

Supplementary Information

Sugar-lectin rich interface between soft tissue and the stiff byssus of *Atrina pectinata*

Jimin Choi ^{a,§}, Elise Hennebert ^{b,§}, Patrick Flammang ^{c,*}, Dong Soo Hwang ^{a,d,*}

a: Division of Integrative Biosciences and Biotechnology, Pohang University of Science and Technology (POSTECH), Pohang, 37673, Republic of Korea

b: Cell Biology Unit, Research Institute for Biosciences, Université de Mons-UMONS, 7000, Mons, Belgium.

c: Research Institute for Biosciences, Université de Mons-UMONS, 7000, Mons, Belgium. E-mail: Patrick.Flammang@umons.ac.be

d: Division of Environmental Science and Engineering, Pohang University of Science and Technology (POSTECH), Pohang, 37673, Republic of Korea. E-mail: dshwang@postech.ac.kr

§ : Both authors contributed equally to this work.

Supplementary Data 1. Peptide sequence of interfacial protein with carbohydrate binding domain. Carbohydrate binding domain highlighted in yellow.

1. m.14349 [5' partial sequence]

KPVPKYKSVVPDYKPPVVPDYKPPVVPDYKPPVVPDYKPPVVPDYKPPVVPDYKPP
VDYKPPVVPDYKPPVVPDYKPVPEVYKKPPVYKPKKPVYKPKKRPAYKPKKKPIYK
PTKPDYKPTETYKPGGNGAVKVGGHGYFPSSYGAGYYGVGYGKGYKQVPDAD
PLSATGGCPHWLPYGGMCYLYSRDKLGWFEASMKCMYMGGYLAIANSAHENSY
FKLMAKKYELKPGVWFGFLNDVLFPSHKWFWGYGKKQCKWFDWGLKEPKYDGA
GYKHCVALWCDYKWQWKVENCYSKKYYICELNPKRPCKCGY*

2. m.8455 [complete sequence]

MYQISR VHQPISYYIPNGLHVGKQVILRGRVTSGETFAINLQETENPGDGEIALHFNP
RPSTDVCVRNSFQGNWGP EETDQPHFPDNNRSFLLRIEVGDEGYRTYVNGKPYVNF
NHRLDMGNVHFLHLTEGAEFYDISFQDRYSLPYRTEIPGRMRAGKAVRVRGASQDN
DGFSINFACDCENETCAFHFNPRNEG VVVRNANLGGWGEEERDYDAEFFPQPNQY
FDALFVAGEDKYHVYVNDKYFSEFNHRQGLE DVSHFHQGNMDIKDVEYMEPLDD
DFVKIIPSGLEKGDVMVFRGFMKPDGDTFSINFMNGYSADDDIAFHFNPRVGQGEVI
MNTCMGGNWGEEDKEDLPSVFANREPF EVKVVTKRNKFKIYVNGKKGKYNARG
NVEDIKGVNVRGDAYVYQVKLQRKLEKPAWERLPGGLREFGWIVVHGI AKK GSEG
FAVNLRCGDDDDSNIALHFNPRLTEECTVRNTMTDGNWGEEERDQSPFPFEKKDTFE
VAINI QPKFVTYVNGNHVYVNYGHRLPLDSVCHLQLTGNADFFFEPEFL*

3. m.9790 [internal sequence]

FTSASMVKLSTVSLLLILNSVVG IASSEIPTGGTYGINNVIFGKNKGNTNTGNKFDGG
KPQIRGKTGDVVFNGSGKLIQNV RKVEPRKTNQKPAVLPTYGKNAPIDQPPVKKNPN
TILKKGDVIIHPFNKEVSIQQPLIKTNTGKEKIVMNKFDKSIPREMPSSNDSPVMVSTEK
LKPISKTTSKVAKNPKPDRFPSGKA IKPIKNKRKPKPTVVVDNNGGKQPQSYIQKNIKFP
APYSGYFPTS DQFSYAAGYRQKYTKVGNVPPEPTDTGCPYGWMSYSGRCYYFSRDK
I

4. m.14568 [5' partial sequence]

PFEIIILTQD SHFKIAINGNHFT EFQHRIPVHRITHLNI SAGIQVSMVRYEPTAAPQVVQ
GGYPQQPMYQPSAPSIGFQPPPAFYPSAAPPMSYPSAAPQMPASPMYNPTVPLTVQIP
GGMTPGKMIFLSGIPKNTRFTVNIQDRAASGGEIAFHFDVRFNVSGNVNEVVRNHFT
HGTWGV EERQKPKFPCPNANFDMIIMAEQQSFRVAVDNQHFIEFKSRLLPV ARFNY
LNVTGNVRLTQVRFQ*

5. m.15895 [complete sequence]

MLVLCILLPGFMKLV DGQMLLQNTGQSKISSYYPWLVDTLTEARKENANNYNNVM
KTLNMCTMQAVGYKQTTNAAGSTGAPKISP KDFYTGITQKLSEMSTDSQNQKQLIN
NITDLMETQMGRIDEIERQNLKLLAALKKLLSVNSIAFGSNSGDGRCSVDDFKCKE
NECKTHSGYVYRRDNAIGVCLKFVSGFSDIN VTRFAEKCAKDDGMLLKIDGYEKHQI
IHKYLGTL SFGDPQIIIQGNDKISEGAWEFDDGTKMTYLPWYSGQPNSYSGEQDHIGM

KKSFNSYKWGDIWRVKS GFSAAYLCEIVV*

6. m.8438 [3' partial sequence]

MWKFYQLCFLMFLLAEAESRSQDCKKPFVSLHRACYKFKEMKSWTEAQRLCSEE
GATLVSIGSMAEKGQIVKQLHKIITHMKHASWWVGLKLNKTIKWKWQEGTDLNL
RVTKWAPGEPNNALNGEHCAELSFRGTINDGNCSFERPYVCELIKITQPPTSSTVATR
LTTTGQPTTVTMTMTSTTSSTPQPTTTEKPITTTQNTTPQTTTNSKITSTTKAVGYTTT
QYTTPETTTNDTTTSATKAAEYTTTQTTKKYNKTTRNYLSGFPVDGNIHHFKTTRST
TVKPTIETTTMKIKIPSRPDATPKPNTRCPLVNAYGTVWPRAVKGETVSQKCKETE
GYATWTCGGYPVHWLGEPNVSECASASFKEIVSMSTSIAGNKSSLTPEVAEDITKIIV
EKTETSNTMSAADIAVSTNLLKITSKVKPKDSKQAKKLVDQFVSAGNNLIEQTKHSL
AIKPFDKSRAASRLLSTMELATSNMVDVINEPTIFQTRTDKISLTLHVLNTTNNENLIYD
DSGTDNTFKIPSSVLKSFNKDGLTKVVFMTDFDVGDLGSKENTGNTDNYKTEVAS
DVISASFGKEKAIHLTEPVTFSMKMTKSVEYGTIPVCSFWNFSIGQIGDWSQVGCERQ
DYNATHTTTCQCDHLTNFAILLDVSGVDVAYHHTRILDYITYIGCIISLSLSTAWLTFQ
CLSTLQGERNSIHKNLVFCFAAELIFVVGISRTEQKVACAAIALLLHFFLSAFTWFMF
MEGIHIFMLVQVFDKAKSKMKYYYLFCGYGAPIIIIVAVTATIDFKAYGTADHCWLTT
ENWFIWSFAGPVALILVMNAGVLA YSLTMVCKHSDYVFSRERAESSNFRA

7. m.8549 [complete sequence]

MLELQKIFIVIIITSDVTHQQTLLAKPDNCPSIDAENGNLQIIRQSLLRNKETQEEMKT
DIAEIWRQLNMLKSLLDQKSNSTSTTKGTTTDMYVSTGTSTKPTTRKPKLACMDPW
LPYEDYCFYIGEERSWSKASGQCQSLDSELATLDDKEKNEFIAGQIKRTENYWIGG
NDIDVEGIWRWKGTGEVIGPYKNWYLNPDGRTNQNCMWYNYQGKPKWIDAPCK
RKLFFVCERKM*

8. m.218 [complete sequence]

MVPKLMILVAVIIIITDVTDQQVLLQKPDNCPKTDNERDLKVLRELTLKNQAKQEE
MINEISELKQQFKLLQKSNLQPTTTTSTTRPTTNKLAPCNPPWLQYEEFCFFISSKKE
DWKAAAENCRQMNAELAAFETKDKNDFMGRKIISSNDYWIGGNDLDKEGDWRWK
KGNKKIGPYTNWYANDPDGRTGQNCMWFNYQNTQQWIDAPCSRSLNFICERNI*

9. m.14097 [complete sequence]

MDLGILRLVPTLSIVILSTVYADPCPSAEWKKDEESRMCYLFSEQRQTFTQAVDFCEK
YETHLLQYKGFQSKQTIRNFLNMNISEMWIKAPDTSEGEYEINSDRNSSTETSQNVST
ETISQIRYRLLSRREESCRVINMTNDKEVIESERCTRKLFPVCEESIKHLEKRRKGMRV
KVVAIIIVGAIIGVLFVIVKTCHKRLKMIEGGEYSIVENKKKKEDEMILRKSEYIHI
GSLRQPDTSGDTKV*

10. m.12209 [complete sequence]

MLLIASVTLVAVIFRPVTGCSCLPNPQDSFCTATFALKATVISDDTEYGDSNYTPLE
RQYTIRFMRKDIYKGSNFFPGANTTASIKTAGYQSTCGIRFEIGQEYFITGSLDNLVVS
LCDWISPLNLTNHQRRGIQYMYEQGCNKCSIMRCFGRSCNNANVNQTCMVDNPT
NINNCHNKYGFLENDQGICKWKDNRLKLLCLNPKEVSRQVVGTQ*

11. m.2564 [internal sequence]

EPIVQIGGPSLDADLVRQNALPLIDPIHTDVCGSVKGCYRVPEHCWEPYCDYIATWRP
YGTANTYLIEMSAMVDGVTDRHVSLGLSGDTRWGGDRVFECVHNGGTGVTQVFQ
AQSVGKSTERYQNSQAGISQRGYADGLKLG

12. m.1338 [internal sequence]

SRRSSDPDKSGYGKKPKSDCSAKCTGDNSQTCGGTWRMNVYQINVEYMGCYLDKS
TRLLEAKFTHGADMTLEKCFKTC SNGKYRFAGLQYSTQCFCGNDISKPGYGIRPESE
CAAACRGDKDQTCG

Supplementary Table S1. Proteins and peptides identified through MS/MS analysis from each gel pieces.

(separate excel file)