TSEN	154	-1	8A	14A	24F	34	4P 🔺	44L	54R	64	0
HS		MEPEPEP	ÅAV	EV <mark>P</mark> A <mark>G</mark> RVI	LSAR <mark>EL</mark> FAA	RSRSQKLP	QRSHGPRD	FLPDGSAA	QAERLRR	REELWQI	LAE 68
MM DR	1	<mark>MEP</mark> E <mark>PEP</mark>	<mark>G</mark> SV	EV <mark>P</mark> A <mark>G</mark> R <mark>V</mark> I	LSASELRAA	RSRSQKLP	QRSH <mark>GPKD</mark>	FLPDGSEA	QAERLRL	CRQELWQI	LAE 68
DM	1	<mark>MD</mark> TKT <mark>Q</mark> D	SEYVE		LSAQ <mark>E</mark> LIA <mark>H</mark>	RHTEFE	PPSGGL <mark>K</mark> R	TKN <mark>EGT</mark> AA	EVEELKR	AQEY <mark>L</mark> RAÇ	LSV 68
SP SC		MDAQDED KKTDQVTTS									
50	1 Mgr A			100000000000000000000000000000000000000	20110 QLINDI	VORMIN			V DDD I IN	HORAFIT DI	
		73R	83 <u>R</u>	935	103G	112	2т.	122E	132-	132	_
HS	69 ORV-1	ER <mark>LG</mark> SLVAAE	WRPEEGEV	ELKSPAGK	WOTMGFS-	EOGRORLH	PEEALYLL	ECGSIHLF	HO	'	DLP 134
MM	69 E <mark>RV</mark> -1	ER <mark>LG</mark> SLVAAE	WKPEEGFV	/ELTSPAGKI	FWQTMGYS-	EEGRQRLHI	PEEALYLL	ECGSIQLF	YQ		DLP 134
DR DM		ER <mark>LGG</mark> RALAI									
SP	75 E <mark>RL</mark> I	S- <mark>A</mark> KHAIIAT	WNA <mark>Q</mark> N <mark>G</mark> MS	CVEKAHGP1	LFK <mark>TMGT</mark> AD	S <mark>Q</mark> N <mark>R</mark> MWLL]	PEETLYLV	ERG <mark>S</mark> MECW	SEE		•GLP 142
SC	81 SIR <mark>G</mark>	TT <mark>V</mark> KSEVR <mark>G</mark> Y	YVPHKHQ <i>I</i>	AVLLK <mark>P</mark> K <mark>G</mark> SI	FMQTMGRAD	ST <mark>GE</mark> LWLDI	FHEFVYLA	ERGTILPY	YRLEAGS	NKSSKHEI	EIL 161
		1 2 0 7	1.4.0	1505	1.000	1.7.0			1000		
HS		139E EAYQLLLTDH	1	158F		I	1	88A SVOHLED			
MM	135 <mark>LSIQ</mark>	EAYQLLL <mark>T</mark> ED	-TLSFLQ	QVFS <mark>H</mark> L <mark>KR</mark> I	L <mark>GYVVRR</mark> FQ	LSSVVS <mark>P</mark> YI	ERQ <mark>L</mark> NLD <mark>G</mark>	YAQCLED	SGKRKRS	S-S <mark>C</mark> RS <mark>V</mark> N	IK <mark>KP</mark> 213
DR DM		DGYERFL <mark>S</mark> SD DAYVLLLGEL									
SP	143 MSLO	AVYSASIPLC	GSLEN	LVYAHLRRO	GFSVIRSN	LVPVKEDE	Y-R		CDS	KIMN	IF <mark>K</mark> D 199
SC	162 LSME	DLYSL-F <mark>S</mark> SQ	QEMDQ	FVFAHL <mark>KR</mark> I	L <mark>G</mark> FILK <mark>PS</mark> N	QEA <mark>AV</mark> KTSI	FF <mark>PL</mark>		KKQ <mark>R</mark> SNL	QA <mark>I</mark> I	WRL 223
			19N	223P	233P				2,56G	266P	
HS MM	215 K 214 K	ALDN VLQN	SI	LQPKSLAASS PPVSLAASS	SPPPCSQPS SSPACDOSS	QCPEEKPQI OYPEEKSOI	ESS DSS	PMKGPG	GPFOLLG LPLOFLG	SL <mark>GPSPGP</mark> SSEPCSDI	ARE 273
DR	111 I	STEE	KTTSH	EA <mark>H</mark> NELAET-				SK	LPADLOE	PVHI	ADA 144
DM SP	209 TYP I 200 T	AEKRMEDLKE	LITSQKAN	ED <mark>P</mark> NI <mark>VV</mark> DP	IDLKF <mark>G</mark> QRK	RMAVEK <mark>P</mark> II	EE <mark>P</mark> ASKKE	LFLGLG	LVDQLKG KASOTLO	EFA	281
SC	224 L							SLFKIQ	E-LSLFS		236
		6G 28	4A	294R	304P	314L			34A	344R	
HS		SWES <mark>G</mark> RAE	N <mark>GVTGAG</mark> H	KR <mark>RW</mark> NFEQIS	S <mark>FP</mark> NMAŠDSI	RHTL <mark>L</mark> RA <mark>P</mark> A					
MM DR		DRESHKIE WWMEEDVQQK									
DM	282		YAP	KFEEIFEKFI	DLVOLE						298
SP SC	214		GFFNA	RKLAF	PFSKRRR	QSIL <mark>L</mark> HD					236
	354	E 364	Е 3	374R	384L	3940	404P	41	4A	424I	
HS	353 REHH	AEAAQEQEDV	NAD <mark>P</mark> EVQF	RCSSWREYK		RSQRRAPHI	LW <mark>GQP</mark> VTP	LLS <mark>PG</mark> QAS	SPAVVLQ	HISVLQTT	HLP 433
MM	352 REQ0	AVVQQ <mark>F</mark> RE <mark>D</mark> V RERDR <mark>Y</mark> RR <mark>D</mark> V	NAD <mark>P</mark> EVR	CSSWQEYK	LLQRRQTQ	KSQPRPPHI	LW <mark>G</mark> QSVT <mark>P</mark>	LLDPDKAD	CPAAVLQ	HISVLQTI	HLA 432
DR DM	299	SODYTKDE									306
SP	237			TFYTYEI	VYHDLQIV	RGYV				PIACN	260
SC	252			1 5P		VR5 <u>V</u>					
	4245	1110		10	4 6 4 33	4740	4045	40.4	<u>^</u>	407	
HS	434D 434 DGGA	444G LLEKS <mark>GGL</mark> E			464N	474C				497-	496
MM	433 D <mark>GG</mark> Y	RLLEKSGGLQ	ISFDVYQA	ADAVATF <mark>RK</mark> I	NS <mark>PG</mark> KPYVR	MCI <mark>S</mark> GFD <mark>D</mark>	PVPDLC <mark>S</mark> L	KCLTYQSG			495
DR DM	307 EGLS	<mark>R-LSP</mark> SDQWK LDT <mark>P</mark> SALK	ISFDVYQH	DTVAEFKKS	SHPGKPYTR SADNADNEN	MCVCSFAGI	PVPDLHAV	KQLSFQSQ			368
SP	261	LITSSDSLFO	ITFHAYK	SASEKKS	SALSEPDER	ICVVSSOD	TLLPTI	FEIDALFS	S		316
SC	279 AQSD	REFQKVKDIP	L <mark>T</mark> FKVWK <mark>I</mark>	H <mark>S</mark> NFKK	RD <mark>PG</mark> L <mark>P</mark> DFQ	VFV <mark>Y</mark> NKN <mark>D</mark> I	DLQHF <mark>PTY</mark>	KELR <mark>S</mark> MFS	SLDYKFE	FLSEIEDD	DDW 357
		497-	497	7- 49	97-	497-	497-	497-	1	02F	512F
HS MM	497								_	ALV <mark>DHG</mark> DI ALV DHG DI	
DR	369								– D <mark>VP</mark> VTF.	AVV <mark>dhg</mark> di	SFY 385
DM SP	362					T <mark>P</mark> T.1	RONM <mark>P</mark> OH-		-APLLVI GYRNIII	SVSESKQI	QAF 378
SC		YVEDI <mark>P</mark> RKEY	IHKRSAKS	QTEKSESS	KASFQKKT	AQSSTRKK	RKAY <mark>PP</mark> HI	QQNRRLKI	GYR <mark>S</mark> FII.	AIMDNGLI	SFV 438
		521P	+5	+15	5						
HS	514 S-FR	DFTL <mark>P</mark> QDVGH									526
MM DR		DFTL <mark>P</mark> RDL <mark>G</mark> H DFKLPTDVY-									525 397
DM	379 LY <mark>Y</mark> I	s									383
SP SC		DVCFEEKVYT EADF <mark>G</mark> S <mark>ES</mark> VW									384 467
20				FULL DAVA							107

<u>Supplementary Data 1</u>: Eukaryotic TSEN54 alignment. Alignment of TSEN54 protein from *Homo sapiens* (HS), *Mus musculus* (MM), *Danio rerio* (DR), *Drosophila melanogaster* (DM), *Schizosaccharomyces pombe* (SP), and *Saccharomyces cerevisiae* (SC). Under the alignment, the structured regions (teal box) and not observed in the structure (black solid line) are noted. Key features important for the structural integrity of the complex are the conserved β -sheet (509-517, a thick black arrow) and the buried L10 like loop (506-509, thin black line in the structured region). PCH mutations (black circle) and some key residues (R36 and K41, teal star) are also annotated.

TSEN15		10T	20G	30A	4 0 M	46Y	56D	59Q	
HS	1 MEER <mark>G</mark> DS	EPTPGCSGLG	PGGVRGFGDG	GGAPSWAPE	DAWMGT-H	PKYLEMMELD	IGDA	-TOVYVAFLVY	L 68
MM	1 MEERSDS	EPTPGCSGPG	PAPVRD	GGAHTWAPE	DA <mark>WMG</mark> TH	PKYLEMMELD	IGDA	-TQVYIAFLVY	L 65
DR	1		M	IDE <mark>AN</mark> PANPP	PNWIQQH	T VYQELMRL <mark>E</mark>	VGDS	-VQV <mark>H</mark> AAFLV <mark>Y</mark>	M 42
DM	1MVRQ	Y <mark>P</mark> I			<mark>AG</mark> CSQT	VMYSLDFQLE	LN <mark>D</mark> FTTVS <mark>P</mark> D	DLTAALGFQIY	N 45
SP	1			-MQ <mark>H</mark> NTFL <mark>P</mark> N	DA <mark>VE</mark> EQKQN	PYY-GIL	<u>-</u>	RAVE	T 29
SC	1		МК		-D <mark>II</mark> NCRREIT	SRRLRK <mark>I</mark> RMA	T	<mark>T</mark> DII <mark>S</mark> LVK	N 32
	69D	79-	83-	84-	921	96E	101G	111T	
HS	69 DLMESKS	WHEVNCV	<mark>G</mark>	L-PEL	OLICL	VGTEI	EGEGLQTVVP	TPITASI TPISASI	S 115
ММ	66 DLMESKS	WHEVNCV	<mark>G</mark>	<mark>I</mark> - <mark>P</mark> EL	Q <mark>LICL</mark>	LGTEI	E <mark>GEG</mark> L <mark>QT</mark> VV <mark>P</mark>	T <mark>PI</mark> SA <mark>S</mark> I	<mark>S</mark> 112
								L <mark>PS</mark> HRSI	
								K <mark>S</mark> LHKKL	
								L <mark>Y</mark> LQE T I	
SC	33 NLLYFQM	WTEVEIL	QDDLS	SWK <mark>G</mark> N <mark>S</mark> LRLL	R <mark>G</mark> R <mark>PP</mark> HKLSND	VDTEHENSLS	SPRPLEFILP	INMSQYKENFL	T 103
	117N	127R	130-	139F	149 <u>T</u>	159F	169L	+8	
HS 1	16 HNRIREI	LKASRKL	QGDPI	LPMSFTLAI	VESDSTIVYYK	LTDGFMLPDP	QNISLRR		171
					VE <mark>SDSTIVYY</mark> K				168
DR	90 HRELRSI	LDRGS		<mark>P</mark> MLLCA	VA <mark>SDST</mark> LVYQR	LC <mark>DGFLTP</mark> DP	PVDIQDQ <mark>G</mark> RR	QHRKRRLQT	147
					VDNTANILYY <mark>Q</mark>				170
					MC <mark>SDSTIVYY</mark> K				143
SC 1	04 LECLSQT	FTHLR	S <mark>I</mark>	STERILLAI	IND <mark>DGTIVYY</mark> F	VYK<mark>GV</mark>RK<mark>P</mark>KR	N		148

<u>Supplementary Data 2</u>: Eukaryotic TSEN15 alignment. Alignment of the TSEN15 proteins from *Homo* sapiens (HS), *Mus musculus* (MM), *Danio rerio* (DR), *Drosophila melanogaster* (DM), *Schizosaccharomyces* pombe (SP), and *Saccharomyces cerevisiae* (SC). Under the alignment is a blue box showing the regions mapped in the structure. Labelled features include the conserved β -sheet (150-155), annotated with a thick black arrow, and the buried L10 like loop (144-150) annotated with a thin line. PCH mutations are indicated by a black circle.

TSEN34	-31	-21	-11	-1	10R	200	30G	40P	
								40P LPRGPRONSRL LPRGPRONSRL LAROPRONVRL ANTKGWSPNOS LPOLPOONVFL LPTAAOONVFL	
	50L	60L	705	75D	82A	92E	102A	111E	
								-RQELLEKITE -RQELLEKIVE	
DR 81	1 R <mark>PLELRPE</mark> I	EALLLLEEEK	TAAVRLDTQC	DFEDRE	AHSEAVRLYEA	ASLES <mark>SYO</mark> EOC	V <mark>LAL</mark> ED <mark>K</mark> KKT	-LNRVM	- 151
								ETER <mark>Y</mark> MD <mark>K</mark> ILT	
								KKIE <mark>IM</mark> K <mark>KLG</mark> R EHIAKLK <mark>KIG</mark> R	
50 5.								EIIIAK	1 1 1 4
									_
	121A	1315	1	149E		1655	175V	185V	
HS 120 MM 120	D QAAKKOKLI	EQASGASSS-(DEAGS SQAAKE	D-ETSDGQA	SGEQEEAG	PSSSQAGPS	NGVAPLPRSA	LLV <mark>Q</mark> LATARPR LLIQLATARPR	P 194
DR 152	2	NEKENG	EGSGSAVRE-			RLEAL	DOSFRFPWSA	MAVOLCTARSG	F 194
DM 120	KRKKLLK-	-Q <mark>G</mark> KNELA	AA		I	LTDKD-VLLEM	SVNFKFDRQN	MAV <mark>O</mark> LC <mark>T</mark> ARS <mark>G</mark> ALL <mark>EVH</mark> CEH	- 164
SP 128	B L <mark>P</mark> L			D-		<mark>PL</mark> N Y D	E-HDSFDL	<mark>SWIP</mark> VD <mark>T</mark> VTTR LFIDIAN <mark>T</mark> SMI	I 155
SC 115	5 NDK		<mark>T</mark> TAEE-	гб-		RLDKS	SNNDQLIESS	LFIDIAN <mark>T</mark> SMI	151
	195V	2050	210P	220R	230R	240F *	250D *	260A IAQCWAPED	
HS 195	5 VKAR <mark>P</mark> LDWI	RVQSKD	-WPHAGRPAHE	LRYSIYRDL	WERGFFLSAA(GKFGGDFLVYP	GDPLRFHAHY	IAQCWAPED	T 268
MM 201 DR 195	1 VKAKPLDWI	RVQSKD	-WPHAGRPAHE	LRYSIYRDL	WERGFFLSAAC	GKFGGDFLVYP	GDPLRFHAHY	IAQCWSA <mark>ED</mark> IALCVSLDE	P 274
DR 19: DM 16:	5 -S <mark>n</mark> CPEER			TRESVIRUL			GDPL RFHANY GDPMTVHASH	IVILHEQP	E 20/ V 231
								LVNSYKWDO	
SC 152	2 RDIRSDSD	S <mark>L</mark> SRD <mark>D</mark> ISDLI	L <mark>F</mark> KQ <mark>Y</mark> RQ <mark>AG</mark> KM	IQT <mark>Y</mark> FLY <mark>K</mark> AL	RD <mark>QGYVL</mark> SPGO	G <mark>R</mark> FGGKFIA <mark>Y</mark> P	GDPL <mark>R</mark> FH <mark>SH</mark> L	TIQDAIDYHNE	P 232
		278G	288L	298K	299-	303 <u>5</u>			
HS 269			LLLCSPOPDC		299- <mark>VVY</mark> 1	SLOWASLO-			310
MM 275	5 IPLODLVSA	AGRLGTSVRK	LLLCSPOPDO	K	VVY	F <mark>SLOWASLQ-</mark>			316
DR 268	B MPVCDILA	IA <mark>RLGS</mark> NVKK	VLLCSP0-DC	D	VV <u>Y</u> S	S <mark>SLOWS</mark> SMV-			308
					EE <mark>IHY</mark> (276
					EELLLL <mark>PVRH</mark> H				284
<i>SC</i> 233	J IDLISMIS	GARLGTTVKKI	LWVI <mark>GG</mark> VAEE1	K	E <mark>TH</mark> FI	FSIEWAGFG-			275

Supplementary Data 3: Eukaryotic TSEN34 alignment.

Alignment of the TSEN34 (3' endonuclease) proteins from *Homo sapiens* (HS), *Mus musculus* (MM), *Danio rerio* (DR), *Drosophila melanogaster* (DM), *Schizosaccharomyces pombe* (SP), and *Saccharomyces cerevisiae* (SC). Labelled conserved features are the β -sheet (298-306) (thick black arrow), and catalytic residues Y247, H255, and K286 (*). The well-established PCH-linked mutation site R58 is annotated (solid black circle) as are the cation- π residues R279 and W306 (pink stars).

TSEN2		10K	20E	30D	37-	37-	37F	47N	
HS	1 MAEAVFH			ODHGPLKE-	۲ <i>۰</i>	۲, 	FKTFRAEM	47N INNNVIVRNAEDI ISNNVVVRGTEDM IGHHVSVTD <mark>P</mark> AHI	57
MS	1 MAEAVFR	A <mark>PKR</mark> KRRV <mark>YE</mark> S	SYESPLPIPF	GODOGPRKE-			FRIFQAEM	ISNNVVVR <mark>G</mark> TEDM	1 57
DR	1 MTDAVFQ	APK <mark>R</mark> RA <mark>R</mark> VYE	CFEAAL <mark>PVP</mark> LI	NSADD <mark>G</mark> D-H-			T LLCRAEI	I <mark>G</mark> HHVSVTD <mark>P</mark> AHI	56
DM SP	1	MSKNHEN	VKDALPISL	AYPLPPIILT	NPLTWIPYIYI	RYLFRKT <mark>P</mark> RO	VOWOC-OLHE:	SDLSCVVTDSEAI	68
SC	1	-MSKGRVNQKI	R <mark>Y</mark> KY <mark>PLPIHP</mark>	VDDL <mark>P</mark> ELILH	N <mark>P</mark> LSWLYWAYI	RYYKSTNALN	DKVHVDFIGD	TTLH ITV QDDKQM	72
		67G	77S	87A	96M	106H	116R	126R	
HS	58 EQ <mark>LYG</mark> KG	1	1			1	1	VRRILKDYTKPL	136
MS	58 EQ <mark>LYG</mark> KG	YFGKGIL <mark>SRS</mark> I	R <mark>PNFTIANP</mark> TI	LAARWK <mark>G</mark> VQ-	TDM <mark>P</mark> IITSEK	QHRVEWARD	FLRRQ <mark>GH</mark> DES	IVQ <mark>KILTDY</mark> TEPL	136
DR DM	57 QQLYHRG	YFGKGVFSRSI	R <mark>PEHMI</mark> S		RCLPVISLSE			SVN <mark>K</mark> ALQILTEPV	. 131
SP								ARRRFORK	
SC	73 LY <mark>LW</mark> NN <mark>G</mark>	FFGTGQF <mark>SR</mark> SI	E <mark>P</mark> TWKART	Е	AR <mark>LGL</mark> NDT <mark>P</mark> L	N-R <mark>GG</mark> TKSN	TETEMT <mark>LE</mark> KV	QQRRLQRL	• 136
									-
		146E	156G	166G	176D	186P	196G	203-	
HS				T <mark>AGG</mark> ER <mark>P</mark> SVV	NGDSGKSGGV			TESFEK	
MS DR	137 ELPCREE	KEETPQHEPLS OEPISKSNCDI	SSKADSSLEGI	TDLTPDO	PGGAGQSDDL TE-SAKP	PGLGTHSDCL	DEGPGHATL	AAA <mark>SP</mark> SSHN <mark>G</mark> <mark>A</mark> EI <mark>YP</mark> EKLDQ	212
DM	1		MOFTFF			CKRKRCE	UKNANFTPI	<mark>P</mark> S	. 24
SP SC	121 OFKAQRA	Y V	RENI	RARER		QLL	LENGKPIPASI	LEEDAEL <mark>P</mark> EYLTK IDEENI <mark>LL</mark> EKQRE	162
50	137 BERKERA	K				L .		IDEENI <mark>ID</mark> ERQRE	- 1/1
		215-	223C	233G	2431	253P	263A	273-	
HS MS								EAA <mark>P</mark> NEE <mark>LV</mark> QR EAAAASDEK <mark>LL</mark> KR	
DR	192 SLKADEK	CRDHNDWISOC	GCRANEERMI	N	RTMH	SOTYEYVLV	EESEDERSOH	I-TOENSSTDOP	254
DM CD	25		<mark>G</mark> QD <mark>G</mark> ER <mark>Y</mark> I	K <mark>G</mark> R	F1	GLSVEIV	-DPEQAKSLHI	DNGCFGKG LQLTF <mark>P</mark> EAF <mark>FL</mark> AS	60
SP SC	172 SLRKFK-				I	LKOTEDVGIV	A-000DI-SE	SNLRDEDNNLLDE	208
		291-	299L	309Y	319E	328L	338T	348Y	
HS	282 NRLICR			LVYALGCLSI				YHYF <mark>RSK</mark> GWVPK	7 358
MS								YHYF <mark>RS</mark> KGWVPKV	
DR DM	255 AELVCRI	OPASCOENET	LOLSYEEAFFI	LVYALGCLSI	YYNGEALSVA-	-QLWTMFRSL	OPNFCFSYAA Nekevenias	YHYY <mark>RSKG</mark> WVPKS YVYLKSKNWVIKS	331
SP	199 L <mark>G-VLR</mark> I	NYENPN	F <mark>E</mark> LL <mark>P</mark> ILKLF <i>i</i>	ANIVAN <mark>S</mark> VAL	THDYS-LQQ <mark>S</mark> -	<mark>H</mark> ED <mark>PI</mark> IE <mark>P</mark> I	DNKFLTELAA	YFYF <mark>R</mark> Q <mark>Q</mark> GWVV <mark>K</mark> N	1270
SC	209 N <mark>G</mark> -D <mark>L</mark>	LPLESI	LELMPVEAMFI	L <mark>T</mark> FAL <mark>P</mark> VLDI	S <mark>P</mark> ACL <mark>AG</mark> KLF-	QFD <mark>A</mark> KYKD	IHSFVR <mark>SYVI</mark>	YHHY <mark>RS</mark> HGWCV <mark>R</mark> S	278
		368L	* ^{378A}	388D	398P	408S	* ^{418L}	428T	
HS	359 GLKYGTD	LLLYRKGPPF	(HA <mark>SY</mark> SVIIE)	LVDDHFEGSL	RR <mark>PLSWK</mark> SLA	ALSRVSVNVS	ELMLCYLIK	PSTMTDKEM	434
MS DR	354 GLKYGTD	LLLYRKGPPF	(HASYSVIIE)	LLDDNYE <mark>G</mark> SL	RRPFSWKSLAA LRPFSWRSLA	ALSRVSGNVS	KELMLCYLIK	PSTMTAEDM PSDVTGDLL	429
DM	139 GIKFGGD	FLIYKQSPRL	THAS FLVIVQ	KPGDTD	HYQSKNLK	VQRVAETS D	KDVLLLTVQP1	KKISS	- 206
SP								PSIE-DFNKIWKN	
SC	279 GIKFGCD	YLLYKRGPPPF(DHAEFCVMGLI	D H	DVSKDYTWYS	SIA <mark>RVVGGA</mark> K	TFVLCYVER	LISEQEAIALWKS	5 3 5 0
									-
		436S	446Q	453-	461D				
HS MS	435 430			WVSSRE					465 460
DR	408								438
DM	207	RDLQEA	rvt <mark>et</mark> ivr <mark>r</mark> fi	NYLT <mark>FVQ<mark>SK</mark>Q</mark>	Q				233
SP SC	347 QASMNEW 351 NNF								380 377
20		IN HOLY							5.,
			-						

<u>Supplementary Data 4</u>: Eukaryotic TSEN2 alignment. Alignment of the TSEN2 (5' endonuclease) proteins from *Homo sapiens* (HS), *Mus musculus* (MM), *Danio rerio* (DR), *Drosophila melanogaster* (DM), *Schizosaccharomyces pombe* (SP), and *Saccharomyces cerevisiae* (SC). The conserved β -sheet (444-452) (thick black arrow), catalytic residues Y369, H377, and K414 (*), cation- π residues 409 and 452 (orange stars), and two residues known to be mutated in some cases of PCH, Y309 and G312 (black solid circles) are all annotated.