

Serpin**Multiple Alignments of R C L Sequences**

P17-----P8-----P1/P1-----P14'PF

<i>Spn27A</i> (CG11331)	mel sim sec yak ere ana pse per wil moj vir gri	EKGTEAYAATVVEIENKFGGSTAIEEFNVNRPF EKGTEAYAATVVEIENKFGGSTAIEEFNVNRPF EKGTEAYAATVVEIENKFGGSTAIEEFNVNRPF EKGTEAYAATVVEIENKFGGSTTIEEFNVNRPF EKGTEAYAATVVEIENKFGGSTLIEEFNVNRPF EKGTEAYAATVVEIENKFGGSTIIIEEFNVNRPF EKGTEAYAATVVEIENKFGGSTTIEEFNVNRPF EKGTEAYAATVVEIENKFGGTTIEEFNVNRPF EKGTEAYAATVVEIENKFGGSIVIEEFNVNRPF EQGTEAYAATVVEIENKFGGSQTIIEEFNVNRPF EQGTEAYAATVVEIENKFGSSQTIEEFNVNRPF EQGTEAYAATVVEIENKFGGPTKIEEFNVNRPF * : * ***** * * * * . * * * * * *
<i>Spn28B</i> (CG6717)	mel sim sec yak ere	EKGGEASAATGVLTRRKKSIDNLIQPPMEFIADHPF EKGGEASAATGILTRRKKSIDNLRLLPPMEFIVDHPF EKGGEASAATGVLTRRKKSIDNLRLLPPMEFIVDHPF ERGTKASAATGVLIRRKKSIDNLKPPPMEFVADHPF ERGAKASAATGVLIRPKKSIDNMKLPPMEFIADHPF * : * : * * * * ; * * * * * : * * * * : . * * * *
<i>Spn28Da</i> (CG31902)	mel sim sec yak ere	EEGSGAGSASASPIRGLSDYATSVVTFTVNSPF EEGSGAGSASANPIRGLSDYAAASLVTFVNSPF EEGSGAGSASANPIQGLSDYAAASVVTFTVNSPF EEGASAGSASASLFRGLSDYPASVASFTVNSPF EECASAGSASASLLRGWSDYPANVASVRVDRPF ***.*****. : : * ***.:.:.:. * : **
<i>Spn28Db</i> (CG33121)	mel sim sec yak ere	HNKKVYVRMMSHVGRFRIADHSYQIIEMPF HNRKAYVRMMSHVGRFRIADHSYQIIELPF HNREAYVRMMSHVGRFRIADHSYQIIELPF HNKKVYVRMMSHVGRFRMADQSYQIIELPF HNKAHVVRMMSQVGRFRMAKQSYQIIELPF ** : : : * * * * : * * * * : * * * * : **
<i>Spn28Dc</i> (CG7219)	mel sim sec yak ere ana pse per wil moj vir gri	EQGTEAAASSVTYLKKSGPDVLFRGDTPF EQGTEAAASSVTYLKKSGPDVLFRGDTPF EQGTEAAASSVTYLKKSGPDVLFRGDTPF EQGTEAAASSVTYLKKSGPDVVFRGDTPF EQGTEAAASSVTYLKKSGPDVVFRGDTPF EQGTEAAAATVTVKKSGPDVLFRGDTPF EQGTEAAAATVTVKKSGPDVLFRGDTPF EQGTEAAAATVTVKKSGPDVLFRGDTPF EQGTEAAAATIAYLKKSGPDILFRVDTPF EQGTEAAAATAILKKSGPEVLFRAETPF EQGTEAAAASAILKKSGPEVVFRAETPF EQGTEAAAASAVLKKSGPDVVFRAETPF *****: : : * * * * : * : * * * *
