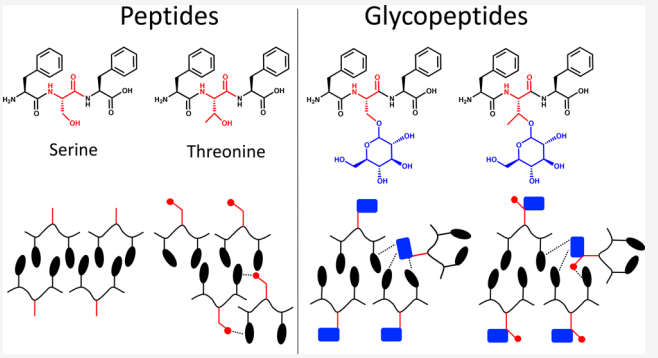
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Article

Expanding the Conformational Landscape of Minimalistic Tripeptides by Their *O*-Glycosylation

Alexandra Brito,[†] Dhwanit Dave,[†] Ayala Lampel, Vânia I. B. Castro, Daniela Kroiss, Rui L. Reis, Tell Tuttle, Rein V. Ulijn,* Ricardo A. Pires, and Iva Pashkuleva*

Cite This: https://doi.org/10.1021/jacs.1c07592
Read Online
ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We report on the supramolecular self-assembly of tripeptides and their *O*-glycosylated analogues, in which the carbohydrate moiety is coupled to a central serine or threonine flanked by phenylalanine residues. The substitution of serine with threonine introduces differential side-chain interactions, which results in the formation of aggregates with different morphology. *O*-glycosylation decreases the aggregation propensity because of rebalancing of the π interactions. The glycopeptides form aggregates with reduced stiffness but increased thermal stability. Our results demonstrate that the designed minimalistic glycopeptides retain critical functional features of glycoproteins and therefore are promising tools for elucidation of molecular mechanisms involved in the glycoprotein interactome. They can also serve as an inspiration for the design of functional glycopeptide-based biomaterials.



INTRODUCTION

Protein aggregation is a supramolecular process often associated with pathological conditions.[1,2] The propensity of a protein to aggregate is primarily coded by the intrinsic properties of the amino acids sequence but also depends on multiple contributing factors from the crowded cellular milieu and post-translational modifications, e.g., glycosylation.[1,3,4] As even the simplest protein is typically composed of hundreds of amino acids, the experimental study and computational modeling of this process is challenging because of the associated combinatorial complexity.[5]

Some years ago, Gazit proposed a reductionist biodesign, which uses the intermolecular self-assembly of minimalistic (less than five amino acids) peptide sequences that can code specific protein bioinformation and transfer it to the assembled system (Chart 1A).[6,7] Such molecular biomimetics are simpler in composition, thus allowing rational and systematic experimental and computational studies to establish connections between the peptide sequence and supramolecular functionality.[5,8,9] Moreover, their simplicity makes them attractive candidates as building blocks for supramolecular materials with designed functions, which may be useful for a variety of biomedical and technological applications.[9−11]

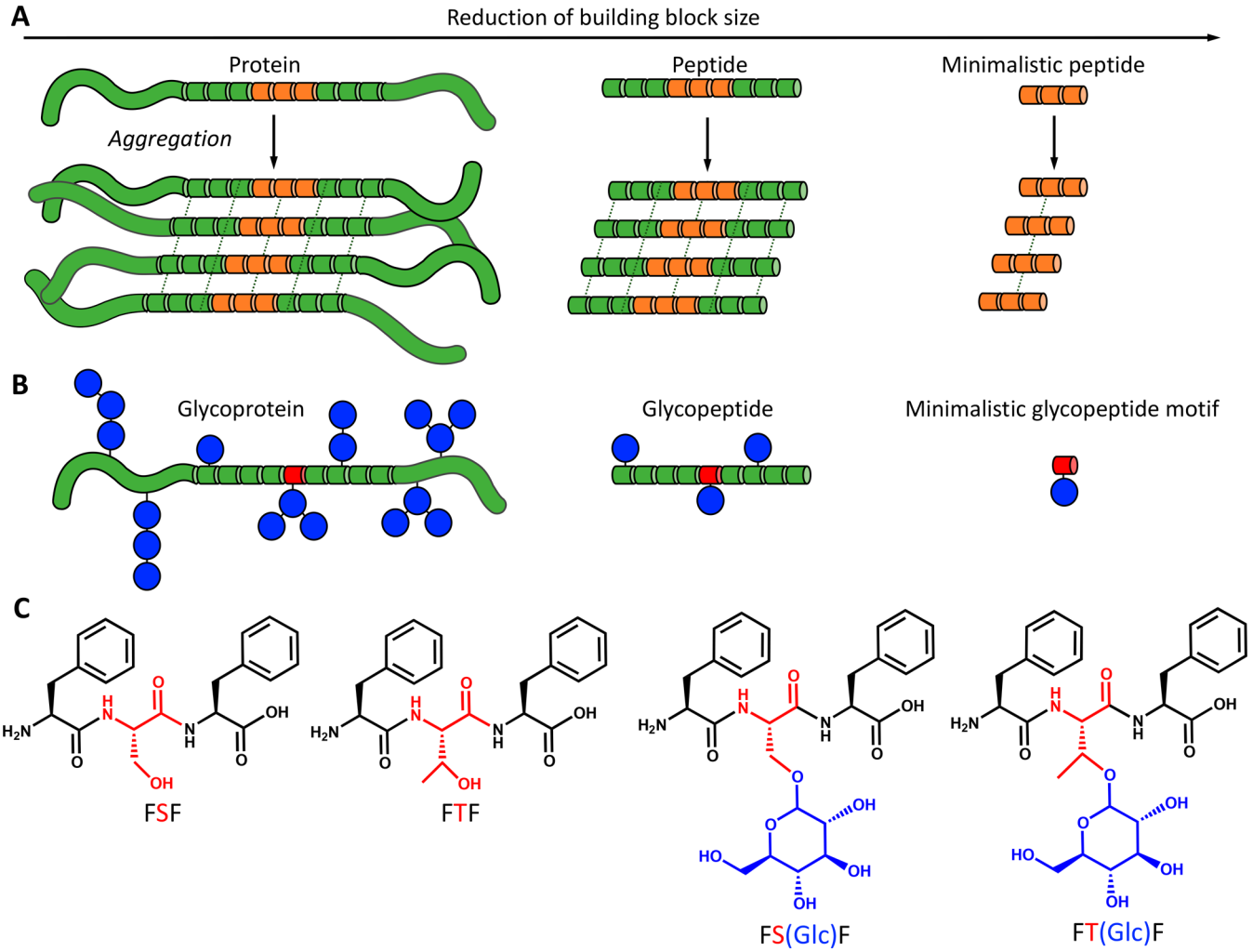
The main rationale of this study is to demonstrate that the reductionist approach proposed for proteins is extendable to glycoproteins (Chart 1B), i.e., that short glycotripeptides can be used as simplified analogues of complex glycoproteins to study and model molecular mechanisms of fundamental properties such as conformational changes and aggregation, and the obtained insights can be applied to rationally modify properties of supramolecular materials based on these motifs.

