSDLEVL	AEEASIPIVN	GLSDDYHPIQ	ILADYLTIQE	HYG	HLKGLTI	SWIGDG.NNV	LHSIMMSAAK
				CKYSSVPIIN						
				AKHASCPVIN						
				CKDSSVPIIN						
	87	97	107	117	127	137	147	153	159	169
OFFICE PROFES										
				AEYASVPVWN						
				ADYAGVPVFN						
				AKHSGIPVWN						
				AKHSTVPVIN						
OTCA_PSEAE	EPIGDSARVM	SRMLDGVMIR	TFAHATLTEF	AAHSKVPVIN	GLSDDLHPCQ	LLADMQTFHE	HRG	SIQGKTV	AWIGDG.NNM	CNSYLEEEMK
OTC2_BACSU	ETVADTAKVL	SGYVDAIMIR	TFEHEKVEEL	AKEADIPVIN	GLTDKYHPCQ	ALADLLTIKE	IKG	KLKGVKV	AYIGDG.NNV	AHSLMIGCAK
OTC_PYRFU	ETIADTARVL	SRYVDAIMAR	VYDHKDVEDL	AKYATVPVIN	GLSDFSHPCQ	ALADYMTIWE	KKG	TIKGVKV	VYVGDG.NNV	AHSLMIAGTK
OTCA_MYCBO	ETLODTAKVL	SRYVDAIVWR	TFGOERLDAM	ASVATVPVIN	ALSDEFHPCO	VLADLOTIAE	RKG	ALRGLRL	SYFGDGANNM	AHSLLLGGVT
										GNSLLLIGAK
				AKYAGVPVFN						
										SYALMYGAAK
OICC_CLOPE	FOIEDIAKAF	GMIIDGIEFK	. Graysiveen		_		n.AnkP	DALKL	Vr IGHIMM	STALIFICAAN
		•		V	$\nabla \nabla \cdot \nabla \nabla$	∇			•	
OTC HUMAN	FGMHLOAATP	KGYEPDASVT	KT.AEOVAKEN	G TK	LLLTMDPLEA	AHCCMM/LT/PD	TWITSMOOTER	K KKBI OAFO	CYOVEMKEDAK	V AASTMITE
				GT.K						
OTC_MOUSE	FGMHLQAATP	KGYEPDPNIV	KLAEQYAKEN	GT.K	LSMINDPLEA	ARGGNVLITD	TWISMGQEDE	K.KKRLQAFQ	GYQVTMKTAK	VAASDWTF
OTC_MOUSE OTC_RAT	FGMHLQAATP FGMHLQAATP	KGYEPDPNIV KGYEPDPNIV	KLAEQYAKEN KLAEQYAKEN	$\begin{array}{l} \texttt{GT.K} \\ \texttt{GT.R} \end{array}$	LSMTNDPLEA LSMTNDPLEA	ARGGNVLITD ARGGNVLITD	TWISMGQEDE TWISMGQEDE	K.KKRLQAFQ K.KKRLQAFQ	GYQVTMKTAK GYQVTMKTAK	VAASDWTF VAASDWTF
OTC_MOUSE OTC_RAT _OTC_RANCA	FGMHLQAATP FGMHLQAATP FGMHLHIATP	KGYEPDPNIV KGYEPDPNIV KGYEPNSSLT	KLAEQYAKEN KLAEQYAKEN EAAKQFSKEC	GT.K GT.R GT.K	LSMTNDPLEA LSMTNDPLEA LLMTNDPLEA	ARGGNVLITD ARGGNVLITD ANGANVLVTD	TWISMGQEDE TWISMGQEDE TWVSMGQEEE	K.KKRLQAFQ K.KKRLQAFQ K.KKRLLDFK	GYQVTMKTAK GYQVTMKTAK GYQITMKTAK	VAASDWTF VAASDWTF LAAPNWIF
OTC_MOUSE OTC_RAT OTC_RANCA OTC_PACTA	FGMHLQAATP FGMHLQAATP FGMHLHIATP SGIDVSIAVP	KGYEPDPNIV KGYEPDPNIV KGYEPNSSLT SGLKFEELIL	KLAEQYAKEN KLAEQYAKEN EAAKQFSKEC SGAKEISAEN	GT.K GT.K GT.K GT.K	LSMTNDPLEA LSMTNDPLEA LLMTNDPLEA ITNDPLEA	ARGGNVLITD ARGGNVLITD ANGANVLVTD INGANVIVTD	TWISMGQEDE TWISMGQEDE TWISMGQEDE TWISMGQEDE	K.KKRLQAFQ K.KKRLQAFQ K.KKRLLDFK R.LQKLKQFE	GYQVTMKTAK GYQVTMKTAK GYQITMKTAK GFQITKEMIS	VAASDWTF VAASDWTF LAAPNWIF KGKAAENWKF
OTC_MOUSE OTC_RAT OTC_RANCA OTC_PACTA OTC_SCHPO	FGMHLQAATP FGMHLQAATP FGMHLHIATP SGIDVSIAVP VGIHTSVAKP	KGYEPDPNIV KGYEPDPNIV KGYEPNSSLT SGLKFEELIL KDVNVRDDIL	KLAEQYAKEN KLAEQYAKEN EAAKQFSKEC SGAKEISAEN SIVNEAANEN	G T.K G T.K G T.K G TTLK G STFE	LSMTNDPLEA LSMTNDPLEA LLMTNDPLEA ITNDPLEA IVNDPKVA	ARGGNVLITD ARGGNVLITD ANGANVLVTD INGANVIVTD VKNADIVVTD	TWISMGQEDE TWISMGQEEE TWISMGQEDE TWISMGQEAE	K.KKRLQAFQ K.KKRLLDFK K.KKRLLDFK R.LQKLKQFE K.EQRLKQFT	GYQVTMKTAK GYQVTMKTAK GYQITMKTAK GFQITKEMIS GFQVTGEIMK	VAASDWTF VAASDWTF LAAPNWIF KGKAAENWKF LAKPSCKF
OTC_MOUSE OTC_RAT OTC_RANCA OTC_PACTA OTC_SCHPO	FGMHLQAATP FGMHLQAATP FGMHLHIATP SGIDVSIAVP VGIHTSVAKP	KGYEPDPNIV KGYEPDPNIV KGYEPNSSLT SGLKFEELIL KDVNVRDDIL	KLAEQYAKEN KLAEQYAKEN EAAKQFSKEC SGAKEISAEN SIVNEAANEN	GT.K GT.K GT.K GT.K	LSMTNDPLEA LSMTNDPLEA LLMTNDPLEA ITNDPLEA IVNDPKVA	ARGGNVLITD ARGGNVLITD ANGANVLVTD INGANVIVTD VKNADIVVTD	TWISMGQEDE TWISMGQEEE TWISMGQEDE TWISMGQEAE	K.KKRLQAFQ K.KKRLLDFK K.KKRLLDFK R.LQKLKQFE K.EQRLKQFT	GYQVTMKTAK GYQVTMKTAK GYQITMKTAK GFQITKEMIS GFQVTGEIMK	VAASDWTF VAASDWTF LAAPNWIF KGKAAENWKF LAKPSCKF
OTC_MOUSE OTC_RAT OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST	FGMHLQAATP FGMHLQAATP FGMHLHIATP SGIDVSIAVP VGIHTSVAKP	KGYEPDPNIV KGYEPDPNIV KGYEPNSSLT SGLKFEELIL KDVNVRDDIL	KLAEQYAKEN KLAEQYAKEN EAAKQFSKEC SGAKEISAEN SIVNEAANEN	G T.K G T.K G T.K G TLK G TTLK G STFE G ATFE	LSMTNDPLEA LSMTNDPLEA LLMTNDPLEA ITNDPLEA IVNDPKVA	ARGGNVLITD ARGGNVLITD ANGANVLVTD INGANVIVTD VKNADIVVTD	TWISMGQEDE TWISMGQEEE TWISMGQEDE TWISMGQEAE	K.KKRLQAFQ K.KKRLLDFK K.KKRLLDFK R.LQKLKQFE K.EQRLKQFT	GYQVTMKTAK GYQVTMKTAK GYQITMKTAK GFQITKEMIS GFQVTGEIMK	VAASDWTF VAASDWTF LAAPNWIF KGKAAENWKF LAKPSCKF
OTC_MOUSE OTC_RAT OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST	FGMHLQAATP FGMHLQAATP FGMHLHIATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179	KGYEPDPNIV KGYEPDPNIV KGYEPNSSLT SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189	KLAEQYAKEN KLAEQYAKEN EAAKQFSKEC SGAKEISAEN SIVNEAANEN DEAKKVAERN 194	GT.K GT.K GTTLK GTTLK GSTFE GATFE	LSMTNDPLEA LSMTNDPLEA LLMTNDPLEA ITNDPLEA IVNDPKVA LTHDSLKA 212	ARGGNVLITD ARGGNVLITD ANGANVLVTD INGANVIVTD VKNADIVVTD STNANILVTD 222	TWISMGQEDE TWISMGQEDE TWVSMGQEEE TWISMGQEDE TWISMGQEAE TFVSMGEEFA 232	K.KKRLQAFQ K.KKRLQAFQ K.KKRLLDFK R.LQKLKQFE K.EQRLKQFT K.QAKLKQFK	GYQVTMKTAK GYQVTMKTAK GYQITMKTAK GFQITKEMIS GFQVTGEIMK GFQINQELVS 252	VAASDWTF VAASDWTF LAAPNWIF KGKAAENWKF LAKPSCKF VADPNYKF 262
OTC_MOUSE OTC_RAT OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC1_ECOLI	FGMHLQAATP FGMHLHIATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLRLVAP	KGYEPDPNIV KGYEPDPNIV KGYEPNSSLT SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP	KLAEQYAKEN KLAEQYAKEN EAAKQFSKEC SGAKEISAEN SIVNEAANEN DEAKKVAERN 194	GT.K GT.K GTTLK GTTLK GSTFE GATFE 202 RALAQQNGGN	LSMTNDPLEA LSMTNDPLEA LLMTNDPLEA ITNDPLEA IVNDPKVA LTHDSLKA 212 ITLTEDVAKG	ARGGNVLITD ARGGNVLITD ANGANVLVTD INGANVIVTD VKNADIVVTD STNANILVTD 222 VEGADFIYTD	TWISMGQEDE TWISMGQEDE TWVSMGQEEE TWISMGQEDE TWISMGQEAE TFVSMGEEFA 232 VWVSMGEAKE	K.KKRLQAFQ K.KKRLLDFK K.KKRLLDFK R.LQKLKQFE K.EQRLKQFT K.QAKLKQFK 242 KWAERIALLR	GYQVTMKTAK GYQVTMKTAK GYQITMKTAK GFQITKEMIS GFQVTGEIMK GFQINQELVS 252 EYQVNSKMMQ	VAASDWTF VAASDWTF LAAPNWIF KGKAAENWKF LAKPSCKF VADPNYKF 262 LTG.NPEVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC1_ECOLI OTC_HAEIN	FGMHLQAATP FGMHLQAATP FGMHLHIATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLRLVAP LGMDVRICGP	KGYEPDPNIV KGYEPDPNIV KGYEPNSSLT SGLKFEELIL KDVNVRDDIL FGIEMDSDIV 189 QACWP KALLP	KLAEQYAKEN KLAEQYAKEN EAAKQFSKEC SGAKEISAEN SIVNEAANEN DEAKKVAERN 1.EAALVTEC .EAALVTEC	G T . K G T . R G T . K G TTLK G STFE G ATFE 202 RALAQQNGGN EKFAKESGAR	LSMTNDPLEA LSMTNDPLEA LLMTNDPLEA ITNDPLEA ITNDPKVA LTHDSLKA 212 ITLTEDVAKG ITVTEDIDKA	ARGGNVLITD ARGGNVLITD ANGANVLVTD INGANVIVTD VKNADIVVTD STNANILVTD 222 VEGADFIYTD VKGVDFIHTD	TWISMGQEDE TWISMGQEDE TWISMGQEDE TWISMGQEDE TWISMGGEAE TFVSMGEEFA 232 VWVSMGEAKE VWVSMGEPLE	K.KKRLQAFQ K.KKRLLDFK R.LQKLKQFE K.EQRLKQFT K.QAKLKQFK 242 KWAERIALLR TWGERIKLLL	GYQVTMKTAK GYQVTMKTAK GYQITMKTAK GFQITKEMIS GFQVIGELWS GFQINQELVS 252 EYQVNSKMMQ PYQVTPELMK	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF Z62 LTG. NPEVKF RTG. NPKVKF
OTC_MOUSE OTC_RAT OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC1_ECOLI OTC_HAEIN OTCP_PSESH	FGMHLQAATP FGMHLQAATP FGMHLHIATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLRLVAP LGMDVRICGP FGYNLRIIAP	KGYEPDPNIV KGYEPDPNIV KGYEPNSSLT SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP KALLP	KLAEQYAKEN KLAEQYAKEN EAAKQFSKEC SGAKEISAEN SIVNEAANEN DEAKKVAERN 194 . EAALVTEC . EAALVTEC . TDAVLAGI	G T . K G T . R G T . K G TTLK G STFE G ATFE 202 RALAQQNGGN EKFAKESGAR YEQTPERNGS	LSMTNDPLEA LSMTNDPLEA LLMTNDPLEA ITNDPLEA ITNDPKVA LTHDSLKA 212 ITLTEDVAKG ITVTEDIDKA IEIFTEVAAG	ARGGNVLITD ARGGNVLITD ANGANVLVTD INGANVIVTD VKNADIVVTD STNANILVTD 222 VEGADFIYTD VKGVDFIHTD VHQADVIYTD	TWISMGQEDE TWISMGQEDE TWISMGQEDE TWISMGQEAE TFVSMGEEFA 232 WWVSMGEAKE VWVSMGEPLE VWISMGESVS	K.KKRLQAFQ K.KKRLLDFK R.LQKLKQFE K.EQRLKQFT K.QAKLKQFK 242 KWAERIALLR TWGERIKLLL V.EERIALLK	GYQVTMKTAK GYQVTMKTAK GYQITMKTAK GFQITKEMIS GFQVIGEIMK GFQINQELVS 252 EYQVNSKMMQ PYQVTPEIMK PYKVTEKMMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF
OTC_MOUSE OTC_RAT OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC1_ECOLI OTC_HAEIN OTCP_PSESH OTC_ASPTE	FGMHLQAATP FGMHLHATP FGMHLHIATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLRLVAP TGLDLRLVAP LGMDVRICGP FGYNLRILAP MGVDLAVATP	KGYEPDPNIV KGYEPDPNIV KGYEPNSSLT SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP KALLP KALLP KGYEIPASMR	KLAEQYAKEN KLAEQYAKEN EAAKQFSKEC SGAKEISAEN SIVNEAANEN DEAKKVAERN 194 . EAALVTEC . EANLVEMC . TDAVLAGI ELIQEAGKGV	GT.K GT.K GTTLK GSTFE GATFE 202 RALAQONGEN EKFAKESGAN YEQTPERNGS ANPGK	LSMTNDPLEA LSMTNDPLEA LLMTNDPLEA ITNDPLEA IVNDPKVA LTHDSLKA 212 ITLTEDVAKG ITVTEDIDKA IBIFTEVAAG LIQTNVPEEA	ARGGNVLITD ARGGNVLITD ANGANVLVTD INGANVIVTD VKNADIVVTD STNANILVTD 222 VEGADFIYTD VKGVDFIHTD VHQADVIYTD VKKADILVTD	TWISMGQEDE TWISMGQEDE TWISMGQEDE TWISMGQEAE TFVSMGEEFA 232 VWVSMGEAKE VWVSMGEPLE VWISMGESVS TWVSMGEEE	K.KKRLQAFQ K.KKRLLDFK K.LQKLKQFE K.LQKLKQFF K.QAKLKQFK 242 KWAERIALLR TWGERIKLLR V.EERIALLK S.LKRMKAFE	GYQVTMKTAK GYQVTMKTAK GYQITMKTAK GFQITKEMIS GFQVTGEIMK GFQINQELVS 252 EYQVNSKMMQ PYQVTPELMK PYKVTEKMMA GFQITSELAK	V. AASDWTF V. AASDWTF L. AAPNWIF EGRAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPEVKF LTG. KADTIF RGGANENWKF
OTC_MOUSE OTC_RAT OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC1_ECOLI OTC_HAEIN OTCP_PSESH OTC_ASPTE OTCA_PSEAE	FGMHLQAATP FGMHLHATP FGMHLHIATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLRLVAP LGMDVRICGP FGYNLRIIAP MGYDLAVATP FDFQLRVACP	KGYEPDPNIV KGYEPDPNIV KGYEPDSIT SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP KALLP NALHP KGYEIPASMR EGYEPKAEFV	KLAEQYAKEN KLAEQYAKEN EAAKQFSKEC SGAKEISAEN SIVNEAANEN DEAKKVAERN 194 . EAALVTEC . EANLVEMC . TDAVLAGT ELIQEAGKGV ALAGDR	GT.K GT.K GTTLK GSTFE GATFE 202 RALAQQNGGN EKFAKESGAR YEQTPERNGS ANPGK	LSMTNDPLEA LSMTNDPLEA ILMTNDPLEA I.TNDPLEA I.VNDPKVA L.THDSLKA 212 ITLTEDVAKG ITVTEDIDKA IEIFTEVAAG LIQTNVPEEA LRVVRDPEEA	ARGGNVLITD ARGGNVLITD ARGANVLVTD INGANVLVTD VKNADIVVTD VKNADIVVTD VKVADIVTD VKGVDFIHTD VKGVDFIHTD VKQADVIYTD VKKADILVTTD VKKADILVTTD VKAGAHLVSTD	TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEAE TRYSMGEEFA 232 VWYSMGEAKE VWISMGEPLE VWISMGEVLE VWISMGEVLE VWISMGEVLE VWISMGEVLE VWISMGEVLE VWISMGEVLE VWISMGEVLE VWISMGEVLE VWISMGEVLE	K.KKRLQAFQ K.KKRLLDFK R.LQKLKQFE R.LQKLKQFE K.BCRLKQFT K.QAKLKQFK 242 KWAERIALLR TWEERIKLLL V.EERIALLK S.LKRMKAFE A.AARIAMFR	GYQVTMKTAK GYQVTMKTAK GYQITMKTAK GFQITKEMIS GFQVTGEIMK GFQINQELVS 252 EYQVNSKMMQ PYQVTPELMK PYKVTEKMMA GFQITSELAK PYQVNAALLD	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LITG. NPEVKF RTG. NPEVKF LTG. KADTIF RGGANENWKF G. AADDVF
OTC_MOUSE OTC_RAT OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC1_ECOLI OTC_HAEIN OTCP_PSESH OTC_ASPTE OTCA_PSEAE OTC2_BACSU	FGMHLQAATP FGMHLHATP FGMHLHIATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLRLVAP LGMDVRICGP FGYNLRI IAP MGCDISIASP	KGYEPDPNIV KGYEPDPNIV KGYEPDNSLT SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP KALLP NALHP KGYEIPASMR EGYEPKAEFV KGYEUPAAM	KLAEQYAKEN KLAEQYAKEN EAAKQFSKEC SGAKEISAEN SIVNEAANEN DEAKKVAERN 194 . EAALVTEC . EAALVTEC . TEAVLAGI ELIQEAGKGY ALAGOR EAAKTYALQS	G T.K G T.R G T.K G T.K G STFE G ATFE 202 RALAQQNGGN EKFAKESGAR YEQTPERNGS A NFGK	LSMTNDPLEA LSMTNDPLEA I.MTNDPLEA I.TNDPLEA I.VNDPKVA L.THDSLKA 212 1TLTEDVAKG ITUTEDIDKA IEIFTEVAAG LIQTNVPEEA LIQTNVPEEA VTLTDDPIEA	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVID VKNADIVVID STNANILVID 222 VEGADFIYTD VKGVDFIHTD VKQADVIYTD VKRADILVITD VKRADILVITD VKRADILVITD VKRADILVITD VKRADIVIYSD	TWISMQQDE TWISMQQDE TWISMQQDE TWISMQQEAE TFYSMGEFA 232 VWYSMGEAK VWSMGEPLE VWISMGESVS TWYSMQEEE VWASMGQEDE VFTSMQQEAE	K.KKRLQAFQ K.KKRLLDFK K.KKRLLDFK K.EQRLKQFE K.EQRLKQFT K.QAKLKQFK 242 242 TWEERIALLR TWEERIALLL V.EERIALLK S.LKRMKAFE A.AARIAMFR E.QERLAVFA	GYQVTMKTAK GYQVTMKTAK GYQVTMKTAK GFQTTKEMIS GFQVTGEIMK GFQTNQELVS 252 EYQVNSKMMQ PYQVTSEIMK PYKVTEKMMA GFQITSELAK PYQVNAALLJ PYQVNAALLJ PYQVNAALUS	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYTF
OTC_MOUSE OTC_RANCA OTC_RANCA OTC_SCHPO OTC_YEAST OTC_HAEIN OTC_HAEIN OTC_ASPTE OTCA_PSEAE OTC2_BACSU OTC_PYRFU	FEMHLQAATP FEMHLHIATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLRLVAP LCMDVRICGP FGYNLRI IAP MGVDLAVATP FDFQLRVACP MGCDISIASP LGADVVATP	KGYEPDENIV KGYEPDENIV KGYEPNSSLT SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP KALLP NALHP KGYEIPASMR EGYEPFKAEFV KGYEVLDEAA EGYEPDEKVI	KLADOYAKEN KLAEOYAKEN KLAEOYAKEN SAKEISAEN SIVNEANNEN 194 . EAALVTEC . EANLVEMC . TDAVLAGI ELIQEAGKGV ALAGOR	G T.K G T.K G T.K G STFE G ATFE 202 RALAQQNGSN EKFAKESGAR YEQTPERNGS A NPGK	LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITHDSLKA 212 ITLTEDVAKG ITVTEDIDKA LIFITEVAKA LICTNVPEEA LRVVNDPREA VTLTDDPIEA	ARGGNVLITD ARGGNVLITD ARGNVLVTD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKGVDFHTTD VKGVDLYTTD VKGADLVSTD VKGADLVSTD VKGADLVSTD VKGDADVIYTD VKDADVIYTD	TWISMGQEDE TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWVSMGEAKE VWVSMGEVS TWVSMGQEE VWASMGGOEE VWASMGQEDE VWASMGQEDE VWASMGQEAE	K.KKRLQAFQ K.KKRLLDFQ K.KKRLLDFQ K.EORLKQFT K.EORLKQFT Z42 KWAERIALLR TWEERIKLLLR TWEERIKLLL S.LKRMKAFE A.AARIAMFR E.QERLAVFA	GYQVTMKTAK GYQVTMKTAK GYQITMKTAK GFQITKEMIS GFQVTOEIIMK GFQINQELVS 252 EYQVNSKMMQ PYQVTPEIMK GFQITSELAK PYGVNAALLD PYQVNAALLD PYQVNAALLS PYQVNAALVS	V. AASDWTF V. AASDWTF L. AAPPWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPEVKF RTG. NPKVKF RGGANENWKF G. AADDVLF H. AKPDYTF H. AKPDYMF
OTC_MOUSE OTC_RAT OTC_RACTA OTC_SCHPO OTC_YEAST OTC_HAEIN OTC_ASPTE OTCA_PSEAE OTCA_PSEAE OTCA_PSEAE OTCA_PYEFU OTCA_MYCBO	FGMHLQAATP FGMHLHIATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLRLVAP LGMDVRICGP FGYNLRIIAP MGYDLAVATP FDFQLRVACP MGCDISIASP AGTHVIVAAP	KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVMVRDDIL PGIEMDSDIV 189 QACWP KALLP NALHP KGYEIPASMR EGYEPEKAEFV KGYEVLDEAA EGYEPDEKVI BGFLPDESVR	KLAEQYAKEN KLAEQYAKEN EAAKQFSKEC SGAKEISAEN SIVNEAANEN 194 . EAALVTEC . EANLVEMC . TDAVLAGI ELIQEAGKGV ALAGDR EAAKTYALQS KWAEQNAAES AAAERRAQDT	G T.K G T.K G T.K G T.K G STFE G ATFE 202 RALAQQNGGN EKFAKESGAR YEQTPENNGS A NPGK G S.S G S.S G A.S	LSMTNDPLEA LSMTNDPLEA LLMTNDPLEA ITNDPLEA IVNDEKVA LTHDSLKA 212 212 ITLTEDVAKG ITVPEDIDKA IEIFTEVAAG IEIFTEVAAG LIQTNVPEEA LRVVRDPREA VTLTDDPIEA FELLHDPVKA VTVTTADAHAA	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVLVTD INGANVLVTD STNANILVTD 222 VEGADFIYTD VKQADVIYTD VKQADVIYTD VKRADILVTT VKGAHLVSTD VKDADVIYTD AAAADVLVTD AAAADVLVTD	TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEAE TRYSMGEEFA 232 VWYSMGEAKE VWYSMGEPLE VWYSMGEPLE VWYSMGQEE VWYSMGQEE VWYSMGQEE VWYSMGQEE VWYSMGQEE VWYSMGQEAE TWYSMGQEAE TWYSMGQEAE	K.KKRLQAFQ K.KKRLQAFQ K.KKRLLDFK R.LQKLKQFE K.EQRLKQFT K.QAKLKQFK 242 KWAERIALLR TWGERIKLLL V.EERIALLK S.LKRWKAFE A.ARIAMFR E.QERLAVFA A.EERRKIFR G.LDRVKPFR	GYQVTMKTAK GYQVTMKTAK GYQITMKTAK GFQITKEMIS GFQVTGEIMK GFQITOBLUS 252 EYQVNSKMMQ PYQVTPEIMK PYKVTEKMMA PYKVTEKMAK PYCVNAALLD PYQVNAALLD PYQVNAALLS PYQVNAALUS PYQUNAALVS PPQLNSKILA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYTF H. AKPDYMF L. ABSDAIV
OTC_MOUSE OTC_RAT OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC1_ECOLI OTC_HAEIN OTCP_PSESH OTC2_EACSU OTC_PYERD OTCA_MYCBO OTCC_PSEAE	FGMHLQAATP FGMHLHATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLRLVAP LGMDVRICGP FGYNLRI IAP MGVDLAVATP FDFOLRVACP MGCDISIASP LGADVVVATP AGTHVVAAP LGADVVIAAP	KGYEPDPNIV KGYEPDPNIV KGYEPDNSLT SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP KALLP NALHP KGYEIPASMR EGYEPKAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDPSVR KALWP	KLAEQYAKEN KLAEQYAKEN EAAKQFSKEC SGAKEISAEN SIVNEAANEN DEAKKVAERN 194 . EAALVTEC . EANLVEMC . TDAVLAGI ELIQEAGKGV ALAGDR EAAKTYALQS KWAEQNAAES AAAERRAQDT HDEFYAQC	G T.K G T.K G T.K G T.K G STFE G ATFE 202 RALAQONGGN EKFAKESGAR YEQTPERNGS A NPGK	LSMTNDPLEA LSMTNDPLEA LLMTNDPLEA ITNDPLEA IVNDPKVA LTHDSLKA 212 ITUTEDVAKG ITUTEDIDKA IELFTEVAAG LLQTNVPEEA VTLTDDPLEA VTLTDDPLEA VTLTDDPLEA FELLHDPVKA VTVTADAHAA LTTTEDPKEA	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVLVID VKNADIVVID STNANILVID 222 VEGADFIYTD VKGVDFIHID VKGADVIYTD VKKADILVTTD VKADAULVTTD VKADAULVTTD VKDADVIYTSD VKDADVIYTSD VKDADVLVTD VKGVDFVHTD VKGVDFVHTD	TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEAE TFYSMGEFFA 232 VWYSMGEAKE VWISMGEVLE VWISMGEVLE VWISMGEVLE VWISMGEVE VWASMGQEDE VFTSMGQEAE VWASMGQEAE VWASMGQEAE VWASMGQEAE VWASMGQEAE VWASMGQEAE VWASMGCEAE VWASMGCEAE VWASMGCEAE	K.KKRLQAFQ K.KKRLLDFK R.LQKLKQFE K.EQRLKQFT K.QAKLKQFK ZAGAKKALKQFK ZAGAKKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAK	GYQVTMKTAK GYQVTMKTAK GYQVTMKTAK GYQITMKTAK GFGITKEMIS GFQVTGEIMK GFGINQELVS 252 EYQVNSKMMQ PYQVTPEIMK PYKVTEKMMA GYGVTSELAK PYQVNAALLD PYQVNAALLD PYQVNAALLS PFQUNKSLLA PYQVNASIIM PYQVNASIIM	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LITG. NPEVKF RTG. NPEVKF RTG. NPKVKF LITG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYTF H. AKPDYMF L. ADSDAIV ATG. NPRAKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_LECOLI OTC_HAEIN OTCP_PSESH OTC_ASPTE OTCA_PSEAE OTC_PYRFU OTCA_MYCBO OTCC_PSEAE OTCC_PSEAE OTCC_NEIGO	FEMHLQAATP FEMHLHIATP FEMHLHIATP SGIDVSIAVP VGHTTSVAKP FETISVSISTP 179 TELDLKLVAP LEMDVRICGP FEGVILRI IAP MGVDLAVATP FDFQLRVACP MGCDISTASP LGADVVVATP AGTHVTVAAP LEMDVRIAAP LEMDVRIAAP	KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACMP. KALLP. NALHP. KGYEIPASMR EGYEPFAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDESVR KALWP. QSLWP.	KLADOYAKEN KLAEOYAKEN KLAEOYAKEN SGAKEISAEN SIVNEANNEN 194 . EAALVYEC . EANLVEMC . TDAVLAGI ELIQEAGKGV ALAGDR	G T.K G T.K G T.K G TT.K G STFE G ATFE 202 RALAQQNGSN EKFAKESGAR A NPGK	LSMINDPLEA LSMINDPLEA ITNDPLEA ITNDPLEA ITNDPKVA LTHDSLKA 212 ITLTEDVAKG ITVTEDIDKA LIQTNVPEEA LRVVNDPREA VTUTDDDIEA FELLHDPVKA VTVTADAHAA LTLTEDPKEA	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKGADFIYTD VKGADIVYTT VKGADILVTTD VAGAHUSTTD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKGVDFVHTD VKGVGFHTTD VKGVGFHTTD VKGVGFHTTD	TWISMGQEDE TWISMGQEDE TWISMGQEE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWISMGEVS TWVSMGEVS TWVSMGEVS TWVSMGEAE VWASMGGEE VWASMGGEE VWASMGGEE VWASMGEE VWASMGEE VWASMGEE VWASMGEE VWASMGEE VWSMGEE VWYSMGEE VWVSMGEE VWSMGEE VWSMGE VSMGE VS	K.KKRLQAFQ K.KKRLLDFK R.LQKLKQFE K.BQRLKQFT K.BQRLKQFT K.QAKLKQFK 242 KWAERIALLR TWGERIKLLL S.LKRMKAFE A.AARIAMFR A.BERRKIFR G.LDRVKPFR AWGERIKELL VWQERIDLLK	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQVITCEIMK GFQITSELWS 252 EYQVNSKMMQ PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNAALLUS PFQVNKDLVK PFQUNSKLLA PYQVNAELLA PYQVNAELLA PYQVNAELLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RAT OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC1_ECOLI OTC_HAEIN OTCP_PSESH OTC2_EACSU OTC_PYERD OTCA_MYCBO OTCC_PSEAE	FEMHLQAATP FEMHLHIATP FEMHLHIATP SGIDVSIAVP VGHTTSVAKP FETISVSISTP 179 TELDLKLVAP LEMDVRICGP FEGVILRI IAP MGVDLAVATP FDFQLRVACP MGCDISTASP LGADVVVATP AGTHVTVAAP LEMDVRIAAP LEMDVRIAAP	KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACMP. KALLP. NALHP. KGYEIPASMR EGYEPFAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDESVR KALWP. QSLWP.	KLADOYAKEN KLAEOYAKEN KLAEOYAKEN SGAKEISAEN SIVNEANNEN 194 . EAALVYEC . EANLVEMC . TDAVLAGI ELIQEAGKGV ALAGDR	G T.K G T.K G T.K G TT.K G STFE G ATFE 202 RALAQQNGSN EKFAKESGAR A NPGK	LSMINDPLEA LSMINDPLEA ITNDPLEA ITNDPLEA ITNDPKVA LTHDSLKA 212 ITLTEDVAKG ITVTEDIDKA LIQTNVPEEA LRVVNDPREA VTVTDDDPEA FELLHDPVKA VTVTADAHAA LTTTEDPKEA	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKGADFIYTD VKGADIVYTT VKGADILVTTD VAGAHUSTTD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKGVDFVHTD VKGVGFHTTD VKGVGFHTTD VKGVGFHTTD	TWISMGQEDE TWISMGQEDE TWISMGQEE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWISMGEVS TWVSMGEVS TWVSMGEVS TWVSMGEAE VWASMGGEE VWASMGGEE VWASMGGEE VWASMGEE VWASMGEE VWASMGEE VWASMGEE VWASMGEE VWSMGEE VWYSMGEE VWVSMGEE VWSMGEE VWSMGE VSMGEE VWSMGE VSMGE VS	K.KKRLQAFQ K.KKRLLDFK R.LQKLKQFE K.BQRLKQFT K.BQRLKQFT K.QAKLKQFK 242 KWAERIALLR TWGERIKLLL S.LKRMKAFE A.AARIAMFR A.BERRKIFR G.LDRVKPFR AWGERIKELL VWQERIDLLK	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQVITCEIMK GFQITSELWS 252 EYQVNSKMMQ PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNAALLUS PFQVNKDLVK PFQUNSKLLA PYQVNAELLA PYQVNAELLA PYQVNAELLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_LECOLI OTC_HAEIN OTCP_PSESH OTC_ASPTE OTCA_PSEAE OTC_PYRFU OTCA_MYCBO OTCC_PSEAE OTCC_PSEAE OTCC_NEIGO	FEMHLQAATP FEMHLHIATP FEMHLHIATP SGIDVSIAVP VGHTTSVAKP FETISVSISTP 179 TELDLKLVAP LEMDVRICGP FEGVILRI IAP MGVDLAVATP FDFQLRVACP MGCDISTASP LGADVVVATP AGTHVTVAAP LEMDVRIAAP LEMDVRIAAP	KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACMP. KALLP. NALHP. KGYEIPASMR EGYEPFAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDESVR KALWP. QSLWP.	KLADOYAKEN KLAEOYAKEN KLAEOYAKEN SGAKEISAEN SIVNEANNEN 194 . EAALVYEC . EANLVEMC . TDAVLAGI ELIQEAGKGV ALAGDR	G T.K G T.K G T.K G TT.K G STFE G ATFE 202 RALAQQNGSN EKFAKESGAR A NPGK	LSMINDPLEA LSMINDPLEA ITNDPLEA ITNDPLEA ITNDPKVA LTHDSLKA 212 ITLTEDVAKG ITVTEDIDKA LIQTNVPEEA LRVVNDPREA VTVTDDDPEA FELLHDPVKA VTVTADAHAA LTTTEDPKEA	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKGADFIYTD VKGADIVYTT VKGADILVTTD VAGAHUSTTD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKGVDFVHTD VKGVGFHTTD VKGVGFHTTD VKGVGFHTTD	TWISMGQEDE TWISMGQEDE TWISMGQEE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWISMGEVS TWVSMGEVS TWVSMGEVS TWVSMGEAE VWASMGGEE VWASMGGEE VWASMGGEE VWASMGEE VWASMGEE VWASMGEE VWASMGEE VWASMGEE VWSMGEE VWYSMGEE VWVSMGEE VWSMGEE VWSMGE VSMGEE VWSMGE VSMGE VS	K.KKRLQAFQ K.KKRLLDFK R.LQKLKQFE K.BQRLKQFT K.BQRLKQFT K.QAKLKQFK 242 KWAERIALLR TWGERIKLLL S.LKRMKAFE A.AARIAMFR A.BERRKIFR G.LDRVKPFR AWGERIKELL VWQERIDLLK	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQVITCEIMK GFQITSELWS 252 EYQVNSKMMQ PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNAALLUS PFQVNKDLVK PFQUNSKLLA PYQVNAELLA PYQVNAELLA PYQVNAELLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_LECOLI OTC_HAEIN OTCP_PSESH OTC_ASPTE OTCA_PSEAE OTC_PYRFU OTCA_MYCBO OTCC_PSEAE OTCC_PSEAE OTCC_NEIGO	FEMHLQAATP FEMHLHIATP FEMHLHIATP SGIDVSIAVP VGHTTSVAKP FETISVSISTP 179 TELDLKLVAP LEMDVRICGP FEGVILRI IAP MGVDLAVATP FDFQLRVACP MGCDISTASP LGADVVVATP AGTHVTVAAP LEMDVRIAAP LEMDVRIAAP	KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACMP. KALLP. NALHP. KGYEIPASMR EGYEPFAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDESVR KALWP. QSLWP.	KLADOYAKEN KLAEOYAKEN KLAEOYAKEN SGAKEISAEN SIVNEANNEN 194 . EAALVYEC . EANLVEMC . TDAVLAGI ELIQEAGKGV ALAGDR	G T.K G T.K G T.K G TT.K G STFE G ATFE 202 RALAQQNGSN EKFAKESGAR A NPGK	LSMINDPLEA LSMINDPLEA ITNDPLEA ITNDPLEA ITNDPKVA LTHDSLKA 212 ITLTEDVAKG ITVTEDIDKA LIQTNVPEEA LRVVNDPREA VTVTDDDPEA FELLHDPVKA VTVTADAHAA LTTTEDPKEA	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKGADFIYTD VKGADIVYTT VKGADILVTTD VAGAHUSTTD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKGVDFVHTD VKGVGFHTTD VKGVGFHTTD VKGVGFHTTD	TWISMGQEDE TWISMGQEDE TWISMGQEE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWISMGEVS TWVSMGEVS TWVSMGEVS TWVSMGEAE VWASMGGEE VWASMGGEE VWASMGGEE VWASMGEE VWASMGEE VWASMGEE VWASMGEE VWASMGEE VWSMGEE VWYSMGEE VWVSMGEE VWSMGEE VWSMGE VSMGEE VWSMGE VSMGE VS	K.KKRLQAFQ K.KKRLLDFK R.LQKLKQFE K.BQRLKQFT K.BQRLKQFT K.QAKLKQFK 242 KWAERIALLR TWGERIKLLL S.LKRMKAFE A.AARIAMFR A.BERRKIFR G.LDRVKPFR AWGERIKELL VWQERIDLLK	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQVITCEIMK GFQITSELWS 252 EYQVNSKMMQ PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNAALLUS PFQVNKDLVK PFQUNSKLLA PYQVNAELLA PYQVNAELLA PYQVNAELLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANT OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST CTC1_ECOLI OTC_HAEIN OTC_PSESH OTCA_PSEAE OTC2_EACSU OTCA_PYFFU OTCA_MYCBO OTCC_NEIGO OTCC_CLOPE	FEMHLQAATP FEMHLHIATP SGIDVSIAVP VGIHTSVAKP FEGISVSISTP 179 TGLDLRLVAP LGMDVRICGP FGYNLRIIAP MGYDLAVATP FDFQLRVACP MGCDISIASP AGGHVVVATP AGGHVVATP AGGHVVALGP MGWDVALGADVVATP AGGHVVALGP MGWDVALGADVVALGP MGWDVALGADVVALGP	KGYEPDPNIV KGYEPDPNIV KGYEPDRNIV SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP KALLP NALHP NALHP KGYEIPASMR EGYEPKAEFV KGYEVLDEAA EGYEPDEKVI EGYEPDESVR KALWP QSLWP QSLWP DSLKP	KLAEQYAKEN KLAEQYAKEN KLAEQYAKEN SIVMEAANEN DEAKKVAERN 194 . EAALVTEC . EANLVEMC . TDAYLAG1 ELIQEAGKGV ALAGDR EAAKTYALQS KWAEQNAAES AAAERRAQDTHDEFVAQCSEGIIAAADEDILKEM	G T.K G T.K G T.K G T.K G STFE G ATFE 202 RALAQQNGGN EKFAKESGAR YEQTPERNGS A NPGK	LSMTNDPLEA LSMTNDPLEA LSMTNDPLEA ITNDPLEA IVNDEKVA LTHDSLKA 212 212 ITLTEDVAKG ITVTEDIDKA IELITTEVAKG IELITTEVAKG LIQTNVPEEA LRVVRDPREA VTUTDDPLEA FELLHDPVKA VTVTADAHAA LTLTEDPKEA ITLTENAHEA IEFSSNVDEA	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVLVTD INGANVLVTD VKNADIVVTD STNANILVTD 222 VEGADFIYTD VKGQDFIHTD VKQADVIYTD VKRADILVTD VKRADILVTD VKDADVIYSD VKDADVIYTD VKGVDFVHTD VKGVGFIHTD VKGVGFIHTD VKGVGVDFVHTD VKGADVIYTD	TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEAE TRYSMGEEFA 232 VWYSMGEAKE VWISMGEPLE	K.KKRLQAFQ K.KKRLQAFQ K.KKRLDAFK R.LOKLKOFE K.EORLKOFT K.OAKLKOFK Z42 KWAERIALLR TWGERIKLLL V.EERIALLK S.LKRMKAFE A.AARIAMFR E.OERLAVFA A.EERRKIFR G.LDRVKPFR AWGERIKELL VWOERIDLLK LYPERVKLLT	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQVITCEIMK GFQITSELWS 252 EYQVNSKMMQ PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNAALLUS PFQVNKDLVK PFQUNSKLLA PYQVNAELLA PYQVNAELLA PYQVNAELLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_LECOLI OTC_HAEIN OTCP_PSESH OTC_ASPTE OTCA_PSEAE OTC2_PACSU OTC_PYRFU OTCA_MYCEO OTCC_NEIGO OTCC_CLOPE OTC_LHUMAN	FEMHLQAATP FEMHLHIATP SGIDVSIAVP VGHHTSVAKP FGISVSISTP 179 TGLDLRLVAP LCMDVRLCGP FGYNLRI IAP MGVDLAVATP FDFQLRVACP MGCDISIASP LGADVVVATP AGIHVTVAAP LCMDVRIAAP	KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP KALLP NALHP KGYEIPASMR EGYEPKAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDESVR KALWP QSLWP DSLKP DSLKP	KLADOYAKEN KLAEOYAKEN KLAEOYAKEN SGAKEISAEN SIVNEANNEN 194 . EAALVTEC . EAALVTEC . TDAVLAGI ELIQEAGKGV ALAGDR . LAGARTALQS KWAEQNAAES AAAERRAQDT . HDEFVAQC . SEGIIAAA . DEDILKEM	G T.K G T.K G T.K G TT.K G STFE G ATFE 202 RALAQQNGN EKFAKESGAR YEQTPERNGS A NPGK	LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPKVA LTHDSLKA 212 ITLTEDVAKG ITVTEDIDKA LIEFTSTVAKG LIQTNVPEFA LRVVRDPREA VTVTLODDELEA FELLHDPVKA VTVTLODAHAA LTTLTEDPKEA LTTLTEDPKEA LTTLTEDPKEA LTTLTEDPKEA LSTLTTLTENAHEA LEFSSNVDEA	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKGVDFHTTD VKGADILVTD VKGADLUSTD VKGADLUSTD VKGADVIYSD VKDADVIYSD VKGVDFWTTD VKGVDFWTTD VKGVGFHTTD VKGVBFHTTD VKGVBFHTTD VKGVBFHTTD VKGVBFHTTD VKGVBFHTTD VKGVBFHTTD VKGVBFHTTD VKGVBFHTTD VKGWBFHTTD VKGWBFH	TWISMQQEDE TWISMQQEDE TWISMQGEE TWISMGGEE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWVSMGEPKE VWISMGEVS TWVSMQQEE VWASMGQEE VWASMGQEE VWASMGQEAE TWISMGEDE VWASMGEE VWASMGEE VWASMGEE VWASMGEE VWASMGEE VWASMGEE VWASMGEE VWASMGEE	K.KKRLQAFQ K.KKRLLDFK R.LQKLKQFE K.BORLKQFT K.BORLKQFT K.QAKLKQFK 242 KWAERIALLR TWEBRIKLLL S.LKRMKAFE A.AARIAMFR G.LDRVKPFR ABGERIKELL VWQERIDLLK LYPERVKLLT	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNAELLA PYQVNAELLA FPQLNSKLLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_HAEIN OTC_ASPTE OTCA_PSEAE OTCA_PSEAE OTCA_PYRFU OTCA_MYCBO OTCC_CLOPE OTC_HUMAN OTC_MOUSE	FEMHLQAATP FEMHLHIATP FEMHLHIATP SGIDVSIAVP VGIHTSVAKP FEGISVSISTP 179 TGLDLRLVAP LCMDVRICGP FGYNLRIIAP FGYNLRIIAP FDFQLRVACP MGCDISIASP LCADVVATP AGIHVTVAAP LCMDVRIAAP LCMDVRIAAP LCMDVRIAAP LCMDVRIAAP LCMDVRIAAP LCMDVRIAAP	KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVMVRDDIL PGIEMDSDIV 189 QACWP KALLP KALLP KALLP KGYEIPASMR EGYEPEKAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDPSVR KALMP QSLWP DSLKP	KLADQYAKEN KLAEQYAKEN KLAEQYAKEN SGAKEISAEN SIVMEANNEN 194 . EAALVTEC . SEGIIAAA . DEDILKEM	G T.K G T.K G T.K G TT.K G STFE G ATFE 202 RALAQQNGGN EKFAKEGAR YEQTPERNGS A NPGK	LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPKVA LTHDSLKA 212 ITLTEDVAKG ITUTEDIDKA LIGHTEVAAG LIGTNVPEEA LRVVRDPREA VTLTDDPIEA LTUTEDPKEA LTUTEDPKEA LTLTEDPKEA LTLTEDPKEA LTLTEDPKEA LTLTEDFKEA SLVFPEAENR SLVFPEAENR	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD STNANILVTD YENGADFIYTD VKGYDFIHTD VKGADVIYTD VKGADLIVTD VKGAHLWSTD VKGADVIYTD AAAADVLVTD VKGYDFYHTD VKGYDFYHTD VKGYDFYHTD VKGYDFYHTD VKGYDFYHTD VKGYMFYHTD VKGYMFYHTD VKGYMFYHTD VKGYMFYHTD VKGYMFYHTD VKGYMFYHTD VKGWMFYHTD	TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEE TWISMGGEAE TVYSMGEFA 232 VWVSMGEAKE VWVSMGEPLE VWISMGEVE VWISMGEVE VWISMGEVE VWISMGEVE VWASMGQEDE VWASMGQEDE VWASMGQEAE TWTSMGQEAE TWTSMGQEAE IWVSMGEPVE VWVSMGEPKE LLITDYSPQLQ LLITDYSPVLQ	K.KKRLQAFQ K.KKRLQAFQ K.KKRLLAFQ R.LQKLKQFE K.BORLKQFT K.QAKLKQFK Z42 KWAERIALLR TWGERIKLLL V.EERIALLK S.LKRMKAFE A.AARIAMFR E.QERLAVFA A.BERRRIFR G.LDRVKPFR AWGERIBLLK LYPERVKLLT KPKF KPKF	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNAELLA PYQVNAELLA FPQLNSKLLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_LECOLI OTC_HAEIN OTCP_PSESH OTC_ASPTE OTCA_PSEAE OTC2_BACSU OTC_PYRFU OTCA_PYRFU OTCA_MYCBO OTCC_NEIGO OTCC_CLOPE OTC_MOUSE OTC_MOUSE OTC_RAT	FEMHLQAATP FGMHLHIATP FGMHLHIATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLRLVAP LGMDVRICGP FGYNLRIIAP MGYDLAVATP FDFQLRVACP MGCDISIASP AGHVIVAAP LGMDVVIAP LGMDVVIAP LGMDVRIGAP MGWHFVALGP LHCLPRKPE LHCLPRKPE	KGYEPDENIV KGYEPDENIV KGYEPNSSLT SGLKFEELIL KDVMVRDDIL 189 QACWP. KALLP. NALHP. KGYEIPASMR EGYEPKAEFV KGYEVLDEAA EGYEPDEKVI EGFLPPDEVI KALLP. CSLWP. CSLWP. CSLWP. DSLKP.	KLABOYAKEN KLABOYAKEN KLABOYAKEN SAAKEISAEN SIVNEANEN 194	G. T.K G. T.K G. T.K G. T.K G. ATFE G. ATFE 202 RALAQONGEN A. NPGK A. NPGK G. S.S G. S G. AS KKFAEESGAK HAAAKETGAK QEYSKETGAT VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR	LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDSLKA 212 ITLTEDVAKG ITUTEDVAKG ITUTEDVAKG LIQTNVPEEA LIQTNVPEEA VTLTDDPLEA FELLHDPVKA VTLTDDPLEA LTLTENAHEA LTLTENAHEA LTLTENAHEA LTLTENAHEA SLVFPEAENR SLVFPEAENR	ARGGNVLITD ARGGNVLITD ARGGNVLIVID INGANVIVID INGANVIVID INGANVIVID STNANILVTD 222 VEGADFIYTD VKGVDFIHTD VKGADVIYTD VKRADILVTD VKGADVIYTD VKGADVIYTD VKGADVLYTD VKGDADVLYTD VKGDGFIHTD VKGGVFHTD VKGADVIYTD KKGVGFHTD VKGADVIYTD KWTIMAVMVS KWTIMAVMVS KWTIMAVMVS	TWISMQQEDE TWISMQQEDE TWISMQQEE TWISMQGEE TWISMGGEAE TYVSMGEAKE 232 VWVSMGEPLE VWVSMGEPLE VWISMGGEE VFISMGQEE VFISMGQEE VFISMGQEE VWSMGEAKQEDE TWISMGEE VWSMGEAKQEE LWASMGQEAE TWYSMGEPLE LUTDYSPOLQ LLTDYSPOLQ LLTDYSPVLQ LLTDYSPVLQ	K.KKRLQAFQ K.KKRLQAFQ K.KKRLDAFK R.LQKLKQFE K.EQRLKQFT K.QAKLKQFK Z42 KWAERIALLR TWEBRIKLLL V.EERIALLK V.EERIALLK S.LKRMKAFE A.ARIAMFR E.QERLAVFA A.EERRKIFR G.LDRVKPFR AWGERIKELL VWQERIDLLK LYPERVKLLT KPKF KPKF KPKF	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNAELLA PYQVNAELLA FPQLNSKLLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_LECOLI OTC_HAEIN OTCP_PSESH OTC_ASPTE OTCA_PSESH OTC_PSEAB OTCC_PSEAB OTCC_NEIGO OTCC_CLOPE OTC_MUMAN OTC_MOUSE OTC_RAT OTC_RANCA	FEMHLQAATP FEMHLHATP FEMHLHIATP SGIDVSIAVP VGHHTSVAKP FGISVSISTP 179 TGLDLKLVAP LEMDVRICGP FGYNLKIIAP MGVDLAVATP AGTHVIVAAP LEMDVVATP AGTHVIVAAP LEMDVRIGAP MCMHFVALGP LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE	KGYEPDENIV KGYEPDENIV KGYEPNSSLT SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACMP. KALLP. NALHP. KGYEIPASMR BGYEPEKAEFV KGYEVLDEAA EGYEPDEKVI EGFELPDESVR KALMP. QSLWP. QSLWP. DSLKP.	KLADOYAKEN KLAEQYAKEN KLAEQYAKEN SGAKEISAEN SIVNEAANEN 194 - EAALVPEC - EANLVEMC - TDAVLAGI ELIQEAGKGV ALAGDR - LAGARRADOR	G T.K G T.K G T.K G T.K G STFE G ATPE 202 RALAQQMGN EKFAKESGAR A NPGK	LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDSLKA 212 ITLTEDVAKG ITVTEDIDKA LIQTNVPEEA LRVVTDPREA FELLHDPVKA VTVTADAHAA LITTTEDPKEA ITLTENAHEA IEFSSNVDEA SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR	ARGGINULITD ARGGINULITD ARGGINULITD INGANVIVID INGANVIVID INGANVIVID STNANILVTD 222 VEGADFIYTD VKGADFIYTD VKADAUIYTD VKADAUIYTD VKADAUIYTD VACAHUNSTD VKDADVIYTD VKDADVIYTD VKGVDFVHTD VKGVGFIHTD VKGVGVFHTD VKGVBFHTD VKGVBFHTD VKGVBFHTD VKGVBVHTD KWTIMAVMVS KWTIMAVMVS KWTIMAVMVS KWTIMAVMVS	TWISMQQEDE TWISMQQEDE TWISMQQEE TWISMGQEE TWISMGGEFA 232 VWYSMGEFA VWISMGEVS TWYSMGEVS TWYSMGEVS TWYSMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGEPVE VWYSMGEPVE VWYSMGEPVE VWYSMGEPVE LUTDYSPVLQ LUTDYSPVLQ LUTDYSPVLQ LUTDYSPVLQ LUTDYSPQLQ	K.KKRLQAFQ K.KKRLLAFK R.LQKLKQFE K.BORLKQFT K.BORLKQFT K.QAKLKQFK 242 KWABRIALLR TWGERIKLLL S.LKRMKAFE A.AARIAMFR G.LDRVKPFR A.BERRKIFR G.LDRVKPFR LVWQERIDLLK LYPERVKLLT KPKF KPKF KPKF RPTF	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNAELLA PYQVNAELLA FPQLNSKLLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_HAEIN OTC_PSESH OTC_ASPTE OTCA_PSEAE OTC2_BACSU OTC_PYRFU OTCA_MYCBO OTCC_PSEAD OTCC_CLOPE OTC_MUMAN OTC_MOUSE OTC_MANCA OTC_RANCA	FEMHLQAATP FEMHLUAATP FEMHLHIATP SGIDVSIAVP VGIHTSVAKP FEGISVSISTP 179 TGLDLRLVAP LCMDVRICGP FEGYNLRIIAP MGVDLAVATP FDFOLRVACP MGCDISIASP LGADVVVATP AGTHVTVAAP LCMDVRIGAP MGMHFVALGP LCHDVRIAAP	KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP KALLP NALHP KGYEIPASMR EGYEPPKAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDESVR KALWP DSLKP DSLKP	KLADOYAKEN KLAEOYAKEN KLAEOYAKEN SGAKEISAEN SIVMEANNEN 194 . EAALVTEC . EAALVTEC . EAALVTEC . EAALVAGI ELIQEAGKGV ALAGOR . TDAVLAGI ELIQEAGKGV ALAGOR . EAAKTYALQS KWAEONAAES AAAERRAQDT . HDEFVAQC . SEGIIAAA . DEDILKEM	G T.K G T.K G T.K G T.K G ATFE 202 RALAQQNGGN EKFAKESGAR YEQTPERNGS A NPGK G S.S G S.S G S.S G S.S G S WKFAESGAK QEYSKETGAT VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYSPE VHDEVFYCPK VHDEVFYCPK	LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITHDSLKA 212 ITLTEDVAKG ITVTEDIDKA IEIFTEVAKG LIQTNVPEFA LRVVRDPREA VTLTDDPIEA FELLHDPVKA VTTADAHAA LTLTEDPKEA ITTTEDPKEA ITTTENAHEA IEFSSNVDEA SLVFPEAENR SLVFPEAENR RSLVFPEAENR RSLVFEEGEN	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKGVDFHTTD VKGADLIVTD VKGADLIVTD VKGADLIVTD VKGADLIVTD VKGVDFVHTD VKGVDFVHTD VKGVDFVHTD VKGVDFHTTD VKGVDFVHTD VKGVDFHTTD VKGVBFHTTD VKGVBFH	TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWVSMGEPKE VWISMGEVS TWVSMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGGEE VWASMGGEE LUTTYSFQLQ LUTTYSFVLQ LUTTYSFVLQ LUTTYSFVLQ LUTTYSFVLL GFVVNKGKLL	K.KKRLQAFQ K.KKRLQAFQ K.KKRLLDAFQ K.KKRLLDAFQ K.BORLKQFT K.BORLKQFT K.BORLKQFT K.BORLKQFT K.WAERIALLR TWGERIKLLL S.LKRMKAFE A.AARIAMFR E.QERLAVFA A.BERRKIFR G.LDRVKPFR AWGERIKELL LYPERVKLLT KPKF KPKF KPKF KPKF KPKF	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNAELLA PYQVNAELLA FPQLNSKLLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RAT OTC_RACTA OTC_SCHPO OTC_YEAST OTC1_ECOLI OTC_HAEIN OTCA_PSESH OTCA_SPTE OTCA_PSEAE OTC2_BACSU OTC_PYRFU OTCA_MYCBO OTCC_NEIGO OTCC_CLOPE OTC_HUMAN OTC_MOUSE OTC_RAT OTC_RANCA OTC_PACTA OTC_SCHPO	FEMHLQAATP FGMHLHIATP FGMHLHIATP FGMHLHIATP FGMHLHIATP FGISVSISTP 179 TGLDLRLVAP LGMDVRICGP FGYNLRIIAP MGYDLAVATP FDFQLRVACP MGCDISIASP LGADVVVATP AGHVTVAAP LGMDVRIGAP LGMDVRIGAP MGMHFVALGP LHCLPRKPE LHCLPRHPE MHCLPRHPE	KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVMVRDDIL PGIEMDSDIV 189 QACWP KALLP. NALHP KGYEIPASMR EGYEPEKAEFV KGYEVLDEAA KGYEVLDEAA KGYEVLDEAA KGYEVLDEAA CGYEPDEKVI EGFLPDPSVR KALMP QSLWP DSLKP	KLAEQYAKEN KLAEQYAKEN KLAEQYAKEN SGAKEISAEN SIVNEANNEN 194 . EAALVTEC . EAALV	G. T.K G. T.K G. T.K G. TTLK G. STFE G. ATFE 202 RALAQQNGGN EKFAKEGGAR YEQTPERNGS A. NFGK G. S.S G. S.S G. A.S KKFAEESGAK HAAAKETGAK VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYCER VDDEVFYCEE VSDEVFYGEN	LSMTNDPLEA LSMTNDPLEA L.MTNDPLEA ITNDPLEA IVNDPKVA L.THDSLKA 212 ITLTEDVAKG ITUTEDIDKA LIGHTVEDIDKA LIGHTVEDIDKA LIGHTVEDEA LRVVRDEREA LVTUTADAHAA LTUTEDPKEA ITLTEDPKEA ITLTEDAHEA SLVFPEAENR	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD YKNADIVVTD STNANILVTD VKGADFIYTD VKGADFIYTD VKGADVIYTD VKGAHLVSTD VKGADVIYTD AAAADVLVTD VKGADVIYTD VKGADVIYTD ** KWTIMAVMVS KWTIMAVMVS KWTIMAVMVS KWTIMAVMVS KWTIMAVWLEA	TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEE TWISMGGEE TYSMGEFA 232 VWSMGEAKE VWSMGEPLE VWISMGEVE VWISMGEVE VWISMGEVE VWISMGEVE VWISMGGEE VWISMGGEE VWISMGEVE LUTDYSPULQ LUTDYSPULD	K.KKRLQAFQ K.KKRLQAFQ K.KKRLDAFQ K.KKRLDAFQ K.DOLKQFE K.DOLKQFE K.DOLKQFE K.QAKLKQFK Z42 KWAERIALLR TWEERIKLLL V.EERIALLK TWEERIKLLL V.EERIALLK AARIAMFR AARIAMFR C.QERLAVFA AARIAMFR G.LDRVKPFR AWGERIKELL WQERIDLLK LYPERVKLLT KPKF KPKF KPKF RPTF PASA	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNAELLA PYQVNAELLA FPQLNSKLLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_HAEIN OTC_PSESH OTC_ASPTE OTC_PSEAE OTC_PYRFU OTC_PYRFU OTC_PYRFU OTC_CLOPE OTC_CLOPE OTC_HUMAN OTC_MOUSE OTC_RANCA OTC_PACTA OTC_RANCA OTC_SCHPO OTC_YEAST	FEMHLQAATP FEMHLQAATP FEMHLHIATP SGIDVSIAVP VGIHTSVAKP FOISVSISTP 179 TOLDLKLVAP LGMDVRICOP FOYNLKIIAP MGVDLAVATP MGVDLAVATP AGIHVIVAAP LGMDVRIAAP LGMDVRIAAP LGMDVRIAAP LCMDVRIAAP LCMCLPRKPE LHCLPRKPE LHCLPRKPE MHCLPRHPE MHCLPRHPE MHCLPRHPE	KGYEPDENIV KGYEPDENIV KGYEPNSSLT SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACUP KALLP NALHP NALHP KGYEIPASMR EGYEPKAEFV EGYEPKAEFV GSYEULDEAA EGYEPDESVR KALWP QSLWP QSLWP QSLWP DSLKP	KLADOYAKEN KLAEQYAKEN KLAEQYAKEN SAAKEISAEN SIVNEAANEN 194 - EAALVTEC - EANLVEMC - TDAVLAGI ELIQEAGKGV ALAGDR EAAKTYALQS KWAEQNAAES AAAERRAQDT - HDEFVAQC - SEGIIAAA - DEDILKEM - E - E - E - E - E - E - E - E - E - E	G T.K G T.K G T.K G T.K G ATFE 202 RALAQQNGGN EKFAKESGAR YEQTPERNGS A NPGK G S.S G S.S G S.S G S.S G S WKFAESGAK QEYSKETGAT VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYSPE VHDEVFYCPK VHDEVFYCPK	LSMTNDPLEA LSMTNDPLEA L.MTNDPLEA ITNDPLEA IVNDPKVA L.THDSLKA 212 ITLTEDVAKG ITUTEDIDKA LIGHTVEDIDKA LIGHTVEDIDKA LIGHTVEDEA LRVVRDEREA LVTUTADAHAA LTUTEDPKEA ITLTEDPKEA ITLTEDAHEA SLVFPEAENR	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD YKNADIVVTD STNANILVTD VKGADFIYTD VKGADFIYTD VKGADVIYTD VKGAHLVSTD VKGADVIYTD AAAADVLVTD VKGADVIYTD VKGADVIYTD ** KWTIMAVMVS KWTIMAVMVS KWTIMAVMVS KWTIMAVMVS KWTIMAVWLEA	TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEE TWISMGGEE TYSMGEFA 232 VWSMGEAKE VWSMGEPLE VWISMGEVE VWISMGEVE VWISMGEVE VWISMGEVE VWISMGGEE VWISMGGEE VWISMGEVE LUTDYSPULQ LUTDYSPULD	K.KKRLQAFQ K.KKRLQAFQ K.KKRLLDAFQ K.KKRLLDAFQ K.BORLKQFT K.BORLKQFT K.BORLKQFT K.BORLKQFT K.WAERIALLR TWGERIKLLL S.LKRMKAFE A.AARIAMFR E.QERLAVFA A.BERRKIFR G.LDRVKPFR AWGERIKELL LYPERVKLLT KPKF KPKF KPKF KPKF KPKF	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNAELLA PYQVNAELLA FPQLNSKLLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_HAELI OTC_PSESH OTC_ASPTE OTCA_PSEAE OTC_PSEACSU OTC_PYRFU OTCA_MYCEO OTCC_CLOPE OTC_CLOPE OTC_MUMAN OTC_MOUSE OTC_MOUSE OTC_RAT OTC_RANCA OTC_SCHPO OTC_SCHPO OTC_SCHPO OTC_YEAST	FEMHLQAATP FEGMHLHIATP FEGMHLHIATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLRLVAP LCMDVRLCGP FGYNLRI IAP MGVDLAVATP FDFQLRVACP MGCDISIASP LGADVVVATP AGIHVIVAAP LCMDVRIAAP LCMDVRIAAP LCMDVRIAAP LCMDVRIAAP LCMDVRIAAP LCHDLPKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE MHCLPRHPE MHCLPRHPE MHCLPRHPE MHCLPRHPE 271	KGYEPDENIV KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL 189 QACWP KALLP NALHP KGYEIPASMR EGYEPFAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDESVR KALWP QSLWP DSLKP DSLKP	KLADOYAKEN KLAEOYAKEN KLAEOYAKEN SGAKEISAEN SIVNEANNEN 194 - EAALVTEC - EAALVTEC - EAALVTEC - EAALVAGI ELIQEAGKGV ALAGDR - EAAKTYALQS KWAEQNAAES AAAERRAQDT - HDEFVAQC - SEGIIAAA - DEDILKEM - E - E - E - E - E - E - E - E - E - E	G T.K G T.K G T.K G T.K G ATFE 202 RALAQQNGSN EKFAKESGAR YEQTPERNGS A NPGK	LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDEKA ITHDSLKA 212 ITLTEDVAKG ITVTEDIDKA IEIFTEVAKG LIQTNVPEEA LRVVRDPREA VTVTJDDELEA FELLHDPVKA VTVTJDDAHAA LTLTEDPKEA ITTTENAHEA IEFSSNVDEA SLVFPEAENR	ARGGNVLITD ARGGNVLITD ARGGNVLITD ARGANVIVTD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKGADFIYTD VKGADILVTD VKGADILVTD VKGADILVTD VKDADVIYTD VKDADVIYTD VKGADVIYTD VKGVDFHTTD * KWTIMAVMVS	TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWVSMGEPKE VWISMGEVS TWVSMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGQEE LUTDYSPUQ LUTDYSPULQ LUTDYSPULD STVNNKGKLL LUNTGELLP FYNNKGNFKD 330	K.KKRLQAFQ K.KKRLQAFQ K.KKRLDAFQ K.KKRLDAFQ K.DOLKQFE K.DOLKQFE K.DOLKQFE K.QAKLKQFK Z42 KWAERIALLR TWEERIKLLL V.EERIALLK TWEERIKLLL V.EERIALLK AARIAMFR AARIAMFR C.QERLAVFA AARIAMFR G.LDRVKPFR AWGERIKELL WQERIDLLK LYPERVKLLT KPKF KPKF KPKF RPTF PASA	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNAELLA PYQVNAELLA FPQLNSKLLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_HAELI OTC_PSESH OTC_ASPTE OTCA_PSEAE OTC_PSEACSU OTC_PYRFU OTCA_MYCEO OTCC_CLOPE OTC_CLOPE OTC_MUMAN OTC_MOUSE OTC_MOUSE OTC_RAT OTC_RANCA OTC_SCHPO OTC_SCHPO OTC_SCHPO OTC_YEAST	FEMHLQAATP FEGMHLHIATP FEGMHLHIATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLRLVAP LCMDVRLCGP FGYNLRI IAP MGVDLAVATP FDFQLRVACP MGCDISIASP LGADVVVATP AGIHVIVAAP LCMDVRIAAP LCMDVRIAAP LCMDVRIAAP LCMDVRIAAP LCMDVRIAAP LCHDLPKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE MHCLPRHPE MHCLPRHPE MHCLPRHPE MHCLPRHPE 271	KGYEPDENIV KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL 189 QACWP KALLP NALHP KGYEIPASMR EGYEPFAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDESVR KALWP QSLWP DSLKP DSLKP	KLADOYAKEN KLAEOYAKEN KLAEOYAKEN SGAKEISAEN SIVNEANNEN 194 - EAALVTEC - EAALVTEC - EAALVTEC - EAALVAGI ELIQEAGKGV ALAGDR - EAAKTYALQS KWAEQNAAES AAAERRAQDT - HDEFVAQC - SEGIIAAA - DEDILKEM - E - E - E - E - E - E - E - E - E - E	G. T.K G. T.K G. T.K G. T.K G. ATFE G. ATFE 202 RALAQQNGN EKPAKESGAR YEQTPERNGS A. NPGK G. S.S G. S.S G. A.S C. S.S G. YEQUELYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYCPK VHDEVFYDEE VSDEVFYCEH VSDEVFYGEH	LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITHDSLKA 212 ITLTEDVAKG ITVTEDIDKA IEIFTEVAKG LIQTNVPEEA LRVVRDPREA VTVTJDDBLEA FELLHDPVKA VTVTJDDAHAA LTLTEDPKEA ITTTENAHEA IEFSSNVDEA SLVFPEAENR	ARGGNVLITD ARGGNVLITD ARGGNVLITD ARGANVIVTD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKGADFIYTD VKGADILVTD VKGADILVTD VKGADILVTD VKDADVIYTD VKDADVIYTD VKGADVIYTD VKGVDFHTTD * KWTIMAVMVS	TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWVSMGEPKE VWISMGEVS TWVSMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGQEE LUTDYSPUQ LUTDYSPULQ LUTDYSPULD STVNNKGKLL LUNTGELLP FYNNKGNFKD 330	K.KKRLQAFQ K.KKRLQAFQ K.KKRLDAFQ K.KKRLDAFQ K.DOLKQFE K.DOLKQFE K.DOLKQFE K.QAKLKQFK Z42 KWAERIALLR TWEERIKLLL V.EERIALLK TWEERIKLLL V.EERIALLK AARIAMFR AARIAMFR C.QERLAVFA AARIAMFR G.LDRVKPFR AWGERIKELL WQERIDLLK LYPERVKLLT KPKF KPKF KPKF RPTF PASA	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNAELLA PYQVNAELLA FPQLNSKLLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_HABEIN OTC_PSESH OTC_ASPTE OTCA_PSEAE OTC2_PACSU OTC_PYRFU OTCA_MYCBO OTCC_CLOPE OTC_MOUSE OTC_RANCA OTC_RANCA OTC_PACTA OTC_RANCA OTC_SCHPO OTC_SCHPO OTC_YEAST	FEMHLQAATP FEMHLUQAATP FEMHLHIATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLRLVAP LCMDVRICGP FGYNLRIIAP MGVDLAVATP FDFQLRVACP MGCDISIASP LCADDVVATP AGIHVIVAAP LCMDVRIGAP MGMHFVALGP LCHDVRIAP LCHDVRIAP LCHDVRIAP LCHDVRIAP LCHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE MHCLPRHPE MHCLPRHPE MHCLPRHPE MHCLPRHQE Z71 LHCLPAFHDD	KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP KALLP NALHP KGYEIPASMR EGYEPEKAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDPSVR KALMP QSLWP QSLWP QSLWP 281 QTILGKKMAE	KLAEQYAKEN KLAEQYAKEN KLAEQYAKEN SIVMEANEN 194 . EAALYTEC . SEGIIAAA . DEDILKEM . E	G. T.K G. T.K G. T.K G. TT.K G. ATFE G. ATFE G. ATFE G. AFFE G. S.S G. S.S G. A.S KKFAEESGAK QEYSKETGAT VDDEVFYSPR	LSMTNDPLEA LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPKVA LTHDSLKA 212 ITLTEDVAKG ITUTEDIDKA IEIFTEVAAG LIQTNVPEEA LRVVRDPREA VTVLTDDPIEA FELLHDPVKA VTVTADAHAA LTLTEDPKEA ITTLTENKEA ITTLTENKEA IEFSSNVDEA SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFQEAENR SLVFQEAENR SIVFQEAENR SIVFQEAENR SIVFQEAENR SIVFQEAENR SIVFQEAENR SIVFQEAENR SIVFQEAENR	ARGGNVLITD ARGGNVLITD ARGGNVLITD ARGANVLVTD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKGOPIHTD VKGADVIYTD VKRADLIVTD VKRADLIVTD VKADADVIYTD AAAADVLVTD VKGADVIYTD KWGYGFIHTD VKGYGFIHTD VKGYGFI	TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEE TWISMGGEE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWVSMGEPLE VWISMGGEE VWASMGGEE VWASMGGEE VWASMGGEE VWASMGGEE VWASMGGEE LUTTYSFQLQ LUTTYSFQLL FYNNKGKLL LLVNRGEILP FYNNKGKLL LLVNRGEILP FYNNKGKFLB 330 TLSK~	K.KKRLQAFQ K.KKRLLAFQ K.KKRLLAFQ R.LQKLKQFE K.BQRLKQFT K.BQRLKQFT K.QAKLKQFK 242 KWAERIALLR TMGERIKLLL S.LKRMKAFE A.AARIAMFR A.ARIAMFR G.LDRVKPFR AWGERIKELL UWQERIDLLK LYPERVKLLT KPKF KPKF KPKF RPKF RPKF RPKF RPKF RPKF	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNAELLA PYQVNAELLA FPQLNSKLLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_PART OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_HAEIN OTC_PSESH OTC_ASPTE OTC_PSEAE OTC_PACTU OTC_PACTU OTC_PACTU OTC_PACTU OTC_PACTU OTC_CLOPE OTC_CLOPE OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_YEAST	FEMHLQAATP FEMHLQAATP FEMHLHATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLKLVAP LGMDVRICGP FGYNLRI IAP MGVDLAVATP FDFQLRVACP MGCDISIASP LGADVVVATP AGIHVTVAAP LCMDVRIGAP MGMHFVALGP LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE MHCLPRHPE MHCLPRHPE MHCLPRHPE MHCLPRHPE LHCLPRHPE	KGYEPDENIV KGYEPDENIV KGYEPNSSLT SGLKFEELIL KDVNVRDDIL 189 QACWP. KALLP. NALHP. KGYEIPASMR EGYEPKAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDPSVR KALLP CSLWP. CSLWP. DSLKP. 281 QYTILGKKMAE ETKVGRQIAE	KLADOYAKEN KLAEQYAKEN KLAEQYAKEN SAAKEISAEN SIVNEANEN 194 ERAALVTEC . EANLVEMC . TDAVLAGI ELICEACKGV ALAGDR	G T.K G T.K G T.K G T.K G T.K G ATFE G ATFE 202 RALAQQNGN EKPAKESQAR YEQTPERNGS A NPGK	LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDSLKA 212 ITLTEDVAKG ITUTEDVAKG ITUTEDVAKG LIQTNVPEEA LIQTNVPEEA VTUTDDPLEA FELLHDPVKA VTVTADAHAA IETSSNVDEA SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SIVFEEAENR 310 SIVFEAENR SIVFEAENR SIVFEEAENR SIVFEEAENR SIVFEAENR SIVFEAENR SIVFEAENR SIVFEAENR	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVLVTD INGANVLVTD STNANILVTD 222 VEGADFIYTD VKGVDFIHTD VKGADVIYTD VKRADILVTD VKGADVIYTD VKGADVIYTD VKGADVIYTD VKGADVLYTD VKGADVLYTD VKGADVLYTD VKGYDFHTD VKGYDFHTD VKGYDFHTD VKGYDFHTD VKGADVIYTD * KWTIMAVMVS MTYAAIAVLE LYAAMSAIDI 320 MHTIKAVMVA MHTIKAVMVA	TWISMGQEDE TWISMGQEDE TWISMGQEE TWISMGGEE TWISMGGEE 232 VWVSMGEPLE VWISMGEVE VWISMGEVE VWISMGEVE VWISMGEE LLITDYSPOLQ LLITDYSPOLQ LLITDYSPOLD LITTYSPOLD LIT	K.KKRLQAFQ K.KKRLLAFQ K.KKRLLAFQ R.LQKLKQFE K.BQRLKQFT K.BQRLKQFT K.QAKLKQFK 242 KWAERIALLR TMGERIKLLL S.LKRMKAFE A.AARIAMFR A.ARIAMFR G.LDRVKPFR AWGERIKELL UWQERIDLLK LYPERVKLLT KPKF KPKF KPKF RPKF RPKF RPKF RPKF RPKF	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNAELLA PYQVNAELLA FPQLNSKLLA FYQVNAELLA FYQVNAELMA DYKVTESLMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_LECOLI OTC_HAEIN OTC_PSESH OTC_ASPTE OTC_ASPTE OTC_PSEAE OTC_PSEAE OTC_PSEAE OTC_CLOPE OTC_HUMAN OTC_MOUSE OTC_SCHPO OTC_YEAST OTC_HAEIN OTC_PSESH	FEMHLQAATP FEMHLQAATP FEMHLHIATP SGIDVSIAVP VGIHTSVAKP FEGISVSISTP 179 TELDLKLVAP LEMDVRICGP FEGYNLRI IAP MGVDLAVATP FDFQLRVACP AGTHVTVAAP LEMDVVATP AGTHVTVAAP LEMDVRIAAP LEMDVRIAAP LEMDVRIAAP LHCLPRKPE MHCLPRHPG 271 LHCLPAFHDD MHCLPAFHDS MHCLPAFHDS MHCLPAFHDS	KGYEPDENIV KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACMP. KALLP. NALHP. KGYEIPASMR EGYEPEKAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDESVR KALMP. QSLMP. DSLKP. 281 QTILGKKMAE ETKVGRQIAE DTEVARE.	KLABOYAKEN KLABOYAKEN KLAEOYAKEN SGAKEISAEN SIVNEANNEN 194 - EAALVTEC - EANLVEMC - TDAVLAGI ELIQEAGKGV ALAGDR - LAGATRALOGE EAAKTYALQS KWAEQNAAES AAAERRAQDT - HDEFVAQC - EA - E - E - E - E - E - E - E - E - E - E	G. T.K G. T.K G. T.K G. T.K G. ATFE G. ATFE 202 RALAQQNGN EKFAKESGAR A. NPGK G. S.S G. S.S G. S.S G. A.S KKFAEESGAK HAAAKETGAK QEYSKETGAT VDDEVFYSPR VDDEVFSPR VDDEVFSPR VTDEVFSSPM VTDEVFSSPM VTDEVFSSPM VTESVFSSPM VTESVFSSPM VEDSVFSSPM	LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITDPLEA II.TDPLEA II.TDPLEA 212 ITLTEDVAKG ITVTEDIDKA LIQTNVPEEA LRVVTDPREA FELLHDPVKA VTVTADAHAA LTTLTEDPKEA ITTLTENAHEA IEFSSNVDEA SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SIVFEEAENR	ARGGNVLITD ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKGADFIYTD VKGADILVTD VKGADILVTD VKADADVIYTD VKADADVIYTD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKGVGFHTTD VKGVADAVIVS KWTIMAVMVS KWTIMAVMVS KWTIMAVMVS KWTIMAVMVS KWTIMAVMVS KWTIMAVMVS MTIKAVMVA MHTIKAVMVA MHTIKAVMVA	TWISMQQEDE TWISMQQEDE TWISMQGEE TWISMGGEE TWISMGGEE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWSMGEPKE VWISMGEVS TWVSMQEEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGCEE VWASMGCEE VWASMGCEE LUTUSPELQ	K.KKRLQAFQ K.KKRLLAFQ K.KKRLLAFQ R.LQKLKQFE K.BQRLKQFT K.BQRLKQFT K.QAKLKQFK 242 KWAERIALLR TMGERIKLLL S.LKRMKAFE A.AARIAMFR A.ARIAMFR G.LDRVKPFR AWGERIKELL UWQERIDLLK LYPERVKLLT KPKF KPKF KPKF RPKF RPKF RPKF RPKF RPKF	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNALLUS FFQVNKDLVK FFQVNKELLA DYKVTEELMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_IECOLI OTC_HAEIN OTC_PSESH OTCA_PSEAE OTCA_PYRFU OTCA_MYCBO OTCC_PSEAD OTC_CLOPE OTC_HUMAN OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_YEAST OTC_YEAST OTC_SCHPO OTC_YEAST OTC_SCHPO OTC_YEAST	FEMHLQAATP FEMHLQAATP FEMHLHIATP SGIDVSIAVP VGIHTSVAKP FEGISVSISTP 179 TGLDLRLVAP LCMDVRICGP FEGYNLRIIAP MGVDLAVATP FDFOLRVACP MGCDISIASP LGADVVVATP AGTHVTVAAP LCMDVRIGAP MGMHFVALGP LHCLPRKPE MHCLPRHPE MHCLPRHPE MHCLPRHPE 271 LHCLPAFHDL MHCLPAFHDL MHCLPAFHDL MHCLPAFHDL MHCLPAFHDL MHCLPAFHDL MHCLPAFHDL	KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP KALLP NALHP KGYEIPASMR EGYEPFKAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDPSVR KALWP DSLKP DSLKP 281 QTTLGKKMAE ETKVGRQIAE	KLADOYAKEN KLAEOYAKEN KLAEOYAKEN SIVMEANEN 194 . EAALVTEC . TDAVLAGI ELIQEAGKGV ALAGOR . LEAAKTYALQS KWAEONAAES AAAERRAODT . HDEFVAQC . SEGIIAAA . DEDILKEM . E . E . E . E . E . E . E . E . E . E	G. T.K G. T.K G. T.K G. T.K G. ATFE G. ATFE 202 RALAQQNGGN EKFAKESGAR YEQTPERNGS A. NFGK G. S.S G. S.S G. A.S KKFAESGAR VDEVFYSPR VDDEVFYSPR	LSMTNDPLEA LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITDPLEA ITDPLEA ITDPLEA ITDPLEA ITTEDVAKG ITTTEDIDAA LITTEDIDAA LITTEDIDAA LTTTEDPLEA FELLHDPVKA VTVTADAHAA LTTTEDPKEA ITTTENAHEA IEFSSNVDEA SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR 310 SIVFDQAENR NIAFEQAENR SIVFDQAENR	ARGGNVLITD ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKGVDFHTTD VKGADLIVTD VKRADILVTD VKRADILVTD VKADADVIYTD VKDADVIYTD AAAADVLVTD VKGVGFHTTD VKGVGF	TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWVSMGEPLE VWISMGEVE VWISMGEVE VWISMGEVE VWISMGEVE LUTDYSPQLQ LLTDYSPQLQ LTDYSPQLQ LTDYSPQLQ LTDYSPQLQ LTDYSPQLQ TSTWORKSLL LLWRGEILP FVNNKGNFKD 330 TLSK- SLA~ TVVP- FVVNKGKIA	K.KKRLQAFQ K.KKRLLQAFQ K.KKRLLAFQ R.LQKLKQFE K.BQRLKQFT K.BQRLKQFT K.QAKLKQFK 242 KWABRIALLR TWGERIKLLL TWGERIKLLL S.LKRMKAFE A.AARIAMFR G.LDRVKPFR A.AERRKIFR G.LDRVKPFR LVWQERIDLLK LYPERVKLLT KPKF KPKF KPKF RPTF PASA LK~~	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNALLUS FFQVNKDLVK FFQVNKELLA DYKVTEELMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_PSESH OTC_ASPTE OTC_PSESH OTC_PSEAE OTC_PSEAE OTC_PSEAE OTC_CLOPE OTC_HUMAN OTC_MOUSE OTC_RANCA OTC_PACTA OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_HOUSE OTC_YEAST	FEMHLQAATP FEMHLQAATP FEMHLHIATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLKLVAP LGMDVRICGP FGYNLRI IAP MGVDLAVATP FDFQLRVACP MGCDISIASP LGADVVVATP AGIHVTVAAP LCMDVRIGAP MCMHFVALGP LHCLPRKPE MHCLPRHPE	KGYEPDENIV KGYEPDENIV KGYEPDENIV KGYEPNSSLT SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACUP KALLP NALHP KGYEIPASMR EGYEPKAEFV EGYEPKAEFV GSYEVLDEAA EGYEPDEKVI EGFLPDFSVR KALWP QSLWP DSLKP 281 281 281 281 27TLGKKMAE ETKVGRQIAE	KLADOYAKEN KLAEQYAKEN KLAEQYAKEN EAAKQFEKEC SGAKEISAEN 194 EAAKUTEC EAAKUTEC EAAKUTEC EAAKUTEC EAAKUTEC EAAKUTEC EAAKUTEC EAAKUTEC EAAKTYALQS KWAEQNAAES AAAERRAQDT HDEFVAQC SEGIIAAA DEDILKEM EE E	G. T.K G. T.K G. T.K G. T.K G. T.K G. ATFE 202 RALAQQNGN EKPAKESGAR YEQTPERNGS A. NPGK G. S.S G. S.S G. A.S VDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYCEN VSDEVFYGEN VSDEVFYGEN VSDEVFYGEN VSDEVFYGEN VSDEVFYGEN VSDEVFYSPR VEDSVFFGEN VSDEVFYSPR VEDSVFFGEN VSDEVFYSPR VEDSVFFGEN VSDEVFYSPR VEDSVFFGEN VSDEVFYSPR	LSMTNDPLEA LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA II.TDPLEA II.TDPLEA II.TDPLEA 212 ITLTEDVAKG LIQTNVPEEA LEVYTDPEA FELLHDPVKA VTVTADAHAA LITTENAHEA IETSSNVDEA SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SIVFQEAENR SIVFPEAENR	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKGVDFIHTD VKGVDFIHTD VKGADVIYTD VKADAUTYTD VKADAUTYTD VKADAUTYTD VKGADVIYTD VKGADVIYTD VKGADVIYTD VKGVFHTD VKGVFHTD VKGVFHTD VKGVFHTD VKGVFHTD VKGADVIYTD	TWISMGQEDE TWISMGQEDE TWISMGQEDE TWISMGQEE TWISMGGEAE TWISMGEFA 232 VWVSMGEPAE VWVSMGEPAE VWSMGQEE VWSMGQEE VWSMGQEE VWSMGEPAE VWSMGEPAE VWSMGEPAE LUTDYSPQLQ LLTDYSPQLQ LTDYSPQLQ LTDYSPQLC TVVNKGKIA LTBHAHYA	K.KKRLQAFQ K.KKRLLQAFQ K.KKRLLDFK R.LQKLKQFE K.BQRLKQFT K.BQRLKQFT K.QAKLKQFK 242 KWABRIALLR TWGERIKLLL VWGERIKLLL S.LKRMKAFE A.AARIAMFR G.LDRVKPFR A.BERRKIFR G.LDRVKPFR G.LDRVKPFR KPKF KPKF KPKF KPKF RPTF PASA LK	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNALLUS FFQVNKDLVK FFQVNKELLA DYKVTEELMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_LECOLI OTC_HAEIN OTC_PSESH OTC_ASPTE OTCA_PSEAE OTC_PSEAE OTC_NEIGO OTC_CLOPE OTC_HAEIN OTC_PACTA OTC_SCHPO OTC_SCHPO OTC_SCHPO OTC_YEAST	FEMHLQAATP FEMHLQAATP FEMHLHIATP SGIDVSIAVP VGIHTSVAKP FEGISVSISTP 179 TELDIKLVAP LEMDVRICEP FGYNLRI IAP MGVDLAVATP FDFQLRVACP AGTHVIVAAP LEMDVRIAAP LEMDVRIAAP LEMDVRIAAP LEMDVRIAAP LEMDVRIAAP LEMDVRIAAP LHCLPRKPE MHCLPAHPD MHCLPAFHD MHCLPAFHD MHCLPAFHD MHCLPAFHD MHCLPAFHD MHCLPAFHD MHCLPAFHD MHCLPAFHD MHCLPAHRGE LHCLPAHRGE LHCLPAHRGE LHCLPAHRGE	KGYEPDENIV KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACMP. KALLP. NALHP. KGYEIPASMR EGYEPEKAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDESVR KALMP. QSLMP. DSLKP. 281 QTTLGKKMAE ETKVGRQIAE DTEVARE.	KLABOYAKEN KLABOYAKEN KLAEOYAKEN SAAKEISAEN SIVNEANNEN 194 - EAALVPEC - EANLVEMC - TDAVLAGI ELIQEAGKGV ALAGDR EAAKTAIQS KWAEQNAAES AAAERRAQDT - HDEFVAQC - EA E E E E E E E E E E E E E E E E E E	G. T.K G. T.K G. T.K G. T.K G. T.K G. ATFE G. ATFE 202 RALAQQNGN EKFAKESGAR A. NPGK G. S.S G. S.S G. A.S KKFAEESGAK HAAAKETGAK QEYSKETGAT VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDEVFYSPR VSDEVFYSPR VSDEVFSPR VSDEVFYSPR VSDEVFYSPR VSDEVFYSPR VSDEVFSSPR VSDEVFYSPR VSDEVFYSPR VSDEVFYSPR VSDEVFSSPR VSDEVFYSPR VSDEVFYSPR VSDEVFYSPR VSDEVFYSPR VSDEVFYSPR VSDEVFSSPR VSDEVFYSPR VSDEVFSSPR VSDEVFYSPR VSDEVFSSPR	LSMTNDPLEA LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA LITUTEDVAKG LITUTEDVAKG LITUTEDVAKG LITUTEDVAKG LITUTEDVAKG LITUTEDLEA FELLHDPVKA VITUTDDPLEA FELLHDPVKA VITUTDDPLEA FELLHDPVKA VITUTDDAEN SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SIVFEAENR SIVFPEAENR SLVPOGENR S	ARGGNVLITD ARGGNVLITD ARGGNVLITD ARGGNVLITD INCANVIVID INCANVIVID INCANVIVID INCANVIVID INCANVIVID INCANVIVID INCANVIVID INCANVIVID INCANVIVID VKGADFIHTD VKGADILYTD VKADADVIYTD VKADADVIYTD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKGVGFHTD VKGVGFHTD VKGVGFHTD VKGVGFHTD INCANVIVID IN	TWISMQQEDE TWISMQQEDE TWISMQQEE TWISMGQEE TWISMGGEFA 232 VWVSMGERKE VWSMGEPKE VWISMGEVS TWVSMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGQEE LUTDYSPULQ LUTDYSPULD LUTDYSPU	K.KKRLQAFQ K.KKRLLAFQ K.KKRLLAFQ R.LQKLKQFE K.BQRLKQFT K.BQRLKQFT K.QAKLKQFK 242 KWABRIALLR TWGERIKLLL V.EERIALLK S.LKRMKAFE A.AARIAMFR G.LDRVKPFR AA.BERRKIFR G.LDRVKPFR AWGERIBLL LYWGERIBLL VWGERIBLL KFKF KFKF KFKF KPKF RPTF PASA LK C	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNALLUS FFQVNKDLVK FFQVNKELLA DYKVTEELMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_HAEIN OTC_PSESH OTC_ASPTE OTCA_PSEAE OTCA_PYRFU OTCA_MYCBO OTCC_CLOPE OTC_HUMAN OTC_MOUSE OTC_TRANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_SCHPO OTC_YEAST OTC_HAEIN OTC_ASPTE OTCA_RSPEEN OTCA_RSPEEN OTCA_PSEAE	FEMHLQAATP FEMHLQAATP FEMHLHIATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLRLVAP LCMDVRICGP FGYNLRIIAP MGVDLAVATP FDFOLRVACP FDFOLRVACP LGADVVATP AGTHVTVAAP LCMDVRIGAP LCMDVRIGAP MGCDISIASP MCGDISIASP MCGDISIASP MCGDIFF LCADVVATP AGTHVTVAAP LCMDVRIAAP LCMDVRIAAP LCMDVRIAP LCMDVRIAP LCMDVRIAP LCMDVRIAP LCMDVRIAP LCMDVRIAP LCMDVRIAP LCMDVRIAP LCMDVRIAP LCHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRHPE LHCLPAFHDI MHCLPAFHDI MHCLPAFHDI MHCLPAFHDI HCLPAHRGE MHCLPAHRGE	KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP KALLP NALHP KGYEIPASMR EGYEPPAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDESVR KALWP OSLKP DSLKP 281 QTTLGKKMAE ETKVGRQIAE	KLAEQYAKEN KLAEQYAKEN KLAEQYAKEN SIVMEANEN 194 - EAAKYEEN - 194 - EAALVTEC - EANLVEMC - EAALVTEC - EANLVEMC - EAALVTEC - EAALVTEC - EAAKYALGI ELIQEAGKGV ALAGDR - EAAKTYALQS KWAEQNASS AAAERRAQDT - HDEFVAQC - SEGIIAAA - DEDILKEM - E - E - E - E - E - E - E - E - E - E	G. T.K G. T.K G. T.K G. T.K G. T.K G. ATFE G. ATFE 202 RALAQQNGN EKFAKESGAR YEQTPERNGS A. NPGK G. S.S G. S.S G. S.S G. A.S KKFAESGAK QEYSKETGAT VDDEVFYSPR VSDDVFYGEN 300 VTDEVFSAA VTEEVFSPM VTDEVFSPM ISEELLDDPR ISEELLDDPR ISEELLDDPR ISEELLDDPR ISEELLDDPR ITABILIGPN VTDDVIDSPN	LSMTNDPLEA LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITDPLEA ITDPLEA ITDPLEA ITDPLEA ITDPLEA ITTPEDLORA LITTEDURA LITTEDURA LITTEDURA LTTTEDHEA TITTENAHEA IEFSSNVDEA SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR 310 SIVFQAENR 310 SIVFQAENR SLVFPEAENR SVWDQAENR SVWDQAENR SVWDQAENR SVWDQAENR SVWDQAENR	ARGGNVLITD ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKGVDFHTTD VKGVDFHTTD VKGADLVSTD VKGADLVSTD VKGADLVSTD VKGADLVSTD VKGVDFHTTD VKGVDFHTD VKGVDF	TWISMGQEDE TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWVSMGEPKE VWSMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGGEE LLTDYSPQLQ LLTDYSPQLQ LLTDYSPVLQ LTDYSPVLQ LTDYSPVLQ TLTMNGGEE FVVNKGKLL LLWNGGEE SLA~ TVVNKGKIA LTHAHYA LTHAHYA LTHAHYA LTHAHYA LTHAHYA LTHAFKESKN VMGGIKF	K.KKRLQAFQ K.KKRLLQAFQ K.KKRLLDFK R.LQKLKQFE K.BQRLKQFT K.BQRLKQFT K.QAKLKQFK 242 KWABRIALLR TWGERIKLLL VWGERIKLLL S.LKRMKAFE A.AARIAMFR G.LDRVKPFR A.BERRKIFR G.LDRVKPFR G.LDRVKPFR KPKF KPKF KPKF KPKF RPTF PASA LK	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNALLUS FFQVNKDLVK FFQVNKELLA DYKVTEELMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_PACTA OTC_SCHPO OTC_SCHPO OTC_SCHPO OTC_HAEIN OTC_PSESH OTC_ASPTE OTC_PSEAE OTC_PSEAE OTC_PSEAE OTC_CLOPE OTC_MOUSE OTC_MOUSE OTC_PACTA OTC_SCHPO OTC_YEAST OTC_SCHPO OTC_PASET OTC_PASET OTC_PASET	FEMHLQAATP FEMHLQAATP FEMHLHATP SGIDVSIAVP VGIHTSVAKP FGISVSISTP 179 TGLDLKLVAP LCMDVRICGP FGYNLRI IAP MGVDLAVATP FDFQLRVACP MGCDISIASP LCMDVVIATP AGHVTVAAP LCMDVRIAAP LCMDVRIAAP LCMDVRIAAP LCMDVRIAAP LCMDVRIAAP LCHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRHPE MGCLPRHPE MGCLPRHPE MGCLPRHPE MGCLPRHPE MGCLPRHPE MHCLPRHPE MHCLPAFHDD MHCLPAFHDE LHCLPAFHDE LHCLPAFHDE MHCLPAFHDE MHCLPAHRGE LHCLPAHRGE LHCLPAHRGE LHCLPAHRGE LHCLPAHRGE LHCLPAHRGE LHCLPAHRGE	KGYEPDENIV KGYEPDENIV KGYEPNSSLT SGLKFEELIL KDVNVRDDIL 189 QACWP. KALLP. NALHP. KGYEIPASME EGYEPKAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDESVR KALLP. OSLWP. DSLKP. 281 QTTLGKKMAE ETKVGRQIAE DTEVARE.	KLABQYAKEN KLABQYAKEN KLABQYAKEN SIVNEANEN 194	G. T.K G. TT.K G. TT.K G. TT.K G. TT.K G. ATFE G. ATFE 202 RALAQONGN REKPAKESGAR YEQTPERNGS A. NPGK T. S.S. G. A.S G. S.S G. A.S KKFAEESGAK HAAAKETGAK QEYSKETGAT VDDEVFYSPR VDDEVFYSPR VDDEVFYSPR VDDEVFYGEN VSDEVFYGEN VSDEVFYGEN VSDEVFYGEN VSDEVFYGEN VSDEVFYSR VEDSVFSGAR VTEEVFESPM VEDSVFSGAR VTEEVFSSAR VTEEVFSSR ISEELLDDPR VTABILIGPN UTDDVIDSPM ITDAVMDGPA	LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA II.TDPLEA II.TDPLEA II.TDPLEA II.TDPLEA III.TDPLEA III.	ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKWADIVTTD VKWADIVTTD VKWADIVTTD VKWADIVTTD VKWADIVTTD VKWADIVTTD VKWADIVTTD VKGADVIYTD AAAADVLVTD VKGDAPVIYTD VKGVGFIHTD VKGVGFIHTD VKGVGFIHTD VKGADVIYTD AWAINVLKWITMAVMVS KWTIMAVMVS KWTIMAVMVS KWTIMAVMVS KWTIMAVMVS KWTIMAVMVS RLYAAIAVLE LYAAMSAIDI 320 MHTIKAUMVA MHTIKALMLE LWAAISALEG LHAQKALLEL LHYOKALLKA LHAQKALLKA LHAQKAVLW	TWISMGQEDE TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEE TWISMGEFA 232 VWVSMGEPLE VWISMGEE VWISMGEP VWISMGEP VWISMGEP VWISMGEP LLITDYSPQLL LLTDYSPQLL LLTDYSPQLL LLTDYSPQLL LLTDYSPQLL LLTDYSPQLL LLTDYSPQLL LLTMSE TVINKGNFKD 330 TLSK- SLA~ TVVP- TVVNKGKIA LITHAHYA ILYKGESKN VMGGIKF LLERS	K.KKRLQAFQ K.KKRLQAFQ K.KKRLDAFQ K.KKRLDAFQ K.DOLKQFE K.DOLKQFE K.DOLKQFE K.QAKLKQFK Z42 KWAERIALLR TWGERIKLLL V.ERIALLL S.LKRMKAFE A.AARIAMFR E.QERLAVFA A.ERRRIFR G.LDRVKPFR AWGERIBLL KYPERVKLLT KPKF KPKF KPKF KPKF RPTF PASA LK	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNALLUS FFQVNKDLVK FFQVNKELLA DYKVTEELMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_SCHPO OTC_HAEIN OTC_PSESH OTC_ASPTE OTC_ASPTE OTC_PSEAE OTC_PSEAE OTC_PSEAE OTC_CLOPE OTC_HUMAN OTC_MOUSE OTC_PACTA OTC_PACTA OTC_SCHPO OTC_PACTA OTC_SCHPO OTC_PACTA OTC_SCHPO OTC_PASE OTC_PACTA OTC_PA	FEMHLQAATP FEMHLQAATP FEMHLHIATP SGIDVSIAVP VGIHTSVAKP FOISVSISTP 179 TGLDLKLVAP LEMDVRICOP FOYNLKI IAP MGVDLAVATP FDFQLRVACP MGCDISIASP LGADVVATP AGIHVTVAAP LEMDVRIAAP LEMDVRIAAP LEMDVRIAAP LEMDVRIAAP LHCLPRKPE LHCLPRHPE MHCLPRHPE MHCLPAHGE LHCLPAFHDL MHCLPAFHDL MHCLPAHRGE LHCLPAHRGE LHCLPAHRGE	KGYEPDENIV KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL RGIEMDSDIV 189 QACMP. KALLP. NALHP. NALHP. KGYEIPASMR EGYEPKAEFV KGYEVLDEAA EGYEPDESVR KALWP. QSLWP. QSLWP. QSLWP. QSLWP. QSLWP. QSTLGKRAEFV ETKVGKQIAE	KLADOYAKEN KLAEQYAKEN KLAEQYAKEN SAAKEISAEN SIVNEANNEN 194 - EAALVTEC - EANLVEMC - TDAVLAGI ELIQEAGKGV ALAGDR EAAKYALQS KWAEQNAAES AAAERRAQDT - HDEFVAQC - SEGIIAAA - DEDILKEM - E - E - E - E - E - F - F - F - F - F - F - F - F - F - F	G. T.K G. T.K G. T.K G. T.K G. T.K G. ATFE G. ATFE G. ATFE 202 RALAQQNGN EKFAKESGAR YEQTPERNGS A. NPGK G. S.S G. S.S G. A.S WORK WODEVFYSPR VDDEVFYSPR VTDEVTESPP VTBUTDEVFSPP VTBUTDEVTSPP VTBUTDEVTSPP	LSMTNDPLEA LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA II.TNDPLEA II.TNDPLEA II.TNDPLEA II.TNDPLEA II.TNDPLEA III.TEDVAKG ITVTEDUAKG ITVTEDUAKG ITVTEDUAKG LIQTNVPEEA FELLHDPVKA VTVTADAHAA IETSSNVDEA SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFPEAENR SLVFQEAENR SIVFEAENR SIVFPEAENR SIVFPEAENR SIVFPEAENR SIVFPEAENR SVAWDQAENR SVAWDQAENR SVAWDQAENR SAVWPQAENR NIAFEQAENR NIAFEQAENR	ARGGNVLITD ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKGVDFIHTD VKGADVIYTD VKAADILVTD VKAADILVTD VKAADILVTD VKAADILVTD VKAADILVTD VKAADILVTD VKAADILVTD AAAADVLVTD VKGADVIYTD KWTIMAVMVS MTIKALMLE LYAAMSAIDI 320 MHTIKAVMVA MHTIKALMLE LWAAI SALEG LHAQKALLEL LHAQKAVLAL LHAQKAVLAL LHAQKAVLAL LHAQKAVLAL LHAQKAVLAL LHAQKAVLAL LHAQKAVLAL LHAQKAVLIVM MHTIKAILVS	TWISMGQEDE TWISMGQEDE TWISMGQEDE TWISMGGEEF TWISMGGEEF 232 VWVSMGEPLE VWISMGEVE VWISMGEVE VWISMGEVE VWISMGEVE VWISMGEVE VWISMGEPLE VWISMGEPLE VWISMGEPLE VWISMGEPLE LUTTYSPULQ L	K.KKRLQAFQ K.KKRLLAFQ K.KKRLLAFQ R.LQKLKQFE K.BQRLKQFT K.BQRLKQFT K.QAKLKQFK 242 KWABRIALLR TWGERIKLLL V.EERIALLK S.LKRMKAFE A.AARIAMFR G.LDRVKPFR AA.BERRKIFR G.LDRVKPFR AWGERIBLL LYWGERIBLL VWGERIBLL KFKF KFKF KFKF KPKF RPTF PASA LK C	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNALLUS FFQVNKDLVK FFQVNKELLA DYKVTEELMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_LECOLI OTC_HAELI OTC_PSESH OTC_ASPTE OTCA_PSEAE OTCC_PSEAE OTCC_NEIGO OTCC_CLOPE OTC_HUMAN OTC_MOUSE OTC_RATO OTC_SCHPO OTC_SCHPO OTC_SCHPO OTC_SCHPO OTC_SCHPO OTC_SCHPO OTC_HAELI OTC_HAELI OTC_HAELI OTC_HAELI OTC_ASPTE OTC_ASPTE OTC_ASPTE OTC_ASPTE OTC_ASPTE OTC_SCHPO	FEMHLQAATP FEMHLQAATP FEMHLHIATP SGIDVSIAVP VGIHTSVAKP FEGISVSISTP 179 TGLDLRLVAP LCHDVRLCGP FEGYNLRI IAP MGVDLAVATP FDFQLRVACP FGYNLRI IAP MGVDLAVATP AGHVTVAAP LCHDVVATP AGHVTVAAP LCHDVVATP AGHHVALGP LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRKPE LHCLPRHPE MHCLPRHPE MHCLPRHPE MHCLPRHPE LHCLPRHPE LHCLPAHRGE LHCLPAHRGE LHCLPAHRGE LHCLPAHRGE LHCLPAHRGE LHCLPAHRGE LHCLPAHRGE	KGYEPDENIV KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP KALLP NALHP KGYEIPASMR EGYEPFAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDESVR KALWP QSLWP DSLKP 281 QTTIGKKMAE ETKVGRQIAE ETKVGRQIAE ETKVGRQIAE	KLAEQYAKEN KLAEQYAKEN KLAEQYAKEN SIVMEANEN 194 - EAAKYESKEC - EANLVEMC - TDAVLAGI ELIQEAGKGV ALAGDR - LIQEAGKGV ALAGDR - EAAKYALQS KWAEQNAAES AAAERRAQDT - HDEFVAQC - SEGIIAAA - DEDILKEM - E - E - E - E - E - E - E - E - E - E	G. T.K G. T.K G. T.K G. T.K G. T.K G. ATFE G. ATFE 202 RALAQQNGN EKFAKESGAR YEQTPERNGS A. NPGK G. S.S G. S. G. S. G. S. G. S. KKFAEESGAK QEYSKETGAT VDDEVFYSPR VSDEVFYSPR SSDEVFYSRA VTEEVFESPA VTDEVIESPA VTDDVIDSPN ISEELDDPR ISEELDDPR ISEELDDPR VTDDVIDSPN ISTEAVMGPA VTEEVFESPA	LSMTNDPLEA LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITDPLEA ITNDPLEA ITDPLEA ITDPLEA ITTEDUAK ITTEDUAK ITTEDUAK ITTEDUAK LTUTEDUAK ITTEDUAK ITTEDUAK ITTEDUAK ITTEDUAK ITTEDAHEA IEFSSNVDEA SLVFPEAENR SVWDQAENR SVWDQAENR SAWDQAENR SAWDGAENR NIAFEQAENR NIAFEQAENR NIAFEQAENR NIAFEQAENR SAWDQAENR SAWDQAENR SAWDQAENR SAWDQAENR SAWDQAENR SAWDQAENR	ARGGNVLITD ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKKADILVTD VKKADILVTD VKAGADILVTD VKADAVIYSD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKGVDFHTTD VKGVDFHTTD VKGVDFHTTD ** KWTIMAVMVS LYAAIAVLE LWAAISALEG LHAQKALLEL LHAQKALLEL LHAQKAVLAL LHAQKAVLAL LHAQKAVLVW MHTIKALIVS	TWISMGQEDE TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWVSMGEPKE VWISMGEVS TWVSMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGGEE VWASMGEDVE TWISMGEPKE LUTDYSPVLQ LLTDYSPVLQ LLTDYSPVLL LINTGEILP FVNNKGKLA LITHAHYA LITHAHYA LITHAHYA LLERS TLADI ALGD-	K.KKRLQAFQ K.KKRLQAFQ K.KKRLDAFQ K.KKRLDAFQ K.DOLKQFE K.DOLKQFE K.DOLKQFE K.QAKLKQFK Z42 KWAERIALLR TWGERIKLLL V.ERIALLL S.LKRMKAFE A.AARIAMFR E.QERLAVFA A.ERRRIFR G.LDRVKPFR AWGERIBLL KYPERVKLLT KPKF KPKF KPKF KPKF RPTF PASA LK	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNALLUS FFQVNKDLVK FFQVNKELLA DYKVTEELMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_SCHPO OTC_HAEIN OTC_PSESH OTC_ASPTE OTC_ASPTE OTC_PSEAE OTC_PSEAE OTC_PSEAE OTC_CLOPE OTC_HUMAN OTC_MOUSE OTC_PACTA OTC_PACTA OTC_SCHPO OTC_PACTA OTC_SCHPO OTC_PACTA OTC_SCHPO OTC_PASE OTC_PACTA OTC_PA	FEMHLQAATP FEMHLQAATP FEMHLHIATP SGIDVSIAVP VGIHTSVAKP FEGISVSISTP 179 TGLDLRLVAP LCHDVRLCGP FEGYNLRI IAP MGVDLAVATP FDFQLRVACP FGYNLRI IAP MGVDLAVATP AGHVTVAAP LCHDVVATP AGHVTVAAP LCHDVVATP AGHHVALGP LHCLPRKPE LHCLPRHPE MHCLPAHRGE LHCLPRHPE LHCLPRHPE MHCLPAHRGE LHCLPRHPE LHCLPRHPE LHCLPRHPE LHCLPRHPE MHCLPAHRGE LHCLPAHRGE	KGYEPDENIV KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP KALLP NALHP KGYEIPASMR EGYEPFAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDESVR KALWP QSLWP DSLKP 281 QTTIGKKMAE ETKVGRQIAE ETKVGRQIAE ETKVGRQIAE	KLAEQYAKEN KLAEQYAKEN KLAEQYAKEN SIVMEANEN 194 . EAAKYEEN 194 . EAALYTEC . EANLVENC . EANLVENC . EAALYTEC . EANLVENC . EAALYTEC . EA	G. T.K G. T.K G. T.K G. T.K G. T.K G. ATFE G. ATFE 202 RALAQQNGN EKFAKESGAR YEQTPERNGS A. NPGK G. S.S G. S. G. S. G. S. G. S. KKFAEESGAK QEYSKETGAT VDDEVFYSPR VSDEVFYSPR SSDEVFYSRA VTEEVFESPA VTDEVIESPA VTDDVIDSPN ISEELDDPR ISEELDDPR ISEELDDPR VTDDVIDSPN ISTEAVMGPA VTEEVFESPA	LSMTNDPLEA LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITDPLEA ITNDPLEA ITDPLEA ITDPLEA ITTEDUAK ITTEDUAK ITTEDUAK ITTEDUAK LTUTEDUAK ITTEDUAK ITTEDUAK ITTEDUAK ITTEDUAK ITTEDAHEA IEFSSNVDEA SLVFPEAENR SVWDQAENR SVWDQAENR SAWDQAENR SAWDGAENR NIAFEQAENR NIAFEQAENR NIAFEQAENR NIAFEQAENR SAWDQAENR SAWDQAENR SAWDQAENR SAWDQAENR SAWDQAENR SAWDQAENR	ARGGNVLITD ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKKADILVTD VKKADILVTD VKAGADILVTD VKADAVIYSD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKGVDFHTTD VKGVDFHTTD VKGVDFHTTD ** KWTIMAVMVS LYAAIAVLE LWAAISALEG LHAQKALLEL LHAQKALLEL LHAQKAVLAL LHAQKAVLAL LHAQKAVLVW MHTIKALIVS	TWISMGQEDE TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWVSMGEPKE VWISMGEVS TWVSMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGGEE VWASMGEDVE TWISMGEPKE LUTDYSPVLQ LLTDYSPVLQ LLTDYSPVLL LINTGEILP FVNNKGKLA LITHAHYA LITHAHYA LITHAHYA LLERS TLADI ALGD-	K.KKRLQAFQ K.KKRLQAFQ K.KKRLDAFQ K.KKRLDAFQ K.DOLKQFE K.DOLKQFE K.DOLKQFE K.QAKLKQFK Z42 KWAERIALLR TWGERIKLLL V.ERIALLL S.LKRMKAFE A.AARIAMFR E.QERLAVFA A.ERRRIFR G.LDRVKPFR AWGERIBLL KYPERVKLLT KPKF KPKF KPKF KPKF RPTF PASA LK	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNALLUS FFQVNKDLVK FFQVNKELLA DYKVTEELMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF
OTC_MOUSE OTC_RANCA OTC_PACTA OTC_SCHPO OTC_YEAST OTC_LECOLI OTC_HAELI OTC_PSESH OTC_ASPTE OTCA_PSEAE OTCC_PSEAE OTCC_NEIGO OTCC_CLOPE OTC_HUMAN OTC_MOUSE OTC_RATO OTC_SCHPO OTC_SCHPO OTC_SCHPO OTC_SCHPO OTC_SCHPO OTC_SCHPO OTC_HAELI OTC_HAELI OTC_HAELI OTC_HAELI OTC_ASPTE OTC_ASPTE OTC_ASPTE OTC_ASPTE OTC_ASPTE OTC_SCHPO	FEMHLQAATP FEMHLQAATP FEMHLHIATP SGIDVSIAVP VGIHTSVAKP FEGISVSISTP 179 TGLDLRLVAP LCHDVRLCGP FEGYNLRI IAP MGVDLAVATP FDFQLRVACP FGYNLRI IAP MGVDLAVATP AGHVTVAAP LCHDVVATP AGHVTVAAP LCHDVVATP AGHHVALGP LHCLPRKPE LHCLPRHPE MHCLPAHRGE LHCLPRHPE LHCLPRHPE MHCLPAHRGE LHCLPRHPE LHCLPRHPE LHCLPRHPE LHCLPRHPE MHCLPAHRGE LHCLPAHRGE	KGYEPDENIV KGYEPDENIV KGYEPDENIV KGYEPDENIV SGLKFEELIL KDVNVRDDIL PGIEMDSDIV 189 QACWP KALLP NALHP KGYEIPASMR EGYEPFAEFV KGYEVLDEAA EGYEPDEKVI EGFLPDESVR KALWP QSLWP DSLKP 281 QTTIGKKMAE ETKVGRQIAE ETKVGRQIAE ETKVGRQIAE	KLAEQYAKEN KLAEQYAKEN KLAEQYAKEN SIVMEANEN 194 - EAAKYESKEC - EANLVEMC - TDAVLAGI ELIQEAGKGV ALAGDR - LIQEAGKGV ALAGDR - EAAKYALQS KWAEQNAAES AAAERRAQDT - HDEFVAQC - SEGIIAAA - DEDILKEM - E - E - E - E - E - E - E - E - E - E	G. T.K G. T.K G. T.K G. T.K G. T.K G. ATFE G. ATFE 202 RALAQQNGN EKFAKESGAR YEQTPERNGS A. NPGK G. S.S G. S. G. S. G. S. G. S. KKFAEESGAK QEYSKETGAT VDDEVFYSPR VSDEVFYSPR SSDEVFYSRA VTEEVFESPA VTDEVIESPA VTDDVIDSPN ISEELDDPR ISEELDDPR ISEELDDPR VTDDVIDSPN ISTEAVMGPA VTDDVIDSPN ITDAVMGPA VTEEVFESPA	LSMTNDPLEA LSMTNDPLEA LSMTNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITNDPLEA ITDPLEA ITNDPLEA ITDPLEA ITDPLEA ITTEDUAK ITTEDUAK ITTEDUAK ITTEDUAK LTUTEDUAK ITTEDUAK ITTEDUAK ITTEDUAK ITTEDUAK ITTEDAHEA IEFSSNVDEA SLVFPEAENR SVWDQAENR SVWDQAENR SAWDQAENR SAWDGAENR NIAFEQAENR NIAFEQAENR NIAFEQAENR NIAFEQAENR SAWDQAENR SAWDQAENR SAWDQAENR SAWDQAENR SAWDQAENR SAWDQAENR	ARGGNVLITD ARGGNVLITD ARGGNVLITD ARGGNVLITD INGANVIVTD INGANVIVTD STNANILVTD 222 VEGADFIYTD VKKADILVTD VKKADILVTD VKAGADILVTD VKADAVIYSD VKDADVIYSD VKDADVIYSD VKDADVIYSD VKGVDFHTTD VKGVDFHTTD VKGVDFHTTD ** KWTIMAVMVS LYAAIAVLE LWAAISALEG LHAQKALLEL LHAQKALLEL LHAQKAVLAL LHAQKAVLAL LHAQKAVLVW MHTIKALIVS	TWISMGQEDE TWISMGQEDE TWISMGQEDE TWISMGGEE TWISMGGEE TWISMGEFA 232 VWVSMGEAKE VWVSMGEPKE VWISMGEVS TWVSMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGQEE VWASMGGEE VWASMGEDVE TWISMGEPKE LUTDYSPVLQ LLTDYSPVLQ LLTDYSPVLL LINTGEILP FVNNKGKLA LITHAHYA LITHAHYA LITHAHYA LLERS TLADI ALGD-	K.KKRLQAFQ K.KKRLQAFQ K.KKRLDAFQ K.KKRLDAFQ K.DOLKQFE K.DOLKQFE K.DOLKQFE K.QAKLKQFK Z42 KWAERIALLR TWGERIKLLL V.ERIALLL S.LKRMKAFE A.AARIAMFR E.QERLAVFA A.ERRRIFR G.LDRVKPFR AWGERIBLL KYPERVKLLT KPKF KPKF KPKF KPKF RPTF PASA LK	GYQVTMKTAK GYQVTMKTAK GYQUTMKTAK GFQITKEMIS GFQVITCEIMK GFQITCEIMK GFQITCEIMK GFQITSELAK PYQVTFEIMK GFQITSELAK PYQVNAALLUS PFQVNKDLVK PFQLNSKLLA PYQVNALLUS FFQVNKDLVK FFQVNKELLA DYKVTEELMA	V. AASDWTF V. AASDWTF L. AAPNWIF KGKAAENWKF LAK. PSCKF V. ADPNYKF 262 LTG. NPEVKF RTG. NPKVKF LTG. KADTIF RGGANENWKF G. AADDVLF H. AKPDYMF H. AKPDYMF L. ADSDAIV ATG. NPRAKF ASG. NPRVKF

FIGURE 2. Sequence alignment of selected OTCase sequences (four ureotelic OTCases, three yeast OTCases, eight anabolic bacterial OTCases, and three catabolic bacterial OTCases, separated by solid lines). For a more complete list, see ref 19. PALO binding residues are indicated by \bullet , and residues proposed to be catalytically important are indicated by ∇ . Sequence numbering is based on the sequence of E. coli OTCase. The organisms used in the alignment were Homo sapiens (human), Mus musculus (mouse), Rattus norvegicus (rat), Rana catesbeiana (bull frog), Pachysolen tannophilus, Schizosaccharomyces pombe, Saccharomyces cerevisiae, Escherichia coli, Haemophilus influenzae, Pseudomonas syringae, Aspergillus terrus, Bacillus subtilis, Pyrococcus furiosus, Mycobacterium bovis, Pseudomonas aeruginosa, Neisseria gonorrheae, and Clostridium perfringens. * indicates sequences not shown for clarity.

		3	13	22	21	200	40			=-
מסטם בירוני	λN7	DIVORBITA	TO	23 VLATAAKL	31	37	43	53	63	73
DVDD CALTY	MY	PLIQNITISI	MDFSKDDFMF	VLATAAKL	KANPQP	ELLKHK	VIASCFFEAS	TRIRLSFETS	MHRLGASVVG	FSDSANTSLG
DVDD CDDMY	AN AN	PLIQKHIISI	NDI CDDDI EI	VLATAAKL VLATAAGL	KANPQP	ELLKHK	VIASCFFEAS	TRIRLSFETS	MHRLGASVVG	FSDSANTSLG
PIND_SERIM	MIN-	PETUVUTTST	NDECKEDIE	VLATAERLER	KANPQP	ELLKHK	VIASCIFEAS	TRIRLSFETS	MHRLGASVVG	FADGSNTSLG
DVDD CTT CO	~~~~~II	MARITTER	KDLZKEDIEL	VLATAERLER IFALTDKYSK	ELKEKGQL	EYAKGK	TLATLFFEPS	TRTRLSFESA	MHRLGGAVIG	FAEA.STSSV
בותם_פסרפס	*MOXCODET V	~~~INTITION	INT SKUELED	IFDVAREMEK	NLNDTRKI	LSGK	TISIAFFEPS	TRTYLSFQKA	TINLGGDVIG	FSGE.ESTSV
PIND_ANAIN	"MQAGIRELK	VL ET2DATER	KOPDKEMESA	IFDVAREMEK	IEKSSSQS	ElLKGY	LMATLFYEPS			
DVDD DACCL	~~~~~~	~~~MTHLFAL	SELPLDEIMR	LLDEAERFRS	GRIWR	PAAPM	YVANLFFEPS		ERKLGLHVIP	
PURD_DACSU	*1 777777777	~~~MKHLTTM	SELSTEEIKD	LLQTAQELKS	GKTDN	QLTGK	FAANLFFEPS	TRTRFSFEVA	EKKLGMNVLN	LDGTSTSV
PIRD_LACLE	~ ARTEDYÖDIA	LLKLPYFVSV	EQLSADDVLH	LLQRAQYFKN	GGEVP	ALSRPI	FCTNMFFENS	TRTHTSFEVA	ERRLGLTVIP	FDPSHSSV
PIRD_LACPL	* AVEDI OF NO	MKTSQNLVSV	EQFSNQDVMA	YLKLAQAFKN	GKTV	QLSQPT	FAMILIFEINS	TRTHTSFEMA	ERRLGLQVIP	FDPKTSSV
DVDD DCDDI	*YKKEPÖTMD	OCOL BURL CT	DGLPRELLTE	ILDTADSFLE	. VGARAVKKV	PLLRGK	TVCNVFFENS	TRTRTTFELA	AQRLSADVIS	LNVSTSST
PIRD_PSEPU	"AKKPDQLIND	MADDIA TAN	ADI CEDENTA	ILDTADSFLE ILDDADRFAQ	.VGARAVKKV	PLLRGK	TVCNVFFENS	TRTRTTFELA	AQRLSADVIS	LNVSTSST
PADD GAMAS	MITMIT 77	~MIPKHLLIA	ADLSKDDATA	1LDDADRFAQ	ALVGRDIKKL	PTLRGR	TOVIMEYENS	TRTRVSFEVA	GKWMSADVIN	VSAAGSSV
FIND_SINIS	MEN	SMIKMUTEDE	TDWKGEELDI	VLQTATTFQQ LFDKTKRLKK	VLSGQ.TKKV	PALQGQ	VVTNLFFEPS	TRTRSSFELA	AKRLSADVMN	FSPGTSSL
FIND_INEDE	~~~~~~~IMEN	KINGKSLIVI	DDF2TDFKKI	LFDKTKKLKK	ATQEDDQKVM	DEFRINDEDF			AKFHQVKL	SDLAAESSSF
							•	•••		•
	0.3	0.3	100	110	110	100	400			
DIED DOOLT	83	93	102	112	118	122	132	142	152	155
PYRB_ECOLI	KKGETLADIT	SVISTY.VDA	IVMRHPQEGA	ARLATE	FSGN	VPVLNAGDGS	NQHPTQTLLD	LFTIQETQGR	LDN	LHVAMVGDLK
PYRB_SALTY	KKGETLADIT	SVISTY.VDA	IVMRHPQEGA	ARLATE	FSGQ	VPVLNAGDGS	NQHPTQTLLD	LFTIQETQGR	LDN	LHIAMVGDLK
PYRB_SERMA	KKGETLADTI	SVISTY.VDA	IVMRHPQEG.	ARMASE	FSGN	VPVLNAGDG.	NQHPTQTLLD	LFTIQETQGR	LSN	LSIAMVGDLK
PYRB_PYRAB	KKGESLRDI'I	KTVEQY.CDV	IVIRHPKEGA	ARLAAE	VA.E	VPVINAGDGS	NQHPTQTLLD	LYTIKKEFGR	IDG	LKIGLLGDLK
PYRB_SULSO	AKGENLADTI	RMLNNY.SDG	IVMRHKYDGA	SRFASE	IS.D	IPVINAGDGK	HEHPTQAVID	IYTINKHFNT	IDG	LVFALLGDLK
PYRB_ARATH	AKGETLEDTI	RTVEGY.SDI	IVMRHFESGA	ARKAAA	TA.N	IPVINAGDGP	GEHPTQALLD	VYTIQSEIGK	LDG	ISVALVGDLA
PYRB_BACCL	OKGETLYDTV	RTLEAIGVDA	VVIRHHEDAY	FEALR	HAVG	IPIINAGDGC	GHHPTQSLLD	LLTIRQEFG.	AFTG	LTVAIIGDIR
				YEELV						
PYRB_LACLE	NKGENLYDTE	LTMASLGIEL	SVIRHPENAY	YNEIIRPKEG	QHLQ	MGLVNAGDGS	GQHPSQSMLD	MMTIYNEFG.	HFDG	LKIMIVGDLT
PYRB_LACPL	TKGESLLDTL	KTIEAIGVNL	AVVRHPRDRY	YQPLLD	$\texttt{A}.\dots.\texttt{GFD}$	${\tt MSLINAGDGS}$	GQHPSQSLLD	MLTIYEEFG.	HFDG	LKIAIVGDLA
PYRB_PSEAE	SKGETLTDTL	RNLEAMAADM	FVVRHSDSGA	AHFIAEHV	SPN	VAVINGGDGR	HAHPTQGMLD	MLTIRRHKGN	FEQ	LSVAIVGDIL
PYRB_PSEPU	SKGETLFDTL	RNLEAMAADM	FVVRHSDSGA	AHFIAEHV	CPD	VAVINGGDGR	HAHPTQGMLD	MLTIRRHKGS	FEN	LSVAIVGDIL
PYRB_MYCTU	GKGESLRDTA	LTLRAAGADA	LIIRHPASGA	AHLLAQWTGA	HNDG	PAVINAGDGT	HEHPTQALLD	ALTIRQRLGG	IEG	RRIVIVGDIL
PYRB_SYNY3	TKGETILDTA	KTYLAMGSDI	FVIRHQQAGV	PHFIASOMDR	LOTG	VKVLNAGDGO	HEHPSOGLLD	LFTICSOFAP	DNPAIOCLOG	KKIAVVGDIL
PYRB_TREDE	NKGESYADTF	NTLAGYQNSI	FIVRSEVEGV	CRWLEDEAQA	FYQRNNLKRK	PAFINAGDGK	HEHPTQELLD	EFTFIED	NNWSFDK	IHIALVGDLY
	•		•			∇ ∇	 ∇ 			
	165	174	184	194	204	214	224	234	242	246
PYRB ECOLI		174 OALAKFDGNR	184 FYFIAPDALA	194 MPOYILDMLD	204 EKGIAWSLHS	214 SIEEVMAEVD	224 ILYMTRVOKE	234 RL., DPSEYA	242 NVKA	246 OFVLRASDLH
	YGRTVH.SLT	QALAKFDGNR	FYFIAPDALA	MPQYILDMLD	EKGIAWSLHS	SIEEVMAEVD	ILYMTRVQKE	RLDPSEYA	NVKA	QFVLRASDLH
PYRB_SALTY	YGRTVH.SLT YGRTVHFAKP	QALAKFDGNR RTLAKFSGNR	FYFIAPDALA FYFIAPDALA	MPQYILDMLD MPQYILDMLD	EKGIAWSLHS EKGMAWSLHG	SIEEVMAEVD SIEEVMADVD	ILYMTRVQKE ILYMTRVQKE	RLDPSEYA RLDPSEYA	NVKA NVKA	QFVLRASDLH QFVLR.PDLN
PYRB_SALTY PYRB_SERMA	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT	QALAKFIJGNR RTLAKFSGNR QALAKFEGNR	FYFIAPDALA FYFIAPDALA FYFIAPDALA	MPQYILDMLD MPQYILDMLD MPAYILKMLE	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG	SIEEVMAEVD SIEEVMADVD SIEEVVPELD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE	RLDPSEYA RLDPSEYA RLDPSEYA	NVKA NVKA	QFVLRASDLH QFVLR.PDLN QFVL.AADLA
PYRB_SALTY PYRB_SERMA PYRB_PYRAB	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA	QALAKFDGNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL	FYFIAPDALA FYFIAPDALA FYFIAPDALA .YLISPELLR	MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKGMKVVETT	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE	RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL	NVKA NVKA NVKA KVKG	QFVLRASDLH QFVLR.PDLN QFVL.AADLA SYQVNLKVLE
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YARTVN.SLL	QALAKFIGNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL	FYFIAPDALA FYFIAPDALA FYFIAPDALA .YLISPELLR VYLISPQLLR	MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL.	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKGMKVVETT NYPVKEVE	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE	RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE	NVKA NVKA NVKA KVKG	QFVLRASDLH QFVLR.PDLN QFVL.AADLA SYQVNLKVLE SYIVSLDLAN
PYRB_SAL/TY PYRB_SERMA PYRB_PYRAB PYRB_SUL/SO PYRB_ARATH	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YARTVN.SLL NGRTVR.SLA	QALAKFDGNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK	FYFIAPDALA FYFIAPDALA FYFIAPDALA .YLISPELLR VYLISPQLLR IYFVSPEIVK	MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKGMKVVETT NYPVKEVE SSGVEWEESS	SIEEVMAEVD SIEEVVPELD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE	RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE	NVKA NVKA NVKA KVKG KIKG	QFVLRASDLH QFVLR.PDLN QFVL.AADLA SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_ARATH PYRB_BACCL	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YARTVN.SLL NGRTVR.SLA HSR.VARSNA	QALAKFDGNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.N	FYFIAPDALA FYFIAPDALA FYFIAPDALA .YLISPELLR VYLISPQLLR IYFVSPEIVK VLFSGPEEWK	MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT DETNPYGT	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKGMKVVETTNYPVKEVE SSGVEWEESS YVE	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVYQTRIQRE	RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK	NVKA	QFVLRASDLH QFVLR.PDLN QFVL.AADLA SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_ARATH PYRB_BACCL PYRB_BACSU	YGRTVH. SLT YGRTVHFAKP YGRTVH. SLT YGRTVH. SLA YARTVN. SLL NGRTVR. SLA HSR. VARSNA HSR. VARSNA	QALAKFIXENR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.R	FYFIAPDALA FYFIAPDALA FYFIAPDALA .YLISPELLR VYLISPQLLR IYFVSPEIVK VLFSGPEEWK VLFSGPSEWQ	MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT DETNPYGT DEENTFGT	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS	SIEEVMAEVD SIEEVVPELD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQNE	RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ	NVKA	QFVLRASDLH QFVLR.PDLN QFVL.AADLA SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_ARATH PYRB_BACCL PYRB_BACSU PYRB_LACLE	YGRTVH. SLT YGRTVHFAKP YGRTVH. SLT YGRTVH. SLA YARTVN. SLL WGRTVR. SLA HSR. VARSNA HSR. VARSNA NSR. VARSNM	QALAKFIXINR RTLAKFSGNR QALAKFEGNR EALIFYDVEL RILITRFRPKL YLLAKFKDVK EVLITRLGA. N EVLITRLGA. R EILNTLGA. E	FYFIAPDALA FYFIAPDALA FYFIAPDALA .YLISPELLR VYLISPELLK IYFVSPEIVK VLFSGPEEWK VLFSGPSEWQ VYFSGPEYWY	MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL, MKDDIKDYLT D. ETNPYGT D. EENTFGT NAEEFSKYGT	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEIPELD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQNE VLMLLRVQHE	RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA	QFVLRASDLH QFVLR.PDLN QFVL.AADLA SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_ARATH PYRB_BACCL PYRB_BACSU PYRB_LACLE PYRB_LACPL	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YARTVN.SLL MGRTVR.SLA HSR.VARSNA HSR.VARSNA MSR.VARSNM HSR.VARSNM	QALAKFIXINR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA. N EVLTRLGA. R EILNTLGA. E ALLTQLGA. Q	FYFIAPDALA FYFIAPDALA FYFIAPDALA .YLISPELLR VYLISPELLR IYFVSPEIVK VLFSGPEEWK VLFSGPSEWQ VYFSGPEYWY VYFGGPKEWY	MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. ETNPYGT D. EENTFGT NAEEFSKYGT .GRDFEAYGE	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS Y.Q	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD LLMEVASKCD . VDEAIARAD . MDEAVESSD NIDDEIPELD TMDQLVATMD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQNE VIMLLRVQHE VMMLLRVQHE	RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ	QFVLRASDLH QFVLR.PDLN QFVL.AADLA SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_ARATH PYRB_BACCL PYRB_BACSL PYRB_LACLE PYRB_LACPL PYRB_PSEAE	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLA YGRTVH.SLA YARTVN.SLL MGRTVR.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM	QALAKFDSNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.R EILNTLGA.E ELLNTLGA.C LALKTLGCPD	FYFIAPDALA FYFIAPDALA FYFIAPDALA .YLISPELLR VYLISPELVK VLFSGPEEWK VLFSGPSEWQ VYFSGPEYWY VYFGGPKEWY IRVIAPATLL	MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. EENTFGT NAEEFSKYGT NAEEFSKYGT GRDFEAYGE PIGLEE	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVK YV, Q,	SIEEVMAEVD SIEEVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEIPELD TMDQLVATMD NADEGLKDVD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQHE VLMLLRVQHE VMMLLRVQHE VVIMLRQRE	RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL	NVKA NVKA NVKA KVKG AARG EEYHA EGYLN QLFDAKDYNA QLFDAKDYNA QTFDASAYHQ .LPSEGEFFK	QFVLRASDLH QFVLR.PDLN QFVL.AADLA SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_ARATH PYRB_BACCL PYRB_BACSU PYRB_LACPL PYRB_LACPL PYRB_PSEAE PYRB_PSEAE	YGRTVH. SLT YGRTVHFAKP YGRTVH. SLT YGRTVH. SLA YARTVN. SLL HSR. VARSNA HSR. VARSNA HSR. VARSNM HSR. VARSNM HSR. VARSNM HSR. VARSNM HSR. VARSNM HSR. VARSNM	QALAKFDSNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.R EILNTLGA.E ALLTQLGA.E LALKTLGC.PD LALKTLGC.PD	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR IYFVSPEIVK VLFSGPEEWK VLFSGPSEWQ VYFSGPEYWY VYFGGPKEWY IRVIAPATILL IRVIGPKTLI	MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEEL MKDDIKDYLT D. ETNPYGT D. EENTFGT NAEEFSKYGT GRDFEAYGE PIGLEE PIGI.E	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKGMKVVETV SSGVEWEESS YVE YVS YVS QYGVRVFTQYGVRVVTT	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEIPELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQNE VLMLLRVQHE VMMLLRVQHE VMMLLRVQHE VVIMLRVQRE VVIMLRLORE	RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVIN RMQGGL RMAGGL	NVKA NVKA NVKA KVKG AARG E EYHA E GYLN QLFDASAYNA QTFDASAYNA QTFDASAYD LPSEGEFFK LPSEGEFYR	QFVLRASDLH QFVLR.PDLN QFVL.AADLA SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAL LYGLTEKRLK LFGLTTARLA
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_ARATH PYRB_BACCL PYRB_LACPL PYRB_LACPL PYRB_LACPL PYRB_PSEAE PYRB_PSEPU PYRB_MYCTU	YGRTVH. SLT YGRTVHFAKP YGRTVH. SLT YGRTVH. SLA YARTVN. SLL HSR. VARSNA HSR. VARSNA HSR. VARSNM	QALAKFDSNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.R EILNTLGA.E ALLTQLGA.E LALKTLGC.PD LALKTLGC.PD MLLDTLGA.E	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR IYFVSPEIVK VLFSGPEEWK VLFSGPSEWQ VYFSGPEYWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL	MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEEL MKDDIKDYLT D. ETNPYGT D. EENTFGT NAEEFSKYGT GRDFEAYGE PIGLEE. PIGI.E PVGVT	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVS QYGVRVFT .QYGVKVYT .GWPATVSH	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEI PELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DFDAELPAAD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQNE VLMLLRVQHE VMMLLRVQHE VVMLLRVQRE VVIMLRLQRE VVIMLRLQRE AVLMLRVQAE	RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMAGGF	NVKA NVKA KVKG KIKG AARG EEYHA EGYLN QLFDAKDYNA QTFDASAYNA QTFDASAYD .LPSEGEFFK .LPSEGEFYR .FPSVREYSV	QFVLRASDLH QFVLR.PDLN QFVL.AADLA SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAL LYGLTERKLA LYGLTERRLA RYGLTERRQA
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACCL PYRB_LACPL PYRB_LACPL PYRB_PSEAC PYRB_PSEPU PYRB_MYCTU PYRB_SYNY3	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLA YGRTVH.SLA YARTVN.SLL MGRTVR.SLA HSR.VARSNA MSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM	QALAKFDGNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRFKL YLLAKFKDVK EVLTRLGA. R EILNTLGA. E ALLTQLGA. Q LALKTLGCPD LALKALGCPD MLLDTLGA. E WSLTTAGA. D	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR IYFVSPEIVK VLFSGPEEWK VLFSGPSEWQ VYFSGPEYWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VHLAGPPTLL	MPQYILDMLD MPQYILDMLD MPAYILKMLD MPAYILKMLE MRHIVEELR ARKEILDEL. MKDDIKDYLT D. EINTPYGT D. EENTFGT NAEEFSKYGT GRDFEAYGE PIGLEE PIGI.E PVGVT PKEFQQLTLA	EKGIAWSLHS EKGMAWSLHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVG QYGVRVFT .QYGVKVVFT .GWPATVSH PGSGKLHCHW	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEI PELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQNE VLMLLRIQNE VLMLLRVQHE VVMLLRVQHE VVIMLRLQRE VVIMLRLQRE VVIMLRVQAE IVMTLRLQKE	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMAGGL RMNXGF RMNXGF RMTAHL	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACCL PYRB_LACPL PYRB_LACPL PYRB_PSEAC PYRB_PSEPU PYRB_MYCTU PYRB_SYNY3	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLA YGRTVH.SLA YARTVN.SLL MGRTVR.SLA HSR.VARSNA MSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM	QALAKFDGNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRFKL YLLAKFKDVK EVLTRLGA. R EILNTLGA. E ALLTQLGA. Q LALKTLGCPD LALKALGCPD MLLDTLGA. E WSLTTAGA. D	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR IYFVSPEIVK VLFSGPEEWK VLFSGPSEWQ VYFSGPEYWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VHLAGPPTLL	MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEEL MKDDIKDYLT D. ETNPYGT D. EENTFGT NAEEFSKYGT GRDFEAYGE PIGLEE. PIGI.E PVGVT	EKGIAWSLHS EKGMAWSLHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVG QYGVRVFT .QYGVKVVFT .GWPATVSH PGSGKLHCHW	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEI PELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQNE VLMLLRIQNE VLMLLRVQHE VVMLLRVQHE VVIMLRLQRE VVIMLRLQRE VVIMLRVQAE IVMTLRLQKE	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMAGGL RMNXGF RMNXGF RMTAHL	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACCL PYRB_LACPL PYRB_LACPL PYRB_PSEAC PYRB_PSEPU PYRB_MYCTU PYRB_SYNY3	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLA YGRTVH.SLA YARTVN.SLL MGRTVR.SLA HSR.VARSNA MSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM	QALAKFDGNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRFKL YLLAKFKDVK EVLTRLGA. R EILNTLGA. E ALLTQLGA. Q LALKTLGCPD LALKALGCPD MLLDTLGA. E WSLTTAGA. D	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR IYFVSPEIVK VLFSGPEEWK VLFSGPSEWQ VYFSGPEYWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VHLAGPPTLL	MPQYILDMLD MPQYILDMLD MPAYILKMLD MPAYILKMLE MRHIVEELR ARKEILDEL. MKDDIKDYLT D. EINTPYGT D. EENTFGT NAEEFSKYGT GRDFEAYGE PIGLEE PIGI.E PVGVT PKEFQQLTLA	EKGIAWSLHS EKGMAWSLHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVG QYGVRVFT .QYGVKVVFT .GWPATVSH PGSGKLHCHW	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEI PELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQNE VLMLLRIQNE VLMLLRVQHE VVMLLRVQHE VVIMLRLQRE VVIMLRLQRE VVIMLRVQAE IVMTLRLQKE	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMAGGL RMNXGF RMNXGF RMTAHL	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACCL PYRB_LACPL PYRB_LACPL PYRB_PSEAC PYRB_PSEPU PYRB_MYCTU PYRB_SYNY3	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLA YGRTVH.SLA YARTVN.SLL MGRTVR.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM	QALAKFDSNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRFKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.E ELINTLGA.E ALLTQLGA.Q LALKTLGCPD LALKALGCPD MLLDTLGA.E WSLTTAGA.D DGLKIFKSVK	FYFIAPDALA FYFIAPDALA FYFIAPDALA .YLISPELLR VYLISPQLLR IYFVSPEIVK VLFSGPEEWK VLFSGPEEWK VYFSGPEYWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VHLAGPPTLL VDLIAPAELA	MPQYILDMLD MPQYILDMLD MPQYILMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. EENTFGT D. EENTFGT GRDFEAYGE PIGLEE. PIGII.E. PVGVT. PKEFQQLTLA MPEYYKVRMQ	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVK QYGVRVFT .QYGVRVFT .QYGVKVYT .GWPATVSH PGSGKLHCHW ENGFTVREFS	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEIPELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DLAEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD SIEEYLRQAD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQNE VLMLLRIQNE VLMLLRVQHE VVMLLRVQHE VVIMLRLQRE VVIMLRLQRE VVIMLRVQAE IVMTLRLQKE	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMAGGL RMNXGF RMNXGF RMTAHL	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_ARATH PYRB_BACCL PYRB_EACSU PYRB_LACPL PYRB_LACPL PYRB_PSEAE PYRB_PSEPU PYRB_MYCTU PYRB_SYNY3 PYRB_TREDE	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YGRTVH.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNV HSR.VARSNV HSR.VARSNL	QALAKFDSNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.R EILNTLGA.R EALLTQLGA.Q LALKTLGCPD LALKALGCPD MLLDTLGA.E WSLTTAGA.D DGLKIFKSVK	FYFIAPDALA FYFIAPDALA FYFIAPDALA .YLISPELLR VYLISPELLR IYFVSPEIVK VLFSGPEEWK VLFSGPSEWG VYFSGPEEWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VHLAGPPTLL VHLAGPPTLL VDLIAPAELA	MPQYILDMLD MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. EENTFGT D. EENTFGT NAEEFSKYGT NAEEFSKYGT GRDFEAYGE PIGLEE PVGVT PKEFQQLTLA MPEYYKVRMQ	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVK YVG QYGVRVFT .QYGVKVYT .GWPATVSH PGSGKLHCHW ENGFTVREFS	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEIPELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD SIEEYLRQAD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVMLLRIQHE VVMLLRIQHE VLMLLRVQHE VMMLLRVQHE VVIMLRLQRE AVLMLRVQAE IVMTLRLQRE VALIWFTRP	RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMAGGL RMNGGF RMTAHL QL. ERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_ARATH PYRB_BACCL PYRB_LACCL PYRB_LACPL PYRB_LACPL PYRB_PSEAE PYRB_PSEPU PYRB_SYNY3 PYRB_TREDE PYRB_ECOLI	YGRTVH. SLT YGRTVHFAKP YGRTVH. SLT YGRTVH. SLT YGRTVH. SLA YARTVN. SLL HSR. VARSNA HSR. VARSNA HSR. VARSNA HSR. VARSNM SR. VARSNW HSR. VARSNW HSR. VARSNW HSR. VARSNL HGRTVH. SKA 256 NAKANMKVLH	QALAKFDSNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.R EILNTLGA.E ALLTQLGA.Q LALKFLGCPD LALKALGCPD MLLDTLGA.E WSLTTAGA.D DGLKIFKSVK	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR IYFVSPEIVK VLFSGPEEWK VLFSGPSEWQ VYFSGPEYWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VHLAGPPTLL VDLIAPAELA 274 ATD. VDKTPH	MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. ETNPYGT D. EENTFGT NAEEFSKYGT GRDFEAYGE PIGLEE PIGI.E PVGVT PKEFQQLTLA MPEYYKVRMQ 283 AWYFQQAGNG	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVK Y.Q QYGVRVFT .QYGVKVYT .GWPATVSH PGSGKLHCHW ENGFTVREFS	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEIPELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD SIEEYLRQAD	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVMLLRIQHE VVMLLRIQHE VVMLLRVQHE VMMLRVQHE VVIMLRLQRE AVLMLRVQAE IVMTLRQKE VALIWYFTRP • •	RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVIN RMQGGL RMAGGL RMNGGF RMNGGF RMTAHL QLERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SYRAB PYRB_PYRAB PYRB_SULSO PYRB_ARATH PYRB_BACCL PYRB_LACLE PYRB_LACPL PYRB_PSEAE PYRB_PSEAE PYRB_SYNY3 PYRB_TREDE PYRB_TREDE PYRB_ECOLI PYRB_SALTY	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YGRTVH.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNW HSR.VARSNW HSR.VARSNW HSR.VARSNV HGRTVH.SKA • 256 NAKANMKVLH GARENMKVLH	QALAKFDGNR RTLAKF SGNR RTLAKF SGNR QALAKFEGNR EALTFYDVEL RILTRFRFKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.R EILNTLGA.E ALLTQLGA.Q LALKTLGCPD MLLDTLGA.D MLDTLGA.D DGLKIFKSVK 266 PLP.RVDEI PLP.RIDEI	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR IYFVSPEIVK VLFSGPSEWQ VYFSGPEWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VDLIAPAELA 274 ATD. VDKTPH TTD. VDKTPH	MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. EENTFGT NAEEFSKYGT .GRDFEAYGE PIGLEE PVGVT PVGVT PVGVT PKEFQQLTLA MPEYYKVRMQ 283 AWYFQQAGNG AWYFQQAGNG AWYFQQAGNG	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVG QYGVRVFT .QYGVKVYT .GWPATVSH PGSGKLHCHW ENGFTVREFS 293 IFARQALLAL IFAAQALLAL	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEI PELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD SIEEYLRQAD 303 VLNRDLVL~~ VLNSELSL~~	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQNE VVMLLRIQNE VMMLLRVQHE VVMLLRVQHE VVIMLRQRE VALIWYFTRP	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMAGGL RMNGGF RMNGGF RMTAHL QLERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACSL PYRB_LACLE PYRB_LACPL PYRB_PSEAE PYRB_PSEAE PYRB_MYCTU PYRB_SYNY3 PYRB_TREDE PYRB_ECOLI PYRB_SALTY PYRB_SALTY PYRB_SERMA	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YGRTVH.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNM HGRTVH.SKA	QALAKFDGNR RTLAKF SGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.E ALLTQLGA.Q LALKTLGCPD LALKALGCPD MLDTLGA.E DGLKIFKSVK 266 PLP.RVDEI PLP.RIDEI PLP.RIDEI	FYFIAPDALA FYFIAPDALA FYFIAPDALA .YLISPELLR VYLISPQLLR IYFVSPEIVK VLFSGPSEW VYFSGPEWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VHLAGPPTLL VHLAGPPTLL VDLIAPAELA 274 ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH	MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D.ENTTGT NAEEFSKYGT GRDFEAYGE PIGLEE PVGVT PVGVT PKEFQQLTLA MPEYYKVRMQ 283 AWYFQQAGNG AWYFQQAGNG AYYFQQAGNG	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVG QYGVRVFT .QYGVKVYT .GYGVKVYT .GYGVKVYT .GYGPATVSH ENGFTVREFS 293 1FARQALLAL IFAAQALLAL IFAAQALLAL IFARSA.LAL	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEI PELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD SIEEYLRQAD 303 VLNRDLVL~~ VLNSELSL~~ VNNADLAL~~	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQNE VMMLLRVQHE VVMLLRVQHE VVIMLRLQRE VVIMLRLQRE VVIMLRVQAE IVMTLRLQRE VALIWYFTRP • •	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLITK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMAGGL RMNGGF QLERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACCL PYRB_LACLE PYRB_LACLE PYRB_PSEAE PYRB_PSEAE PYRB_PSEPU PYRB_MYCTU PYRB_SYNY3 PYRB_TREDE PYRB_ECOLI PYRB_SALTY PYRB_SERMA PYRB_PYRAB	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YGRTVH.