<i>Spn28F</i> (CG8137)	mel sim yak ere ana (16627) ana (21984)	EEGAEEAAAATALLFVRYSMP-MPSSQMVFVNADHPF EEGAEEAAAATALLTVRYSKP-RPSSQMVFVNADHPF EEGAEEAAAATKIVVYPLS---AHPIWMDFNVDHPF EEGAEEAAAATKIDMVFKSAASRPQPI--DFIADHPF EEGAEEAAAATAVVAVHICANCATPSRMSFIANHPF EEGAEEASAASAVSIVFYE-----MDFHADHPF *****: : : * . : * * *

<i>Spn31A</i> (CG4804)	mel sim sec yak ere ana pse per wil moj vir gri	ESGSGSGPELPKNATEYKPIVISNSSRQKFFRADH PF ESGSGP--ELPKNATEYKPIVISNSSRQKFFRADH PF ESGSGP--ELPKNATEYKPIVISNSSRQKFFRADH PF ESGSGP--ELPKNATEYKPIVISNSSRQKFFRADH PF ESGSGP--EPRVAAEYKPIVISNSSRQKFFRADH PF ESGSGS--EPPRAAEYKPIVISNSSRQKFFRADH PF ESGSGS--EPPRAAEYKPIVISNSSRQKFFRADH PF ESGSGS--ESPQPAAYKPIVISNSSRQKFFRADH PF ESGSGS--ELPPPATGYKPIVISNSSRQKFFRADH PF ESGSGS--ELPQPAAGYKPIVISNSSRQKFFRADH PF ESGSGS--ESPQPAAGYKPIVISNSSRQKFFRADH PF ***** . * : *****:*****:*****:*****:*****:
<i>Spn38F</i> (CG9334)	mel	EEGSEAAAATAVVFYRK S IRSP-PMDFNVNH PF
<i>Spn42Da</i> (CG9453)	mel sim sec yak ere ana pse per wil moj vir gri	EEGTEAAAATGMAVRRKRAIMSPEEPIEFFADH PF EEGTEAAAATGMAVRRKRAIMSLEEPIEFLADH PF EEGTEAAAATGMAVRRKRAIMSLEEPIEFLADH PF EEGTEAAAATGMVVRKKRAIVSLEEPLEFLADH PF EEGTEAAAATGMVARTKRAIYSLEEPIDFFADH PF EEGTEAAAATGFRIMPVAGFRR----KHTVNR PF EEGTEAAAATGAVVRMKRSIVSLTEPIEFHADH PF EEGTEAAAATGAVVRMKRSIVSLAEPIEFHADH PF EQGTEAAAATGFVVRKKRAIVSLTEPVEFFADH PF LGTEAAAATAAVVRMKRSVISIEQPIQFHADH PF LGTEAAAATVFRIMPVLAFRR----KKFLATH PF EHGTEAAAATAMVMCFASMPMFQPEPIRFHAEH PF * ***** . * . :**
<i>Spn42Db</i> (CG9454)	mel sim sec yak ere ana pse per wil moj vir gri	EVGTEAAAATGEPACKK S -KTFGNLKASYI---- EVGTEAAAATAAVATFRSM-PAREVPPKVFHANR PF EVGTEAAAATAAVATFRSM-PARESPPKVFHANR PF EVGT K AAAATVTTPMVFYL-PVILSLQIENS---- EVGTEAAAATAVVATFRSM-PPPQGSPKVFHANR PF ELGTEAAAATAVVLMVRSL-PAP-EPHQVFIANR PF EVGTEAAAATAAVMVMRSLPATPVDRPKAFHANR PF EVGTEAAAATAAVMVMRSLPATPVDRPKAFHANR PF ELGSEAAAATGKFLNM K GYDCDG----- EGTEAAAATAMLVYF E MMPTED-EPKIFHADH PF EEGTEAAAATAAVMMMRSAAPAPESMPVEFHAN-H PF EEGTEAAAATALVVKM R SALEPKIQIFHAN--H PF *:***** :
