nature research

Corresponding author(s): DBPR and NCHEMB-A190609720

Last updated by author(s): Nov 12, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The protein crystals were diffracted and collected at Diamond (Oxford). SAXS data was collected at Soleir synchrotron.
Data analysis	NMR, AFM, crystallography and SAXS data were processed as described in the Online Methods. For X-ray data, we used CCP4 software, XDS, Refmac5, ARP/wARP and Procheck. For SAXS data we used FOXTROT, ATSAS 2.7.3, Ranch2.0, EOM2.0 and DADIMODO. For STD-NMR data, we used the computer Aided Resonance Assignment (CARA) software. For AFM data, we used WSxM software. For Kinetics determinations, we used GraphPad Prism 7.03 All the software used in this manuscript are published and references to the different programs are cited in the manuscript.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The datasets generated and/or analysed during this study are available from the corresponding author on reasonable request. Structure datasets generated during the current study are available in the PDB repository under accession numbers 6S24 and 6S22.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. We used three independent experiments and in other experiments we performed a larger number of experiments.
Data exclusions	none
Replication	all attempts for replication were successful
Randomization	not applicable - unnecessary for this work
Blinding	not applicable - unnecessary for this work

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Primary antibodies used: anti 6xHis tag (Thermo, Cat#MA1-21315, lot#TK269015), anti-myc tag (Invitrogen, Cat#MA1-980, 9E10), anti GALNT2 (4C4), anti GALNT3 (2D10), anti GALNT4 (4G2) Secondary antibodies used: HRP-conjugated rabbit anti mouse Ig (Dako, P0260, lot20051347), FITC-labelled rabbit anti mouse Ig (Dako, P0261, lot#20043928)
Validation	Commercially purchased antibodies against standard protein tags were tested on cell lines transfected with proteins with or without the relevant tag. In house antibodies against GALNT enzymes were validated by loss of staining in the KO in relation to positive staining. See PMIDs 31172184 and 31040225 for further information.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	CHOZN®GS-/- ZFN-modified CHO cell line (Sigma-Aldrich), CHO-K1 (ATCC) and CHO-ldlD , HEK293F and SMD1168 were used.
Authentication	All these strains are commercial
Mycoplasma contamination	A representative set of growing cell lines in the lab selected randomly is subjected to mycoplasma screening bi-monthly, and within the last 10 yrs no infected cells have been found.
Commonly misidentified lines (See ICLAC register)	None of the cell lines used are listed in the ICLAC database.