Supplementary Information:

Structural elucidation of recombinant *Trichomonas vaginalis* 20S proteasome bound to covalent inhibitors

Jan Silhan^{1,#}, Pavla Fajtova^{1,2,#,*}, Jitka Bartosova¹, Brianna M. Hurysz², Jehad Almaliti³, Yukiko Miyamoto⁴, Lars Eckmann⁴, William Gerwick³, Anthony J. O'Donoghue^{2,5*}, Evzen Boura^{1,*}

¹Institute of Organic Chemistry and Biochemistry AS CR, v.v.i., Flemingovo nam. 2., 166 10 Prague 6, Czech Republic

²Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA, 92037, USA

³Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, 92037, USA
⁴Division of Gastroenterology, School of Medicine, University of California San Diego, La Jolla, CA, 92037, USA

⁵Center for Discovery and Innovation in Parasitic Diseases, Skaggs School of Pharmacy and pharmaceutical Sciences, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093, USA

[#]these authors contributed equally

*correspondence to: fajtova@uochb.cas.cz or ajodonoghue@health.ucsd.edu or boura@uochb.cas.cz

2		1	10	20	30	40	50	60	7 (
a	Tv_alfa_1	MSSGAD	RYLTVESA	EGRLWOVEY	SFKAVKQAE	VTAVAVKSKNA	VCVAVQKKVS	DKLIDPST	VTHMYRITI	NVGACLV
	Tv_alfa_2	MGDSD	SLTT <mark>FS</mark> S	G <mark>GKLNQIE</mark> S	ALKAVSIG.	GQCVGVKAKNG	AVIACESKPS	SPLVEKVT	NLKVQKINI	NVGIVYS
	Tv_alfa_3	MTYR <mark>YD</mark>	AGTTT <mark>FS</mark> S	D <mark>GRILO</mark> VEY	AIQSINQA.	GTAIGVQFTNG	VVL <mark>A</mark> AEKKNI	GR <mark>LV</mark> DYLF	PEKMAKIDO	GHIVTAVA
	Tv_alfa_4	MSDYTI	RSITR <mark>FS</mark> P	^D GRLFQIDH	I <mark>AHAAV</mark> QRG .	TTVVATRSKDM	IIVI <mark>A</mark> VEKTAV	AKLQDPHT	FSKICSLDI	KHVMCAFA
	Tv_alfa_5	MFNSGSEYDI	RNVNTESP	DGRLLQVEY	AIEAVKLG.	SSAVAILCPEG	VIFAVEKRLS	SQLLIASS	VEKVYAIDI	HVGVVMA
	Tv_alfa_6	MFRSKYDI	NATTESP	PEGRILOVEN	AMKAVQQG.	MPTVGLKSKTH	AVIA. GVMI	ISPSEFSSH	QPKIFKIDQ	2HIGVAIS
	TV_alla_/	MSGAGSGIDI	NPITESP		ATRAVERD.	SLALGVACADG	LLAASKNLI	STLLTPGG	NPRIFWINI	JSIACATI
		80	90	100	110		120	130	140	150
	Tv_alfa_1	GLPSDVNFI	MLLRSFA	NNFEYKQGF	SIPVSILAC	MLSE	RHQLESQLVY	VRPSAVSA	ILFGLDGPS	SDSFALYK
	Tv_alfa_2	GVNTDFHVIJ	LKS <mark>LR</mark> KAS	SIKYSLRLGV	VEMPTREVV K	:н <mark>А</mark> АН	KMQYY <mark>TQ</mark> IGG	V <mark>RP</mark> FGVSL	LIIGWEEL	GPT <mark>LW</mark> Q
	Tv_alfa_3	GLTADANTL	/ DL <mark>MR</mark> TSA	AQKYLKT <mark>Y</mark> DE	QMPVEQLVF	MVCDE	. KHSY <mark>TQ</mark> YGG	GL <mark>RP</mark> YGVSF	LIAGYDRHI	K.GCQ <mark>LY</mark> L
	Tv_alfa_4	GLHADARRL:	QSGQRQC	QSHRLTYED	PISIENIAR	YIATL	.QLKNTQSGG	ARPYGVST	LICGFDDM	CSQPHIYE
	TV_alia_5	GLAADGRTM	VEHMRVEA	I UURROR CE	IPIGIKAVTÇ	SVCDLALAFGE	GREEKGDGQM	ISRPFGTAL	LVAGIENG	CDDIFF
	Tv_alia_0	CHPPDCVST	TEOSPNRA	ETETSNEGI	KITVPOLAS	FVSO	OFHLAHYYOR	VPPECCTV	TFASYKDD	
	IV_dIId_/						21 m D m m 1 2 m		11.101.100	
		160	?	170	180	190	200	210	22	2 Q
	Tv_alfa_1	I E <mark>P S G</mark> Y S N G I	FRAVAC GV	KEIE <mark>A</mark> MSA <mark>I</mark>	EKKMEDFEI	PEATAEFTLST	LQTVCGVDFE	AQDVEVS L	LTRD1	ISKFSKLP
	Tv_alfa_2	VDPSGTFWA	VKATALGK	RSDGSRTF	ERRYSEDQS	VDDAIHTAIST	LKEGFDGQL.	TAELI	EIG.VVDE	TRKFRTLS
	TV_alfa_3	TDPSGNFGG	VKATAIGE VKATAIGE	NNQTAQSI	KSQYKDNM1	ATEAMDLTVKV	LCKTLDSTSI	SADKLEFA	VLQFREEYC	SPKVRILT COMPTNA
	TV_alia_4	TDPSGTIAL	TRARIER	CSECAEAL	PDLVKDCMT	TUFAFDIAIGA		IN LEVA	ADV C	GIMEIMA
	Tv alfa 6	TCPSGOHWE	NAOAIGR	RAOAAKTYI	ETNLNEFPC	C. TRDOLIBHA	LRALNDCKSF	ESDSLEAT	ALG. VVGII	DEPETILE
	Tv alfa 7	IEPSGAFYG	FASCEGK	NSNLARAEI	OKTEWKNIT	VREAVPEVARI	IKSLHESOFF	KWEIEMF.	WLC.EETNO	GRPOKVPE
					~					
		0.2.0	2.4.0							
		230	240							
	Tv_alfa_1	NDKVNEILH	AVAEKD							
	Tv_alfa_2	TAEIRDFLTH	SV							
	TV_alfa_3	FEVIDALTE	XIEETIKK STVAV	SALEKE						
	Tv_alfa_5	SEOROEIVAL	31. P P P T T P	· · · · · · ·						
	Tv alfa 6	GPELOKYID								
	Tv_alfa_7	DVFQSRFVNB	ENPQN							
n						1	1.0	2.0	3.0	
D						1	10	20	30	
D	Tv_beta_1			MEACLO		1 MSEYQ	10 FP.KESMGTI	20 LLAIQCTD	30 GVVMASDSI	RTS SGSFI
D	Tv_beta_1 Tv_beta_2 Tv_beta_3				FDFSNYARN	1 MSEYQ IKSLEPKLG MSD	10 FP.KESMGTT KPQLLSTGTT ISTYNGSC	20 LLAIQCTD NAAAIFDG	30 GVVMASDSI GVVLGADTI CVATASDBI	RTSSGSFI RATAGPIV RLGVNMLT
D	Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_4			MEAGLG	FDFSNYARN	1 MSEYQ IKSLEPKLG MSD	10 FP.KESMGTT KPQLLSTGTT .ISTYNGSC	20 LLAIQCTD LAAAIFDG VLAMAGDH	30 GVVMASDSI GVVLGADTI CVAIASDRI WVLIAADSS	RTSSGSFI RATAGPIV RLGVNMLT SVSSSIIC
D	Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_5	MQSLYLKPRI	DEVESEET	MEAGLG	FDFSNYARN YPDPCQFVKN	1 MSEYQ IKSLEPKLG MSD IHISLSYTNEPG	1 0 PFP.KESMGTT KPQLLSTGTT .ISTYNGSC MLS KM.AAVHGTT	20 TLLAIQCTD TIAAAIFDG VLAMAGDH SIVGLQGPD TLSFIYNG	30 GVVMASDSI GVVLGADTI CVAIASDRI WVLIAADSS GIVVAVDSI	RTSSGSFI RATAGPIV RLGVNMLT SVSSSIIC RATGGQFI
U	Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_5 Tv_beta_6	MQSLYLKPRI	DEVESEET	MEAGLG	FDFSNYARN 	1 MSEYQ IKSLEPKLG MSD IHISLSYTNEPG IKKGQ	10 SFP.KESMGTT KPQLLSTGTT .ISTYNGSC MLS KM.AAVHGTT WSPYEMHGGT	20 TLLAIQCTD TIAAAIFDG TVAMAGDH SIVGLQGPD TLSFIYNG TAIGICGDD	30 GVVMASDSI GVVLGADTI CVAIASDRI WVLIAADSS GIVVAVDSI YVVIGADTI	RTSSGSFI RATAGPIV LGVNMLT SVSSSIIC RATGGQFI RLSVDYSI
U	Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7	MQSLYLKPRI	DEVESEET	TKALEPAHV	FDFSNYARN PDPCQFVKN GEFREN	1 SLEPKLG MSD 	10 FP.KESMGTT KPQLLSTGTT .ISTYNGSC MLS KM.AAVHGTT WSPYEMHGGT .MQVITASGZ	20 TLAIQCTD TAAAIFDG VLAMAGDH TVGLQGPD TLSFIYNG AIGICGDD AIVAAKYDG	30 GVVMASDSI GVVLGADTI CVAIASDRI WVLIAADSS GIVVAVDSI YVVIGADTI GILLASDLS	RTSSGSFI RATAGPIV RLGVNMLT SVSSSIIC RATGGQFI RLSVDYSI SITYGSMF
D	Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7	MQSLYLKPRI	DEVESEET	TKALEPAHV	FDFSNYARN YPDPCQFVKN GEFREN	1 MSEYQ KSLEPKLG MSD HIISLSYTNEPG KKGQ	10 FP.KESMGTT KPQLLSTGT .ISTYNGS MLS KM.AAVHGT WSPYEMHGGT .MQVITASG	20 TLLAIQCTD TIAAAIFDG VLAMAGDH SIVGLQGPD TLLSFIYNG TAIGICGDD AIVAAKYDG	30 GVVMASDSI GVVLGADTI CVAIASDRI WVLIAADS GIVVAVDSI YVVIGADTI GILLASDLS	RTSSGSFI RATAGPIV LGVNMLT SVSSSIIC RATGGQFI LSVDYSI SITYGSMF
D	Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7	MQSLYLKPRI	DEVESEET	TKALEPAHV	FDFSNYARN PDPCQFVKN GEFREN	1 MSEYQ IKSLEPKLG MSD 	10 FP.KESMGTT ISTYNGSC MLS KM.AAVHGTT WSPYEMHGGT .MQVITASCP 90	20 TLLAIQCTD TIAAIFDG VLAMAGDH SIVGLQGPD TLSFIYNG AIGICGDD AIVAAKYDG 100	30 GVVMASDSI GVVLGADTI CVALASDRI WVLLAADSS GIVVAVDSI YVVIGADTI GILLASDLS	RTSSGSFI RATAGPIV LGVNMLT SVSSSIIC RATGGQFI LSVDYSI SITYGSMF
U	Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7	MQSLYLKPRI 40 PNRatinktit	50 STOPKTFA	TKALEPAHV	FDFSNYARN PDPCQFVKN GEFREN 70 70	1 	10 FP.KESMGTT KPQLLSTGT STYNGSC MLS KM.AAVHGTT .MQVITASGP 90 TDDSTTLVAS	20 TLLAIQCTD TIAAIFDG VLAMAGDH SIVGLOGPD TLSFIYNG AIGICGDD AIVAAKYDG 100 NVTRSIT	30 GVVMASDSF GVVLGADTI CVALASDRI WVLLAADSS GIVVAVDSF YVVIGADTI GILLASDLS	ATSSGSFI ATGQPIV LGVNMLT SVSSSIIC ATGGQFI LSVDYSI SIT YGSMF
U	Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_1 Tv_beta_1 Tv_beta_2	MQSLYLKPRI 40 PNRATNKITR AVKDEMKLH	50 SIQPKIFA	TKALEPAHU ME	FDFSNYARN YPDPCQFVKN GEFREN 70 CQFLARAVKN NDNINAVISA	1 MSEYQ KSLEPKLG HISLSYTNEPG KKGQ 80 NYLNALNITREN KURLFOMNTG.	10 FP.KESMGTT KPQLLSTGT .ISTYNGSC 	20 TLAAIQCTD TIAAAIFDG CVLAMAGDH SIVGLQGPD TLSFIYNG AIGICGDD AIGICGDD AIGICGDD IO0 SNVIRSLI.	30 GVVMASDSI GVVLGADTI CVAIASDRI WVLIAADSS GIVVAVDSI GILLASDLS VRYRQY FOYMGY	ATSSGSFI ATAGPIV LGVNMLT SVSSSIC ATGGQFI LSVDYSI 110 LSAGVIV IOAALIV
b	Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_1 Tv_beta_1 Tv_beta_2 Tv_beta_3	40 PNRATNKIT AVKDEMKLH VSKDFKRIF(50 Siqpkifa Ziqpkifa Jisdniwv Zindriyi	TKALEPAHU 	FDFSNYARN FDPCQFVKN GEFREN 70 QFLARAVKN IDNINAVISA LIVREQLRF	1 MSEYQ IKSLEPKLG MSD IHISLSYTNEPG IKKGQ KGQ 	10 FP.KESMGTT KPQLISTGTT ISTYNGSC MLS KM.AAVHGT WSPYEMHGGT .MQVITASG2 90 TDDSTTLVAS .LQPRVDQCT .RPLDPRKFM	20 TLLATQCTD TLAAIFDG VLAMAGDH SIVGLQGPD TLSFIYNG AIGICGDD IVAAKYDG 100 SNVIRSLI. NILASRL. NILASRL.	30 GVVMASDS GVVLGADT CVAIASDR WVLIAADS GIVLADT GILLASDL VRYRQY FQYMGY .EKRFSPF	ATSSGSFI ATAGPIV LGVNMLT SVSSSIIC ATGGQFI LSVDYSI IIY LSAGVIV IQAALIV FVTPVI
b	$\begin{array}{c} Tv_beta_1\\ Tv_beta_2\\ Tv_beta_4\\ Tv_beta_5\\ Tv_beta_6\\ Tv_beta_7\\ \end{array}$	40 PNRATNKITI AVKDEMKLHI VSKDFKRIFO MSENVDRIA	50 SIQPKIFA ZISDNIW JINDRIYI JLDDRHAI	60 ARCGNAADT CGACIAADN GLAGLATDV AMSGETGDC	PDPCQFVKN GEFREN 70 CQFLARAVKN IDNINAVISA VLTVREQLRF	1 MSEYQ IKSLEPKLG MSD IHISLSYTNEPG IKKGQ IYLNALNITREN .KLRLFQMNTG. DVNLLELREE. NVALYKFRNGV	10 FP.KESMGTT KPQLLSTGTT .ISTYNGSC MLS KM.AAVHGTT WSPYEMHGGT .MQVITASGF 90 TDDSTILVAS .LQPRVDQCT .RPIDPKKFM E.LSSDALA	20 LIATQCTD INAAIFDG VLAMAGDH IVGLQGPD ITLSFIYNG AIGICGDD IVAAKYDG 100 SNVIRSLI. NNILASRL. MLVKSTLY HFIRHTMA	30 GVVNASDSS GVVLGADT CVAIASDRI WVLIAADS GIVVAVDSS YVVIGADTI GILLASDL VRYRQY FQYMGY EKRFSPF KAVRKSP.	ATAGPIV ATAGPIV LGVNMLT SVSSSIIC ATGGQFI LSVDYSI IIYGSMF LSAGVIV LSAGVIV IQAALIV YEVIPVI YEVIMLL
D	$\begin{array}{c} Tv_beta_1\\ Tv_beta_2\\ Tv_beta_3\\ Tv_beta_5\\ Tv_beta_6\\ Tv_beta_7\\ Tv_beta_1\\ Tv_beta_2\\ Tv_beta_3\\ Tv_beta_4\\ Tv_beta_5\\ \end{array}$	40 PNRATNKIT AVKDEMKLH VSKDFKRIF MSENYDRIA FSQTVMKILH	50 2TQPKIFA ZISDNIW 2INDRIYI 2LDDRHAI 2LDRHAI 2LDRHAI	TKALEPAHU 60 ARCGNAADT CGAGIAADN GLAGLATDV TMAGGAADC	FDFSNYARN PDPCQFVKN . GEFREN . GEFREN . GEFREN . JNAVISA . LIVREQLRF . LQLSEVLQG . QYWLRNLSF	1 MSEYQ K. SLEPKLG HISLSYTNEPG K. S. KGQ IYINALNITREN KLRLFQMNTG. DVNLLELREE. NVALYKFRNGV LIQLHKFRYQ.	10 FP.KESMGTT KPQLLSTGT .ISTYNGSC MLS KM.AAVHGTT WSPYEMHGGT .MQVITASGP 90 TDDSTILVAS .LQPRVDQCT .RPIDPKKFM E.LSSDALA .QPLTVAAAS	20 TIAAIQCTD TIAAIFDG VIAMAGDH IVCLQGPD TISFIYNG AIGICGDD IVAAKYDG 100 SNVTRSLI. SNILASRL. NILVKSTLY HFIRHTMA SKILVNEL.	30 GVVMASDS GVVLGADT CVAIASDR WVLIAADS GIVVAVDSI GILLASDL 	ATSSGSFI ATAGPIV ALGVNMLT SVSSSIC ATGGQFI LSVDYSI IIQ LSAGVIV IQAALIV .FVVTPVI VEVNML VEVNML
D	$\begin{array}{c} Tv_beta_1\\ Tv_beta_2\\ Tv_beta_4\\ Tv_beta_5\\ Tv_beta_6\\ Tv_beta_7\\ \end{array}$	40 PNRATNKIT AVKDEMKLH VSKDFKRIF MSENYDRIA FSQTVMKIL DSRHKARIFI	50 SIQPKIFA ISDNIW ZINDRIYI LAPNMIG (MNSNCMI	60 ARCCNAADT CGAGIAADN GLAGLATDV AMSCETGADC SATGFDGDI	FDFSNYARN 70 CQFLARAVKN IDNINAVISA 7LTVREQLRF LQLSEYLQG QYWLRNISP DAFITRMRS	1 MSEYQ KSLEPKLG HISLSYTNEPG KKGQ NYLNALNITREN KLRLFQMNTG. DVNLLELREE. NVALYKFRNGV LITQLHKFRYQ.	10 FP.KESMGTT KPQLLSTGT .ISTYNGSC MLS SKM.AAVHGT .MQVITASCP 90 TDDSTTLVAS .LQPRVDQCT .RPIDPKKFM E.LSSDALP FHEMSVESVP	20 TLATOCTO TLAAIPCG VLAMAGDH IVCLQGPD TLSFIYNG AIGICGDD AIVAAKYDG 100 SNVTRSLI. NILASRL. MLVKSTLY HHFIRHTMA KSILVNEL. RCVSNTLY	30 GVVLASDES GVVLGADT CVAIASDES WVLIAADS GIVVAVDSI GILLASDLS VRYRQY FQYMGY FQYMGY YRYKGY YRYKGY	ATSSGSFI ATAGPIV LGVNMLT SVSSSIC ATGGQII LSVDYSI IIQ ISAGVIV .FVTPVI JQALIV .FVTPVI VEVNMLL .XINILV
D	$\begin{array}{c} Tv_beta_1\\ Tv_beta_2\\ Tv_beta_4\\ Tv_beta_5\\ Tv_beta_6\\ Tv_beta_7\\ \end{array}$	40 PNRATNKITI AVKDEMKLH MSENYDRIA FSQTVMKIL DSRHKARIF RHNNVSHFVI	50 STQPKIFA ZISDNIW ZINDRIYL ZINDRHAL ZLADRHAL ZAPNMIG ZWAPNIII	60 ARCENAADT CGAGIAADX AMSGETGDC TMAGGAADC SATGFOGDI GASGEFADF	PDPCQFVKN GEFREN CQFLARAVKN IDNINAVISA VLTVREQLRF LQLSEYLQG QYWLRNLSP DAFITRMRS QTLIEVIKS	1 MSEYQ IKSLEPKLG IKSYTNEPG IKKGQ IKKGQ IVINALNITREN KURLFQMNTG. DVNLLEREE. IVVALYKFRNGV LIQLHKFRYQ. IILL.NYENQH VILQQQCKHN.	10 FP.KESMGTT KPQLLSTGTT .ISTYNGSC MLS KM.AAVHGTT WSPYEMHGGT .MQVITASGP 90 TDDSTILVAS .LQPRVDQCT .RPIDPKKFM 'E.LSSDALP .QPLTVAAAS GEYLTASEVE	20 TIAAIFOG VIAAAFDG VIAAAGDH IVGLQGPD TISFIYNG AIGICGDD IVAAKYDG SNVIRSLI. NNILASRL. NIVKSTLY HFIRHTMA KKILVNEL. RCVSNTLY NNIKRYMY	30 GVVNASDSS GVVLGADT CVAIASDRI WVLIAADS GIVVAVDSS YVVIGADTI GILLASDLS VRYRQY FQYMGY EKRFSPF KAVRKSP SKRFFY QCRSNMK.I	ATSSGSFI ATAGPIV LGVNMLT SVSSSIIC ATGGQFI LSVDYSI LSAGVIV IQAALIV .FVTPVI YEVNMLL ULSIGSMI .YINIL 2LSCKVIV
U	$\begin{array}{c} Tv_beta_1\\ Tv_beta_2\\ Tv_beta_4\\ Tv_beta_5\\ Tv_beta_6\\ Tv_beta_7\\ \\ Tv_beta_2\\ Tv_beta_2\\ \\ Tv_beta_3\\ Tv_beta_4\\ \\ Tv_beta_5\\ Tv_beta_6\\ \\ Tv_beta_7\\ \end{array}$	40 PNRATNKITI AVKDEMKLITI SKDFKRIFO MSENYDRIA FSQTVMKILI DSRHKARIFV	50 SIQPKIFA ZIQPKIFA ZINDNIW ZINDRIYI ZLDRHAI ZLDPNMIG KMNSNCMI ZVAPNIII	60 TKALEPAHV ME CGAGIAADT CGAGIAADN GLAGLATD JAMSGETGDC SATGFDGDI GASGEFADF	FDFSNYARN FDPCQFVKN GEFREN CQFLARAVKN IDNINAVISA VITVREQLRF IQLSEYLQG QYWLRNLSP DAFITRMRS QTLIEVIKS	1 MSEYQ IKSLEPKLG MSD 	10 FP.KESMGTT KPQLLSTGTT .ISTYNGSC MLS KM.AAVHGT WSPYEHGGT .MQVITASGP 90 TDDSTILVAS .LQPRVDQCT .RPIDPKKFM E.LSSDALA .QPLTVAAS FHEMSVESVF GEYLTASEVF	20 ILAIQCTD INAAIFDG VLAMAGDH IVGLQGPD ITLSFIYNG AIGICGDD IVAAKYDG IO0 SNVIRSLI. NILASRL. NILKSTLY HFIRHTMA SKILVNEL. ARCVSNTLY INYIKRYMY	30 GVVMASDSS GVVLGADTI CVAIASDRI WVLIAADS GIVVAVDSI YVVIGADTI GILLASDLS VRYRQY FQYMGY FKRFSPF YRYKGYI SKRFFPY QCRSNMK.I	ATSSGSFI ATAGPIV LGVNMLT SVSSSIC LSVDYSI SITYGSMF 110 LSAGVIV FVTPVI VEVNMLL VEVNMLL VEVNMLL VEVNMLL VEVNMLL VEVNMLL VEVNMLL
J	$\begin{array}{c} Tv_beta_1\\ Tv_beta_2\\ Tv_beta_4\\ Tv_beta_5\\ Tv_beta_6\\ Tv_beta_7\\ \\ Tv_beta_2\\ Tv_beta_3\\ Tv_beta_3\\ Tv_beta_4\\ Tv_beta_5\\ Tv_beta_6\\ Tv_beta_7\\ \end{array}$	40 PNRATNKIT AVKDEMKIH VSKDFKRIF MSENYDRIA FSQTVMKIL DSRHKARIF RHNNVSHFVI 120	50 2TQPKIFA 2ISDNIW 2INDRIYL 2LDDRHAI 2LDPNMIG KMNSNCMI 2VAPNIII 313	TKALEPAHU GO CGACIAADT CGACIAADN CLACLATDV TMACGAADC SATGFDGDI GASCEFADF	FDFSNYARN PDPCQFVKN GEFREN 	1 MSEYQ IKSLEPKLG MSD IHISLSYTNEPG KKGQ IXINALNITREN KLRLFQMNTG. DVNLLELREE. DVNLLELREE. IJQLHKFRYQ. IIIQLHKFRYQ. IIIQLHKFRYQ. IIIQLHKFRYQ. IIIQLHKFRYQ. IIIQ	10 FP.KESMGTT KPQLLSTGT .ISTYNGSC MLS KM.AAVHGTT .MQVITASCP 90 TDDSTILVAS .LQPRVDQCT .RPIDPKKFM F.LSSDAL2 .QPLTVAAAS FHEMSVESVA GEYLTASEVF	20 TIAAIQCTD TIAAIFDG VIAMAGDH TUSFIYNG TISFIYNG 100 NVTRSLT. NILASRL. NILASRL. NILASRL. RFIRHTMA SKILVNEL. RCVSNTLY INYLKRYMY 60	30 GVVMASDS GVVLGADT CVAIASDR WVLIADS GIVVAVDS GIVAVDS GILLASDLS 	ATSSGSFI ATAGPIV ALGVNMLT SVSSSIIC ATGGQFI LSVJYSI IIQ LSAGVIV IQAALIV .FVVTPVI VEVNU VEVNU VEVNU VEVNU 180
J	$\begin{array}{c} Tv_beta_1\\ Tv_beta_2\\ Tv_beta_3\\ Tv_beta_5\\ Tv_beta_6\\ Tv_beta_7\\ Tv_beta_1\\ Tv_beta_2\\ Tv_beta_3\\ Tv_beta_3\\ Tv_beta_5\\ Tv_beta_6\\ Tv_beta_7\\ Tv_beta_1\\ Tv_beta_1\\ \end{array}$	40 PNRATNKIT AVKDEMKIH VSKDFKIF MSENYDRIA FSQTVMKIL DSRHKARIFI RHNNVSHFVI 120 GEWDSAG	50 EIQPKIFA ZIQPKIFA ZISDNIW ZINDRIYL LDDRHAI LAPNMIG ZWAPNIII ZVAPNIII 13 20 VYSIEV	TKALEPAHU TKALEPAHU ME	FDFSNYARN FDPCQFVKN GEFREN 70 CQFLARAVKN IDNINAVISA VLTVREQLRF LQLSEYLQG QYWLRNLSP DAFITRMSS QTLIEVIKS 140 IIASNESGST	1 MSEYQ IKSLEPKLG MSD 	10 FP.KESMGTT KPQLISTGTT STYNGSC STYNGSC MLS WSPYEMHGGT .MQVITASGP TDDSTTLVAS .LQPRVDQCT LSSDALP .QPLTVAAAS GEYLTASEVF	20 LIATOCTU TAAAIFDG VLAMAGDH IVGLOGPD IIVGLOGPD IIVAAKYDG 100 SNVTRSLT. NILASRL. NILASRL. KILVNEL. KILVNEL. NYLKRYMY 60 DMTMEEAT	30 GVVMASDSS GVVLGADT CVAIASDRI WVLIAADS GIVVAVS YVVIGADT GILLASDLS VRYRQY FQYMGY FQYMGY SKRFFPY QCRSNMK.I 170 KFAIAAVTC	ATSSGSFI ATAGPIV LGVNMLT SVSSSIC ATGGQFI LSVDYSI LSAGVIV IQAALIV .FVTPVI .YEVNMLL ULSIGSMI LSCKVIV 180 AIIIRDGS
d	$\begin{array}{c} Tv_beta_1\\ Tv_beta_2\\ Tv_beta_3\\ Tv_beta_5\\ Tv_beta_6\\ Tv_beta_7\\ \\ Tv_beta_1\\ Tv_beta_2\\ Tv_beta_4\\ Tv_beta_6\\ Tv_beta_6\\ Tv_beta_6\\ Tv_beta_7\\ \\ \\ Tv_beta_1\\ Tv_beta_1\\ Tv_beta_2\\ \end{array}$	40 PNRATNKITH AVKDEMKLH SENYDRIA FSQTVMKIL DSRHKARIF RHNNVSHFVI 120 GEWDS.AG GEIDF.QG	50 STQPKIFA VISDNIWU LDDRHAL LDDRHAL VAPNNII VAPNIII VAPNIII	60 ARCGNAADT CGAGIAADT CGAGIAADT CGAGIAADC TMAGGAADC SATGFOGDI GASGEFADF	PDPCQFVKN 	1 MSEYQ IKSLEPKLG IKSLEPKLG IK	10 FP.KESMGTT KPQLLSTGTG .ISTYNGSC MLS KM.AAVHGT WSPYEMHGGT .MQVITASGP TDDSTILVAS .LQPRVDQCT .RPIDPKKFM E.LSSDALA .QPLTVAAAS GEYLTASEVF	20 TIAAIGCTD TIAAIGCTD VIAMAGDH IVGLQGPD ITLSFIYNG AIGICGDD IVAAKYDG 100 SNVIRSLI. NNIRSLI. NNIRSTLY NVKSTLY NVKSTLY NYKRYMY 60 CDMTMEEAT	30 GVVMASDSS GVVLGADT CVAIASDRI WVLIAADS GIVVAVDSS YVVIGADTI GILLASDLS VRYRQY FQYMGY SKRFSPF KAVRKSP SKRFSPF XAVRKSP QCRSNMK.I 170 KFAIAAVTC EMVADATY	ATSSGSFI ATAGPIV LGVNMLT SVSSSIIC ATGGQFI LSVDYSI IIQ LSAGVIV IQAALIV VSIGSMI VSIGSMI LSCKVIV 180 SAIIRDGS
D	Tv_beta_1 Tv_beta_2 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_7	40 PNRATNKITI AVKDEMKLIHI VSKDFKRIFG MSENYDRIA FSQTVMKILI DSRHKARIF HNNVSHFVI 120 GEWDS.AGG GIDF.QG ACLPETNE	50 SIQPKIFA VISDNIWU JINDRIYI LDDRHAL SLAPNMIG (MNSNCMI I VAPNIII 13 QVYSIEV 20VYQVAP YLAASDS	TKALEPAHU TKALEPAHU ME CGAGIAADN GLACLATDV MSGETGDC TMAGGAADC SATGFDGDI GASGEFADF SGMA.IKKK HCSF.SKQF	FDFSNYARN FDFCQFVKN GEFREN	1 MSEYQ IKSLEPKLG MSD 	10 FP.KESMGTT KPQLLSTGTT .ISTYNGSC 	20 ILA IQCTD ILAA IQCTD VLAMAGDH IVGLQGPD ITLSFIYNG AIGICGDD IVAAKYDG IO0 SNVIRSLI. NILASRL. NILVKSTLY HFIRHTMA SKILVNEL. RCVSNTLY INYIKRYMY 60 DMTMEEAT IKMNEHDCM	30 GVVMASDSS GVVLGADTI CVAIASDR WVLIAADSS YVVIGADTI GILLASDLS VRYRQY FQYMGY FQYMGY SKRFSPF QCRSNMK.I 170 KFAIAAVTC ECTAKCLIJ	ATSSGSFI ATAGPIV ALGVNMLT SVSSSIC ATGQCI LSVDYSI SITYGSMF 110 LSAGVIV IQAALIV VEVNMLL VEVNMLL VEVNMLL VEVNMLL VEVNMLL VEVNMLL VEVNMLL SGITNDLG AIIRDGS GITNDLG AVERDSI
