

Supplementary Materials

Identification of the Ricin-B-lectin LdRBLk in the Colorado potato beetle and an analysis of its expression in response to fungal infections

Table S1. Amino acid and nucleotide sequences of LdRBLk.

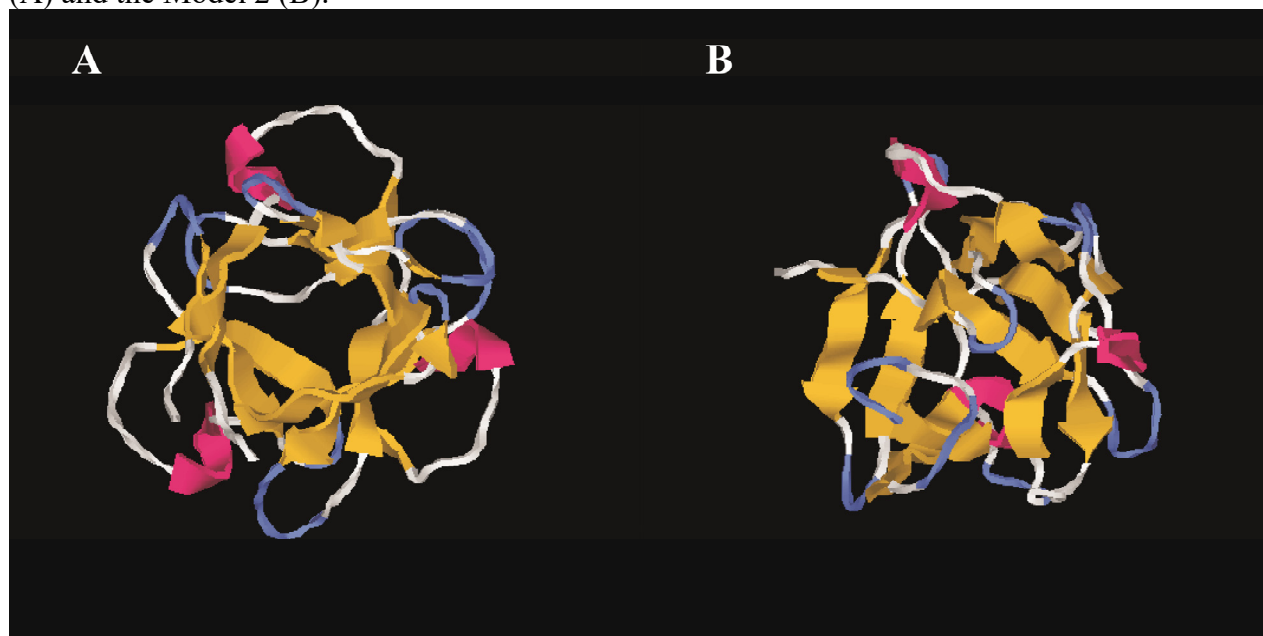
Title	Length	Sequence
ORF sequence	127aa	MSGIRIISKLSGLALDGNGSNVKAAMPNNTVGQIWSLESL GDGTFVIHSYSDKNMVL DIEYGCHGENIILYPRHGDDNQQ FYINSDGTIVSAGDRQKVLQMVRGKV VVRPVNISRNVDHI FNIEQLF
DNA sequence	381bp	ATGTCAGGTATTCGGATCATATCGAAATTATCTGGATTAG CTTTGGATGGAAATGGGAGCAACGTCAAAGCAGCAATGC CTAATAACACCGTTGGCCAAATTTGGTCTCTGGAAAGTTT GGGAGATGGAAC TTTTGT CATAACACAGTTATTCAGACAAG AACATGGTACTAGATATTGAATATGGCTGCCATGGCGAAA ATATCATACTATATCCAAGACATGGTGACGACAATCAACA GTTCTACATCAATTCTGACGGGACAATTGTATCAGCTGGTG ACAGACAGAAGGTTTTGCAAATGGTTAGAGGAAAAGTCGT GGTGCGACCAGTGAACATTTTCGAGGAATGTCGATCACATTT TCAACATCGAGCAACTATTT
qPCR product sequence	134bp	TGGAAAGTTTGGGAGATGGAAC TTTTGT CATAACACAGTTATT CAGACAAGAACATGGTCCTAGATATTGAATATGGCTGCCAT GGTGAAAATATCATACTATATCCAAGACATGGTGACGACAA TCAACAGTTC

Table 2. List and description of genes and primers sequences used in the qPCR.

