

*S. acutus* (ACB06751): pI=3.95, no putative glycosylation site

MKLFAAVVLVTCALASVQGRTLQQAAAAPTTVAQALAQTPLSSTLNAAVQAAGIDIPADAAW**TIF**  
**APKNEAF**SDDDIREETGLTAQQLLQOPENQALTQLQYHIVPSGALRAAQQLQDGQQLTTALAGAA  
PLKVDIEDEKVEIEVPDAGNDDGDDADVEQADMTVG**NSVIHVIDEVL**IPASLRGSRQPARG

*C. vacuolata* (CvacAstaP1 MZ063685): pI=6.12, 10 putative glycosylation sites

MRTLALFVLACTIGAAYGRELLQAAAPAASYNTLAQAVAAANASAPNLSILLAAVKAANVS GALGP  
NTT**WTILAPTNQAF**TSRLNESLGITPQQLLQPNKCTLVEVLSYHIVPSGAVYSSNLTNGQEVT  
ALAGADPLTVAINRTVTFEGAENNATVTVANIRAG**RSVIHVINDVL**LPAGVG NSSNTTAEGEST  
APGAAGSRDVAAGSAVGGGSTSAPAPAAKGAAGAAQVPAVLLASMLLIGMLF

*C. vacuolata* (CvacAstaP2 MZ063686): pI=10.3, 2 putative glycosylation sites

MARNAIVFLAIVCLVAGATAADKADPAKKKSAPSYPPLSDAVKAAATAPQLTLLAIAAGVASA  
LKPDTAW**TILAPTNKAF**ETRLAKLNVTADALLKNKDLLVKVLSYHIVPSGAVYSSQLDKQTVAT  
ALKNATLTVKIYGKNAFKGPVNKAKVVAADIKAG**KSVVHVIDDVL**VPPGVVSPEVAQWKAEYK  
SKKDASMKKMDKP

*Crl. variabilis* (EFN56922): pI=4.15, 1 putative glycosylation site

MARHSAQSYDNIGELLDTPKLSILKAALEAAGLDELASDDALEA**TLFAPTDKAF**ADAI TALNLT  
AEEVLGDKDLTTVLSYHIVPGAVEAAALEEGQILQTILEGAGQLKYYTFKGLPRLRTTADQV  
ARITETDISAG**KAIVHLIDAVL**IPGDEFFSAPAPGPSP

*Cmy. reinhardtii* (XP001700577): pI=6.33, 15 putative glycosylation sites

MARWTRPGASTILVASCLLLAVAGSSAQTILOAIQSA**NLT**LLAAGLQVADLATTLNSTGPNITLF  
APSDDALLAMTAALNTTPNALLTLPKVAPVFTYHAITPLLAADIPTGDPYATLNTGL**NLT**VS  
AAANGSVVVRALGSDANVIRTDITAGGAVVHIDSVLLPFLSASAATRTPQLSSLLS  
PALSNPNLNV**TVFAPVNDAF**NASSAYLTSVNASITDVLTYHVATSRLV**NGSE**VTASPILT  
ANLT**VARDGNVVVT**PVGATAATVLAANI**PG**LDLTGGP**RTFVHLVNAVLL**PFYTTVANAERAG  
LTTLVAAVAASDPAFLAAVTDPTFRG**TILAPSNAAFA**ATLASL**NVT**AAQLLADKDNRRILNAHII  
PNAAVFSNQLSN**NQTVNTTGALLTV**SIGNGIVRFNAAKSSAKVVAADVSIGNGRAVVHVIDRVLL  
PADLTLPTPYVTVAGAAQAAANLSTLLAAUTASDASFLTTLTDPDFNGTVLAPTPLSDCHVALLMPN  
AAFTAALTALGV**NAT**QLLADKDNRRILNAHII**TPGAVRSSQLVNNQNIT**TLGVLTVR**NGT**GVF  
FVAAKSTAKVVTVDVIAG**NAVVHTIDFVL**LPASVTLTTGGSGGGGGGGAAAGMASPSVLTLLSSA  
LAFLALFAFGRF

**Figure S1.** The deduced amino acid sequences of the AstaP orthologs. N-terminal signal peptides are highlighted by yellow. H1 and H2 motifs of the fasciclin domains are shown in boldface italics. Potential sites for N-linked glycosylation are shown in red fonts.

CvacAstaP1	-----MRTLALFVLACTIG--AAYGRELQAAAPA	27
CvacAstaP2	MGRPCLLTYTQTQSYLIDFVVTPTSQQSAMARNAIVFLAIVCLVAGATADKADPAKKSA	60
	. : : : : . * : : . : * : : : : . *	
CvacAstaP1	ASYNTLAQAVAAAANASAPNLSILLAAVKAAANVSGALGPNTTWILAPTNQAFTSRLNESL	87
CvacAstaP2	PSYPSLSDAVKAA-ATAPQLTTLLAAINAAGVASALKPDTAWTILAPTNKAFETRLAK-L	118
	. ** : * : : ** * * : * : * : * : *** : : * . * : * : : * : * : * :	
CvacAstaP1	GITPQQLLQPANKKTLVEVLSYHVIPSGAVYSSNLNTNGQEVPPTALAGADPLTVAINNRTV	147
CvacAstaP2	NVTADALLK--NKDLLVKVLSYHVIPSGAVYSSQLDKQTVATALKNAT-LTVKIYGKNV	175
	. : * . : * : * . * : * : * : * : * : * . : * * . * . * * * * . : * . : *	
CvacAstaP1	TFEGAENNATVTVANIRAGRSVIHVINDVLLPAGVGNSSNTTAEGESTAPGAAGSRDVA	207
CvacAstaP2	AFKGPNKAKVVVADIKAAGKSVVHVIDDVLVPPGV-----VSPEVAQKWKA	223
	: * : * . * : * . * : * : * : * : * : * : * : * . * : * . * . . .	
CvacAstaP1	GSAVGGGSTSAPAPAAKGAAGAAQVPAVLLASMLLIGMLF	248
CvacAstaP2	KSKKDASMKKMDKP-----	237
	* . . . . *	

**Figure S2.** Comparison of amino acid sequences among AstaP-proteins in *C. vacuolata*. Amino acid sequence similarity between CvacAstaP1 and CvacAstaP2 showed 44% identity in overlapping regions.