

Synthesis of a glycan hairpin

Received: 21 June 2022

Accepted: 26 May 2023

Published online: 3 July 2023

Check for updates

Giulio Fittolani^{1,2,7}, Theodore Tyrikos-Ergas^{1,2,8}, Ana Poveda³, Yang Yu^{1,2,9},
Nishu Yadav^{1,2}, Peter H. Seeberger^{1,2}, Jesús Jiménez-Barbero^{3,4,5,6}
& Martina Delbianco¹✉

The primary sequence of a biopolymer encodes the essential information for folding, permitting to carry out sophisticated functions. Inspired by natural biopolymers, peptide and nucleic acid sequences have been designed to adopt particular three-dimensional (3D) shapes and programmed to exert specific functions. In contrast, synthetic glycans capable of autonomously folding into defined 3D conformations have so far not been explored owing to their structural complexity and lack of design rules. Here we generate a glycan that adopts a stable secondary structure not present in nature, a glycan hairpin, by combining natural glycan motifs, stabilized by a non-conventional hydrogen bond and hydrophobic interactions. Automated glycan assembly enabled rapid access to synthetic analogues, including site-specific ¹³C-labelled ones, for nuclear magnetic resonance conformational analysis. Long-range inter-residue nuclear Overhauser effects unequivocally confirmed the folded conformation of the synthetic glycan hairpin. The capacity to control the 3D shape across the pool of available monosaccharides has the potential to afford more foldamer scaffolds with programmable properties and functions.

Linear amino acid sequences encode the information needed to generate three-dimensional (3D) motifs such as helices, sheets and turns, which can combine to give complex functional macromolecules (that is, proteins). Using the same alphabet of 20 α -amino acids, it is possible to design molecules with engineered 3D structures¹. The development of unnatural amino acids expanded the catalogue even further, and molecules with size, shape and arrangement beyond the natural ones could be generated (that is, foldamers)^{2–5}. Following the principle ‘form follows function’, foldamers were designed to recognize specific ligands⁶ or penetrate cell membranes⁴ as well as to perform catalysis^{7,8} or undergo supramolecular assembly⁹. Moreover, these designed systems largely improved our understanding of the factors influencing the stability of protein folding^{1,10}.

In contrast, glycan structures capable of folding into defined 3D shapes have never been designed, despite the potential numerous advantages. Glycans are constructed from a vast alphabet of more than 100 natural monosaccharides that generates linear or branched polymers, suggesting an enormous potential to access new conformations. The presence of many hydroxyl groups with well-defined orientations provides opportunities for site-specific functionalization. The tendency of naturally occurring polysaccharides to self-assemble into hierarchically organized materials is another attractive feature¹¹. However, so far, the complex chemical synthesis of glycans and the lack of design principles have prevented access to glycans with predictable 3D shapes. Glycans are historically considered flexible molecules rarely exhibiting stable secondary structures¹², with few exceptions for

¹Department of Biomolecular Systems, Max Planck Institute of Colloids and Interfaces, Potsdam, Germany. ²Department of Chemistry and Biochemistry, Freie Universität Berlin, Berlin, Germany. ³CICbioGUNE, Basque Research and Technology Alliance, Derio, Spain. ⁴Ikerbasque, Basque Foundation for Science, Bilbao, Spain. ⁵Department of Organic Chemistry II, Faculty of Science and Technology, University of the Basque Country, Leioa, Spain. ⁶Centro de Investigación Biomedica en Red de Enfermedades Respiratorias, Madrid, Spain. ⁷Present address: Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA. ⁸Present address: Department of Chemistry, University of Illinois, Urbana, IL, USA. ⁹Present address: Simpson Querrey Institute, Northwestern University, Evanston, IL, USA. ✉e-mail: martina.delbianco@mpikg.mpg.de

polysaccharides capable of adopting helix- (for example, amylose) or ribbon-like (for example, cellulose) conformations in solution¹³. Nevertheless, the central role of glycans in a biological setting is often tightly related to the conformation and presentation they adopt¹⁴. Recent studies report increasing evidence that even small glycans can adopt relatively well-defined conformation in solution, often stabilized by intramolecular hydrogen bonds^{15–25}. The increased understanding of glycan behaviour suggests that glycans capable of adopting stable 3D architectures can be created.

In this Article, we rationally designed a glycan capable of folding autonomously into a secondary structure motif to challenge the common view that sees glycans exclusively as flexible molecules^{12,26}. Inspired by peptide model systems, we constructed a glycan that adopts a hairpin secondary structure motif in aqueous solution. The design was based on the combination of natural glycan structural elements and aided by molecular dynamics (MD) simulations. Automated glycan assembly (AGA) provided rapid access to a series of well-defined glycan sequences, including ¹³C-labelled analogues to facilitate the structural and conformational analysis. Nuclear magnetic resonance (NMR) spectroscopy detected long-range inter-residue nuclear Overhauser effects (NOEs) that unequivocally confirmed the folded conformation of the synthetic glycan hairpin. This work demonstrates that it is possible to program glycans adopting defined conformation in aqueous solution²⁷, opening opportunities for endowing glycans with new properties and functions. Analogous to the discovery of peptide-based foldamers, we envision applications of folded glycans in several areas, including materials science^{28,29}, biology³⁰ and catalysis³¹.

Results and discussion

Rational design of a glycan hairpin

Peptide model systems have largely improved our understanding of the factors influencing protein folding and stability^{1,10}. This is especially true in the case of protein β -sheets, for which folding rules could be deduced from model systems consisting of a hairpin (that is, strand–loop–strand). In a hairpin, the β -turn acts as a β -sheet nucleator, placing the two strands adjacent to each other and favouring inter-strand interactions in a parallel³² or antiparallel³³ fashion (Fig. 1a). Intramolecular hydrogen bonds and side-chain interactions stabilize the turn unit and hold the two strands in a closed hairpin conformation (Fig. 1a). Model systems such as double^{34,35} or triple-stranded^{36,37} β -hairpins served as blueprints to gain molecular insights into protein folding^{32,33,38–40} and have proven essential to study structural aspects of amyloid aggregation⁴¹. Beyond applications in biology, these systems have been employed to generate self-assembling nanomaterials^{42–44} and hydrogels^{45,46}, or as catalysts for asymmetric transformations^{47,48}.

Given the breadth of the research on peptide hairpins, we decided to explore the design of an analogous secondary structure motif in glycans. The molecular nature of glycans, composed of hydroxyl-rich backbones, makes the design of a glycan folded secondary structure highly challenging in comparison with peptides and imposes the exploration of new design rules. Interactions with water are strong and drastically impact the conformation in solution, disrupting intramolecular hydrogen bonds. Long-range inter-residue interactions are scarce, and the absence of hydrophobic side chains limits the stability of glycan secondary structures¹². Still, within the glycan realm, subtle non-covalent interactions can take place (Fig. 1b), which have a large impact on macroscopic properties⁴⁹ or bioactivity¹⁴. These non-covalent interactions, combined with design rules developed for peptides, served as inspiration for the creation of a glycan hairpin.

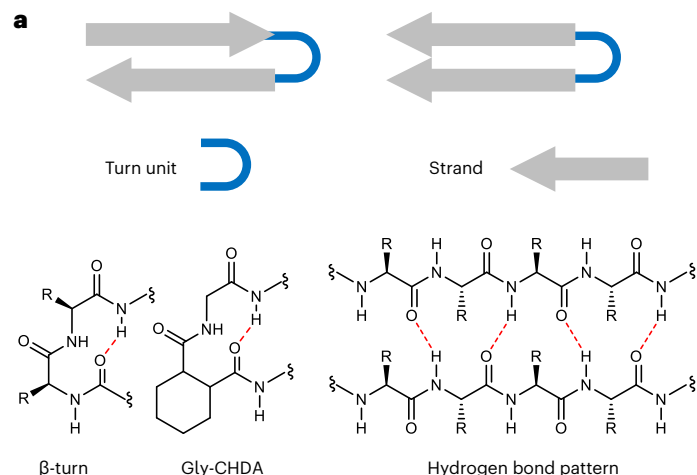
A parallel glycan hairpin structure, composed of two strands attached to a turn, is a secondary structure motif absent in nature. The design requires (1) a suitable glycan loop adopting a turn conformation and (2) two stacking strands⁵⁰. The similarity of the Lewis X (Le^x , α -L-Fuc-(1,3)-[β -D-Gal-(1,4)]-D-GlcNAc) trisaccharide conformation to that of naturally occurring peptide β -turns prompted us to