<i>Spn42Dc</i> (CG9455)	mel sim sec yak ere ana pse per wil moj vir gri	EAGSEAAA V FSFMKIVPMMLNMNKLFKADH PF EAGSEAAA V FSFMKIVPMMLNMNKLFKADH PF EAGSEAAA V FSFMKIVPMMLNMNKLFKADH PF EAGSEAAA V FSFMKIVPMMLNMNMTKHFKA DH EAGSEAAA V FSFMKIVPMMLNMNKLFKANH PF EAGSEAAA V FSFMKIVPMMLNMNKQFKADH PF EAGSEAAA V FSFMKIVPMSLNMSKKL F KVDH PF EAGSEAAA V FSFMKIVPMSLNMSKKL F KVDH PF EAGSEAAA V FSFMKIVPMMLNMNKKSF K VDH PF EACCEAAA T YLKIVPM S MIMHQKF K VDH PF EAGSEAAA V FLKVVPM S LNHMKKTF K VDH PF EAGCEAAA V LKVVPM S LNMNKET F KVDH PF ***.****.: :*: * : * : **. :***
<i>Spn42Dd</i> (CG9456)	mel sim sec yak ere ana pse per	EEGA E AAGATSVAVTN R AGFS----TFLMADH PF EEGA E AAGAT-VAVTN R AGFS----TFLMADH PF EEGA E AAGATSVAVTN R AGFS----TFLMADH PF EEGA E AAA A ATKIVVYPLSAHP---IWMDFNVDH PF EEGA E AAA A ATSVAVTN R AGFS----MFLAADH PF EEGA E AAA A ATVAMVSPRS G FS----QS F VADH PF EEGA E ASA A ATYG M ITNR S SFT----MVLSD D H PF EEGV E ASA A ATCTCDLLIRPDYGS M IILTYINVY I ***.***.*** : . : . : . : * : *

<i>Spn42De</i> (CG9460)	mel sim sec yak ere ana pse per wil	EKGTTASGATFIVKSVVESLIGEEVFEFIADHPF EKGTTASGATFVKVSVESLLIGEEVFEFIADHPF EKGTTASGATFVKVSVESLLIGEEVFEFIAYHPF EKGT SASGAT YVNAAVESLLIGEQFEGTADHPF EKGTTASGATFVQVAEVESLLIGEEVFEFSADHPF EKGTTASGATFVKAELLESLVIGEQTVEFVADHPF ENGTTASA AT AVKFSLESAFMG-EVQQFTADHPF ENGTTASA AT AVKFSLESAFMG-EVQQFTADHPF EKGTTATGASFSKVVIELLEVIGEVTYNFTVDHPF *:***:***: : * : * : * . : * . ***
<i>Spn43Aa</i> (CG12172)	mel sim sec yak ere ana pse per wil moj vir gri	EIGCEAAGASYAAGVPMSLPLDPKTFVADHPF EIGCEAAGASYAAGVPMSLPLDPKTFVADHPF EIGCEAAGASYAAGVPMSLPLDPKTFVADHPF EIGCEAAGASYAAGVPMSLPLDPKTFVADHPF EIGCEAAGASYAAGVPMSLPLDPKTFVADHPF EIGCEAAGVSYAAGVPMSLPLDPKTFVADHPF EIGCEAAGVSYAAGVPMSLPLDPKTFVADHPF EIGCEAAGVSYAAGVPMSLPLDPKTFVADHPF EIGCEAAGVSYAAGVPMSLPLDPKTFVADHPF EIGCEAAGVSYAAGVPMSLPLDPKTFVADHPF EIGCEAAGVSYAAGVPMSLPLEPKTFVADHPF *****:*****:*****:*****:*****
<i>Spn43Ab</i> (1865)	mel sim sec yak ere ana pse per wil moj vir gri	EAGVDQPLETGLLKGLFSRSKKFEADHPF EAGVEQPLETGLLKGLFSRSKKFEADHPF EAGVEQPLETGFLKGLFSRSKKFEADHPF EAGLEQPLETGLLKGLFSRSKKFEADHPF EAGLEQPLETGLLKGLFSRSKKFEADHPF EAGVEQPLETGVLKGLFSRSKKFEADHPF EAGVEQPLESGVLKGLFSRSKKFEADHPF EAGVEQPLESGVLKGLFSRSKKFEADHPF EAGVEPSLHQGVLKGLFSRSKKFEADHPF EAGVEQPLETGVLKGLFSRSKKFEADHPF EAGVEQPLESGLVLSKFFSRTKKFEADHPF EAGVEQPLESGLLKGGLFSRNKKFEADHPF *****:*****:*****:*****:*****
<i>Spn43Ac</i> (CG1857)	mel sim sec yak ere ana pse per wil moj vir gri	EAGTEASAASYAKFVPLSLPPKPTEFVANRPF EAGTEASAASYAKFVPLSLPPKPTEFVANHPF EAGTEATAASYAKFVPLSLPPKPTEFVANRPF EAGTEASAASYAKFVPLSLPPKPKEFIANRPF EAGTEAAAASYAKFVPLSLPAKSPEFTADHPF EAGSEAAAASYAKFVPLSLPVKSREFNADHPF EAGSEAAAASYAKFVPLSLPVKSREFNADHPF EAGSEASASSYAKFVPLSLPVKSHEFTADHPF EAGSEAAAATFAKFVPLLILPMKSREFIADHPF EAGSEAAAASYAKFVPLSLPVKSREFVADHPF EAGSEAAAASYAKFVPLSLPVQSLEFTADHPF *****:*****:*****:*****:*****
<i>Spn43Ad</i> (CG1859)	mel sim sec yak ere ana pse per moj vir gri	EDGGNADDDFS--FGDLFRRALPLVINHPF EEGGNADDSSF--FGDLFRRALPLVINHPF EEGGNADDSSF--FGDLFRRALPLVINHPF EEGGNAEDSFS--FGDLFRRALPLVINHPF EEGGHADDSSF--FGDLFRRALPLVINHPF EKGGTGDSFS--FGDLFRRALPLVINHPF EEGGSAGDSFS--FGDLFRRALPLVINHPF EEGGSAGDSFS--FGDLFRRALPLVINHPF EAGGEAEQTFFVAFTAIDLFRSTLSQLVINHPF EAGGEAEQTFFVAFTAIDLFRSTLSQLVINHPF EQGAEAEQTFFATDIDLFRSTLSQLVINHPF * . : * * : * : * : * : * : *****

<i>Spn47C</i> (CG7722)	mel sim sec yak ere ana pse per wil moj vir gri	EFGCEVAPEAEVQPEVLKK-NPDRKFFKADR PF EFGCEVAPEAEVQPEVLKK-NPDRKFFKADR PF EFGCEVAPEAEVQPEVLKK-NPDRKFFKADR PF EFGCEVAPESDVQPEVLKK-NPDRKVFKADR PF EFGCEVAPEPDTHIHKTKN-NPDRKIFRADR PF ESGCETDPDKPARAAAFVQ-NPDRKLFMANR PF ESGCETDPDTPSRAAAFVQ-NPDRKLFTMNR PF ENGCEVDIENKPGPTVPIVPDAERKIFRANR PF EAGCTTNIDESTKNRFVKA-NPERKVFVADH PF EAGCEIDANTPAETSTTS-NPERKVFVTADH PF ESGCGTADHTQGRVAAAKV-NPERKVFVKANQ PF * * * . : * * . * : : : * * *
<i>Spn53F</i> (CG10956)	mel sim sec yak ere ana	EATYPREFRVNATKSVMIPMMHEDSKFAFGILGNLKATAVLV PF EATYPREFRVNAARSVMI PMMHEDRSP----- EATYPREFRVNAAKSVMIPMMHEDSKFAFGILGNLKATAILV PF EATYPREFRVSAARKVMIPMMHEDSKFAFGNLATLKASAVL PF EATYPQRFRVSAAKEVLIPMMHEDSKFAFGTLGHLKATAVLV PF GATYDREFRVSSQRTINVPMMHEDSKFAFGDLEKLQATALLLP PF * * * . * * * : * * * . : : : * * * * * : .
<i>Spn55B</i> (CG10913)	mel sim sec yak ere ana pse per wil moj vir gri	EEGTEAAAATGMIMMTTRMMTFPLQFQADR PF EEGTEAAAATGMIMMTTRMMTFPLQFQADR PF EEGTEAAAATGMIMMTTRMMTFPLQFQADR PF EEGTEAAAATGMIMMTTRMMTFPLQFQADR PF EEGTEAAAATGMIMMTTRMMTFPLQFQADR PF EEGTEAAAATGMIMMTTRMMTFPLQFQADR PF EEGTEAAAATGMIMMTTRMMTFPLQFQADR PF EEGTEAAAATGMIMMTTRMMTFPLQFQADR PF EEGTEAAAATGMIMMTTRMMTFPLQFQADR PF EEGTEAAAATGIIMMTTRMMMLPLQFQADR PF EEGTEAAAATGMIMMTTRMMMLPLQFQADR PF * * * * * : * * * * : * : * * * * :
<i>Spn75F</i> (CG32203)	mel sim sec yak ere	ETALTNRTFTLLRQNKKPFVYTTQMMYTEAPMDFFNNNDQVRGV V EKALTNRTFTLLRQNRRKPFVYTVEMMYTEAPMEFFNNNDQVRGV V EKALTNRTFTLLRQNRRKPFVYTVEMMYTEAPMFQFFNNNDQVRGV V EKAMTNRLFSYIRPNRKPVYKVEMIYTEAPMFQFFNEDQVRGV V EKAMTNRTFRWIRPNRRTFVYRVQMIFYIEAPMFSGDQCRCGV V *: * * * * : * * : . * * : * : * * * * * . * * * * :
<i>Spn76A</i> (CG3801)	mel sim sec yak	FKDSAFKSKAKIKINNFRVNHGIRFQPILRLEVDDIDTGKTET F FKDSAFKSKAKIKINDFRVNHGIRFQPILRLDVVDDIDTGKTE K FKDSAFKSKAKIKINDFRVNHGIRFQPILRLDLVDDIDTEKE K FKDSGFSKPKNIKINDFRVNHGQVQFEPTQRLNVVEDIDTQNA Q * * * . * : * . * * * * * . * : * * * * * . * : * * * :
<i>Spn77Ba</i> (CG6680)	mel sim sec yak ere ana pse per wil moj vir gri	EQGTTAGAVTEAALANKATPPKFLLNR PF EQGTTAGAVTEAALANKATPPKFLLNR PF EQGTTAGAVTEAALANKATPPKFLLNR PF EQGTTAGAVTEAALSNKATPPKFQLNR PF EQGTTAGAVTEAALANKATPPKFQLNR PF EQGTTAGAVTAALINKATPPKFQLNK PF EQGTTAGAVTAALINKATPPKFQLNK PF EQGTTAGAVTAALINKATPPKFQLNK PF EKGTTAAAVTSASLVNKATPPKFVLNK PF EKGTTAAAVTASVLSNKSTPPKFQLNR PF EKGTTAAAVTASVLSNKSTPPKFQLNR PF EQGTTAAGTVVSSLNSKSTPPKFHLNR PF *: * * * . * * : * * : * * * * * : * :
<i>Spn77Bb</i> (CG6663)	mel sim wil	VDEEGLPNAVPQKSSGKN---NIKFHVNR PF VDEEGLPDAVPQKSSGTN---NIKFHMTR PF VDEKGTAACAFSAATLSNKAAPPKFMIN KPF * * * . * . * . * . * . * . * . * .

<i>Spn77Bc</i> (CG6289)	mel sim sec yak	EQFYNAGGSPAGKVEMMVQTGKYAYVNNVKGLQADVLELPF EQFYDSSGKPKGKVNMMPQTGKFAYIRS IKRL DADFLMLPF EQFYDSSGKPKGVDMMVQTGKFAYIRS IRRL DADFLMLPF EQFYNDGSPAGKVHMMVQTGKFAYAKNVKDLQADVLELPF *****: .*.*****.*****: ** ..:: *:***.* ***
<i>Spn85F</i> (CG12807)	mel sim sec yak ere ana pse per wil moj vir gri	YHTDIVAAAASLKLGP TLRLMRKQLKPR ----- RTAASLKLGP TLRLMRKQLKPR YHTDIVAAAASLKLGP TLRLMRKQLKPR YHTDIVAAAASLKLGP TLRLMRKQLKPR YHTDIVAAAASLKLGP TLRLMRKQLKPR YHTDIVAAAASLKLGP TLRLMRKQLKPR YHTDIVAAAASLKLGP TLRLMRKQLKPR YHTDIVAAAASLKLGP TLRLMRKQLKPR YHMDIVAAAASLKLGP TLRLMRKQLKPR YHMDIVAAAASLKLGP TLRLMRKQLKPR YHMDIVAAAASLKLGP TLRLMRKQLKPR ** :***:*****:*****:*****
<i>Spn88Ea</i> (CG18525)	mel sim sec yak ere ana pse per wil moj vir gri	EEGSTAAAATVLFTYRSARPVEPAKFECNHPF EEGSTAAAATVLFTYRSARPVEPAKFECNHPF EEGSTAAAATVLFTYRSARPVEPAKFECNHPF EEGSTAAAATVLFTYRSARPVEPAKFECNHPF EEGSTASAATVLFTYRSARPVEPAKFECNHPF EEGSTAAAATVLVSYRSARPIEPTKFE CNH PF EEGSTAAAATVLVSYRSARPIEPTKFE CNH PF EEGSTAAAATVLVSFRSARPIEPTKFECKH PF EEGSTAAAATVLVSFRSARPVEPTKFE CNH PF EEGSIAAAATVLVSFRSARPIEPTKFE CNH PF EEGSTAAAATVLVSFRSARPAEPTKFE CNH PF ***** *:*****. :***** *:***** :***
<i>Spn88Eb</i> (CG6687)	mel sim sec yak ere	EVGSTAAAATILLVSRSSRQPDP TKFNCNHPF EVGSTAAAATILFVSR SARQPDP TKFNCNHPF EVGSTAAAATILFVSR SARQPDP TKFNCNHPF EVGSTAAAATILFASRSARQPDP SKFNCNHPF EVGSTAAAATILFVSR SARQPDP TKFNCNHPF *****:*****: .*****:*****:*****
<i>Spn100A</i> (CG1342)	mel sim sec yak ere ana pse per wil moj vir gri	EGGSSANSLSAATMQARTPSV---ESTVLPVPEPEPEELP-GVERFEVNRPF EGGSSANSLSAATMQARTPSV---ESTVLPVPEPEPEELP-GVERFEVNRPF EGGSSANSLSAATMQARTPSV---ESTVLPVPEPEPEELP-GVERFEVNRPF EGGSSANSLSAATMQARTPSV---ESTVLPVPEPEPEELP-GVERFEVNRPF EGGSSANSLSAATMQARTPSV---ESTVLPVPEPEPEELP-GVERFEVNRPF EGGSSTNALS AANMQARTPSV---ESTVLPVPEPEPE-P-GVERFEVNRPF EGGSANALSAAT-QARSPQAEDSASVLPVPEPEPE-P-GVERFEVNRPF EGGSANALSAAT-QARSPQAEDSASVLPVPEPEPE-P-GVERFEVNRPF EGGSANALSAANMQARSP LIAAE-SATELPVPEPEPE-P-GVERFEVNRPF EDGSSTNSLSAGNIQARSP LIAAE-AAA VLPVPEPEPE-P-GVERFDVNRPF EGGSSANSLSAATMQGRSP LIAAE-AAATLPVPEPEPE-P-GVELFDVNRPF EGGSSANSLSAATTQARSP L FIL---QEPEPEPEPE-PEGVERFEVNRPF * .*****:*****. *.*: * ***** * *** * :*****

Additional Figure 1: Multiple alignment of Serpin RCL sequences.

Alignment annotation: * Identical residues, : closely similar residues, . broadly similar residues. Inhibitory serpins, identified in red, like *Spn27A* (CG11331), are

characterised by a consensus sequence of small side-chain residues: E[EKR]G[TSG][ET][AGS][YAGS][AGS][VAGS][TS]. These residues form a flexible hinge region, 17-8 residues N-terminal to the putative protease cleavage site residues, P1/P1', marked in green. Both inhibitory serpins and non-inhibitory serpin folds have conserved PF residues at the C-terminal "shutter region" of the RCL. Note that the *Spn42Da* orthologues contain a mixture of several splice variants and the putative P1/P1' sites are unmarked. *Spn100A* represents a novel protein fold of unknown function, with two stretches of homology to the serpin-fold, one of which contains a putative RCL with partial conservation of the inhibitory serpin flexible hinge-region residues.