Glycosylation is a common post-translational modification that effectively enriches the protein repertoire beyond the bioactivities coded by the amino acids sequence and alter the energy landscape associated with the protein aggregation.[4,12−15] However, the exact mechanism of this process is poorly understood and mainly based on in silico models.[4] In eukaryotic cells, *O*-glycosylation takes place at the endoplasmic reticulum or Golgi, where a monosaccharide (usually *N*-acetylgalactosamine but also fucose and glucose) is coupled to the hydroxyl of serine (S) or threonine (T) of newly synthesized polypeptides.[4,13,15] Previous studies have demonstrated that the torsion angle (Ψ) of the glycosidic linkage that determines the orientation of the carbohydrate chain is different for S and T glycopeptides, but the consequence of this difference for distinct biological functions is not clear.[16,17] In here, we applied a reductionist approach to study the effect of S vs T and their glycosylation on conformer selection and molecular aggregation.

Received: July 21, 2021

Downloaded via UNIVERSIDADE DO MINHO 01000 on November 19, 2021 at 13:44:20 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

Chart 1. Schematic Presentation of (A) the Reductionist Approach Using Short Peptides As Analogues of Proteins in Aggregation Studies and (B) the Herein Proposed Approach That Uses Minimalistic Glycopeptide Motifs; (C) Chemical Structure of the Peptides and Glycopeptides Used in This Study[a]



[a]Orange, amino acids involved in the aggregation; green, other amino acids; blue, carbohydrates; red, glycosylated amino acid

RESULTS AND DISCUSSION

Our design of minimalistic *O*-glycoproteins mimics is based on the simplest *O*-glycoprotein motif,[16] S or T that are functionalized with glucose (Glc). Although our study is focused on a deliberately simple model system, *O*-glucosylation is biologically relevant: it is essential for Notch trafficking/signaling and has been associated with defects in neurogenesis, cardiovascular remodeling, somitogenesis, and aberrant gastrulation.[18,19] To promote the aggregation of this motif, we have also included phenylalanine (F) in the peptide sequence (Chart 1C) because previous experimental and computational approaches have shown that the presence of aromatic amino acids enhances the aggregation propensity of short sequences (i.e., sequences with a limited amount of H-bonding between backbone elements).[5,9,20,21]

F has an ability to aggregate alone or when inserted into short (di- and tri-) peptides.[22−24] Tripeptides with flanked aromatic amino acids, e.g., FXF, where X is a hydrophilic amino acid, adopt conformations that allow intramolecular stacking of the two aromatic rings, thus exposing the central amino acid to water.[25] Such sequences self-assemble in water because of the formation of aromatic zippers and a hydrophobic collapse.[23,24,26] In our molecular design, the hydrophilic amino acid is also introduced in the middle of the peptide sequence, and thus the *O*-glycosylation of the short peptide chains at S or T generates minimalistic *O*-glycopeptides, which differ from previously described self-assembling glycopeptides that are end-on glycosylated.[27−29]

We used all-atom molecular dynamics simulations (MDS) with explicit water to investigate the conformational space of the designed peptides and their glycosylated analogs. In agreement with previous studies that include the aromatic−X−aromatic motif,[24−26] we found that in the predominant conformations of FSF and FTF the aromatic amino acids adopt arrangement that allow their intramolecular stacking (Figure 1A).

The F/F dihedral distribution showed similar peaks in the 90° region for FSF and FTF that correspond to intramolecular stacking interactions and was suggestive of molecular preorganization for supramolecular self-assembly (Figure 1C, black). When MDS was applied to 50 molecules instead of one, we observed a small shift in the dihedral distribution to lower angles for both peptides (Figure 1C, red vs black). This shift is indicative of reorganization of the F/F intramolecular

https://doi.org/10.1021/jacs.1c07592
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

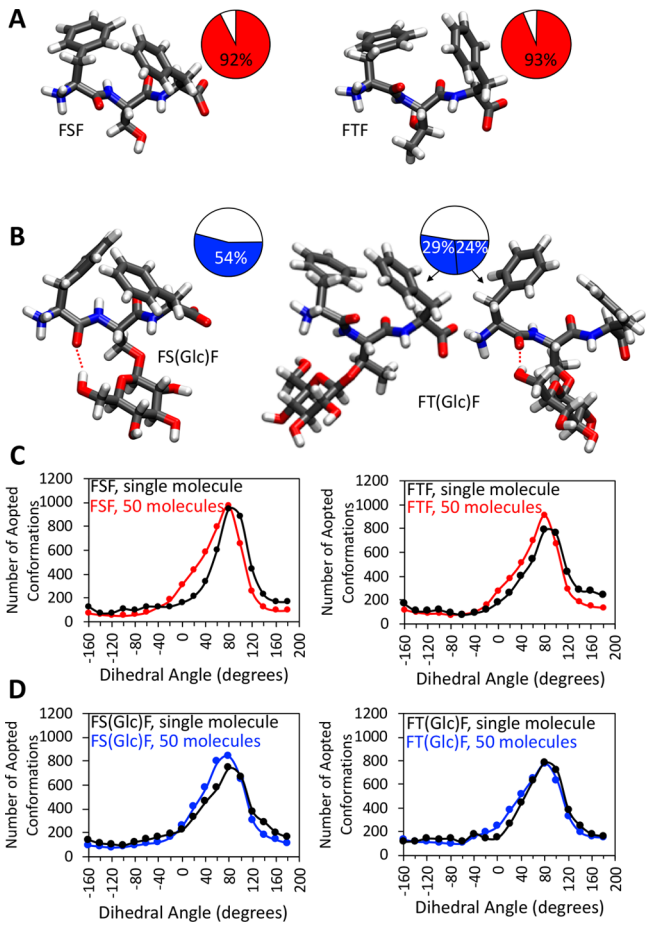


Figure 1. Molecular dynamics analyses showing (A, B) the representative central structures of the largest clusters obtained for (A) model tripeptides and (B) their *O*-glycosylated analogues (the pie-charts and percentages show the fraction of 5000 structures adopting these conformations; a root-mean-square deviation (RMSD) linkage cutoff of 0.1 nm was used for the analyses; supporting data are provided in Figure S17). (C, D) Comparative dihedral analysis (CZ(F)-CA(F)-CA(F)-CZ(F)) of single-molecule and 50 molecule simulations of (C) tripeptides and (D) glycotripeptides. Data for the alpha anomers are presented in Figure S28.

stacks to allow formation of intermolecular aromatic zippers (Figure 2A vs B), i.e., the stacking of interdigitated F side chains from cross-strand peptides leading to self-assembly.

MDS of the glycopeptides revealed that the glycosylation widens the conformational landscape (Figure 1B and Figures S25−S27). The flexibility of the glycosidic bond provides additional modes of interactions with contributions from H-bonding, CH−π, and electrostatic salt-bridge type interactions,[30] leading to different conformer distributions. A comparison of the F/F dihedral distributions in the tripeptides and the glycopeptides (Figure 1D) reveals a reduction in the mean dihedral angle, indicating a wider distribution of glycopeptides conformations that are stabilized by non $\pi-\pi$ type interactions. Of note, the data obtained for the alpha and beta anomers were very similar (Figure S28). Ramachandran plots of these (glyco)peptides (Figure S27) showed that the backbone conformations are similar to the reported for FXF peptides[26] and a conformational diversity arises from side chains and their glycosylation. Additionally, MDS showed that Glc anomers are involved in different intramolecular H-bonding, e.g., the alpha anomer forms H-bond with the carboxylate oxygen, whereas in beta stereochemistry, the Glc interacts with the amide oxygen in FS(Glc)F (Figure S28), thus influencing the glycopeptide conformation.

The computational results were verified experimentally. The aggregation of the tripeptides and their glycosylated analogs was studied in water at 40 mM, i.e., above the critical aggregation concentration, giving rise to transparent viscous liquids. The S to T exchange in these peptides introduces a methyl group into the structure, which affects the morphology of the generated assemblies: FSF forms nanotapes, whereas nanofibrils are observed for FTF (Figure 3 and Figure S33). These results are consistent with previous observations on S/T substitution in self-assembling Fmoc-dipeptides, where planar structures were observed for Fmoc-SF-OMe and an extended network of twisted fibers was obtained for Fmoc-TF-OMe.[31]

The MDS data provided insights in the supramolecular interactions dictating the organization within these structures. A substantial decrease in the solvent-accessible surface area (SASA) for the aromatic F groups over time (Figure 2E, black) was observed, indicating that, as expected, these groups participate in the self-assembly and are mostly buried in the core of the assembled structures (Figure 2D, Movies S1 and S2). We also observed a change in SASA for the amino acid residues during the self-assembly process: SASA reduced less for S and T compared to F (Figure 2E, red), confirming greater exposure of these amino acids on the surface of the peptide assemblies (Figure 2D, Movies S1 and S2). A comparison between S and T revealed differences: S has a higher propensity to form hydrogen bonds with water (Table 1, Figure S30A), whereas T has a higher tendency to interact with phenyl rings via CH−π interactions (Table 1, Figure 2B, and Figure S31A radial distribution peak at 4.5 Å).

Despite the greater hydrophobicity of FTF compared to FSF due to the additional methyl group, the MDS showed a counterintuitive reduction in aggregation propensity for FTF. This result indicates possible disruption of the primarily $\pi-\pi$ driven aggregation by formation of CH-π interactions in FTF aggregates (Figure 2D), i.e. the methyl groups protruding from the FTF peptide chains disturb the assembly of the aromatic zippers (Figure 2B, FSF vs FTF) and can explain the different morphology of FSF and FTF assemblies.

Circular dichroism (CD) data further supported the MDS data. The CD spectra of the peptides have an intensive, positive signal at ∼220 nm for the n−π* transition (Figure 4B).[32,33] In the FTF spectrum, there is an additional positive peak at ∼200 nm that was assigned to the $\pi-\pi$* transition and confirms that the methyl group of T affects the relative spatial orientation of F and thus the supramolecular interactions and the aggregation process.

The in silico models showed lower aggregation propensity for *O*-glycotripeptides (Table 1). Experimentally, we studied mixtures of alpha and beta anomers (alpha:beta was 58:42 for FS(Glc)F and 47:53 FT(Glc)F, Figures S10 and S18). We observed that the glycosylation affected the fiber diameter as observed by TEM (Figure 3), as well as the mechanical properties of the aggregates (Figure 4A), whereas the overall morphologies between peptides and the respective glycosylated analogues appeared similar.

Native *O*-glycoproteins have a high capacity to capture water, which is essential for their viscoelastic properties and physiological functions. At the molecular level, protein glycosylation usually causes higher hydration that can lead to enhanced steric bulk, i.e., steric hindrance around the protein

https://doi.org/10.1021/jacs.1c07592
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

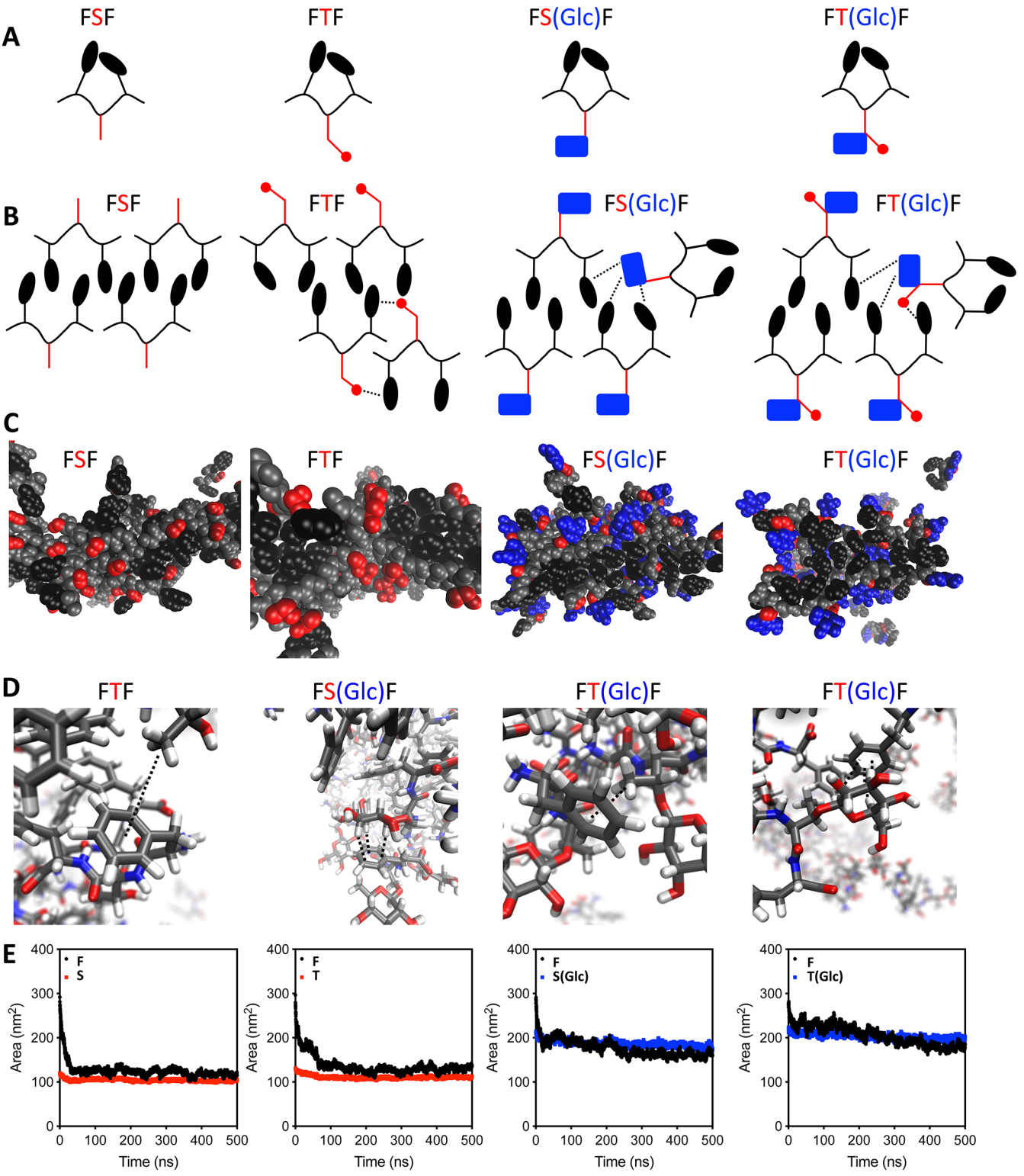


Figure 2. Schematic presentation of (A) predominant conformations of single (glyco)tripeptides based on computational modeling and (B) the CH−π interactions (black dotted lines) involved in the formation of aggregates as shown by the calculated probability $g(r)$; F are shown in black, S/T in red (the red circles represent the methyl group of T), and glucose in blue. (C) Representative van der Waals structures of the (glyco)peptide aggregates observed during MDS: F forms the core of the aggregates with the backbone shown in gray, whereas the polar amino acid (S/T) and glucose are primarily water exposed with some incorporation into aggregates due to (D) CH−π interactions (black dotted lines). (E) Solvent-accessible surface area (SASA) analysis. More details are provided in Table 1 and the Supporting Information.

backbone, which can prevent aggregation, including β-sheet formation.[34] Thus, the decreased Young's modulus of *O*-glycopeptides (Figure 4A, 2-fold as compared with the nonglycosylated tripeptides) is likely due to the increased hydration capacity and/or structural changes caused by the conformational distortions and supramolecular forces, such as

https://doi.org/10.1021/jacs.1c07592
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

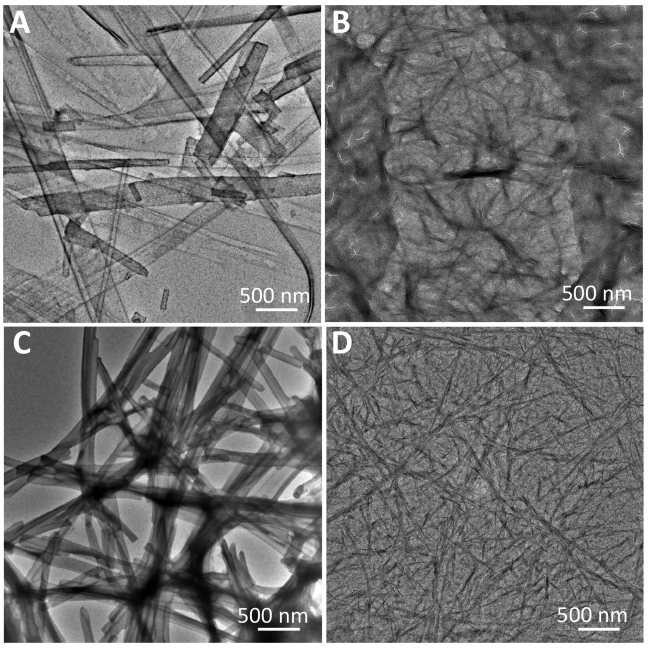


Figure 3. Transmission electron microscopy (TEM) images of (A) FSF, (B) FTF, (C) FS(Glc)F, and (D) FT(Glc)F assemblies formed in water (40 mM, room temperature, 24 h).

Table 1. Computational Data for Aggregation Propensity (50 Molecule Simulations, Average of Three 500 ns Runs), Number of Hydrogen Bonds Formed between the Aggregates and the Solvent, Probability ($g(r)$) of CH−π Interactions between the Aromatic F and S, T, and Glc and β-sheet-like H-Bonds in the Aggregates of the Studied (Glyco)peptides

	aggregation propensity[a]	H-bonds with water	CH−π interactions ($g(r)$)[b]	β-sheet-like H-bonds
FSF	2.23 ± 0.26	313 ± 22	2.8 ± 0.3	28 ± 2
FTF	2.18 ± 0.05	277 ± 20	4.8 ± 1.7	22 ± 6
FS(βGlc)F	1.90 ± 0.14	776 ± 37	3.5 ± 0.3	10 ± 4
FS(αGlc)F	1.82 ± 0.17	740 ± 38	4.1 ± 0.2	14 ± 7
FT(βGlc)F	1.62 ± 0.08	784 ± 37	5.6 ± 1.1	4 ± 2
FT(αGlc)F	1.76 ± 0.06	797 ± 30	6.6 ± 0.4	8 ± 1

[a] $SASA_{initial}/SASA_{final}$. [b] Peak $g(r)$ was measured between heavy atoms of the polar amino acid and the aromatic side chain.

H-bonding and π-interactions, impaired by the introduced Glc.[30,35] As discussed, the carbohydrate moiety is predominantly exposed on the surface of the assemblies, contributing to their increased hydration when compared with the respective tripeptides. The MDS showed that, as expected, the glycosylated peptides have higher SASAs (Figure 2E) combined with an enhanced propensity to form hydrogen bonding with water (Table 1, Figure S29B,C). However, the results also indicated the presence of carbohydrate moieties in the core of the aggregates (Figure 2C, D, Movies S3 and S4), suggesting their involvement in the aggregation process (Figure 2B), beyond simple hydration. Indeed, the glycosylation led to an ∼1.2-fold increase of the propensity to form CH−π interactions (Table 1), which in turn affects the n−π and π−π interactions. This rebalancing of the π-interactions was confirmed by the CD spectra (Figure 4B), where a decrease in the 220 nm signal intensity for both glycopeptides

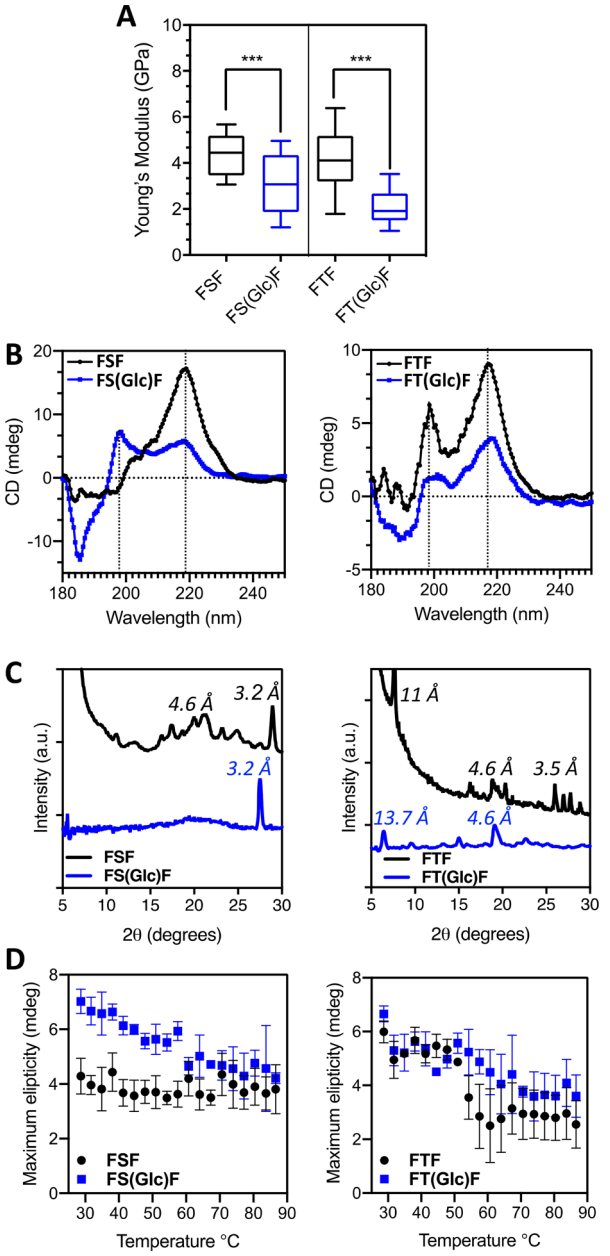


Figure 4. Effect of glycosylation on the aggregation: (A) Young's modulus of the aggregates determined by atomic force microscopy (***$p < 0.0005$); (B) circular dichroism (CD) spectra; (C) X-ray diffraction patterns; and (D) thermal stability of the generated aggregates obtained from the CD spectra at different temperatures.

was observed. The introduction of the carbohydrate group at S has a similar stereochemical effect as the T's methyl group (Figure 2B, FS(Glc)F vs FTF) shown by the appearance of the π−π* signal (198 nm) in the FS(Glc)F spectrum (FSF vs FS(Glc)F in Figure 4B).

X-ray diffraction (XRD, Figure 4C) corroborated the rebalancing of the supramolecular interactions upon *O*-glycosylation of the tripeptides that was observed by CD: in the case of FSF, the peak associated with β-sheet formation (4.6 Å) vanishes for the respective glycosylated analogues, and this change can be explained with the above-mentioned steric bulk, whereas in the case of FTF, the peak associated with aromatic interactions (3.5 Å) disappears upon glycosylation. MDS analysis showed a reduction in H-bonds between the

https://doi.org/10.1021/jacs.1c07592
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

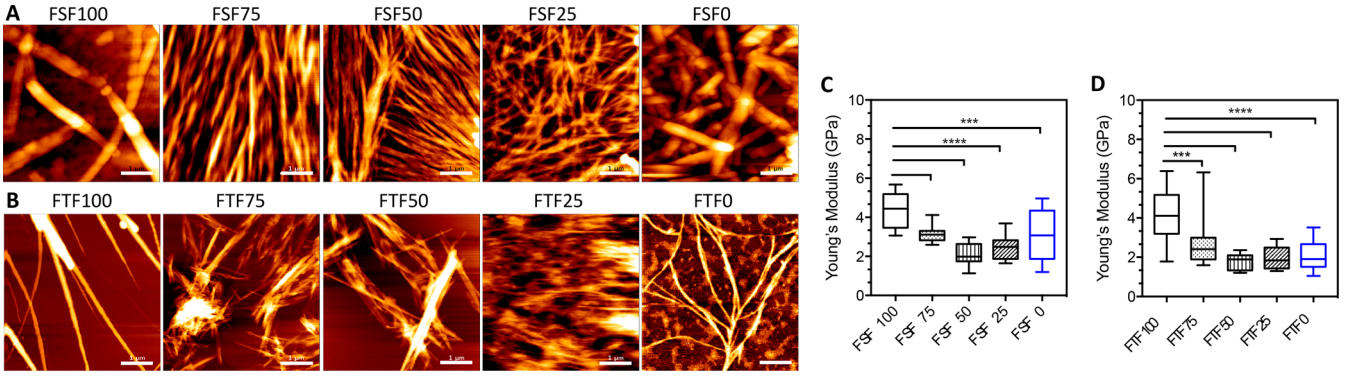


Figure 5. Aggregation of peptide/glycopeptide mixtures at different ratios: (A, B) atomic force microscopy (AFM) images of the assemblies obtained from mixtures at different ratios and (C, D) Young's modulus for these assemblies measured by AFM. FXF100 (FXF:FX(Glc)F = 1:0); FXF75 (FXF:FX(Glc)F = 3:1); FXF50 (FXF:FX(Glc)F = 1:1); FXF25 (FXF:FX(Glc)F = 1:3); FXF0 (FXF:FX(Glc)F = 0:1). Scale bares: 1 μm.

peptide backbones (β-sheet-like interactions, Table 1) upon glycosylation. Moreover, a significant decrease in the FT(Glc)F vs FTF aggregation propensity was also observed.