Gene Name	NCBI GenBank Accession number /Conserved Protein Domain Family/	Gene symbol	Primer sequence (5' – 3')	Product size (bp)	PCR efficiency (±SD)	Tm in qPCR, °C	Primer's source
60S ribosomal protein L4	XM_023165859.1	<i>Rp4</i>	For GAAACGAGCATTGCCCTTCC Rev TCGCTGACACTGTAGGGTTG	119	1,97±0,06	62	Shi et al (2013)*
60S ribosomal protein L18	XM_023172940.1	<i>Rp18</i>	For TAGAATCCTCAAAGCAGGTGGC Rev CTGGACCAAAGTGTT CACTGC	133	1,96±0,04	62	Shi et al (2013)*
ADP-ribosylation factor- like protein 1	XM_023169879.1	<i>Arf19</i>	For CGGTGCTGGTAAAACGACAATATT Rev TGACCTCCCAAATCCCAAACCTT	135	1,95±0,06	62	Shi et al (2013)*
ADP-ribosylation factor 4	KC190027.1	<i>Arf2</i>	For GTGCTCGCGAACCATGTGA Rev AAACCTCCAATCCCTCGTGAAG	140	2,00±0,01	62	Shi et al (2013)*
Ricin-B-lectin k	GEEF01084863.1 /pfam00652/	<i>LdRBlk</i>	For GAACTGTTGATTGTCGTCACCATG Rev TGGAAAGTTTGGGAGATGGAAC	140	1,96±0,04	62	^a ISaEA
Attacin_C domain contained protein	XM_023168834.1 /pfam03769/	<i>attacin</i>	For TGAGAACTCCACAAAATATTCACCTCG Rev CTAAGGGTATGGCAGCAACAAC	98	1,90±0,07	62	^a ISaEA
Heat shock protein 90	KC556802.1 /pfam00183 /	<i>hsp90</i>	For GGGTGTAGTCGACTCTGAAGAC Rev AGAGCTCCTCAAACAGTTCCAA	125	1,90±0,06	62	^a ISaEA

Shi, X.Q., Guo, W.C., Wan, P.J., Zhou, L.T., Ren, X.L., Ahmat, T., Fu, K.Y., Li, G.Q. Validation of reference genes for expression analysis by quantitative real-time PCR in *Leptinotarsa decemlineata* (Say). *BMC Res Notes* **2013**, 6, 93 doi: 10.1186/1756-0500-6-93 * Primers were modified for this work by Rotskaya U. ^a Primers were designed by Rotskaya U. in programs Primer-BLAST <https://www.ncbi.nlm.nih.gov/tools/primer-blast/>, IDT OligoAnalyser Tool and IDT UNA Fold Tool <https://eu.idtdna.com> .

Figure S1. Tertiary structure of LdRBLk peptide constructed by I-TASSER program. The Model 1 (A) and the Model 2 (B).



