Molecular Cell, Volume 29

Supplemental Data

Crystal Structure of Bacillus stearothermophilus

UvrA Provides Insight into ATP-Modulated

Dimerization, UvrB Interaction, and DNA Binding

Danaya Pakotiprapha, Yoshihiko Inuzuka, Brian R. Bowman, Geri F. Moolenaar, Nora Goosen, David Jeruzalmi, and Gregory L. Verdine

Supplemental Experimental Procedures

Expression of Bacillus stearothermophilus UvrA, UvrB, and UvrC

The genes for *Bacillus stearothermophilus* UvrA, UvrB, and UvrC were identified by searching with the *Escherichia coli* sequences against the genome of *B. stearothermophilus* strain 10 (The University of Oklahoma Advance Center for Genome Technology). PCR was used to amplify the genes from genomic DNA, and the resulting PCR products were cloned into pET-28a (+) (Novagen) (Table S1). The expression constructs contained full-length Uvr proteins, an N-terminal His₆ tag and a thrombin cleavage site and were confirmed by sequencing. In the present work, protein residues are numbered starting from the initiator methionine in the wild-type sequence.

UvrA, UvrB, and UvrC proteins were expressed in *E. coli* BL21(DE3) pLysS. The cells were grown in LB broth at 37°C until OD₆₀₀ reached 0.5-0.6, at which point expression was induced by addition of 1 mM isopropyl- β -D-thiogalactopyranoside (IPTG). The cells were allowed to grow at 30°C for 4 hours and harvested by centrifugation. The cell pellet was resuspended in lysis buffer (50 mM NaPO₄ pH 8.0, 500 mM NaCl, 10 mM imidazole, 5 mM β -ME), flash frozen in liquid nitrogen, and stored at -80°C.

Selenomethionine-substituted UvrA was expressed in *E. coli* BL21(DE3) pLysS using a protocol adapted from Van Duyne *et al* (Van Duyne et al., 1993). Cells were

grown in M9 minimal medium at 37°C until OD₆₀₀ reached 0.5-0.6, at which point amino acid mix and selenomethionine were added. After additional 30 minutes at 30°C, expression was induced with 1 mM IPTG. The cells were allowed to grow at 30°C for 14 hours and processed as the native protein except that lysis buffer contained 15 mM β -ME.

Protein Purification

Thawed cells were supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF) and lysed by sonication. The clarified cell lysate was incubated for 45 minutes with nickelnitrilotriacetic acid (Ni-NTA) agarose beads (Qiagen). Unbound proteins were washed away with 20 mM imidazole in 50 mM NaPO₄ pH 8.0, 500 mM NaCl, 5 mM β -ME. Partially purified UvrA was eluted with 250 mM imidazole in the same buffer.

UvrA was further purified using Heparin HiTrap column (GE Healthcare) with a linear gradient of 0.3-1.2 M NaCl in Buffer A (25 mM Tris·HCl pH 7.4, 5 mM β -ME), and size exclusion chromatography (Superdex200, GE Healthcare) using 400 mM NaCl in Buffer A. Ni-NTA purified UvrB was buffer-exchanged into 100 mM NaCl in buffer A before further purified over Mono-Q (GE Healthcare) using a linear gradient of 100-600 mM NaCl in buffer A and completed using size exclusion chromatography (Superdex200, GE Healthcare) using a linear gradient of 100-600 mM NaCl in buffer A and completed using size exclusion chromatography (Superdex200, GE Healthcare) using 400 mM NaCl in Buffer A. UvrC was purified in the same way as UvrA except that 0.5-1.5 M NaCl linear gradient was used for Heparin HiTrap column and 0.5 M NaCl in Buffer A was used for the size exclusion step. Isolated domain constructs of UvrA and UvrB (UvrA131-245 and UvrB149-250) were purified by Ni-NTA and size exclusion chromatography (Superdex75, GE Healthcare, 200 mM NaCl in Buffer A).

For biochemical experiments, the proteins were dialyzed into 25 mM Tris-HCl pH 7.4, 250 mM NaCl, 20% (v/v) glycerol, 5 mM β -ME, flash frozen in liquid nitrogen and stored at -80°C.

ATPase Assay

The rate of ATP hydrolysis was measured using a coupled enzyme assay system consisting of pyruvate kinase and lactate dehydrogenase, in which ATP hydrolysis is coupled to the oxidation of NADH (Kiianitsa et al., 2003). The reaction mixture (200 µI) contained 50 mM K-HEPES pH 7.5, 150 mM potassium acetate, 8 mM magnesium acetate, 5 mM β -ME, 250 µg/ml bovine serum albumin (BSA), 0.3 mM NADH, 2 mM PEP, 1.5 U pyruvate kinase, 1.8 U lactate dehydrogenase, 50-200 mM *Bst*UvrA and 1-2 mM ATP. The mixture was incubated at 55°C for 5 minutes before addition of ATP. The reaction was followed by monitoring the decrease in A₃₄₀ over a 30-minute period using SpectraMax M5 microplate reader (Molecular Devices). The rate of ATP hydrolysis was calculated from the linear change in A₃₄₀, with correction for NADH oxidation in the absence of protein. The data are reported as mean turnover number (*k*_{cat}) (mol ATP/min/mol UvrA) ± standard error of the mean (*n*=8). The effect of DNA on ATP hydrolysis was studied by adding 0.2 µM 50-mer dsDNA containing fluorescein at the central position to the assay mixture.

Incision and Electrophoretic Mobility Shift Assays (EMSA)

The incision assay employed a 50-bp DNA fragment containing an N3-menthol lesion (Verhoeven et al., 2002) that was labeled at the 5' end of the damaged strand using $[\gamma^{-3^2}P]$ -ATP and T4 polynucleotide kinase. Damaged DNA fragment (0.2 nM) was incubated with 0.675 nM UvrA, 50 nM UvrB, and 12.5 nM UvrC in 20 µl Uvr-endo buffer (50mM Tris-HCl pH 7.5, 10 mM MgCl₂, 100 mM KCl, 0.1 µg/µl BSA, 1 mM ATP) for 15 minutes at 57°C. The reactions were terminated by using 3 µl EDTA/SDS (0.33 M EDTA, 3.3% SDS) and 2.4 µl glycogen (4 µg/µl) followed by ethanol precipitation. The incision products were visualized on a 15% denaturing polyacrylamide gel.

Loading of UvrB onto damaged DNA was measured using EMSA. Damaged DNA (0.2 nM) was incubated with 2 nM UvrA and 100 nM UvrB in 10 µl Uvr-endo buffer for 10 minutes at 57°C. The reaction mixtures were analyzed on a cooled 3.5% native polyacrylamide gel containing 1 mM ATP and 10 mM MgCl₂ in 1X Tris-borate/EDTA as described (Visse et al., 1992).

Analysis of DNA binding by UvrA mutants was carried out using EMSA. Fluorescein-containing 50-mer DNA (0.2 nM) was titrated with UvrA in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM MgCl₂, 1 mM ATP, 5% glycerol, 0.1 mg/mg BSA at 55°C for 30 minutes and then chilled on ice. The reaction mixtures were analyzed on a 6% native polyacrylamide gel containing 1 mM ATP and 10 mM MgCl₂ in 1X Tris-borate/EDTA. For the reactions with DNA containing N3-menthol lesion (Verhoeven et al., 2002), UvrA was preincubated at 57°C in the Uvr-endo buffer for 5 minutes prior to addition of DNA. The reactions were incubated at 57°C for 10 minutes, chilled on ice, and analyzed on a cooled 3.5% native polyacrylamide gel containing 1 mM ATP and 10 mM MgCl₂ in 1X Trisborate/EDTA.

Supplemental References

Kiianitsa, K., Solinger, J.A., and Heyer, W.D. (2003). NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities. Anal Biochem *321*, 266-271.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22, 4673-4680.

- Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L., and Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. J Mol Biol 229, 105-124.
- Verhoeven, E.E., van Kesteren, M., Turner, J.J., van der Marel, G.A., van Boom, J.H.,
 Moolenaar, G.F., and Goosen, N. (2002). The C-terminal region of Escherichia coli
 UvrC contributes to the flexibility of the UvrABC nucleotide excision repair system.
 Nucleic Acids Res *30*, 2492-2500.
- Visse, R., de Ruijter, M., Moolenaar, G.F., and van de Putte, P. (1992). Analysis of UvrABC endonuclease reaction intermediates on cisplatin-damaged DNA using mobility shift gel electrophoresis. J Biol Chem 267, 6736-6742.



PakotipraphaS1.tif

Figure S1. Structural comparison of the nucleotide-binding domains (NBDs) of *Bst*UvrA and *E.coli* MalK. NBD-I and NBD-II of UvrA were superimposed with the NBD of the maltose transporter MalK (ADP-bound, PDB code 2AWO) The conserved ATPase motifs are colored as follows: Walker A/P-loop, green; Walker B and D-loop, orange; ABC signature motif, blue; Q-loop, magenta; and H-loop/switch, cyan. For clarity, the structures are shown side-by-side. The bound ADP molecules are shown as sticks and the Zn atoms in the UvrA structure are shown as spheres. The UvrB-binding domain and the insertion domain, which are inserted in the NBD-I of UvrA, are shown in pale yellow and pale green, respectively.



PakotipraphaS2.tif

Figure S2. SDS gel electrophoretic analysis of size exclusion chromatography of the following: wild-type UvrA·UvrB or UvrA_{Δ 131-245}·UvrB. We observe that only wild-type UvrA complexed with UvrB, whereas UvrA_{Δ 131-245} failed to interact with UvrB. The UV absorption profiles are shown in Figure 5C.



