

Pathogen recognition of a novel C-type lectin from *Marsupenaeus japonicus* reveals the divergent sugar-binding specificity of QAP motif

Rod Russel R. Alenton¹, Keiichiro Koiwai¹, Kohei Miyaguchi¹, Hidehiro Kondo¹, Ikuo Hirono¹

¹Graduate School of Marine Science and Technology, Tokyo University of Marine Science and Technology, Minato-ku, Tokyo, 108-8477

SUPPLEMENTARY INFORMATION:

Fig. 1 Sequence alignment of MjGCTL and LvCTLD. The amino acid sequence of MjGCTL and LvCTLD (AEH05998) were also aligned using ClustalW in Genious 6.1.8 (Biomatters Ltd., New Zealand) software. The predicted amino acid sequences of MjGCTL and LvCTLD have the same length share the same QAP motif (Fig. 1B) with a sequence identity of 77%.

MjGCTL	1	MKAIVLLLCIGFATALECTGDEVACGSAERCVPYRYICDFSDSCSDGSEDDPYLCWAWNN	60
		MK VLLLC+GFA+AL+CTGDE+AC S ERVCPYRY+CD D+DC+DGSDE P LC AW N	
LvCTLD_AEH05998	1	MKRFLVLLCLGFASALDCTGDEIACTSGERCVPYRYLCDSNDNDCADGSESPDLCLAWRN	60
MjGCTL	61	TECERGSAQCLTNGRAECIPIETYCHRTQPACSGSLNRRVCSIIEDKKLVPLASIKFIPD	120
		T+CE+G AQC NG +CI IE YCHRTQPAC GSL+RR+CSII++K LVPL+SI+ P	
LvCTLD_AEH05998	61	TQCEKGQAQCHANGDDQCISIEAYCHRTQPACDGSDDRRCISIIKNKSLVPLSSIRLPPS	120
MjGCTL	121	NEPADAYNRSVSLGAELRTNLNNTLSHPDCPDFYTRVGDQCLSVFYVGRSSWGGEARAFCK	180
		N+P AYN+SV LG+ELR NLN+TLSHPDCE FYTRVG QCLSVFYVG SSWGGEAR+FCCK	
LvCTLD_AEH05998	121	NDPGVAYNKSVELGSELRLNLNSTLSHPDCPRFYTRVGGQCLSVFYVGSWSWGGEARFCK	180
MjGCTL	181	HIGGDLLSIQNASHYIDLNVNHLSENQITSDFWLGGRYELDDLWMLDGTMPQGTPFWFS	240
		HIGGDLLSIQN +HY+DLVNHL +N+ITSDFWLGGRYE+DDLW WLDGTMP+GTPFWFS	
LvCTLD_AEH05998	181	HIGGDLLSIQNVNHYVDLVNHLVDNRITSDFWLGGRYEVDLWTLWLDGTMPMPRGTPFWFS	240
MjGCTL	241	LRRYHHCNTRNVTAGTYQVLEANNGECEHYTQAEEDPPRGFCAAITYGKHFYMSDEDCL	300
		LRRY C+ RNVT+ GT +V EANNGECEHYTQAE PP+GFCAAITY KHFYMSDEDCL	
LvCTLD_AEH05998	241	LRRYDSCNPRNVTLTGTSEVREANNGECEHYTQAEETPPKGFCAAITYDKHFYMSDEDCL	300
MjGCTL	301	ADMSPLCVTSV	311
		ADMSPLCVT+V	
LvCTLD_AEH05998	301	ADMSPLCVTAV	311

Fig. 2 Comparison of 3d model of the carbohydrate recognition domains of MjGCTL and LvCTL D. The 3D-structures of MjGCTL and LvCTL D predicted by Swiss Model are shown in Fig. 5. The z-scores 19.637 and 19.2970, respectively, indicate good matches to the template (bovine mincle). Black arrows indicate the locations of the CRDs and green spheres indicate the locations of the bound calcium atoms predicted by Swiss Model. Note that the CRD of MjGCTL is co-located with the Ca²⁺-binding site while the CRD of LvCTL D is not near Ca²⁺-binding site. The carbohydrate-binding QAP motif is located in β 5-sheet.