explore this glycan motif as a starting point for the design of the turn unit. Le^x adopts a fairly rigid closed conformation in aqueous solution stabilized by hydrophobic interactions between the methyl group of L-Fuc and the β -face of Gal. Further stabilization arises due to a non-conventional $\text{CH}\cdots\text{O}$ hydrogen bond⁵¹ between the H-5 of L-Fuc (L-Fuc-5) and the O-5 of Gal branches (Fig. 1b)^{19,20,24}. This $\text{CH}\cdots\text{O}$ hydrogen bond forms a ten-membered atom ring analogous to naturally occurring peptide β -turns. Looking at the 3D shape of Le^x , we noticed that further monosaccharide extensions at the OH-4 positions of the stacked residues could keep a parallel arrangement between the two new moieties, provided that OH-4 displays an equatorial orientation²⁰. Therefore, the canonical Le^x trisaccharide loop was modified by converting β -1,4-Gal into β -1,4-Glc and α -1,3-L-Fuc into α -1,3-L-Rha, both presenting OH-4 in the equatorial orientation (Fig. 1c, **3mer-I**). We hypothesized that the non-conventional hydrogen bond could still act as a key stabilizing force to keep in spatial proximity the two branches²⁰. Two additional turn units were also designed: (1) **3mer-II** is based on GlcNAc substituted with two Glc residues attached as β -1,4 and β -1,3 branches, lacking the stabilizing non-conventional H-bond, and (2) **3mer-III**, in which the branched GlcNAc residue was substituted by Glc, ideally preserving an optimal spatial arrangement of the two branches to engage in the $\text{CH}\cdots\text{O}$ hydrogen bond (Fig. 1c).

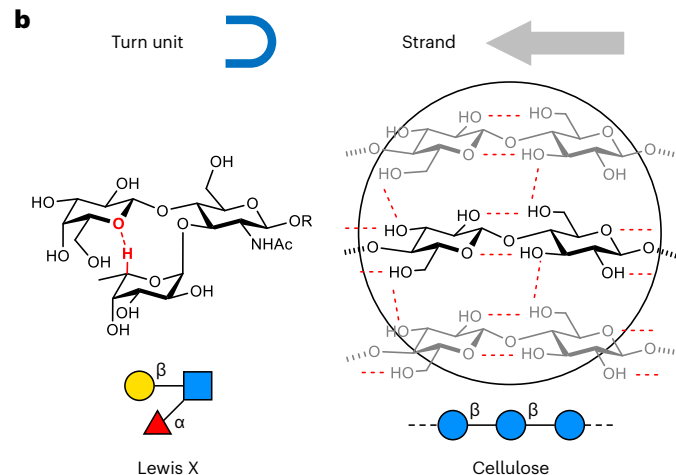
As model for the stacking strands, we were inspired by the cellulose backbone, a polysaccharide consisting of a β -1,4-Glc repeating sequence (Fig. 1b). Cellulose is a linear polysaccharide adopting a rigid rod conformation in solution stabilized by intramolecular hydrogen bonds between the OH-3 and O-5 of neighbouring Glc residues. Strong intermolecular interactions such as hydrogen bonds and hydrophobic interactions between the C-H-rich faces of Glc drive amphiphilic cellulose chains to self-assemble into highly ordered (insoluble) crystallites^{49,52}. Thus, we selected oligomers of cellulose as strand sequences due to their propensity to assemble in a regular manner^{53,54}.

Overall, our design includes a rigid turn-like glycan, stabilized by a non-conventional $\text{CH}\cdots\text{O}$ hydrogen bond, carrying two cellulose oligomer strands. Natural glycan structural elements are thus combined to generate an unnatural glycan hairpin. As target structures, we set a short pentasaccharide (**5mer-I**) and two nonasaccharide analogues (**9mer-I** and **9mer-II**) (Fig. 1d).

Atomistic MD simulations were carried out to screen the tendency of the designed glycans to adopt the desired conformation. All the modelled structures were simulated for 500 ns, employing a modified version of the GLYCAM06 (ref. 55) carbohydrate force field. The systems were solvated with TIP5P⁵⁶ water model to avoid excessive interactions between the monomers⁵⁷. We began by comparing **3mer-I**, **3mer-II** and **3mer-III** to identify the optimal sequence for the turn unit. The MD trajectory suggested a clear difference in overall conformation between **3mer-I** and **3mer-II**. The latter displayed enhanced flexibility and extended conformation compared to **3mer-I**, as clearly indicated by the large fluctuations in the root-mean-square-deviation plot and the higher value of the radius of gyration (R_g) (Supplementary Fig. 7). In contrast, **3mer-I** adopted a more rigid and compact turn-like conformation with a single conformer dominating nearly all the simulation time (Fig. 2a, top). MD simulations pointed at typical exo-syn- Φ /syn- Ψ conformations for all glycosidic linkages in **3mer-I** (Fig. 2a, bottom, and Supplementary Fig. 8). On the other hand, Ramachandran plots for **3mer-II** suggested a two-state situation (that is, closed versus open conformation; Fig. 2a, top, and Supplementary Fig. 12) with the β -1,4 glycosidic linkage between Glc and GlcNAc fluctuating between exo-syn- Φ /syn- Ψ and non-exo- Φ /syn- Ψ conformations (Fig. 2a, bottom, and Supplementary Fig. 9). Similarly to **3mer-I**, **3mer-III** displayed a rigid closed conformation albeit slightly more flexible, probably due to the lack of the NHAc moiety (Supplementary Figs. 7, 10 and 11). Importantly, **3mer-I** displayed the hydroxyls planned for elongation of the two strands in an ideal orientation (Supplementary Fig. 13).

Peptide β -sheet model systems

Natural glycan structural elements



This work: a synthetic glycan hairpin

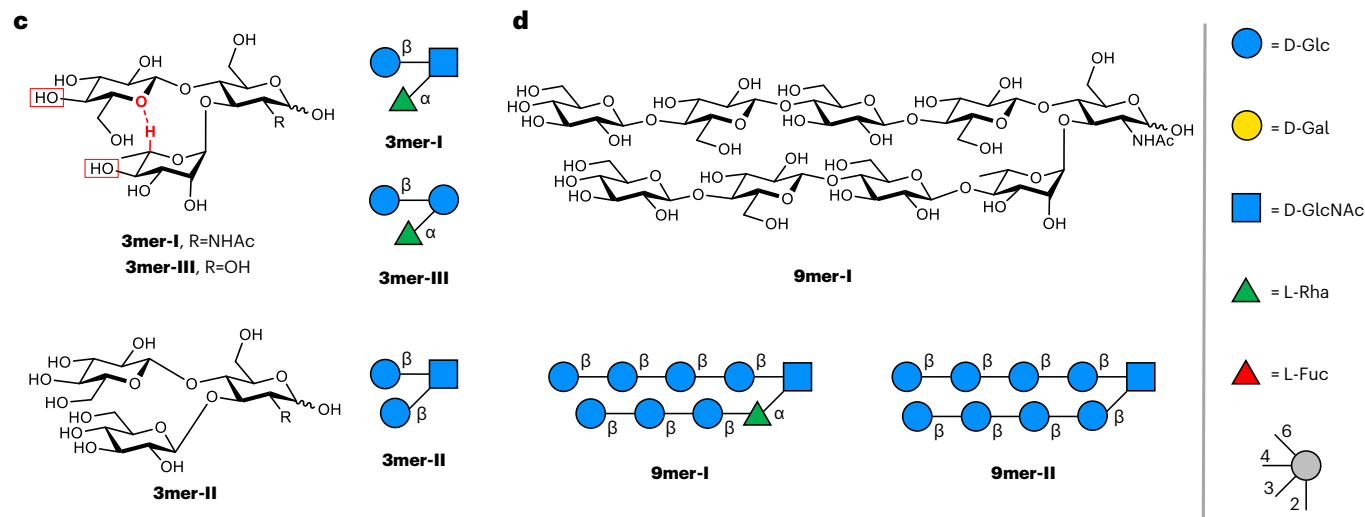


Fig. 1 | Design of a glycan hairpin. a, Autonomously folding antiparallel or parallel peptide β -sheet model systems are composed of a turn unit (for example, β -turn or Gly-CHDA)³² and two strands interacting via hydrogen bonds (red dotted lines, exemplified for parallel β -sheets). **b**, Naturally occurring glycans with turn-like conformation (that is, LeX trisaccharide) and strand-like conformation (that is, cellulose) inspired the design of an artificial glycan hairpin. **c**, The glycan turn unit **3mer-I**, and **3mer-III**, was designed to preserve the non-conventional hydrogen bond (red dotted line) and optimally

align two hydroxyl groups (red boxes) to elongate the strands of the hairpin. **d**, Synthetic glycan hairpin model system composed of two cellulose strands attached as branches to a trisaccharide turn unit. The following abbreviations are used for monosaccharides: Glc, glucose; Gal, galactose; GlcNAc, *N*-acetyl glucosamine; Rha, rhamnose; Fuc, fucose. The monosaccharide residues are represented following the Symbol Nomenclature for Glycans (SNFG) graphical representation⁷⁴.