PakotipraphaS3.tif

Figure S3. Zinc structural modules in UvrA. Each Zn structural module of UvrA (green) is shown in an overlay with its closest structural neighbor (gray). The zinc atoms in UvrA are colored magenta. (A) UvrA Zn module 1 and the Zn module from YfgJ (PDB code 2JNE), RMSD 0.798 Å. (B) and (C) UvrA Zn modules 2 and 3 and the Zn ribbon 2 from DnaJ (PDB code 1EXK), RMSD 0.655 and 0.574 Å, respectively. Residue numbers are indicated for UvrA. The amino and carboxy termini are labeled. The RMSD was calculated using Zn; S_γ, C_β, and C_α (for cysteine residues); N_δ, C_γ, C_β, and C_α (for histidine residues); as well as C_α of the intervening residues.

	β1	β:	2 β3		α1		α2	β4	β5	α.3	
BstUvrA	MDKIIVKGAŘ	AHNLKNIDVÊ	IPRGKLVVLŤ	GESGSGKSSL	AFDTIYAEGÔ	RRYVESLSAŸ	AROFLGOMEŘ	PDVDAIEGLŠ	PAISIDOKTT	SRNPRSTVGT	100
BcaUvrA	MDKIVVKGAR	AHNLKNIDVE	IPRGKLVVLT	GLSGSGKSSL	AFDTIYAEGQ	RRYVESLSAY	ARQFLGQMEK	PDVDAIEGLS	PAISIDOKTT	SRNPRSTVGT	100
<i>Eco</i> UvrA	MDKIEVRGAR	THNLKNINLV	IPRDKLIVVT	GLSGSGKSSL	AFDTLYAEGQ	RRYVESLSAY	ARQFLSLMEK	PDVDHIEGLS	PAISIEQKST	SHNPRSTVGT	100
TthUvrA	MDRIVIRGAR	EHNLKNISLE	LPRGKFIVIT	GVSGSGKST <mark>L</mark>	AFDTIYAEGQ	RRYVESLSSY	ARQFLGVMDK	PEVESIEGLS	PAISIDQKTT	SHNPRSTVGT	100
TmaUvrA	MNEIVVKGAR	VHNLKNITVR	IPKNRLVVIT	GVSGSGKSSL	AMDTIYAEGQ	RRYLESLSTY	ARQFLGNLKK	PDVDEIEGLS	PAIAIDOKTV	SHNPRSTVGT	100
	α.4 β.6 α.5					87					
					p	→					
BstUvrA	VTEIYDYLRĽ	LFARIGRPIČ	PT <mark>H</mark> GIEIQSQ	TIEQMVDRLL	S.YPERTKMQ	ĪLAPIV SGKK	ĞTHAKTLEDI	ŘKQGYVRVRI	DREMRELTGD	IELEKNKK	197
BcaUvrA	VTEIYDYLRL	LFARIGRPVC	PTHGIEIQSQ	TIEQMVDRLL	A. YPERTKMQ	ILAPIVSGKK	GTHAKTLEDI	RKQGYVRVRI	DGEMRELTED	I. ELEKNKK	197
Tthura	VTETHDYLEL	LEARVGOARC	PERCEPTERO	SASETTDRUL	K PPDGTRAT	LMAPLIKERK	GEVEKLEOOL	LKEGVARVRV	DGVTVLLEEA	OGLSLEKVEK	197
TmaUvrA	VTEIYDYLRV	LYARIGKAHC	PECGRPLEKK	SIDEILQDLF	NSFKEGSRIY	ILAPVATEKK	GTFKKEIEEF	ISKGFARIEI	DGEIYRLEEV	P ELDKNKR	198
		2	'n 1								
	β8		α6	β9	β 10		α.7	α 8	β 11	θ.9	
BstUvrA	HSİDVVVDRI	IIÅDGIAARL	ADŠLETALKL	ADĞKVVVDVI	GEĜELLF	SEKHACPYCG	FSIGELEPRL	FSFNSPFGAC	pd <mark>C</mark> dglgakl	EVDLDLVIPN	294
BcaUvrA	HSIDVVVDRI	IIKDGIAARL	ADSLETALKL	ADGKVVVDVI	GEGELLF	SEKHACPYCG	FSIGELEPRL	FSFNSPFGAC	PDCDGLGAKL	EVDLDLVIPN	294
EcoUvrA	HTIEVVVDRF	KVRDDLTQRL	AESFETALEL	SGGTAVVADM	DDPKAEELLF	SANFACPICG	YSMRELEPRL	FSFNNPAGAC	PTCDGLGVQQ	YFDPDRVIQN	297
TthUvrA	HDIDLVIDRV	VLKEEERPRI	AEAVELALLR	GEGLLRVLYP	DTGEEELF	SEKFACPEHG	SVLEELEPRI	FSFNSPYGAC	PACSGLGYRQ	EFDPELVV.N	296
ImauvrA	HIVKLVVDRL	ILETRNEHRI	LDSLELAMKE	GKGFVEIKNV	DIG. ESKIF	Zn 1	IGFPEITPKL	FSFNSPIGAC	In 2	EVDPSLVI.D	295
	α 10			α 11	α 12	β 12	β13	α 13	α 14		
RatHvrA	DELTLŘEHAT	APWEPÔSSO	VYPOLLÊA	VCRHYGIPMD	VEVKDLEKÊO	LDKTLYGSÅG	EPIVERYTND	FGOVREOVTA	FEGUTENVÊR	RVRETSSDŸT	391
BcaUvrA	DELTLKEHAI	APWEPQSSQ.	YYPQLLEA	VCRHYGIPMD	VPVRDLPKEQ	LDKILYGSGG	EPIYFRYTND	FGQVREQVIA	FEGVIPNVER	RYRETSSDYI	391
<i>Eco</i> UvrA	PELSLAGGAI	RGWD.RRNF.	YYFQMLKS	LADHYKFDVE	APWGSLSANV	HKVVLYGSGK	ENIEFKYMND	RGDTSIRRHP	FEGVLHNMER	RYKETESSAV	393
TthUvrA	PELSLAEGAI	LPWSRGRDTG	RSYLWDRLRA	LAEHLGFDLK	TPFKDLPEEA	KRAVLYGLP.	EPFEVVFRRG	GKETFRVEVR	YEGVIPWLEK	RYQESDSEGV	395
TmaUvrA	EEKSVLEGAI	IPYRWDRRLS	RWVAR	EIEKRGVSPH	LPFKDLPEDV	KEFILYGDD.		R	FEGVVPKVQR	WHRETESPEM	370
	β	14	β 15	α 15	α 16	α	17	α 18	_	α 19	
RatHurb	REOMEK VMÅ	FOPOPTCOGY	RI.KKESI.AVÎ.	VGGKHIGEVŤ	AMSVTEALAP	FDGLELTEK	ÊAOTARI.TI.	RÊTRORLGRI.	ONVGLOVI.TL	SRSAGTISGG	488
BcaUvrA	REOMEK . YMA	EOPCPTCOGY	RLKKESLAVL	VGGKHIGEVT	AMSVTEALAF	FDGLELTEK.	EAOIARLIL	REIRDRLGFL	ONVGLDYLTL	SRSAGTLSGG	488
<i>Eco</i> UvrA	REELAK.FIS	NRPCASCEGT	RLRREARHVY	VENTPLPAIS	DMSIGHAMEF	FNNLKLAGQ.	RAKIAEKIL	KEIGDRLKFL	VNVGLNYLTL	SRSAETLSGG	490
TthUvrA	REALEG.FMS	LRP <mark>C</mark> PACGGT	RYKREVLSVK	VAGRNIAEVS	ALPVREALAF	FQGLEKTLPP	FQAQIARPIL	REIVERLGFL	VDVGLDYLTL	DRAANTLSGG	494
TmaUvrA	KEWLEKNFIV	QRTCSVCGGR Zn 2	RLNREALSVK	INGLNIHEFT	ELSISEELEF	LKNLNLTER.	.EREIVGELL	KEIEKRLEFL	VDVGLEYLTL	SRSATTLSGG ABC signature	468 e-l
	α 20		3 16	α.21		β17, α2	2β18	α 23	β19 α24	α2	5
Datitur?		TOOPI TOULY	W DEDCTOLU	OPDNDRLTAT	LÉCMEDICNE	LTUUEEDEDT	MAADVLIDT		100A A CTEREN	MUDDING LTTC	507
BCaUVIA	FAORIRLATO	IGSRLTGVLY	VLDEPSIGLH	ORDNDRLIAT	LKSMRDLGNT	LIVVEHDEDT	MLAADYLIDI	GPGAGIHGGE	VISAGTPEEV	MEDPNS.LTG	587
EcoUvrA	EAQRIRLASQ	IGAGLVGVMY	VLDEPSIGLH	QRDNERLLGT	LIHLRDLGNT	VIVVEHDEDA	IRAADHVIDI	GPGAGVHGGE	VVAEGPLEAI	MAVPES.LTG	589
TthUvrA	EAQRIRLATQ	VGSGLTGVLY	VLDEPSIGLH	PRDNQRLIRT	LKRLRDLGNT	LIVVEHDEET	MRAADWIVDM	GPGAGIHGG <mark>E</mark>	VVAQGTLEDI	LKSPQS.LTG	593
TmaUvrA	ESQRIRLATQ	IGSGLTGV <mark>IY</mark>	VLDEPTIGLH	PRDTERLIKT	LKKLRDLGNT	VIVVEHDEEV H-loop-l	IRNADHIIDI	GPGGGTNGGR Gly-rich loop-l	VVFQGTVDEL	LKNPDSSLTG	568
		B 20 . B 2			1 β 22		α 26		β 230	<u>ε 27</u> β 24	
D = +17 3				most upper up			•		O		
BCallyra	QILSGKKFIP SVLSGKKFIP	LPAKKRRPD.		MINDLE NUSSAK	A A A A A A A A A A A A A A A A A A A		TRATING TO STATE T	OVI U DEVICE	DODUBDITDOT	BUT DUTTOTO	Ac. 12 Sc.
EcoUvrA	DITTO OLULE TE	L.P.P.F.R.K.PD	GRWLEVVGAR	FHNLKNVSVK	TPLGTEVAVT	GVSGSGKSTL GVSGSGKSTL	VNEVLYKALA VNEVLYKALA	QKLH. RAKAK	PGEHRDIRGL PGEHRDIRGL	EHLDKVIDID	685
TthUvrA	OYMSGKRKIE	LPPERRKPD. VPKKRVPANP	GRYIEIKGAR EKVLKLTGAR	EHNLKNVSVK GNNLKDVTLT	IPLGTFVAVT IPLGTFVAVT LPVGLFTCIT	GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL	VNEVLYKALA VNEVLYKALA INDTLFPIAO	QKLH.RAKAK QKLH.RAKAK ROLNGATIAE	PGEHŘDIRGL PGEHRDIRGL PAPYRDIOGL	EHLDŘVIDID EHLDKVIDID EHFDKVIDID	685 689
	QYMSGKRKIE AYLRGEKRIP	LPPERRKPD. VPKKRVPANP VPKERRKGN.	GRYIEIKGAR EKVLKLTGAR GKWLVLKGAR	EHNLKNVSVK GNNLKDVTLT AHNLKNVTLR	IPLGTFVAVT IPLGTFVAVT LPVGLFTCIT IPLGRFVAIT	GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GPSGSGKSTL	VNEVLYKALA VNEVLYKALA INDTLFPIAQ VHDVLYAALA	QKLH.RAKAK QKLH.RAKAK RQLNGATIAE QRLM.RAKTT	PGEHŘDIRGL PGEHRDIRGL PAPYRDIQGL PGPYEALEGV	EHLDŘVIDID EHLDKVIDID EHFDKVIDID EHLDKVIEID	685 689 691
TmaUvrA	QYMSGKRKIE AYLRGEKRIP EYLSGKRKIT	LPPERRKPD. VPKKRVPANP VPKERRKGN. VNKTRRLPY.	GRWLEVVGAR GRYIEIKGAR EKVLKLTGAR GKWLVLKGAR .ASLKIKGVR	EHNLKNVSVK GNNLKDVTLT AHNLKNVTLR HNNLKNIDVE omatic-II	IPLGTFVAVT IPLGTFVAVT LPVGLFTCIT IPLGRFVAIT IPLGVFVCVT	GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GPSGSGKSTL GVSGSGKSSL Walker A-II	VNEVLYKALA VNEVLYKALA INDTLFPIAQ VHDVLYAALA VMETLYPALM	QKLĤ. RAKAK QKLH. RAKAK RQLNGATIAE QRLM. RAKTT NLLH. KTKLP	PGEHRDIRGL PGEHRDIRGL PAPYRDIQGL PGPYEALEGV AGEFDSIEGH	EHLDŘVIDID EHLDKVIDID EHFDKVIDID EHLDKVIEID ENIDKMIAID	685 689 691 665
TmaUvrA	QYMSGKRKIE AYLRGEKRIP EYLSGKRKIT	LPPERRKPD. VPKKRVPANP VPKERRKGN. VNKTRRLPY.	GRVIEVVGAR GRVIEIKGAR EKVLKLTGAR GKWLVLKGAR .ASLKIKGVR AI	EHNLKNVSVK GNNLKDVTLT AHNLKNVTLR HNNLKNIDVE romatic-II α 30	IPLGIFVAVI IPLGTFVAVI LPVGLFTCIT IPLGRFVAIT IPLGVFVCVT	GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GPSGSGKSTL <u>GVSGSGKSS</u> L Walker A-II β 2	VNEVLYKALA VNEVLYKALA INDTLFPIAQ VHDVLYAALA VMETLYPALM	QKLĤ.RAKAK QKLH.RAKAK RQLNGATIAE QRLM.RAKTT NLLH.KTKLP	PGEHRDIRGL PGEHRDIRGL PAPYRDIQGL PGPYEALEGV AGEFDSIEGH	EHLDŘVIDID EHLDKVIDID EHFDKVIDID EHLDKVIEID ENIDKMIAID	685 689 691 665
TmaUvrA	QYMSGKRKIE AYLRGEKRIP EYLSGKRKIT	LPPERRKPD. VPKKRVPANP VPKERRKGN. VNKTRRLPY.	GRWLEVVGAR GRVIEIKGAR EKVLKLTGAR GKWLVLKGAR .ASLKIKGVR Ar	EHNLKNVSVK GNNLKDVTLT AHNLKNVTLR HNNLKNIDVE comatic-II α 30	IPLGIFVAVT IPLGTFVAVT LPVGLFTCIT IPLGRFVAIT IPLGVFVCVT	GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GPSGSGKSTL GVSGSGKSSL Walker A-II β2	VNEVLYKALA VNEVLYKALA INDTLFPIAQ VHDVLYAALA VMETLYPALM	QKLÅ. RAKAK QKLH. RAKAK RQLNGATIAE QRLM. RAKTT NLLH. KTKLP	PGEHŘDIRGL PGEHRDIRGL PAPYRDIQGL PGPYEALEGV AGEFDSIEGH	EHLDŘVIDID EHLDKVIDID EHFDKVIDID EHLDKVIEID ENIDKMIAID	685 689 691 665
TmaUvrA BstUvrA BcaUvrA	QYMSGKRKIE AYLRGEKRIP EYLSGKRKIT OSPIĞRTPRS	LPPERRKPD. VPKKRVPANP VPKERRKGN. VNKTRRLPY.	GRVIEKVGAR EKVLKLTGAR GKWLVLKGAR .ASLKIKGVR A DIRDVFASTN DIRDVFASTN	EHNLKNVSVK GNNLKDVTLT AHNLKNVTLR HNNLKNIDVE omatic-II a 30 EAKVRGYKKG EAKVRGYKKG	IPLGIFVAVT IPLGFVAVT LPVGLFTCIT IPLGVFVCVT RFSFŇVKGGR	GVSGSGKSTL GVSGSGKSTL GPSGSGKSTL GPSGSGKSTL GPSGSGKSSL Walker A-II β2 CEACHGDGII GEACHGDGII	VNEVLYKALA VNEVLYKALA INDTLFPIAQ VHDVLYAALA VMETLYPALM 5 KIEMÅFLPDV KIEMÅFLPDV	QKLH.RAKAK QKLH.RAKAK RQLNGATIAE QRLM.RAKTT NLLH.KTKLP 326 YVPCEVCHGK	PGEHŘDIRGL PGEHRDIRGL PGPYRDIQGL PGPYEALEGV AGEFDSIEGH & 31 RYNRĚTLEVT RYNRĚTLEVT	EHLDŘVIDID EHLDKVIDID EHPDKVIDID EHLDKVIEID ENIDKMIAID VKGŘÍAEVL VKGŘÍAEVL	685 689 691 665 785 785
TmaUvrA BstUvrA BcaUvrA EcoUvrA	QYMSGKRKIE AYLRGEKRIP EYLSGKRKIT OSPIĞRTPRS OSPIGRTPRS	LPPERRKPD. VPKKRVPANP VPKERRKGN. VNKTRRLPY. Φ 28 Φ PATΥTGVPD NPATΥTGVPD NPATYTGVPT	GRVIEKUGAR EKVLKLTGAR GKWLVLKGAR .ASLKIKGVR A 0 29 DIRDŮFASTN DIRDŮFASTN DIRDVFASTN PVRELFAGVP	EHNLKNVSVK GNNLKDVTLT AHNLKNVTLR HNNLKNIDVE romatic-II α 30 EAKVŘGYKKG EAKVRGYKKG ESRARGYTPG	IPLGIFVAVT IPLGFVAVT LPVGLFTCIT IPLGVFVAVT IPLGVFVCVT RFSFNVKGGR RFSFNVKGGR	GVSGSGKSTL GVSGSGKSTL GPSGSGKSTL GPSGSGKSTL GPSGSGKSSL Walker A-II β2 CEACHGDGII CEACHGDGII CEACHGDGII CEACGGDGVI	VNEVLYKALA VNEVLYKALA INDTLFPIAQ VHDVLYAALA VMETLYPALM 5 KIEMÅFLPDV KIEMÅFLPDV KVEMHFLPDI	QKLH.RAKAK QKLH.RAKAK RQLNGATIAE QRLM.RAKTT NLLH.KTKLP 326 YVPCÉVCHGK YVPCÉVCHGK YVPCQCKGK	PGEHŘDIRGL PGEHRDIRGL PAPYRDIQGL PGPYEALEGV AGEFDSIEGH «31 RYNRĚTLEVT RYNRĚTLEVT RYNRETLEVT	EHLDŘVIDID EHLDKVIDID EHPDKVIDID EHLDKVIEID ENIDKMIAID VKGKŇIAEVL YKGKNIAEVL YKGKNIAEVL	685 689 691 665 785 785 789
TmaUvrA BstUvrA BcaUvrA EcoUvrA TthUvrA	QYMSGKRKIE AYLRGERRIP EYLSGKRKIT OSPIĞRTPRS OSPIGRTPRS OSPIGRTPRS	LPPERRKPD. VPKKRVPANP VPKERRKGN. VNKTRRLPY.	GRVIEIKGAR EKVLKLTGAR GKWLVLKGAR .ASLKIKGVR Δ α 29 0 DIRDŮFASTN PVRELFAGVP EIRDLFAKTP	EHNLKNVSVK GNNLKDVTLT AHNLKNVTLR HNNLKNIDVE romatic-il α 30 EAKVŘGYKKG EAKVRGYKKG ESRARGYTPG EARKRGYGPG	IPLGIFVAVT IPLGIFVAVT IPLGRFVAIT IPLGVFVCVT RFSFŇVKGGR RFSFNVKGGR RFSFNVKGGR	GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSSL Walker A-II β2 CEACHGDGII CEACHGDGII CEACGGDGVI CEACGGDGVI	VNEVLYKALA VNEVLYKALA INDTLFPIAQ VHDVLYAALA VMETLYPALM S KIEMÅFLPDV KVEMHFLPDU KVEMHFLPDI	QKLÅ. RAKAK QKLH. RAKAK RQLNGATIAE QRLM. RAKTT NLLH. KTKLP 326 YVPCÉVCHGK YVPCOQCKGK YVPCOQCKGK	PGEHŘDIRGL PGEHRDIRGL PAPYRDIQGL PGPYEALEGV AGEFDSIEGH RYNRĚTLEVT RYNRETLEVT RYNRETLEVK	EHLDŘVIDID EHLDKVIDID EHPDKVIDID EHLDKVIEID ENIDKMIAID YKGŘÍAEVL YKGKŇIAEVL YKGKNIAEVL JKGKSIADVL	685 689 691 665 785 785 789 791
TmaUvrA BstUvrA BcaUvrA EcoUvrA TthUvrA TmaUvrA	QYMSGKRKIE AYLRGEKRIP EYLSGKRKIT OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS	LPPERRKPD. VPKKRVPANP VPKERRKGN. VNKTRRLPY. NPATYTGVPD NPATYTGVPD NPATYTGVPD NPATYTGVPD NPATYTKVPD	GRVIEIKGAR GKVLEIKGAR GKWLVLKGAR .ASLKIKGVR AI (29 DIRDVPASTN DIRDVPASTN DVRELPAGVP EIRDLPAKTP EIRSLFAMTP	EHNLKNVSVK GINLKDVTLT AHNLKNVTLR HINLKNVTLR UNDELEN EAKVŘGYKKG EAKVRGYKKG EARVRGYKKG EARKRGYGPG AAKARGYNKS	IPLGTFVAVT IPLGTFVAVT IPLGRFVAIT IPLGVFVCVT RFSFNVKGGR RFSFNVKGGR RFSFNVKGGR	GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSSL Walker A-II ΔΕΑΦΗGDGII ΔΕΑΦΗGDGII ΔΕΑΦΗGDGIY ΔΕΑΦGGDGVI ΔΕΑΦGGQGVY ΔΕΑΦGQQVY ΔΑ	VNEVLYKALA INDTLPFIAQ VHDVLYKALA INDTLPFIAQ VHDVLYAALA VMETLYPALM 5 5 KIEMÅPLPDV KIEMÅPLPDU KIEMLFLPDL KIEMLFLPDV	QKLÅ. RAKAK QKLH. RAKAK RQLNGATIAE QRLM. RAKTT NLLH. KTKLP 326 YVPCEVCHGK YVPCEVCHGK YVPCOQCKGK YVPCOQCKGK YVPCOVCKGK 70 3	PGEHŘDIRGL PGEHRDIRGL PAPYRDIQGL PGEYEALEGV AGEPDSIEGH AGEPDSIEGH RYNRĚTLEVT RYNRĚTLEVT RYNRĚTLEIK RYNKETLEIK	EHLDŘVIDID EHLDKVIDID EHPDKVIDID ENIDKMIAID ENIDKMIAID YKGKŇIAEVL YKGKNIAEVL LRGKSIADVL YKGKNISDIL	685 689 691 665 785 785 789 791 765
TmaUvrA BstUvrA BcaUvrA EcoUvrA TthUvrA TmaUvrA	QYMSGKRKIE AYLRGEKRIP EYLSGKRKIT OSPIĞRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS QSPIGRTPRS QSPIGRTPRS QSPIGRTPRS	LPPERRKPD. VPKERRKGN. VNKTRRLPY. A28 NPATYTGVPD NPATYTGVPD NPATYTGVPT NPATYTGVFD NPATYTKVFD	GRVIELKAAR EKVLKLTGAR GKWLKLTGAR GKWLVLKAGAR ASLKIKGVR AU 229 DIRDVPASTN DIRDVPASTN DIRDVPASTN PVRELPAGVP EIRDLPAKTP EIRSLFAMTP	EHNLKINSVK GINLKDVTLT AHNLKDVTLT AHNLKNVTLR MINLKNUTVE romatic-II a 30 EAKVRGYKKG EAKVRGYKKG EARKRGYGPG AAKARGYNKS a 35	IPLGTPVAVT IPLGTPVAVT IPLGRFVAIT IPLGVFVCVT RFSFNVKGGR RFSFNVKGGR RFSFNVKGGR RFSFNVKGGR	GVSGSGKSTL GVSGSGKSTL GPSGSGKSTL GPSGSGKSTL GPSGSGKSSL Walker A-II β2 CEACHGGII CEACHGGII CEACGGCGVI CEACGGCGVI CEACGGCGVV CEACGGCGVV Zn 3	VNEVLYKALA VNEVLYKALA INDTLPFIAQ VHDVLYAALA VMETLYPALM 5 5 KIEMÅFLPDV KVEMÅFLPDV KVEMÅFLPDV KIEMLFLPDL KIEMLFLPDV β 27.	QKLÅ. RAKAK QKLH. RAKAK RQLNGATIAE QRLM. RAKTT NLLH. KTKLP 328 YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCKGK YVPCEVCKGK ZN 3	PGEHÅDIRGL PGEHÅDIRGL PAPYRDIQGL PGPYEALEGV AGEPDSIEGH «31 RYNRĚTLEVT RYNRĚTLEVT RYNRĚTLEVK RYNRETLEIT «37	EHLDŘVIDID EHLDKVIDID EHPDKVIDID EHLDKVIEID ENIDKMIAID YKGKNIAEVL YKGKNIAEVL LRGKSIADVL YKGKNISDIL B28	685 689 691 665 785 785 789 791 765
TmaUvrA BstUvrA BcaUvrA EcoUvrA TthUvrA TmaUvrA	QYMSGKRKIE AYLRGERRIP EYLSGKRKIT OSPIĞRTPRS OSPIGRTPRS OSPIGRTPRS QSPIGRTPRS QSPIGRTPRS Q-loop-II	LPPERRKPD. VPKERRKGN. VNKTRRLPY. «28 NPATYTGVPD NPATYTGVPD NPATYTGVPD NPATYTGVPD NPATYTKVPD	GRVIELKAAR EKVLKLTGAR GKWLLKAAR .ASLKIKGVR 4 29 DIRDŮPASTN DIRDŮPASTN DIRDŮPASTN PVRELPAGVP EIRDLPAKTP EIRSLFAMTP	EHNLKINSVK GINILKDVTLT AHNLKDVTLT AHNLKNVTLR HINLKNIDVE romatic-II a 30 EAKVÄGYKKG ESRARGYTPG EARKRGYTPG AAKARGYNKS a 35	IPLGTPVAVT IPLGTPVAVT IPLGRFVAIT IPLGVFVCVT RFSPŇVKGGR RFSFNVKGGR RFSFNVKGGR RFSFNLKGGR	GVSGSGKSTL GVSGSGKSTL GPSGSGKSTL GPSGSGKSTL GPSGSGKSTL Maiker A-II β2 CEACHGDGII GEACHGDGII GEACHGDGII GEACGGGGVV CEACGGGGVV Zn 3	VNEVLYKALA INDTLFPIAQ VHDVLYAALA VMETLYPAAM 5 KIEMÅFLPDV KIEMÅFLPDV KIEMLFLPDV KIEMLFLPDV KIEMLFLPDV	QKLÅ. RAKAK QKLH. RAKAK RQLNGATIAE QRLM. RAKTT NLLH. KTKLP 326 YVPCEVCHGK YVPCEVCHGK YVPCEQCKGK YVPCEQCKGK YVPCEVCKGK Zn 3	PGEHÅDIRGL PGEHÅDIRGL PAPYRDIQGL PGPYEALEGV AGEPDSIEGH «31 RYNRĚTLEVT RYNRĚTLEVT RYNRĚTLEIK RYNKETLEVK RYNRETLEIT «37	EHLDŘVIDID EHLDKVIDID EHPDKVIDID EHLDKVIEID ENIDKMIAID 2432 YKGKNIAEVL YKGKNIAEVL LRGKSIADVL YKGKNISDIL β28	685 689 691 665 785 785 789 791 765
TmaUvrA BstUvrA EcaUvrA EcoUvrA TthUvrA TmaUvrA BstUvrA BstUvrA	QYMSGKRKIE AYLRGERRIP EYLSGKRKIT OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS QSPIGRTPRS QSPIGRTPRS QSPIGRTPRS DMTVEDALDP DMTVEDALDP	LPPERRKPD. VPKERRKGN. VNKTRRLPY. VNKTRRLPY. VNKTRRLPY. NPATYTGVPD NPATYTGVPD NPATYTGVPD NPATYTKVPD	GRVIEIKGAR EKVLKLTGAR GKWLKLTGAR GKWLVLKGAR .ASLKIKGVR AI 22 DIRDŮPASTN DIRDŮPASTN PVRELPASTN PVRELPASTN EIRDLPAKTP EIRSLPAMTP	EHNLKINSVK GINILKDVTLT AHNLKDVTLT AHNLKNVTLR HINLKNIDVE romatic-II 4 30 EAKVÄGYKKG EARVRGYKKG ESRARGYTPG AAKARGYNKS 4 35 YMKLGORATT	IPLGTPVAVT IPLGTPVAVT IPLGTPVAVT IPLGVPVCVT IPLGVPVCVT RFSFNVKGGR RFSFNVKGGR RFSFNVKGGR RFSFNVKGGR RFSFNLKGGR LSGGEAQPVK	GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GPSGSGKSTL GPSGGKSGKSTL GPSGGKSGKSTL GPSGGKSTL GPSGGKSGKSTL GPSGGKSTL GPSGGKSGKSTL GPSGGKSGKSTL GPSGGKSGKSTL GPSGGKSGKSTL GPSGGKSGKSTL GPSGGKSGKSTL GPSGGKSGKSTL GPSGGKSGKSTL GPSGGKSGKSTL GPSGGKSGKSTL GPSGGKSGKSTL GPSGGKSGKSTL GPSGGKSGKSTL GPSGGKSGKSTL GPSGGKSGKSTL GPSGKSTL GPSGKSTL GP	VNEVLYKALA VNEVLYKALA INDTLFPIAQ VHDVLYAALA VMETLYPALM 5 KIEMÅFLPDV KVEMHFLPDI KIEMLFLPDV KIEMLFLPDV KIEMLFLPDV RTLFVILDEP QRTLFVILDEP	QKLÅ. RAKAK QKLH. RAKAK RQLMGATTAE QRLM. RAKTT NLLH. KTKLP 326 YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK Zn 3	PGEHŘDIRGL PGBHRDIRGL PGPYEALEGV AGEPDSIEGH AGEPDSIEGH AGIT RYNRĚTLEVT RYNRĚTLEVT RYNRĚTLEVT RYNRĚTLEVT RYNRĚTLEVT A 37 RLLDÝLHRLY	EHLDŘVIDID EHLDKVIDID EHPDKVIDID EHPDKVIDID ENIDKMIAID YKGKNIAEVL YKGKNIAEVL YKGKNIAEVL YKGKNIAEVL YKGKNIAEVL YKGKNIADIL BAGDÝUVUE	685 689 691 665 785 785 789 791 765
TmaUvrA BstUvrA EcoUvrA TthUvrA TmaUvrA BstUvrA BcaUvrA EcoUvrA	QYMSGKRKIE AYLRGERRIP EYLSGKRKIT OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS O-loop-II OMTVEDALDP DMTVEDALDP DMTVEDALDP	LPPERRKPD. VPKERKGN. VPKERKGN. VNKTRRLPY. «28 NPATYTGVPD NPATYTGVPD NPATYTGVPD NPATYTGVPD NPATYTKVFD PASIPKIKRK PASIPKIKRK PASIPKIKRK	GRVIEIKGAR EKVLKLTGAR GKWLKLTGAR GKWLVLKGAR .ASLKIKGVR AI 0 IRDVPASTN DIRDVPASTN DVRLFAGVP EIRDLFAGVP EIRDLFAGVP EIRSLFAMTP 34 LETLÝDVGLG LETLÝDVGLG LETLÝDVGLG	AHNLKNVSVK GNNLKDVTLT AHNLKNVTLR HNNLKNIDVE romatic-II a 30 EAKVÄGYKKG EARVRGYKKG EARKRGYAPG AAKARGYNKS AAKARGYNKS YMKLÅQPATT YMKLÅQPATT YMKLÅQPATT	IPLGTPVAVT IPLGTPVAVT IPLGRPVAIT IPLGVPVCVT IPLGVPVCVT RPSPNVKGGR RPSPNVKGGR RPSPNVKGGR RPSPNVKGGR RFSPNLKGGR LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK	GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GEAGHGDGII GEAGHGDGII GEAGHGDGII GEAGGGGVI GEAGGGGVV Zn 3 LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRSN	VNEVLYKALA VNEVLYKALA INDTLFPIAQ VHDVLYAALA VMETLYPALM 5 KIEMHPLPDV KIEMHPLPDV KIEMLPLPDU B 27 GRTLVILDEP GRTLVILDEP GQTLVILDEP	QKLÅ. RAKAK QKLH. RAKAK RQLMAATIAE QRLM. RAKTT NLLH. KTKLP 326 УVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCKGK Zn 3 TTGLÅVDDIA TTGLÅVDDIA TTGLÅVDDIA	PGEHŘDIRGL PGBHRDIRGL PGPYEALEGV AGEPDSIEGH a31 RYNRĚTLEVT RYNRETLEVT RYNRETLEIK RYNRETLEIK RYNRETLEIT a37 RLLDŮLHRLW RLLDŮLHRLW	EHLDŘVIDID EHLDŘVIDID EHPDKVIDID ENDKMIAID ENDKMIAID VKGKNIAEVL YKGKNIAEVL YKGKNIAEVL YKGKNISDIL β28 DNGDŤVLVIE DNGDŤVLVIE DNGDŤVLVIE	685 689 691 665 785 785 789 791 765 885 885
TmaUvrA BstUvrA EcoUvrA TthUvrA TmaUvrA BstUvrA BcaUvrA EcoUvrA TthUvrA	QYMSGKRKIE AYLRGERRIP EYLSGKRKIT OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS O-loop-II DMTVEDALDF DMTVEDALDF DMTVEEAREF DMTVEEAREF	LPPERRKPD. VPKERVPANP VPKERKGN. VNKTRRLPY. 4.28 NPATYTGVPD NPATYTGVPD NPATYTGVPD NPATYTGVPD NPATYTKVFD PASIPKIKKK PASIPKIKKK PDAVPALARK	GRVIEIKGAR GKVLEIKGAR GKVLKLTGAR GKWLVLKGAR ASLKIKGVR AU COMPASIN DIRDVPASIN DIRDVPASIN DIRDVPASIN DIRDVPASIN PVRELPAVP EIRSLFAMTP 34 LETLÝDVGLG LQTLMVVGLG LQLMVDVGLG	AND	IPLGTPVAVT IPLGTPVAVT IPLGVPVAVT IPLGVPVCVT IPLGVPVCVT RFSFNVKGGR RFSFNVKGGR RFSFNVKGGR RFSFNLKGGR LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK	GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GEACHGDGII CEACHGDGII CEACHGDGII CEACGGGVI CEACGGGVV Zn 3 LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN	VNEVLYKALA INDTLFPIAQ VHDVLYKALA INDTLFPIAQ VHDVLYKALA VMETLYPALM 5 5 KIEMHFLPDV KIEMHFLPDV KIEMLFLPDV B 27 GRTLYILDEP GRTLYILDEP GRTLYILDEP	QKLÅ. RAKAK QKLH. RAKAK RQLMAATTAE QRLM. RAKTT NLLH. KTKLP 326 YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCKGK Zn 3 TTGLÅVDDIA TTGLÅVDDIA TTGLHADDIA	PGEHŘDIRGL PGBHRDIRGL PGPYEALEGV AGEPDSIEGH AGEPDSIEGH RYNRĚTLEVT RYNRETLEVT RYNRETLEIT A 37 RLLDŮLHRLV RLLDŮLHRLV RLLDVLHRLV	EHLDŘVIDID EHLDŘVIDID EHPDKVIDID ENDKMIAID ENIDKMIAID VKGKNIAEVL YKGKNIAEVL YKGKNIAEVL YKGKNISDIL B28 DNGDŤVLVIE DNGDŤVLVIE DNGDŤVLVIE DQGNTIVVIE	685 685 691 665 785 785 785 789 791 765 885 885 885 885 889
TmaUvrA BstUvrA BcaUvrA ThAUvrA TmaUvrA BstUvrA BcaUvrA BcaUvrA ThUvrA TmaUvrA	QYMSGKRKIE AYLRGERRIP EYLSGKRKIT OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS C-loop-II OMTVÉDALDF DMTVÉDALDF DMTVEEAREF DMTVEEAREF DMTVEEALDF	LPPERRKPD. VPKERVPANP VPKERKGN. VNKTRRLPY. VNKTRRLPY. VNKTRLPY. NPATYTGVPD NPATYTGVPD NPATYTGVPD NPATYTKVFD PASIPKIKRK PASIPKIKRK FQAVPALARK FQNVPSIARK FKNIPSIKRT	GRVIEIKGAR GKVLEIKGAR GKVLKLTGAR GKVLKLTGAR GKVLVLKGAR ASLKIKGVR A C C C C C C C C C C C C C	AHNLKNUSVK GMNLKDUTLT AHNLKNUTLR HMNLKNUTUR romatic-II a 30 EAKVÄGYKKG EARKRGYAPG AAKARGYNKS CARGYAPG AAKARGYNKS YMKLÅQPATT YMKLGQPATT YMKLGQPATT YMKLGQPATT	IPLGTPVAVT IPLGTPVAVT IPLGVPVAVT IPLGVPVCVT IPLGVPVCVT RPSPNVKGGR RPSPNVKGGR RPSPNVKGGR RFSPNVKGGR LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK	GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GEACHGDGII CEACHGDGII CEACHGDGII CEACGGGVV Zn 3 LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN	VNEVLYKALA INDTLFPIAQ VHDVLYKALA INDTLFPIAQ VHDVLYKALA VMETLYPALM 5 5 KIEMHFLPDV KIEMHFLPDV KIEMLFLPDV B27 GRTLYILDEP GRTLYILDEP GRTLYILDEP GRTLYILDEP GRTLYILDEP Malker R-II	QKLÅ. RAKAK QKLH. RAKAK RQLMAATTAE QRLM. RAKTT NLLH. KTKLP 326 YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCKGK Zn 3 TTGLÅVDDIA TTGLÅVDDIA TTGLHVDDIA TTGLHPDIQ TTGLHFEDVR	PGEHŘDIRGL PGBHRDIRGL PGPYEALEGV AGEPDSIEGH a31 RYNRĚTLEVT RYNRETLEVT RYNRETLEIK RYNRETLEIT a37 RLLDŮLHRLV RLLDŮLHRLV RLLDVLHRLW KLLSVLHRLV	EHLDŘVIDID EHLDŘVIDID EHPDKVIDID ENIDKMIAID ENIDKMIAID YKGRNIAEVL YKGRNIAEVL YKGRNIAEVL YKGRNISDIL B28 DNGDŤVLVIE DNGDŤVLVIE DNGDŤVLVIE DAGNTVVVIE DRGNTVVIE	685 689 669 665 785 785 789 791 765 885 885 885 889 891 865
TmaUvrA BstUvrA BcaUvrA EcoUvrA ThUvrA TmaUvrA BstUvrA BscUvrA EcoUvrA ThUvrA TmaUvrA	QYMSGKRKIE AYLRGERRIP EYLSGKRKIT OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS O-loop-II OMTVEDALDF DMTVEDALDF DMTVEEALDF DMTVEEALDF DMTVEEALDF DMTVEEALDF DMTVEEALDF	LPPERRKPD. VPKERRKGN. VPKERRKGN. VNKTRRLPY. VNKTRRLPY. MPATYTGVPD NPATYTGVPD NPATYTGVPD NPATYTGVPD NPATYTKVFD PASIPKIKRK FASIPKIKRK FDAVPALARK FQNVPSIARK FKNIPSIKRT β29.	GRVIEIKGAR GKVLEIKGAR GKVLKLTGAR GKVLKLTGAR GKVLVLKGAR AJ 4 29 DIRDÝPASTN DIRDÝPA	EHNLKINSVK GMNLKDVTLT AHNLKNVTLR AHNLKNTVE romatic-II 4 30 EAKVÄGYKKG EARVRGYKKG EARKRGYAPG AAKARGYNKS 4 35 YMKLÅQPATT YMKLGQPATT YMKLGQPATT YMKLGQPATT YMKLGQPATT YMKLGQPATT	IPLGTPVAVT IPLGTPVAVT IPLGVPVAVT IPLGVPVCVT IPLGVPVCVT RPSPNVKGGR RPSPNVKGGR RPSPNVKGGR RPSPNVKGGR RFSPNVKGGR RFSPNLKGGR LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK	GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GPSGSGKSTL GPSGSGKSTL GPSGGGSGKSTL GPSGGGSGKSTL GPSGGGSGKSTL GPSGGGSG GPSGGGSG GPSGGGGGG CEACGGGGGT CEACGGGGT CEACGGGGT CEACGGGGT CEACGGGGT CEACGGGGT CEACGGGGGT CEACGGGGGT CEACGGGT CEACGGGT CEACGGGT CEACGGGT CEACGGGT CEACGGGT CEACGGGT CEACGGGT CEACGGGT CEACGGGT CEACGGGT CEACGGGT CEACGGGT CEACGGGT CEACGGGT CEACGGT CEACGGGT CEACGGGT CEACGGGT CEACGGT C	VNEVLYKALA VNEVLYKALA INDTLFPIAQ VHDVLYAALA VMETLYPALM 5 KIEMÅPLPDV KIEMÅPLPDV KIEMLPLPDI KIEMLPLPDU β 27 GRTLYILDEP GRTLYILDEP GRTLYILDEP Walker B-II	QKLÅ. RAKAK QKLH. RAKAK RQLMAATTAE QRLM. RAKTT NLLH. KTKLP 326 YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCKGK Zn 3 TTGLÅVDDIA TTGLÅVDDIA TTGLÅVDDIA TTGLHPDIQ TTGLHPDIQ TTGLHPDIQ	PGEHŘDIRGL PGBHRDIRGL PGPYEALEGV AGEPDSIEGH a31 RYNRĚTLEVT RYNRETLEVT RYNRETLEVK RYNRETLEIK RYNRETLEVK RYNRETLEVK RYNRETLEVK RYNRETLEVK RYNRETLEVK RUNKETLEVK	EHLDŘVIDID EHLDŘVIDID EHPDKVIDID ENIDKMIAID ENIDKMIAID VKGRNIAEVL YKGRNIAEVL YKGRNIAEVL YKGRNISDIL β28 DNGDŤVLVIE DNGDŤVLVIE DAGNTVVIE DAGNTVVIE DRGNTVVIE	685 689 691 665 785 785 789 791 765 885 885 885 885 889 891 865
TmaUvrA BstUvrA BcaUvrA ThJVvrA TmaUvrA BstUvrA BcaUvrA TmaUvrA TmaUvrA	QYMSGKRKIE AYLRGERRIP EYLSGKRKIT OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS O-loop-II 0MTVEDALDF DMTVEDALDF DMTVEEAREF DMTVEEAREF DMTVEEALDF DMTVEEALDF	LPPERRKPD. VPKERRKON. VNKTRRLPY. VNKTRRLPY. NPATYTGVPD NPATYTGVPD NPATYTGVPD NPATYTGVFD NPATYTKVFD PATYTK	GRVIEIKGAR GKVLEIKGAR EKVLKLTGAR GKVLKLAGAR ASLKIKGVR AU 4 29 DIRDÝPASTN DIRDÝPASTN DIRDÝPASTN DIRDÝPASTN PVRELPASTP EIRDLFAKTP EIRDLFAKTP EIRSLFANTP 34 LETLÝDVGLG LQTLMDVGLG LQULHDVGLG B 30 DIRDÝPASTN	EHNLKINSVK GMNLKDVTLT AHNLKNVTLR HINLKNITVE romatic-II 4 30 EAKVÄGYKKG EARVRGYKKG EARKRGYAPG AAKARGYNKS 4 35 YMKLGQPATT YMKLGQPATT YMKLGQPATT YMKLGQPATT YMKLGQPATT	IPLGTPVAVT IPLGTPVAVT IPLGVPVAVT IPLGVPVCVT IPLGVPVCVT RPSPNVKGGR RPSPNVKGGR RPSPNVKGGR RPSPNVKGGR RPSPNLKGGR LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK	GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GEAGHGDGII GEAGHGDGII GEAGHGDGII GEAGGGGVI GEAGGGGVV Zn 3 LAAELHRRSN MANGAG	VNEVLYKALA VNEVLYKALA INDTLFPIAQ VHDVLYAALA VMETLYPALM 5 KIEMHPLPDV KIEMHPLPDV KIEMLPLPDU KIEMLPLPDU B 27 GRTLYILDEP GRTLYILDEP GRTLYILDEP GRTLYILDEP Walker B-II	QKLÅ. RAKAK QKLH. RAKAK RQLMAATTAE QRLM. RAKTT NLLH. KTKLP 326 YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCKGK Zn 3 TTGLÅVDDIA TTGLÅVDDIA TTGLÅVDDIA TTGLÅVDDIA TTGLÅPDIQ TTGLÅPDIQ	PGEHŘDIRGL PGBHRDIRGL PGPYEALEGV AGEPDSIEGH a31 RYNRĚTLEVT RYNRETLEVT RYNRETLEIT a37 RLLDŮLHRLV RLLDŮLHRLV RLLDŮLHRLV KLVEVLHRLV	EHLDŘVIDID EHLDŘVIDID EHPDKVIDID ENIDKMIAID ENIDKMIAID VKGRNIAEVL YKGRNIAEVL YKGRNIAEVL YKGRNISDIL B28 DNGDŤVLVIE DNGDŤVLVIE DNGDŤVLVIE DRGNTVVIE DRGNTVVIE	685 689 691 665 785 785 789 791 765 885 885 889 891 865
TmaUvrA BstUvrA EcoUvrA TthUvrA TmaUvrA BstUvrA EcoUvrA TmaUvrA BstUvrA BstUvrA	QYMSGKRKIE AYLRGEKRIP EYLSGKRKIT OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS DMTVEDALDF DMTVEDALDF DMTVEDALDF DMTVEDALDF DMTVEDALDF DMTVEDALDF DMTVEDALDF DMTVEALF DMTVEALF DMTVF DMTVEALF DMTVF DMT	LPPERRKPD. VPKERRKGN. VNKTRRLPY. VNKTRRLPY. NPATYTGVPD NPATYTGVPD NPATYTGVPT NPATYTGVPT NPATYTGVPT NPATYTGVPT NPATYTKVFD Ca PASIPKIKRK PASIPKIKRK PASIPKIKRK FDAVPALARK FKNIPSIKRT \$29 YIIDLGPEGG	GRVIEIKAAR EKULKLTGAR GKWLVLKJAGAR GKWLVLKJAGAR A.SLKIKGVR A.29 DIRDŮPASTN DIRDŮPASTN DIRDŮPASTN DIRDŮPASTN EIRSLFAMTP 34 LETLÝDVGLG LQTLMDVGLG LQVLHDVGLG LQVLHDVGLG DRGGQIVAVG	EHNLKINSVK GINLKDVTLT AHNLKDVTLT AHNLKDVTLT AHNLKDVTLT AHNLKDVTLT COMMERCIAL	IPLGTPVAVT IPLGTPVAVT IPLGRPVATT IPLGRPVATT IPLGRPVKGGR RFSFNVKGGR RFSFNVKGGR RFSFNVKGGR RFSFNVKGGR RFSFNVKGGR RFSFNLKGGR LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK SGGEAQRVK SGGEAQRVK SGGEAQRVK SGGEAQRVK SGGEAQRVK SGGEAQRVK SGGEAQRVK SGGEAQRVK SGGEAQRVK SGGEAQRVK SGGEAQRVK SGGEAQRVK SGGEAQRVK SGGEAQRVK SGGEAQRVK SGGEAQRVK SGGEAQRVK	GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GVSGSGKSTL GEACHGDGII CEACHGDGII CEACHGDGII CEACGGGVI CEACGGGVV Zn 3 LAAELHRRSN	VNEVLYKALA INDTLFPTAQ VHDVLYKALA INDTLFPTAQ VHDVLYKALA VMETLYPALM 5 5 KIEMHPLPDV KIEMHPLPDV KIEMLPLPDV B 27 GRTLYILDEP GRTLYILDEP GRTLYILDEP GRTLYILDEP Walker B-II	QKLÅ. RAKAK QKLH. RAKAK RQLMARTAR QRLM. RAKTT NLLH. KTKLP 326 YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCKGK Zn 3 TTGLÅVDDIA TTGLÅVDDIA TTGLHVDIA TTGLHVDIA TTGLHPDIVA	PGEHŘDIRGL PGBHRDIRGL PGPYEALEGV AGEPDSIEGH AGEPDSIEGH RYNRĚTLEVT RYNRĚTLEVT RYNRETLEIK RYNRETLEIT G 37 RLLDŮLHRLV QLLDVLHRLV QLLDVLHRLV KLVEVLHRLV	EHLDŘVIDID EHLDŘVIDID EHPDKVIDID ENIDKMIAID ENIDKMIAID VKGRNIAEVL YKGRNIAEVL YKGRNIAEVL YKGRNISDIL β28 DNGDŤVLVIE DNGDŤVLVIE DAGNTVLVIE DAGNTVLVIE	685 689 691 665 785 785 785 785 785 785 785 785 785 885 8
TmaUvrA BstUvrA BcaUvrA CoOVrA ThUvrA TmaUvrA BstUvrA BstUvrA BstUvrA BstUvrA BcaUvrA	QYMSGKRKIE AYLRGEKRIP EYLSGKRKIT OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS DMTVEDALDF DMTVEDALDF DMTVEDALDF DMTVEDALDF DMTVEDALDF DMTVEDALDF DMTVEDALDF DMTVEALDF MTVDEALEF	LPPERRKPD. VPKERRKGN. VPKERRKGN. VNKTRRLPY. 428 NPATŤTGVPT NPATŤTGVPT NPATŤTGVPT NPATŤTGVPT NPATŤTGVPT NPATŤTKVPD 4 6 4 7 4 7 4 7 8 29 YIIDLGPEGG YIIDLGPEGG YIIDLGPEGG	GRVIEIKGAR GRVIEIKGAR EKVLKLTGAR GKWLVLKGAR GKWLVLKGAR A: ASLKIKGVR A: 29 DIRDVPASTN DIRDVPASTN DIRDVPASTN DIRDVPASTN PVRELPASTN PVRELPASTN PVRELPASTN PVRELPASTN A: 29 24 LETLÝDVGLG LETLÝDVGLG LQUVHDVGLG LQUVHDVGLG LQUVHDVGLG DRGGQIVAVG DRGGQIVAVG DRGGQIVAVG	AHULKIVSVK GMNLKDVTLT AHNLKNVTLT AHNLKNVTLT AHNLKNVTLT AHNLKNUTLT COMMERCIAL	IPLGTPVAVT IPLGTPVAVT IPLGRPVAVT IPLGRPVAIT IPLGRPVAIT IPLGVPVCVT RPSPNVKGGR RPSPNVKGGR RPSPNVKGGR RPSPNVKGGR RPSPNLKGGR LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK SHTGRYLKPI SHTGRYLKPI SHTGRYLKPI	GVSGSGKSTL GVSGSGKSTL GPSGSGKSTL GPSGSGKSTL GPSGSGKSSL Walker A-II β2 CEACHGDGII CEACHGDGII CEACGGDGV CEACGGGGV CEACGGGGV CEACGGV CEACGUV CEACGUV CEACGUV CEACGUV CACUV CACGUV CACGUV CACU	VNEULYKALA INDTLPPIAQ VHDULYAALA VMETLYPALM 5 KIEMÅFLPDV KIEMÅFLPDV KIEMLFLPDV KIEMLFLPDV KIEMLFLPDV B 27 GRTLFILDEP GRTLFILDEP GRTLFILDEP Walker B-II RYEAÅKA 95: QYEAVKA 95: 944	QKLÅ. RAKAK QKLH. RAKAK QKLH. RAKAT QRLM. RAKTT NLLH. KTKLP 326 YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCKGK ZN 3 TTGLÅVDDIA TTGLÅVDDIA TTGLHPDIA TTGLHPDIA TTGLHPDIA	PGEHŘDIRGL PGBHRDIRGL PGPYEALEGV AGEPDSIEGH AGEPDSIEGH RYNRĚTLEVT RYNRĚTLEVT RYNRETLEIK RYNRETLEIT 437 RLLDŮLHRLV RLLDŮLHRLV RLLDVLHRLV KLLSVLHRLV	EHLDŘVIDID EHLDŘVIDID EHPDKVIDID ENIDKMIAID ENIDKMIAID YKGRŇIAEVL YKGRŇIAEVL YKGRNIAEVL YKGRNISDIL β28 DNGDŤVLVIE DNGDŤVLVIE DQGNTVLVIE DRGNTVVVIE DRGNTVIVIE	685 689 691 665 785 785 785 785 785 785 785 785 785 78
TmaUvrA BstUvrA BcaUvrA CoUvrA ThUvrA TmaUvrA BstUvrA BstUvrA BstUvrA BstUvrA BcaUvrA BcaUvrA CoOUrA	QYMSGKRKIE AYLRGERRIP EYLSGKRKIT OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OMTVEDALDF DMTVEDALDF DMTVEDALDF DMTVEDALDF DMTVEDALDF DMTVEDALF DMTVEDALF DMTVEALF	LPPERRKPD. VPKERRKGN. VNKTRRLPY. VNKTRRLPY. A28 NPATYTGVPD NPATYTGVPD NPATYTGVPT NPATYTGVPT NPATYTGVPT NPATYTKVPD A51PKIKRK FDAVPALARK FQNVPSIARK FKNIPSIKRT β29 YIIDLGPEGG VIIDLGPEGG WIVDLGPEGG WIVDLGPEGG	GRVIEIKGAR EKVLKLTGAR GKWLKLTGAR GKWLVLKGAR . ASLKIKGVR 4 29 DIRDVPASTN DIRDVPASTN PVRELPAGVP PVRELPAGVP EIRDLPAKTP EIRDLPAKTP EIRSLFAMTP 34 LETLÝDVGLG LQTLMDVGLT LQUHDVGLG LQVLHDVGLG LQVLHDVGLG DRGGQIVAUG DRGGQIVAUG DRGGQIVAUG DRGGQIVAUG	EHNLKINSVK GINILKDVTLT AHNLKDVTLT AHNLKDVTLT AHNLKDVTLT a30 EAKVRGYKKG EAKVRGYKKG EAKVRGYKKG EARKRGYGPG AAKARGYNKS 435 YMKLGQPATT YMKLGQPATT YMKLGQPATT YMKLGQPATT YMKLGQPATT YMKLGQPATT YMKLGQPATT YMKLGQPATT TPEEVAEVE TPEEVAEVE	IPLGTPVAVT IPLGTPVAVT IPLGRPVAVT IPLGRPVATT IPLGVFVCVT RFSFNVKGGR RFSFNV	GVSGSGKSTL GVSGSGKSTL GPSGSGKSTL GPSGSGKSTL GPSGSGKSTL GPSGSGKSSL Walker A-II β2 CEACHGDGII CEACGGGII CEACGGGGV CEACGGGGV CEACGGGGV Zn 3 LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELRRSN	VNEULYKALA VNEULYKALA INDTLFPIAQ VHDULYAALA VMETLYPALM 5 KIEMÅFLPDV KIEMÅFLPDV KIEMLFLPDV KIEMLFLPDV KIEMLFLPDV RTLYILDEP QTLYILDEP QRTLYILDEP QRTLYILDEP QRTLYILDEP QYEAÅKA AD AD	QKLH. RAKAK QKLH. RAKAK RQLMGATIAE QRLM. RAKTT NLLH. KTKLP 328 YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCKGK YVPCEVCKGK YVPCEVCKGK TTGLHVDDIA TTGLHVDDIA TTGLHFDDVA TTGLHFDDVA TTGLHFDDVA TTGLHFDVA	PGEHŘDIRGL PGBHRDIRGL PGPYEALEGV AGEPDSIEGH AGEPDSIEGH RYNRĚTLEVT RYNRĚTLEVT RYNRĚTLEVT RYNRĚTLEVT RYNRĚTLEVT RYNRĚTLEVT RYNRĚTLEVT RYNRĚTLEVK RYNRĚTLEVK RYNRĚTLEVK RYNRĚTLEVK RULDVLHRLV QLLDVLHRLV	EHLDŘVIDID EHLDKVIDID EHPOKVIDID ENIDKMIAID ENIDKMIAID VKGKNIAEVL YKGKNIAEVL YKGKNIAEVL YKGKNIAEVL YKGKNISDIL β28 DNGDŤVLVIE DNGDŤVLVIE DNGDŤVLVIE DAGNTVVVIE DRGNTVIVIE	685 689 691 665 785 789 791 765 885 885 885 885 885 885 885 885 885 8
TmaUvrA BstUvrA BcaUvrA EcoUvrA ThUvrA TmaUvrA BstUvrA BstUvrA BstUvrA BstUvrA BcaUvrA TmaUvrA ThUvrA TmaUvrA	QYMSGKRKIE AYLRGERRIP EYLSGKRKIT OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS OSPIGRTPRS DMTVEDALDF DMTVEDALDF DMTVEBALF DMTVEBALF DMTVEF DMTVEF DMTVEF DMTVEF DMTVEF DMTVEF DMTVEF DMTVEF DMTVEF DMTVEF DMTVEF DMTVEF DMTVEF DMTVEF DMTVF DMTVEF DMTVF DM	LPPERRKPD. UPPERRKPD. UPKERRKGN. UNKTRRLPY. UNKTRRLPY. UNKTRRLPY. UNKTRRLPY. UNKTRRLPY. UNKTRRLPY. UNKTRRLPY. UNKTRRLPY. Δ28 0 0 0 0 0 0 0 0 0 0 0 0 0	GRVIEIKGAR GRVIEIKGAR EKVLKLTGAR GKWLVLKGAR ASLKIKGVR AI (22) DIRDVPASTN DIRDVPASTN PVRELPAGVP EIRDLPAKTP EIRDLPAKTP EIRDLPAKTP EIRDLPAKTP EIRDLPAKTP EIRDLPAKTP EIRDLPAKTP EIRDLPAKTP EIRDLPAKTP EIRDLPAKTP B30 DRGGQIVAGL GGGQIVAGL SGGGQIVAG SGGQ SGGQ SGGQ SGGGQIVAG SGGGQIVAG SGGGQIVAG SGGGQIVAG SGGGQIVAG SGGGQIVAG SGGQ SGGGQIVAG SGGQ SGGQ SGGGQIVAG SGGGQIVAG SGGGQIVAG SGGGQIVAG SGGGQI SGGQ SGGGQI SGGGQI SGGQ SGGGQI SGGGQI SGGQ SGGGQ SGGGQI SGGQ	AHULANUSUK GMNLKDVTLT AHNLKNVTLR AHNLKNVTLR AHNLKNUTVE comatic-II a 30 EAKVRĞYKKG EARVRĞYKKG EARVRĞYKKG EARVRĞYKG EARKRĞYGPĞ AAKARĞYNKS AKARĞYNKS AKARĞYNKS AKARĞYNKS AKARĞYNKS AKARĞYNKS AKARĞYNKS AKARĞYNKS AKARĞYNKS AKARĞYNKS AKARĞYNKS TYELÖVACIA TPEEVABCEA TPEEVALTG. TPEELAKNPH	IPLGTPVAVT IPLGTPVAVT IPLGRFVAIT IPLGRFVAIT IPLGVFVCVT RFSFNVKGGR RFSFNVKGGR RFSFNVKGGR RFSFNVKGGR RFSFNKGGR LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK LSGGEAQRVK SGGBAQ SGGBAQ S	GVSGSGKSTL GVSGSGKSTL GPSGSGKSTL GPSGSGKSTL GPSGSGKSSL Walker A-II β2 CEACHGDGII CEACHGDGII CEACGGDGII CEACGGDGV CEACGGDGV CEACGGDGV CEACGGDGV CEACGGQGVV Zn 3 LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELHRRSN LAAELRRCAT LASELRKRDT LASELRKRDT LASELRKRDT LASELRKRDT LASELRKRDT LASELRKRDT LASELRKRDT LASELRKRDT LASELRKRDT LASELRKRDT LASELRKRDT LASELRKRDT LASELRKRDT LASELRKRDT	VNEULYKALA VNEULYKALA INDTLFPIAQ VHDULYAALA VMETLYPALM 5 KIEMHFLPDV KIEMHFLPDV KIEMHFLPDV KIEMHFLPDV KIEMLFLPDV KIEMLFLPDV GRTLYILDEP GRTLYILDEP GRTLYILDEP Walker B-II RYEAÅKA 95: QYEAVKA 95:	QKLÅ RAKAK QKLH.RKAK RQLMGATIAE QRLM.RAKTT NLLH.KTKLP 326 YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK YVPCEVCHGK TYGLHVDUA TTGLHVDUA TTGLHPDVA TYGLHPEDVR D-loop-I	PGEHŘDIRGL PGBHRDIRGL PGPYEALEGV AGEPDSIEGH AGEPDSIEGH RYNRĒTLEVT RYNRĒTLEVT RYNRĒTLEVT RYNRĒTLEIT A 37 RLLDVLHRLV RLLDVLHRLV KLLSVLHRLV KLVEVLHRLV	EHLDŘVIDID EHLDKVIDID EHPDKVIDID EHPDKVIDID ENIDKMIAID VKGKNIAEVL YKGKNIAEVL YKGKNIAEVL YKGKNIAEVL YKGKNIAEVL DNGDŤVLVIE DNGDŤVLVIE DAGNTVVVIE DRGNTVVIE	685 689 691 665 785 789 791 765 885 885 885 885 885 889 891 865

PakotipraphaS4.tif

Figure S4. Multiple sequence alignment of UvrA orthologs. Sequences of 135 UvrA orthologs were aligned using ClustalW v.1.83 (Thompson et al., 1994). Only sequences

from *Bacillus stearothermophilus* and four other bacteria in which NER is most extensively studied are shown: *Bst, Bacillus stearothermophilus; Bca, Bacillus caldotenax; Eco, Escherichia coli; Tth, Thermus thermophilus; Tma, Thermotoga maritima.* The secondary structure according to the *Bst*UvrA model is shown above the sequences, colored by domains as in Figure 1. Disordered regions are depicted as dashed lines. Locations of the conserved ABC ATPase motifs, glycine-rich loops, and zinc-coordinating residues are shown on the amino acid sequences.



PakotipraphaS5.tif

Figure S5 Binding of wild-type and mutant *Bst*UvrA to the 50-bp duplex containing the fluorescein lesion. The positions of the free DNA and the UvrA·DNA complex are indicated. Mutants 1, 2, 3, and 4 contain mutations in regions a, b, a+b, and c, respectively (see Figure 6C). The 5 nM experimental point (light gray boxes) prominently illustrates the difference in affinity for damaged DNA between the wild-type and mutant proteins.



Figure S6 Binding of wild-type and mutant *Bst*UvrA to the 50-bp duplex containing the N3-menthol lesion (Verhoeven et al., 2002) (A) and the undamaged 50-bp duplex (B). The positions of the free DNA and the UvrA·DNA complex are indicated. Mutants 1, 2, 3, and 4 contain mutations in regions a, b, a+b, and c, respectively (see Figure 6C).

Table S1. Sequences of primers used in the amplification of *Bacillus stearothermophilus uvr* genes and the construction of UvrA mutants.

Primers	Sequences $(5' \rightarrow 3')^{a, b}$
<i>uvrA</i> fwd	GGTAATC <u>CATATG</u> GATAAAATTATCGTCAAAGGGGCGCGCGCCCACAACTTG
uvrA rev	GGTAATC <u>AAGCTT</u> TCACGCCTTCGCCGCTTCATACCGCGCCTGCATGCGCG
<i>uvrB</i> fwd	GGTAATC <u>CATATG</u> GGTCCGAAGAAAGTGGAGGGCCGTTTTCAATTAGTGTCGC
uvrB rev	GGTAATC <u>AAGCTT</u> TCACCCTTCCGCTTTCAATTCGAAAATGATATCGCGC
<i>uvr</i> C fwd	GGTAATC <u>GCTAGC</u> AAAAAGAACGAGCGGCTGAAAGAAAAGCTGG
uvrC rev	GGTAATC <u>AAGCTT</u> TCATTCATGCAGTTTTTCATAGATTTTCTCCGCCACC
UvrA _{A131-245} fwd	CACGGCATTGAAATCCAATCGCAG GGTACC GAAAAGCACGCTTGTCCGTACTG
UvrA _{ΔI131-245} rev	CAGTACGGACAAGCGTGCTTTTCGGTACCCTGCGATTGGATTTCAATGCCGTG
UvrA ₂₂₈₅₋₄₀₀ fwd	CTGCGACGGGCTCGGGGCGAAGCTC GGTACC GAACAACCATGTCCGACATGCCAAG
UvrA _{A285-400} rev	CTTGGCATGTCGGACATGGTTGTTC GGTACC GAGCTTCGCCCCGAGCCCGTCGCAG
UvrA 131-245 fwd	CGCGGCAGC <u>CATATG</u> CCCATTTGCCCGACGCAC
UvrA 131-245 rev	GGCCGCAAGCTTTTACGAAAAGCCGCAGTACGGAC
UvrB 149-250 fwd	CGCGGCAGC <u>CATATG</u> GGGTCGCCGGAAGAATATCGG
UvrB 149-250 rev	GGCCGC <u>AAGCTT</u> TTACACGAAGTGCGACGCCGG
UvrA-R708A fwd	ACCGGGGTGTTTGACGACATC GCG GATGTGTTTGCCTCGACGAAC
UvrA-R708A rev	GTTCGTCGAGGCAAACACATC CGC GATGTCGTCAAACACCCCGGT
UvrA-K718A/R720A/R726A fwd	GCCTCGACGAACGAAGCGGCGGTGGCGGGCTACAAAAAAGGGGCGTTCAGCTTCAATGTCAAAGG
UvrA-K718A/R720A/R726A rev	
UvrA-K732A/R735A fwd	GGGCGGTTCAGCTTCAATGTC GCG GGCGGG GCG TGCGAGGCCTGCCATGGCGAT
UvrA-K732A/R735A rev	ATCGCCATGGCAGGCCTCGCA CGC CCCGCC CGC GACATTGAAGCTGAACCGCCC
UvrA-K765A/R766A/R769A fwd	CCGTGCGAAGTGTGCCACGGCGCGCGTACAACGCGGAGACGCTCGAGGTGACGTAT
UvrA-K765A/R766A/R769A rev	ATACGTCACCTCGAGCGTCTCCGCGTGTACGCCGCGCGCG

^a Recognition sequences of the restriction enzymes used for cloning of the PCR products are underlined: *uvrA* and *uvrB*, *Nde*I and *Hind*III; *uvrC*, *Nhe*I and *Hind*III

^b The sequence encoding the GT dipeptide inserted in place of the deleted domains, and the positions of $K \rightarrow A$, $R \rightarrow A$ mutations are in bold.