Table 2. Midpoint Transition (T_M), Enthalpy (ΔH) and Heat Capacity Change (ΔC_p) of the Disassembly Calculated from the CD Spectra (Signal at ~220 nm) at Different Temperatures

	T_M (°C)	ΔH (kJ mol^{-1})	ΔC_p (J °C^{-1})
FSF	52.1 ± 3.1	18.7 ± 2.3	84.9 ± 22.9
FTF	57.2 ± 3.1	16.7 ± 2.9	100.5 ± 32.3
FS(Glc)F	70.0 ± 3.6	28.3 ± 3.5	165.3 ± 28.2
FT(Glc)F	66.2 ± 3.0	20.8 ± 2.8	122.2 ± 29.1

Together, these data confirm the disruption of the aromatic zippers in the glycotripeptides due to the formation of stronger CH-π interactions and explain the disappearance of the peak associated with aromatic stacking. In the case of FS(Glc)F, the aromatic interactions are preserved, as evidenced by the smaller decrease in the aggregation propensity, but the significant reduction in the backbone H-bonds explains the disappearance of the β-sheet peak upon glycosylation.

The performed *O*-glycosylation also affected the thermal stability of the aggregates (Figure 4D, Table 2).[36] Upon heating, the aggregates of the glycosylated FS(Glc)F were more stable with a melting temperature that was 27 °C higher compared to that of FSF. The difference was less pronounced for the FT(Glc)F/FTF couple (~10 °C). These results agree with previous studies with glycoproteins showing that the glycosylation generally improves the thermal stability of the proteins and the magnitude of this effect depends on the size of the carbohydrate chain, the position of glycosylation, and the protein crystallinity.[37−39]

Finally, we also studied the aggregation of mixtures of peptides and the respective glycosylated analogues at different molar ratios to simulate a scenario in which proteins and glycoproteins coexist. The morphology of the assemblies obtained from the mixtures was different from the single-component systems and we observed the formation of entangled nanofibers for all mixtures (Figure 5A, B). The Young's modulus gradually decreased upon addition of the glycopeptides and reached a minimum at ratio 1:1 after which further enrichment of the mixtures with *O*-glycopeptides did not affect the modulus significantly (Figure 5C, D). These results are consistent with coassembly and indicate that in a crowded environment glycosylation affects not only the aggregation of the protein to which the carbohydrate unit is bound but also to the close neighbors.

CONCLUSIONS

In summary, we showed a distinct role of F, S/T, and Glc in the glycopeptides supramolecular interactome and consequently in the characteristics of the generated aggregates. The introduced glycosylation clearly influenced the aggregation, giving rise to enhanced disorder and dynamics in the assembled structures, due to the introduction of CH−π interactions. Such interactions are often challenging to quantify because they are usually inaccessible in native glycoproteins but are crucial for protein synthesis, trafficking, and function.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.1c07592.

Description of the used experimental procedures; data on (glyco)peptides characterization: ^{1}H NMR, ^{13}C NMR, MS, HPLC; supplementary data obtained from the molecular dynamic simulations; supplementary microscopy images (TEM and SEM). (PDF)

Movie S1, representative 500 ns molecular dynamics trajectories of the self-assembly of 50 FSF molecules in a periodic cubic box starting from a random dispersed configuration. The F residues are colored grey and the polar amino acid S is presented in red. Explicit TIP3 water molecules are not shown for clarity. (MPG)

Movie S2, representative 500 ns molecular dynamics trajectories of the self-assembly of 50 FTF molecules in a periodic cubic box starting from a random dispersed configuration. The F residues are presented in grey and the polar amino acid T is presented in red. Explicit TIP3 water molecules are not shown for clarity. (MPG)

Movie S3, representative 500 ns molecular dynamics trajectories of the self-assembly of 50 FS(Glc)F molecules in a periodic cubic box starting from a random dispersed configuration. Color code: F (grey), S (red), and β-D-Glc (blue). Explicit TIP3 water molecules are not shown for clarity. (MPG)

Representative 500 ns molecular dynamics trajectories of the self-assembly of 50 FT(Glc)F molecules in a

https://doi.org/10.1021/jacs.1c07592
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

periodic cubic box starting from a random dispersed configuration. Color code: F (grey), T (red), and β-D-Glc (blue). Explicit TIP3 water molecules are not shown for clarity. (MPG)

AUTHOR INFORMATION

Corresponding Authors

Iva Pashkuleva − *3B's Research Group, I3Bs Research Institute on Biomaterials, Biodegradables and Biomimetics, University of Minho, Headquarters of the European Institute of Excellence on Tissue Engineering and Regenerative Medicine, 4805-017 Barco, Portugal; ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal;* orcid.org/0000-0001-6818-3374; Email: pashkuleva@i3bs.uminho.pt

Rein V. Ulijn − *Advanced Science Research Center (ASRC) at the Graduate Center, City University of New York (CUNY), New York, New York 10031, United States; Department of Chemistry, Hunter College, City University of New York, New York, New York 10065, United States; Ph.D. Program in Chemistry and Ph.D. program in Biochemistry, The Graduate Center of the City University of New York, New York, New York 10016, United States;* orcid.org/0000-0002-7138-1213; Email: rulijn@gc.cuny.edu

Authors

Alexandra Brito − *3B's Research Group, I3Bs Research Institute on Biomaterials, Biodegradables and Biomimetics, University of Minho, Headquarters of the European Institute of Excellence on Tissue Engineering and Regenerative Medicine, 4805-017 Barco, Portugal; ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal; Advanced Science Research Center (ASRC) at the Graduate Center, City University of New York (CUNY), New York, New York 10031, United States;* orcid.org/0000-0002-5027-4080

Dhwanit Dave − *Advanced Science Research Center (ASRC) at the Graduate Center, City University of New York (CUNY), New York, New York 10031, United States; Department of Chemistry, Hunter College, City University of New York, New York, New York 10065, United States; Ph.D. Program in Chemistry, The Graduate Center of the City University of New York, New York, New York 10016, United States;* orcid.org/0000-0002-6218-4584

Ayala Lampel − *Advanced Science Research Center (ASRC) at the Graduate Center, City University of New York (CUNY), New York, New York 10031, United States;* Present Address: Shmunis School of Biomedicine and Cancer Research, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv 9978, Israel

Vânia I. B. Castro − *3B's Research Group, I3Bs Research Institute on Biomaterials, Biodegradables and Biomimetics, University of Minho, Headquarters of the European Institute of Excellence on Tissue Engineering and Regenerative Medicine, 4805-017 Barco, Portugal; ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal;* orcid.org/0000-0002-3061-3954

Daniela Kroiss − *Advanced Science Research Center (ASRC) at the Graduate Center, City University of New York (CUNY), New York, New York 10031, United States; Department of Chemistry, Hunter College, City University of New York, New York, New York 10065, United States; Ph.D. Program in Chemistry, The Graduate Center of the City University of New York, New York, New York 10016, United States;* orcid.org/0000-0003-1483-9096

Rui L. Reis − *3B's Research Group, I3Bs Research Institute on Biomaterials, Biodegradables and Biomimetics, University of Minho, Headquarters of the European Institute of Excellence on Tissue Engineering and Regenerative Medicine, 4805-017 Barco, Portugal; ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal*

Tell Tuttle − *Department of Pure and Applied Chemistry, University of Strathclyde, Glasgow G1 1XL, United Kingdom;* orcid.org/0000-0003-2300-8921

Ricardo A. Pires − *3B's Research Group, I3Bs Research Institute on Biomaterials, Biodegradables and Biomimetics, University of Minho, Headquarters of the European Institute of Excellence on Tissue Engineering and Regenerative Medicine, 4805-017 Barco, Portugal; ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal;* orcid.org/0000-0002-9197-0138

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.1c07592

Author Contributions

†A.B. and D.D. contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge the EU's H2020 program (Forecast 668983) and the Portuguese FCT (BD/113794/2015; PTDC/BTM-MAT/28327/2017 CARDIOHEAL; M-ERA-NET2/0001/2016 INCIPIT) for the financial support. Part of this research was supported by National Science Foundation (NSF) grant CHE-1808143 and a grant of computer time from the City University of New York High Performance Computing Center under NSF grants CNS-0855217, CNS-0958379, and ACI-1126113. D.D. thanks Maithreyi Ramakrishnan for provision of a dihedral analysis script and Mateusz Marianski for discussions.

REFERENCES

(1) Vendruscolo, M.; Zurdo, J.; MacPhee, C. E.; Dobson, C. M. Protein folding and misfolding: a paradigm of self-assembly and regulation in complex biological systems. *Philos. Trans. R. Soc., A* **2003**, *361* (1807), 1205−1222.

(2) Gazit, E. The "Correctly folded" state of proteins: Is it a metastable state. *Angew. Chem., Int. Ed.* **2002**, *41* (2), 257−259.

(3) White, D. A.; Buell, A. K.; Knowles, T. P. J.; Welland, M. E.; Dobson, C. M. Protein Aggregation in Crowded Environments. *J. Am. Chem. Soc.* **2010**, *132* (14), 5170−5175.

(4) Shental-Bechor, D.; Levy, Y. Effect of glycosylation on protein folding: A close look at thermodynamic stabilization. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (24), 8256−8261.

(5) Pappas, C. G.; Shafi, R.; Sasselli, I. R.; Siccardi, H.; Wang, T.; Narang, V.; Abzalimov, R.; Wijerathne, N.; Ulijn, R. V. Dynamic peptide libraries for the discovery of supramolecular nanomaterials. *Nat. Nanotechnol.* **2016**, *11* (11), 960−967.

(6) Brahmachari, S.; Arnon, Z. A.; Frydman-Marom, A.; Gazit, E.; Adler-Abramovich, L. Diphenylalanine as a Reductionist Model for the Mechanistic Characterization of beta-Amyloid Modulators. *ACS Nano* **2017**, *11* (6), 5960−5969.

(7) Gazit, E.; Kornberg, R. D. Reductionist Approach in Peptide-Based Nanotechnology. *Annu. Rev. Biochem.* **2018**, *87*, 533−553.

(8) Yoshida, M.; Hinkley, T.; Tsuda, S.; Abul-Haija, Y. M.; McBurney, R. T.; Kulikov, V.; Mathieson, J. S.; Galinanes Reyes, S.;

Castro, M. D.; Cronin, L. Using Evolutionary Algorithms and Machine Learning to Explore Sequence Space for the Discovery of Antimicrobial Peptides. *Chem-Us* **2018**, *4* (3), 533−543.
(9) Frederix, P. W. J. M.; Scott, G. G.; Abul-Haija, Y. M.; Kalafatovic, D.; Pappas, C. G.; Javid, N.; Hunt, N. T.; Ulijn, R. V.; Tuttle, T. Exploring the sequence space for (tri-) peptide self-assembly to design and discover. *Nat. Chem.* **2015**, *7* (1), 30−37.
(10) Clemons, T. D.; Stupp, S. I. Design of materials with supramolecular polymers. *Prog. Polym. Sci.* **2020**, *111*, 101310.
(11) Levin, A.; Hakala, T. A.; Schnaider, L.; Bernardes, G. J. L.; Gazit, E.; Knowles, T. P. J. Biomimetic peptide self-assembly for functional materials. *Nature Reviews Chemistry* **2020**, *4* (11), 615−634.
(12) Soares da Costa, D.; Reis, R. L.; Pashkuleva, I. Sulfation of Glycosaminoglycans and Its Implications in Human Health and Disorders. *Annu. Rev. Biomed. Eng.* **2017**, *19*, 1−26.
(13) Xu, C. C.; Ng, D. T. W. Glycosylation-directed quality control of protein folding. *Nat. Rev. Mol. Cell Biol.* **2015**, *16* (12), 742−752.
(14) Dobson, C. M. Protein folding and misfolding. *Nature* **2003**, *426* (6968), 884−890.
(15) Pratt, M. R.; Bertozzi, C. R. Synthetic glycopeptides and glycoproteins as tools for biology. *Chem. Soc. Rev.* **2005**, *34* (1), 58−68.
(16) Corzana, F.; Busto, J. H.; Jimenez-Oses, G.; Garcia de Luis, M.; Asensio, J. L.; Jimenez-Barbero, J.; Peregrina, J. M.; Avenoza, A. Serine versus threonine glycosylation: The methyl group causes a drastic alteration on the carbohydrate orientation and on the surrounding water shell. *J. Am. Chem. Soc.* **2007**, *129* (30), 9458−9467.
(17) Coltart, D. M.; Royyuru, A. K.; Williams, L. J.; Glunz, P. W.; Sames, D.; Kuduk, S. D.; Schwarz, J. B.; Chen, X. T.; Danishefsky, S. J.; Live, D. H. Principles of mucin architecture: Structural studies on synthetic glycopeptides bearing clustered mono-, di-, tri-, and hexasaccharide glycodomains. *J. Am. Chem. Soc.* **2002**, *124* (33), 9833−9844.
(18) Yu, H. J.; Takeuchi, H. Protein O-glucosylation: another essential role of glucose in biology. *Curr. Opin. Struct. Biol.* **2019**, *56*, 64−71.
(19) Acar, M.; Jafar-Nejad, H.; Takeuchi, H.; Rajan, A.; Ibrani, D.; Rana, N. A.; Pan, H.; Haltiwanger, R. S.; Bellen, H. J. Rumi is a CAP10 domain glycosyltransferase that modifies notch and is required for notch signaling. *Cell* **2008**, *132* (2), 247−258.
(20) Pawar, A. P.; DuBay, K. F.; Zurdo, J.; Chiti, F.; Vendruscolo, M.; Dobson, C. M. Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. *J. Mol. Biol.* **2005**, *350* (2), 379−392.
(21) Gazit, E. A possible role for pi-stacking in the self-assembly of amyloid fibrils. *FASEB J.* **2002**, *16* (1), 77−83.
(22) Adler-Abramovich, L.; Vaks, L.; Carny, O.; Trudler, D.; Magno, A.; Caflisch, A.; Frenkel, D.; Gazit, E. Phenylalanine assembly into toxic fibrils suggests amyloid etiology in phenylketonuria. *Nat. Chem. Biol.* **2012**, *8* (8), 701−706.
(23) Shlomo, Z.; Vinod, T. P.; Jelinek, R.; Rapaport, H. Stacking interactions by two Phe side chains stabilize and orient assemblies of even the minimal amphiphilic beta-sheet motif. *Chem. Commun.* **2015**, *51* (15), 3154−3157.
(24) Vargiu, A. V.; Iglesias, D.; Styan, K. E.; Waddington, L. J.; Easton, C. D.; Marchesan, S. Design of a hydrophobic tripeptide that self-assembles into amphiphilic superstructures forming a hydrogel biomaterial. *Chem. Commun.* **2016**, *52* (35), 5912−5915.
(25) Lampel, A.; McPhee, S. A.; Park, H. A.; Scott, G. G.; Humagain, S.; Hekstra, D. R.; Yoo, B.; Frederix, P.; Li, T. D.; Abzalimov, R. R.; Greenbaum, S. G.; Tuttle, T.; Hu, C. H.; Bettinger, C. J.; Ulijn, R. V. Polymeric peptide pigments with sequence-encoded properties. *Science* **2017**, *356* (6342), 1064−1068.
(26) Garcia, A. M.; Iglesias, D.; Parisi, E.; Styan, K. E.; Waddington, L. J.; Deganutti, C.; De Zorzi, R.; Grassi, M.; Melchionna, M.; Vargiu, A. V.; Marchesan, S. Chirality Effects on Peptide Self-Assembly Unraveled from Molecules to Materials. *Chem-Us* **2018**, *4* (8), 1862−1876.
(27) Restuccia, A.; Seroski, D. T.; Kelley, K. L.; O'Bryan, C. S.; Kurian, J. J.; Knox, K. R.; Farhadi, S. A.; Angelini, T. E.; Hudalla, G. A. Hierarchical self-assembly and emergent function of densely glycosylated peptide nanofibers. *Communications Chemistry* **2019**, *2*, 53.
(28) Lee, S. S.; Fyrner, T.; Chen, F.; Alvarez, Z.; Sleep, E.; Chun, D. S.; Weiner, J. A.; Cook, R. W.; Freshman, R. D.; Schallmo, M. S.; Katchko, K. M.; Schneider, A. D.; Smith, J. T.; Yun, C. W.; Singh, G.; Hashmi, S. Z.; McClendon, M. T.; Yu, Z. L.; Stock, S. R.; Hsu, W. K.; Hsu, E. L.; Stupp, S. I. Sulfated glycopeptide nanostructures for multipotent protein activation. *Nat. Nanotechnol.* **2017**, *12* (8), 821−829.
(29) Liu, R. Y.; Zhang, R.; Li, L.; Kochovski, Z.; Yao, L. T.; Nieh, M. P.; Lu, Y.; Shi, T. F.; Chen, G. S. A Comprehensive Landscape for Fibril Association Behaviors Encoded Synergistically by Saccharides and Peptides. *J. Am. Chem. Soc.* **2021**, *143* (17), 6622−6633.
(30) Brito, A.; Kassem, S.; Reis, R. L.; Ulijn, R. V.; Pires, R. A.; Pashkuleva, I. Carbohydrate amphiphiles for supramolecular biomaterials: Design, self-assembly, and applications. *Chem-Us* **2021**, *7* (11), 2943−2964.
(31) Hughes, M.; Frederix, P. W. J. M.; Raeburn, J.; Birchall, L. S.; Sadownik, J.; Coomer, F. C.; Lin, I. H.; Cussen, E. J.; Hunt, N. T.; Tuttle, T.; Webb, S. J.; Adams, D. J.; Ulijn, R. V. Sequence/structure relationships in aromatic dipeptide hydrogels formed under thermodynamic control by enzyme-assisted self-assembly. *Soft Matter* **2012**, *8* (20), 5595−5602.
(32) Brito, A.; Abul-Haija, Y. M.; da Costa, D. S.; Novoa-Carballal, R.; Reis, R. L.; Ulijn, R. V.; Pires, R. A.; Pashkuleva, I. Minimalistic supramolecular proteoglycan mimics by co-assembly of aromatic peptide and carbohydrate amphiphiles. *Chemical Science* **2019**, *10* (8), 2385−2390.
(33) Naganagowda, G. A.; Gururaja, T. L.; Satyanarayana, J.; Levine, M. J. NMR analysis of human salivary mucin (MUC7) derived O-linked model glycopeptides: comparison of structural features and carbohydrate-peptide interactions. *J. Pept. Res.* **1999**, *54* (4), 290−310.
(34) Kramer, J. R.; Onoa, B.; Bustamante, C.; Bertozzi, C. R. Chemically tunable mucin chimeras assembled on living cells. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (41), 12574−12579.
(35) Birchall, L. S.; Roy, S.; Jayawarna, V.; Hughes, M.; Irvine, E.; Okorogheye, G. T.; Saudi, N.; De Santis, E.; Tuttle, T.; Edwards, A. A.; Ulijn, R. V. Exploiting CH-pi interactions in supramolecular hydrogels of aromatic carbohydrate amphiphiles. *Chemical Science* **2011**, *2* (7), 1349−1355.
(36) Greenfield, N. J. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protoc.* **2006**, *1* (6), 2527−2535.
(37) Chen, W. T.; Enck, S.; Price, J. L.; Powers, D. L.; Powers, E. T.; Wong, C. H.; Dyson, H. J.; Kelly, J. W. Structural and Energetic Basis of Carbohydrate-Aromatic Packing Interactions in Proteins. *J. Am. Chem. Soc.* **2013**, *135* (26), 9877−9884.
(38) Wang, C. Q.; Eufemi, M.; Turano, C.; Giartosio, A. Influence of the carbohydrate moiety on the stability of glycoproteins. *Biochemistry* **1996**, *35* (23), 7299−7307.
(39) Sola, R. J.; Al-Azzam, W.; Griebenow, K. Engineering of protein thermodynamic, kinetic, and colloidal stability: Chemical glycosylation with monofunctionally activated glycan. *Biotechnol. Bioeng.* **2006**, *94* (6), 1072−1079.

https://doi.org/10.1021/jacs.1c07592
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX