

## Evolutionary conservation between cysteine synthase (CS) and cystathionine $\beta$ -synthase (CBS) enzymes from various organisms

The boxed area represents a highly conserved region present in both the CBS and the CS enzymes; this region also displays significant structural conservation with several other members of the  $\beta$  family of the PLP-dependent enzymes such as serine/threonine deaminases and  $\beta$  subunit of tryptophan synthase. In human CBS, this region comprises amino acid residues 84-382. The positions of the heme binding residues in the human CBS enzyme (Cys52, His65) are marked with a filled triangle. The PLP binding residue in human CBS (Lys119) is marked with a filled box. The residues are colored according to the following criteria: **Small + hydrophobic, including aromatic residues**; **acidic residues**; **basic residues**; **hydroxyl + amine + basic residues**. Asterisks indicate identical amino acid, dots indicate semi-conservative replacement.

CBS HS	-----MPSETPQAEVGPPTGCPHRSGPHSAKGSLEKGSPEDEKAKEPLWIRPDAPSRC	52
CBS OC	-----MPSETAQAGEGPAQCPLHSGAQGSDRSLDQRPPGNKDAPERVWIRPDVPSRC	52
CBS RN	-----MPSGTSQCEDGSAGCPDLEVPQEKQLEKASGDK---ERVWISPDTPSRC	49
CBS MM	-----MPSGTSQCEDGSAGGFQHLDMHSEKRQLEKGPSGDK---DRVWIRPDTPSRC	49
CBS FR	MAGFHRIISLLSAFSQRGEMSAKQKQETDKGVQDEAHSEGPESSEL---KDWIRPDPSRC	56
CBS DM	-----MPQPKPYERP-----ADFDIDPGKPSKC	22
CBS DD	-----MSAPEGPKSKC	10
CBS SC	-----	
CBS TC	-----	
CS LT	-----	
CS SO	-----MASLVNNAAYALRTSKLELREVKNLANFRVGGPPSSLSCNNFKK	43
CS AT	-----	
CS CE	-----	
CS ST	-----	

		84	
	▲	↓	
CBS HS	TWQLGRPAS <b>ES</b> PHHHTAPAKSPKILPDILK <b>I</b> GDTPMVRINKIGKFFGLKCELLAKCEFF	112	
CBS OC	TWELGRPVADSPHQHAALAKSPKILPDIL <b>Q</b> IGDTPMVRINKIGKFFGLKCELLAKCEFF	112	
CBS RN	TWQLGRPMADSPHYHTVPTKSPKILPDIL <b>R</b> IGNTPMVRINRISKNAGLKCELLAKCEFF	109	
CBS MM	TWQLGRAMADSPHYHTVLT <b>K</b> SPKILPDIL <b>R</b> IGNTPMVRINKISKNAGLKCELLAKCEFF	109	
CBS FR	TWELGAPASKSPHSHQPRIKPPSILPNILGNIGNTPLVRLNKIPKQFGIKCDILVKCEFF	116	
CBS DM	KWHLG-TAEKSPHIHRGIAHRQOITPNI <b>EV</b> IGCTPLVKLNIPASDGI <b>EC</b> EMYAKCEFL	81	
CBS DD	TWTPN-TTENTPHTTRRTP <b>KK</b> -LIMDNILDNIGGTPLVRVNVK <b>SS</b> --D <b>LE</b> CELVAKCEFF	66	
CBS SC	---MTK <b>EQ</b> Q <b>AD</b> SRH-----NVID <b>I</b> VGNTPLI <b>AK</b> KL <b>PK</b> ALG <b>IK</b> PQ <b>IY</b> AKLELY	46	
CBS TC	---MAQ <b>NP</b> SAVNN <b>KKE</b> V <b>PD</b> RR <b>TC</b> ILDT <b>VLD</b> AI <b>IG</b> ST <b>PC</b> IRLN <b>HVP</b> KK <b>HG</b> VCC <b>DV</b> VAKCEFL	57	
CS LT	-----MT <b>SAT</b> PHDH-----ILAN <b>AL</b> E <b>AI</b> GN <b>TP</b> CI <b>RL</b> N <b>RV</b> P <b>Q</b> K <b>HGI</b> Q <b>CE</b> V <b>V</b> AKCEFF	46	
CS SO	VSS <b>SPIT</b> CA <b>V</b> SL <b>SP</b> ST <b>IE</b> GL <b>NI</b> A <b>ED</b> V <b>SQ</b> L <b>IG</b> KT <b>PM</b> V <b>YL</b> N--N <b>V</b> SK <b>GS</b> V <b>AN</b> IA <b>AK</b> LE <b>SM</b>	101	
CS AT	-----MED <b>RC</b> L <b>IK</b> ND <b>IT</b> EL <b>IG</b> NT <b>PM</b> V <b>YL</b> N--N <b>V</b> D <b>GC</b> V <b>AR</b> IA <b>AK</b> LE <b>MM</b>	41	
CS CE	-----MS <b>REL</b> M <b>VE</b> T <b>GG</b> EL <b>IG</b> NT <b>PL</b> L <b>KL</b> N-- <b>KIG</b> KL <b>GAS</b> I <b>AV</b> K <b>VE</b> Y <b>M</b>	40	
CS ST	-----SK <b>IY</b> ED <b>NS</b> L <b>T</b> IG <b>HT</b> PL <b>VRL</b> N-- <b>RIG</b> NG--- <b>RIL</b> AK <b>VE</b> SR	34	
		* * * * *	

CBS HS	NAGGSVKDRISLRMI <b>EDA</b> ERD <b>GT</b> L <b>KPG</b> -D <b>T</b> IE <b>PT</b> SGNT <b>GI</b> GLALAA <b>AVR</b> GY <b>RC</b> I <b>IV</b> MP <b>E</b>	171
CBS OC	NAGGSVKDRISLRMI <b>EDA</b> ERAG <b>T</b> L <b>RPG</b> -D <b>T</b> IE <b>PT</b> SGNT <b>GI</b> GLALAA <b>AVK</b> GY <b>RC</b> I <b>IV</b> MP <b>E</b>	171
CBS RN	NAGGSVKDRISLRMI <b>EDA</b> ERAG <b>T</b> L <b>KPG</b> -D <b>T</b> IE <b>PT</b> SGNT <b>GI</b> GLALAA <b>AVK</b> GY <b>RC</b> I <b>IV</b> MP <b>E</b>	168
CBS MM	NAGGSVKDRISLRMI <b>EDA</b> ERAG <b>N</b> L <b>KPG</b> -D <b>T</b> IE <b>PT</b> SGNT <b>GI</b> GLALAA <b>AVK</b> GY <b>RC</b> I <b>IV</b> MP <b>E</b>	168
CBS FR	SAGGS <b>IK</b> DRI <b>AL</b> RM <b>VED</b> AERAG <b>L</b> L <b>KPG</b> -D <b>T</b> IE <b>PT</b> SGNT <b>GI</b> GLAL <b>VAS</b> V <b>KY</b> RC <b>VI</b> T <b>MP</b> D	175
CBS DM	NP <b>GS</b> VKDRIGY <b>RM</b> V <b>QDA</b> E <b>EQ</b> L <b>LKPG</b> -Y <b>T</b> IE <b>PT</b> SGNT <b>GI</b> GLAMAC <b>AVK</b> GY <b>KC</b> I <b>IV</b> MP <b>E</b>	140
CBS DD	NAGGSVKDRIG <b>HR</b> M <b>IVDA</b> E <b>ES</b> GR <b>IK</b> KG-D <b>T</b> IE <b>PT</b> SGNT <b>GI</b> GLAL <b>TAA</b> I <b>KY</b> K <b>MII</b> T <b>LP</b> E	125
CBS SC	NP <b>GS</b> IKDRI <b>AK</b> SM <b>VEE</b> A <b>AS</b> GR <b>I</b> H <b>PSR</b> ST <b>LI</b> E <b>PT</b> SGNT <b>GI</b> GLAL <b>LIG</b> A <b>IKY</b> R <b>TII</b> T <b>LP</b> E	106
CBS TC	NP <b>GS</b> VKDR <b>IA</b> R <b>Q</b> M <b>ILDA</b> E <b>AS</b> GR <b>L</b> RP <b>N</b> -Q <b>T</b> IE <b>AT</b> SGNT <b>GI</b> GL <b>SL</b> V <b>AAV</b> K <b>Y</b> PM <b>II</b> T <b>MP</b> K	116
CS LT	NP <b>GS</b> VKDRIG <b>Q</b> M <b>VLD</b> A <b>EN</b> GT <b>LKPG</b> -S <b>V</b> VE <b>AT</b> SGNT <b>GI</b> GL <b>SMA</b> A <b>AI</b> R <b>GY</b> RM <b>VI</b> R <b>MP</b> K	105
CS SO	E <b>PC</b> SS <b>VK</b> DRIG <b>YS</b> M <b>IDD</b> A <b>EQ</b> K <b>V</b> IT <b>PG</b> K <b>T</b> TL <b>VE</b> PTSGNT <b>GI</b> GL <b>AF</b> I <b>AAAR</b> GY <b>KIT</b> L <b>TP</b> MA	161
CS AT	E <b>PC</b> SS <b>VK</b> DR <b>IA</b> Y <b>SM</b> IK <b>DA</b> E <b>DK</b> L <b>IT</b> PG <b>K</b> ST <b>LI</b> E <b>PT</b> AGNT <b>GI</b> GL <b>AC</b> M <b>GAAR</b> GY <b>KV</b> IL <b>V</b> MP <b>S</b>	101
CS CE	NP <b>AC</b> S <b>VK</b> DR <b>IA</b> F <b>N</b> M <b>IDT</b> A <b>E</b> K <b>AG</b> L <b>IT</b> PG <b>K</b> T <b>V</b> LI <b>E</b> PTSGN <b>MG</b> I <b>ALAY</b> CG <b>KLR</b> GY <b>KV</b> IL <b>TP</b> MA	100
CS ST	NP <b>S</b> F <b>S</b> V <b>K</b> CR <b>IG</b> AN <b>M</b> I <b>WDA</b> E <b>KR</b> G <b>V</b> L <b>KPG</b> -V <b>EL</b> VE <b>PT</b> NGNT <b>GI</b> AL <b>AY</b> V <b>AAAR</b> GY <b>KL</b> TL <b>TP</b> MA	93
	* * * * *	

CBS HS KMSSEKVDVLRALGAEIVRTPPTNARFDSPE SHVGVAVRLKNEIPNSHILDQYRNASNPLA 231

CBS OC KMSLEKVDVLRALGAEIVRTPPTNARFDSPE SHVGVAVRLKQEIIPNSHILDQYRNASNPLA 231

CBS RN KMSMEKVDVLRALGAEIVRTPPTNARFDSPE SHVGVAVRLKNEIPNSHILDQYRNASNPLA 228

CBS MM KMSMEKVDVLRALGAEIVRTPPTNARFDSPE SHVGVAVRLKNEIPNSHILDQYRNASNPLA 228

CBS FR RMSMEKVDVLRALGAEVHTPSSAPFDSPE SQVGMAWCLNKIPNSHILDQYRNASNPLA 235

CBS DM KMSNEKVSALRTLGAKIIRTPTEAAYDSEGLIYVAQQLRQETPNSIVLDQYRNAGNPLA 200

CBS DD KMSQEKVDVLKALGGEIIRTPTEAAFDAPESHIGVAKKLNSEIPNSHILDQYGNPSNPLA 185

CBS SC KMSNEKVSVLKALGAEIIRTPTEAAWDSPE SHIGVAKKLEKEIPGAVILDQYNNMNPPEA 166

CBS TC KMSHEKEVTLQALGAEVIRTE TALPWDHPDSLIGMARRLEKEK-GYVLLDQYRNPSNPKA 175

CS LT KMSHEKETTLQARGAEVIRTE TSLPNHPE SLIGVATPLRDEK-GYVLLDQYR-PQQP-D 162

CS SO SMSMERRVILKAFGAELVLTDPAKGMKG--AVEKAAEILKKT PDSY-MLQQFDNPANPKI 218

CS AT TMSLERRIILRALGAEHLSDQRIGLKG--MLEKTEAILSKT PGGY-IPQQFENPANPEI 158

CS CE SMSIERRCLLKAYGAEVILTDPATAVKG--AVQRAEELRDVIPNAY-ILNQFGNPANPEA 157

CS ST TMSIERRKLLKALGANLVLTEGAKGMKG--AIQKAAEIVASDPQKYLLQLQFNSNPANPEI 151

\*\* \* . \* . . . \* . . . \*

CBS HS HYDTTADEILQQCDG-----KLDMLVASVGTGGTITGIARKLK--EKCPGCRIIIGVDPEG 284

CBS OC HYDTTAEEILQQCDG-----KLDMLVASAGTGGTITGIARKLK--EKCPGQIIIGVDPEG 284

CBS RN HYDDTAEEILQQCDG-----KVDMLVASAGTGGTITGIARKLK--EKCPGCKIIIGVDPEG 281

CBS MM HYDDTAEEILQQCDG-----KLDMLVASAGTGGTITGIARKLK--EKCPGCKIIIGVDPEG 281

CBS FR HYDATAEEILEQCDG-----KLDMLVAGITGGTLTGVARLKL--EKCP--TIVAVDPEG 286

CBS DM HYDGTAAEILWQLDN-----KVDMIVVSAGTAGTISGIGRKIK--EQVPSQCQIVGVDPYG 253

CBS DD HYDGTAAEELLEQCEG-----KIDMIVCTAGTGGTITGIARKIK--ERLPNCIVVGVDPHG 238

CBS SC HYFGTGREIQRLLEDLNLFDNLRAVVAGAGTGGTISGISKYLK--EQNDKIQIVGADPF 224

CBS TC HYGTAQEIYDQCGG-----KVDMAVFTTGTGGTMAGVAKRLK--ELLPNIVIVGVDPYG 228

CS LT PLRVDRSEIYDQCGG-----KVDMVVVSAGTGGTITGAAKKQK--ELIPEIIVVGVDPIG 215

CS SO HYETTGP EIWEDTKG-----KVDIFVAGITGGTISGVGRYLK--ERNPGVQVIGIPE 271

CS AT HYRTTGPEIWRDSAG-----KVDILVAGVGTGGTATGVGKFLK--EQNKDIKVCVPEPE 211

CS CE HYKTTGPEIWRQTQGG-----KVDIVCFVGVSGGTCTGVGRFLK--EKNPSVQVFPVEPE 210

CS ST HEKTTGPEIWEDTDG-----QVDVFISGVGTGGTLTGVTTRYIKGTGKTDLITVAVEPTD 206

\* . . . . \* . . . \*

CBS HS S-ILAPEEELNQTTEQTYEVEGIGYDFIPTVLDRTVVDKWFKSNDEEAF FARMLIAQEG 343

CBS OC S-ILAPEEELNQTTEVTA YEVEGIGYDFIPTVLDRTVVDKWFKSTDKAEAFARMLIAQEG 343

CBS RN S-ILAPEEELNQTTEQTA YEVEGIGYDFIPTVLDRAVVDKWFKSNDDDSFAFARMLISQEG 340

CBS MM S-ILAPEEELNQTTEQTA YEVEGIGYDFIPTVLDRAVVDKWFKSNDEDSFAFARMLIAQEG 340

CBS FR S-LLIHSDDHCK---SSYEVEGIGYDFVPTVLDKSLVDYFYKVNDKETFNMSRRLIRDEG 342

CBS DM S-ILARPAELNKT DVQFYEVEGIGYDFPPTVFDVVDVWTKIGSDCFPMSRRLNAEEG 312

CBS DD S-ILAQPESLNNTN-KSYKIEGIGYDFIPNVLERKLV DQWKTTDDKESFIMARRLIKEEG 296

CBS SC S-ILAQPENLNKT DITDYKVEGIGYDFVQVLDRLKIDVWYKTDDKPSFKYARQLISNEG 283

CBS TC S-ILADP-SVPLDPKPY-LVEGIGYDFVDPV CERKYVDRWVK SADKESFALASEVHREEA 285

CS LT S-VVADP-EHPCEPVMYQEV EGIYDFVPAVCERKYVDRWVKTRDQQSFDLARELHDEG 273

CS SO SNILSGG-----KPGPHKIQGLGAGFVPSNLDLGVMDEVIEVSSEEAVEMAKQLAMKEG 325

CS AT SPVLSGG-----QPGPHLIQGI GSGIVPFNLDLTIVDEIIQVAGEEAIETAKLLALKEG 265

CS CE SSVINGL-----PHSPHKIQMGMTGMI PDILDLTLFSEALRVHSDDATAMAKKLAD EES 264

CS ST SPVIAQALAGEEIKPGPHKIQGI GAGFIPGNLDLKLIDKVVGITNEEAISTARRLMEEEG 266

\* . . . . \* . . . \*

382



CBS HS LLCGGSAGSTVAVAVKAAQ---ELQEGQRCVVILPDSVRNMYTKFSLDRWMLQKGF LKEE 400

CBS OC LLCGGSAGS AVAVAVKAAQ---ELQEGQRCVVILPDSVRNMYSKFSLDRWMLQKGF LLEE 400

CBS RN LLCGSSSGSAMAVAVKAAQ---ELKEGQRCVVILPDSVRNMYSKFSLDKWMLQKGF MK-E 396

CBS MM LLCGSSSGSAMAVAVKAA R---ELQEGQRCVVILPDSVRNMYSKFSLDKWMLQKGF MK-E 396

CBS FR LLCGSSSGSAMAAAVKVAQ---HLKEGQRCVVILADSVRNMYSKFSLDKWMSEKGF L VPE 399

CBS DM LLCGSSSGGAMHAALEHAR---KLKKGQRCVVILPDGIRNMYTKFVSDNWM EARNFKE-- 367

CBS DD LLCAGSSSGSAMVGALLAAK---QLKKGQRCVVLLADSI RNMYTKHLNDDWLVDNGF V DPE 353

CBS SC VLVGSSSGSAFTAVVKYCEDHPELTEDDVIVAIFPDSIRS YLTKFVDD EWLKKNLWDDD 343

CBS TC LLVGSSSGAAMWV LQAAK---DLRPDQRCVVVFPDGI RNMYTKFPDKNWLVENGLEEGE 342

CS LT LLVGSSSGAAMAGVLEAAK---DLRADQRCVVLMADGIRNMYMGKFADTNWMI EHGFRQGE 330

CS SO LLVGISSGAAAAA AVRIKGR--PENAGKLI AVVFP SFGERY LSSILFQS IREECENMKPE 383

CS AT LLVGISSGAAAAA ALKVAKR--PENAGKLI VVFP SGGERY LSTKLFDS IRYEANLPIE 323

CS CE ILGGISSGANVCAAVQLAKR--PENKGLIVTTVNSFGERY LSTALYAE LRDNAANMKQL 322

CS ST ILAGISSGA AVAAALKLOED--ESFTNKNI VVILPSSGERY LSTALFADLFT EKELQQ-- 322

\* \* . \* . . . \*

CBS HS	DLTEKKP-----WWWHLRVQELGLSAPLTVLPT-ITCGHTIEILREKGFQAP	447
CBS OC	ELSVKRP-----WWWHLRVQELSLSVPLTVLPG-VTCSDTIDILRGKGFQAP	447
CBS RN	ELSVKRP-----WWWHLRVQELSLSAPLTVLPT-VTCEHTIAILREKGFQAP	443
CBS MM	ELSVKRP-----WWWRLRVQELSLSAPLTVLPT-VTCEDTIAILREKGFQAP	443
CBS FR	APVEPKP-----WWWGTTVNCLEHLSPPFSTLPS-VSCQEAIIEILKERAIHLP	446
CBS DM	PVNEHGH-----WWWSLAIAELELPAPPVILKSDATVGEAIALMKKHRVDQLP	415
CBS DD	YTKDQQE-----EEKYHGATVKDLTLPKPITISAT-TTCAAQVLLQQYGFQQLP	403
CBS SC	VLARFDSSKLEASTTKYADVFGNATVKDLHLKPVVSVKET-AKVTDVIKILKDNQFDQLP	402
CBS TC	VTRP-----	346
CS LT	VTRP-----	334
CS SO	-----	
CS AT	-----	
CS CE	NLDDSIKIAKEYLGI-----	337
CS ST	-----	

CBS HS	VVD-EAGVILGMVTLGNMLSSLLAGKVQPS-----DQVGKVIY----K	485
CBS OC	VVD-ETGEILGMVTLGNMLSSLLAGKVQPS-----DQVCKVLY----K	485
CBS RN	VVN-ESGAILGMVTLGNMLSSLLAGKVRPS-----DEVCKVLY----K	481
CBS MM	VVN-ESGAILGMVTLGNMLSSLLAGKVRPS-----DEVCKVLY----K	481
CBS FR	VVD-DSGLAQSVVTTDTILSSVSSGRVKLS-----DPI SKVVC-----K	484
CBS DM	VVDQDDGSVLGVVQETLITQIVSMNRQQS-----DPAIKALN----K	454
CBS DD	VVS-ESKKVLVNSLLVTSLT-YASKKAVPT-----DAVSKVMFRFTKNE	445
CBS SC	VLT-EDGKLSGLVTLSELLRKL SINNSNDNTIKGKYLDFKKLNNFNDVSSYENKSGKK	461
CBS TC	-----TYGSLAQLEEAKKLS-----EYEAKCGG-----	371
CS LT	-----TYDALKKELEGVAKLA-----QYESAAK-----	358
CS SO	-----	
CS AT	-----	
CS CE	-----	
CS ST	-----	

CBS HS	QFKQIRLTDTLGRLSHILEMDHFALVVHEQIQYH-----STGKSSQRQMVF	531
CBS OC	QFKQIRLTDTLGALSHILEMDHFALVVHEQIQYG-----GDEQPSKRQTVF	531
CBS RN	QFKPIHLTDTLGMLSHILEMDHFALVVHEQIQR-----NNGVSSKQLMVF	526
CBS MM	QFKPIHLTDTLGTLSHILEMDHFALVVHEQIQSRDQAWSGVVGPTDCSNGMSSKQQMVF	541
CBS FR	PFKEVRMSDSLGLLSQILETDRFALVVQEPQQNK-----SKGSAHQEKML	529
CBS DM	RVIRLNESEILGKLARVLEVDPVSLILG-----KNPAGKVELK	492
CBS DD	KYIPITQSTSLATLSKFFENHSSAIVTE-----NDEII	478
CBS SC	KFIKFDENSKLSDLNRFEEKNSSAVITDG-----LKPI	494
CBS TC	---SVEEGKRCEKAK-----	384
CS LT	-----	
CS SO	-----	
CS AT	-----	
CS CE	-----	
CS ST	-----	

CBS HS	GVVTAIDLLNFVAAQ-----ERDQK-----	551
CBS OC	GVVTAMDLLHFVASR-----GQDQQ-----	551
CBS RN	GVVTAIDLLNFVAAR-----EQTRK-----	546
CBS MM	GVVTAIDLLNFVAAR-----EQTQT-----	561
CBS FR	SVVTAMELLSYITAKREQVVLPASTSQTQLHHPPFRGLRMLQGDLLKKPTAFYLTV	585
CBS DM	ALATKLDVTFIAAG-----KQPKANGTTNGGSH-----	522
CBS DD	SIVTKIDLLTYLMKS-----QQKN-----	497
CBS SC	HIVTKMDLLSYLA-----	507
CBS TC	-----	
CS LT	-----	
CS SO	-----	
CS AT	-----	
CS CE	-----	
CS ST	-----	

**Sequences:**

CBS HS, Homo sapiens, (Swiss-Prot accession no. P35520)  
CBS OC, Oryctolagus cuniculus, (Swiss-Prot accession no. Q9N0V7)  
CBS RN, Rattus norvegicus, splice isoform III, (Swiss-Prot accession no. P32232)  
CBS MM, Mus musculus, (Swiss-Prot accession no. Q91WT9)  
CBS FR, Fugu rubripes, (Swiss-Prot accession no. Q9YHU3)  
CBS DM, Drosophila melanogaster, (Swiss-Prot accession no. Q9VRD9)  
CBS DD, Dictyostelium discoideum, (Swiss-Prot accession no. P46794)  
CBS SC, Saccharomyces cerevisiae (Swiss-Prot accession no. P32582)  
CBS TC, Trypanosoma cruzi (Swiss-Prot accession no. Q9BH24)  
CBS LT, Leishmania tarentolae, (Swiss-Prot accession no. Q9NG81)  
CS SO, Spinacia oleracea, (Swiss-Prot accession no. Q00834)  
CS AT, Arabidopsis thaliana, (Swiss-Prot accession no. P47998)  
CS CE, Caenorhabditis elegans, (Swiss-Prot accession no. O45679)  
CBS ST, Salmonella typhimurium, (Swiss-Prot accession no. P12674)

*The sequences were aligned using the CLUSTALW 1.82 multiple sequence alignment software (<http://www2.ebi.ac.uk/clustalw>) using clustalw\_mp CPU mode at the default settings.*