SignalP-5.0 prediction (Eukarya): LdRBLk

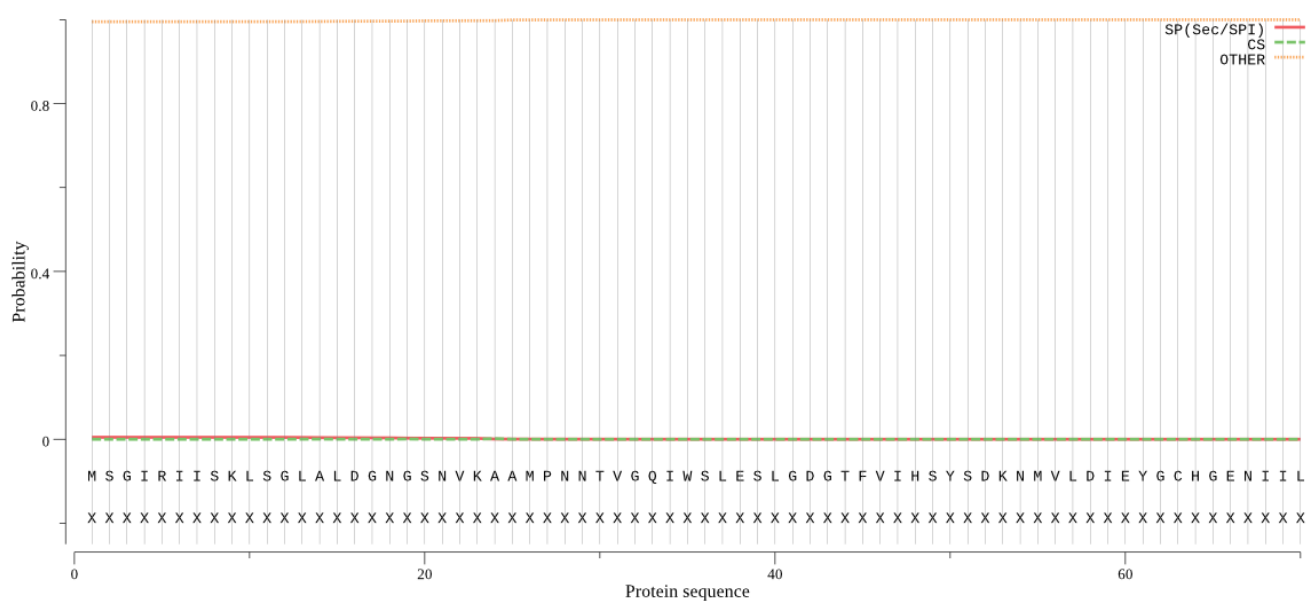


Figure S3. Prediction of the presence of a transmembrane domain in the LdRBLk peptide in the TMHMM online program. Possible transmembrane helices: inside to outside helices (i->o) 0 found, outside to inside helices (o->i) 0 found. Suggested models for transmembrane topology: no possible model found.