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNL HGRTVH.SKA 256 NAKANMKVLH GARENMKVLH GARENMKVLH KAKDELRIMH	QALAKFDSNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.E ELLNTLGA.E ALLTQLGA.Q LALKTLGCPD LALKALGCPD MLLDTLGA.E WSLTTAGA.D DGLKIFKSVK 266 PLP.RVDEI PLP.RIDEI PLP.RIDEI PLP.RVDEI	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR IYFVSPEIVK VLFSGPSEWK VLFSGPSEWK VLFSGPSEWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VHLAGPPTLL VHLAGPPTLL VDLIAPAFLA 274 ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH HPE. VDNTKH	MPQYILDMLD MPQYILDMLD MPQYILMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. EENTFGT TO. EENTFGT NAEEFSKYGT NAEEFSKYGT OGRDFEAYGE PIGLEE PVGVT PKEFQQLTLA MPEYYKVRMQ 283 AWYFQQAGNG AYYFQQAGNG AYYFQQAGNG AIYFRQVFNG	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHE EKKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVK QYGVRVFT .QYGVKVYT .GWPATVSH PGSGKLHCHW ENGFTVREFS 293 IFARQALLAL IFARQALLAL IFARSA.LAL VPVRMALLAL	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEIPELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD SIEEYLRQAD 303 VLNRDLVL~~ VLNSELSL~~ VUNSDLAL~~ VLGVI~~~~~	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVMLLRIQHE VVMLLRIQHE VVMLLRVQHE VMMLLRVQHE VVIMLRUQRE VVIMLRUQRE VVIMLRUQRE IVMTLRUQRE VALIWYFTRP • •	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMAGGL RMNGGF RMTAHL QLERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_ARATH PYRB_BACCL PYRB_BACCL PYRB_LACPL PYRB_PSEAE PYRB_PSEAE PYRB_PSEPU PYRB_SYNY3 PYRB_TREDE PYRB_ECOLI PYRB_SALTY PYRB_SALTY PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YGRTVH.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM GSR.VARSNM HSR.VARSNM GSR.VARSNW HSR.VARSNW HKRLYWH KAKDELRIMH KMKKDSILH	QALAKFDSNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.R EILNTLGA.C LALKTLGCPD MLLDTLGA.E WSLTTAGA.D DGLKIFKSVK 266 PLP.RVDEI PLP.RIDEI PLP.RVDEI PLP.RVDEI PLP.RVNEI	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR IYFVSPEIVK VLFSGPESWK VLFSGPESWK VLFSGPESWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VHLAGPPTLL VDLIAPAELA 274 ATD. VDKTPH TTD. VDKTPH HPE. VDNTKH DRK. VDKTTK	MPQYILDMLD MPQYILDMLD MPQYILTMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. EENTFGT D. EENTFGT NAEEFSKYGT NGRDFEAYGE PIGLEE PKGYT PKEFQQLTLA MPEYYKVRMQ 283 AWYFQQAGNG AWYFQQAGNG AYYFQQAGNG AYYFQQAGNG AKYFEQASYG	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVK YQGVRVFT .QYGVKVYT .GWPATVSH PGSGKLHCHW ENGFTVREFS 293 IFARQALLAL IFARSA.LAL VPVRMALLAL VPVRMSILTK	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEIPELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD SIEEYLRQAD 303 VLNRDLVL VINSELSL VVNADLAL VLGVI IYGE	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQHE VVMLLRVQHE VVMLLRVQHE VVIMLRQRE AVIMLRVQAE IVMTLRQKE VALIWYFTRP	RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMNGGF RMNGGF QLERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_ARATH PYRB_BACCL PYRB_LACCL PYRB_LACPL PYRB_PSEAE PYRB_PSEPU PYRB_SYNY3 PYRB_TREDE PYRB_ECOLI PYRB_SALTY PYRB_SALTY PYRB_SERMA PYRB_SERMA PYRB_SULSO PYRB_ARATH	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YGRTVH.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNA HSR.VARSNM HSR.VARSNL HGRTVH.SKA 256 NAKANMKVLH GARENMKVLH GARENMKVLH KAKDELRIMH KAKDELRIMH KMKKDSIILH VMQKKAIIMH	QALAKFDGNR RTLAKFSGNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.R EILNTLGA.R EILNTLGA.G LALKTLGCPD MLDTLGA.E WSLTTAGA.D DGLKIFKSVK 266 PLP.RVDEI PLP.RIDEI PLP.RIDEI PLP.RVDEI PLP.RVDEI PLP.RVDEI PLP.RVDEI PLP.RVDEI	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR VYLISPELVK VLFSGPSEW VYFSGPSEW VYFSGPEWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VDLIAPAELA 274 ATD.VDKTPH ATD.VDKTPH ATD.VDKTPH ATD.VDKTPH ATD.VDKTTK DRK.VDKTTK TAD.VDADPR	MPQYILDMLD MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. EENTFGT NAEEFSKYGT GRDFEAYGE PIGLEE PIGLIE PVGVT PKEFQQLTLA MPEYYKVRMQ 283 AWYFQQAGNG AWYFQQAGNG AYYFQQAGNG AIYFRQVFNG AKYFEQASYG AAYFRQAKNG	EKGIAWSLHS EKGMAWSLHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVQ QYGVRVFT .QYGVKVYFT .GWPATVSH PGSGKLHCHW ENGFTVREFS 293 IFARQALLAL IFAAQALLAL IFAARA. LAL VPVRMALLAL VPVRMASILTK LFIRMALLKL	SIEEVMAEVD SIEEVMADVD SIEEVVPEUD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEI PELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD SIEEYLRQAD 303 VLNRDLVL~ VLNSELSL~ VVNADLAL~ VLGVI~~ LLVGW~~~~	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQNE VVMLLRIQNE VMMLLRVQHE VVMLLRVQHE VVIMLRVQAE IVMTLRVQAE IVMTLRVQAE VALIWYFTRP	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLITK RHAETMGLITK RLSQVNN RMQGGL RMAGGL RMAGGL QLERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACSL PYRB_LACLE PYRB_LACPL PYRB_PSEAE PYRB_MYCTU PYRB_SYNY3 PYRB_TREDE PYRB_ECOLI PYRB_SALTY PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_PYRAB PYRB_SALTH PYRB_BACCL	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YGRTVH.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNW HSR.VARSNW HSR.VARSNW HSR.VARSNW HSR.VARSNW HSR.VARSNW HSR.VARSNW HSR.VARSNW HGRTVH.SKA 256 NAKANMKVLH GARENMKVLH GA.ANLKVLH KAKDELRIMH KMKKDSILH VMQKKAIIMH RMKSGAIILH	QALAKFDGNR RTLAKF SGNR RTLAKF SGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.E ALLTQLGA.Q LALKTLGCPD LALKALGCPD MLDTLGA.D DGLKIFKSVK 266 PLP.RVDEI PLP.RIDEI PLP.RIDEI PLP.RUDEI PLP.RVDEI	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR VYLISPELVK VLFSGPSEWQ VYFSGPEWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VDLIAPAELA 274 ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDKTTK DRK. VDKTTK TAD. VDADPR ASELVEAKA.	MPQYILDMLD MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. EENTFGT NAEEFSKYGT GRDFEAYGE PIGLEE PVGVT PVGVT PKEFQQLTLA MPEYYKVRMQ 283 AWYFQQAGNG AWYFQQAGNG AYYFQQAGNG AYYFQAGNG AYYFQAGNG AYYFQAGNG SRIFKQMENG SRIFKQMENG	EKGIAWSLHS EKGMAWSLHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVSQYGVRVFT .QYGVKVYT .GWPATVSH PGSGKLHCHW ENGFTVREFS 293 IFARQALLAL IFARSA.LAL VPVRMSILTK LPTVRMSILTK LFTRMALLKL VYVRMAVLKR	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD MDEAVESSD NIDDEIPELD TMDQLVATMD NADEGLKDVD DFDAELPAAD QLQPALEGAD SIEEYLRQAD 303 VLNRDLVL~~ VINSELSL~ VVNADLAL~~ VLGVI~~~~ IYGE~~~~~ AMEGRMEHGR	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQNE VMMLLRVQHE VVMLLRVQHE VVIMLRLQRE AVLMLRVQAE AVLMLRVQAE VALIWYFTRP • •	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RL.DPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNKJAEAKTE RLSQVNN RMQGGL RMAGGL RMNGGF QLERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACCL PYRB_BACSU PYRB_LACPL PYRB_PSEAE PYRB_PSEPU PYRB_SYNY3 PYRB_TREDE PYRB_SCOLI PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACCL PYRB_BACCL	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YGRTVH.SLA YARTVN.SLL MGRTVR.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNL HGRTVH.SKA 256 NAKANMKVLH GARENMKVLH GARENMKVLH KAKDELRIMH KMKKDSIILH VMQKKAIIMH RMKSGAIILH RMKSGAIILH RMKRHAIIMH	QALAKFDSNR RTLAKFSGNR QALAKFEGNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRFKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.R ELLNTLGA.Q LALKTLGCPD LALKALGCPD MLLDTLGA.D MSLTTAGA.D DGLKIFKSVK 266 PLP.RVDEI PLP.RVDEI PLP.RIDEI PLP.RVDEI	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPQLLR IYFVSPEIVK VLFSGPEEWK VLFSGPEEWK VYFSGPEYWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VHLAGPPTLL VHLAGPPTLL ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDKTTK TAD. VDADPR ASELVEAKA. DDSLVESEK.	MPQYILDMLD MPQYILDMLD MPQYILMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. EENTFGT D. EENTFGT GRDFEAYGE PIGLEE PIGI.E PVGVT PKEFQQLTLA MPEYYKVRMQ 283 AWYFQQAGNG AWYFQQAGNG AYYFQQAGNG AYYFQQAGNG AKYFEQASYG AKYFEQASYG ARYFRQAKNG SRIFKQMENG SRIFKQMENG SRIFKQMENG	EKGIAWSLHS EKGMAWSLHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS QYGVRVFT .QYGVKVYT .GYGVKVYT .GYGVKVYT .GYFATVSH ENGFTVREFS 293 1FARQALLAL IFAAQALLAL IFAAQALLAL IFAAQALLAL VPVRMALLAL VPVRMALLAL VPVRMSILTK LFIRMALLKL VYVRMAVLKR VFIRMAVLICR	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD MDEAVESSD NIDDEIPELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DLAEGLKDVD DLAEGLKDVD OFDAELPAAD SIEEYLRQAD 303 VLNRDLVL~~ VLNSELSL~~ VVNADLAL~~ VLGVI~~~~ LLVGW~~~~~ LLVGW~~~~~ AMEGRMEHGR ALQTNVKRGE	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQHE VVMLLRVQHE VVMLRLQRE VVIMLRVQHE VVIMLRVQHE VVIMLRVQRE VVIMLRVQRE VVIMLRVQRE VVIMLRVQRE VVIMLRVQRE VALIWYFTRP	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMMGGFL RMTAHL QLERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACCL PYRB_LACLE PYRB_LACLE PYRB_PSEAE PYRB_PSEAE PYRB_SYNY3 PYRB_TREDE PYRB_SALTY PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACCL PYRB_BACCL PYRB_BACCL PYRB_LACLE	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YGRTVH.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNL HGRTVH.SKA	QALAKFDSNR RTLAKFSGNR QALAKFEGNR RTLAKFSGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.R EILNTLGA.E ALLTQLGA.Q LALKTLGCPD LALKALGCPD MLLDTLGA.E WSLTTAGA.D DGLKIFKSVK 266 PLP.RVDEI PLP.RIDEI PLP.RIDEI PLP.RUDEI PLP.RUDEI PLP.RVNEI PLP.RVNEI PLP.RVNEI PAPVNRGVEI PAPVNRGVEI PAPVNRGVEI PGPINRGVEW	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR IYFVSPEIVK VLFSGPSEWG VYFSGPEWK VLFSGPEWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VHLAGPPTLL VHLAGPPTLL VHLAGPPTLA 274 ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH HPE. VDNTKH DRK. VDKTTK TAD. VDADTK TAD. VDADTK TAD. VDADTK TAD. VDADTK TAD. VDATFH DRK. VDKTPH ATD. VDATFH DRK. VDKTTK TAD. VDADTK TAD. VDATTK TAD. VDAT	MPQYILDMLD MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. EENTFGT TO. EENTFGT NAEEFSKYGT GRDFEAYGE PIGLEE PVGVT PKEFQQLTLA MPEYYKVRMQ 283 AWYFQQAGNG AYYFQQAGNG AYYFQQAGNG AYYFQQAGNG AYYFQQAGNG AYYFQQAGNG AYYFQQAGNG AYYFQQAGNG AYYFQAGNG SRYFRQMKNG SRIFKQMKNG SRIFKQMKNG SRYAVQMHNG	EKGIAWSLHS EKGMAWSLHG EKGIEYSSHE EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVK QYGVRVFT .QYGVKVYT .GWPATVSH PGSGKLHCHW ENGFTVREFS 293 IFARQALLAL IFARSA.LAL VPVRMALLAL IFARSA.LAL VPVRMSILTK LFIRMALLKL VYVRMAVLKR VYFIRMAVIQC VFVRMAMIEA	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEIPELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DLAEGLKDVD DFDAELPAAD SIEEYLRQAD 303 VINNDLVL~~ VINSELSL~~ VVNADLAL~~ VLGVI~~~~ IYGE~~~~~ LLVGW~~~~ AMEGRMEHGR ALQTNVKRGE VLRGRKLGGL	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQHE VVMLLRVQHE VVMLRLQRE VVIMLRUQRE VVIMLRUQRE VVIMLRUQRE VVIMLRUQRE VVALIWYFTRP • •	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMNGGF RMNGGF QLERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACCL PYRB_BACCL PYRB_BACCL PYRB_PSEAE PYRB_PSEAE PYRB_PSEPU PYRB_SYNY3 PYRB_TREDE PYRB_SALTY PYRB_SALTY PYRB_SALTY PYRB_SYNAB PYRB_PYRAB PYRB_PYRAB PYRB_ARATH PYRB_BACCSU PYRB_LACLE PYRB_LACLE PYRB_LACLE	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YGRTVH.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNA HSR.VARSNM HSR.VARSNL HGRTVH.SKA 256 NAKANMKVLH GARENMKVLH GARENMKVLH GARENMKVLH HKAKDELRIMH KMKKDSIILH VMQKKAIIMH RMKKGAIILH RMKRHAIIMH RMKKHAIIMH	QALAKFDGNR RTLAKFSGNR RTLAKFSGNR QALAKFEGNR QALAKFEGNR EALTFYDVEL RILTRFRFKL YLLAKFKDVK EVLTRLGA.R EILNTLGA.R EILNTLGA.G LALKTLGCPD MLDTLGA.E WSLTTAGA.D DGLKIFKSVK 266 PLP.RVDEI PLP.RIDEI PLP.RIDEI PLP.RVDEI PAPVNRGVEL PAPVNRGVEL	FYFIAPDALA FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELV VLISPEEV VLFSGPSEW VYFSGPEW VYFSGPEW VYFSGPEW IRVIAPATLL IRVIAPATLL VLLAGPPTLL VLLAGPPTLL VLLAGPPTLL VLLAPAELA 274 ATD. VDKTPH TTD. VDKTPH ATD. VDADPR ASELVEARA. DDSLVESEK. ASDLVEAPQ.	MPQYILDMLD MPQYILDMLD MPQYILDMLD MPAYILKNLE MRHIVEELR ARKEILDEL. MKDDIKDYLT D.ENTFGT NAEEFSKYGT GRDFEAYGE PIGIEE PIGI.E PKEFQQLTLA MPEYYKVRMQ 283 AWYFQQAGNG AWYFQQAGNG AYYFQQAGNG AYYFQQAGNG AYYFQAGNG AKYFEQASYG AAYFRQAKNG SRIFKQMENG SRIFKQMENG SRYAVQMING SRIFQQMTNG	EKGIAWSLHS EKGMAWSLHG EKGMKVVETT .NYPVKEYE SSGVEWEESS YVE YVS YVQ QYGVRVFT .QYGVKVYT .GWPATVSH PGSGKLHCHW ENGFTVREFS 293 IFARQALLAL IFAAQALLAL IFAAQALLAL IFAARSA. LAL VPVRMALLAL VPVRMANIEA	SIEEVMAEVD SIEEVMAEVD SIEEVMADVD SIEEVVPEUD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEI PELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD SIEEYLRQAD 303 VLNRDLVL~~ VLNSELSL~~ VVNADLAL~~ VLGVI~~~~ LLVGW~~~~~ AMEGRMEHGR ALQTNVKRG LLAHQGLISA	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQNE VMMLLRVQHE VMMLLRVQHE VVIMLRVQAE IVMTLRQRE VALIWYFTRP • • MAEKWHVVQ- AAYVISH TQVEV	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMAGGL RMAGGL QLERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SYRAB PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACCL PYRB_BACSU PYRB_LACPL PYRB_PSEAE PYRB_MYCTU PYRB_SYNY3 PYRB_TREDE PYRB_SALTY PYRB_SALCL PYRB_BACCL PYRB_BACCL PYRB_BACSU PYRB_LACPL PYRB_LACPL PYRB_PSEAE	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YGRTVH.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNL HGRTVH.SKA 256 NAKANMKVLH GARENMKVLH GA.ANLKVLH KAKDELRIMH KMKKDSILH VMQKKAIIMH RMKKGAIILH RMKRHAIIMH MLKDDAIIMH RMPKHAIIMH LAKPDAIVMH	QALAKFDGNR RTLAKF SGNR RTLAKF SGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.R EILNTLGA.E ALLTQLGA.Q LALKTLGCPD MLDTLGA.D DGLKIFKSVK 266 PLP.RVDEI PLP.RIDEI PLP.RIDEI PLP.RVDEI PAPVNRGVEI PAPVNRGVEI PAPVNRGVEI	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR VYLISPELVK VLFSGPSEWQ VYFSGPEWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VDLIAPAELA 274 ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDADPR ASELVEAKA. DDSLVESEK. DGDLVEAPK. ASDLVEAPR ESAVADGAQ.	MPQYILDMLD MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. EENTFGT NAEEFSKYGT GRDFEAYGE PIGLEE PVGVT PVGVT PVGVT PVKFYQQLTLA MPEYYKVRMQ 283 AWYFQQAGNG AWYFQQAGNG AYYFQQAGNG AYYFQQAGNG AYYFQAGNG SKYFQAKNG SRIFKQMENG SRIFKQMENG SRIFKQMENG SRIFKQMING SRIFKQMING SRIFKQMING SRIFKQMING SRIFKQMING SRIFKQMING SRIFKQMING SRIFKQMING SVILNQVTYG	EKGIAWSLHS EKGMAWSLHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVQQYGVRVFT .QYGVKVYT .GWPATVSH PGSGKLHCHW ENGFTVREFS 293 IFARQALLAL IFARSA.LAL VPVRMALLAL IFFARSA.LAL VPVRMALLAL LFIRMALLKL VYVRMAVLKR VFIRMAVIÇC VFVRMAMIEA VYIRMAMVAS IAIRMAVLSM	SIEEVMAEVD SIEEVMAEVD SIEEVMADVD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD MDEAVESSD NIDDEI PELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD SIEEYLRQAD 303 VLNRDLVL~~ VLNSELSL~~ VVNADLAL~~ VLGVI~~~ LLVGW~~~~ AMEGRMEHGR ALQTNVKRGE VLRGRKLGGL VLAHQGLISA AMSGQNTQRQ	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQNE VMMLLRVQHE VVMLLRVQHE VVIMLRVQAE IVMTLQRE VALIWYFTRP MAEKWHVVQ- AAYVISH E	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMAGGL RMAGGL QLERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SYRAB PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACCL PYRB_BACSU PYRB_LACPL PYRB_PSEAE PYRB_PSEAE PYRB_TREDE PYRB_SCTU PYRB_SCTU PYRB_SCTU PYRB_SALTY PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_ARATH PYRB_BACCL PYRB_BACCL PYRB_BACCL PYRB_LACPL PYRB_PSEAE PYRB_PSEAE PYRB_PSEAE	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YGRTVH.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNW HSR.VARSNW HSR.VARSNW HSR.VARSNV HGRTVH.SKA 256 NAKANMKVLH GARENMKVLH GARENMKVLH KMKKDSILH VMÇKKAIIMH KMKKDSILH RMKRHAIIMH MLKDDAIIMH LAKPDAIVMH CAKPDAIVMH	QALAKFDGNR RTLAKF SGNR RTLAKF SGNR RALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA. N EVLTRLGA. R EILNTLGA. E ALLTQLGA. Q LALKTLGCPD LALKALGCPD MLDTLGA. D DGLKIFKSVK 266 PLP. RVDEI PLP. RIDEI PLP. RIDEI PLP. RVDEI	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPQLLR IYFVSPEIVK VLFSGPSEWC VYFSGPEWY VYFGGPKEWY IRVIAPATLL IRVIAPATLL IRVIAPATLL VLLIAPAELA 274 ATD. VDKTPH ATD.	MPQYILDMLD MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. EENTFGT NAEEFSKYGT GRDFEAYGE PIGLEE PVGVT PVGVT PKEFQQLTLA MPEYYKVRMQ 283 AWYFQQAGNG AWYFQQAGNG AYYFQQAGNG AYYFQAGNG AKYFEQASYG AKYFEQASYG AKYFEQASYG AKYFEQASYG SKIFKQMENG SRIFKQMENG SRIFKQMENG SRIFKQMING SRIFKQMING SRIFKQMING SVILNQVTYG SVILNQVTYG	EKGIAWSLHS EKGMAWSLHG EKGMAWSLHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS QYGVRVFT .QYGVKVYT .GWPATVSH PGSGKLHCHW ENGFTVREFS 293 IFARQALLAL IFAAQALLAL IFAASA.LAL VPVRMSILTK LPIRMALLAL VPVRMSILTK LPIRMAVLKR VFIRMAVICC VFVRMAMIEA ZIARMAVLSM IAVRMAVLSM	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD MDEAVESSD NIDDEIPELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD SIEEYLRQAD 303 VLNRDLVL VLNSELSL VVNADLAL VLGYUI LIVGW AMEGRMEHGR ALQTNVKRGE VLRGRKLGGL LLAGKUGGL LLAGKUGGL AMSGQNTQRQ AMSGQNAQRQ	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQNE VMMLLRVQHE VVMLLRVQHE VVIMLRLQRE VVIMLRVQAE IVMTLRLQRE VALIWYFTRP • • MAEKWHVVQ- AAYVISH E TQVEV FDQENAQ	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMAGGL RMNGGF QLERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACCL PYRB_BACSU PYRB_LACPL PYRB_PSEAE PYRB_PSEAE PYRB_TREDE PYRB_SYLYO PYRB_STREDE PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SHACTU PYRB_SALTY PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_LACLE PYRB_BACSU PYRB_BACSU PYRB_LACLE PYRB_LACLE PYRB_LACLE PYRB_PSEAE PYRB_PSEAE PYRB_PSEAE PYRB_PSEAE PYRB_PSEAE PYRB_PSEAE PYRB_MYCTU	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YGRTVH.SLA YGRTVH.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNL HGRTVH.SKA 256 NAKANMKVLH GARENMKVLH KAKDELRIMH KMKKDSIILH VMCKKAIIMH KMKKDSIILH RMKRHAIIMH MLKDDAIIMH RMKSGAIILH RMKRHAIIMH LAKPDAIVMH CARPDAIVMH CARPDAIVMH MLPGHAVVLH	QALAKFDSNR RTLAKF SGNR QALAKFEGNR RTLAKF SGNR QALAKFEGNR EALTFYDVEL RILTRFRFKL YLLAKFKDVK EVLTRLGA.N EVLTRLGA.R EVLTRLGA.Q LALKTLGC PD LALKALGC PD LALKALGC PD LALKALGC PD MLLDTLGA.D GSLKIFKSVK 266 PLP.RVDEI PLP.RVDEI PLP.RIDEI PLP.RVDEI PAPVNRGVEI PAPVNRGVEI PGPINRGVEN PAPVNRGVEI PGPINRGVEI PGPMVRGMEI	FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR IYFVSPEIVK VLFSGPEEWK VLFSGPEEWK VYFSGPEYWY VYFGGPKEWY IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VHLAGPPTLL VHLAGPPTLL ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDKTPK TAD. VDATPK ASDLVEAPK ASDLVEAPK ASDLVEAPK ASDLVEAPK ASDLVEAPQ ESAVADGKH TSSVADSSQ.	MPQYILDMLD MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. EENTFGT D. EENTFGT GRDFEAYGE PIGLEE PIGI.E PVGVT PKEFQQLTLA MPEYYKVRMQ 283 AWYFQQAGNG AYYFQQAGNG AYYFQQAGNG AYYFQQAGNG AYYFQQAGNG SRIFKQMENG SRIFKQMING SRIFLQWING SVILNQVTYG SVULNQVTYG SAVLQQVSNG	EKGIAWSLHS EKGMAWSLHG EKGMAWSLHG EKGIEYSSHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS QYGVRVFT .QYGVKVYT .GYGVKVYT .GYGVKVYT .GYGVKVYT .GYGVKVYT .GYGVKVYT .GYGVKVYT .GYGVKVYT .GYGVATVH ENGFTVREFS 293 IFARQALLAL IFAAQALLAL IFAAQALLAL IFAAQALLAL VPVRMALLAL VPVRMALLAL VPVRMALLAL VPVRMALLAL VPVRMALLAL VPVRMAVLKR VFIRMAVLKR VFIRMAVIC VFVRMAWIEA VYIRMAWVLSM IAVRMAVLSM VQVRMAVLFH	SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD MDEAVESSD NIDDEIPELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DLAEGLKDVD DLAEGLKDVD SIEEYLRQAD 303 VLNRDLVL~~ VLNSELSL~~ VVNADLAL~~ VLGVI~~~~ LLVGW~~~~~ LLVGW~~~~~ LLVGW~~~~~ AMEGRMEHGR ALQTNVKRGE VLRGKLGGL VLAHQGLISA AMSGQNTQRQ AMSGQNAQRQ VLVGAQDAGK	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQHE VVMLLRVQHE VVMLRLQRE VVIMLRUQRE VVIMLRUQRE VVIMLRUQRE VVIMLRUQRE VALIWYFTRP	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMTGGF RMTAHL QLERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACCL PYRB_BACSU PYRB_LACPL PYRB_PSEAE PYRB_PSEAE PYRB_SYNY3 PYRB_TREDE PYRB_SALTY PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_ARATTH PYRB_BACCL PYRB_BACCL PYRB_BACSU PYRB_LACLE PYRB_LACPL PYRB_PSEAE PYRB_PSEAE PYRB_PSEAE PYRB_PSEAE PYRB_PYRTU PYRB_MYCTU PYRB_SYNY3	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YGRTVH.SLA YGRTVH.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNM HSR.VARSNL HGRTVH.SKA	QALAKFDGNR RTLAKFSGNR RTLAKFSGNR QALAKFEGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.R EILNTLGA.R EILNTLGA.G LALKTLGCPD MLDTLGA.G WSLTTAGA.D DGLKIFKSVK 266 PLP.RVDEI PLP.RIDEI PLP.RIDEI PLP.RVDEI PROVNRGVEI PGPINRGVEU PAPVNRGVEI PGPINRGVEI PGPWNRGWEI PGPWNRGWEI	FYFIAPDALA FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR VYLISPEEW VYFSGPEEW VYFSGPEW VYFSGPEWY VYFGGPKEW IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VHLAGPPTLL VDLIAPAELA 274 ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDRTPH ASELVEAKA. DDSLVESEK. DGDLVEAPR. ASDLVEAPQ. ESAVADGKH. TSSVADSSQ. SSELMDDPDI	MPQYILDMLD MPQYILDMLD MPQYILDMLD MPAYILKMLE MRHIVEELR ARKEILDEL. MKDDIKDYLT D.ENTFGT NAEEFSKYGT GRDFEAYGE PIGIEE PIGI.E PKEFQQLTLA MPEYYKVRMQ 283 AWYFQQAGNG AYYFQQAGNG AYYFQQAGNG AYYFQQAGNG AYFFQAFNG AKYFEQASYG AAYFRQAFNG SRIFKQMENG SRIFKQMENG SRIFKQMENG SRIFKQMING SRIFQQMTNG SVILNQVTYG SAVLQVSNG SAVLQVSNG SAVLQVSNG SAVLQQVSNG SAVLQQVSNG SAVLQQVSNG SALIQDQVTSG	EKGIAWSLHS EKGMAWSLHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVS QYGVRVVT .QYGVKVYT .GWPATVSH PGSGKLHCHW ENGFTVREFS 293 IFARQALLAL IFAAQALLAL IFAAQALLAL IFARSA.LAL VPVRMSILTK LFIRMALLKL VYVRMAVLKR VFIRMAVURQ VFVRMAMUEA VYIRMAMVAS IAIRMAVLSM IAVRMAVLSH VQVRMAVLFH VAIRMALLYL	SIEEVMAEVD SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEI PELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD SIEEYLRQAD 303 VLNRDLVL~~ VINSELSL~~ VVNADLAL~~ VLGVI~~~~ LLVGW~~~~ AMEGRMERG ALQTNVKRG ALQTNVKRG ALQTNVKRG AMSGQNAQRQ AMSGQNAQRQ LGTVQE~~~~~	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQNE VMMLLRVQHE VMMLLRVQHE VVIMLRLQRE AVIMLRVQAE IVMTLRLQKE VALIWYFTRP MAEKWHVVQ MAEKWHVVQ LEQEDAE FDQENAQ ILYMTRVQAE CONTROL TQVEV	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMAGGL QLERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK
PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_SULSO PYRB_BACCL PYRB_BACCL PYRB_BACSU PYRB_LACPL PYRB_PSEAE PYRB_PSEAE PYRB_SYNY3 PYRB_TREDE PYRB_SALTY PYRB_SALTY PYRB_SERMA PYRB_PYRAB PYRB_ARATTH PYRB_BACCL PYRB_BACCL PYRB_BACSU PYRB_LACLE PYRB_LACPL PYRB_PSEAE PYRB_PSEAE PYRB_PSEAE PYRB_PSEAE PYRB_PYRTU PYRB_MYCTU PYRB_SYNY3	YGRTVH.SLT YGRTVHFAKP YGRTVH.SLT YGRTVH.SLA YGRTVH.SLA HSR.VARSNA HSR.VARSNA HSR.VARSNA HSR.VARSNM	QALAKFDGNR RTLAKFSGNR RTLAKFSGNR QALAKFEGNR QALAKFEGNR EALTFYDVEL RILTRFRPKL YLLAKFKDVK EVLTRLGA.R EILNTLGA.R EILNTLGA.G LALKTLGCPD MLDTLGA.G WSLTTAGA.D DGLKIFKSVK 266 PLP.RVDEI PLP.RIDEI PLP.RIDEI PLP.RVDEI PROVNRGVEI PGPINRGVEU PAPVNRGVEI PGPINRGVEI PGPWNRGWEI PGPWNRGWEI	FYFIAPDALA FYFIAPDALA FYFIAPDALA FYFIAPDALA YLISPELLR VYLISPELLR VYLISPEEW VYFSGPEEW VYFSGPEW VYFSGPEWY VYFGGPKEW IRVIAPATLL IRVIGPKTLI VVLVAPPTLL VHLAGPPTLL VDLIAPAELA 274 ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDKTPH ATD. VDRTPH ASELVEAKA. DDSLVESEK. DGDLVEAPR. ASDLVEAPQ. ESAVADGKH. TSSVADSSQ. SSELMDDPDI	MPQYILDMLD MPQYILDMLD MPQYILDMLD MPAYILKMLE MPRHIVEELR ARKEILDEL. MKDDIKDYLT D. EENTFGT D. EENTFGT GRDFEAYGE PIGLEE PIGI.E PVGVT PKEFQQLTLA MPEYYKVRMQ 283 AWYFQQAGNG AYYFQQAGNG AYYFQQAGNG AYYFQQAGNG AYYFQQAGNG SRIFKQMENG SRIFKQMING SRIFLQWING SVILNQVTYG SVULNQVTYG SAVLQQVSNG	EKGIAWSLHS EKGMAWSLHG EKGMKVVETT .NYPVKEVE SSGVEWEESS YVE YVS YVS QYGVRVVT .QYGVKVYT .GWPATVSH PGSGKLHCHW ENGFTVREFS 293 IFARQALLAL IFAAQALLAL IFAAQALLAL IFARSA.LAL VPVRMSILTK LFIRMALLKL VYVRMAVLKR VFIRMAVURQ VFVRMAMUEA VYIRMAMVAS IAIRMAVLSM IAVRMAVLSH VQVRMAVLFH VAIRMALLYL	SIEEVMAEVD SIEEVMAEVD SIEEVMADVD SIEEVVPELD TLEDVIGKLD NPFEVINEVD DLMEVASKCD .VDEAIARAD .MDEAVESSD NIDDEI PELD TMDQLVATMD NADEGLKDVD DLAEGLKDVD DLAEGLKDVD DFDAELPAAD QLQPALEGAD SIEEYLRQAD 303 VLNRDLVL~~ VINSELSL~~ VVNADLAL~~ VLGVI~~~~ LLVGW~~~~ AMEGRMERG ALQTNVKRG ALQTNVKRG ALQTNVKRG AMSGQNAQRQ AMSGQNAQRQ LGTVQE~~~~~	ILYMTRVQKE ILYMTRVQKE ILYMTRVQKE VLYVTRIQKE VLYVTRIQKE VVYQTRIQRE VVMLLRIQHE VVMLLRIQNE VMMLLRVQHE VMMLLRVQHE VVIMLRLQRE AVIMLRVQAE IVMTLRLQKE VALIWYFTRP MAEKWHVVQ MAEKWHVVQ LEQEDAE FDQENAQ ILYMTRVQAE CONTROL TQVEV	RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RLDPSEYA RFPDEQE.YL RFVDEME.YE RFGERLDLYE RHAETMGLTK RHQSAVSQ RHNGAEAKTE RLSQVNN RMQGGL RMAGGL QLERMGEQ	NVKA NVKA NVKA KVKG KIKG AARG E EYHA E GYLN QLFDAKDYNA QTFDASAYHQ LPSEGEFFK LPSEGEFFY LPSEGEFFYR LPSEGEFYSV LPSLREYHH	QFVLRASDLH QFVLR.PDLN QFVL.ADLL SYQVNLKVLE SYIVSLDLAN KFIVDKDLLG RYGLTLERAR KYGLTVERAE AYGLNQRRYD QYGLTAERAA LYGLTEKRLK LFGLTTARLA YFGITHDRLK

FIGURE 3. Sequence alignment of selected ATCase catalytic subunits. PALA binding residues are indicated by •, and residues proposed to be catalytically important are indicated by ▽. Sequence numbering is based on the sequence of *E. coli* ATCase. The organisms used in the alignment were *Escherichia coli*, *Salmonella trphimurum*, *Serratia marcescens*, *Pyrococcus abyssi*, *Sulfolobus solfataricus*, *Arabidopsis thaliana*, *Bacillus caldolyticusd*, *Bacillus subtilis*, *Lactobacillus leichmannii*, *Lactobacillus plantarum*, *Pseudomonas aeruginosa*, *Pseudomonas putida*, *Mycobacterium tuberculosis*, *Synechocystis* sp., and *Treponema denticola*. * indicates sequences not shown for clarity.

indicated.) The positive charges of the active site complement the negative charge of the ligand. In both families, at least one residue at this binding site is contributed by

a second subunit. In *E. coli* ATCase, these residues are Ser 80 and Lys 84. In *E. coli* OTCase, Gln 82 (His 117 in human OTCase), the counterpart of Ser 80 in *E. coli*

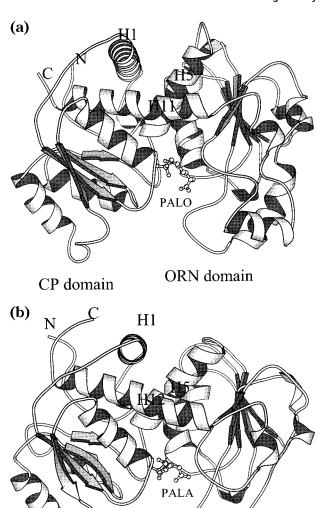


FIGURE 4. Ribbon drawing of subunit of (a) *E. coli* OTCase and (b) ATCase. The bisubstrate analogues PALA or PALO, which are located in the cleft between the CP domain and the L-Asp or L-Orn domain, are shown as ball-and-stick models. The α 9a loop, present in *E. coli* OTCase adjacent to the 240s loop, is not present in *E. coli* ATCase and some other OTCases. In general, the tertiary structures of the ATCase catalytic subunit and OTCase subunit are very similar.

CP domain

ASP domain

ATCase, forms a hydrogen bond with one of the phosphonate oxygens of the bound bisubstrate analogue PALO. Lys 86, which corresponds to Lys 84 in *E. coli* ATCase, is adjacent to both the carbonyl O of Gln 82 and the carboxy O of PALO, but not close enough to form direct bonds. However, these interactions appear to be functionally significant, since mutating Lys 86 to Gln lowers the enzyme's maximal velocity by 2 orders of magnitude.³² These side chains may be bridged by solvent molecules.

Aspartate or Ornithine Binding Site

L-Orn and L-Asp have side chains with different lengths and different charges, and their binding sites have evolved to reflect these differences (Figure 7). Several side chains

at the L-Asp binding site of ATCase are positively charged, such as Lys 84, Arg 167, and Arg 229, while others such as Gln 231 are neutral. The binding site for L-Orn in OTCase has a negatively charged side chain, Asp 231, and two long hydrophobic side chains, Met 236 and Leu 125, which form a hydrophobic pocket for the methylene group of L-Orn, as well as two hydrogen-bonding side chains, Ser 235 and Asn 167.

Electrostatic Features of Active Site

Both OTCase and ATCase have a high density of charged residues at their active sites which create electrostatic potentials that can guide incoming substrates to their binding sites and assist in their binding. The electrostatic potentials of the catalytic subunits of E. coli ATCase and OTCase are shown in Figure 8. In both ATCase and OTCase, substrate binding and product release are ordered, with CP binding first, L-Asp or L-Orn binding second, carbamoylaspartate or citrulline dissociating first, and phosphate dissociating second. 33,34 When ATCase and OTCase are unliganded, their active sites are dominated by positive electrostatic potentials that will guide their common substrate, CP, with two negative charges at physiological pH, into the CP binding pocket. After CP binds, the electrostatic potentials around the active sites of ATCase and OTCase have the opposite sign. While the active site of ATCase is still dominated by a positive potential, the electrostatic potential of the active site in OTCase has become negative. The positive potential of ATCase will help dock negatively charged L-Asp, while the negative potential of OTCase will play the same role in docking the positively charged L-Orn.

Catalytic Mechanisms

The critical step in the catalytic mechanisms of both ATCases and OTCases is nucleophilic attack of the carbonyl C of CP by an amino group. In both enzymes, the crystal structures indicate that attack is facilitated by interactions between the protein and its substrates, which increase the nucleophilicity of the amino group and the positive charge on the carbonyl C. However, the greater basicity of the δ -amino group of L-Orn requires general acid—base catalysis in OTCase, while L-Asp does not.

The group which appears to function as a general acidbase catalyst in OTCase is the sulfhydryl group of a Cys residue, which is found in a conserved HCLP motif. HCLP is replaced by HPLP in *E. coli* ATCase, which does not require a general acid—base catalyst. The S atom of this Cys in the crystal structure of the complex of the bisubstrate analogue PALO with *E. coli* OTCase is 4.3 Å from the ϵ -N of the L-Orn moiety in PALO. It thus has the potential to act as a general acid—base catalyst and to abstract a proton from the δ -amino group of L-Orn. Its basicity is increased by a hydrogen bond to the side chain of Asp 231. The Cys side chain is also close to the imidazole group of His 272, which in turn interacts with Glu 299. Although this Cys-His-Glu triad is reminiscent

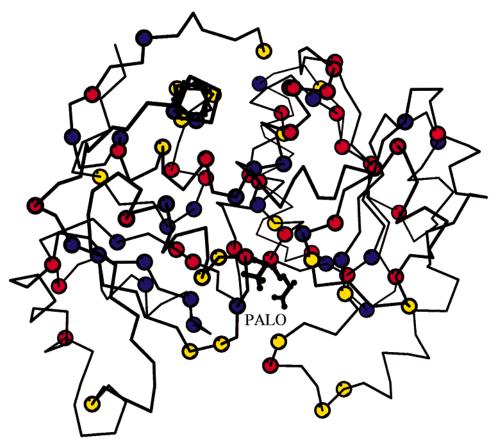


FIGURE 5. Deleterious mutations found in patients with OTCD mapped on the structure of human OTCase. Mutations which produce neonatal OTCD are shown in red, those identified only in females are in blue, and those which produce late onset OTCD are in yellow.

of the catalytic triads in Cys and Ser proteases, the distance between the γ -S of Cys and the imidazole ring is too long for a strong interaction; however, a stronger interaction in some other steps of the catalytic cycle cannot be ruled out at this point.

The catalytic mechanisms of both ATCase and OTCase involve a tetrahedral intermediate (Figure 1). Gln 136 in *E. coli* OTCase, which is 4.1 Å from the carbonyl C of the CP moiety, is positioned so as to be able to stabilize this intermediate, while Gln 137 plays the same role in *E. coli* ATCase. The sequence motif NxLxxxxHxxQxxD around Gln 136 is strongly conserved in OTCases, as is the corresponding motif NxGDGxxxHxxQxxD in ATCases. The similarity between these motifs suggests that both OTCase and ATCase may use a similar mechanism to stabilize the tetrahedral intermediate. However, replacement of the GDG motif in ATCase by L in OTCase reduces backbone flexibility and enables the long side chain of L-Orn to be held in place by the Leu side chain of the protein.

Subunit Interfaces

The interface between subunits in the trimers of both OTCase and ATCase is formed primarily by residues from the CP binding domains. In *E. coli* OTCase, these residues are 57–68, 72–98, and 278–317; the residues at the subunit interfaces of ATCase are analogous. Binding of substrate analogues to the active site triggers localized

conformational changes in the vicinity of residues 52 and 84 which strengthen subunit interactions; the restructuring is greater in ATCase, where it contributes to the T \rightarrow R quaternary transition, 35 than it is in OTCase (Figure 6). Although most residues involved in intersubunit interactions are variable, the salt bridge between Arg 57 and Glu 87 is conserved in both OTCase and ATCase, as is the interaction between Arg 59 and the carbonyl O of residue 75. These interactions are likely to be important in positioning Gln 82 in OTCase and Ser 80, Lys 84 in ATCase to interact with bound CP.

Other Protein—Protein Interfaces

Although both *E. coli* and human OTCase are trimeric in vitro, the OTCases of *Pseudomonas aeruginosa* and *Pyrococcus furiosus*, a thermophilic archaeabacterium, are dodecameric, with 23-point group symmetry. The trimers are organized tetrahedrally, with their convex surfaces in contact. However, a 6° rotation of the trimers of the *Pseudomonas* and *Pyrococcus* OTCases with respect to each other results in the intertrimer interfaces being quite different in the two proteins. In *P. aeruginosa*, a cluster of charged residues rich in Arg forms a 3-fold channel, which may bind negatively charged ions such as sulfate or phosphate.²⁶ Contacts between monomers from adjacent catalytic trimers around the 2-fold symmetry axes appear to play a role in the allosteric behavior of this

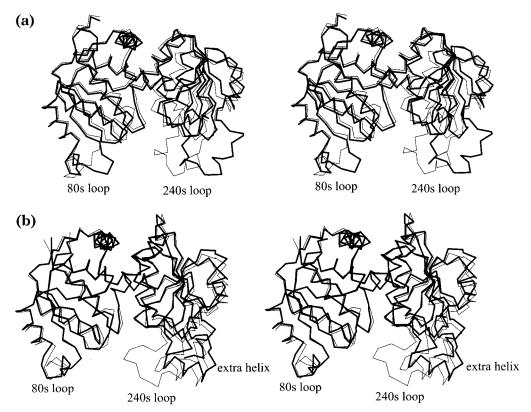


FIGURE 6. Superposition of liganded (thin line) and unliganded (thick line) catalytic subunits of (a) *E. coli* ATCase and (b) OTCase. The Protein Data Bank references for the structures are 8ATC (liganded *E. coli* ATCase), 1RAI (unliganded *E. coli* ATCase), 2ORT (liganded *E. coli* OTCase), and 1AKM (unliganded *E. coli* OTCase).

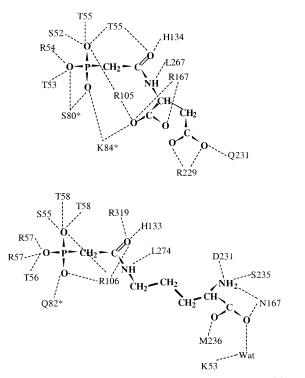


FIGURE 7. Schematic drawing showing the interaction of (a) the bisubstrate analogue PALA with active-site residues in *E. coli* ATCase and (b) the bisubstrate analogue PALO with active-site residues in *E. coli* OTCase. The residue indicated with * is from an adjacent subunit.

enzyme, since eliminating these residues eliminates allosteric regulation. In *Py. furiosus* OTCase, the Arg residues

at this interface are replaced by Trp, giving it a hydrophobic character. Since hydrophobic interactions strengthen as temperature increases, this feature of Py. furiosus OTCase, together with an increase in the number of electrostatic interactions within each monomer, probably accounts for its ability to retain 50% activity, even after being heated at 100 °C for as long as 1 h. 36

E. coli ATCase also exists as a dodecamer, however, composed of six catalytic chains and three regulatory chains, organized as two catalytic subunits and three regulatory subunits with 32-point group symmetry. This complex organization creates three unique subunit interfaces for each catalytic chain, one between catalytic subunits, and two types of interfaces between catalytic and regulatory subunits. In contrast to the prokaryotic OTCases discussed above, contacts between the catalytic trimers of *E. coli* ATCase in the T state involve the concave surfaces of the trimers, with residues in the 240s loops interacting, primarily through ionic or polar interactions. All subunit interfaces undergo major changes in the $T \rightarrow$ R allosteric transition. When the transition occurs, the contact between the catalytic subunits and one of the contacts between catalytic and regulatory subunits are eliminated.

The homology between E. coli OTCase and ATCase raises the question of why OTCase does not associate with the regulatory subunit of ATCase, since both coexist in the cytoplasm. This interaction may be prevented by an extra helix (α 9a) present in both E. coli and P. aeruginosa catabolic OTCase which interacts with the 240s loop,

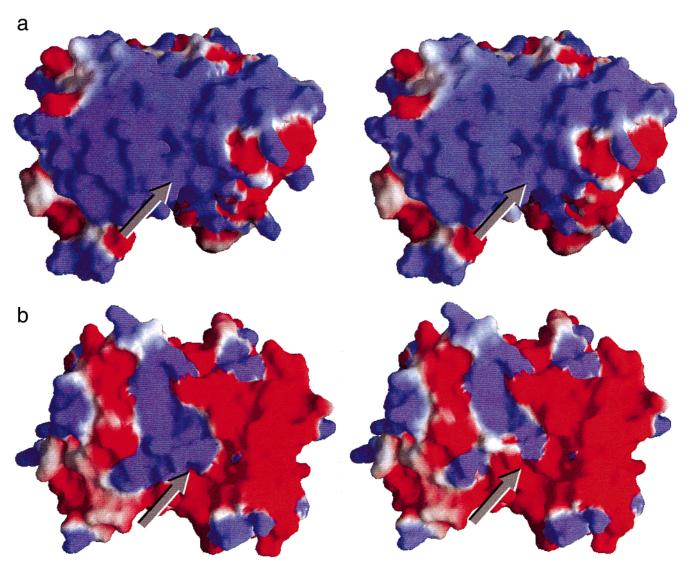


FIGURE 8. Electrostatic potential of (a) E. coli ATCase and (b) OTCase catalytic subunit before and after binding the first substrate CP. Potentials greater than +1 kT are shown in blue, and those less than -1 kT in are shown in red, mapped onto the molecular surface. The subunit is viewed in the same orientation as that in Figure 5. The active site is in the cleft between two domains, as indicated by the arrow. This figure was produced with GRASP.41

mainly through hydrophobic bonds, and with the 80s loop. The position of this helix, combined with sequence differences in a number of residues, would disrupt the potential interface between E. coli OTCase and the regulatory subunit of E. coli ATCase.15

Ureotelic OTCases have a C-terminus extension which folds back on the L-Orn domain and forms a ridge on the convex face of the trimer. This extension has a surprising homology with several membrane associated proteins, including two yeast mitochondrial inner membrane carrier proteins. 16 In models of these carrier proteins, this sequence occurs in a loop between two transmembrane helices in a region of the molecule which has been shown to be important in carrier-specific transport. Since OTCase has been shown to be associated with the inner mitochondrial membrane,37 the sequence of this extension may be a protein recognition motif which enables OTCase and other membrane-associated proteins to interact with other proteins with related functions.

Relationship between State of Assembly and Regulation

Although trimeric OTCases undergo domain closure when substrates bind, wild-type trimeric OTCases appear to be unregulated, although allostery can be induced through site-directed mutagenesis. Higher order assemblies of OTCase may or may not be regulated. Catabolic P. aeruginosa OTCase is regulated by the binding of negatively charged allosteric activators such as phosphate or nucleoside monophosphates at the interfaces between trimers in the dodecamer, while anabolic Py. furiosus OTCase, also a dodecamer, is not.

In yeast, a regulatory complex consisting of one molecule of OTCase and one molecule of arginase forms to inactivate OTCase, effectively uncoupling the biosynthetic and catabolic pathways in arginine metabolism.^{38,39} The conformational changes which promote association of OTCase with arginase have been proposed to be linked to binding of substrates at the active site.

ATCases comprised of both catalytic and regulatory subunits are highly regulated, although the details of the regulation vary between species. Chimeric enzymes have provided considerable insight into the elements of secondary structure that determine the nature of the regulation. For example, CTP stimulates the catalytic activity of Serratia marcescens ATCase, while CTP and UTP synergistically inhibit *E. coli* ATCase. When five divergent residues in the S5′ β -strand at the junction of the allosteric and zinc domains of the regulatory chain of *S. marcescens* ATCase were replaced by the corresponding residues in *E. coli* ATCase, the chimeric ATCase acquired the allosteric properties of *E. coli* ATCase.

Conclusions

Structural and functional comparisons between the OT-Cases and ATCases provide an unusual opportunity to compare how two families of enzymes recognize their substrates and assemble to form large aggregates. The primary and tertiary structures of the domains that bind the common substrate CP are similar, while those of the domains that bind the second substrate, L-Asp or L-Orn, are different. The binding pockets for the second substrate are complementary to the substrates and therefore differ between OTCases and ATCases, despite some common features. The similarity in the active sites of ATCase and OTCase implies that they use similar mechanisms to stabilize the tetrahedral intermediate in the reaction mechanism. Differences in the active site, particularly replacement of the HPLP motif in ATCase by HCLP in OTCase, compensate for the difference in basicity of the α -amino group of L-Asp and the δ -amino group of L-Orn. The electrostatic properties of the active site are consistent with the ordered reaction mechanism for both enzymes, with the change of electrostatic potential in OTCase after binding the first substrate, CP, reflecting the different net charge of the second substrate, L-Asp or L-Orn. Variability in primary structure within and between the two enzyme families determines their ability to associate with other enzymes or to self-assemble to form larger aggregates. Association and aggregation of the enzymes create new properties, such as allosteric regulation, thermal stabilization, or inactivation.

This work was supported by public health service grant DK47870 from the National Institute of Diabetes Digestive and Kidney Disease.

References

- (1) Knier, B.; Allewell, N. M. Calorimetric analysis of aspartate transcarbamylase from *Escherichia coli*: Binding of substrates and substrate analogs to the native enzyme and catalytic subunit. *Biochemistry* **1978**, *17*, 784–790.
- (2) Edge, V.; Allewell, N. M.; Sturtevant, J. M. Differential Scanning calorimetric study of the thermal denaturation of aspartate transcarbamylase of *Escherichia coli. Biochemistry* **1988**, *27*, 8081–8087.

- (3) Burz, D. S.; Allewell, N. M. Interactions of ionizable group in *Escherichia coli* aspartate transcarbamylase with adenosine and cytosine 5'-triphosphates. *Biochemistry* **1982**, *21*, 6647–6655.
- (4) McCarthy, M. P.; Allewell, N. M. Thermodynamics of assembly of *Escherichia coli* aspartate transcarbamylase. *Proc. Natl. Acad. Sci. U.S.A.* 1983, 80, 6824–6828.
- (5) Oberoi, H.; Trika, J.; Yuan, X.; Allewell, N. M. Identification and analysis of long-range electrostatic effects in proteins by computer modeling: Aspartate transcarbamylase. *Proteins: Struct., Funct. Genet.* **1996**, *25*, 300–314.
- (6) Hariharan, M.; Allewell, N. M. Effects of the T→R transition on the electrostatic properties of *Escherichia coli* aspartate transcarbamylase. *Proteins: Struct., Funct. Genet.* **1998**, *32*, 200–210.
- (7) Lennick, M.; Allewell, N. M. Changes in the hydrogen exchange kinetics of *Escherichia coli* aspartate transcarbamylase produced by effector binding and subunit association. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *21*, 6759–6763.
- (8) Mallikarachchi, D.; Burz, D. S.; Allewell, N. M. Effects of CTP and ATP on the conformation of the regulatory subunit of *Escherichia coli* aspartate transcarbamylase: A medium resolution hydrogen exchange study. *Biochemistry* 1989, 28, 5386-5391.
- (9) LiCata, V. J.; Allewell, N. M. Functionally linked hydration changes in *Escherichia coli* aspartate transcarbamylase. *Biochemistry* **1997**, *36*, 10161–10167.
- (10) LiCata, V. J.; Burz, D. S.; Moerke, N. J.; Allewell, N. M. The magnitude of the allosteric conformational transition of aspartate transcarbamylase is altered by mutations. *Biochemistry* 1998, 37, 17381–17385.
- (11) Allewell, N. M.; Hofmann, G. E.; Zaug, A.; Lennick, M. The Bohr effect in *Escherichia coli* aspartate transcarbamylase: Linkages between proton binding, ligand binding, and conformational transitions. *Biochemistry* 1979, 18, 3008–3015.
- (12) Bromberg, S.; LiCata, V. J.; Mallikarachchi, D.; Allewell, N. M. Ligation alters the pathway of ureainduced denaturation of the catalytic trimer of *Escherichia coli* aspartate transcarbamylase. *Protein Sci.* **1994**, *3*, 1236–1244.
- (13) Allewell, N. M.; LiCata, V. J. Thermodynamic approaches to understanding aspartate transcarbamylase. *Methods Enzymol.* **1995**, *259*, 608–628.
- (14) Ha, Y.; Allewell, N. Intersubunit hydrogen bond acts as a global molecular switch in *Escherichia coli* aspartate transcarbamoylase. *Proteins: Struct., Funct. Genet.* **1998**, *33*, 430–443.
- (15) Ha, Y.; McCann, M. T.; Tuchman, M.; Allewell, N. M. Substrate-induced conformational change in a trimeric ornithine transcarbamoylase. *Proc. Natl. Acad. Sci. U.S.A.* 1997, 94, 9550–9555.
- (16) Shi, D.; Morizono, H.; Ha, Y.; Aoyagi, M.; Tuchman, M.; Allewell, N. M. 1.85 Å resolution crystal structure of human ornithine transcarbamoylase complexed with N-phosphonacetyl-L-ornithine. J. Biol. Chem. 1998, 273, 34247–34254.
- (17) Morizono, H.; Tuchman, M.; Rajagopal, B. S.; Mc-Cann, M. T.; Listrom, C. D.; Yuan, X.; Venugopal, D.; Barany, G.; Allewell, N. M. Expression, purification and kinetic characterization of wild-type human ornithine transcarbamylase and a recurrent mutant that produces 'late onset' hyperammonaemia. *Biochem. J.* 1997, 322, 625–631.

- (18) Morizono, H.; Listrom, C. D.; Rajagopal, B. S.; Aoyagi, M.; McCann, M. T.; Allewell, N. M.; Tuchman, M. 'Late onset" ornithine transcarbamylase deficiency function of three purified recombinant mutant enzymes. *Human Mol. Genet.* 1997, 6, 963–968.
- (19) Tuchman, M.; Morizono, H.; Rajagopal, B. S., Plante, R. J.; Allewell, N. M. Biochemical and molecular spectrum of ornithine transcarbamylase deficiency. J. Inh. Met. Dis. 1998, 21, 40-58.
- (20) Schachman, H. K. From allostery to mutagenesis: 20 years with aspartate transcarbamoylase. *Biochem. Soc. Trans.* **1987**, *15*, 772–775.
- (21) Allewell, N. M. *Escherichia coli* aspartate transcarbamoylase: structure, energetics, and catalytic and regulatory mechanisms. *Annu. Rev. Biophys. Chem.* **1989**, *18*, 71–92.
- (22) Hervé, G. Aspartate transcarbamylase; In *Allosteric Enzymes*; CRC Press: Boca Raton, FL, 1989; pp 61–79.
- (23) Kantrowitz, E. R.; Lipscomb, W. N. *Escherichia coli* aspartate transcarbamoylase: the molecular basis for a concerted allosteric. *Trends Biochem. Sci.* **1990**, *15*, 53–59.
- (24) Wild, J. R.; Wales, M. E. Molecular evolution and genetic engineering of protein domains involving aspartate transcarbamoylase. *Annu. Rev. Microbiol.* **1990**, *44*, 193–218.
- (25) Lipscomb, W. N. Aspartate transcarbamylase from *Escherichia coli*.: activity and regulation. *Adv. Enzyme Relat. Areas Mol. Biol.* **1994**, *68*, 67–151.
- (26) Vileret, V.; Tricot, C.; Stalon, V.; Dideberg, Crystal structure of *Pseudomonas aeruginosa* catabolic ornithine transcarbamoylase at 3.0-Å resolution: a different oligomeric organization in the transcarbamoylase family. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 10762–10766.
- (27) Jin, L.; Seaton, B. A.; Head, J. F. Crystal structure at 2.8 Å resolution of anabolic ornithine transcarbamylase from *Escherichia coli*. *Nature Structural Biology* **1997**, *4*, 622–625.
- (28) Villeret, V.; Clantin, B.; Tricot, C.; Legrain, C.; Roovers, M.; Stalon, V.; Glansdorff, N.; Jozef van Beeumen. The crystal structure of *Pyrococcus furiosus* ornithine carbamoyltransferase reveals a key role for oligomerization in enzyme stability at extreme high temperature. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 2801–2806.
- (29) Tuchman, M.; Morizono. H.; Reish, O.; Yuan, X.; Allewell, N. M. The molecular basis of ornithine transcarbamylase deficiency: Modeling the human enzyme and the effects of mutations. *J. Med. Genet.* **1995**, *32*, 680–688.

- (30) Houghton, J. E.; O'Donovan, G. A.; Wild, J. R. Reconstruction of an enzyme by domain substitution effectively switches substrate specificity. *Nature* **1989**, *338*, 172–174.
- (31) Murata, L. B.; Schachman, H. K. Structural similarity between ornithine and aspartate transcarbamoylases of *Escherichia coli*: Characterization of the active site and evidence for an interdomain carboxy-terminal helix in ornithine transcarbamoylase. *Protein Sci.* **1996**, *5*, 709–718.
- (32) Murata, L. B.; Schachman, H. K. Structural similarity between ornithine and aspartate transcarbamoylase of *Escherichia coli*: Implications for domain switching. *Protein Sci.* **1996**, *5*, 719–728.
- (33) Porter, R. W.; Modebe, M. O.; Stark, G. R. Aspartate transcarbamylase: kinetic studies of the catalytic subunit. *J. Biol. Chem.* **1969**, *244*, 1846–1859.
- (34) Legrain, C.; Stalon, V. Ornithine carbamoyltransferase from *Escherichia coli*: purification, structure and steady-state kinetic analysis. *Eur. J. Biochem.* **1976**, *63*, 289–301.
- (35) Ke, H.; Lipscomb, W. N.; Cho, Y.; Honzatko, R. B. Complex of *N*-phosphonacetyl-L-aspartate with aspartate carbamoyltransferase. *J. Mol. Biol.* **1988**, *204*, 725–747.
- (36) Legrain, C.; Demarez, M.; Glansdorff, N.; Piérard, A. Ammonia-dependent synthesis and metabolic channelling of carbamoyl phosphate in the hyperthermophilic archaeon *Pyrococcus furiosus*. *Microbiology* 1995, 141, 1093–1099.
- (37) Powers-Lee, S. G.; Mastico, R. A.; Bendayan, M. The interaction of rat liver carbamoyl phosphate synthetase and ornithine transcarbamoylase with inner mitochondrial membranes. *J. Biol. Chem.* **1987**, *262*, 15683–15688.
- (38) Eisenstein, E.; Hensley, P. Ligand binding-promoted conformational changes in yeast ornithine transcarbamoylase. J. Biol. Chem. 1986, 261, 6192–6200.
- (39) Hensley, P. Ligand binding and multienzyme complex formation between ornithine carbamoyltransferase and arginase from *Saccharomyces cerevisiae*. *Curr. Top. Cell Regul.* **1988**, *29*, 35–75.
- (40) Liu, L.; Wales, M. E.; Wild, J. R. Conversion of the allosteric regulatory patters of aspartate transcarbamoylase by exchange of a single β -strand between diverged regulatory chains. *Biochemistry* **1997**, *36*, 3126–32.
- (41) Nichols, A. GRASP: Graphical Representation and Analysis of Surface Properties, Columbia University: New York, 1993.

AR950262J