We then analysed the hairpin analogues, **9mer-I** and **9mer-II**, to monitor how the turn unit impacted the overall conformation of the glycan hairpins. According to MD, only **9mer-I** adopted a closed hairpin conformation with the two cellulose strands in close proximity (Fig. 2b). The turn unit based on **3mer-I** held the two cellulose strands in an optimal orientation resulting in small inter-residue distances across the strands (between 5.0 Å and 7.2 Å; Fig. 2c, left, and Supplementary Fig. 14). In contrast, in **9mer-II** the turn unit based on **3mer-II** did not hold the two strands in proximity and the two cellulose strands tended to stay further apart (up to 18.0 Å; Fig. 2c, left, and Supplementary Fig. 14). This tendency was also reflected in the R_g and RMSD plots, which pointed at a rigid and compact shape for **9mer-I** versus a flexible, more extended shape for **9mer-II** (Fig. 2c, right, and Supplementary Fig. 15). MD simulations suggested also that larger glycan hairpins based on the **3mer-I** turn unit could in principle be obtained, even

though larger fluctuations were observed for the residues located at the non-reducing termini (Supplementary Figs. 16 and 17).

Synthesis of a glycan hairpin

The target compounds were synthesized by AGA (Fig. 3), a solid-phase automated synthetic method⁵⁸. Each oligosaccharide was assembled in an overnight run using protected monosaccharide building blocks (BBs). **BB1** and **BB4** were equipped with two orthogonal protecting groups—levulinoyl ester (Lev) and 9-fluorenylmethoxycarbonyl (Fmoc)—allowing for the preparation of the three different turn units. The synthesis proceeded from the branching monosaccharide towards the non-reducing end, elongating one strand at a time, following cycles of glycosylation (modules C1 or C2; see Supplementary Section 3.3) and Fmoc deprotection (module E1). Once the first strand of the hairpin was constructed using **BB2a** or **BB2b** (refs. 59,60), the terminal hydroxyl

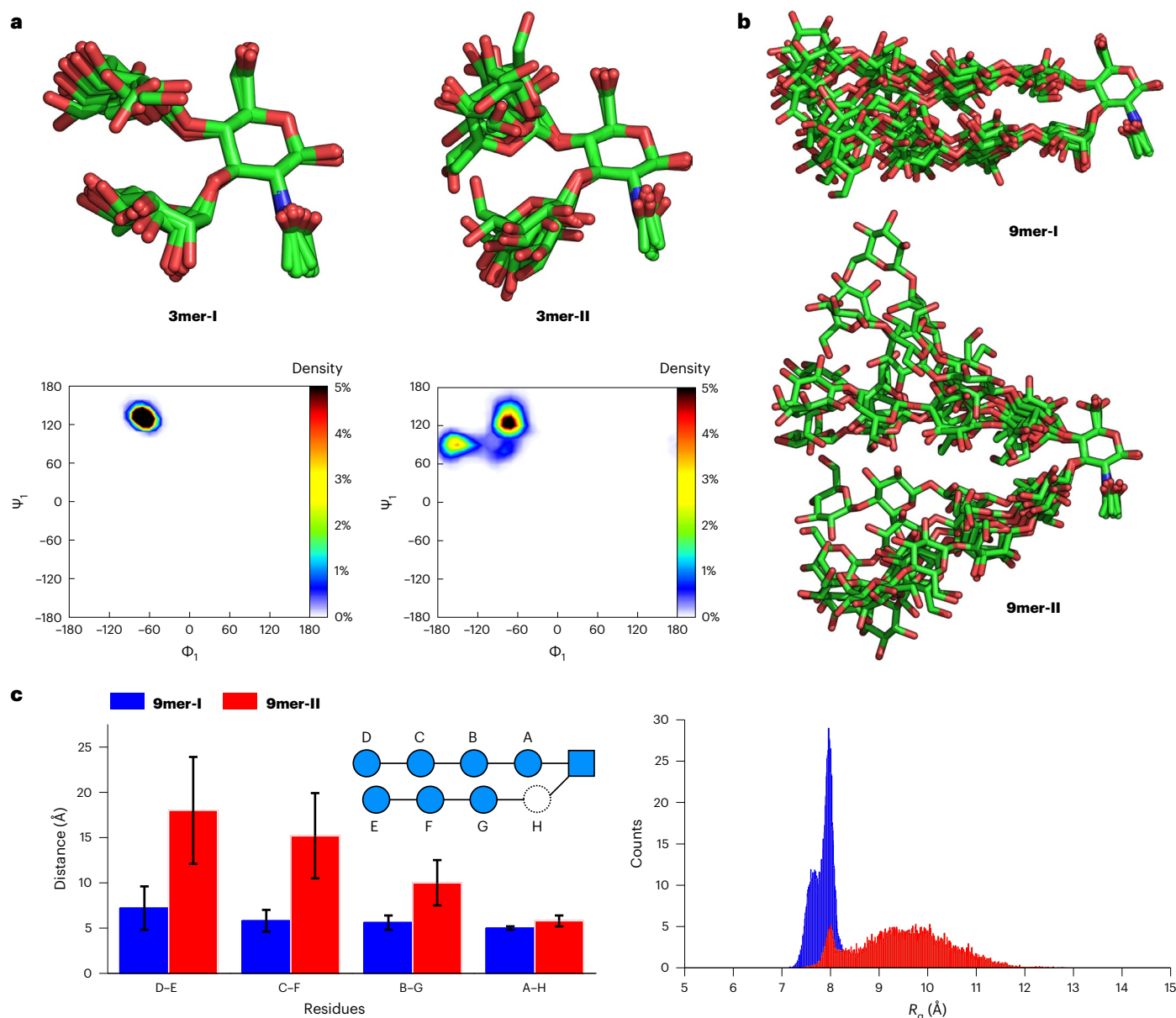


Fig. 2 | MD simulations of the turn units and glycan hairpins.

a, Over-imposition of seven representative snapshot (top) and Ramachandran plots for the β -1,4-Glc linkage (bottom) extracted from the MD simulation show the larger degree of flexibility of **3mer-II** compared with **3mer-I**.

b, Over-imposition of seven representative snapshot shows a closed hairpin conformation for **9mer-I** (top) in contrast to a highly flexible and open

conformation for **9mer-II** (bottom). **c**, Average inter-residue distances (left) and R_g (right) plots showing opposite trends for the two nonasaccharides regarding overall conformation and flexibility. Average inter-residue distances and R_g were calculated from the MD simulation trajectory (25,001 frames, 500 ns). Error bars represent the standard deviation. Overall, MD suggests that a rigid turn is needed to preserve the closed hairpin conformation.

group was capped (module D). Elongation of the second strand proceeded after Lev deprotection (module E2). **BB3a** was first employed for the construction of **5mer-I** and **9mer-I**, but proved to be a poor acceptor during the glycosylation of the L-Rha-OH-4. Low reactivity was also observed when we attempted the elongation of the α -1,3 branch of the hairpin glycan in the absence of the bulky β -1,4 branch, proving that the poor reactivity was not due to steric hindrance (see the AGA of **3mer-IV**, Supplementary Section 3.5). To increase the nucleophilicity of the L-Rha acceptor, we switched to **BB3b**, bearing a C-3 *O*-benzyl group instead of the EWG benzoyl group. The efficiency of the elongation of the second cellulose strand drastically improved, and no major deletion sequence side-products were observed in the synthesis of **5mer-I** and **9mer-I** (Supplementary Figs. 4 and 5). Post-AGA steps included solid-phase methanolysis, photocleavage from the

solid support, and hydrogenolysis (Supplementary Section 3.4). The latter proved increasingly challenging for **5mer-I** and **9mer-I**, requiring multiple hydrogenolysis cycles to convert the *N*-trichloroacetyl protecting group to the *N*-acetyl moiety. The turn units **3mer-I**, **3mer-II** and **3mer-III** as well as the longer hairpin analogues **5mer-I**, **9mer-I** and **9mer-II** were obtained after a single final purification step in overall yields of 10–27%.