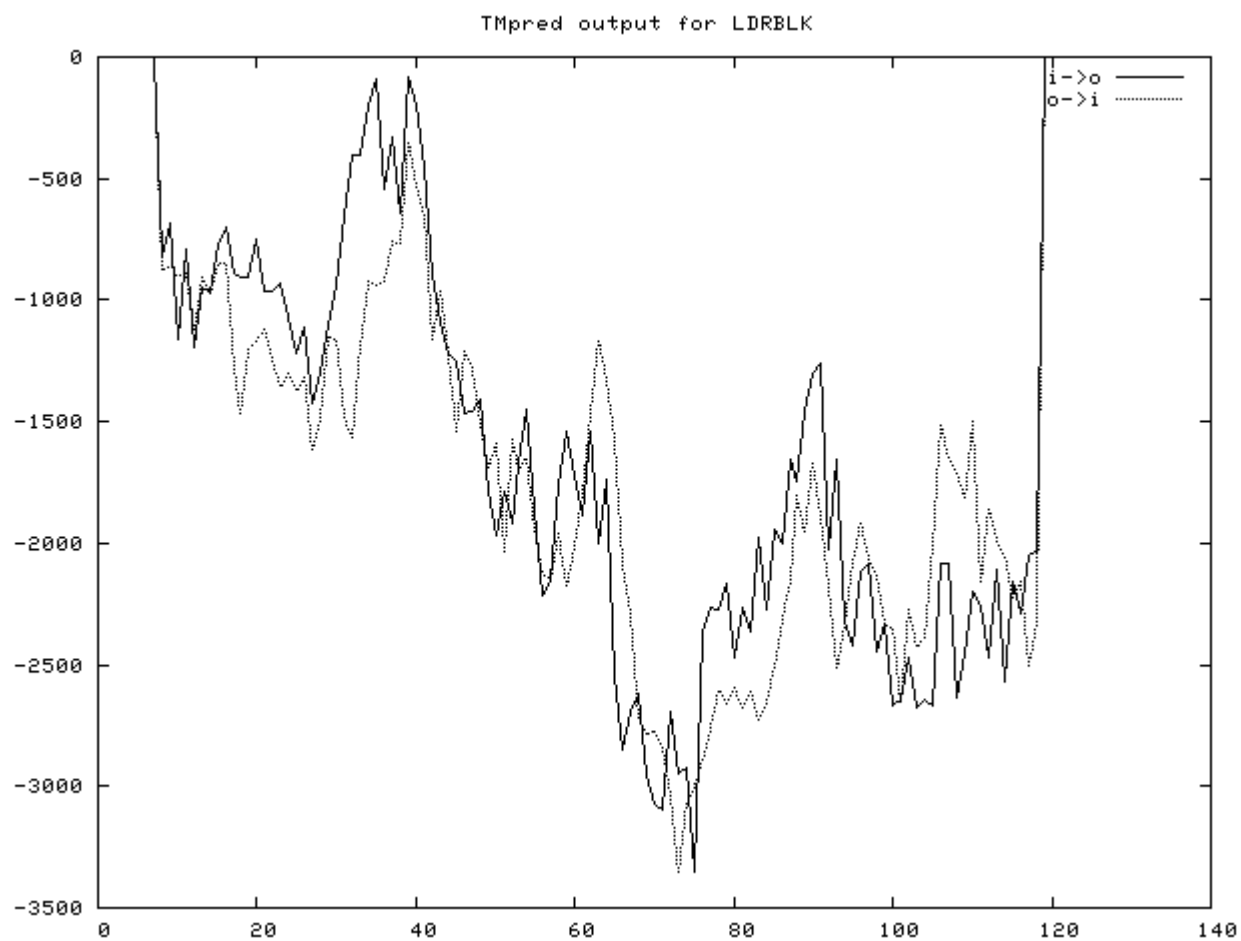


Figure S4. The analyze of hydrophilic property of LdRBLk peptide by the online program ProtScale using the Kyte-Doolittle (A) and the Hopp-Woods (B) algorithms.

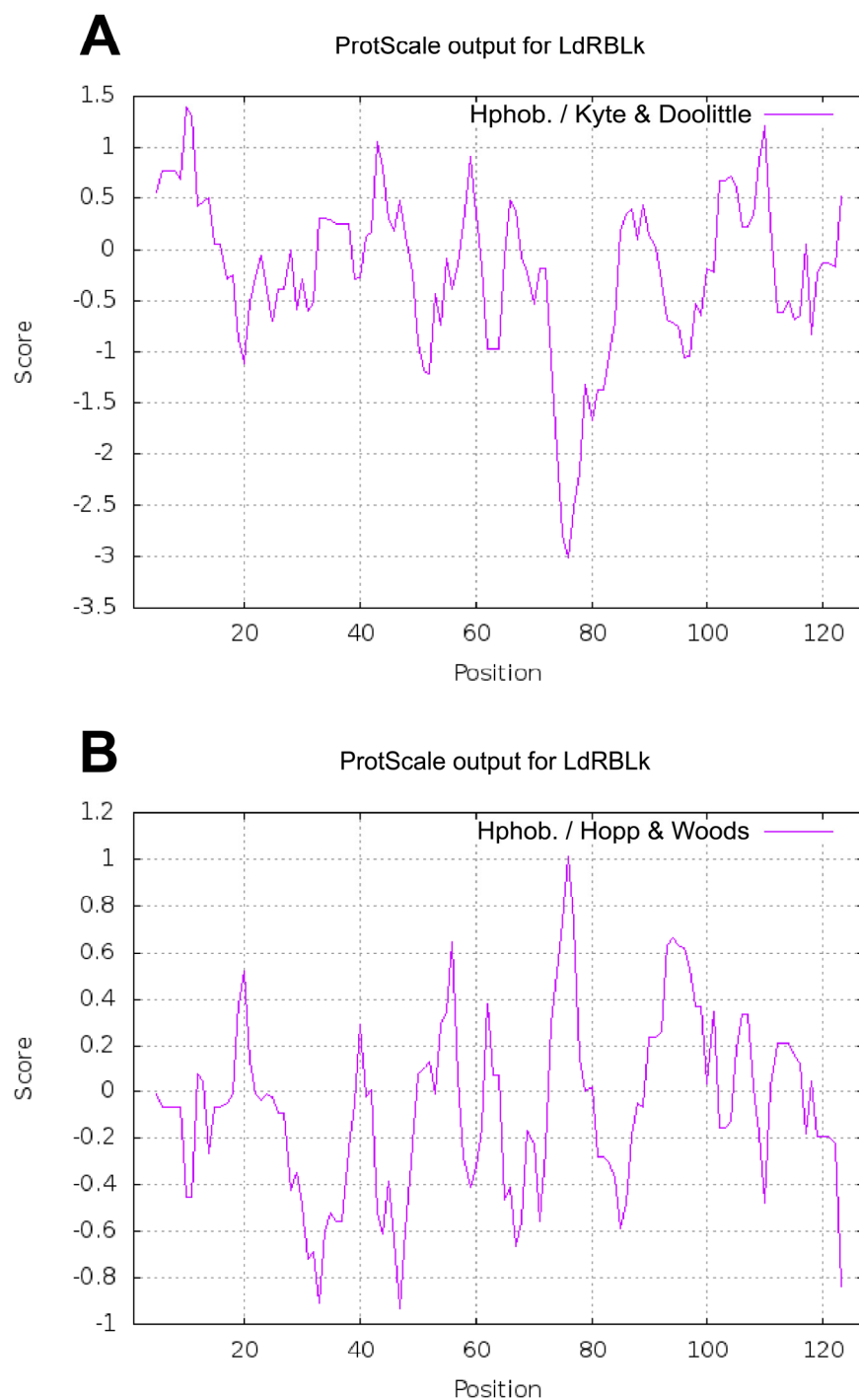


Figure S5. Multiple alignment of amino acid sequences of LdRBLs peptides of *L. decemlineata*. It was performed using the CLUSTAL algorithm in the MAFFT online program (v7.471).

CLUSTAL format alignment by MAFFT FFT-NS-i (v7.471)

```
LdRBLk          LVPS-----TFNIKMSG-----IRIISKLSGLALDGNNGSNVKAAMPNNTVGQIWSLE
XP_023029736.1  MATS-PVILFFIICISSMYCLADDYVIIQSKTSGMVLDGNHYYVRARLADGSIGQKWNLT
XP_023029739.1  MASN-SAILCFFICISSMYCLANAYIIIESKNSGLVLDGNEYNVRAMPADGSIGQKWTLT
XP_023029750.1  MASNFSSYLFFLVVIAILTCLTEAQVRIKSKKTGLVFQYNGTNVEVAEPNDSEGQKWILE
                :...      * : ::      : * ** :*::: *   *..   ::: ** * *
```

```
LdRBLk          SLGDGTFVIHSYSDKNMVLIDIEYGCH-GE--NIILYPRHGDDNQQFYINSDGTIVSAGDR
XP_023029736.1  STKIGRFIITN-ARNGHVLDVQHGCKLGNTSNVIVYKNHFGRNQEFYINSDGTIVSSCDN
XP_023029739.1  PSDTGRFVITN-ARNGHALDVQDGCEKGN--NLIVYQKHSGTNQLFYINSDGTIGSACNT
XP_023029750.1  CSQKGHFIINS-NKQGLAMDIKKGCKIGQ--SVMVYDKHSRQNQQFFINSDATIGTACDA
                * *:* .   :. .:*:: ** . *:   :::* .*   ** *:****.** :: :
```

```
LdRBLk          QKVLQMVRGKVVRPVNIS-RNVDHIFNIEQLF
XP_023029736.1  NEIIEMKDGKVEVCGKEEPCSSGDVFFKILETK
XP_023029739.1  DLIVEMKDSNGQLGDIQNSCSSGDVIFEIQAAQ
XP_023029750.1  NLILLVKDRNVEVAPRDNY-FDFDIHFEIEPIE
                : :: :   :   :   . * *:*
```

Figure S6. Multiple sequences alignment of Ricin-type β -trefoil lectin domain (pfam00652) contained proteins by hierarchical clustering algorithm.

Corpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **1988**, *16*, 10881-10890. doi:10.1093/nar/16.22.10881

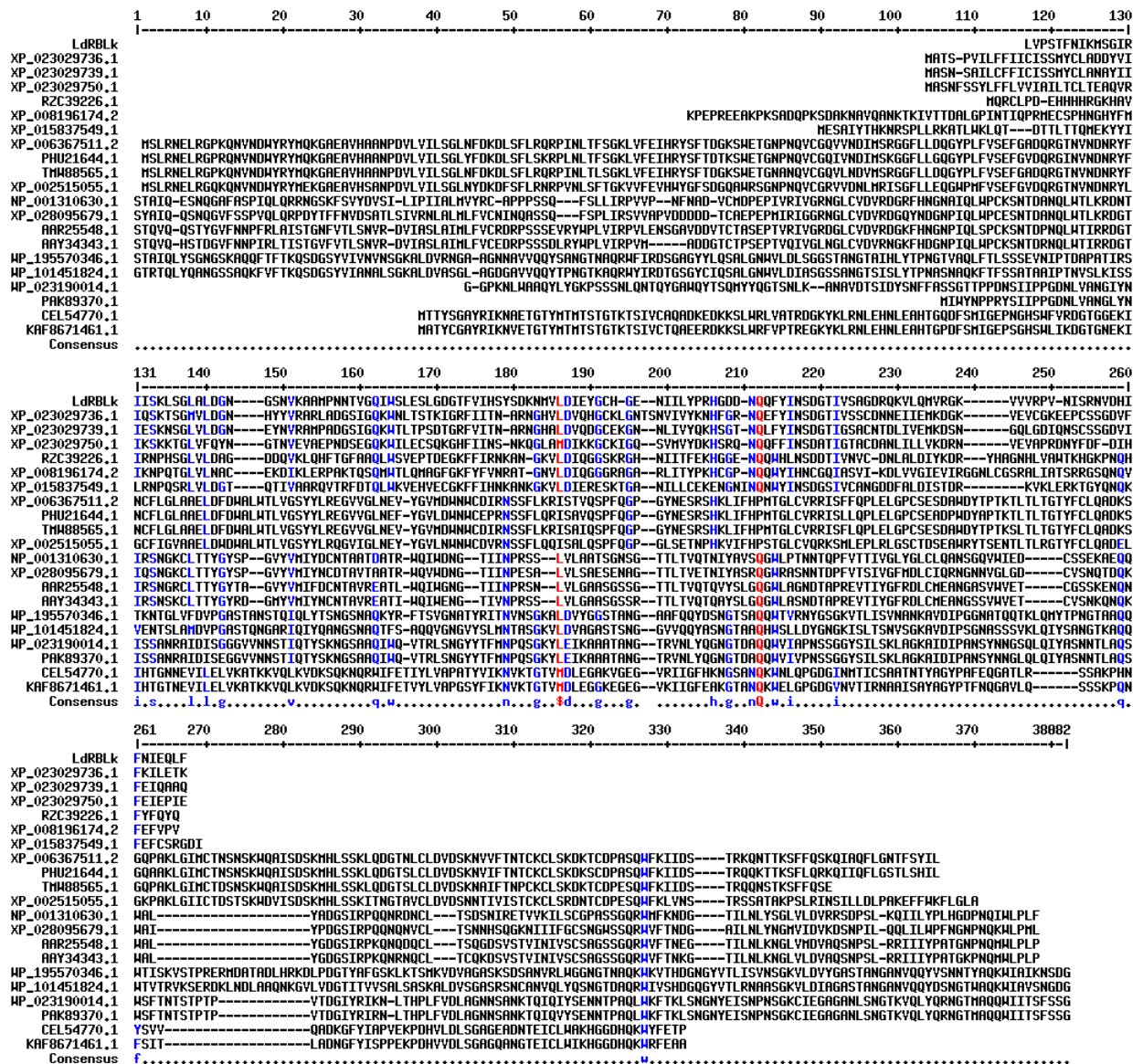
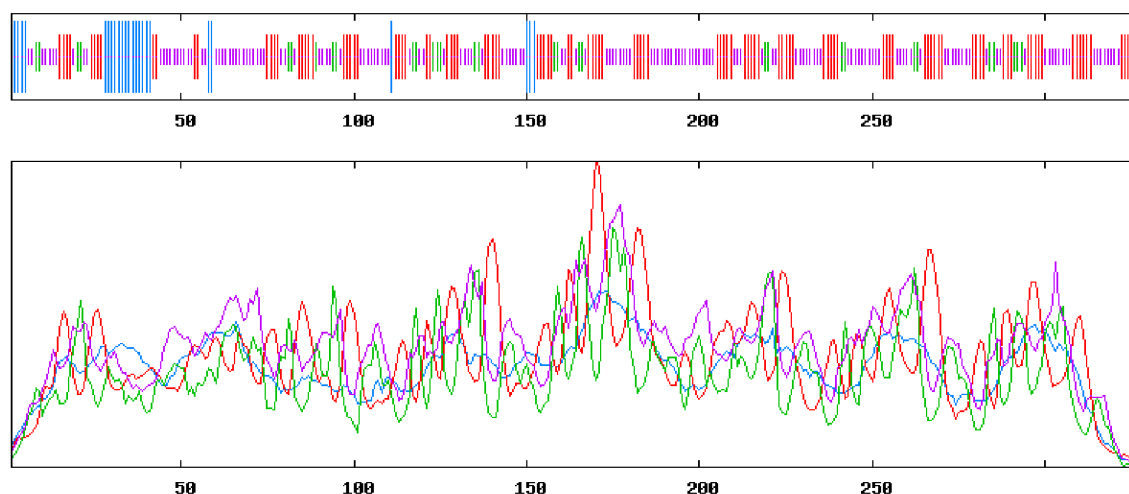


Figure S7. Predicted secondary and tertiary structure of β -lectin from *R. communis* ricin toxin (RBL). **A.** The secondary structure in the online SOPMA program. Strokes and lines are shown in blue, red, green and purple, representing an alpha helix, an extended strand, a beta turn, and a random coil, respectively. **B.** Tertiary structure of peptide constructed by I-TASSER program. The Model 1. **C.** Prediction of the ligand binding site for the RBL peptide by the I-TASSER program. The target ligand is galactose.

Sphyris, N., Lord, J.M., Wales, R., Roberts, L.M. Mutational analysis of the Ricinus lectin B-chains. Galactose-binding ability of the 2 gamma subdomain of Ricinus communis agglutinin B-chain. *J Biol. Chem.* **1995**, 35, 20292-20297. doi: 10.1074/jbc.270.35.20292. PMID: 7657599.

A



B



C