D	Tv_beta_1 Tv_beta_2 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_3 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_2 Tv_beta_3 Tv_beta_2 Tv_beta_3 Tv_beta_3 Tv_beta_3 Tv_beta_3 Tv_beta_3 Tv_beta_3 Tv_beta_3 Tv_beta_3 Tv_beta_3	40 PNRATNKIT AVKDERKIH VSKDFKRIF MSENYDRIA FSQTVMKIL DSRHKARIFI RHNNVSHPVI 120 GEWDS.AG GEIDF.QG GEIDF.QG GEIDF.QG GIDF.QG GIDF.QG	50 SIQPKIFA SISDNIWU SISDNIWU LISDNIWU LAPNMIG SANSNCMI SVAPNIII SVAPNIII SVAPNIII SVYSIEV SQVYSIEV SVYQVAP SVLASSS HLYFMY	TKALEPAHU TKALEPAHU GQ CGAGIAADI CGAGIAADI GLAGLATDV AMSCETGDC TMAGGAADC SATGFDGDI GASGEFADF S CGASLATAF S C CGASLATAF S C C C C C C C C C C C C C	FDFSNYARN FDPCQFVKN GEFREN CQFLARAVKN IDNINAVISA VITVREQLRF LQLSEYLQG QYWLRNISF DAFITRMRS QTLIEVIKS 140 KIASNESGSI FFAQGSCSI DFAVAGSCSI DFAVAGSCSI DFAVAGSCSI	1 MSEYQ IKSLEPKLG MSD HISLSYTNEPG IKKGQ KGQ IYINALNITREN KLRLFQMNTG. DVNLLELREE. IIQLHKFRYQ. IIL.NYENGV 	1 0 FP.KESMGTT KPQLISTGTT 	20 LIATOCTPO TIAAITOC VIAMAGDH IVGLOGPD IIVGLOGPD IIVGLOGPD IIVGLOGPD IIVAKYDG IO0 IVAKYDG IO0 IVAKYDG INVIRSLI. NILASRL. KILVNEL. KILVNEL GO DMTMEEAT KKNEHDCM IMNPDELF GLILEDGK	30 GVVNASDSS GVVLGADT CVAIASDRI WVLIAADS GIVVAVDSS YVVIGADT GILLASDL CVAIASDL GIVAVVS SVRFSPF EKRFSPF XAVRSP- SKRFFY QCRSNMK- I 170 KFAIAAVTC ECTAKCLII ELMKLALN	ATSSGSFI ATAGPIV LGVNMLT SVSSSIC ATGGQFI LSAGQFI LSAGVIV IQALIV FVTPVI VEVINLI VEVNMLI VEVNMLI VEVNMLIV LSCKVIV 180 AGIIRDGS AGIIRDGS AVERDSI
u	Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_7 Tv_beta_2 Tv_beta_3 Tv_beta_3 Tv_beta_3 Tv_beta_4 Tv_beta_7 Tv_beta_1 Tv_beta_2 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_4 Tv_beta_5	40 PNRATNKITH AVKDEMKLIH SKDFKRIF MSENYDRIA FSQTVMKILI DSRHKARIFI RHNNVSHFVI 120 GEWDS.ACG GCIDF.QG ACLPETNE SGYDG.KI CGYDN.TGG	50 2 TQP KIFA 2	TKALEPAHU TKALEPAHU ME GQ CGAGIAADU GLACLADU AMSGETGDC TMAGGAADC TMAGGAADC GASGEFADF GS SATGFOGDI GASGEFADF GS SGMA.IKKK HGSF.SKQP IGAFAFPKD IGAFAFPKD IGAFAFPKD	PDPCQFVKN PDPCQFVKN GEFREN 	1 MSEYQ IKSLEPKLG IKSLEPKLG IKKGQ IKKGQ IKKGQ IKKGQ IVILALNITREN KLRLFQMNTG. DVNLLELREE. IVVALYKFRNGV LIQLHKFRYQ. IILL.NYENQH VILQQQCKHN. 150 YYIQAYIDQNYR .AAISVLENRWH SLYGICESAWR FVMSVFDKHYK IAAGVLDTCYR	1 0 FP.KESMGTT KPQLLSTGT .ISTYNGSC MLS KM.AAVHGTT WSPYEMHGGT .MQVITASGP 90 TDDSTILVAS .LQPRVDQCT .RPIDPKKFM 'E.LSSDALP .QPLTVAAAS GEYLTASEVF I 	20 TIAAIFOG VIAAAGDH IVGLQGPD IIVGLQGPD IVAAKYDG 100 SNVIRSLI. NILASRL. NIVKSLY HFIRHTMA KILVNEL. KIVNEL COMTMEEAT KMNEHDCM NMNPDELF SCHLEDGK	30 GVVNASDSS GVVLGADT CVAIASDRI WVLIAADS GIVVAVDS YVVIGADT GILLASDL EKRFSPF KAVRKSP. SKRFFSPF XAVRKSP. 2CRSNMK.I 170 KFÄIAAVT ENVADAIYJ ECTAKCLII ELMKLALM	ATSSGSFI ATAGQFIV SVSSSIIC ATGGQFI LSAGQVIV IQAALIV .FVTPVI VEVNMLL ULSIGSMI VEVNMLL ULSIGSMI AIIRDGS AGITNDLG AAVERDSI IKQRFTV
d	Tv_beta_1 Tv_beta_2 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_2 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_5 Tv_beta_7 Tv_beta_7 Tv_beta_1 Tv_beta_2 Tv_beta_2 Tv_beta_3 Tv_beta_3 Tv_beta_3 Tv_beta_3 Tv_beta_5 Tv_beta_5 Tv_beta_5 Tv_beta_5 Tv_beta_5 Tv_beta_5 Tv_beta_7	40 PNRATNKITI AVKDEMKLHT AVKDEMKLHT SENYDRIA FSQTVMKILI DSRHKARIF RHNNVSHFVI 120 GEWDS.AGG GEIDF.QG AGLLPETNE SYDGKG GGINSEG.KK	50 EIQPKIFA ZIQPKIFA ZISDNIWU ZINDRIYI ZLDPRHAI ZLDPNHIG MNSNCMI ZVAPNIII 20VYQVAP YLAASDS HIFYIDN SKIYGYDP	TKALEPAHU TKALEPAHU ME	70 70 70 70 70 70 70 70 70 70	1 MSEYQ IKSLEPKLG NSD IK.S.YTNEPG IKKGQ IK.N.KLRLFQMNTG. DVNLLELREE. NVALYKFRNGV LTQLHKFRYQ. ILL.NYENQH VVILQQQCKHN. 150 ILL.NYENQH SLYGICESAWR FVMSVFDKHYK ILAZYLDTCYR ILAZYLDTCYR	10 FP.KESMGTT KPQLLSTGTT .ISTYNGSC .ISTYNGSC .WSPYEMHGGT WSPYEMHGGT .MQVITASGP 90 TDDSTILVAS .LQPRVDQCT .RPIDPKKFM E.LSSDALA .QPLTVAAS .QPLTVAAS .GEYLTASEVF 	20 LIATQCTD IAAAIFDG VIAMAGDH IVGLQGPD IIVGLQGPD IVAAKYDG 100 SNVIRSLI. NNILASRL. MNVKSTLY HFIRHTMA SKILVNEL. ARCVSNTLY COMMEEAT NNYLKRYMY COMMEEAT NNNPDELF GLTLEDGK CALQADKEEAC	30 GVVMASDSS GVVLGADTI CVAIASDRI WVLIAADSS YVVIGADTI GILLASDLS VRYRQY FQYMGY KRSPF YRYKGYI SKRFFY QCRSNMK_I 170 KFAIAAVTC ECTAKCLII ELGRRANY NIVRDAICS	ATSSGSFI ATAGPIV LGVNMLT SVSSSIIC LSVDYSI IIYGSMF IIQ ISAGVIV IQAALIV VEVIMLL VEVIMLL VEVIMLL VEVIMLL SGITNILO AIIRDGS AIIRDGS AIIRDGS IKQRFTV IATYRDSG
D	Tv_beta_1 Tv_beta_2 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_2 Tv_beta_3 Tv_beta_3 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_1 Tv_beta_2 Tv_beta_4 Tv_beta_5 Tv_beta_7	40 PNRATNKIT AVKDEKLIH VSKDFKRIF(MSENYDRIA(MSENYDRIA(SRHKARIF) RHNNVSHFVI 120 GEWDS.AG GGIDF.QG GGIDF.QG GGIDF.QG GGIDF.QG GGIDF.QG GGIDF.QG GGIDF.QG GGIDF.QG GGIDF.QG GGINSGG.K AGINFDG.SS	50 21 QP KIFA VISDNIWU 21 NDRIYI 21 NDRIYI 21 APNMIG (MNSNCMI 24 APNMIG (MNSNCMI 24 APNIII 24 APNIII 25 APNIII 25 APNIII 25 APNIII 26 APNIII 26 APNIII 27 APNIII 27 APNIII 28 APNIII 28 APNIII 29 APNIII 29 APNIII 20 APNIIII 20 APNIIII 20 APNIII 20 APNIII 20 APNIII 20 APNIII 20 APNI	TKALEPAHU TKALEPAHU ME CGAGIAADN CGAGIAADN GLAGLATDV TMAGGAADC SATGFDGDI TMAGGAADC SATGFDGDI GASGEFADF CGASIEFA	FDFSNYARN FDFCQFVKN QFLARAVKN IDNINAVISA ILTVREQLRF LQLSEVLQG QYWLRNLSP DAFITRMRS QTLIEVIKS 140 IAQCSGSI FFIAQCSGSI FFIAQCSGSS VGAQCYCQY CRFSVGSGSS HIGTGFGKY	1 MSEYQ IKSLEPKLG MSD HISLSYTNEPG IKKGQ IYINALNITREN .KLRLFQMNTG. DVNLLELREE. DVNLLELREE. IVQLYKFRNGV LIQLHKFRYQ. IILL.NYENQH VILQQQCKHN. 150 YYIQAYIDQNYR .AAISVLENRWH SLYGICESAWR FVMSVFDKHYK HAYGVLDTCYR IAAFLDSAFG LQGLQIADV	10 FP.KESMGTT KPQLLSTGT .ISTYNGSC MLS KM.AAVHGTT WSPYEMHGT .MQVITASGP 90 TDDSTILVAS .LQPRVDQCT .RPIDPKKFM E.ISSDALP .QPLTVAAAS FHEMSVESVF GEYLTASEVF 	20 TIAAIQCTD TIAAIFDG VUAMAGDH TUSFIYNG TISFIYNG 100 SNVTRSLI. NIUASRL. NIUKSTLY HFIRHMA SKILVNEL. RCVSNTLY INYIKRYMY 60 DMTMEEAT IKMNEHDCM NMNPDELF SCILEDGK SDMTKEEAC	30 GVVMASDS GVVLGADTI CVAIASDR WVLIAADS: GIVVAVDS; YVVIGADTI GILLASDLS VRYRQY FQYMGY FQYMGY FQYMGY KRSP.T YRYKGY YRYKGY KRSPFFY QCRSNMK.I 170 KFAIAAVTT EMVADAIYJ ECTAKCLIJ ELGRRAIY NIVRDAICZ KGITEVFRJ	ATSSGSFI ATAGPIV ALGVNMLT SVSSSIC LSVDYSI SITYGSMF 110 LSAGVIV IQAALIV JQAALIV JQAALIV VEVNMLI VEVNMLI VEVNML VEVNLI VEVNLI SGITNDLG AVERDSI IKQRFTV IATYRDSG VTERDIY VNARNTT
d	Tv_beta_1 Tv_beta_2 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_3 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_3 Tv_beta_2 Tv_beta_7	40 PNRATNKIT AVKDEMKIH YSKDFRRIF MSENYDRIA FSQTVMKIL DSRHKARIFT RHNNVSHFVI 120 GCWDS.AGI GCIDF.QGG ACLIPETNE SCYDGKI CGYDN.TGG KACINPDG.SI	50 2 TQPKIFA 2 TQPKIFA 2 TDQNU 2 TDDRHAT 2 LAPNMIG 2 MNSNCMI 3 2 QVYSIE 9 QVYSIE 9 YLAASDS 9 HLYFMDY 9 HLYGDP 4 KLYGDP 4 KLYGDP	TKALEPAHV TKALEPAHV ME	FDFSNYARN FDPCQFVKN FDPCQFVKN GEFREN 70 CQFLARAVKN IDNINAVISA LTVREQLRF LQLSEYLQG QYWLRNLSP QYWLRNLSP QYULRVKS 140 IIASNCSGSI PFAVAGTCEE YGAQGYCQY RFSVGSGSI DHIGTGFGKY	1 MSEYQ IKSLEPKLG IKSLEPKLG IKKGQ IKKGQ IKKGQ IYINALNITREN KKLRLFQMNTG. DVNLLELREE. NVALYKFRNGV LIQLHKFRYQ. IILL.NYENQH VILQQQCKHN. 150 YIQAYIDQNYR .AAISVLENRWH SLYGICESAWR FVMSVFDKHYK HAYGVLDTCYR HAAYGVLDTCYR ILAPLLDSAFG ILQGLQIADV	1 0 FP.KESMGTT KPQLISTGTT 	20 TIAAITOCTU TIAAITOC VIAMAGDH IVGLQGPD IIVGLQGPD IVGAKYDG NVTRSLT. NILKSTLY NIVKSTLY NIVKSTLY NIVKSTLY NYKRYMY 60 DMTMEEAT KKNYMELDCK NMNPDELF GGITLEDGK VNGSFDDVK	30 GVVMASDSS GVVLGADT CVAIASDRI WVLIAADS GIVVAVDS YVVIGADT GILLASDLS VRYRQY FQYMGY SKRFFPY QCRSNMK.I 170 KFAIAAVTC EMVADAIY ECTAKCLIJ ELMKLAIM ELGRRAIYI NIVRDA	ATSSGSFI ATAGPIV LGVNMLT SVSSSIC ATGGQFI LSAGQVIV IQAALIVV ICAALIVV ISIGSMI VLSIGSMI ISSCKVIV 180 SAIIRDGS AVERDSI IKQRFTV IATYRDSG VTERDIY
d	$\begin{array}{c} Tv_beta_1\\ Tv_beta_2\\ Tv_beta_3\\ Tv_beta_5\\ Tv_beta_6\\ Tv_beta_7\\ Tv_beta_2\\ Tv_beta_3\\ Tv_beta_3\\ Tv_beta_3\\ Tv_beta_5\\ Tv_beta_6\\ Tv_beta_7\\ Tv_beta_1\\ Tv_beta_2\\ Tv_beta_2\\ Tv_beta_3\\ Tv_beta_6\\ Tv_beta_7\\ Tv_beta_3\\ Tv_beta_6\\ Tv_beta_3\\ Tv_beta_7\\ Tv_beta_4\\ Tv_beta_5\\ Tv_beta_5\\ Tv_beta_6\\ Tv_beta_7\\ Tv_b$	40 PNRATNKITH AVKDEMKIH AVKDEMKIH FSQTVMKIJ DSRHKAIF RHNNVSHFVI 120 GEWDS.AC GCIDF.QC ACLPETNE SYDG.K CYDN.TC GCINSEG.K CYDN.TC GCINFOG.S 190	50 2 TQP KIFA 2	TKALEPAHU TKALEPAHU ME	PDPCQFVKN PDPCQFVKN GEFREN	1 MSEYQ IKSLEPKLG IKSLEPKLG IKKGQ IK	10 FP.KESMGTT KPQLLSTGTG .ISTYNGSC MLS KM.AAVHGT WSPYEMHGGT .MQVITASGP TDDSTILVAS .LQPRVDQCT .RFIDPKKFM E.LSSDALP .QPLTVAAAS GEYLTASEVF GEYLTASEVF 	20 TIAAIQCTD TIAAIFDG VIAMAGDH IVGLQGPD ITJSFIYNG AIGICGDD IVAAKYDG 00 SNVIRSLI. NILASRL. NIVRSLIY HFIRHTMA KILVNEL. UNUKSTLY HFIRHTMA KILVNEL. COMTMEEAT KMNEHDCM NMNPDELF COMTKEEAC CAVSLQDAK NGSFDDVK	30 GVVMASDSS GVVLGADT CVAIASDRI WVLIAADS GIVVAVDS YVVIGADTI GILLASDL VRYRQY FQYMGY EKRFSPF KAVRKSP SKRFFSPF XAVRKSP SKRFG VRKSP QCRSNMK.I 170 KFAIAAVTT ENVADAIYJ ECTAKCLI ELMKLALM ELGRRAIYH NIVRDAICS	ATSSGSFI ATAGPIV LGVNMLT SVSSSIIC ATGGQFI LSAGQVIV IQAALIV .FVTPVI .FVTPVI LSIGSMI .SITROGS AGITNDLG AAVERDSI IKQRFTV IATYRDSG SVTERDIY
d	Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_1 Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_3 Tv_beta_3 Tv_beta_5 Tv_beta_7	40 PNRATNKITT AVKDEMKLITT AVKDEMKLITT AVKDEMKLIT SENYDRIAG FSQTVMKILI DSRHKARIF RHNNVSHFVT 120 GEWDS.AG GEIDF.QG AGLUPETNE SCYDG.K GYDN.TG GGINSEG.K AGINPDG.S 190 SGGVVNIVQ	50 EIQPKIFA VISDNIWU LDDRIYI LDDRIYI LDDRHAI LAPNMIG WAPNIII VAPNIII QVYQVAP YLAASDS HIFFUDN SKLYGYDP ZFLACTDP 20 10 20 20 20 20 20 20 20 20 20 2	TKALEPAHV TKALEPAHV MEAGLADT CGACIAADT CGACIAADT GLAGLATDV GLAGLATDV GLAGLATDV GLAGLATDV GSATGFOGI SATGFOGI SATGFOGI GASGEFADF GO SGMA.IKKK HGSF.SKQF HGSR.IAGK VGTI.EDLH VGTI.EDLH VGTI.EDLH	70 70 70 70 70 70 70 70 70 70	1 MSEYQ IKSLEPKLG NSD IKSYTNEPG IKKGQ IKKGQ IVINALVITREN KKLRLFQMNTG. DVNLLELREE. NVALYKFRNGV LITQLHKFRYQ. IILL.NYENQH VVILQQQCKHN. 150 IYIQAYIDQNYR AAISVLENRWH SLYGICESAWR FVMSVFDKHYK IAAFLLDSAFG IQGLQIADV	10 FP.KESMGTT KPQLLSTGT .ISTYNGSC .ISTYNGSC WSPYEMHGGT WSPYEMHGGT .MQVITASGP 90 TDDSTILVAS .LQPRVDQCT .RPIDPKKFM E.ISSDALA .QPLTVAAAS FHEMSVESVP GEYLTASEVF 	20 ILAIQCTD INAAIFDG VIAMAGDH IVGLQGPD IIVAAKYDG IOO SNVIRSLI. NNILASRL. MNUKSTLY HFIRHTMA KILVNEL. ARCVSNTLY CDMTMEEAT NMNYLKRYMY COMTMEEAT NMNPDELF GLTLEDGK VMNPDELF GLTLEDGK VNGSFDDVK	30 GVVMASDSS GVVLGADTI CVAIASDRI WVLIAADS GIVVAVDS YVVIGADTI GILLASDLS VRYRQY FQYMGY KRFSPF YRYKGYI YRYKGYI YRYKGYI KFFFY QCRSNMK I 170 KFAIAAVTC ELGRRANYI NIVRDAICS KGITEVFRJ	ATSSGSFI ATAGPIV LGVNMLT SVSSSIC LSVDYSI II LSAGVIV IVAALIV VEVIMLL VEVIML VEVIM
d	Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_2 Tv_beta_3 Tv_beta_3 Tv_beta_4 Tv_beta_5 Tv_beta_7 Tv_beta_5 Tv_beta_7 Tv_beta_6 Tv_beta_1 Tv_beta_4 Tv_beta_7 Tv_beta_1 Tv_beta_7	40 PNRATNKIT AVKDEMKIH VSKDFKRIFC MSENYDRIAC MSENYDRIAC MSENYDRIAC MSENYDRIAC MSENYDRIAC MSENYDRIAC MSENYDRIAC GGIDF.QGG GGIDF.QGG GGIDF.QGG GGIDF.QGG GGIDF.QGG GGIDF.QG GGIDF.QG GGIDF.QG GGINFAC GGINFAC GGINFAC GGINFAC GGINFAC GGINFAC GGINFAC GGINFAC SGGVVNIVQ SGGVVNIVQ SGSHVNICV	50 21 QP KI F A 21 SDNIW 21 NDRIYL 21 DDRHAI 22 DDRHAI 24 PNII 24 PN	TKALEPAHV TKALEPAHV ME GQ GCGACIAADT GCGACIAADT GCGACIATDV AMSCETGDC TMAGGAADC SATGFDGDI GASGEFADF GSGMA.IKKK HGSF.SKQF HGSF.SKQF HGSF.IAGK YGAS.WESD VGTI.QSIF HCSR.IAGK YGAS.WESD	FDFSNYARN FDFCQFVKN GEFREN	1 MSEYQ IKSLEPKLG IKSLEPKLG IKKGQ IKKGQ IYINALNITREN KLRLFQMNTG. DVNLLELREE. IVALYKFRNGV IIL.NYENQH IIL.NYENQH IJLQQQCKHN. 150 YIQAYIDQNYR AAISVLENRWH SLYGICESAWR FVMSVFDKHYK HAYGVLDTCYR IAAFUCISAWU IAAPLDSAFG IQGLQIADV	1 0 FP.KESMGTT KPQLISTGTT 	20 TIAAITOCTU TIAAITOG VIAMAGDH IVGLQGPD ITLSFIYNG AIGICGDD IVAAKYDG IOO SNVTRSLI. NIIASRL. NIIASRL. KILVNEL. KILVNEL SKILVNEL SKILVNEL SKILVNEL SKILVNEL SKILVNEL SKILVNEL SCANTA SKILVNEL SCANTA SKILVNEL SCANTA SKILVNEL SCANTA SKILVNEL SCANTA SKILVNEL SCANTA S	30 GVVNASDSS GVVLGADT CVAIASDRI WVLIAADS GIVVAVGADT GILLASDLS VRYRQY FQYMGY FQYMGY YRYKGY YRYKGY YRYKGY YRYKGY CRSNMK.B 170 KFAIAAVTC ECTAKCLII ECTAKCLII ELGRRAIY KGITEVFRJ	ATSSGSFI ATAGPIV LGVNMLT SVSSSIC ATGGQFI LSAGQFI LSAGVIV IQALIV FVTPVI LSAGVIV IQALIV VINILV SAGITNDLG AGITNDLG SAIIRDGS AVERDSI SVTERDSI SVTERDSI VVTARNTT
d	Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_3 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_2 Tv_beta_2 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_2 Tv_beta_3 Tv_beta_2 Tv_beta_3 Tv_beta_2 Tv_beta_3 Tv_beta_2 Tv_beta_3 Tv_bet	40 PNRATNKITH AVKDEMKITH AVKDEMKITH AVKDEMKITH SENYDRIA FSQTVMKIL DSRHKARIFI RHNNVSHFVI 120 GCWDS.AG GCIDF.QG AGIDF.QG AGINFOG.SI SCYDN.TG SGSHVNICV SGSHVNICV SGSHVNICV	50 2 TQP KIFA 2 TQP KIFA 2 TQP KIFA 2 TQP KIFA 2 TDD RHAT 2 TDD RHAT 2 TDD RHAT 3 TO 2 QVYS IE V 9 QVYQVAP 9 YLAASDS 9 QVYQVAP 9 YLAASDS 9 HIFY IDN 3 KIFY IDN 3 KIFY IDN 3 KIFY IDN 3 KIFY IDN 4 KIFY IDN 2 O 1 NADGA KR 1 KRENPED 1 TQD KYII 1 TQU YYI 1	TKALEPAHU TKALEPAHU MEAGLADU CGAGIAADU CGAGIAADU GLACLATDU AMSGETGDC TMACGAADC SATGFOGDI GASCEFADF CGASCEFADF CGASSCEFADF CGASAFAFKD CGASAFAFKD CGASAFAFKD CGASAFAFKD CGASAFAFKD CGASAFAFKD CGASAFAFKD CO CGASAFAFKD CO CGASAFAFKD CO CGASAFAFKD CO CGASAFAFKD CO CGASAFAFKD CO CGASAFAFKD	PDPCQFVKN PDPCQFVKN GEFREN	1 MSEYQ IKSLEPKLG SLEPKLG IKSLEPKLG IKKGQ NVALYNTREN IKLRLFQMNTG. DVNLLELREE. NVALYKFRNGV LIQLHKFRYQ. ILL.NYENQH VILQQQCKHN. 150 YTQAYIDQNYR .AAISVLENRWH SLYGICESAWR FVMSVFDKHYK HAYGVLDTCYR ILAPLLDSAFG LQGLQIADV	10 FP.KESMGTT KPQLLSTGT .ISTYNGSC .ISTYNGSC WSPYEMHCGT WSPYEMHCGT .MQVITASGP 90 TDDSTILVAS .LQPRVDQCT .RPIDPKFM YE.LSSDALP .QPLTVAAAS GEYLTASEVF I 	20 TIAAITQCTD TIAAITQCTD TIAAITQCTD TIAAITQCTD TIAAITQCTD TIAAITQCTD TISFIYNG	30 GVVMASDSS GVVLGADT CVAIASDRI WVLIAADS GIVVAVDS YVVIGADT GILLASDL EKRFSPF KAVRKSP. SKRFRSPF XAVRKSP. SKRFRSPF XAVRKSP. QCRSNMK.I 170 KFAIAAVT ELMKLALM ELMKLALM ELMKLALM SGITEVFRI DVHLEILDI	ATSSGSFI ATAGPIV LGVNMLT SVSSSIIC ATGGQFI LSVDYSI II LSAGVIV IQAALIV .FVTPVI LSIGSMI LSCKVIV 180 AIIRDGS AGITNDLG AAVERDSI IKQRFTV IATYRDSG VTERDIY VNARNTT
d	Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_3 Tv_beta_5 Tv_beta_7 Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_6 Tv_beta_7	40 PNRATNKITH AVKDEMKIHT AVKDEMKIHT SENYDRIA FSQTVMKIL DSRHKARFF RHNNVSHFVI 120 GWDS.AG GCIDF.QG ACLPETNE SYDG.XC CYDN.TG GGINSEG.XC CYDN.TG GGINSEG.S 190 SGGVVNIVQ SGSHVNICV SGWGIVYII APHGFIVKVI	50 STQPKIFA STQPKIFA VISDNIWU LDDRHAL LDDRHAL LDDRHAL LDRHAL LDRHAL LAPNMIG VAPNIII VAPNIII SVAPNIII SVAPNIII VAPNIII VALASDS HIYFYIDN SKLYGYDP 20 INADGAKR KRENPED INADGAKR	TKALEPAHU TKALEPAHU ME	PDPCQFVKN GEFREN	1 MSEYQ IKSLEPKLG MSD HISLSYTNEPG IKKGQ KGQ NVALYKFRNGV LIQLHKFRYQ NVALYKFRNGV LIQLHKFRYQ 150 	10 FP.KESMGTT KPQLLSTGTG .ISTYNGSC .ISTYNGSC .WSPYEMHGGT WSPYEMHGGT .MQVITASGF 90 TDDSTILVAS .LQPRVDQCT .LQPRVDQCT 	20 TIAAIQCTD TIAAIFDG VIAMAGDH IVGLQGPD IIVGLQGPD IVAAKYDG AIGICGDD IIVAAKYDG NVIRSLI NILASRL NILASRL NIVKSTLY NYIKSTLY NYIKRYMY COMTMEEAT KMNEHDCM NMNPDELF GDTKEEAC AVSLQDAK VNGSFDDVK TTERPLTLP	30 GVVMASDSS GVVLGADT CVAIASDRI WVLIAADS GIVVAVDS YVVIGADTI GILLASDL VRYRQY FQYMGY FXRFSPF KAVRKSP YRYKGYY QCRSNMK J 170 KFAIAAVTC KFAIAAVTC ELGRRAIYI NIVRDAICSKGITEVFRJ	ATSSGSFI ATAGPIV LGVNMLT SVSSSIIC LSVDYSI IIQ LSAGQVIV IQAALIV IQAALIV VSIGSMI JYEVNMLL VLSIGSMI JSCKVIV ISO AIIRDGS AIIRDGS AIIRDGS VTERDIY IXQFTV AVNARMTT
d	Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_2 Tv_beta_3 Tv_beta_3 Tv_beta_4 Tv_beta_5 Tv_beta_7 Tv_beta_7 Tv_beta_6 Tv_beta_1 Tv_beta_1 Tv_beta_1 Tv_beta_2 Tv_beta_3 Tv_beta_4 Tv_beta_5 Tv_beta_3 Tv_beta_4 Tv_beta_7	40 PNRATNKIT AVKDEKLIH VSKDFKRIF MSENYDRIA FSQTVMKIL DSRHKARIF RHNNVSHFVI 120 GEWDS.AG GGIDF.QG GGINFDG.XX AG NPDG.SS 190 SGGVVNIVQ SGGVVNIVQ SGGVVVII SGGRVSVH CGNVC	50 STQPKIFA STQPKIFA STQPKIFA VISDNIWU UINDRIYL LAPNMIG CMNSNCMI SUAPNII SUAPNIII SUAPNII SUAPNIII SUAPNIII SUAPNIII SUAPNIII SUAPNIII SUAPNIII SUAPNIII SUAPNIII SUAPNIII SUAPNIII SUAPN	TKALEPAHU TKALEPAHU ME	FDFSNYARN FDFCQFVKN GEFREN	1 MSEYQ K. SLEPKLG MSD. HISLSYTNEPG K. KQQ VINALVITREN KLRLFQMNTG DVNLLELREE DVNLLYKFRNGV LIQLHKFRYQ IIL. NYENQH VILQQCKHN. 150 YIQQYIDQNYR AAISVLENRWH SLYGICESAWR FVMSVFDKHYK LAAPLLDSAFG LQGLQIADV.	10 FP.KESMGTT KPQLLSTGT .ISTYNGSC MLS KM.AAVHGT WSPYEMHGGT .MQVITASGP 90 TDDSTILVAS .LQPRVDQCT .RPIDPKKFM C.LSSDALA .QPLTVAAS .QPLTVAAS 	20 ILAIQCTD ILAAIGCTD VLAMAGDH IVGLQGPD IVJAKGDH IVGLQGPD IVAAKYDG 100 SNVIRSLI. NILASRL. NILKSTLY HFIRHTMA KILVNEL. ARCVSNTLY 60 SDMTMEEAT SDMTMEEAT SDMTKEEAC SDMTKEEAC SAVSLQDAK NGSFDDVK	30 GVVMASDS: GVVLASDR: VVIGADS: CVAIASDR: WVLIAADS: SVVVGADTI GILLASDL: EKRESPE YRYKGYI KRYKKSP. YRYKGYI KFAIAAVTC CCRSNMK.I 170 KFAIAAVTC ELGRRANATI ELGRRANATI NIVRDAIC KGITEVFRJ	ATSSGSFI ATAGPIV ALGVNMLT SVSSSIC ATGQCFI LSVDYSI SITYGSMF 110 LSAGVIV FVTPVI VEVINLL
	Tv_beta_1 Tv_beta_2 Tv_beta_4 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_2 Tv_beta_3 Tv_beta_3 Tv_beta_5 Tv_beta_6 Tv_beta_7 Tv_beta_1 Tv_beta_4 Tv_beta_6 Tv_beta_7 Tv_beta_1 Tv_beta_2 Tv_beta_4 Tv_beta_2 Tv_beta_7	40 PNRATNKIT AVKDEMKIH VSKDFKRIFC MSENYDRIAC MSENYDRIAC MSENYDRIAC MSENYDRIAC MSENYDRIAC MSENYDRIAC SCHUDS COMDS.ACC GGINS.ACC GGINS.ACC GGINSCA AGINPDG.S 190 SCGVVNIVQ SCSHVNLCV SCSHVSVH TCDALQLCV	50 SIQPKIFA SI	TKALEPAHV TKALEPAHV MEAGLAPY ARCGNAADT CGAGIAADV GLAGLATDV AMSCETGDC SATGFDGC SATGFDGC SATGFDGC SATGFDGC SGMA.IKKK HGSF.SKQF JGAFAFPKD JGTL.QSIF HGSF.IAGK YGAS.WESD QC QC MTVRPAQQF KQSKVIYTF KEIKTRMD JDKTDVFDM JEFFLPRH.	FDFSNYARN FDFCQFVKN GEFREN	1 MSEYQ IKSLEPKLG IKSLEPKLG IKKGQ IHISLSYTNEPG IKKGQ IYINALNITREN KLRLFQMNTG. DVNLLELREE. NVALYKFRNGV LIQLHKFRYQ. IILNYENQH VILQQQCKHN. 150 YIQAYIDQNYR .AAISVLENRWH SLYGICESAWR FVMSVFDKHYK HAYGVLDTCYR IAAPLDSAFG IQGLQIADV	10 FP.KESMGTT KPQLISTGTT 	20 TIAAITOCTU TIAAITOG VLAMAGDH IVGLQGPD ITLSFIYNG AIGICGDD IVAAKYDG SNVTRSLI. NILASRL. NILASRL. KILVNEL. KILVNEL. GO DMTMEEAT KKNEHDCM NMNPDELF GGLTLEDGK VNGSFDDVK	30 GVVMASDSS GVVLGADT CVAIASDRI WVLIAADSS GIVVAVSS YVVIGADT GILLASDLS VRYRQY FQYMGY FQYMGY SKRFFPY QCRSNMK.I 170 KFAIAAVTC ECTAKCLIJ ELMKLALMC ELGRRAIYI NIVRDADATYY 	ATSSGSFI ATAGPIV LGVNMLT SVSSSIC ATGGQFI LSAGQVIV IQAALIV IQAALIV ISIGSMI VISIGSMI VISIGSMI ISSCKVIV 180 SAIIRDGS AVERDSI IKQRFTV AVERDSI SAVERDSI SAVERDSI SAVERTSI

С	Tv_alfa_1 Hs_alfa_1	1 M <mark>S</mark> SG MSRGS <mark>S</mark> AG	10 AdryLTVF FdrhITIF	20 S <mark>AEGRLWQVE</mark> SPEGRLYQVE	30 YSFKAVKO YAFKAINO	4 0 AEVTAVAVKSK GGLTSVAVRGK	50 NAVCVAVOKK DCAVIVTOKK	6 0 VSDKLIDPST VPDKLLDSST	7 0 VTHMYRITDNVG VTHLFKITENIG	ACL CVM
	Tv_alfa_1 Hs_alfa_1	80 VGLPSDVN TGMTADSR	90 FIVMLLRSI SQVQR <mark>AR</mark> YI	100 FANNFEYKOG EAANWKYKYG	110 FSIPVSIL YEIPVDML	120 AQMISERHOLE CKRIADISOVY	130 Solvyvrpsa Tonaemrplo	140 VS <mark>AILFGLD</mark> G CCMILIGIDE	150 PSDSFALYKIEP EQ.GPQVYKCDP	S <mark>gy</mark> A <mark>gy</mark>
	Tv_alfa_1 Hs_alfa_1	160 SN <mark>GFRAV</mark> A YC <mark>GFKA</mark> TA	170 C <mark>gvkeie</mark> ai Agvkqtes	180 M <mark>SALEKKM</mark> ED TSF <mark>LEKKV</mark> KK	19 FETPEA KFDWTFEQ	20 20 CAEFTLSTLQT VETAITCLST	0, 21 VCGVDFEAQE VLSIDFKPSE	0 22 VEVSLLTRDN IEVGVVTVEN	0, 230 S <mark>KF</mark> SKLPNDKVN PKFRILTEAEID	EIL AHL
	Tv_alfa_1 Hs_alfa_1	240 H <mark>avaekd</mark> V <mark>alaerd</mark>								
	Tv_alfa_2 Hs alfa 2	1 MGDSDF MAERGYSF	10 SLTTFSSG SLTTFSPS	20 SKLNQIESAI SKLVOIEXAI	30 KAVSLGGQQ AAVAGGAPS	40 CVGVKAKNGAV	50 TACESKPSSE	60 LVEKVTNLKV LYDERSVHKV	70 QKINDN <mark>VGIVYS</mark> EPITKHIGLVYS	GVN GMG
	Tv_alfa_2 Hs_alfa_2	80 T <mark>dfhvilk</mark> Pdyrvlvh	90 Slrkasik R <mark>arkl</mark> ago	100 YSLRLGVEMP YYLVYQEPIP	110 TREVVKHAZ TAQLVQRVZ	120 HKMQYYTQIG SVMQEYTQSG	130 GVRPFGVSLI GVRPFGVSLI	140 IIGWEELGPT ICGWNEGRPY	150 LWQVDPSGTFWA LFQSDPSGAYFA	WKA WKA
	Tv_alfa_2 Hs_alfa_2	160 TALGKRSD TAMGKNYV	170 GS <mark>RTFLER</mark> NGKTFLEK	180 RysedQsvdd Rynedleled	190 AIHTAISTI AIHTAILTI	200 KEGFDGQLTA KESFEGQMTE	210 ELIEIGVVDE DNIEVGICNE	220 TRK <mark>FRTLS</mark> TA A.G <mark>FR</mark> RLTPT	230 BIRDFLTEV. Evkdylaaia	
	Tv_alfa_3	1 MTYRYDAG	10 TTTFSSDG	20 RILOVEYAIQ	30 S <mark>INQAGT</mark> A	4 0 COVQFINGVVI	50 AAEKKNTGRI	eo eo	7 Q AKIDGHIVTAVA	GLT
	Hs_alfa_3 Tv_alfa_3	MSRRYDSR 80 Adantlvd	TTIFSPEGI 90 L <mark>MR</mark> TS <mark>AO</mark> K	RLYQVEYAME 100 YLKTYDEQMP	IIQHAGTCI 110 VEQLVRMV(LGILANDGVLI 120 DEKHS <mark>YTQY</mark> G	AAERRNIHKI 130 GLRPYGVSFI	LDEVFESEKI 140 I <mark>agydrh</mark> kgc	YKLNEDMACSVA 150 QLYLTDPSGNEG	GIT
	Hs_alfa_3 1 Tv_alfa_3	SDANVLTN 60 ATAIGENN	ELRLIAOR 170 OTAOSILK	180 180 SQXKDN.MTA	190 TEAMDLTVE	200 200 VLCKTLDSTS	210 LSADKLEFAV	YIGWDKHYGF 220 Lofreeygpk	QIYQSDPSGNYG 230 VRIITSEVDTL	GWK M <mark>K</mark> R
	Tv_alfa_3	240 YEETIKKS	250 AEEKE		K SELAMAT			TR. DAGATV	INVINUNDVEQU	
		1	10	20	30	40	50	6 Q	70	8 Q
	Tv_alfa_4 Hs_alfa_4	MSDYTRSI .MS <mark>Y</mark> D <mark>R</mark> AI	TRFSPDGR TVFSPDGH 90	LFOIDHAHAA LFOVEYAQEA 100	VQRGTTVV VKKGSTAV 110	120 120	VEKTAVAKLO VEKKSVAKLO 130	DPHTFSKICS DERTVRKICA 140	LDKHVMCAFAGL LDDNVCMAFAGL 150	HAD TAD
	Tv_alfa_4 Hs_alfa_4	ARRLIQSG ARIVINRA	QRQ <mark>CQSHR</mark> RVE <mark>CQSHR</mark>	LTYEDPISIE LTVEDPVTVE	NIARYIATI YITRYIASI	LQLKNTQSGGA KQRYTQSNGR	RPYGVSTLIC RPFGISALIV	GFDDMTSOPH GFD.FDGTPR	IYETLPSGTYAE LYQTDPSGTYHA	WKA WKA
	Tv_alfa_4 Hs_alfa_4	RT <mark>IGR</mark> HDQ NA <mark>IGR</mark> GAK	TVMEYLEK Svreflek	HYKDDMTD Nytdeaietd	EEAQKLAIC DLTIKLVI	ALLEVVENGS ALLEVVQSGG	ZIU KNLEVAYMKE KNIELAVMRE	GGTMEIMAEE DQSLKILNPE	VLDALIESTKAK EIEKYVAEIEKE	KEE
	Tv_alfa_4 Hs_alfa_4	NEKKKQKK	AS							
	Tv_alfa_5 Hs_alfa_5	1 MFNSGSEY MFLTRSEY	1 0 DR <mark>NVNTFSI DR</mark> GVNTFSI	20 PDGRLLQVEY PEGRLFQVEY	30 AIEAVKLGS AIEA <mark>I</mark> KLGS	40 SAVAILCPEG TAIGIQTSEG	50 VIF <mark>AVEKR</mark> LS VCLAVEKRIT	60 Soliiassve Splmepssie	70 Kvyaiddhvgvv Kiveidahigca	80 MAG MSG
	Tv_alfa_5 Hs_alfa_5	LAADGRTM LIADAKTL	90 VEHMRVEA IDKARVET	100 DNHRFSFDEP DNHWFTYNET	110 IGIKAVTOS MTVESVTO	120 S <mark>VCDLALAFGE</mark> AVSNLALQFGE	130 GRRKKGDGO EDADPGA	140 SRPFGTALLV SRPFGVALLF	150 A <mark>GIE</mark> NGKCH <mark>LFH</mark> G <mark>GVD</mark> EKGPQLFH	160 TDP MDP
	Tv_alfa_5 Hs_alfa_5	1 SGTYTECR SGTFVQCD	70 ARAIGGGSI ARAIGSASI	180 EGAEALLRDL EGAQSSLQEV	190 YKDGMTLH YHKSMTLK	200 Aedlalstlr Aikssliilk	210 QVIQEKLNEN QVMEEKLNAI	220 NVEVACARVS NIELATVOP.	230 TGKFEIYTSBQR GQNFHMFTKBEL	240 QEI EEV
	Tv_alfa_5 Hs_alfa_5	2 VARLPPPI IKDI	50 IPE							



Tv_beta_4 Hs_beta_4	1 10 MLS <mark>IVGLQGPT</mark> MEYLIGIQGPT	20 WVLIAADSSVSS VVLVASDRVAAS	30 STICMSENYDRI NTVQMKDDHDKM	4 0 AQLDDRHALA IFKMSEKILLI	50 MSCETCDCLO CVCEACDTVO	бо LS <mark>EYLQGNVA</mark> FAEYIQKNVQ	7 º Lykfrngvel Lyk <mark>mrngyel</mark>	8 0 SSD <mark>AL</mark> SPT <u>AA</u>
Tv_beta_4 Hs_beta_4	90 AHFIRHTMAK ANFT <mark>RRNLA</mark> DO	100 VR.KSPYEVNMI LRSRTPYHVNLI	110 LSCYDGKPHI LACYDEHEGPAI	120 YFMDYLGTLÇ YYMDYLAALA	130 2SIPYGAQGYC AKAPFAAHGYG	140 QYFVMSVFDK AFLTLSILDR	150 HYKEGLTLED YYTPTISRER.	GK <mark>EL</mark> M AV <mark>EL</mark> L
Tv_beta_4 Hs_beta_4	160 1 KLALNQIKQRE RKCLEELQKRE	TVAPHGFIVKLV ILNLPTFSVRII	190 DKNGITKIDLE. DKNGIHDLDNIS	FPKQGS				
Tv_beta_5 Hs_beta_5	1 10 TTTLSFIYNG TTTLAFKFRHC	20 IVVAVDSRATGO VIVAADSRATAG	30 QFIF <mark>SQTVMKII</mark> A <mark>YIASQTV</mark> KKVI	4 0 Plapnmigtm Einpyllgtm	50 Aggaadcoyw Aggaadcsfw	60 L <mark>RNISR</mark> LIQL ERLLARQCRI	70 HKF <mark>R</mark> YQQPLT YEL <mark>R</mark> NKER <mark>IS</mark>	8 0 VAAAS VAAAS
Tv_beta_5 Hs_beta_5	9 Q KILVNELYRYR KLLANMVYQYR	100 GYNLSIGSMICO GMGLSMGTMICO	110 YDNTGPHIFYII WDKRGPGLYYVI	120 NH <mark>GSRIAG</mark> KE SEGNRISGAT	130 RESVGSGSTHA FSVGSGSVYA	140 YGVLDTCYRE YGVMDRGYSY	150 DMTKEEACEL DLEVEQAYDL	160 G <mark>RRAI</mark> A <mark>RRAI</mark>
Tv_beta_5 Hs_beta_5	170 YH <mark>ATYRD</mark> SGSG YQATYRDAYSG	180 GR <mark>VSVVHI</mark> TQNG GAV <u>NL</u> YHVREDG	190 VEWIDKTDVFD WIRVSSDNVADI	200 HDFSKTTF. HEKYSG <mark>S</mark> TP				
Tv_beta_6 Hs_beta_6	MLSSTAMYSAF	1 ME.GEFR GRDLGMBPHRAA	10 ENKKGQ <mark>WSPY</mark> EM GPLQLR <mark>FSPY</mark> VF	20 H <mark>ggtaigi</mark> cg N <mark>ggtilaiag</mark>	30 DDYVVIGADT EDFAIVASDT	40 RLSVD <mark>YSIDS</mark> RLSEGFSIHT	50 RHKARIFKMN RDSPKCYKLT	60 SNC <mark>MI</mark> DKT <mark>VI</mark>
Tv_beta_6 Hs_beta_6	70 SATGFDGDIDA GC <mark>SGF</mark> HGDCLI	80 FITRMRSILLNY LTKILEARLKMY	90, 1 ENQHFHEMSVES KHSNNKAMTTGA	00 1 VARCVSNTLY IAAMLSTILY	10 1 SKRFFPYYIN SRRFFPYYVY	20 1 I <mark>LVGGI</mark> NS <mark>EGI</mark> N <mark>IIGGL</mark> DEEGI	30 1 KGKLYGYDPV KGAVYSFDPV	4 0 STIED SSYQR
Tv_beta_6 Hs_beta_6	150 LHYDSNGSGSS DSFKAGGSASA	160 LAAPLLDSAFGT MLQPLLDNQVGF	170 1 IHHNTRPFPAVS KNMQNVEHVPLS	80, 1 LQDAKNIVRD LDRAMRLVKD	90 2 AICSVTERDI VFISAAERDV	00022 YTGDALQLCVI YTGDALRICI	10 2: FTKDGFAQEEI VTKEGIREET	2 0 F P <mark>L</mark> P R V S <mark>L</mark> R <mark>K</mark>
Tv_beta_6 Hs_beta_6	H D							
Tv_beta_7 Hs_beta_7	MEAFLGSRSGL	WAGGPAPGQFYR	IPSTPDSFMDPA	SALYRGPITR	1 1 . MQVITASGA TQNPMVTGTS	0 20 IVAAKYDGGII VLGVKFEGGVV) 30 LLASDLSITYO VIAADMLGSYO	SMFR SLAR
Tv_beta_7 Hs_beta_7	40 HN <mark>NVSHFVEVA</mark> FR <mark>NISR</mark> IMRVN	50 PNIIIGASGEFA NSTMLGASGDYA	60 7 DFQTLIEVIKSV DFQYLKQVLGQM	0.8 ILQQQCKHNG VIDEELLGDG	O 90 EYLTASEVHN HSYSPRAIHSI	0 100 YIKRYMYQCRS WLTRAMYSRRS	110 NMKPLSCKV SKMNPLWNTMV	IVAGI /IGG.
Tv_beta_7 Hs_beta_7	120 NPDGSKFLACT YADGESFLGYV	130 DPYGASWESDHI DMLGVAYEAPSL	140 15 G <mark>TGFG</mark> K <u>YI</u> QGLQ ATGYGAYLAQPL	º I A D V V N G S F D L R E V L E K Q P V	160 DVK KGI LSQT <mark>EAR</mark> DL <mark>V</mark> I	170 TEVF <mark>RAV</mark> NARI ERCM <mark>RVL</mark> YY <mark>R</mark> I	180 NTTANGKIEF DARSYN <mark>R</mark> FQI	190 TVTP TVTE
Tv_beta_7 Hs_beta_7	200 Q <mark>ginhlapeq</mark> i K <mark>gv</mark> eieg <mark>pls</mark> t	210 DP <mark>nWevvegtwd</mark> Et <mark>nwdia</mark> hmisg	Q. FE					

e 1 10 20 30 40 50 60 Tv_Ump-1 Hs_Ump-1 MNARGLGSELKDSIPVTELSASGFESHDLERKEFSCVKNELLPSHPISLSEKNFOLNODKMNSTERNIOGLFAPIKLQ 70 80 90 100 110 120 Tv_Ump-1 MLYKIKSTRIGFRTYERPDDFATEVFTCDIDDVDRNDMFAPNGMRADIEFDFETCEKRINTE Hs_Ump-1 MEFKA..VQQVQRLFFLSSSNLSLDVLRGNDETIGEDILNDPSQ.SEVMGEPHLMVSYKLGLL



g

Subunit	<i>Tv</i> Uniprot ID	Hs Uniprot ID
α1	A2F568	P60900
α2	A2FJV7	P25787
a3	A2FT79	P25789
α4	A2DTN3	O14818
α5	A2FCM7	P28066
a6	A2E1I9	P25786
α7	A2D8G5	P25788
β1	A2E7Z2	P28072
β2	A2F2T6	Q99436
β3	A2F3H9	P49720
β4	A2F8W4	P49721
β5	A2DD57	P28074
β6	A2F716	P20618
β7	A2F3X4	P28070
Ump-1	A2FJW0	Q9Y244

Supplementary Figure 1. Sequence alignment of *Tv*20S and human 20S proteasomes and Ump-1 chaperone.

Panel **a**) represents multiple sequence alignments of all seven α subunits (α 1- α 7) and **b**) all seven β subunits (β 1- β 7) from *Tv*20S, with β 1, β 2 and β 5 shown as proenzymes, and **c**) & **d**) aligned individually with their human $\alpha \& \beta$ subunits (*Hs*20S) counterparts. Conserved residues are boxed, and identical residues are highlighted in red. Conservation patterns and variations between species are illustrated. **e**) Sequence alignment of *Tv* Ump-1 and with *Hs* Ump-1, highlighting conserved regions. **f**) Sequence alignment of mature β 2 subunit from *T. vaginalis* (*Tv*), *Giardia lamblia* (*Gl*), *Plasmodium falciparum* (*Pf*), *Trypanosoma brucei* (*Tb*), *Homo sapiens* (*Hs*), and *Ixodes ricinus* (*Ir*). The asterisk denotes highly conserved residues Lys33 and a loop consisting of Ala46, Ala49, and Asn52. The aligned sequences were visualized using ESP3 software according to Robert, X. and Gouet, P. (2014).¹ g) Uniprot accession codes of each gene used.



d Insert BamHI_ α 1_ α 2_ α 3_ α 4_ α 5_ α 6_ α 7_ NotI - sequences encoding Kozak sequence, start and stop codon, <u>gene</u>, polyhedrin promoter, SV40 polyA tail and twin-strep tag. are highlighted.

ggatccGCCACCATGAGCAGCGGTGCAGATCGTTATCTGACCGTTTTTAGTGCCGAAGGTCGTCTGTGGCAGGTTGAATATAGTTTTAA AGCAGTTAAACAGGCCGAAGTTACCGCAGTTGCAGTTAAAAGCAAAAATGCAGTTTGTGTTGCCGTGCAGAAAAAGGTTAGCGATAAAC ATTGTTATGCTGCTGCGTAGCTTTGCCAACAACTTTGAATATAAACAGGGCTTTAGCATCCCGGTTTCAATTCTGGCACAGATGCTGAG CGAACGTCATCAGCTGGAAAGCCAGCTGGTTTATGTTCGTCCGAGCGCAGTTAGCGCAATTCTGTTTGGTCTGGATGGTCCGAGCGATA GCTTTGCACTGTATAAAATCGAACCGAGCGGTTATAGCAATGGTTTTCGTGCAGTTGCATGTGGCGTTAAAGAAATTGAAGCAATGAGC JCACTGGAAAAGAAGATGGAAGATTTTGAAACACCGGAAGCAACCGCAGAATTTACCCTGAGCACCCTGCAAACCGTTTGTGGTGTTGA TTTTGAAGCACAGGATGTTGAAGTTAGCCTGCTGACCCGTGATAATAGCAAATTTTCAAAACTGCCGAACGACAAGGTGAACGAAATTC TGCATGCCGTTGCCGAAAAAGAT**TAA**AGATCTAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACC TCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTGTTGTTGTTGTTGTTGCAGCTTATAATGGTTACAAATAA AGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCA CAATAACTATCATAACCCCTAGGGTATACCCATCTAAGGTAGCGAGTTTAAACACTAGTATCGATTCGCGACCTACTCCGGAATATTAA ATAAATATCCGGATTATTCATACCGTCCCACCATCGGGCGCGAGCTCGCCACCATGGGCGATAGCGATTTTAGCCTGACCACCT CAGCGGTGGTAAACTGAATCAGATTGAAAGCGCACTGAAAGCAGTTAGCTTAGGTGGCCAGTGTGTTGGTGTTAAAGCCAAAAATGGTG CAGTTATTGCCTGTGAAAGCAAACCGAGCAGTCCGCTGGTTGAAAAAGTTACCAATCTGAAAGTGCAGAAAATCAACGATAATGTGGGC ATTGTTTATAGCGGTGTGAACACCGATTTTCACGTTATTCTGAAAAGCCTGCGTAAAGCCAGCATCAAATATAGCCTGCGTCTGGGTGT TGAAATGCCGACACGTGAAGTTGTTAAACATGCAGCACATAAGATGCAGTATTATACCCAGATTGGTGGTGTTCGTCCGTTTGGTGTTA GCCTGCTGATTATTGGTTGGGAAGAACTGGGTCCGACACTGTGGCAGGTTGATCCGAGCGGCACCTTTTGGGCATGGAAAGCAACCGCA CTGGGTAAACGTAGTGATGGTAGCCGTACCTTTCTGGAACGTCGTTATAGCGAAGATCAGAGCGTTGATGATGCAATTCATACCGCAAT TAGCACCCTGAAAGAAGGTTTTGACGGCCAGCTGACCGCAGAACTGATTGAAATTGGTGTTGTTGATGAAACCCGTAAATTTCGTACCC TGAGCACCGCAGAAATCCGCGATTTTCTGACCGAAGTT**TAA**GAATTCAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTAC AATGGTTACAAATAAAGCAATAGCATCACAAAATTTCACAAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCAT **CAATGTATCTTATCATGTCTGGATC**TGATCACTGCTTGAGCCTAGAAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCT GCTGCCACCGCTGAGCAATAACTATCATAACCCCTAGGGTATACCCATCTAAGGTAGCGAGTTTAAACACTAGTATCGATTCGCGACCT **GTAATAAAAAAACCTATAAAT**ATTCCGGATTATTCATACCGTCCCACCATCGGGCGCCCGCGG<mark>GCCACC</mark>ATGACCTATCGTTATGATGC TTCAGTTTACCAATGGTGTTGTTCTGGCAGCCGAAAAGAAAAATACCGGTCGTCTGGTTGATTACCTGTTTCCTGAGAAAATGGCCAAA ATTGATGGTCATATTGTTACCGCAGTTGCAGGTCTGACCGCAGATGCAAATACCCTGGTTGATCTGATGCGTACCAGCGCACAGAAATA TCTGAAAACCTATGATGAGCAGATGCCGGTTGAACAGCTGGTTCGTATGGTTTGTGATGAAAAACATAGCTATACCCAGTATGGTGGTC TGCGTCCGTATGGTGTTAGCTTTCTGATTGCAGGTTATGATCGTCATAAAGGTTGTCAGCTGTATCTGACCGATCCGAGCGGTAATTTT <u>GGTGGTTGGAAAGCAACCGCCATTGGTGAAAATAATCAGACCGCACAGAGCATTCTGAAAAGCCAGTATAAAGATAATATGACCGCAAC</u> CGAAGCAATGGATCTGACCGTTAAAGTTCTGTGTAAAACCCTGGATAGCACCAGCCTGAGCGCAGATAAACTGGAATTTGCAGTTCTGC AGTTTCGCGAAGAATATGGTCCGAAAGTGCGTATTCTGACCACCAGTGAAGTTGATACCCTGATGAAACGTTATGAGGAAACCATTAAA AAGTCAGCCGAGGAAAAAGAA**TAA**GGTACCAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAAACCTC CCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAG CAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATG ATAACTATCATAACCCCTAGGGTATACCCATCTAAGGTAGCGAGTTTAAACACTAGTATCGATTCGCGACCTACTCCGGAATATTAATA AAATATTCCGGATTATTCATACCGTCCCACCATCGGGCGCCTCGAG<mark>GCCACC</mark>ATGAGCGATTATACCCGTAGCATTACCCGTTTTAGTC CGGATGGTCGTCTGTTTCAGATTGATCATGCACATGCAGCAGTTCAGCGTGGCACCACCGTTGTTGCAACCCGTAGTAAAGATATGATT GTTATTGCCGTTGAGAAAACCGCAGTTGCAAAACTGCAAGATCCGCATACCTTTAGCAAAATTTGTAGCCTGGATAAACATGTGATGTG TGCATTTGCAGGTCTGCATGCCGATGCACGTCGCCTGATTCAGAGCGGTCAGCGTCAGTGTCAGAGCCATCGTCTGACCTATGAAGATC CGATTAGCATTGAAAACATTGCCCGTTATATTGCAACCCTGCAACTGAAAAATACCCAGAGCGGTGGTGCACGTCCGTATGGTGTTAGC ACCCTGATTTGTGGTTTTGATGATATGACCAGCCAGCCGCATATTTATGAAACCCTGCCGAGCGGCACCTATGCAGAATGGAAAGCACG TACCATTGGTCGTCATGATCAGACCGTTATGGAATATCTGGAAAAACACTACAAAGACGATATGACCGATGAAGAAGCACAGAAACTGG CAATTGGTGCACTGCTGGAAGTTGTTGAAAATGGTAGCAAAAATCTGGAAGTGGCCTATATGAAACGTGGTGGTACAATGGAAATTATG GCCGAAGAGGTTCTGGATGCACTGATTGAAAGCACCAAAGCAAAATAAACGCGTAGAGGATCATAATCAGCCATACCACATTTGTAGAG AGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCA AACTCATCAATGTATCTTATCATGTCTGGATCTGATCACTGCTTGAGCCTAGAAGATCCGGCTGCTAACAAAGCCCCGAAAGGAAGCTGA GTTGGCTGCCACCGCTGAGCAATAACTATCATAACCCCTAGGGTATACCCATCTAAGGTAGCGAGTTTAAACACTAGTATCGATTC **CAGTTTTGTAATAAAAAAACCTATAAAT**ATTCCGGATTATTCATACCGTCCCACCATCGGGCGCAAGCTT<mark>GCCACC</mark>ATGTTTAATAGCG GCAGCGAATATGATCGCAATGTGAATACCTTTAGTCCGGATGGTCGTCTGCTGCAAGTTGAATATGCAATTGAAGCAGTTAAACTGGGT TGAAAAAGTTTATGCCATCGATGATCATGTTGGTGTTGTTATGGCAGGTCTGGCAGCAGATGGTCGTACAATGGTTGAACACATGCGTG TTGAAGCACAGAATCATCGTTTTAGCTTTGATGAACCGATTGGCATTAAAGCAGTTACCCAGAGCGTTTGTGATCTGGCACTGGCATTT TCACCTGTTTCATACCGATCCGAGCGGCACCTATACCGAATGTCGTGCACGTGCCATTGGTGGTGGTGGTAGCGAAGGTGCCGAAGCACTGC TGCGTGATCTGTATAAAGATGGTATGACCCTGCATGAAGCAGAGGATCTGGCCCTGAGCACCCTGCGTCAGGTTATTCAAGAAAAACTG AATGAGAACAATGTGGAAGTTGCATGTGCACGTGTTAGCACCGGTAAATTTGAAATCTATACCAGCGAACAGCGCCAAGAAATTGTTGC CCGTCTGCCTCCGCCTATTATTCCGGAA**TAA**GTCGACAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAA

AATAAAGCAATAGCATCACAAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCCAAACTCATCAATGTATCT TATCATGTCTGGATCTGATCACTGCTTGAGCCTAGAAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCCACCG ${\tt CTGAGCAATAACTATCATAACCCCTAGGGTATACCCATCTAAGGTAGCGAGTTTAAACACTAGTATCGATTCGCGACCTACTCCGGAAT$ **AACCTATAAAT**ATTCCGGATTATTCATACCGTCCCACCATCGGGCGCCCCGGGGCCACCATGTTCCGCAGCAAATATGATGAAAATGCC ACCACCTTTAGTCCGGAAGGTCGTATTCTGCAAGTTGAAAATGCAATGAAAGCAGTTCAGCAGGGTATGCCGACCGTTGGTCTGAAAAG CAAAACCCATGCAGTTATTGCCGGTGTTATGCATAGCCCGAGCGAATTTAGCAGCCATCAGCCGAAAATCTTTAAAAATCGATCAGCATA GGCACCGAAATTCGTGTTGCCGATCTGGCAGATACCGTTGCACTGCAAAGCCAGAAAAAGACCAGCAAAGTTGGTAAACGTCCGTATGG TGTTGGTCTGCTGATGATTGGTGCGGGTGTTGATGGTCCGCGTCTGTTTGAAACCTGTCCGAGCGGTCAGCATTGGGAATATAATGCAC CGTCATGCACTGCGTGCACTGAATGATTGTAAAAGCCGTGAAAGCGATAGCCTGGAAGCAATTGCACTGGGTGTTGTTGGTATTGATGA <u>CGTTTACCATTCTGGAAGGTCCGGAATTACAGAAATATATCGAT**TAA**AGGCCTAGAGGATCATAATCAGCCATACCACATTTGTAGA</u> $\mathsf{CAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTCACTGCATTCTAGTTGTGGTTTGTCC$ AAACTCATCAATGTATCTTATCATGTCTGGATCTGATCACTGCTTGAGCCTAGAAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTG AGTTGGCTGCCGCCGCCGAGCAATAACTATCATAACCCCTAGGGTATACCCATCTAAGGTAGCGAGTTTAAACACTAGTATCGATT **ACAGTTTTGTAATAAAAAAACCTATAAAT**ATTCCGGGATTATTCATACCGTCCCACCATCGGGCGCCATATG GGTAGCGGTTATGATTTTAATCCGATTACCTTTAGTCCGGATGGTCGTCAGTTTCAGGTTGAATATGCAACCAAAGCCGTTGAAAAAGA TAGCCTGGCACTGGGTGTTAAATGCAAAGATGGTATTCTGCTGGCAGCCGAGAAAAATCTGACCAGCACACTGCTGACCCCTGGTGGTA ATCCGCGTATCTTTTGGATTAATGATAGCATTGCATGTGCCACCATTGGTCATCGTCCGGATTGTTATAGCATTGTTGAACAGAGCCGT AATCGTGCAGAAACCTTTACCAGCAATTTTGGCATCAAAATTACCGTTCCGCAGCTGGCAAGCGAAGTTAGCCAGCAGTTTCATCTGGC ACATTATTATCAGGCATATCGTCCGTTTGGTTGTACCGTTATTTTTGCCAGCTATAAAGATGATGCCCTGTATGCAATTGAACCGAGCG GTGCCTTTTATGGTTATTTTGCAAGCTGCTTTGGCAAGAATAGCAATCTGGCACGTGCAGAATTACAGAAAAACCGAATGGAAAAATATC ACCGTTCGTGAAGCAGTTCCGGAAGTTGCACGTATTATCAAAAGCCTGCATGAAAGCCAGTTTAAAAAGTGGGAAATCGAAATGTTTTG GCTGTGCGAAGAAACCAATGGTCGTCCGCAGAAAGTGCCGGAAGATGTTTTTCAGAGCCGTTTTGTTAATGAAAACCCGCAGAAT**TAA**g caaccac

Insert BamHI_β1_β2_β3_β4_β5_β6_β7+tag_ NotI - sequences encoding *Kozak sequence*, <u>start</u> and <u>stop</u> codon, <u>gene</u>, **polyhedrin promoter**, SV40 polyA tail and twin-strep tag.are highlighted.

ggatcc
GCCACC
ATGAGCGAATATCAGTTTCCGAAAGAAAGCATGGGTACAACCCTGCTGGCAATTCAGTGTACCGATGGTGTTGTTAT CACGTTGTGGTAATGCAGCAGATACCCAGTTTCTGGCACGTGCAGTTAAAAACTATCTGAATGCACTGAACATCACCCGTGAAAATACC GATGATAGCACCATTCTGGTTGCCAGCAATGTTATTCGTAGCCTGATTGTTCGTTATCGTCAGTATCTGAGTGCCGGTGTTATTGTTGG GCACCTATATTCAGGCATATATTGATCAGAACTATCGCGAGGATATGACAATGGAAGAAGCAACCAAATTTGCCAATTGCCGCAGTTACC **GGTGCAATTATTCGTGATGGTAGCAGCGGTGGTGTTGTGAATATTGTTCAGATTAATGCAGATGGTGCCAAACGTATGACCGTTCGTCC** GGCACAGCAGCCGTTTAACTATGATATTGTTAAAGGT**TAA**AGATCTAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACT **ATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATC AATGTATCTTATCATGTCTGGATC**TGATCACTGCTTGAGCCTAGAAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTG CTGCCACCGCTGAGCAATAACTATCATAACCCCTAGGGTATACCCATCTAAGGTAGCGAGTTTAAACACTAGTATCGATTCGCGACCTA **TAATAAAAAAACCTATAAAT**ATTCCGGATTATTCATACCGTCCCACCATCGGGCGCGAGCTC<mark>GCCACC</mark>ATGGAAGCAGGTCTGGGTTTT GATTTTAGCAATTATGCCCGTAACAAAAGCCTGGAACCGAAACTGGGTAAACCGCAGCTGCTGAGCACCGGCACCACCATTGCAGCAGC AATTTTTTGATGGTGGTGTTGTTCTGGGTGCAGATACCCGTGCAACCGCAGGTCCGATTGTTGCAGTTAAAGATGAAATGAAACTGCACT ACATCAGCGATAACATTTGGGTTTGTGGTGCAGGTATTGCAGCCGATAATGATAACATTAATGCAGTGATTAGCGCCCAAACTGCGTCTG TTTCAGATGAATACCGGTCTGCAACCGCGTGTTGATCAGTGTACCAATATTCTGGCAAGTCGCCTGTTTCAGTATATGGGTTATATTCA GGCAGCACTGATTGTTGGTGGTATTGATTTTCAGGGTCCGCAGGTTTATCAGGTTGCACCGCATGGTAGCTTTAGCAAACAGCCGTTTA GTTGCAGATGCAATTTATGCCGGTATTACCAATGATTTAGGTTCAGGTAGCCATGTTAATCTGTGCGTGATTAAACGTGAAAACCCGGA AGATAAACAGAGCAAAGTGATCTACACCTTCTATAAAGATTATCGCGTGCCGCATGAAAACGATCGTAATTTTCGTCTGGAACCGCAGA TCAATAACATTGATGTGGAAGTGATTAAAACCACCGAACGTCCGCTGACACTGCCGGATGTTCATCTGGAAATTCTGGATGATGCAC GCATAAGAATTCAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAAACCTCCCACACCTCCCCCTGAAC CTGAAACATAAAATGAATGCAATTGTTGTTGTTGTT**AACTTGTTTATTGCAGCTTATAATGGTTACAAAATAAAGCAATAGCAATAGCAACAAATTT** CACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGGATC TTGAGCCTAGAAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCCACCGCTGAGCAATAACTATCATAACCCCT AGGGTATACCCATCTAAGGTAGCGAGTTTAAACACTAGTATCGATTCGCGACCTACTCCGGAATATTAATAGATCATGGAGATAATTAA **AATGATAACCATCTCGCAAATAAATAAGTATTTTACTGTTTTCGTAACAGTTTTGTAATAAAAAAACCTATAAAT**ATTCCGGATTATTC ATACCGTCCCACCATCGGGCGCCCGCGGGCCACCATGAGCGATATCAGCACCTATAATGGTAGCTGTGTTCTGGCAATGGCAGGCGATC ATTGTGTTGCAATTGCCAGCGATCGTCGTCGGGTGTTAATATGCTGACCGTTAGCAAAGATTTCAAACGCATCTTTCAGATCAACGAC CGCATTTATCTGGGTTTAGCAGGTCTGGCAACCGATGTTCTGACCGTTCGTGAACAGCTGCGTTTTGATGTTAATCTGCTGGAACTGCG **TGAAGAACGTCCGATTGATCCGAAAAAGTTTATGAATCTGGTTAAAAGCACCTGTACGAAAAACGCTTTAGCCCGTTCTTTGTTACAC** GGTTATTGCAGGTCTGCTGCCGGAAACCAATGAACCGTATCTGGCAGCAAGCGATAGCATTGGTGCATTTGCATTTCCGAAAGATTTT GCAGTTGCAGGCACCTGTGAAGAAAGCCTGTATGGTATTTGTGAAAGCGCATGGCGTCCGAATATGAATCCGGATGAACTGTTTGAATG TACCGCCAAATGTCTGATTGCAGCAGTTGAACGTGATAGCATTAGCGGTTGGGGTGGTATTGTTTATATCATCACCCAGGATAAAGTGA TCATCAAAGAAATCAAAACCCGCATGGATTAAGGTACCAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAA AAATAAAGCAATAGCATCACAAAATTTCACAAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATC

TTATCATGTCTGGATCTGATCACTGCTTGAGCCTAGAAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCCGCCACC GCTGAGCAATAACTATCATAACCCCTAGGGTATACCCATCTAAGGTAGCGAGTTTAAACACTAGTATCGATTCGCGACCTACTCCGGAA **AAACCTATAAAT**ATTCCGGATTATTCATACCGTCCCACCATCGGGCGCCTCGAG<mark>GCCACC</mark>ATGCTGAGCATTGTTGGTCTGCAAGGTCC GGATTGGGTTCTGATTGCAGCAGATAGCAGCGTTAGCAGCAGCATTATTTGTATGAGCGAAAACTATGATCGTATCGCACAGCTGGATG ATCGTCATGCACTGGCAATGAGCGGTGAAACCGGTGATTGTCTGCAACTGAGCGAATATTTACAGGGTAATGTTGCCCTGTACAAATTT CGTAATGGTGTTGAACTGAGCAGTGATGCCCTGGCACATTTTATTCGTCATACAATGGCAAAAGCCGTTCGTAAAAGCCCGTATGAAGT TAATATGCTGCTGAGCGGTTATGATGGTAAACCGCATCTGTATTTCATGGATTATCTGGGCACCCTGCAAAGCATTCCGTATGGTGCAC AGGGTTATTGTCAGTATTTTGTGATGAGCGTGTTCGATAAGCACTATAAAGAAGGTCTGACCCTGGAAGATGGTAAAGAACTGATGAAA CTGGCACTGAACCAGATTAAACAGCGTTTTACCGTTGCACCGCATGGCTTTATTGTTAAACTGGTGGATAAAAACGGCATCACCAAAAT CGATCTGGAATAAACGCGTAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAAACCTCCCACACCTCCC CCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTT**GACTTGTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCA CAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGGATCTGA** TCACTGCTTGAGCCTAGAAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCCGCCGCCGAGCAATAACTATCAT AACCCCTAGGGTATACCCATCTAAGGTAGCGAGTTTAAACACTAGTATCGATTCGCGACCTACTCCGGAATATTAATAGATCATGGAGA TAATTAAAATGATAACCATCTCGCAAATAAATAAGTATTTTACTGTTTTCGTAACAGTTTTGTAATAAAAAAACCTATAAATATTCCGG ATTATTCATACCGTCCCACCATCGGGCGCAAGCTTGCCACCATGCAGAGCCTGTATCTGAAACCGCGTGATGAAGTTGAAAGCGAAGAA <u>ACCACCAAAGCACTGGAACCGGCACATGTTCCAGATCCGTGTCAGTTTGTTAAAAACCATATTAGCCTGAGCTATACCAACGAACCGGG</u> TAAAATGGCAGCAGTTCATGGCACCACCACCACGAGGTTTTATCTATAATGGTGGTATTGTTGTTGCCGTTGATAGCCGTGCAACCGGTG GTCAGTTTATCTTTAGCCAGACCGTTATGAAAATTCTGCCGCTGGCACCGAATATGATTGGTACAATGGCAGGCGGTGCAGCAGATTGT **TCTGGTTAATGAACTGTATCGCTACAAGGGCTATAATCTGAGTATTGGTAGCATGATTTGCGGCTATGATAATACCGGTCCGCACATCT** TTTATATCGATAATCATGGTAGCCGCATTGCCGGTAAACGTTTTAGCGTTGGTAGCGGTAGCACCCATGCCTATGGTGTTCTGGA TACC TGTTATCGTGAAGATATGACCAAAGAAGAAGCCTGCGAACTGGGTCGTCGTGCCATTTATCATGCAACCTATCGTGATAGCGGTAGTGG TGGTCGTGTTAGCGTTGTTCATATTACCCAGAATGGTGTGGAATGGATCGATAAAACCGATGTGTTTGATATGCACGACTTTAGCAAAA CCACCTTTTAAGTCGACAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAAACCTCCCACACCTCCCCC TGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTGTT**AACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACA AATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGGATC**TGATC ACTGCTTGAGCCTAGAAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCCGCCGCCGAGCAATAACTATCATAA CCCCTAGGGTATACCCATCTAAGGTAGCGAGTTTAAACACTAGTATCGATTCGCGACCTACTCCGGAATATTAATAG**ATCATGGAGATA ATTAAAATGATAACCATCTCGCAAATAAATAAGTATTTTACTGTTTTCGTAACAGTTTTGTAATAAAAAAACCTATAAAT**ATTCCGGAT TATTCATACCGTCCCACCATCGGGCGCCCCGGG<mark>CCCACC</mark>ATGGAAGGTGAATTTCGCGAGAATAAGAAAGGTCAGTGGTCACCGTATGA AATGCATGGTGGCACCGCAATTGGTATTTGTGGTGATGATTATGTTGTGATGGTGCAGATACCCGTCTGAGCGTTGATTATAGCATTG ATAGCCGTCATAAAGCCCGTATCTTTAAGATGAATAGCAACTGTATGATTAGCGCCACCGGTTTTGATGGTGATATTGATGCATTTATT ACCCGCATGCGTAGCATTCTGCTGAATTATGAAAACCAGCACTTTCACGAAATGAGCGTGGAAAGCGTTGCACGTTGTGTATGCAATAC ATCCGGTTGGCACCATTGAGGATCTGCATTATGATAGCAATGGTAGCGGTAGCAGCCTGGCAGCACCGCTGGATAGCGCATTTGGT ACAATTCATCATAATACCCGTCCGTTTCCCGCCAGTTTCACTGCAAGATGCCAAAAACATTGTTCGTGATGCAATTTGTAGCGTTACCGA ${\tt ACGTGATATCTATACCGGTGATGCACTGCAACTGTGTGTTTTTACCAAAGATGGTTTTGCCCAAGAAGAATTTCCGCTGCCTCGTCAT{\tt T}$ AAAGGCCTAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAAACCTCCCACACCTCCCCCTGAACCTGA AACATAAAATGAATGCAATTGTTGTTGTTGTTAACTGGTTACAAAATGGTTACAAAATAAAGCAATAGCAACAAATTTCACA AATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGGTTTGTCCAAACTCATCATGTATCTTATCATGTCTGGATC GCCTAGAAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCCACCGCTGAGCAATAACTATCATAACCCCTAGGG TATACCCATCTAAGGTAGCGAGTTTAAACACTAGTATCGATTCGCGACCTACTCCGGAATATTAATAGATCATGGAGATAATTAAAATG ATAACCATCTCGCAAATAAATAAGTATTTTACTGTTTTCGTAACAGTTTTGTAATAAAAAAACCTATAAAATATTCCGGATTATTCATAC GCTGGCAAGCGATCTGAGCATTACCTATGGTAGCATGTTTCGCCATAATAACGTTAGCCATTTTGTTGAAGTTGCACCGAACATTATTA TCGGTGCCAGCGGTGAATTTGCAGATTTTCAGACCCTGATTGAAGTGATCAAAAGCGTTATTCTGCAGCAGCAGTGTAAACATAATGGT GAATATCTGACCGCCAGCGAAGTGCATAATTACATTAAACGCTATATGTATCAGTGCCGCAGCAATATGAAACCGCTGAGCTGTAAAGT TATTGTGGCAGGTATTAATCCGGATGGCAGCAAATTTCTGGCATGTACCGATCCGTATGGTGCAAGCTGGGAAAGCGATCACATTGGCA CCGGTTTTGGTAAATATCTGCAGGGTCTGCAGATTGCCGATGTTGTTAATGGTAGCTTTGATGATGTGAAAAAGGGCATCACCGAAGTT TTTCGTGCAGTTAATGCCCGTAATACCACCGCAAATGGTAAAATCGAATTTATCACCGTTACACCGCAGGGCATTAATCATCTGGCACC GGAACAAATTGATCCGAATTGGGAAGTTGTTGAAGGCACCTGGGATCAG<mark>AGCGCTTGGAGCCACCCGCAGTTCGAAAAAGGTGGAGGTT</mark> CTGGCGGTGGATCGGGAGGTTCAGCGTGGAGCCACCCGCAGTTCGAGAAA**TAA**gcggccgc

Insert BamHI_Ump-1_NotI

ggatccGCCACCATGTACGAACAGTGGATTCCGGAAGAACCGGTTGATCGTCTGCGTGAAGGTCTGCCGAATCTGCGTCATGGTAGGGT TAATCAGCATCCGCTGGAAATTGCCATTGAAGAACGTCGTAAAAACCCAGTTCAAAGACAAATTTGATGAACTGGCACTGCTGTATGGTG AAGGTTTTGCAAATCATGAGAAGATGCTGTACAAACTGATTAAAAGCACCCGTATTGGTTTTCGCACCTATGAACGTCCGGATGATCTG GCAATTGAAGTTTTTACCGGTGATATCGATGATGTGGGATTTCAACGATATGTTTGCACCGAATGGTATGCGTGCCGATATTGAATTTGA TCCGCATGAAATCCAAGAGAAACGCCTGAATATTGAATATGAA**TAA**gcggccgc

Supplementary Figure 2. Cloning plasmids and proteasome gene sequences. a-c) Schematic representation of the cloning of the three plasmids used to generate recombinant baculoviruses. The figure was created using SnapGene® (snapgene.com). d) The inserted DNA sequences correspond to the seven $Tv20S \alpha$ subunits (BamHI_ $\alpha1_\alpha2_\alpha3_\alpha4_\alpha5_\alpha6_\alpha7_NotI$), seven Tv20S β subunits (BamHI_ $\beta1_\beta2_\beta3_\beta4_\beta5_\beta6_\beta7+$ tag_NotI) and Tv20S Ump-1 (BamHI_Ump-1_NotI). The accession numbers for each subunit are listed in Supplementary Table 3. Sequences are labelled as follow: *restriction enzyme cleavage site*, Kozak sequence, start and stop codon, gene, polyhedrin promoter, SV40 polyA tail, twin-strep tag.



Supplementary Figure 3. Confirmation of recombinant Ump-1 function and 20S stability. Panel a) shows a native protein gel demonstrating Me4BodipyFL-Ahx3Leu3VS probe labelling of the proteasome. The left panel is a silver-stained gel, while the right panel is a fluorescent scan of the same gel at 470 nm excitation and 530 nm emission. Panel b) displays a denaturing gel with the same protein samples as panel a lane 1 represents Sf9 cell lysate, lane 2 represents Sf9 cell lysate infected with alpha, beta, and Ump-1 baculovirus, lane 3 represents Sf9 cell lysate infected with alpha and beta baculovirus, lane 4 represents native Tv20S, and lane 5 represents purified rTv20S. Panel c) shows the time-dependent assembly of the 20S proteasome in the presence and absence of Ump-1, while panel d) shows the enzymatic activity of rTv20S in the presence and absence of unp-1. Assays were performed in technical replicates (n=3). e) Silver-stained gel showing the mass of rTv20S is unaltered in the presence of inhibitors while the fluorescent gel shows that reduced labelling of the active sites. All gels were repeated independently two times with similar results. 'ST' stands for the molecular weight ladder.



Supplementary Figure 4. Activation of $\beta 1$ in the presence of CP-17. rTv20S was preincubated with 5 μ M of CP-17 to inactivate $\beta 5$ only and then assayed with Ac-RYFD-amc. Activity was compared to a vehicle control. Assays were performed in triplicate, with each dot representing the mean of 3 technical replicates. The green arrow shows an example of the apparent increase in substrate concentration (from 126 μ M to 200 μ M) within the proteasome core in the presence of CP-17.



Supplementary Figure 5. Structural comparison of cryo-EM maps of the inhibitor molecules MZB and CP-17 covalently bound in the active sites of Tv20S

The top views of both *Tv*20S structures with **a**) MZB and **b**) CP-17 inhibitors (both shown as blue sticks). Cryo-EM Maps are carved around inhibitors for clarity as green mesh at sigma 4. Detailed views on the active sites are shown as cartoon and sticks for inhibitors. MZB (panels **c**), **d**) and **e**) and CP-17 (panels **f**, **g** and) are displayed with active sites in order: β 1 (panels **c**&f shown in yellow), β 2 (panels **d**) & **g**), shown in cyan), and β (panel **e**) & **h**), shown in magenta).



Supplementary Figure 6. Cryo-EM maps in all three active sites for both Tv-20S structures

Comparison of cryo-EM maps for both Tv20S structures with MZB (**a**,**b**,**c**,**g**,**h**) and CP-17 (**d**,**e**,**f**,**i**,**j**) inhibitors. Maps are shown as grey mesh at sigma 4 and density around inhibitors is depicted in green. Active sites are displayed in order: $\beta 1$ (panels **a**) & **d**) are shown in yellow), $\beta 2$ (panels **b**) & **e**), shown in cyan), and β (panel **c**) & **f**), shown in magenta. Panels **g**) - **j**) additionally show the density around Tv20S residues that form the binding site, and ligand, shown simultaneously as sticks and also cryo-EM maps at the same sigma level. The map around Tv20S residues is depicted in blue.



Supplementary Figure 7. C-terminal end being properly inserted between the $\beta 1$ and $\beta 2$. a) Tv20S with MZB cartoon representation where $\beta 1$ is shown in yellow, $\beta 2$ cyan and $\beta 7$ ' grey with its the C-terminus marked by arrow. Panels b) shows local cryoEM map (green mesh) and c) show detailed view with an overlay with superimposed human 20S proteasome (PDB ID = 7PG9) shown in orange. The residues of $\beta 7$ ' Asn204-Val208 of Tv20S shown as grey sticks inserted between $\beta 1$ and $\beta 2$ chains.



Supplementary Figure 8. Prospects for developing inhibitors that target Cys46 of the β 1 subunit of *Tv*20S proteasome.

a) The active site of the β 1 subunit in the *Tv*20S-MZB proteasome structure is highlighted in yellow, while MZB is shown as blue sticks. This region includes neighbouring residues, including Cys46. **b)** The structural overlay of β 1 and β 2 (cyan sticks) subunits of *Tv*20S-MZB proteasome. **c)** An overlay comparison is made between the β 1 active site of *Tv*20S-MZB (shown in yellow, with MZB represented as blue sticks) and the human 20S structure, PDB ID = 7PG9 (shown in orange). **d)** Another overlay is presented, showcasing only the CP-17 inhibitor (depicted in white) of the β 2 subunit onto the β 1 active site of *Tv*20S-MZB (represented in yellow, with MZB shown as blue lines).



Supplementary Figure 9. Model of CP-17 "virtual" clashes within the *Tv*20S proteasome β1 active site pocket

a) The β 1 active site of *Tv*20S-MZB structure (shown in yellow, MZB in blue) with neighbouring residues including Cys46. b) The surface of β 1 with MZB (yellow and blue sticks) and β 2 with and overlay of surface and CP-17 inhibitor (white and white sticks) from the β 2 of *Tv*20S-CP-17 proteasome. CP-17 is shown as white sticks for clarity. c) Detail of β 2 with CP-17 with highlighted residues mainly surrounding indole rings. d) The inset of the figure shows surface of β 2 S3 pocket of *Tv*20S-CP-17 proteasome. The main panel shows surface of β 1 site with CP-17 (white) modelled by superimposition of β 2 and β 1 subunits. The β 1 pocket is lined with much bulkier residues such as Pro27 and polar residues Ser118, Gln112 and Gln127. Unlike in the β 2 pocket lined with smaller and more hydrophobic residues Ala (22,27,124,126,132) and Va128.



Supplementary Figure 10. Cryo-EM workflow of data processing. This image processing workflow was employed to reconstruct the *Tv*20S structures in cryo-EM. Both datasets were acquired at Titan Krios with the Falcon 4i detector under identical conditions using the same setup (refer to Supplementary Table 1 for details). The *ab initio model* served as the starting point for 3D homogeneous refinement (Homogeneous Refinement) to enhance the quality of the maps. Multiple iterations, including 3D classification and 2D classification, were

carried out in several rounds to eliminate unwanted particles, refine the resolution, and improve the maps. The unsharpened maps of the final reconstruction and the gold-standard Fourier Shell Correlation (FSC) curve using different masks are shown. The figures of the maps were generated by ChimeraX². The resolution in italics corresponds to a particular Fourier Shell Correlation Cryosparc and was used to navigate the process of data analysis. The final resolution was estimated by Fourier Shell Correlation job in Cryosparc³ when a tight mask was applied and was estimated to be 2.86 Å for 8IOX and 2.60 Å for 8POT. An example of the *Tv*2OS-MZB route of processing is as follows: 1,436,978 particles were extracted from 6135 processed images. After several rounds of 2D classification to sort out unwanted classes, only 40,916 particles remained. These were then reclassified with independent 2D classifications for the top views and the side views. Separately, classes with 13,724 particles of both views were used for an ab initio model. This initial cryo-EM map was used as a starting model for homogeneous reconstruction, where 13,933 side view particles were used. Next, 4381 particles corresponding to the top views were added in the next round. For the final model, additional particles were removed during several rounds of homogeneous reconstruction and 3D classifications. Note: clear top views of the full proteasome were assumed to be the particles with the highest contrast. Representative Cryo-EM images are shown in subsequent Supplementary Figure 11.



Supplementary Figure 11. Cryo-EM maps coloured according to the calculated local resolution and example of Cryo-EM images. The estimate of local resolution for both *Tv*20S structures a) 80IX and b) 8P0T are compared. c&d show average map correlation coefficient for each chain. Source data are provided as a Source Data file Figures.xlsx. Panels e&f demonstrate the particle density and quality of Cryo-EM images for both datasets.



Supplementary Figure 12. Detailed interactions between small molecular inhibitors and active site pockets of Tv20S. The interactions between amino acid residues surrounding the active sites of Tv20S with the inhibitors; a) & b) MZB and c)-f) CP-17 with highlighted residues shown as sticks suggest interactions with inhibitors. Dashed lines represent hydrogen bridges in panels a)-d) and "hydrophobic" interactions are shown in panels e) & f).



Supplementary Figure 13. MZB 2D representation of interactions between small molecular inhibitor MZB and *Tv*20S active site. a) & b) highlight residues in close contact with inhibitor in β 1 and β 2 active sites. Dashed lines represent hydrogen bridges, and "fan-like structures" demonstrate possible hydrophobic interaction.

Supplementary Information



Supplementary Figure 14. The lig plot CP-17 2D representation of interactions between molecular inhibitor CP-17 and the active site of Tv20S. a) & b) highlight residues in close contact with inhibitor in $\beta 2$ and $\beta 5$ active sites. Dashed lines represent hydrogen bridges, and "fan-like structures" demonstrate possible hydrophobic interaction.



Supplementary Figure 15. The main differences of *Tv*20S and potential clashes CP-17 in human 20S proteasome β 2 and β 5 active site pockets. Structural superimposition of *Tv*20S-CP-17 (8P0T) with human 20S (PDBID = 4R3O, and 5LF3 with 7AWE in right hand panels). a) The rotamer of human 20S (PDB ID = 4R3O) Met131 (β 3) is relatively close and bulky compared to Ala of *Tv*20S, b) another view to the β 2 active site besides 4R3O structure with the overlay of additional two structures 5LF3, 7AWE shows potential clash between CP-17 and Met131. Additionally, the Glu106 (β 3) and Glu22 (β 2) of human 20S suggest additional site of steric hindrance for CP-17 Panels c) & d) show similar contact of possible clash yet the rotamers in different structures such as 7AWE show deflection from the possible clash. Human 20S proteasome is shown in orange.

Supplementary Table 1. Structural alignment of 20S proteasome *Tv*20-CP1 7ZYJ with leishmania and human 20S proteasomes.

The table represents RMSD values for alignments of C α of individual chains *Tv*20-CP17 with *Leishmania tarentolae* 20S proteasome and human proteasome:

	RMSD Leismania7ZYJ	Chain ID	RMSD Human 7PG9		RMSD Leismania 7ZYJ	Chain ID	RMSD Human 7PG9
α1	1.252	А	1.153	β1	0.975	Н	0.976
α2	1.219	В	1.031	β2	0.807	Ι	0.989
α3	0.985	С	1.260	β3	0.886	J	0.898
α4	1.048	D	1.028	β4	0.960	K	1.289
α5	1.252	Е	1.099	β5	0.870	L	1.010
α6	1.116	F	1.144	β6	1.228	М	1.050
α7	1.121	G	1.205	β7	1.178	N	1.489

Overall RMSD values:

20S proteasome Tv20-CP17 with *Leishmania tarentolae* 20S proteasome **RMSD** = 2.729

20S proteasome Tv20-CP17 with human 20S proteasome **RMSD = 2.541**

The average RMSD values:

for all chains of *Leishmania tarentolae* 20S proteasome RMSD = 1.064

20S proteasome Tv20-CP17 with human 20S proteasome RMSD = 1.116*

(RMSD values were calculated for Tv20-CP17 (8P0T as a fix model) and human crystal structure (7PG9) for C α atoms in the Pymol [The PyMOL Molecular Graphics System, Version 2.5.5 Schrödinger, LLC] using the command:align)

Supplementary Table 2	. Cryo-EM statistics for	[.] data collection, re	efinement, and validation.

Microscope Titan Krios Titan Krios Detector Falcon 4i Falcon 4i Magnification (nominal) 165,000x 165,000x Voltage (kV) 300 300 Spherical aberration 2.7 mm 2.7 mm Total electron dose (e') 40 40 Defocus range (µm) (-2.4)-(-0.9) (-2.4)-(-0.9) Exposure (s) 2.25 2.25 Pixel size (Å) 0.75 0.75 Number of Micrographs 7.983 10,037 Final number of particles 14.257 80,145 Map resolution (Å) 2.86 [FSCa44] 2.60 [FSCa44] IffSC threshold] RCI 80IX Symmetry during reconstruction C2 C2 Validation 11.17 11.90 Rotamer outliers 0.93 % 1.21 % Ramachandran plot 1 1 Favoured 97,73 % 97,92 % Outliers 0.00 % 0.004 Angles (°) 0.611 0.659 Mode	Deposited Cryo-EM structure	Tv20S-MZB PDBID: 80IX	Tv20S-CP-17 PDBID: 8P0T		
Detector Falcon 4i Falcon 4i Magnification (nominal) 165,000x 165,000x Voltage (kV) 300 300 Spherical aberration 2.7 mm 2.7 mm Total electron dose (e) 40 40 Defocus range (µm) (-2.4)-(-0.9) (-2.4)-(-0.9) Exposure (s) 2.25 2.25 Pixel size (Å) 0.75 0.75 Number of Micrographs 7,983 10,037 Final number of particles 14,257 80,145 Map resolution (Å) 2.86 [FSC_0.143] 2.60 [FSC_0.143] IPSC threshold] 7ZYJ 8OIX Symmetry during reconstruction C2 C2 Validation 11.17 11.90 Rotamer outliers 0.93 % 1.21 % Ramachandran plot 1 1 Favoured 97,73 % 97,92 % Outliers 0.00 % 0.00 % RMSD bonds 1 0.659 Model vs. Data 1 1.67 Ligands (no.) </td <td>Microscope</td> <td>Titan Krios</td> <td>Titan Krios</td>	Microscope	Titan Krios	Titan Krios		
Magnification (nominal) 165,000x 165,000x Voltage (kV) 300 300 Spherical aberration 2.7 mm 2.7 mm Total electron dose (e) 40 40 Defocus range (µm) (-2.4)–(-0.9) (-2.4)–(-0.9) Exposure (s) 2.25 2.25 Pixel size (Å) 0.75 0.75 Number of Micrographs 7.983 10.037 Final number of particles 14.257 80,145 Map resolution (Å) 2.86 (FSC.n43] 2.60 (FSC.n43] [FSC threshold] 7ZYJ 80IX Symmetry during reconstruction C2 C2 Validation 11.17 11.90 Rotamer outliers 0.93 % 1.21 % Ramachandran plot 1 1 Favoured 97.73 % 97.92 % Outliers 0.00 % 0.00 % RMSD bonds 0.611 0.659 Idea (r) 0.611 0.659 Model vs. Data 1 1 Ligands (no.)	Detector	Falcon 4i	Falcon 4i		
Voltage (kV) 300 300 Spherical aberration 2.7 mm 2.7 mm Total electron dose (e') 40 40 Defocus range (µm) (-2.4)-(-0.9) (-2.4)-(-0.9) Exposure (s) 2.25 2.25 Pixel size (Å) 0.75 0.75 Number of Micrographs 7.983 10.037 Final number of particles 14.257 80.145 Map resolution (Å) 2.86 (FSC_htas] 2.60 (FSC_htas] Pixe trebhold] 7ZYJ 80IX Symmetry during reconstruction C2 C2 Validation 11.17 11.90 Rotamer outliers 0.93 % 1.21 % Ramachandran plot 1 1 Favoured 97.73 % 97.92 % Outliers 0.00 % 0.004 Angles (°) 0.611 0.659 Model vs. Data 1 1 Ligands (no.) 4 (MZB) 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 9.85 / 40.13 0.007 / 8.57 / 29.79 <	Magnification (nominal)	165,000x	165,000x		
Spherical aberration 2.7 mm 2.7 mm Total electron dose (e') 40 40 Defocus range (µm) (-2.4)-(-0.9) (-2.4)-(-0.9) Exposure (s) 2.25 2.25 Pixel size (Å) 0.75 0.75 Number of Micrographs 7.983 10,037 Final number of particles 14,257 80,145 Map resolution (Å) (FSC fireshold) 2.86 [FSC _{0.143}] 2.60 [FSC _{0.143}] Refinement 2 C2 Initial model used (PDBID) 7ZYJ 80IX Symmetry during reconstruction C2 C2 Validation 11.17 11.90 Rotamer outliers 0.93 % 1.21 % Ramachandran plot 1.21 % Ramachandran plot Favoured 97.73 % 97.92 % Outliers 0.00 % 0.00 % RMSD bonds 1 1.659 Iegnst (Å) 0.004 0.004 Angles (°) 0.611 0.659 Model vs. Data 1 1	Voltage (kV)	300	300		
Total electron dose (e) 40 40 Defocus range (µm) (-2.4)-(-0.9) (-2.4)-(-0.9) Exposure (s) 2.25 2.25 Pixel size (Å) 0.75 0.75 Number of Micrographs 7,983 10.037 Final number of particles 14.257 80,145 Map resolution (Å) [FSC threshold] 2.86 [FSC_0.143] 2.60 [FSC_0.143] Refinement 2 C2 C2 Validation C2 C2 Validation MolProbity score 1.62 1.67 1.69 Clashscore, all-atom 11.17 11.90 Refinement Favoured 97.73 % 97.92 % 0.00 % Outliers 0.004 0.004 0.004 Angles (°) 0.611 0.659 Model vs. Data Ligands (no.) 4 (MZB) 0.04 (N.64 / 0.77) P-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00/78.57 / 29.79	Spherical aberration	2.7 mm	2.7 mm		
Defocus range (µm) (-2.4)-(-0.9) (-2.4)-(-0.9) Exposure (s) 2.25 2.25 Pixel size (Å) 0.75 0.75 Number of Micrographs 7.983 10.037 Final number of particles 14,257 80,145 Map resolution (Å) 2.86 [FSC0.143] 2.60 [FSC0.143] FFSC threshold] 7ZYJ 80IX Symmetry during reconstruction C2 C2 Validation 11.17 11.90 Rotamer outliers 0.93 % 1.21 % Ramachandran plot 1 1 Favoured 97.73 % 97.92 % Outliers 0.00 % 0.00 % RMSD bonds 1 1 Lenght (Å) 0.004 0.004 Angles (°) 0.611 0.659 Model vs. Data 1 1 Ligands (no.) 4 (MZB) 4 (CP-17) CC (mask/box/ligand) 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79	Total electron dose (e ⁻)	40	40		
Exposure (s) 2.25 2.25 Pixel size (Å) 0.75 0.75 Number of Micrographs 7.983 10,037 Final number of particles 14,257 80,145 Map resolution (Å) [FSC threshold] 2.86 [FSC0.143] 2.60 [FSC0.143] Refinement 2 C2 Rufinement 72YJ 80IX Symmetry during reconstruction C2 C2 Validation 11.62 1.67 Clashscore, all-atom 11.17 11.90 Rotamer outliers 0.93 % 1.21 % Rumachandran plot 97.73 % 97.92 % Outliers 0.000 % 0.000 % RMSD bonds 0.004 0.004 Ruspl s (°) 0.611 0.659 Model vs. Data 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 Ligands (no.) 4 (MZB) 0.007 / 8.57 / 29.79 0.611 0.659 Model vs. Data 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 0.61 / 0.657 Ligands (no.) <	Defocus range (µm)	(-2.4)–(-0.9)	(-2.4)–(-0.9)		
Pixel size (Å) 0.75 0.75 Number of Micrographs 7,983 10,037 Final number of particles 14,257 80,145 Map resolution (Å) [FSC threshold] 2.86 [FSC 0.143] 2.60 [FSC 0.143] Refinement 2.86 [FSC 0.143] 2.60 [FSC 0.143] Refinement 7ZYJ 80IX Symmetry during reconstruction C2 C2 Validation 11.17 11.90 Rotamer outliers 0.93 % 1.21 % Ramachandran plot 97.73 % 97.92 % Outliers 0.000 % 0.004 Angles (°) 0.611 0.659 Model vs. Data 1 1.90 Ligands (no.) 4 (MZB) 4 (CP-17) CC (mask/box/ligand) 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79 B-factor ligand (min/max/mean) 4.24 / 57.82 / 49.23 12.71 / 39.10 / 24.94	Exposure (s)	2.25	2.25		
Number of Micrographs 7,983 10,037 Final number of particles 14,257 80,145 Map resolution (Å) [FSC threshold] 2.86 [FSC.0.143] 2.60 [FSC.0.143] Refinement 2.86 [FSC.0.143] 2.60 [FSC.0.143] Refinement 72YJ 80IX Symmetry during reconstruction C2 C2 Validation 11.17 11.90 Rotamer outliers 0.93 % 1.21 % Ramachandran plot 97.73 % 97.92 % Outliers 0.000 % 0.004 Angles (°) 0.611 0.659 Model vs. Data 1 1.90 Ligands (no.) 4 (MZB) 4 (CP-17) CC (mask/box/ligand) 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 7.857 / 29.79 B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94	Pixel size (Å)	0.75	0.75		
Final number of particles14,25780,145Map resolution (Å) [FSC threshold]2.86 [FSC h143]2.60 [FSC h143]RefinementInitial model used (PDBID)7ZYJ80IXSymmetry during reconstructionC2C2ValidationMolProbity score1.621.67Clashscore, all-atom11.1711.90Rotamer outliers0.93 %1.21 %Amachandran plotFavoured97.73 %97.92 %Outliers0.00%0.00 %RMSD bondsLenght (Å)0.0040.004Angles (°)0.6110.659Model vs. Data4 (MZB)4 (MZB)4 (CP-17)CC (mask/box/ligand)0.87 / 0.68 / 0.800.84 / 0.61 / 0.71B-factor protein (min/max/mean)3.38 / 99.85 / 40.130.00 / 78.57 / 29.79B-factor ligand (min/max/mean)4.24 / 57.82 / 49.2312.71 / 39.10 / 24.94	Number of Micrographs	7,983	10,037		
Map resolution (Å) [FSC threshold] 2.86 [FSCu.143] 2.60 [FSC0.143] Refinement	Final number of particles	14,257	80,145		
Refinement Initial model used (PDBID) 7ZYJ 80IX Symmetry during reconstruction C2 C2 Validation C2 C2 Validation 1.62 1.67 Clashscore, all-atom 11.17 11.90 Rotamer outliers 0.93 % 1.21 % Ramachandran plot 97.73 % 97.92 % Outliers 0.00 % 0.00 % Outliers 0.00 % 0.00 % RMSD bonds U U Lenght (Å) 0.004 0.004 Angles (°) 0.611 0.659 Model vs. Data U U Ligands (no.) 4 (MZB) 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79 B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94	Map resolution (Å) [FSC threshold]	2.86 [FSC _{0.143}]	2.60 [FSC _{0.143}]		
Initial model used (PDBID) 7ZYJ 80IX Symmetry during reconstruction C2 C2 Validation	Refinement				
Symmetry during reconstruction C2 C2 Validation	Initial model used (PDBID)	7ZYJ	80IX		
Validation MolProbity score 1.62 1.67 Clashscore, all-atom 11.17 11.90 Rotamer outliers 0.93 % 1.21 % Ramachandran plot 97.73 % 97.92 % Outliers 0.00 % 0.00 % Outliers 0.00 % 0.00 % RMSD bonds U 10.659 Lenght (Å) 0.004 0.004 Angles (°) 0.611 0.659 Model vs. Data U 10.679 Ligands (no.) 4 (MZB) 4 (CP-17) CC (mask/box/ligand) 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79 B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94	Symmetry during reconstruction	C2	C2		
MolProbity score 1.62 1.67 Clashscore, all-atom 11.17 11.90 Rotamer outliers 0.93 % 1.21 % Ramachandran plot 97.73 % 97.92 % Outliers 0.00 % 0.00 % Outliers 0.00 % 0.00 % RMSD bonds 1 Lenght (Å) 0.004 0.004 Angles (°) 0.611 0.659 Model vs. Data 4 (MZB) 4 (CP-17) CC (mask/box/ligand) 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79 B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94	Validation				
Clashscore, all-atom 11.17 11.90 Rotamer outliers 0.93 % 1.21 % Ramachandran plot 97.73 % 97.92 % Favoured 97.73 % 97.92 % Outliers 0.00 % 0.00 % RMSD bonds 0.00 % 0.00 % Lenght (Å) 0.004 0.004 Angles (°) 0.611 0.659 Model vs. Data 11.90 4 (MZB) 4 (CP-17) CC (mask/box/ligand) 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79 B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94	MolProbity score	1.62	1.67		
Rotamer outliers 0.93 % 1.21 % Ramachandran plot 97.73 % 97.92 % Favoured 97.73 % 97.92 % Outliers 0.00 % 0.00 % RMSD bonds 0.00 % 0.00 % RMSD bonds 0.004 0.004 Angles (°) 0.611 0.659 Model vs. Data 1 1 Ligands (no.) 4 (MZB) 4 (CP-17) CC (mask/box/ligand) 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79 B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94	Clashscore, all-atom	11.17	11.90		
Ramachandran plot Favoured 97.73 % 97.92 % Outliers 0.00 % 0.00 % RMSD bonds 0.00 % 0.00 % Lenght (Å) 0.004 0.004 Angles (°) 0.611 0.659 Model vs. Data 4 (MZB) 4 (CP-17) CC (mask/box/ligand) 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79 B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94	Rotamer outliers	0.93 %	1.21 %		
Favoured 97.73 % 97.92 % Outliers 0.00 % 0.00 % RMSD bonds 0.004 0.004 Lenght (Å) 0.004 0.004 Angles (°) 0.611 0.659 Model vs. Data 1 1 Ligands (no.) 4 (MZB) 4 (CP-17) CC (mask/box/ligand) 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79 B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94 Resolution Estimates (Å) 5 5	Ramachandran plot				
Outliers 0.00 % 0.00 % RMSD bonds	Favoured	97.73 %	97.92 %		
RMSD bonds Lenght (Å) 0.004 0.004 Angles (°) 0.611 0.659 Model vs. Data Ligands (no.) 4 (MZB) 4 (CP-17) CC (mask/box/ligand) 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79 B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94 Resolution Estimates (Å)	Outliers	0.00 %	0.00 %		
Lenght (Å) 0.004 0.004 Angles (°) 0.611 0.659 Model vs. Data 4 (MZB) 4 (CP-17) Ligands (no.) 4 (MZB) 0.84 / 0.64 / 0.77 CC (mask/box/ligand) 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79 B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94 Resolution Estimates (Å) Kesolution Estimates (Å) Kesolution Estimates (Å)	RMSD bonds				
Angles (°) 0.611 0.659 Model vs. Data 4 4 Ligands (no.) 4 (MZB) 4 (CP-17) CC (mask/box/ligand) 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79 B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94 Resolution Estimates (Å) Ketor Ketor	Lenght (Å)	0.004	0.004		
Model vs. Data Ligands (no.) 4 (MZB) 4 (CP-17) CC (mask/box/ligand) 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79 B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94 Resolution Estimates (Å) Estimates (Å) Estimates (Å)	Angles (°)	0.611	0.659		
Ligands (no.) 4 (MZB) 4 (CP-17) CC (mask/box/ligand) 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79 B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94 Resolution Estimates (Å) Example 100 / 1	Model vs. Data				
CC (mask/box/ligand) 0.87 / 0.68 / 0.80 0.84 / 0.64 / 0.77 B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79 B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94 Resolution Estimates (Å) Example 10.000 10.000	Ligands (no.)	4 (MZB)	4 (CP-17)		
B-factor protein (min/max/mean) 3.38 / 99.85 / 40.13 0.00 / 78.57 / 29.79 B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94 Resolution Estimates (Å)	CC (mask/box/ligand)	0.87 / 0.68 / 0.80	0.84 / 0.64 / 0.77		
B-factor ligand (min/max/mean) 42.42 / 57.82 / 49.23 12.71 / 39.10 / 24.94 Resolution Estimates (Å)	B-factor protein (min/max/mean)	3.38 / 99.85 / 40.13	0.00 / 78.57 / 29.79		
Resolution Estimates (Å)	B-factor ligand (min/max/mean)	42.42 / 57.82 / 49.23	12.71 / 39.10 / 24.94		
	Resolution Estimates (Å)				
FSC model unmasked (0/0.143/0.5) 2.7 / 2.7 / 3.1 2.5 / 2.5 / 2.8	FSC model unmasked (0/0.143/0.5)	2.7 / 2.7 / 3.1	2.5 / 2.5 / 2.8		

Refinment and model vs data values were calculated using Phenix 1.20.1-4487

Supplementary references

- 1 Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320-324 (2014). https://doi.org:10.1093/nar/gku316
- 2 Meng, E. C. *et al.* UCSF ChimeraX: Tools for structure building and analysis. *Protein science* : *a publication of the Protein Society* **32**, e4792 (2023). https://doi.org:10.1002/pro.4792
- 3 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature methods* **14**, 290-296 (2017). https://doi.org:10.1038/nmeth.4169

Supplementary Information

Uncropped gels for Supplementary Figure 3a, b

denaturing gel - silver stained

fluorescence



- 1. Sf lysate non-infected cells
- 2. alpha, beta, ump-1 coexpression
- 3. alpha, beta coexpression
- 4. 6. native Tv20S
- 5.-7. rTv20S
- 8. c20S





- 1. Sf lysate non-infected cells
- 2. alpha, beta, ump-1 coexpression
- 3. alpha, beta coexpression
- 4. 6. native Tv20S
- 5.-7. rTv20S
- 8. c20S

Uncropped gel for Supplementary Figure 3c



rTv20S + Ump time 0
 rTv20S + Ump time 24 h
 rTv20S + Ump time 72 h
 rTv20S-Ump time 0
 rTv20S - Ump time 24 h
 rTv20S - Ump time 72 h

Uncropped gels for Supplementary Figure 3e

Native silver stain



1) rTv20S+DMSO 2) rTv20S+CP17 3) rTv20S+MZB Fluorescence