NMR structural analysis

The structural analysis has been a major obstacle to the complete description of peptide β -sheet models. Circular dichroism provides simple and rapid information on the overall conformation, but lacks molecular details. X-ray crystallography affords molecular resolution, but fails to represent the range or ensemble of conformations

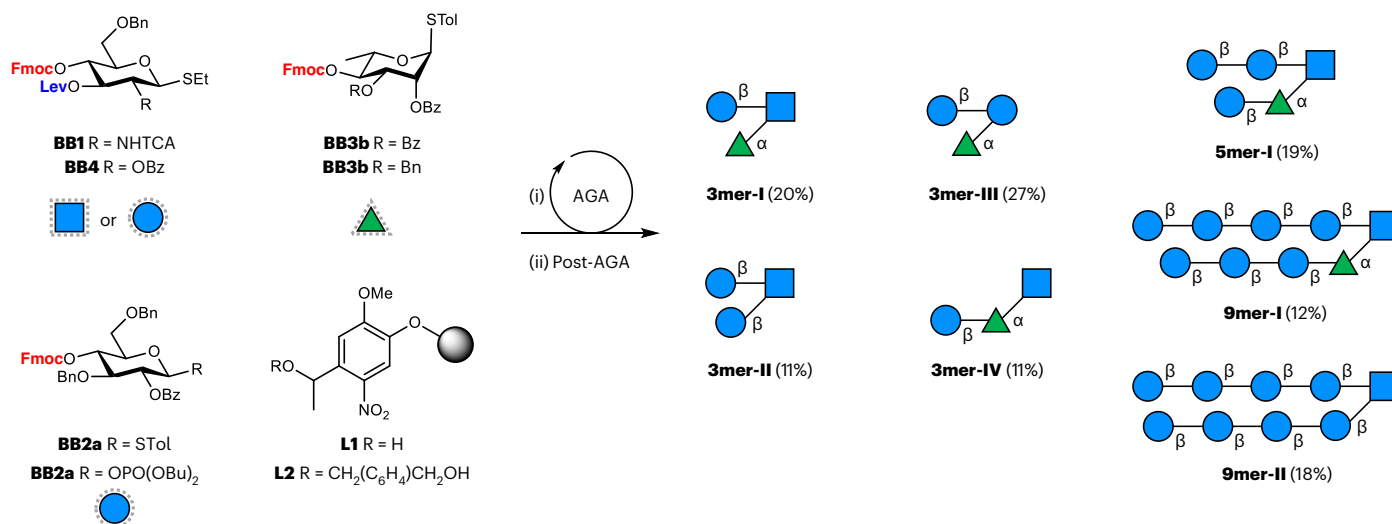


Fig. 3 | Synthesis of glycan structures. The glycan hairpin model structures were prepared by AGA using protected monosaccharide BBs (Supplementary Section 2). Overall yields are reported in parentheses. Reaction conditions for AGA and Post-AGA are reported in Supplementary Section 3.

that may be present in solution. NMR spectroscopy can capture the solution conformation(s) with molecular resolution and is so far the most reliable technique to characterize peptide β -sheet models⁴⁰. Glycan structural studies are even more challenging because the lack of chromophores limits the use of circular dichroism, while difficulties in obtaining single crystals hinder the application of X-ray crystallography to glycans beyond a certain size (that is, tetrasaccharides). NMR may provide some structural information, but suffers from severe overlap between the resonances of different residues and the scarcity of inter-residue and long-range NOEs^{26,61,62}.

A key feature of our turn is the non-conventional CH \cdots O hydrogen bond forming a ten-membered ring that holds the two branches in an ideal parallel orientation. Such non-conventional hydrogen bonding has been hypothesized for rhamnosylated motifs²⁰, but not investigated further^{63,64}. We set out to verify this hypothesis using the chemical shift deviation ($\Delta\delta$) of Rha-5 as an indicator^{18–20}. We compared **3mer-I** and **3mer-III** with **3mer-IV**, in which no hydrogen bond can occur, observing a substantial downfield shift of Rha-5 for the first two compounds ($\Delta\delta = 0.36$ ppm, Fig. 4a; for details, see Supplementary Section 5.2) providing solid ground to our hypothesis.

To confirm the spatial proximity between key residues at both sides of the putative hairpin, nuclear Overhauser effect spectroscopy (NOESY) experiments were employed. The analysis of the NOESY spectra provided additional evidence that **3mer-I** adopts a closed turn-like conformation. In particular, besides the typical NOEs between protons across glycosidic linkages (Glc-1/GlcNAc-4 and Rha-1/GlcNAc-3), we observed key NOEs between the Glc and Rha moieties (Glc-2/Rha-5, Glc-2/Rha-6, Glc-4/Rha-5 and Glc-4/Rha-6) (Supplementary Figs. 29 and 30). Inter-proton distances of 2.9, 3.6, 3.6 and >4 Å, estimated by applying the isolated spin pair approximation to the NOE intensities, satisfactorily matched those calculated in the MD simulations (Supplementary Tables 1 and 2). Similar NOEs between the key proton pairs were also observed for **3mer-III**, suggesting a related conformation (Supplementary Figs. 41 and 42). The NOE pattern observed for **3mer-I** (and **3mer-III**) is fully consistent with previous reports on analogous glycan motifs (including Le^x) (refs. 19,20,63), further corroborating the presence of a closed conformation. In contrast, only the standard NOEs across the glycosidic linkages were observed for **3mer-II**, in agreement with distances calculated in the MD model, indicating that the two branches are further apart compared with **3mer-I** (Fig. 4b and Supplementary Figs. 35 and 36).

The short hairpin model **5mer-I** was then scrutinized. The observed downfield shift of Rha-5 confirmed the presence of the non-conventional hydrogen bond ($\Delta\delta = 0.38$ ppm; Fig. 4a). The non-conventional hydrogen bond was preserved over a wide temperature range (Supplementary Fig. 24). The ROESY⁶⁵ experiment confirmed the presence of the key inter-residue NOEs between Glc A and Rha, identical to those observed for **3mer-I** (Fig. 4c and Supplementary Figs. 47 and 52). These data prove that the closed conformation of the turn unit remains stable upon elongation of the two strands. To confirm that **5mer-I** globally folds into a hairpin conformation, additional evidence of the proximity between Glc B and Glc C on the two strands was required. The NOEs between Glc B-1 and Glc C-2 were evident in the 2D t-ROESY (Fig. 4d and Supplementary Fig. 48) and selective 1D t-ROESY (Supplementary Figs. 49–51), corresponding to an inter-proton distance of ca. 3.0 Å, in full agreement with the MD simulations. As commonly recognized for peptides, the NOEs between residues distant in sequence are a strong indication that **5mer-I** adopts a hairpin conformation (Fig. 4e).

The next challenge was to demonstrate that the hairpin conformation is also kept in longer glycan structures. Thus, long-range inter-residue NOEs were also searched within **9mer-I**, to ultimately prove that this oligomer autonomously folds into a hairpin conformation. However, the severe chemical shift degeneracy hampered any standard structural analysis besides the observation of the hydrogen bond-induced downfield shift of Rha-5 ($\Delta\delta = 0.40$ ppm; Fig. 4a) and a key Rha-6/Glc A-2 NOE confirming the stability of the closed turn unit conformation over a wide temperature range (Fig. 5b and Supplementary Fig. 24).

To break the chemical shift degeneracy, and thus perform the required NMR analysis, two selectively ¹³C-labelled analogues of **9mer-I** were synthesized, employing ¹³C-**BB2b**, using AGA (Supplementary Sections 2.3 and 3.5). **9mer-I-B** and **9mer-I-F** present a ¹³C₆-labelled Glc at residues B and F of the hairpin structure, respectively (Fig. 5a). The use of site-selective labelled analogues allowed extraction of the key inter-residue NOE information otherwise inaccessible. Half-filtered ¹H[¹³C], ¹H[¹²C] NOESY experiments, showing only inter-residue NOEs, demonstrated the presence of the Glc B-1/Glc G-2 medium size NOE, corresponding to an inter-proton distance of 2.9–3.0 Å on **9mer-I-B** (estimated by applying the isolated spin pair approximation to the NOE intensities, Fig. 5c), which was further confirmed by HSQC-NOESY (Supplementary Fig. 58). Next, we recorded HSQC-NOESY on **9mer-I-F**

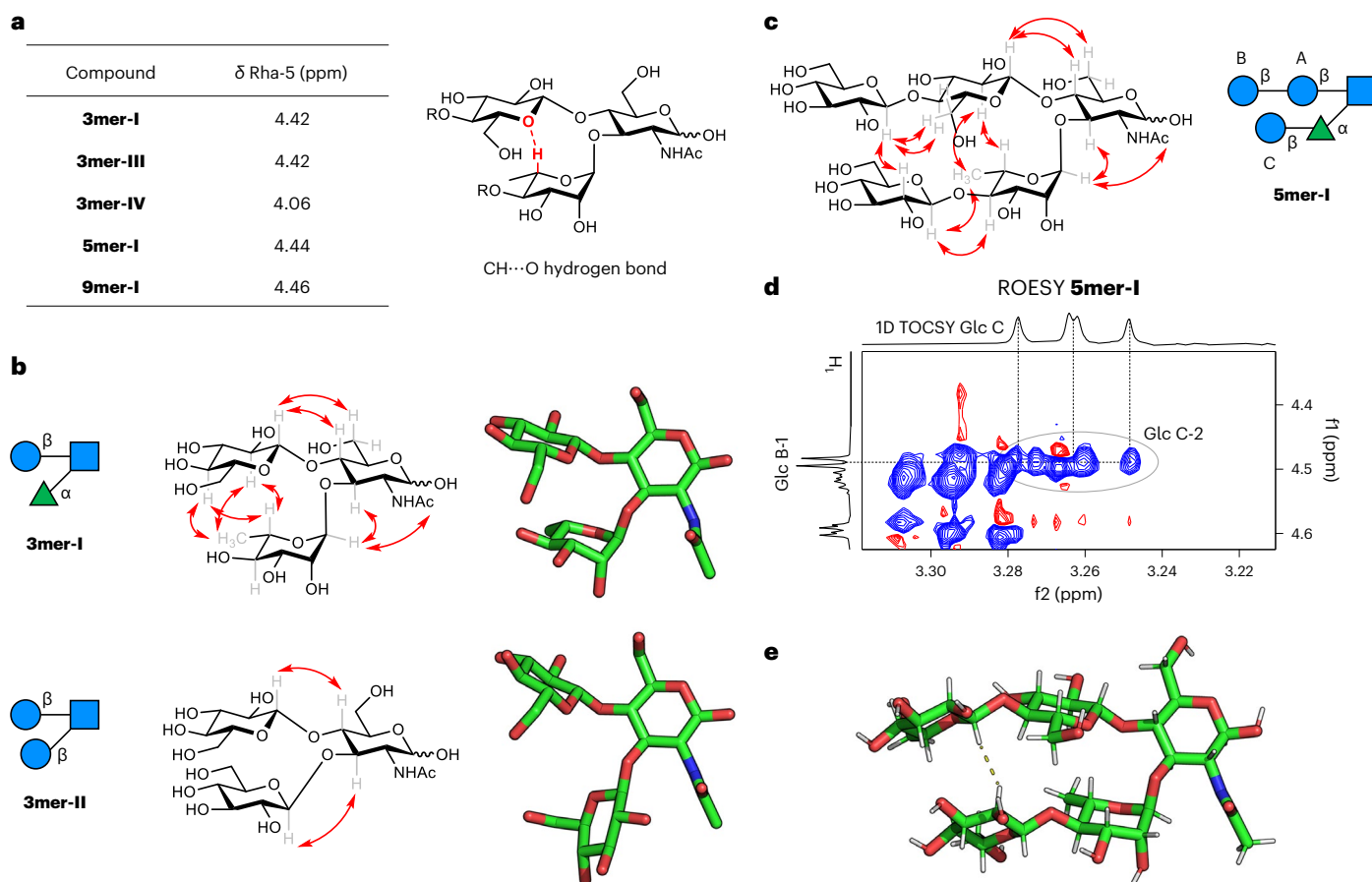


Fig. 4 | NMR analysis of turn units and 5mer-I. **a**, The chemical shift of Rha-5 reflects the presence of a non-conventional CH \cdots O hydrogen bond, confirming the closed conformation of the turn unit in those structures. **b**, Experimental NOEs extracted from NOESY NMR experiments for **3mer-I** and **3mer-II** (red arrows). **c**, Experimental NOEs extracted from ROESY NMR experiments for

5mer-I (red arrows). **d**, Excerpt of the 2D ROESY NMR of **5mer-I** showing the key NOE between the two strands of the hairpin (300 K, mixing time 300 ms, D₂O, 600 MHz). **e**, Snapshot obtained from the MD simulation of **5mer-I** showing a hairpin conformation with the key experimental NOE indicated with a dotted line.

and detected another key medium size NOE (Glc F-6/Glc C-2, Fig. 5d), corresponding to an inter-proton distance of ca. 4.0 Å. These data unambiguously show that, even for the nonasaccharide, the hairpin conformation is kept in solution (Fig. 5e,f).

Further demonstration of the different 3D shape of **9mer-I** versus the **9mer-II** analogue was deduced from diffusion-ordered NMR spectroscopy (DOSY) experiments independently carried out for both molecules at the same concentration (Fig. 5g). A substantial difference in the diffusion coefficients estimated for both molecules ($2.03 \times 10^{-10} \text{ m}^2 \text{ s}^{-1}$ for **9mer-I** and $1.77 \times 10^{-10} \text{ m}^2 \text{ s}^{-1}$ for **9mer-II**) was observed, indicating a much more compact structure for **9mer-I**. According to the Stokes–Einstein equation that describes DOSY experiments, the effective hydrodynamic volume of **9mer-II** is 50% larger than that of **9mer-I** (the ratio of hydrodynamic volumes is 1.52; Supplementary Fig. 64).

Conclusions

We present a glycan that folds into a secondary structure not found in nature. Employing design principles developed for creating peptide hairpin structures, we designed and synthesized a glycan hairpin. We combined natural glycan structural elements to access a turn unit and two stacking strands. The glycan turn unit was inspired by the Le^x trisaccharide motif, adopting a closed conformation stabilized by a non-conventional hydrogen bond. The two strands were based on cellulose chains, adopting a rigid rod conformation stabilized by intramolecular hydrogen bonds and hydrophobic interactions.

Atomistic MD simulations aided the design, by providing 3D models of the glycans as a blueprint for synthesis and structural analysis. A collection of synthetic hairpin analogues, including ¹³C-labelled oligosaccharides, was rapidly synthesized by AGA. The detection of long-range NOEs unambiguously demonstrated the hairpin conformation adopted by the nonasaccharide **9mer-I**, and its shorter analogues, in an aqueous solution. DOSY experiments further confirmed the propensity to adopt a closed and compact conformation.

Our work illustrates that glycan sequences capable of adopting specific secondary structure motifs can be designed. Glycans with predictable folded shapes will expand the catalogue of foldamer scaffolds. The vast pool of monosaccharides available might result in many structural motifs, beyond the commonly known natural geometries. Previously unknown natural polysaccharides are continuously discovered, and their sequences could provide a source of secondary structure motifs^{66,67}. In addition, non-canonical monosaccharides⁶⁸ and/or glycomimetics^{69,70} could be used to bring in proximity different glycan motifs or to induce folding of even shorter glycans⁷¹. Improvements in glycan synthesis will open up opportunities to design folded glycans⁷², allowing for the synthesis of long polysaccharides, for the construction of distinct linkages and for the production of larger quantities of materials⁷³. We imagine that the ability to control the conformation of glycans could lead to predictable functions and properties, with applications in catalysis and nanotechnology.

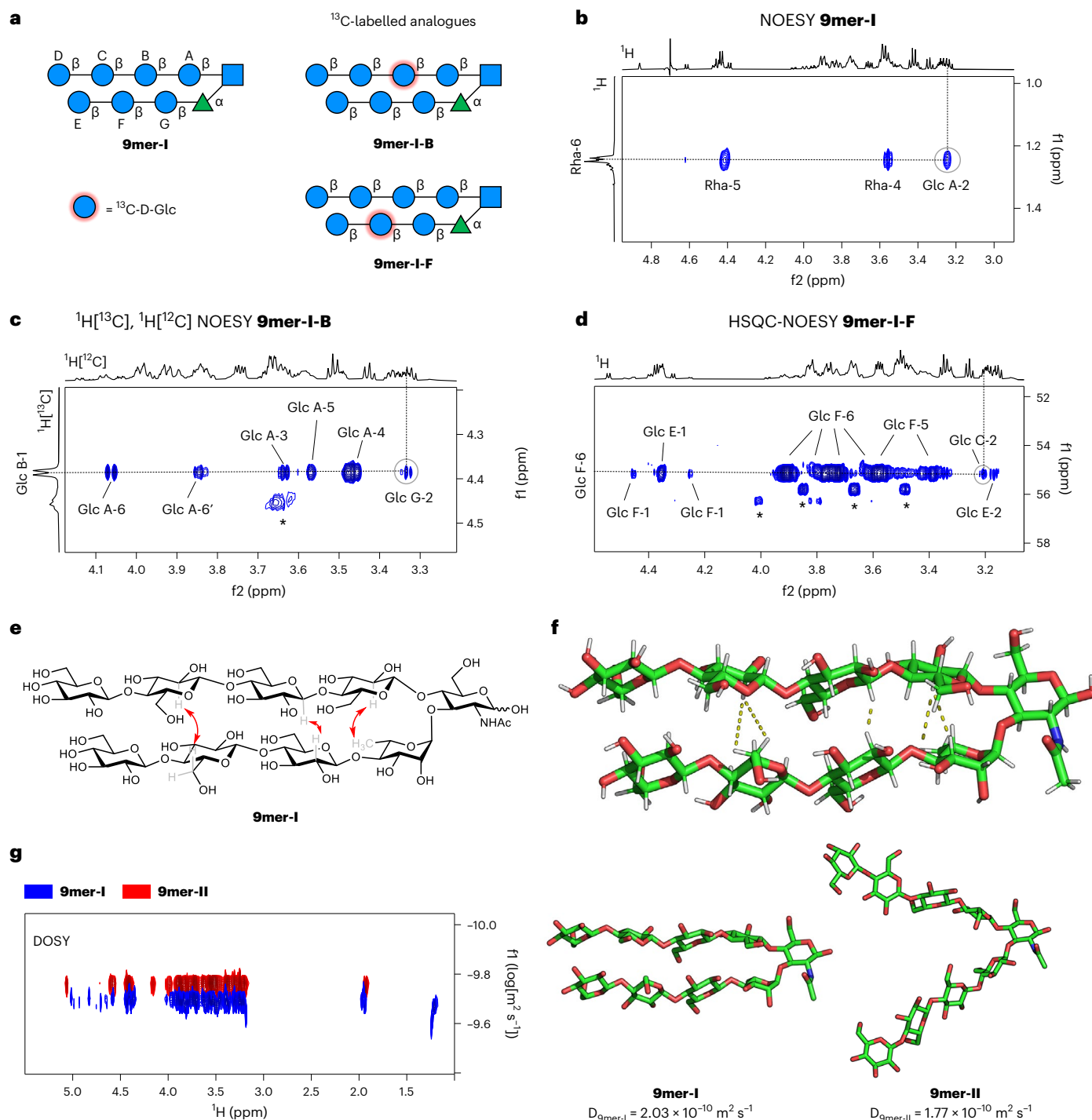


Fig. 5 | NMR analysis of the glycan hairpin. a, AGA enabled the rapid synthesis of two labelled analogues of **9mer-I**. For details about BBs, reaction conditions and AGA, see Supplementary Sections 2.3 and 3.5. **b**, Excerpt of the 2D NOESY NMR of **9mer-I** showing the key NOE (Rha-6/Glc A-2) at the turn unit (298 K, mixing time 300 ms, D₂O, 800 MHz). **c**, Excerpt of the half-filtered 2D ¹H[¹³C], ¹H[¹²C] NOESY NMR of **9mer-I-B** showing the typical inter-residue NOEs across the glycosidic linkages and the key NOE (Glc B-1/Glc G-2) between the two strands of the hairpin (298 K, mixing time 300 ms, D₂O, 800 MHz). The asterisk indicates signals belonging to an impurity. **d**, Excerpt of the 2D HSQC-NOESY NMR of **9mer-I-F** showing the expected intra-residue NOEs and the key NOE

(Glc F-6/Glc C-2) between the two strands of the hairpin (298 K, mixing time 300 ms, D₂O, 800 MHz). The asterisks indicate signals belonging to an impurity (Supplementary Figs. 67 and 68). **e**, Selected experimental NOEs extracted from NMR conformational experiments for **9mer-I** (red arrows) that confirm its folded conformation. **f**, Snapshot of the MD simulation of **9mer-I** showing a hairpin conformation with selected experimental NOEs indicated with dotted lines. **g**, Superimposed DOSY spectra of **9mer-I** (blue) and **9mer-II** (red) (295 K, D₂O, 600 MHz). The corresponding diffusion coefficients are indicated, confirming that the effective hydrodynamic volume of **9mer-II** is 50% larger than that of **9mer-I**.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41557-023-01255-5>.

References

1. Venkatraman, J., Shankaramma, S. C. & Balam, P. Design of folded peptides. *Chem. Rev.* **101**, 3131–3152 (2001).
2. Gellman, S. H. Foldamers: a manifesto. *Acc. Chem. Res.* **31**, 173–180 (1998).
3. Guichard, G. & Huc, I. Synthetic foldamers. *Chem. Commun.* **47**, 5933 (2011).
4. Goodman, C. M., Choi, S., Shandler, S. & DeGrado, W. F. Foldamers as versatile frameworks for the design and evolution of function. *Nat. Chem. Biol.* **3**, 252–262 (2007).
5. Hill, D. J., Mio, M. J., Prince, R. B., Hughes, T. S. & Moore, J. S. A field guide to foldamers. *Chem. Rev.* **101**, 3893–4012 (2001).
6. Ferrand, Y. & Huc, I. Designing helical molecular capsules based on folded aromatic amide oligomers. *Acc. Chem. Res.* **51**, 970–977 (2018).
7. Girvin, Z. C., Andrews, M. K., Liu, X. & Gellman, S. H. Foldamer-templated catalysis of macrocycle formation. *Science* **366**, 1528–1531 (2019).
8. Girvin, Z. C. & Gellman, S. H. Foldamer catalysis. *J. Am. Chem. Soc.* **142**, 17211–17223 (2020).
9. Yoo, S. H. & Lee, H.-S. Foldectures: 3D molecular architectures from self-assembly of peptide foldamers. *Acc. Chem. Res.* **50**, 832–841 (2017).
10. Gellman, S. H. Minimal model systems for β sheet secondary structure in proteins. *Curr. Opin. Chem. Biol.* **2**, 717–725 (1998).
11. Diener, M. et al. Primary, secondary, tertiary and quaternary structure levels in linear polysaccharides: from random coil, to single helix to supramolecular assembly. *Biomacromolecules* **20**, 1731–1739 (2019).
12. Woods, R. J. Predicting the structures of glycans, glycoproteins, and their complexes. *Chem. Rev.* **118**, 8005–8024 (2018).
13. Fittolani, G., Seeberger, P. H. & Delbianco, M. Helical polysaccharides. *Pept. Sci.* **112**, e24124 (2020).
14. Topin, J. et al. The hidden conformation of Lewis X, a human histo-blood group antigen, is a determinant for recognition by pathogen lectins. *ACS Chem. Biol.* **11**, 2011–2020 (2016).
15. Wyss, D. F. et al. Conformation and function of the N-linked glycan in the adhesion domain of human CD2. *Science* **269**, 1273–1278 (1995).
16. Canales, A. et al. Breaking the limits in analyzing carbohydrate recognition by nmr spectroscopy: resolving branch-selective interaction of a tetra-antennary N-glycan with lectins. *Angew. Chem. Int. Ed.* **56**, 14987–14991 (2017).
17. Delbianco, M. et al. Well-defined oligo- and polysaccharides as ideal probes for structural studies. *J. Am. Chem. Soc.* **140**, 5421–5426 (2018).
18. Zhang, Y. et al. Synthesis and structural analysis of *Aspergillus fumigatus* galactosaminogalactans featuring A-galactose, A-galactosamine and A-N-acetyl galactosamine linkages. *Angew. Chem. Int. Ed.* **59**, 12746–12750 (2020).
19. Zierke, M. et al. Stabilization of branched oligosaccharides: Lewis X benefits from a nonconventional C–H \cdots O hydrogen bond. *J. Am. Chem. Soc.* **135**, 13464–13472 (2013).
20. Aeschbacher, T. et al. A secondary structural element in a wide range of fucosylated glycoepitopes. *Chemistry* **23**, 11598–11610 (2017).
21. Zhang, Q. et al. Synthetic, zwitterionic Sp1 oligosaccharides adopt a helical structure crucial for antibody interaction. *ACS Cent. Sci.* **5**, 1407–1416 (2019).
22. Battistel, M. D., Shangold, M., Trinh, L., Shiloach, J. & Freedberg, D. I. Evidence for helical structure in a tetramer of A2-8 sialic acid: unveiling a structural antigen. *J. Am. Chem. Soc.* **134**, 10717–10720 (2012).
23. Yu, Y. & Delbianco, M. Conformational studies of oligosaccharides. *Chemistry* **26**, 9814–9825 (2020).
24. Battistel, M. D., Azurmendi, H. F., Frank, M. & Freedberg, D. I. Uncovering nonconventional and conventional hydrogen bonds in oligosaccharides through NMR experiments and molecular modeling: application to sialyl Lewis-X. *J. Am. Chem. Soc.* **137**, 13444–13447 (2015).
25. Lycknert, K., Edebrink, P. & Widmalm, G. A conformational carbohydrate scaffold is present in the short-chain lipopolysaccharides of *Moraxella catarrhalis*. *Angew. Chem. Int. Ed.* **43**, 2288–2290 (2004).
26. Freedberg, D. I. & Kwon, J. Solution NMR structural studies of glycans. *Isr. J. Chem.* **59**, 1039–1058 (2019).
27. Fisher, B. F. & Gellman, S. H. Impact of γ -amino acid residue preorganization on α/γ -peptide foldamer helicity in aqueous solution. *J. Am. Chem. Soc.* **138**, 10766–10769 (2016).
28. Su, L., Hendrikse, S. I. S. & Meijer, E. W. Supramolecular glycopolymers: how carbohydrates matter in structure, dynamics, and function. *Curr. Opin. Chem. Biol.* **69**, 102171 (2022).
29. Delbianco, M. & Seeberger, P. H. Materials science based on synthetic polysaccharides. *Mater. Horizons* **7**, 963–969 (2020).
30. Ernst, B. & Magnani, J. L. From carbohydrate leads to glycomimetic drugs. *Nat. Rev. Drug Discov.* **8**, 661–677 (2009).
31. Breslow, R. & Dong, S. D. Biomimetic reactions catalyzed by cyclodextrins and their derivatives. *Chem. Rev.* **98**, 1997–2012 (1998).
32. Freire, F. et al. Diacid linkers that promote parallel β -sheet secondary structure in water. *J. Am. Chem. Soc.* **130**, 7839–7841 (2008).
33. Khakshoor, O. & Nowick, J. S. Artificial β -sheets: chemical models of β -sheets. *Curr. Opin. Chem. Biol.* **12**, 722–729 (2008).
34. Blanco, F. J. et al. NMR evidence of a short linear peptide that folds into a β -hairpin in aqueous solution. *J. Am. Chem. Soc.* **115**, 5887–5888 (1993).
35. Blanco, F. J., Rivas, G. & Serrano, L. A short linear peptide that folds into a native stable β -hairpin in aqueous solution. *Nat. Struct. Mol. Biol.* **1**, 584–590 (1994).
36. Kortemme, T., Ramírez-Alvarado, M. & Serrano, L. Design of a 20-amino acid, three-stranded β -sheet protein. *Science* **281**, 253–256 (1998).
37. Schenck, H. L. & Gellman, S. H. Use of a designed triple-stranded antiparallel β -sheet to probe β -sheet cooperativity in aqueous solution. *J. Am. Chem. Soc.* **120**, 4869–4870 (1998).
38. Fisk, J. D. & Gellman, S. H. A parallel β -sheet model system that folds in water. *J. Am. Chem. Soc.* **123**, 343–344 (2001).
39. Fisk, J. D., Powell, D. R. & Gellman, S. H. Control of hairpin formation via proline configuration in parallel β -sheet model systems. *J. Am. Chem. Soc.* **122**, 5443–5447 (2000).
40. Nowick, J. S., Smith, E. M. & Pairish, M. Artificial β -sheets. *Chem. Soc. Rev.* **25**, 401–415 (1996).
41. Samdin, T. D., Kreutzer, A. G. & Nowick, J. S. Exploring amyloid oligomers with peptide model systems. *Curr. Opin. Chem. Biol.* **64**, 106–115 (2021).
42. Wei, G. et al. Self-assembling peptide and protein amyloids: from structure to tailored function in nanotechnology. *Chem. Soc. Rev.* **46**, 4661–4708 (2017).
43. Ke, P. C. et al. Half a century of amyloids: past, present and future. *Chem. Soc. Rev.* **49**, 5473–5509 (2020).

44. Levin, A. et al. Biomimetic peptide self-assembly for functional materials. *Nat. Rev. Chem.* **4**, 615–634 (2020).
45. Schneider, J. P. et al. Responsive hydrogels from the intramolecular folding and self-assembly of a designed peptide. *J. Am. Chem. Soc.* **124**, 15030–15037 (2002).
46. Rajagopal, K. & Schneider, J. P. Self-assembling peptides and proteins for nanotechnological applications. *Curr. Opin. Struct. Biol.* **14**, 480–486 (2004).
47. Metrano, A. J. et al. Asymmetric catalysis mediated by synthetic peptides, version 2.0: expansion of scope and mechanisms. *Chem. Rev.* **120**, 11479–11615 (2020).
48. Metrano, A. J. & Miller, S. J. Peptide-based catalysts reach the outer sphere through remote desymmetrization and atroposelectivity. *Acc. Chem. Res.* **52**, 199–215 (2019).
49. Medronho, B., Romano, A., Miguel, M. G., Stigsson, L. & Lindman, B. Rationalizing cellulose (in)solubility: reviewing basic physicochemical aspects and role of hydrophobic interactions. *Cellulose* **19**, 581–587 (2012).
50. Sebaoun, L., Maurizot, V., Granier, T., Kauffmann, B. & Huc, I. Aromatic oligoamide β -sheet foldamers. *J. Am. Chem. Soc.* **136**, 2168–2174 (2014).
51. Newberry, R. W. & Raines, R. T. Secondary forces in protein folding. *ACS Chem. Biol.* **14**, 1677–1686 (2019).
52. Glasser, W. G. et al. About the structure of cellulose: debating the Lindman hypothesis. *Cellulose* **19**, 589–598 (2012).
53. Bernet, B., Xu, J. & Vasella, A. Oligosaccharide analogues of polysaccharides, part 20, NMR analysis of templated cellodextrins possessing two parallel chains: a mimic for cellulose I? *Helv. Chim. Acta* **83**, 2072–2114 (2000).
54. Murty, K. V. S. N., Xie, T., Bernet, B. & Vasella, A. Oligosaccharide analogues of polysaccharides. Part 26. *Helv. Chim. Acta* **89**, 675–730 (2006).
55. Kirschner, K. N. et al. GLYCAM06: a generalizable biomolecular force field. *Carbohydrates. J. Comput. Chem.* **29**, 622–655 (2008).
56. Mahoney, M. W. & Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **112**, 8910–8922 (2000).
57. Sauter, J. & Grafmüller, A. Predicting the chemical potential and osmotic pressure of polysaccharide solutions by molecular simulations. *J. Chem. Theory Comput.* **12**, 4375–4384 (2016).
58. Guberman, M. & Seeberger, P. H. Automated glycan assembly: a perspective. *J. Am. Chem. Soc.* **141**, 5581–5592 (2019).
59. Yu, Y. et al. Systematic hydrogen-bond manipulations to establish polysaccharide structure–property correlations. *Angew. Chem. Int. Ed.* **58**, 13127–13132 (2019).
60. Wilsdorf, M. et al. A traceless photocleavable linker for the automated glycan assembly of carbohydrates with free reducing ends. *Chem. Commun.* **52**, 10187 (2016).
61. Battistel, M. D., Azurmendi, H. F., Yu, B. & Freedberg, D. I. NMR of glycans: shedding new light on old problems. *Prog. Nucl. Magn. Reson. Spectrosc.* **79**, 48–68 (2014).
62. Novakovic, M. et al. The incorporation of labile protons into multidimensional NMR analyses: glycan structures revisited. *J. Am. Chem. Soc.* **143**, 8935–8948 (2021).
63. Broecker, F. et al. Multivalent display of minimal *Clostridium difficile* glycan epitopes mimics antigenic properties of larger glycans. *Nat. Commun.* **7**, 11224 (2016).
64. Martin, C. E., Weishaupt, M. W. & Seeberger, P. H. Progress toward developing a carbohydrate-conjugate vaccine against *Clostridium difficile* ribotype O27: Synthesis of the cell-surface polysaccharide PS-I repeating unit. *Chem. Commun.* **47**, 10260 (2011).
65. Bothner-By, A. A., Stephens, R. L., Lee, J., Warren, C. D. & Jeanloz, R. W. Structure determination of a tetrasaccharide: transient nuclear overhauser effects in the rotating frame. *J. Am. Chem. Soc.* **106**, 811–813 (1984).
66. Macdonald, S. S. et al. A synthetic gene library yields a previously unknown glycoside phosphorylase that degrades and assembles poly- β -1,3-GlcNAc, completing the suite of β -linked GlcNAc polysaccharides. *ACS Cent. Sci.* <https://doi.org/10.1021/acscentsci.1c01570> (2022).
67. Speciale, I. et al. Biopolymer skeleton produced by *Rhizobium radiobacter*: stoichiometric alternation of glycosidic and amidic bonds in the lipopolysaccharide O-antigen. *Angew. Chem.* **132**, 6430–6436 (2020).
68. Tyrikos-Ergas, T., Fittolani, G., Seeberger, P. H. & Delbianco, M. Structural studies using unnatural oligosaccharides: toward sugar foldamers. *Biomacromolecules* **21**, 18–29 (2020).
69. Bernardi, A., Checchia, A., Brocca, P., Sonnino, S. & Zuccotto, F. Sugar mimics: an artificial receptor for cholera toxin. *J. Am. Chem. Soc.* **121**, 2032–2036 (1999).
70. Guzzi, C. et al. Insights into molecular recognition of LewisX mimics by DC-SIGN using NMR and molecular modelling. *Org. Biomol. Chem.* **9**, 7705 (2011).
71. Schneider, J. P. & Kelly, J. W. Templates that induce α -helical, β -sheet, and loop conformations. *Chem. Rev.* **95**, 2169–2187 (1995).
72. Wu, Y., Qiu, Y., Feng, Y. & Stoddart, J. F. Automating glycan assembly in solution. *ACS Cent. Sci.* <https://doi.org/10.1021/acscentsci.2c01043> (2022).
73. Yao, W. et al. Automated solution-phase multiplicative synthesis of complex glycans up to a 1,080-mer. *Nat. Synth.* <https://doi.org/10.1038/s44160-022-00171-9> (2022).
74. Varki, A. et al. Symbol nomenclature for graphical representations of glycans. *Glycobiology* **25**, 1323–1324 (2015).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

Methods

Synthesis

The oligosaccharides were prepared using a home-built synthesizer designed at the Max Planck Institute of Colloids and Interfaces⁵⁸. All details concerning BB synthesis^{75,76}, AGA modules^{60,77} and post-AGA manipulations⁷⁸ can be found in Supplementary Sections 2 and 3.

MD simulations

All-atom MD simulations were performed using gromacs 5.1.2 (ref. 79). The oligosaccharides were modelled using a modified version of GLYCAM06_{OSMO,r14} force field^{55,57}, and the system was solvated with TIP5P⁵⁶ water molecules to avoid excessive interactions between the monomers. The topology was converted to gromacs format using the glycam2gmx.pl script and solvated with 2,100 water molecules using gromacs tools. The systems were kept at a constant temperature of 303 K using a Nosé–Hoover thermostat^{80,81} and at constant pressure of 1 bar with the Parrinello–Rahman barostat^{82,83}. Non-bonded interactions were cut off at 1.4 nm, and long-range electrostatics were calculated using the particle mesh Ewald method⁸⁴. Bonds involving hydrogens were constrained using the LINCS⁸⁵ to allow a 2 fs timestep algorithm; water molecules were kept rigid with SETTLE⁸⁶. After energy minimization (steepest descent algorithm) and before the production run, the systems were equilibrated at 300 K for 50 ns in a canonical (NVT) ensemble (constant number of particles N, volume V and temperature T) and subsequently at 300 K and 1 atm for 50 ns in an isothermal–isobaric (NPT) ensemble (constant number of particles N, pressure P and temperature T). All the modelled structures were simulated for 500 ns. Further details on MD simulations are reported in Supplementary Section 4.

NMR analysis

¹H, ¹³C, HSQC, 1D and 2D TOCSY, 1D and 2D ROESY, 2D NOESY NMR spectra were recorded on Varian 400-MR (400 MHz), Varian 600-NMR (600 MHz), Bruker Biospin AVANCE700 (700 MHz) Bruker AVANCE III 800 (800 MHz) spectrometers. Samples were prepared by dissolving lyophilized samples in D₂O (concentration ~1–6 mM). Proton resonances of the oligosaccharides were assigned using a combination of ¹H, 2D COSY, HSQC, 1D and 2D TOCSY. Selective 1D TOCSY (HOHAHA, pulse program: seldigpzs) spectra were recorded using different mixing times to assign all the resonances (d₉ = 40, 80, 120, 160 and 200 ms). 2D TOCSY (pulse program: mlevphpp) spectra were recorded using different mixing times (d₉ = 80 or 120 ms). Selective 1D t-ROESY (pulse program: selrogp.2) spectra were recorded using different mixing times (p15 = 100, 200, or 300 ms). 2D t-ROESY (pulse program: roesyph.2) and 2D NOESY (pulse program: noesygpshpp) spectra were recorded using different mixing times (p15 = 100, 200 or 300 ms for ROESY and d₈ = 600, 800 or 1,000 ms for NOESY). The full NMR analysis of the glycans reported in the manuscript can be found in Supplementary Section 5.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The authors declare that all data supporting the findings of this study are available within the article and in the supplementary information files. Raw data for NMR analysis and MD simulations can be downloaded from <https://doi.org/10.17617/3.EG7I36>, Edmond. Data are also available from the corresponding author upon request.

References

75. Dallabernardina, P., Schuhmacher, F., Seeberger, P. H. & Pfrengle, F. Mixed-linkage glucan oligosaccharides produced by automated glycan assembly serve as tools to determine the substrate specificity of lichenase. *Chemistry* **23**, 3191–3196 (2017).

76. Le Mai Hoang, K. et al. Traceless photolabile linker expedites the chemical synthesis of complex oligosaccharides by automated glycan assembly. *J. Am. Chem. Soc.* **141**, 9079–9086 (2019).
77. Huang, J.-Y. & Delbianco, M. Recent developments in solid-phase glycan synthesis. *Synthesis* **55**, 1337–1354 (2023).
78. Hurevich, M. et al. Continuous photochemical cleavage of linkers for solid-phase synthesis. *Org. Lett.* **16**, 1794–1797 (2014).
79. Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
80. Hoover, W. G. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697 (1985).
81. Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **81**, 511–519 (1984).
82. Parrinello, M. & Rahman, A. Crystal structure and pair potentials: a molecular-dynamics study. *Phys. Rev. Lett.* **45**, 1196–1199 (1980).
83. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
84. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: an $N \cdot \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
85. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
86. Miyamoto, S. & Kollman, P. A. Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **13**, 952–962 (1992).

Acknowledgements

We thank the Max Planck Society, the MPG-FhG Cooperation Project Glyco3Dysplay and the German Federal Ministry of Education and Research (BMBF, grant number 13XP5114, M.D.) for generous financial support. The CIC bioGUNE team acknowledge Agencia Estatal de Investigación of Spain (MCIN/AEI/10.13039/501100011033) for grants PDI2021-1237810B-C21 and CEX2021-001136-S and CIBERES, an initiative of Instituto de Salud Carlos III (ISCIII, Madrid, Spain). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. We also thank M. Heydenreich (Universität Potsdam), J. Imanol Quintana (CIC bioGUNE) and O. Niemeyer (Max Planck Institute of Colloids and Interfaces) for help in recording NMR spectra. We thank Y. Zhu (Max Planck Institute of Colloids and Interfaces) for providing solid support **L2**.

Author contributions

M.D. and G.F. conceived the project. G.F. optimized the synthetic routes. G.F. and N.Y. performed the syntheses. T.T.-E. performed the MD simulations. G.F., A.P. and J.J.-B. performed the NMR experiments and analysed the data. Y.Y. performed the synthesis of intermediate **S-7**. M.D. supervised the project. G.F. and M.D. wrote the paper with contributions from T.T.-E., A.P., Y.Y., N.Y., P.H.S. and J.J.-B.

Funding

Open access funding provided by Max Planck Society.

Competing interests

The authors declare no competing interest.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41557-023-01255-5>.

Correspondence and requests for materials should be addressed to Martina Delbianco.

Peer review information *Nature Chemistry* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our [Editorial Policies](#) 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 |
|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> 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) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> 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

Data analysis

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The authors declare that all data supporting the findings of this study are available within the article and in the Supplementary Information files. Raw data for NMR analysis and MD simulations can be downloaded from <https://doi.org/10.17617/3.EG7136>, Edmond. Data are also available from the corresponding author upon request.

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender	<input type="text" value="n/a"/>
Population characteristics	<input type="text" value="n/a"/>
Recruitment	<input type="text" value="n/a"/>
Ethics oversight	<input type="text" value="n/a"/>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

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](https://www.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	<input type="text" value="The graph in figure 2c represent data obtained from MD simulations (25.001 frames, 500 ns)"/>
Data exclusions	<input type="text" value="n/a The graph in figure 2c represent data obtained from MD simulations (25.001 frames, 500 ns)"/>
Replication	<input type="text" value="n/a The graph in figure 2c represent data obtained from MD simulations (25.001 frames, 500 ns)"/>
Randomization	<input type="text" value="n/a The graph in figure 2c represent data obtained from MD simulations (25.001 frames, 500 ns)"/>
Blinding	<input type="text" value="n/a The graph in figure 2c represent data obtained from MD simulations (25.001 frames, 500 ns)"/>

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	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging