Supporting information to

The Ig-like domain of Punctin/MADD-4 is the primary determinant for interaction with the ectodomain of neuroligin NLG-1

Semeli Platsaki[1], Xin Zhou[2], Bérangère Pinan-Lucarré[2], Vincent Delauzun[1], Haijun Tu[2,5], Pascal Mansuelle[3], Patrick Fourquet[4], Yves Bourne[1], Jean-Louis Bessereau[2], and Pascale Marchot[1]*

[1]CNRS / Aix-Marseille Univ, Laboratory 'Architecture et Fonction des Macromolécules Biologiques' (AFMB), Marseille - France.
[2]Univ Lyon / Univ Claude Bernard Lyon 1 / CNRS / INSERM, Institut NeuroMyoGène (INMG), Lyon - France.
[3]CNRS / Aix-Marseille Univ, Institut de Microbiologie de la Méditerranée (IMM), Marseille Proteomics (MaP), Marseille – France
[4]Aix-Marseille Univ / INSERM / CNRS, Institut Paoli-Calmettes, Centre de Recherche en Cancérologie de Marseille (CRCM), Marseille Proteomics (MaP), Marseille - France

[5]Present address: Institute of Neuroscience, Hunan University, China

***Corresponding author:** Lab 'Architecture et Fonction des Macromolécules Biologiques' (AFMB, UMR-7257 CNRS/AMU), Faculté des Sciences Campus Luminy - Case 932, 163 avenue de Luminy, 13288 Marseille CEDEX 09, France. Tel: +33 (0) 491 825 579. Email: pascale.marchot@univ-amu.fr

Running title: Molecular bases of Punctin/MADD-4 interaction with NLG-1

Supporting Figures

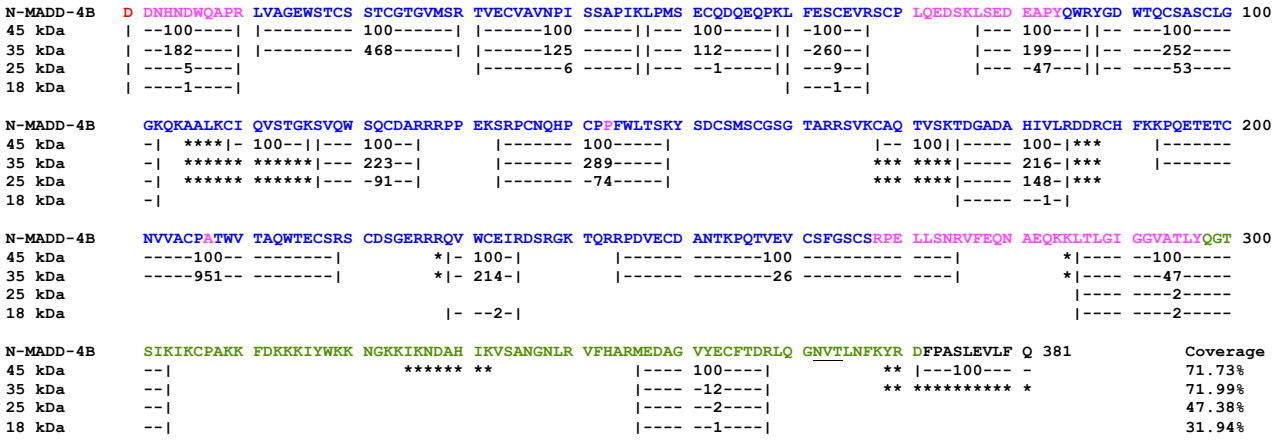

Figure S1. LC-MS/MS analysis of intact and processed N-MADD-4B proteins from distinct SDS-PAGE bands. For each of the four entities with apparent masses of ca. 45 kDa, 35 kDa, 25 kDa and 18 kDa (see gels in Figs. 2A and 3A), the identified tryptic peptides (framed by vertical bars) and their respective intensity values are indicated below the N-MADD-4B sequence, coloured as in Fig. 1. Additional residues revealed from miscleavage peptides are indicated by stars. The coverage values (in %) are indicated on the right of the last lines. The intensity corresponds to the area of the monoisotopic precursor ion calculated by Proteome Discoverer. The N-terminal, unnumbered Asp residue (in red) arising from the cloning procedure and evidenced by N-terminal sequencing (see Experimental procedures) and intact mass analyses (see Supp. Fig. S2) was considered for data processing. The putative N-glycosylation site is underlined. Absence of coverage for peptide Y^{150}-R^{163} or -R^{164} or -K^{167} and peptide L^{359}-K^{368} may denote O-glycosylation at Ser^{156} and N-glycosylation at Asn^{362}.

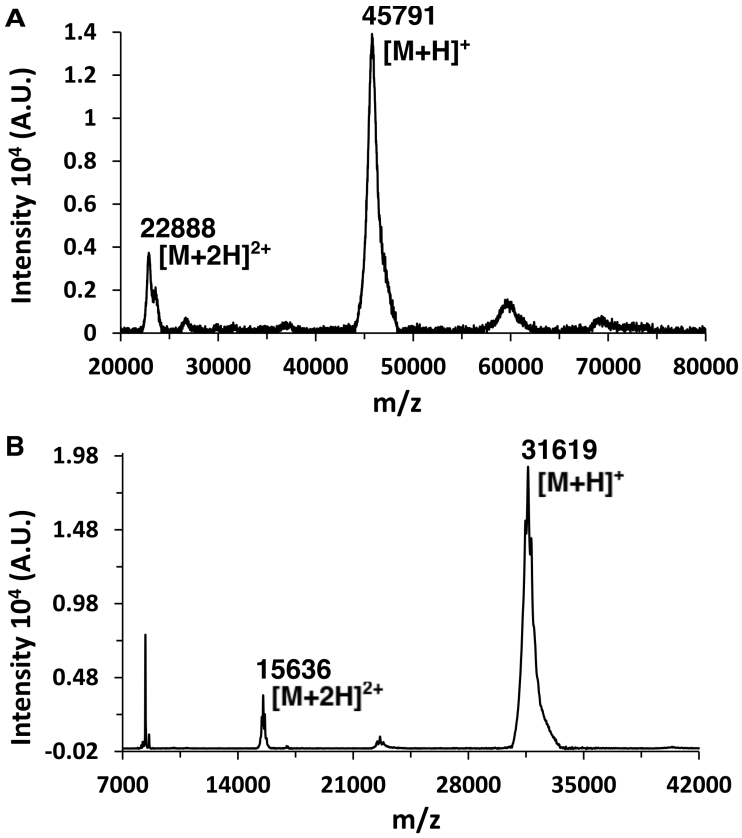

Figure S2. MALDI-TOF-TOF mass analysis of intact and processed N-MADD-4B separated by chromatography. ***A***, Intact N-MADD-4B generated a mean mass of 45 781 Da for a theoretical mass of 42 915 Da. The 2866 Da difference suggests occupancy of Asn362 by a fairly large N-glycan moiety. ***B***, Processed N-MADD-4B generated a mean mass of 31 345 Da for a theoretical mass of 31 337 Da. Reduction of the mass difference to 8 Da only likely reflects the loss of the N-glycan moiety in processed N-MADD-4B. Mean masses were calculated from the measured $[M+H]^+$ and $[M+2H]^{2+}$ values. Theoretical masses were calculated from the protein sequence using the ProtParam tool in Expasy.

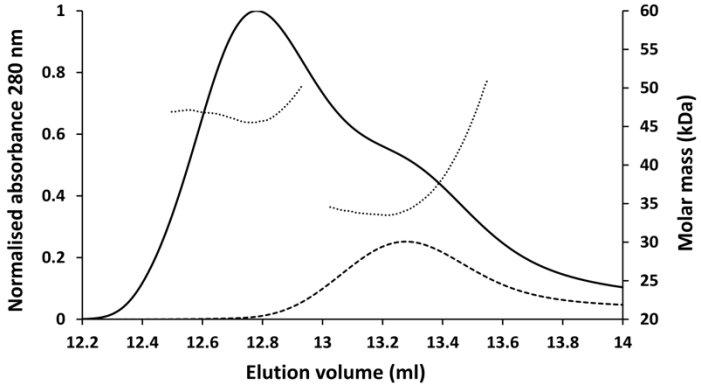

Figure S3. N-MADD-4B and N-MADD-4B$_{\Delta Ig}$ behave as monomers in solution. SEC-MALS analysis of a processed N-MADD-4B solution sample (solid line) and its N-MADD-4B$_{\Delta Ig}$ component purified by cation-exchange (dashed line), both injected at 2 mg/ml, yielded calculated molecular masses of 45.7 kDa for the intact protein and 33.7 kDa for the fragment (dotted lines), consistent with monomers.

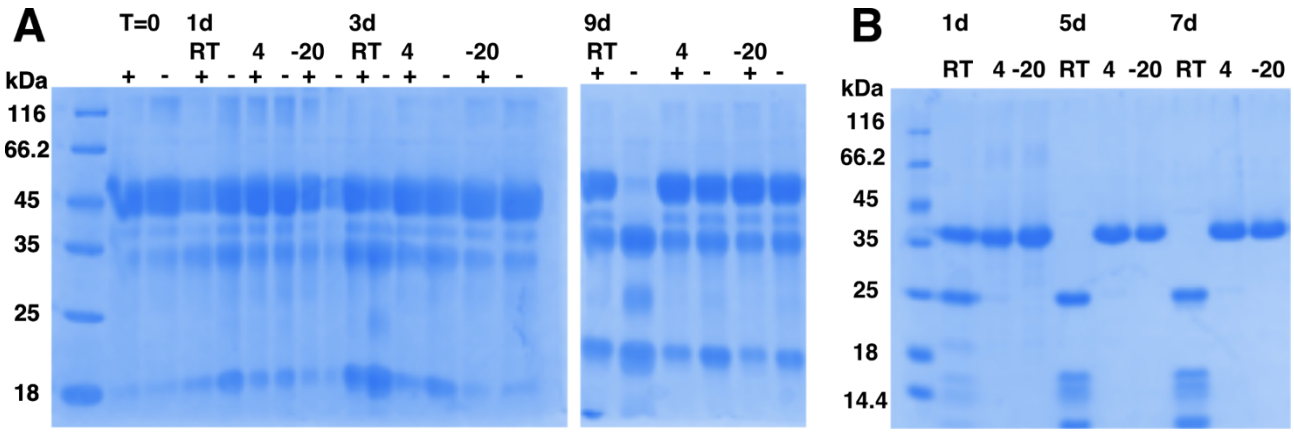

Figure S4. N-MADD-4B and TSP1-4 maturation is time- and temperature-dependent. SDS-PAGE analysis of the time-course of: ***A***, N-MADD-4B maturation after 1, 3 and 9 days, in the presence or absence of a protease inhibitor cocktail and 2 mM EDTA; ***B***, TSP1-4 maturation after 1, 5 and 7 days, both at room temperature (RT), 4°C or -20°C.

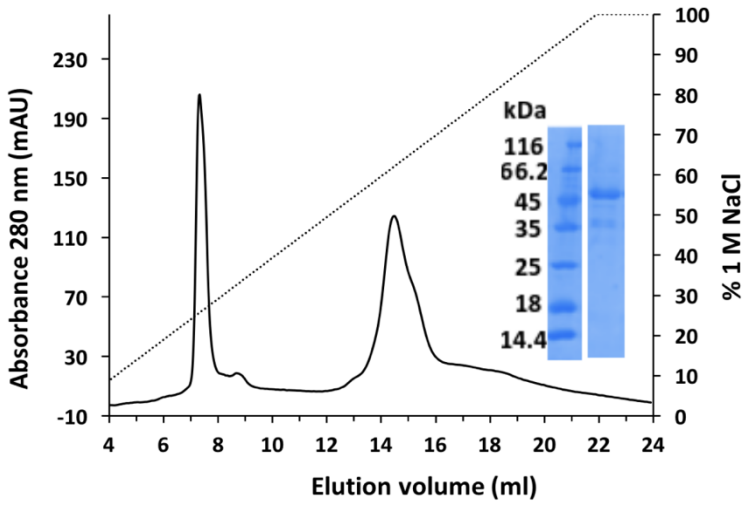

Figure S5. The N-MADD-4B double mutant ($N^{321}A$, $N^{336}A$) undergoes similar processing as N-MADD-4B. Cation-exchange analysis (MonoS, NaCl gradient) of the affinity-purified N-MADD-4B double mutant generates a similar profile as for the original N-MADD-4B protein (see Fig. 2B). Inset: SDS-PAGE analysis of the loaded (input) sample confirms presence of intact and processed protein populations as found for N-MADD-4B.

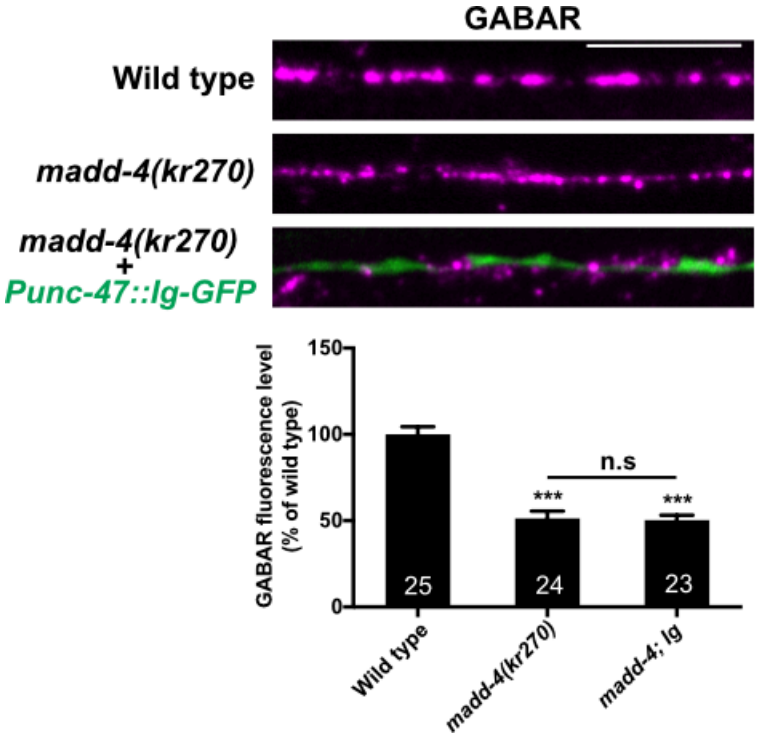

Figure S6. The Ig domain of MADD-4 is not sufficient to rescue $GABA_AR$ clustering defect in the *C. elegans madd-4* null mutant. Confocal detection of RFP-$GABA_AR$ (*kr296* knock-in allele, magenta) in wild-type animals or *madd-4(kr270)* null mutants or upon expression of an Ig-GFP fusion protein (*Punc-47::Ig-GFP*, green) in GABAergic motoneurons of the *madd-4(kr270)* mutant. $GABA_AR$ fluorescence levels were quantified and normalized to the wild type (mean ± SEM). One-way ANOVA followed by Tukey's multiple comparisons test. $^{***}p<0.001$. n.s, not significant. The numbers of animals are indicated inside the histogram boxes. Scale bar, 10 µm.

Supporting Experimental procedures

Mass spectrometry (MS)

Peptide mass fingerprinting - Proteins in SDS-PAGE gel-excised bands were submitted to trypsin digestion using a slightly modified Shevchenko protocol (Shevchenko et al., 1996). Briefly, the bands were washed and destained using NH_4HCO_3 100 mM pH 8.0; reduced using DTT 10 mM (56°C, 45 min) and alkylated with iodoacetamide 55 mM (25°C, 30 min in the dark), both in NH_4HCO_3 100 mM pH 8.0; dried at room temperature; and reswollen into NH_4HCO_3 25 mM supplemented with trypsin 12.5 ng/µl (V5111, Promega) (37°C, overnight). Peptide extracts collected from the digestion solution, from a first extract in formic acid 5%, and from a second one in formic acid 5% / CH_3CN (40:60 v/v) were pooled and dried in a centrifugal vacuum system.

Samples were reconstituted in TFA 0.1% / CH_3CN (80:4 v/v) and analysed by LC-MS/MS using a LTQ-Orbitrap Velos mass spectrometer (Thermo Electron, Bremen, Germany) online with a nanoLC Ultimate 3000 chromatography system (Dionex, Sunnyvale, CA). Peptides were first concentrated on a pre-column (C18 PepMap100, 2 cm x 100 µm, 100 Å, 5 µm, Dionex) in loading buffer ((0.05% TFA / CH_3CN (98:2 v/v)), then they were separated on a reverse-phase LC EASY-Spray C18 column (Acclaim PepMap RSLC C18, 15 cm x 75 µm, 100 Å, 2 µm, Dionex; flow rate 300 nL/min) equilibrated in 4% of solvent B ((0.1% formic acid / CH_3CN (20:80 v/v)) in solvent A (0.1% formic acid) and eluted using a linear gradient of 4-55% B in A in 30 min.

For peptide ionisation in the EASY-Spray source, the spray voltage was set at 1.9 kV and the capillary temperature at 275°C. The spectrometer was operated in the data-dependent mode to switch consistently between MS and MS/MS. MS spectra were acquired in the 300-1700 m/z range at a full width at half maximum (FWHM) resolution of 30 000 measured at 400 m/z. For internal mass calibration, the 445.12 m/z ion was used as lock mass. The 10 most abundant precursor ions were selected and collision-induced dissociation fragmentation was performed in the ion trap to insure both maximum sensitivity and a maximum amount of MS/MS data. The signal threshold for an MS/MS event was set to 500 counts. Charge state screening was enabled to exclude precursors with 0 and 1 charge states. Dynamic exclusion was enabled with a repeat count of 1, exclusion list size of 500 MS runs and exclusion duration of 30 s.

The acquired raw data were processed using Proteome Discoverer (version 1.4.1.14, Thermo Fisher Scientific). Spectra were searched using SEQUEST (Thermo Fisher Scientific) against a homemade database comprising 20 150 human and 4306 E. coli sequences and the N-MADD-4B sequence. To detect contaminants, an additional search against the Swiss-Prot database (https://www.uniprot.org/uniprot; version 2014.02; 20 284 entries) supplemented with a set of 245 frequently observed contaminants (e.g., human keratin) was performed using MASCOT (http://www.matrixscience.com). Search parameters were: (i) trypsin; cleavage before Pro and one miscleavage allowed; (ii) mass tolerance of 6 ppm for monoisotopic precursor ions and 0.8 ppm for fragment ions from MS/MS; (iii) Cys carbamidomethylation (+57.02146 Da) as a fixed modification; Met oxidation (+15.99491 Da) and N-terminal acetylation (+42.0106 Da) as variable modifications; (iv) minimum peptide length of five residues. Only high-score peptides were selected. Proteins were identified with a false discovery rate (FDR) of 1%.

Intact protein masses - Protein masses were determined on solution samples. 1 µl of protein at ca. 20 µM was mixed with 1 µl of a saturated solution of sinapinic acid in 0.1% TFA / CH_3CN (50:50 v/v). 1 µl of the mix was spotted on the target and analysed by MALDI-TOF-TOF on a Ultraflex III spectrometer (Bruker Daltonics, Wissembourg, France) controlled by the Flexcontrol 3.0 package (Build 51) and operated in the linear mode, using a maximum accelerating potential of 25 kV and a 20 000-80 000 m/z range (LP_66kDa Method). The laser frequency was fixed to 100 Hz and ca. 1000 shots per sample were cumulated. Four external standards (Protein Calibration Standard II, Bruker Daltonics) were used to calibrate each spectrum to a mass accuracy within 100 ppm. Peak picking was performed using the FlexAnalysis 3.0 software with an adapted analysis method. Parameters used were: centroid peak detection algorithm, S/N threshold fixed to 5 and a quality factor threshold of 30. Mean masses were calculated from the measured $[M+H]^+$ and $[M+2H]^{2+}$ values.

Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal . Chem.* **68**, 850–858.

Microscale thermophoresis (MST)

Equations used by the NanoTemper MO.Affinity Analysis software v2.2.4

for the Kd model:

$$F(A) = \text{Free} + \frac{(\text{Bound} - \text{Free}) * (A + B + Kd - \sqrt{(A + B + Kd)^2 - 4 * A * B})}{2 * B}$$

for the Hill model:

$$F(A) = \text{Free} + \frac{(\text{Bound} - \text{Free})}{1 + (\frac{EC50}{A})^n}$$

where F is the fluorescence signal as a function of A; A is the *titrated protein* concentration, which varies within a selected range; B is the *labelled protein* total concentration, which is kept constant; Bound and Free are the *labelled protein* concentrations in the bound and unbound states, respectively; and n is the Hill coefficient.

End of Supporting information