Electronic Supplementary Material (ESI) for RSC Medicinal Chemistry.
This journal is © The Royal Society of Chemistry 2020

SUPPORTING INFORMATION (SI) for

# Scaffold Diversity for Enhanced Activity of Glycosylated Inhibitors of Fungal Adhesion

Harlei Martin[1], Tara Somers[2], Mathew Dwyer [2], Ryan Robson[3], Frederick M. Pfeffer [3], Ragnar Bjornsson[4], Tobias Krämer[1,5], Kevin Kavanagh[2,6*], and Trinidad Velasco-Torrijos[1,6*].

[1]Department of Chemistry, Maynooth University, Maynooth, Co. Kildare, Ireland.

[2]Department of Biology, Maynooth University, Maynooth, Co. Kildare, Ireland.

[3]School of Life and Environmental Sciences, Deakin University, Geelong, Victoria 3217, Australia.

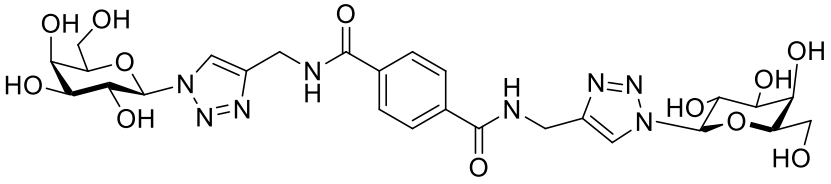
[4] Department of Inorganic Spectroscopy, Max Planck Institute for Chemical Energy Conversion, Stiftstrasse 34–36, 45470 Mülheim an der Ruhr, Germany.

[5]The Hamilton Institute, Maynooth University, Maynooth, Co. Kildare, Ireland.

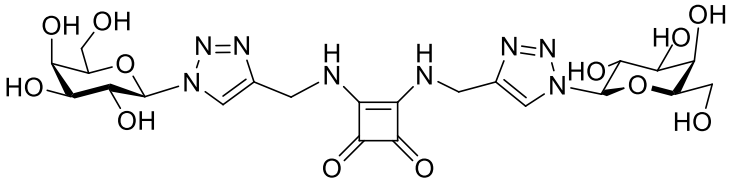
[6]The Kathleen Lonsdale Institute for Human Health Research, Maynooth University, Maynooth, Co. Kildare, Ireland.

## Contents:

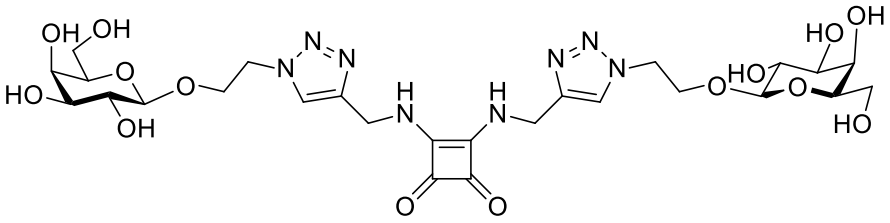
## 1. Chemical Structures of Compounds Evaluated



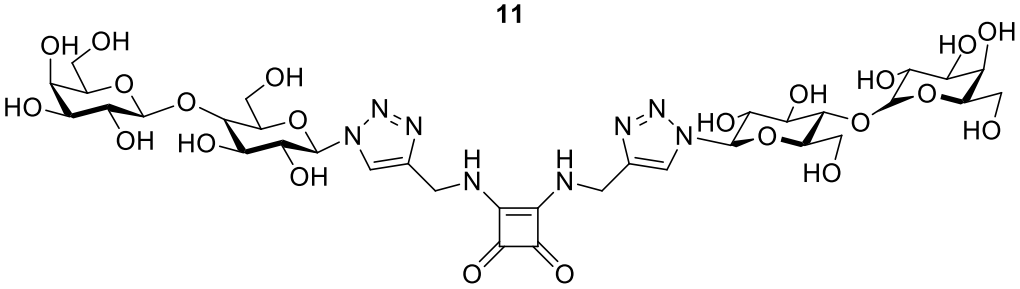
**6**



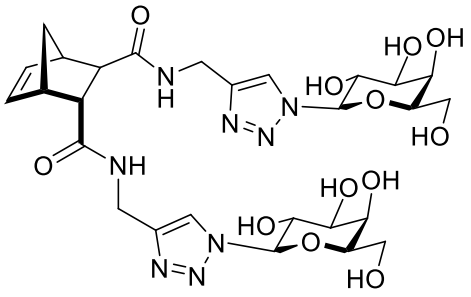
**9**



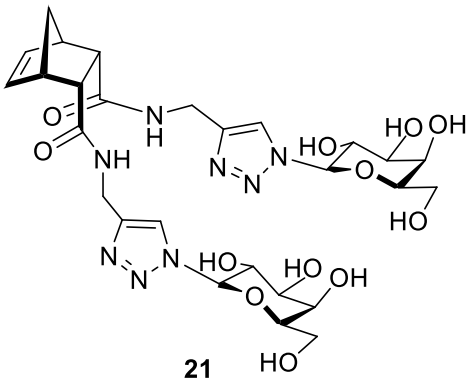
**11**



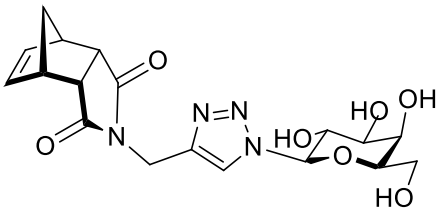
**13**



**17**



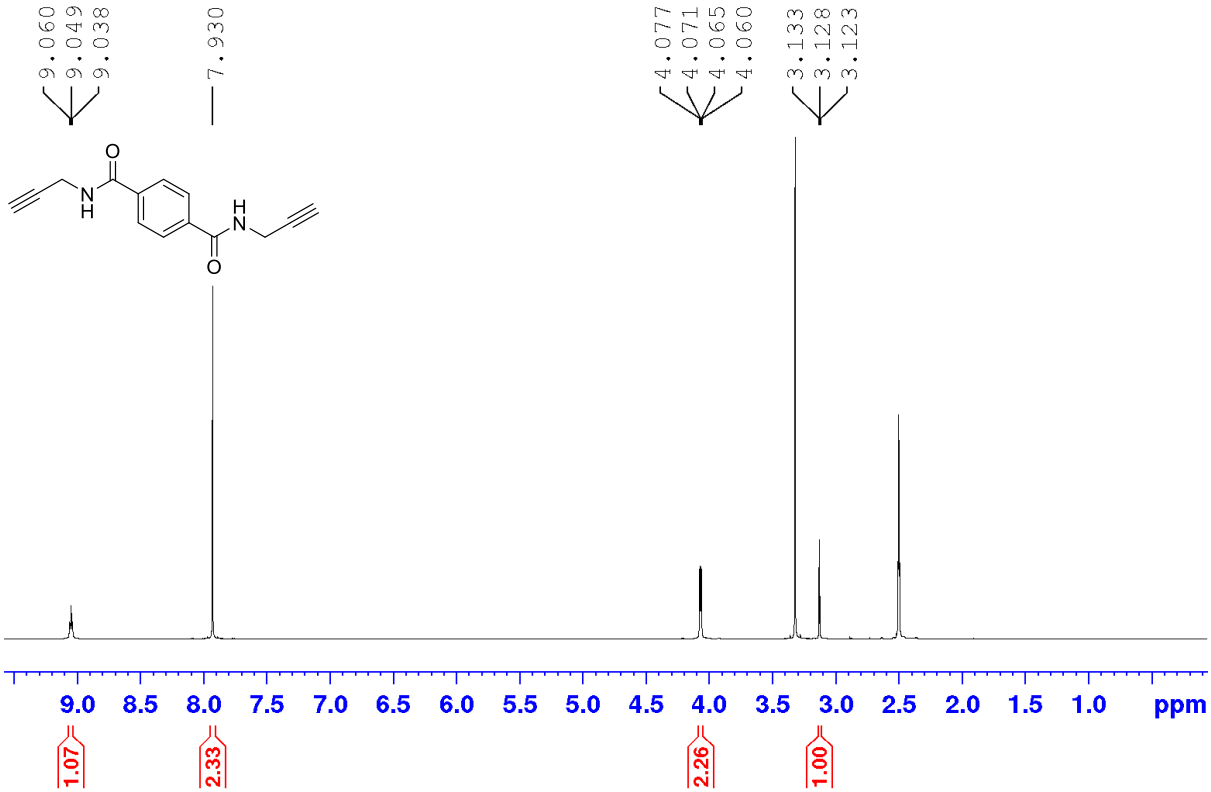
**21**



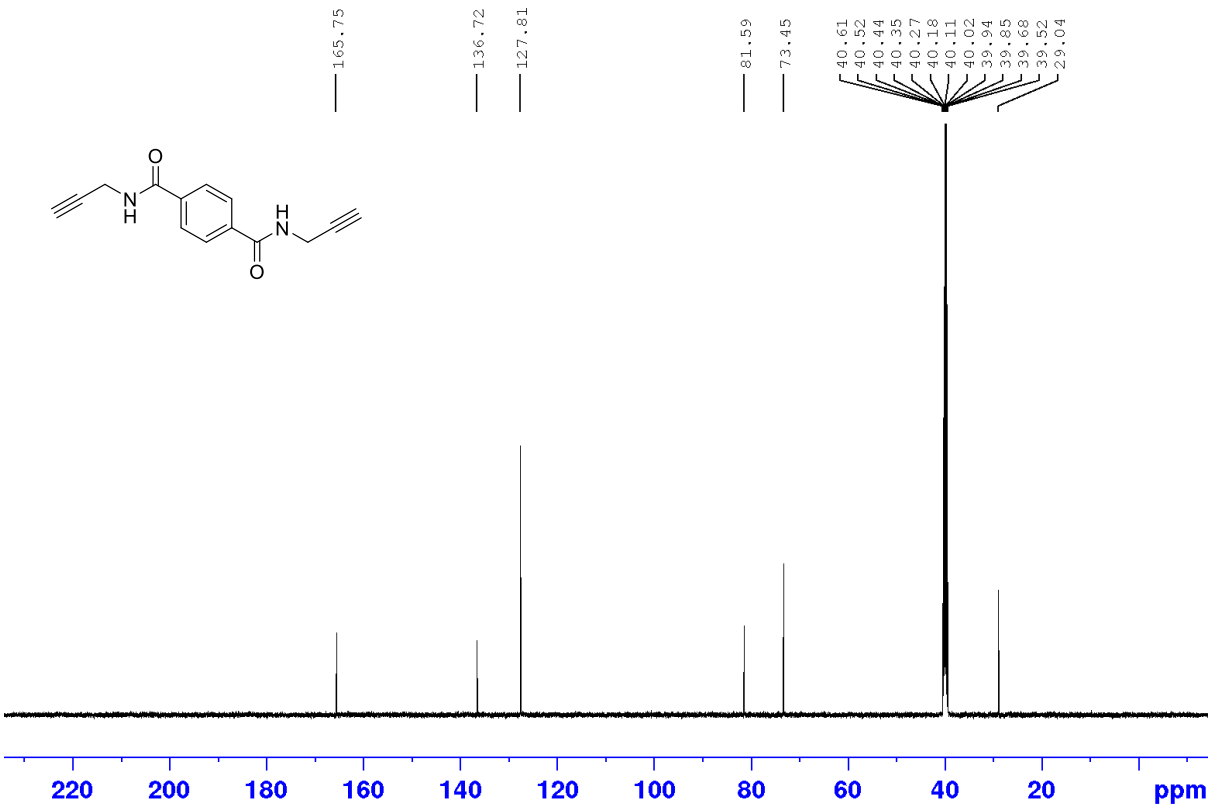
**22**

## 2. Spectroscopic Data

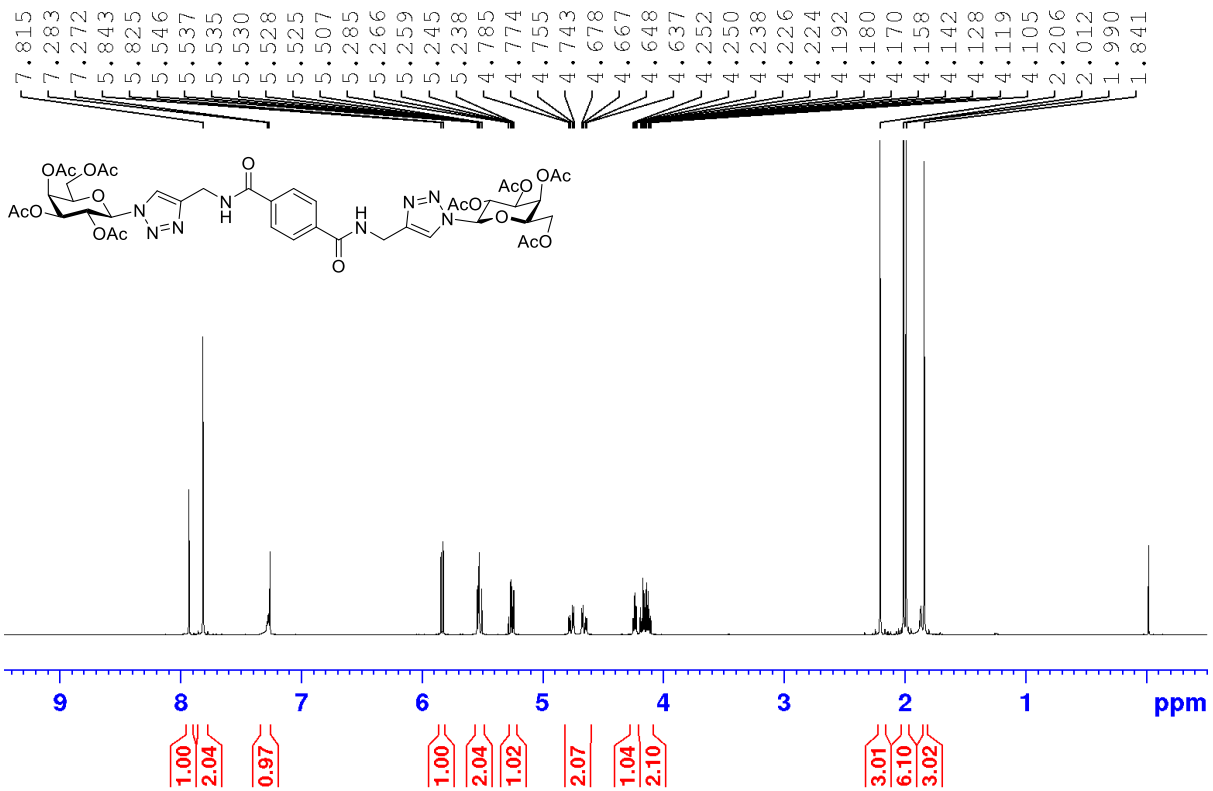
*$^{1}$H NMR spectrum of N,N'-di(prop-2-yn-1-yl)terephthalamide* **4**



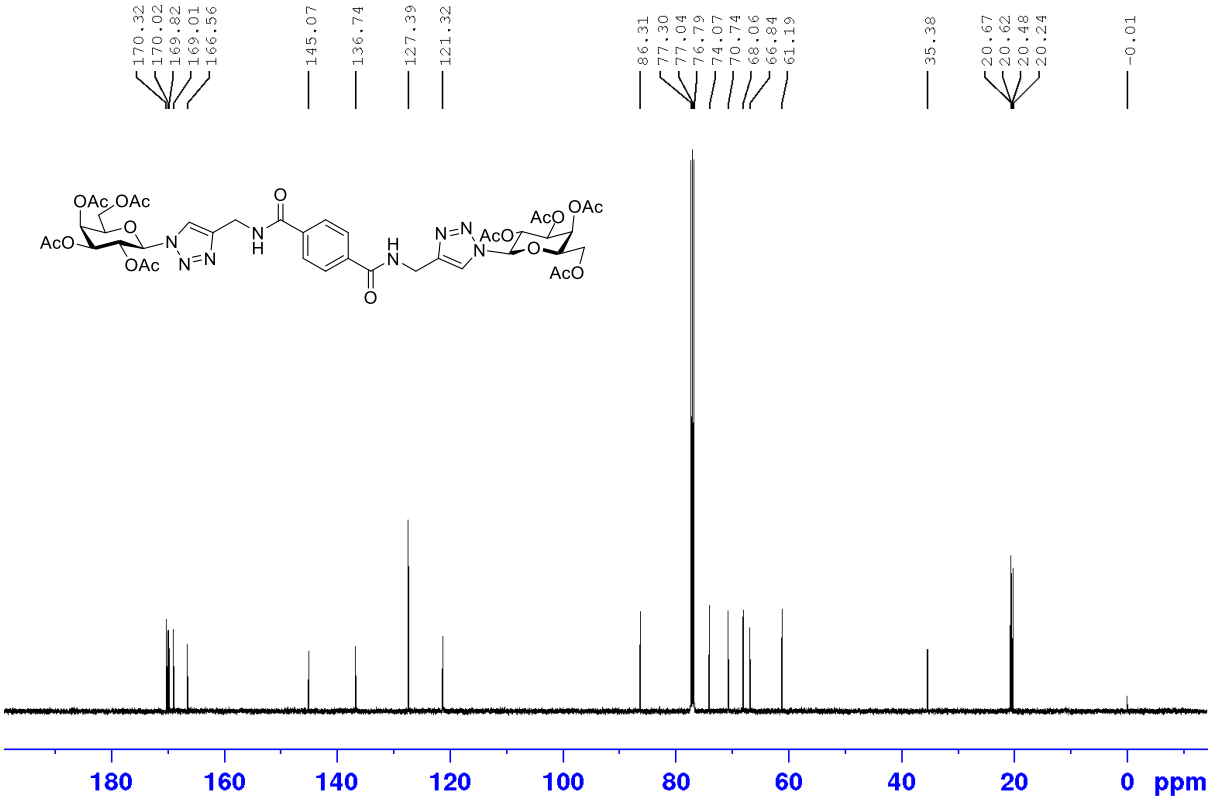

*$^{13}$C NMR spectrum of N,N'-di(prop-2-yn-1-yl)terephthalamide* **4**



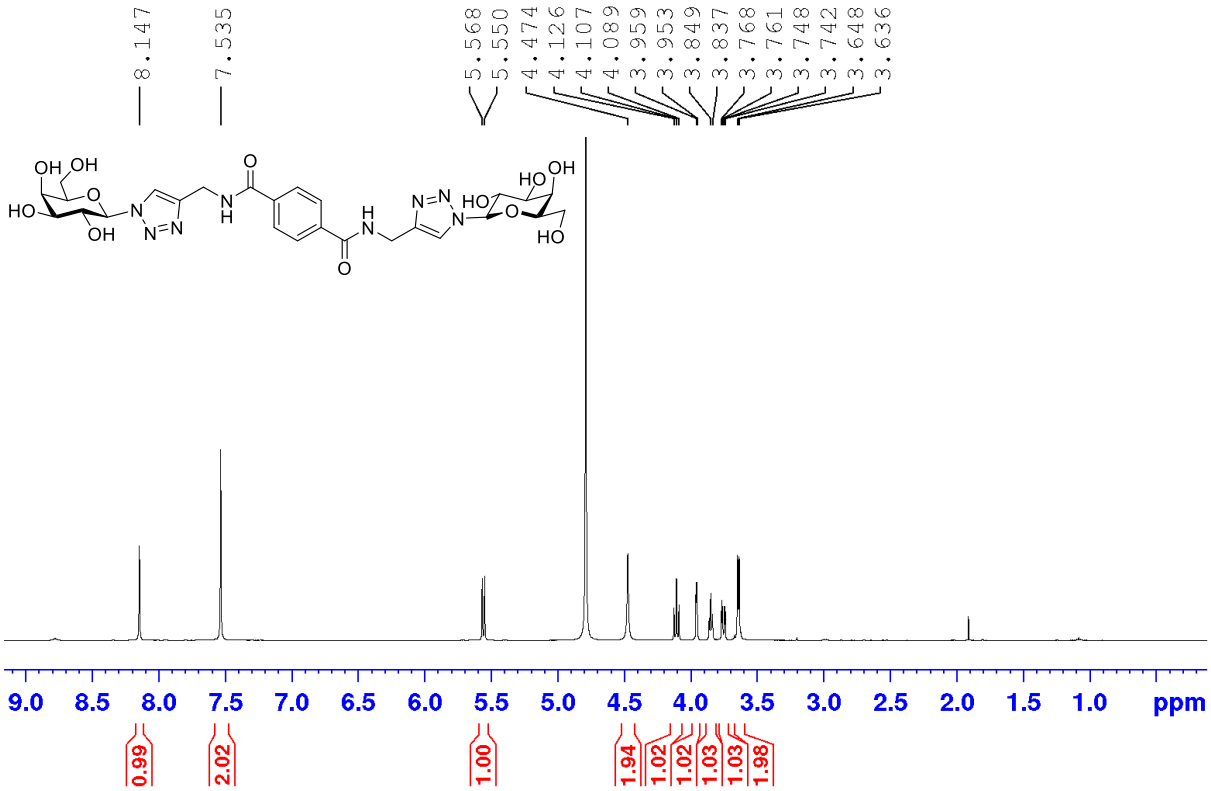
*[1]H NMR spectrum of N, N'-di-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-1,2,3-triazol-4-ylmethyl)-terephthalamide* **5**



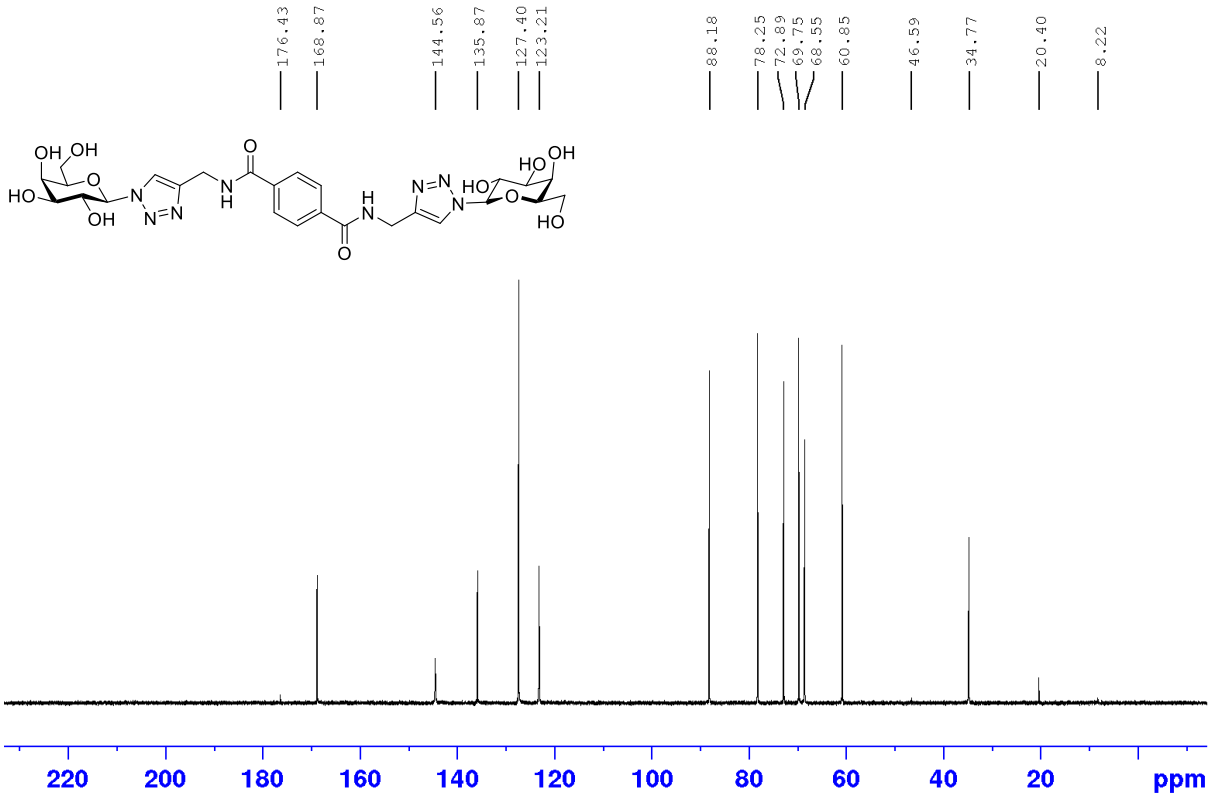

*[13]C NMR spectrum of N, N'-di-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-1,2,3-triazol-4-ylmethyl)-terephthalamide* **5**



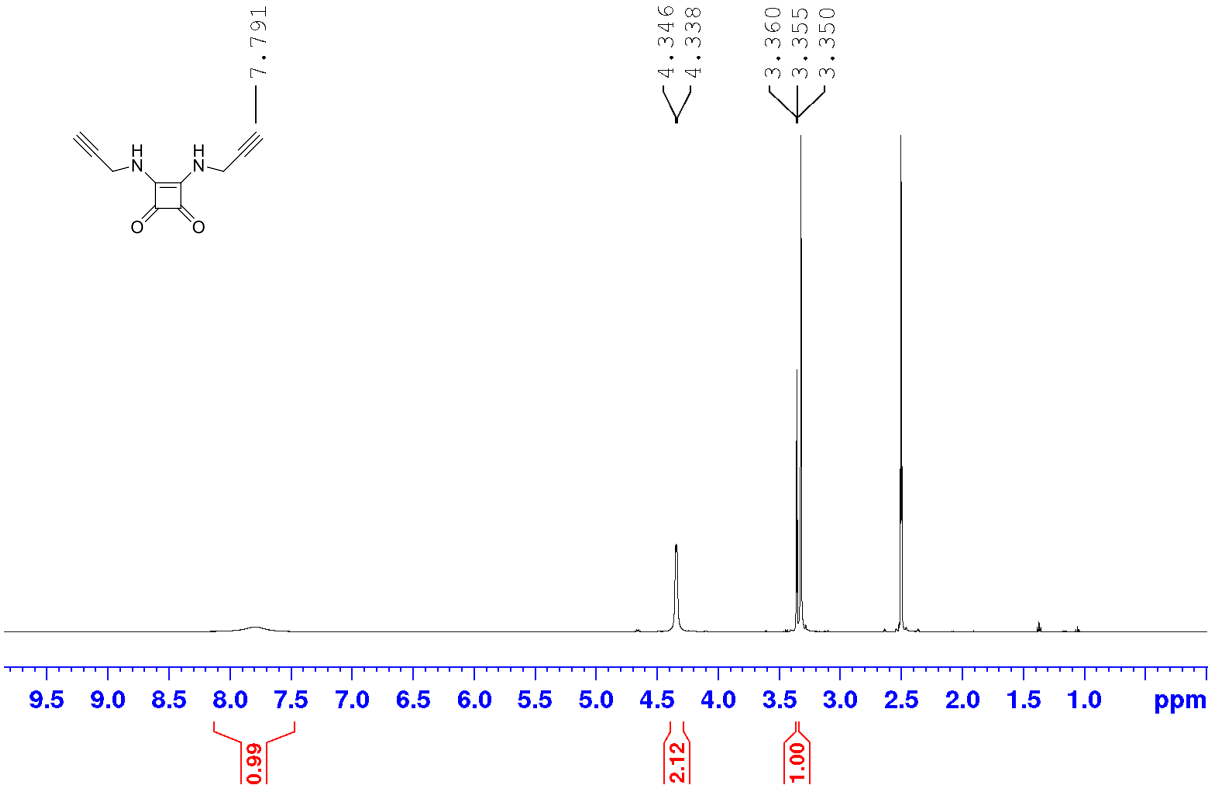
*[1]H NMR spectrum of N, N'-di-(β-D-galactopyranosyl-1,2,3-triazol-4-ylmethyl)-terephthalamide* **6**



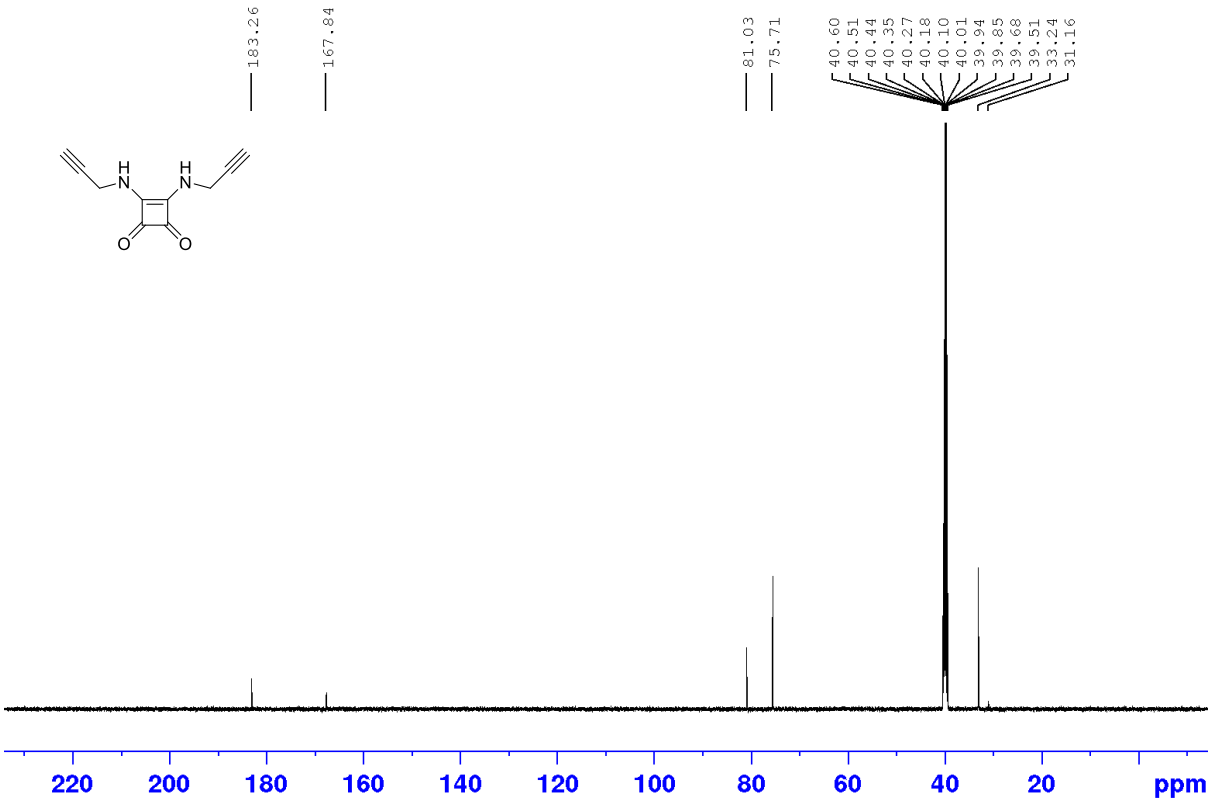

*[13]C NMR spectrum of N, N'-di-(β-D-galactopyranosyl-1,2,3-triazol-4-ylmethyl)-terephthalamide* **6**



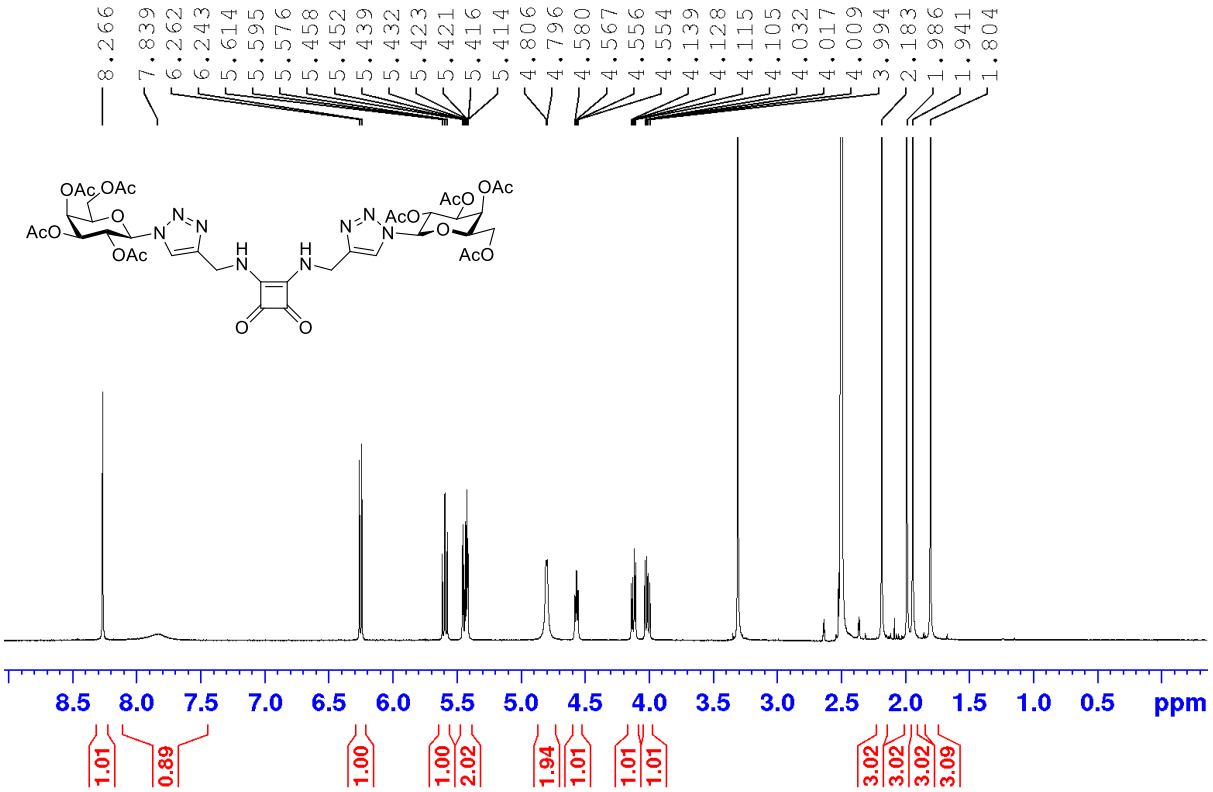
*[1]H NMR spectrum of 3,4-di(prop-2-yn-1-ylamino)cyclobut-3-ene-1,2-dione* **7**



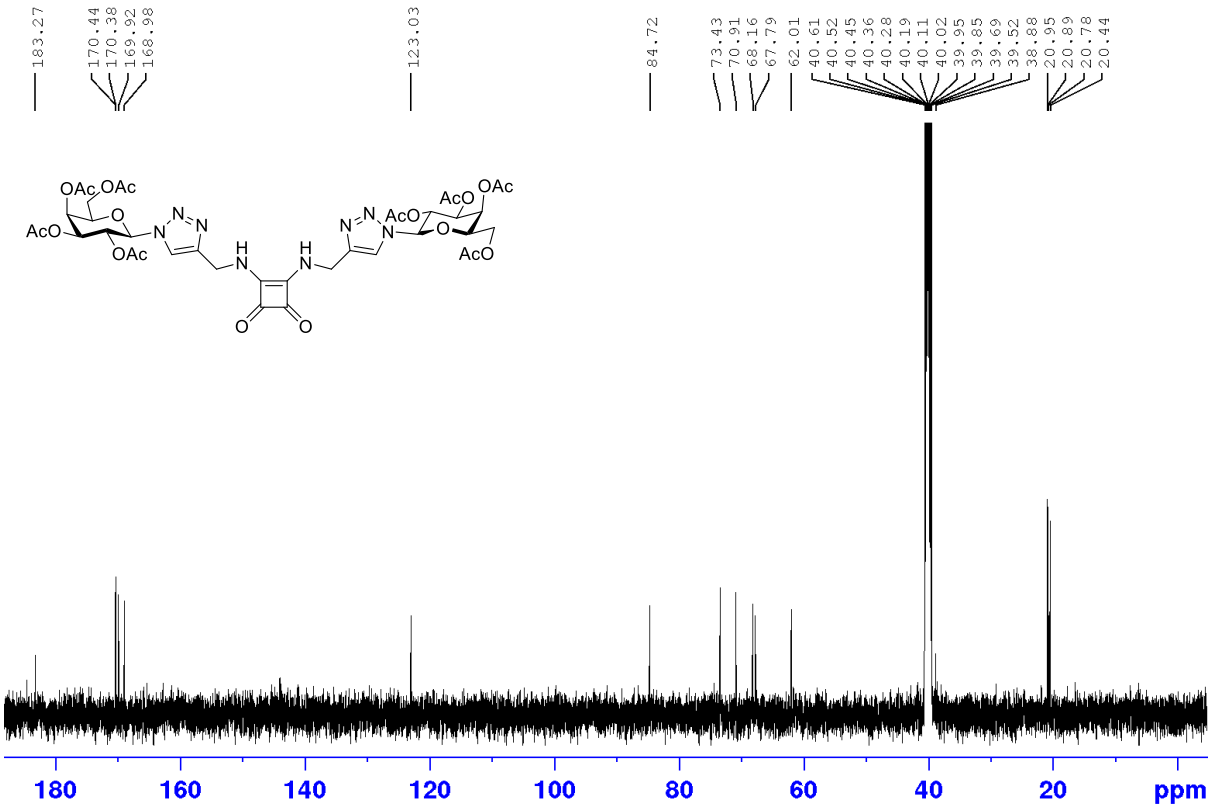

*[13]C NMR spectrum of 3,4-di(prop-2-yn-1-ylamino)cyclobut-3-ene-1,2-dione* **7**



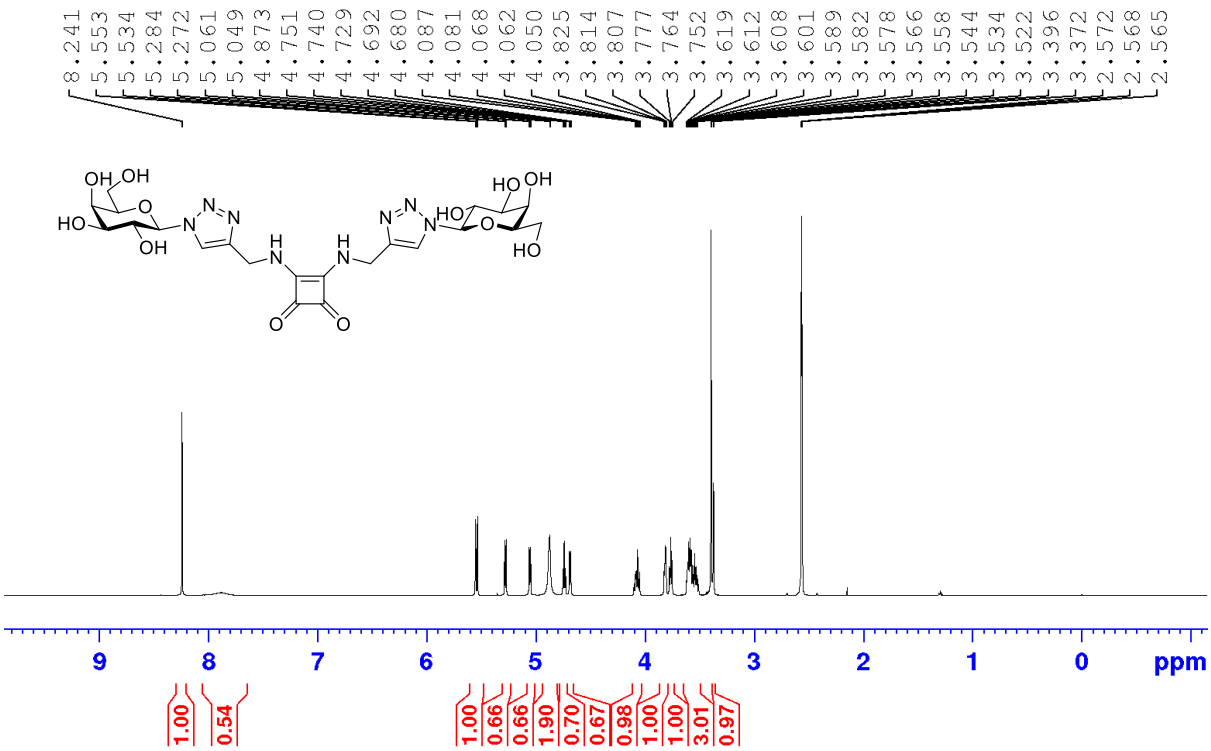
*[1]H NMR spectrum of 3,4-di(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-1,2,3-triazol-4-ylmethylamino)cyclobut-3-ene-1,2-dione* **8**



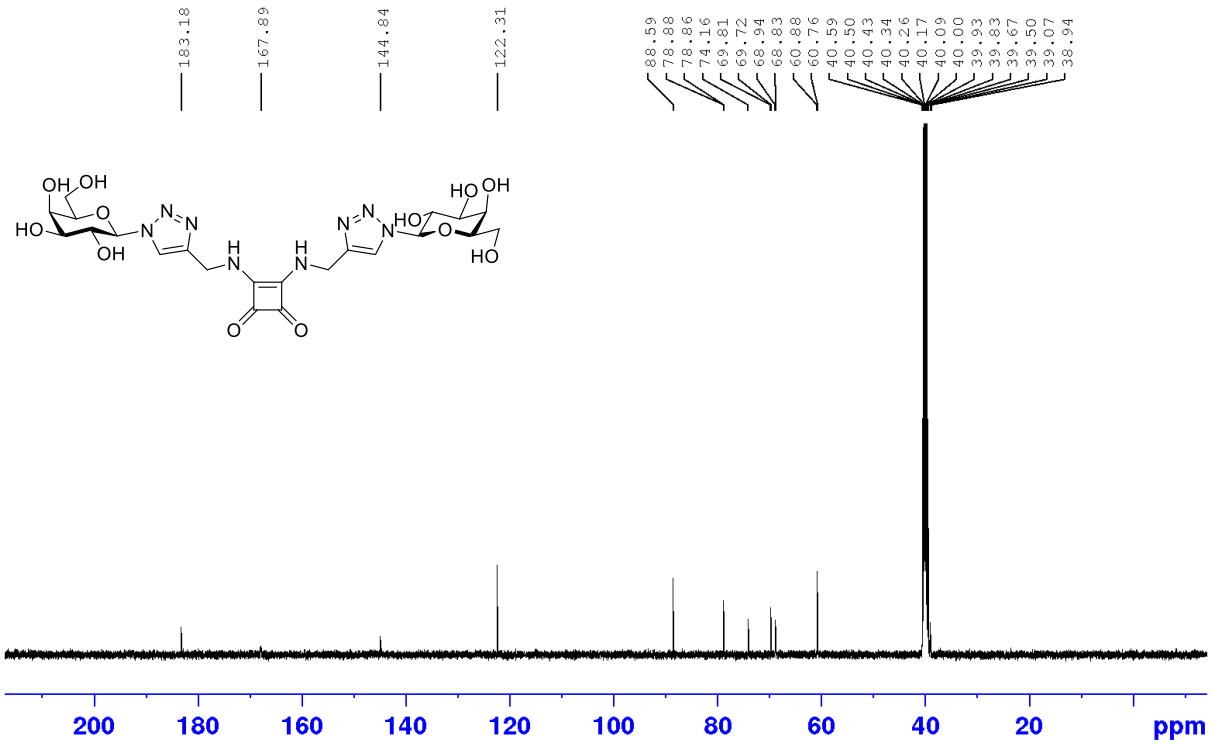

*[13]C NMR spectrum of 3,4-di(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-1,2,3-triazol-4-ylmethylamino)cyclobut-3-ene-1,2-dione* **8**



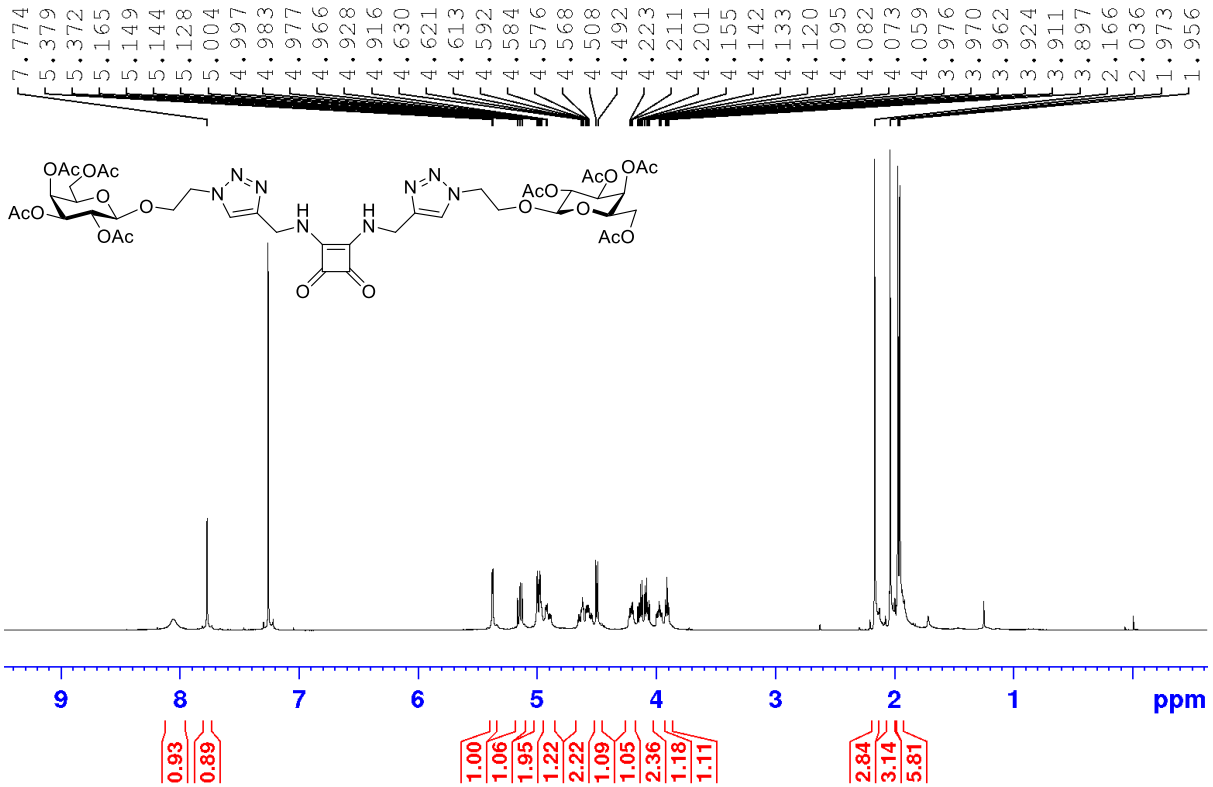
*[1]H NMR spectrum of 3,4-di(β-D-galactopyranosyl-1,2,3-triazol-4-ylmethylamino)cyclobut-3-ene-1,2-dione* **9**



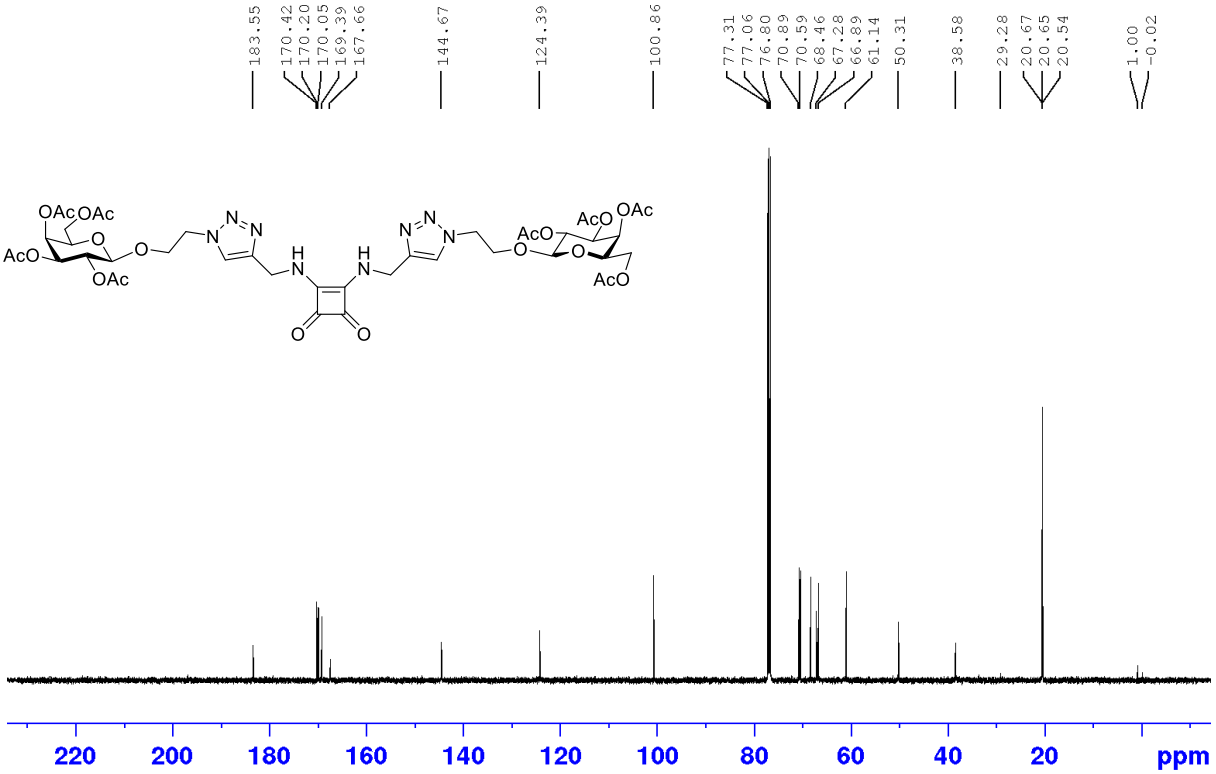

*[13]C NMR spectrum of 3,4-di(β-D-galactopyranosyl-1,2,3-triazol-4-ylmethylamino)cyclobut-3-ene-1,2-dione* **9**



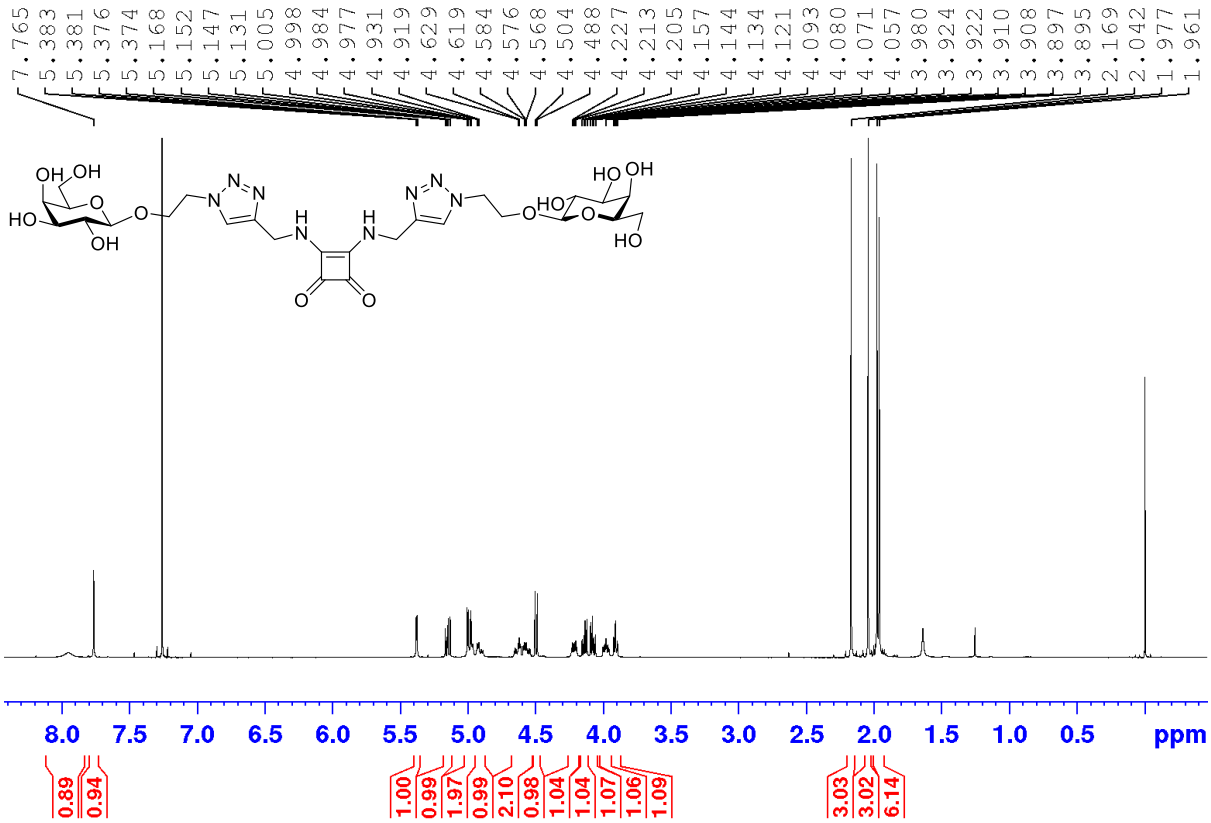
*[1]H NMR spectrum of 3,4-di(2-O-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl)-ethyl-1,2,3-triazol-4-ylmethylamino)cyclobut-3-ene-1,2-dione* **10**



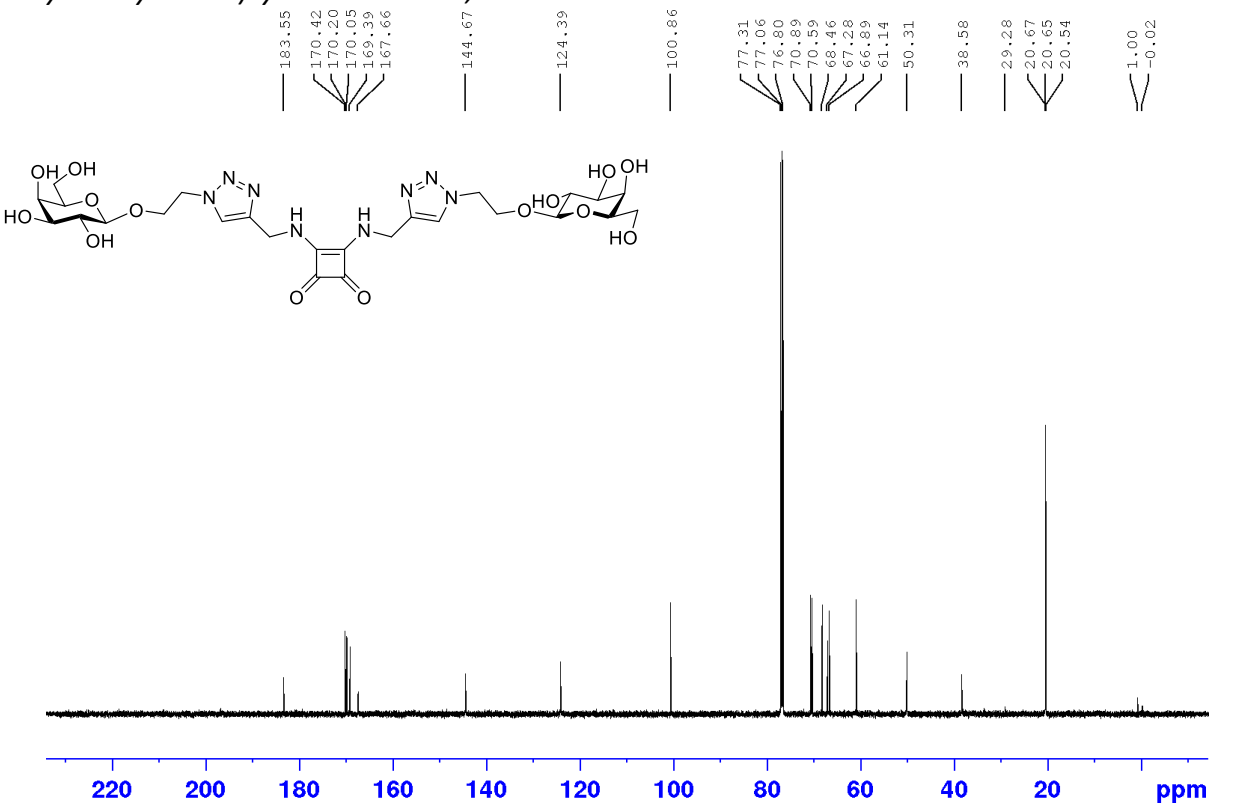

*[13]C NMR spectrum of 3,4-di(2-O-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl)-ethyl-1,2,3-triazol-4-ylmethylamino)cyclobut-3-ene-1,2-dione* **10**



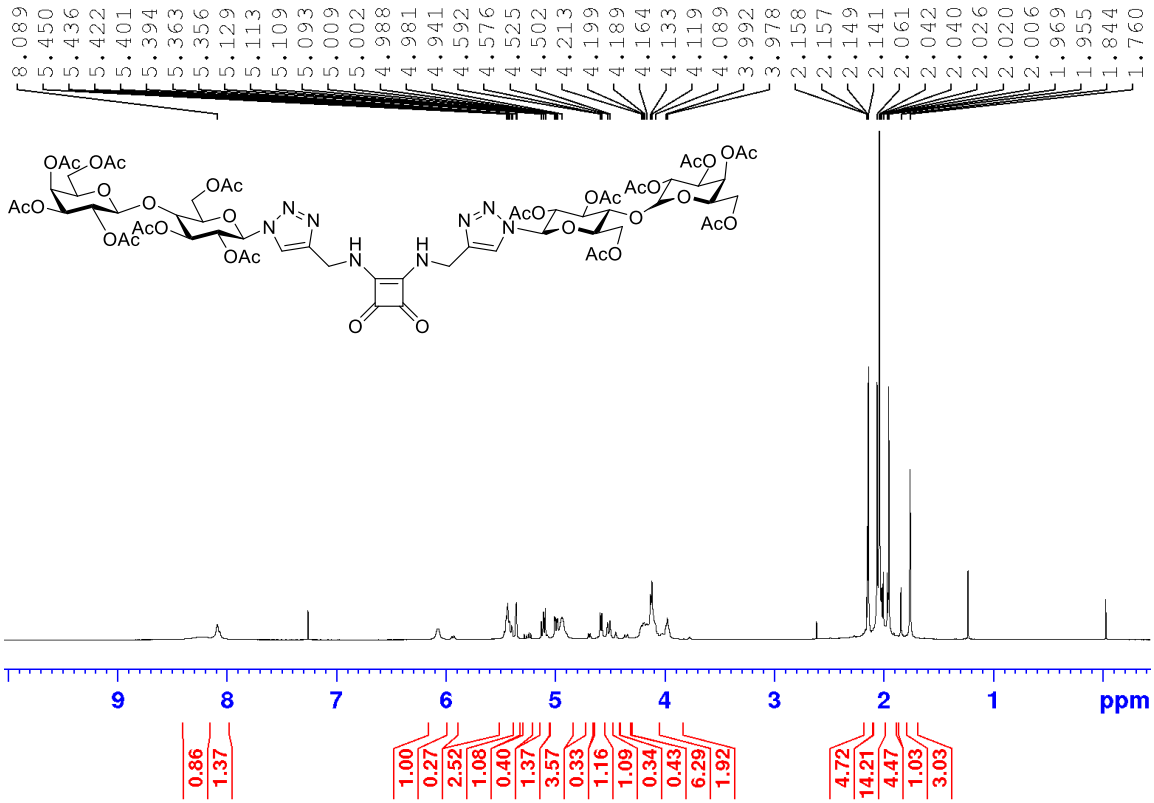
*$^{1}$H NMR spectrum of 3,4-di(2-O-(β-D-galactopyranosyl)-ethyl-1,2,3-triazol-4-ylmethylamino)cyclobut-3-ene-1,2-dione* **11**



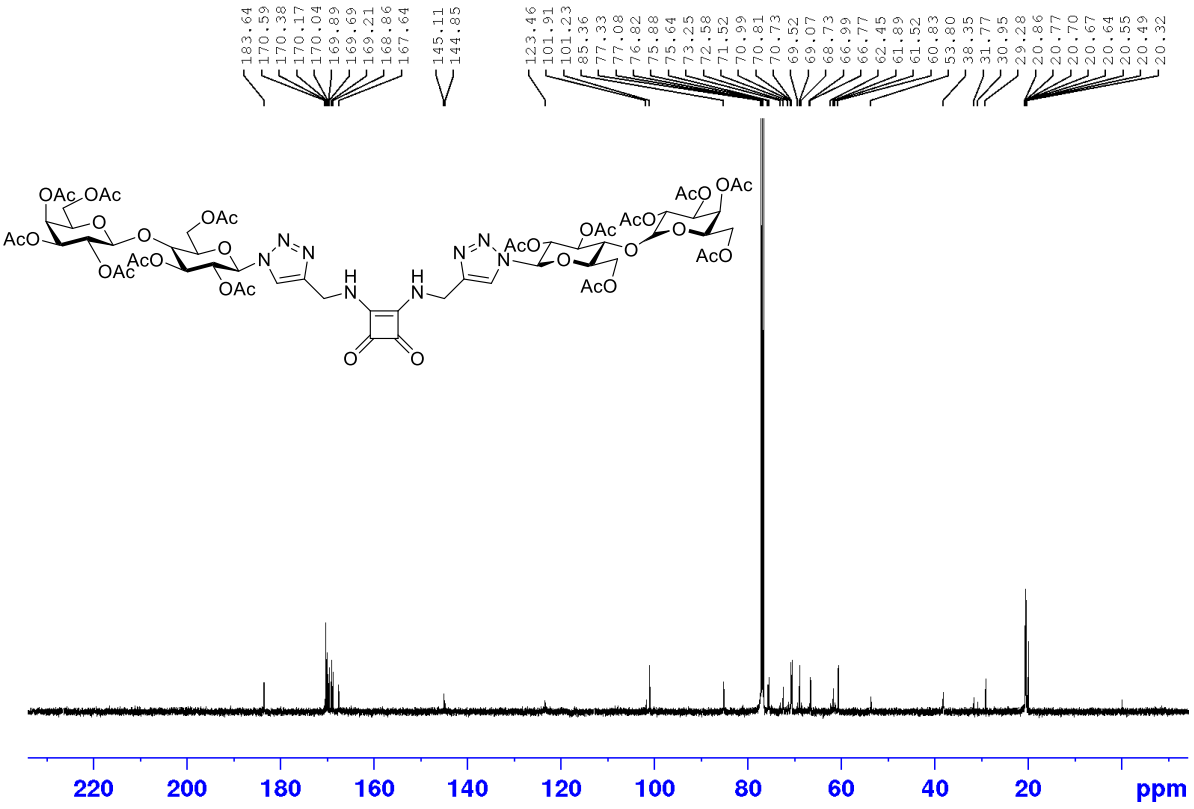

*$^{13}$C NMR spectrum of 3,4-di(2-O-(β-D-galactopyranosyl)-ethyl-1,2,3-triazol-4-ylmethylamino)cyclobut-3-ene-1,2-dione* **11**



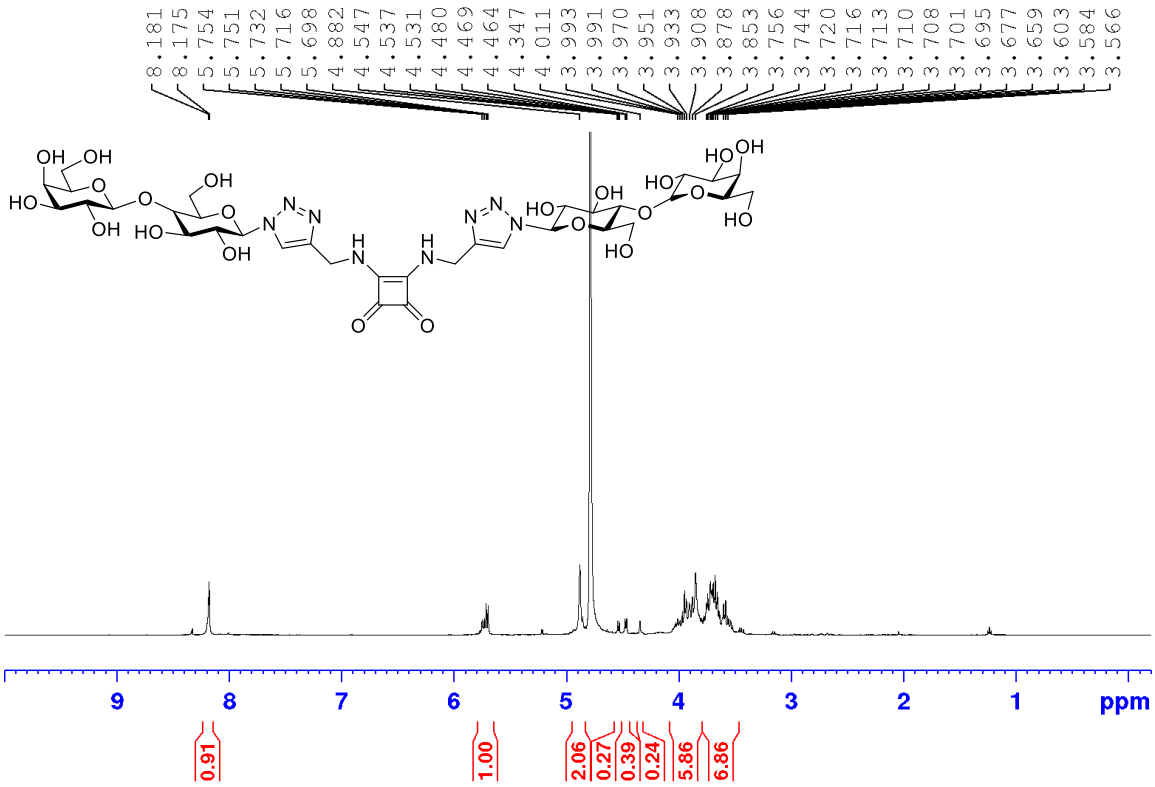
*[1]H NMR spectrum of 3,4-di-[{4-O-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl)-2,3,6-tri-O-acetyl-β-D- glucopyranosyl}-1,2,3-triazol-4-ylmethylamino]cyclobut-3-ene-1,2-dione* **12**



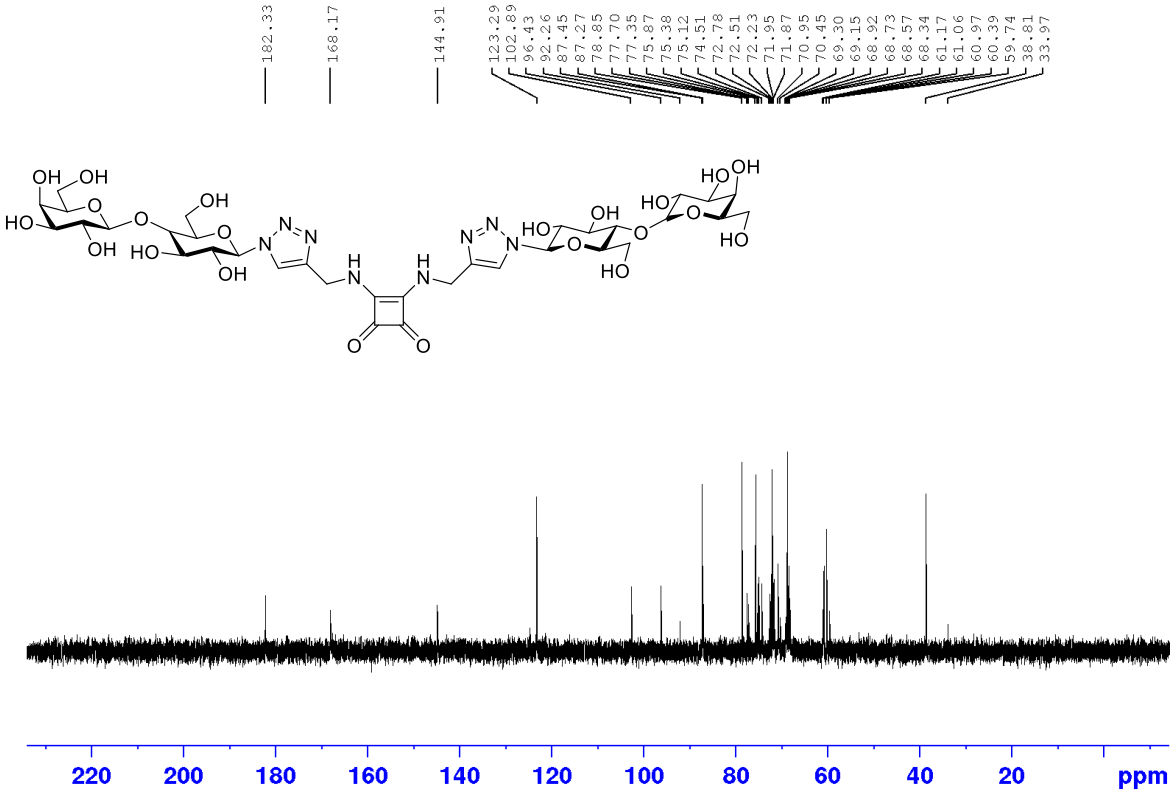

*[13]C NMR spectrum of 3,4-di(2-O-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl)-ethyl-1,2,3-triazol-4-ylmethylamino)cyclobut-3-ene-1,2-dione* **12**



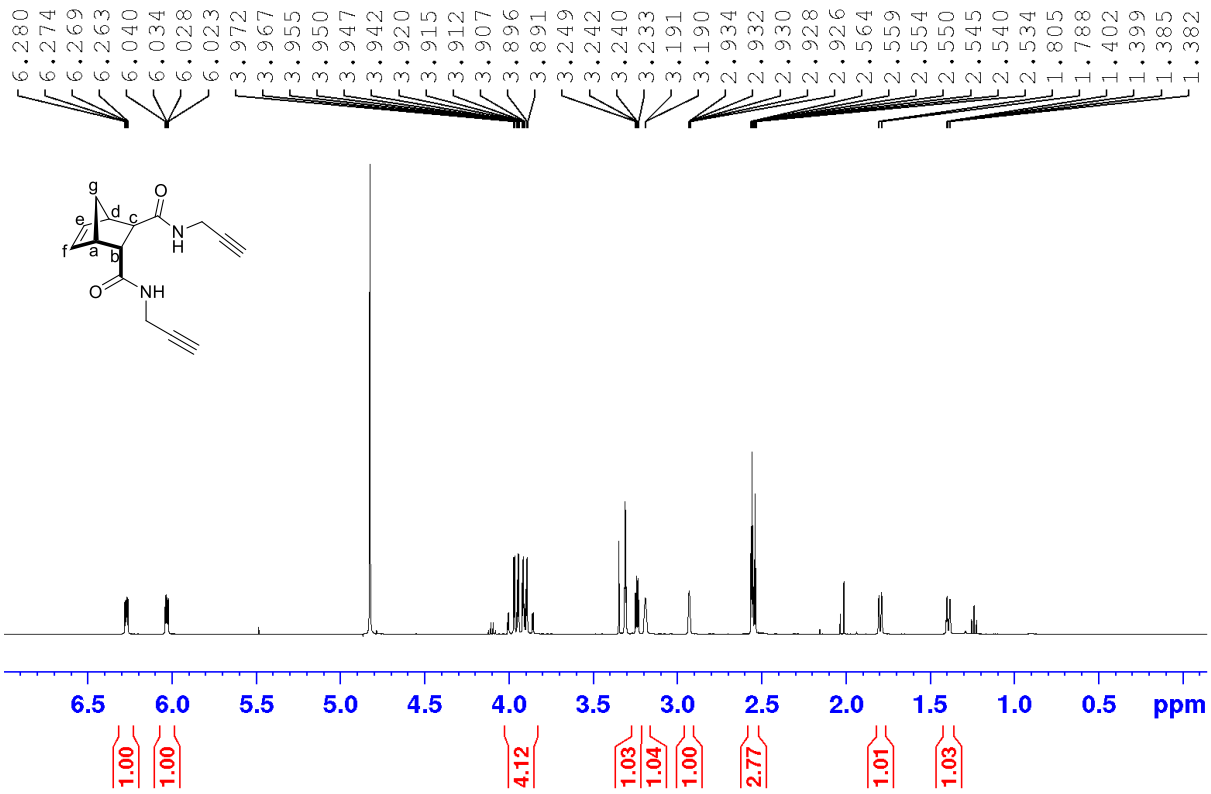
*[1]H NMR spectrum of 3,4-di-[{4-O-(β-D-galactopyranosyl)-β-D-glucopyranosyl}-1,2,3-triazol-4-ylmethylamino]cyclobut-3-ene-1,2-dione* **13**



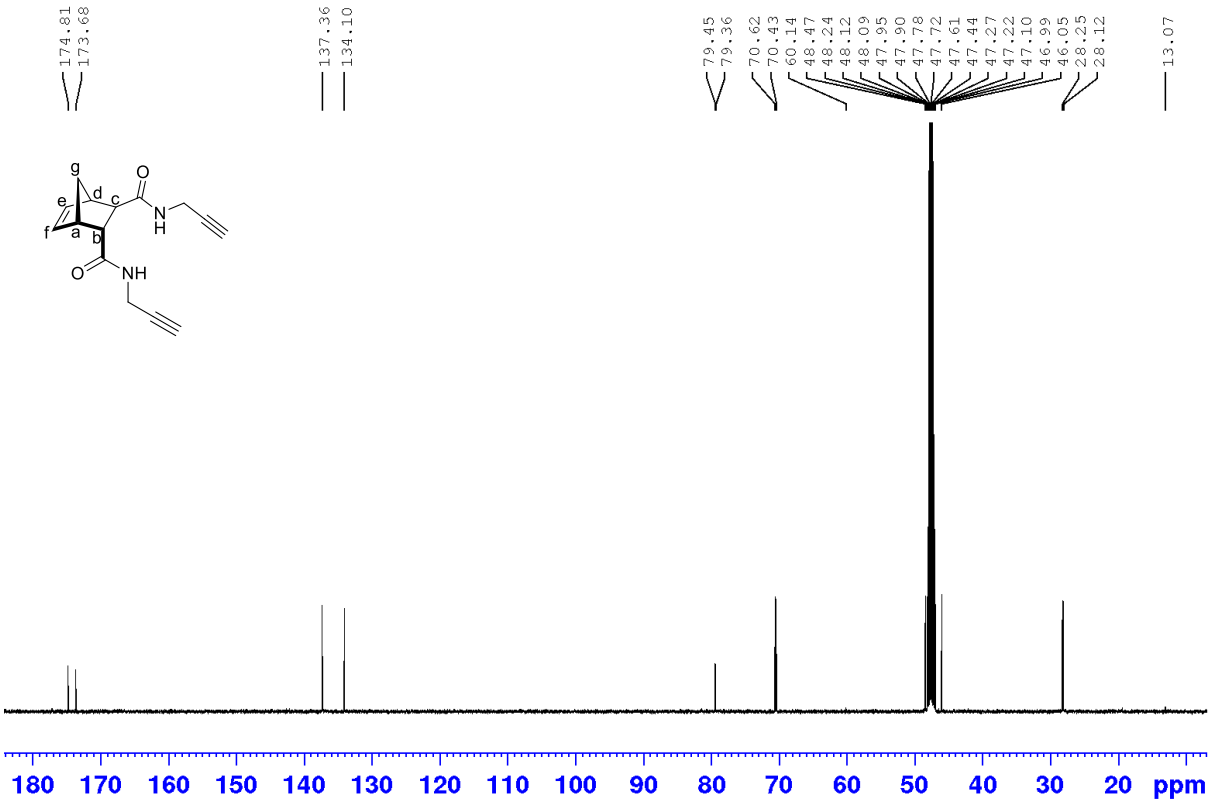

*[13]C NMR spectrum of 3,4-di-[{4-O-(β-D-galactopyranosyl)-β-D-glucopyranosyl}-1,2,3-triazol-4-ylmethylamino]cyclobut-3-ene-1,2-dione* **13**



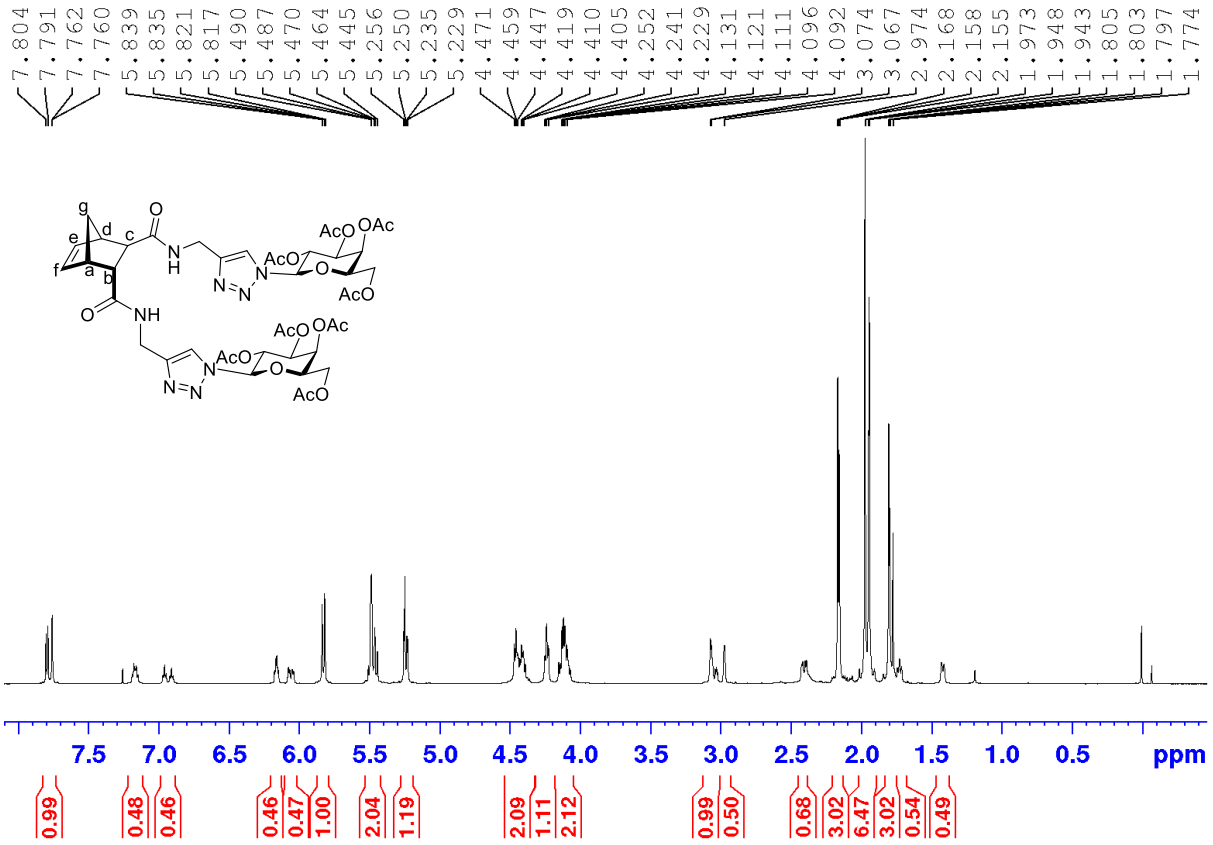
*[1]H NMR spectrum of bicyclo[2.2.1]hept-5-ene-2-endo,3-exo-2,3-dicarboxamide, N-(prop-2-yn-1-yl)* **15**



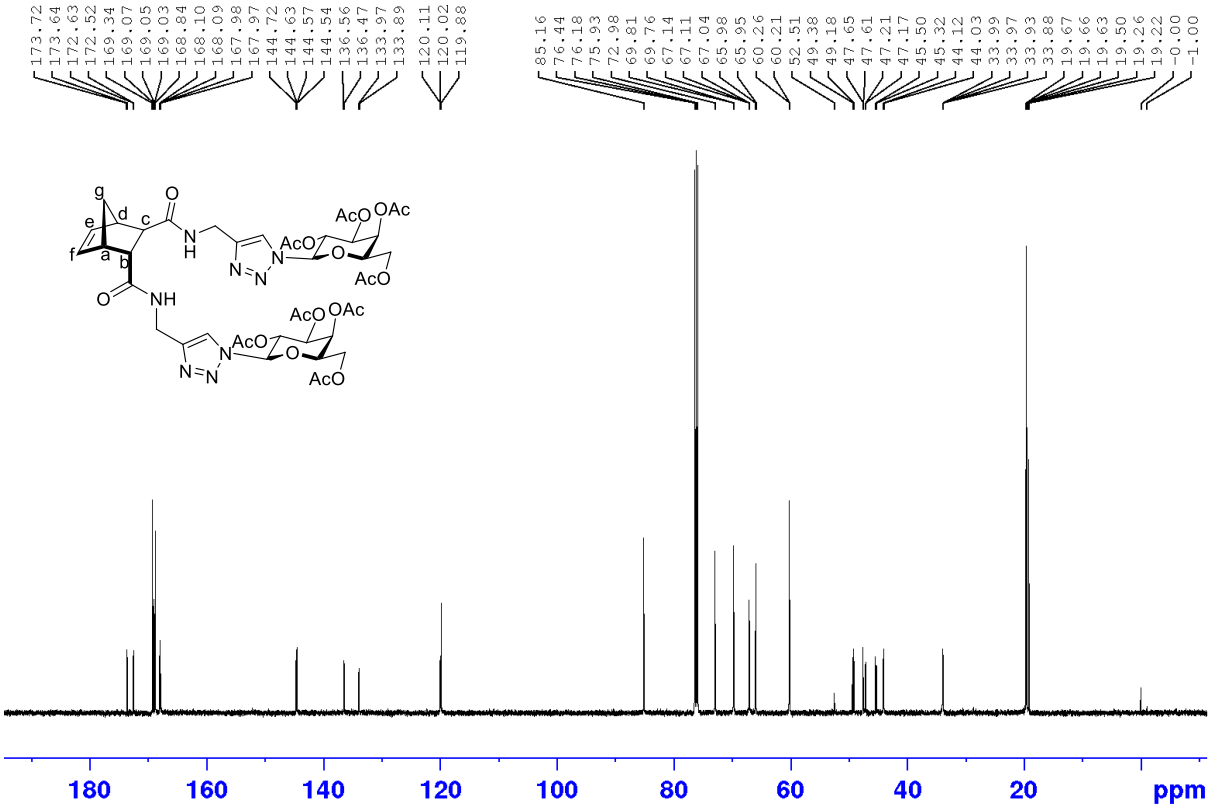

*[13]C NMR spectrum of bicyclo[2.2.1]hept-5-ene-2-endo,3-exo-2,3-dicarboxamide, N-(prop-2-yn-1-yl)* **15**



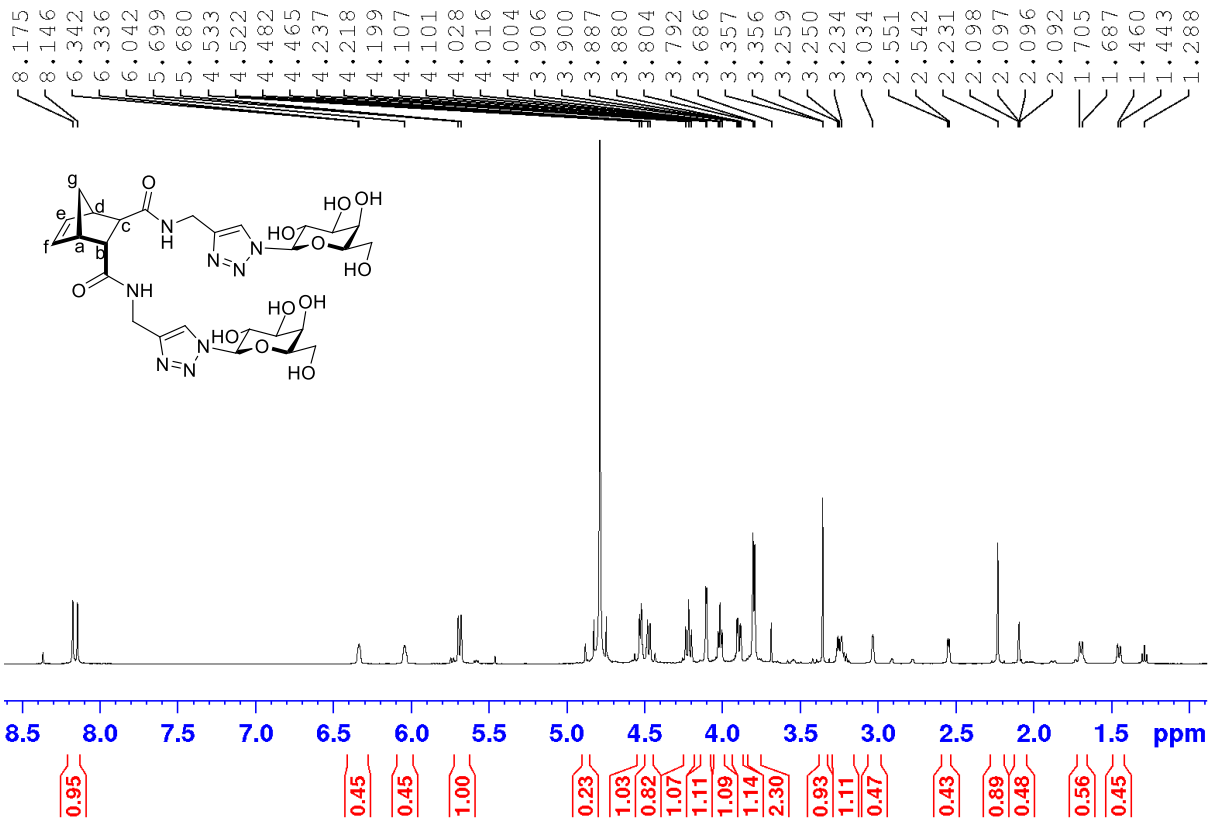
*[1]H NMR spectrum of bicyclo[2.2.1]hept-5-ene-2-endo,3-exo-2,3-dicarboxamide, N-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-1,2,3-triazol-4-ylmethyl)* **16**



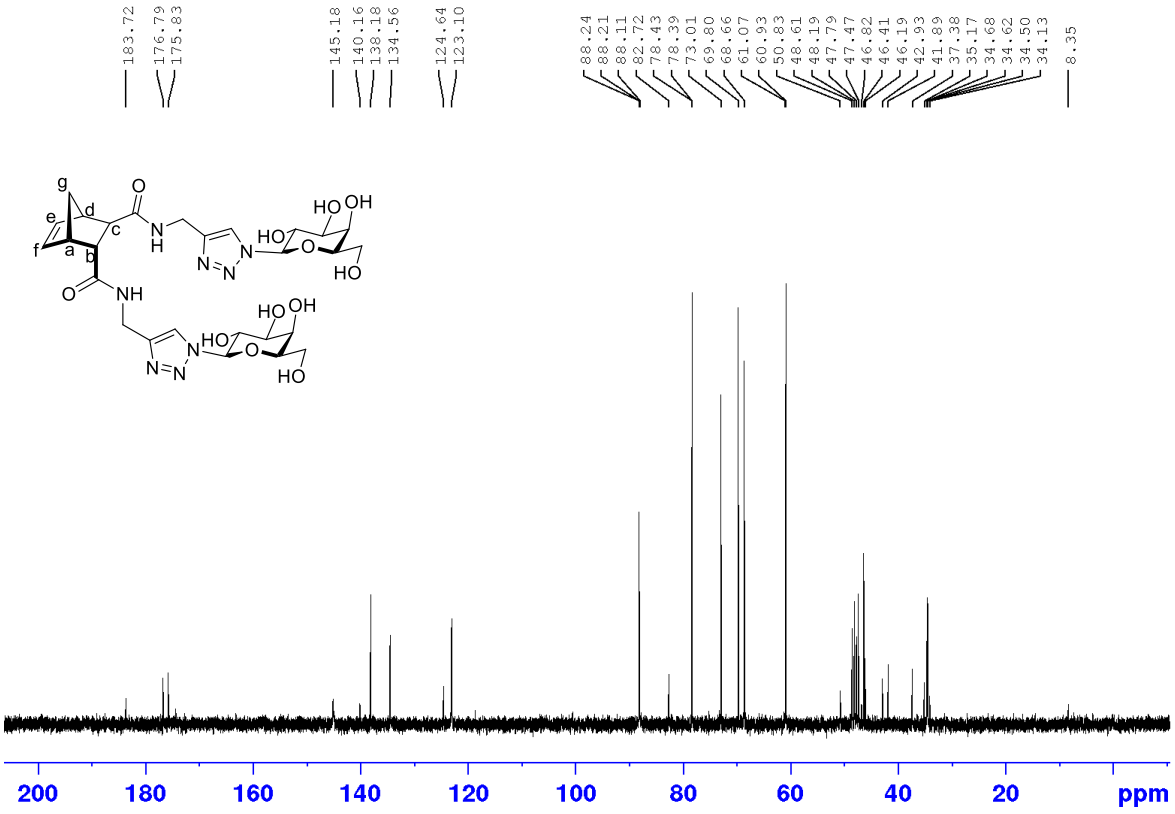

*[13]C NMR spectrum of bicylco[2.2.1]hept-5-ene-2-endo,3-exo-2,3-dicarboxamide, N-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-1,2,3-triazol-4-ylmethyl)* **16**



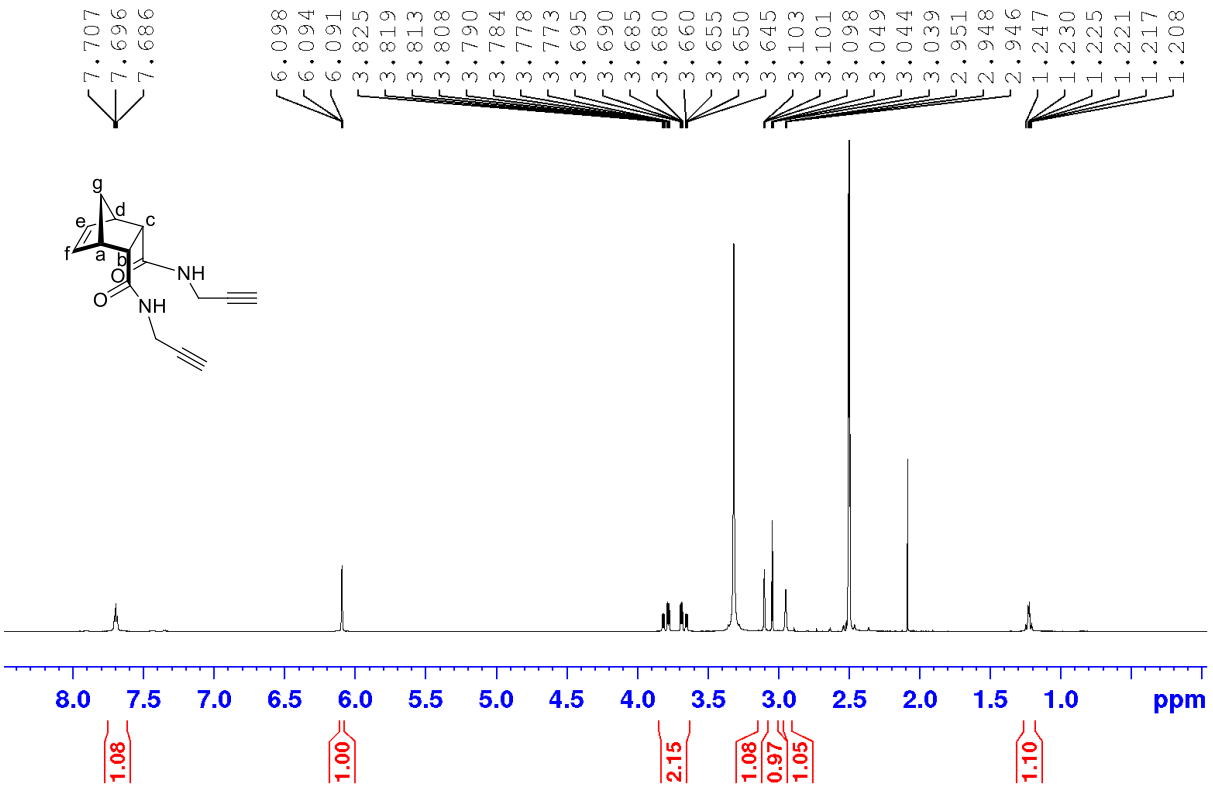
*[1]H NMR spectrum of bicyclo[2.2.1] hept-5-ene-2-endo,3-exo-2,3-dicarboxamide, N-(β-D-galactopyranosyl-1,2,3-triazol-4-ylmethyl)* **17**



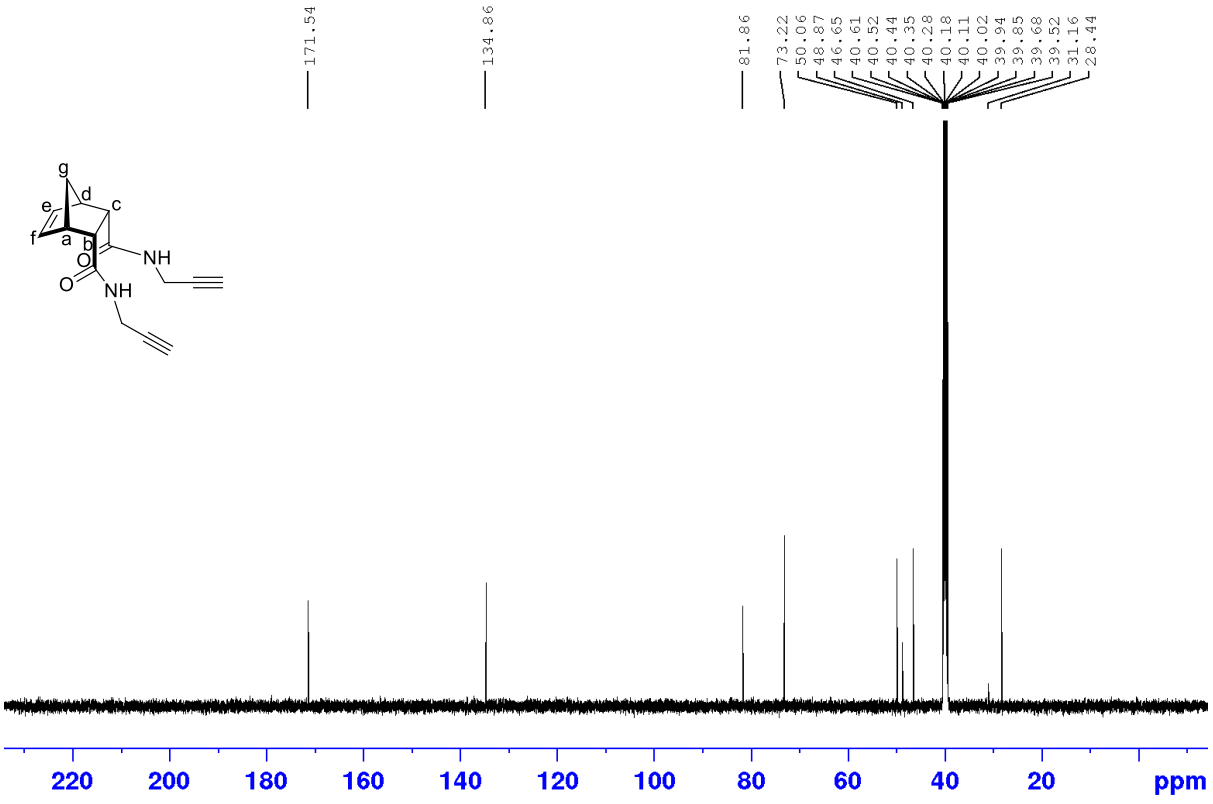

*[13]C NMR spectrum of bicyclo[2.2.1]hept-5-ene-2-endo,3-exo-2,3-dicarboxamide, N-(β-D-galactopyranosyl-1,2,3-triazol-4-ylmethyl)* **17**



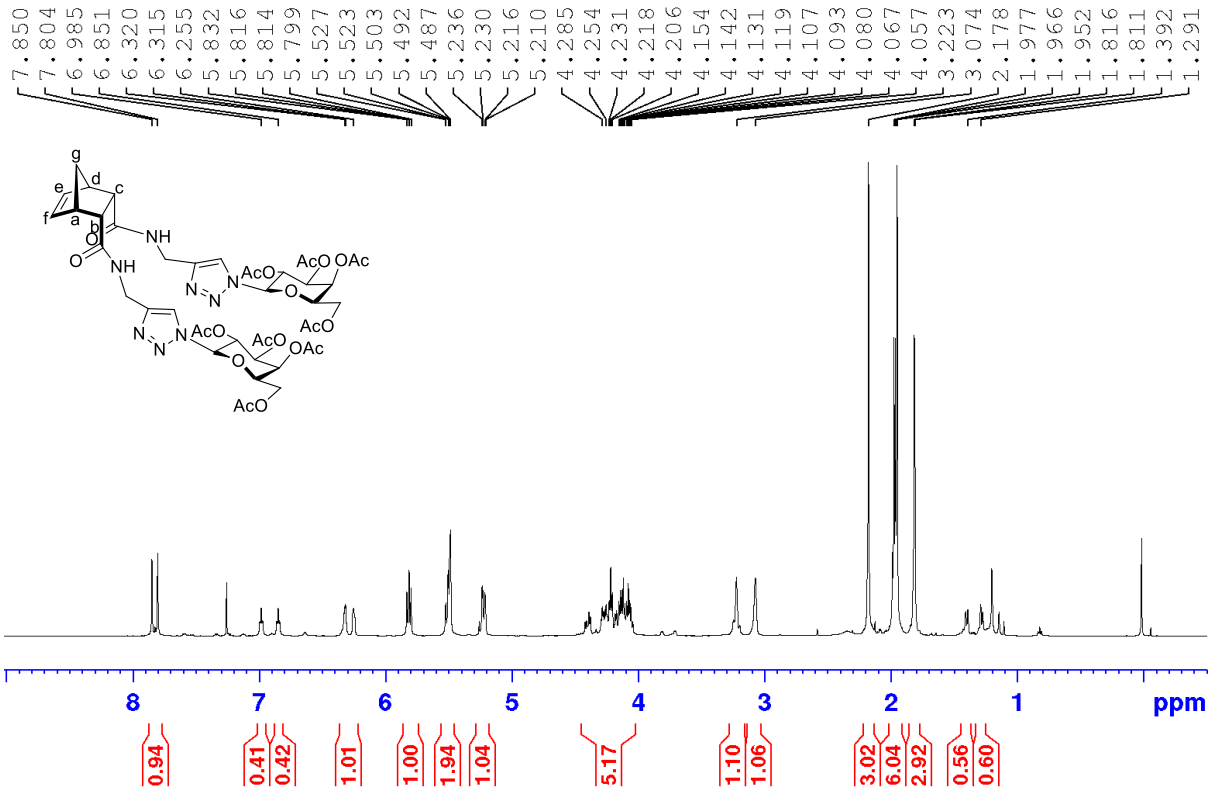
*[1]H NMR spectrum of bicyclo[2.2.1]hept-5-ene-2,3-endo-2,3-dicarboxamide, N-(prop-2-yn-1-yl)* **19**



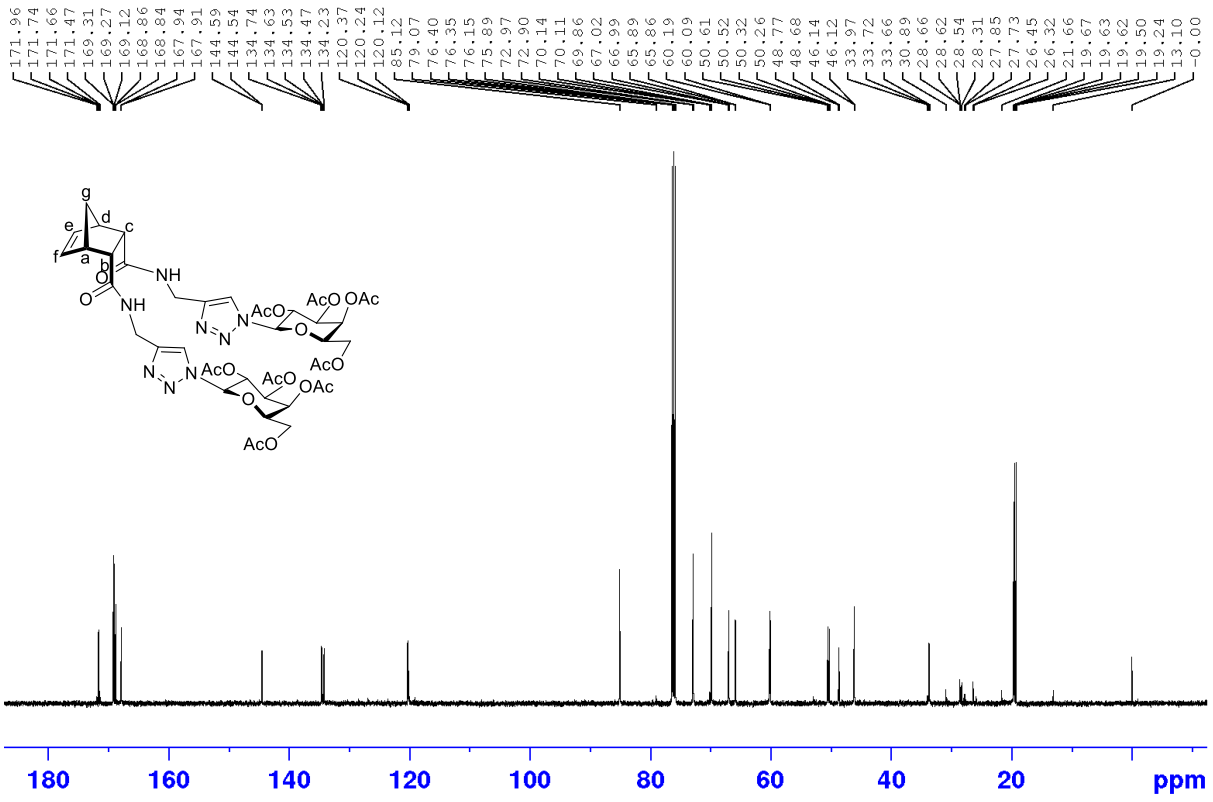

*[13]C NMR spectrum of bicyclo[2.2.1]hept-5-ene-2,3-endo-2,3-dicarboxamide, N-(prop-2-yn-1-yl)* **19**



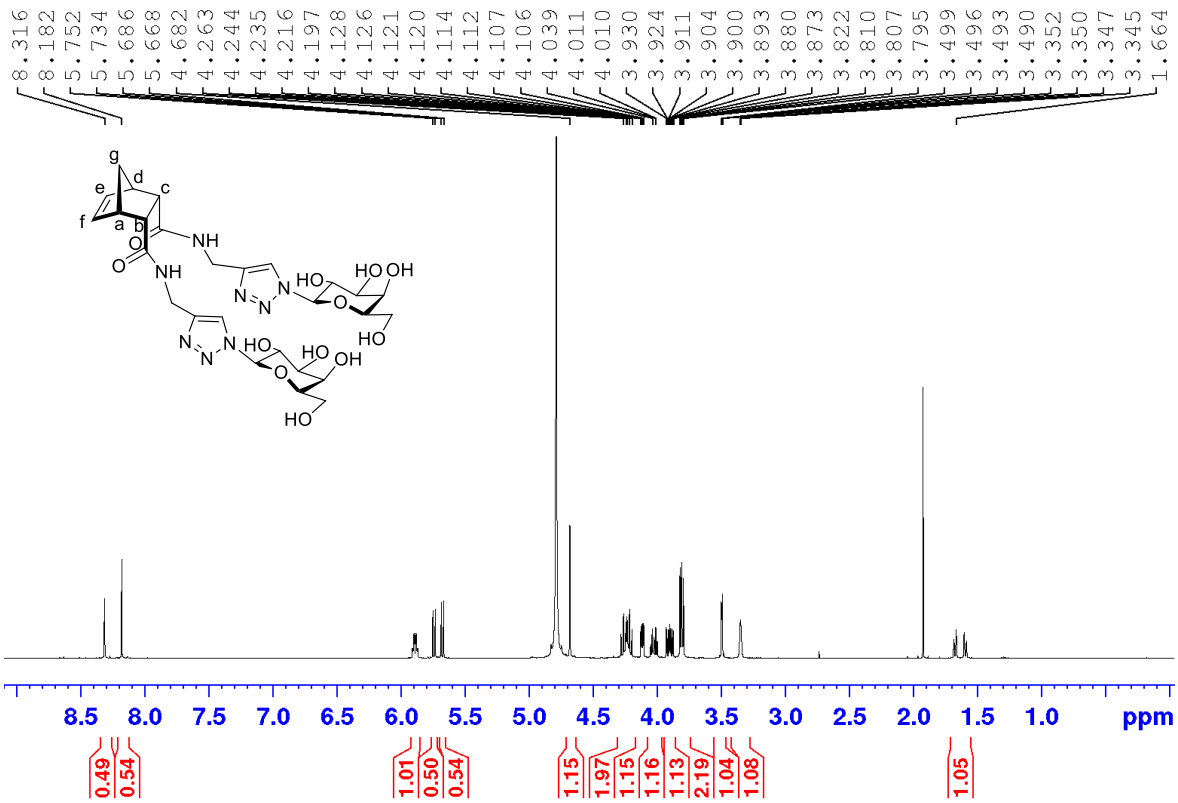
*$^{1}$H NMR spectrum of bicylco[2. 2.1]cis-hept-5-ene-2,3-endo-2,3-dicarboxamide, N-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-1,2,3-triazol-4-ylmethyl)* **20**



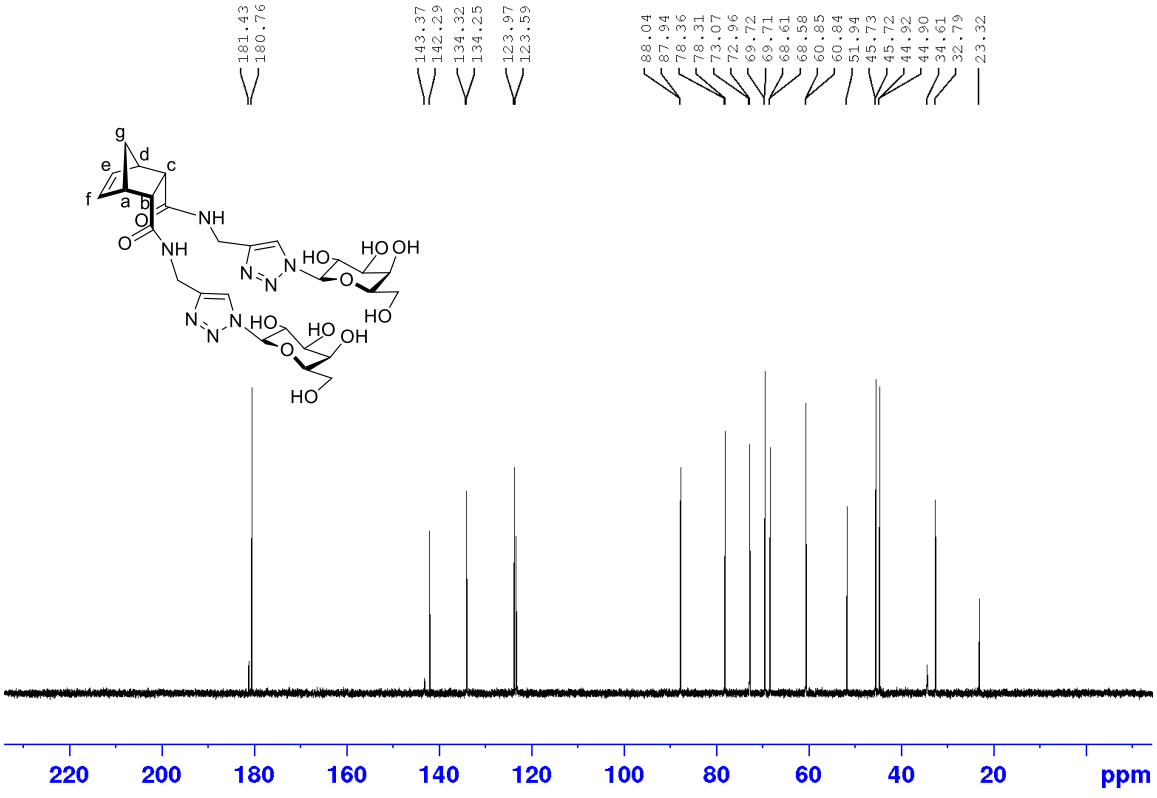

*$^{13}$C NMR spectrum of bicylco[2.2.1]cis-hept-5-ene-2,3-endo-2,3-dicarboxamide, N-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-1,2,3-triazol-4-ylmethyl)* **20**



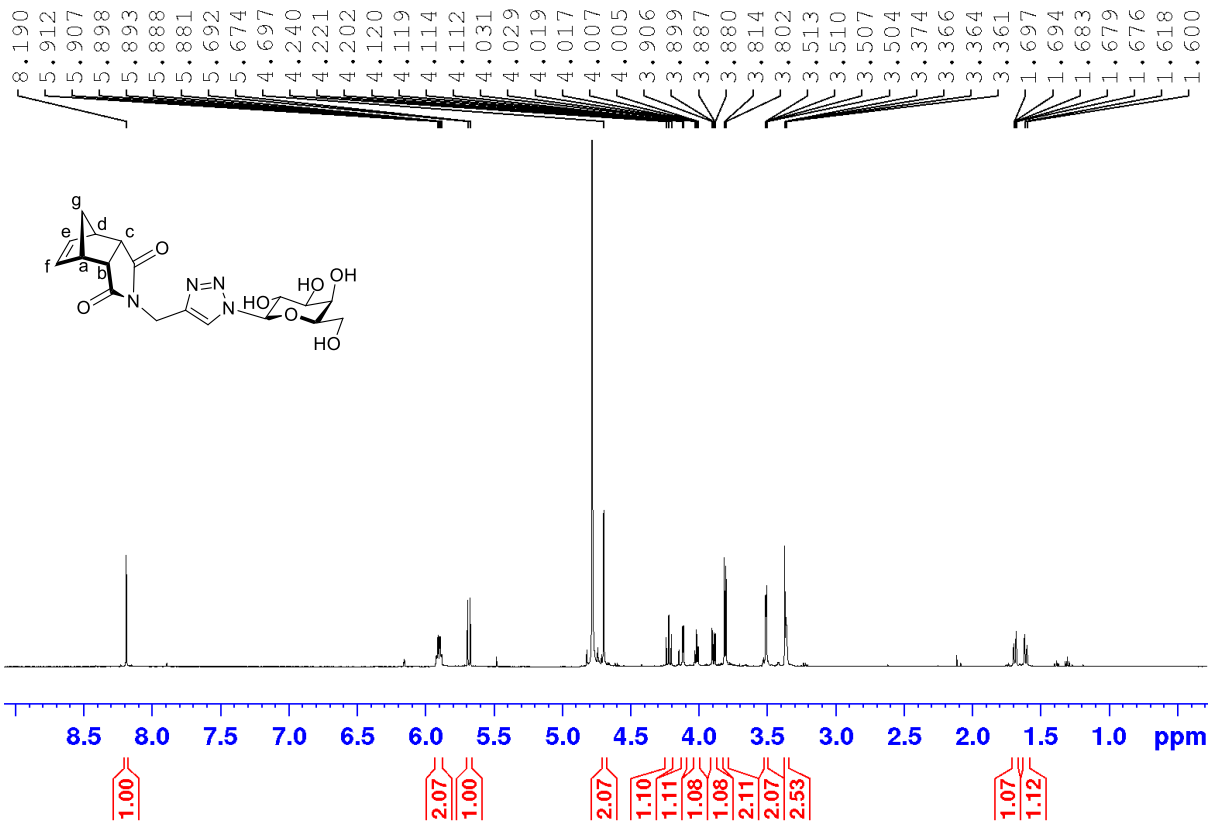
*[1]H NMR spectrum of bicylco[2.2.1]cis-hept-5-ene-2,3-endo-2,3-dicarboxamide, N-(β-D-galactopyranosyl-1,2,3-triazol-4-ylmethyl)* **21**



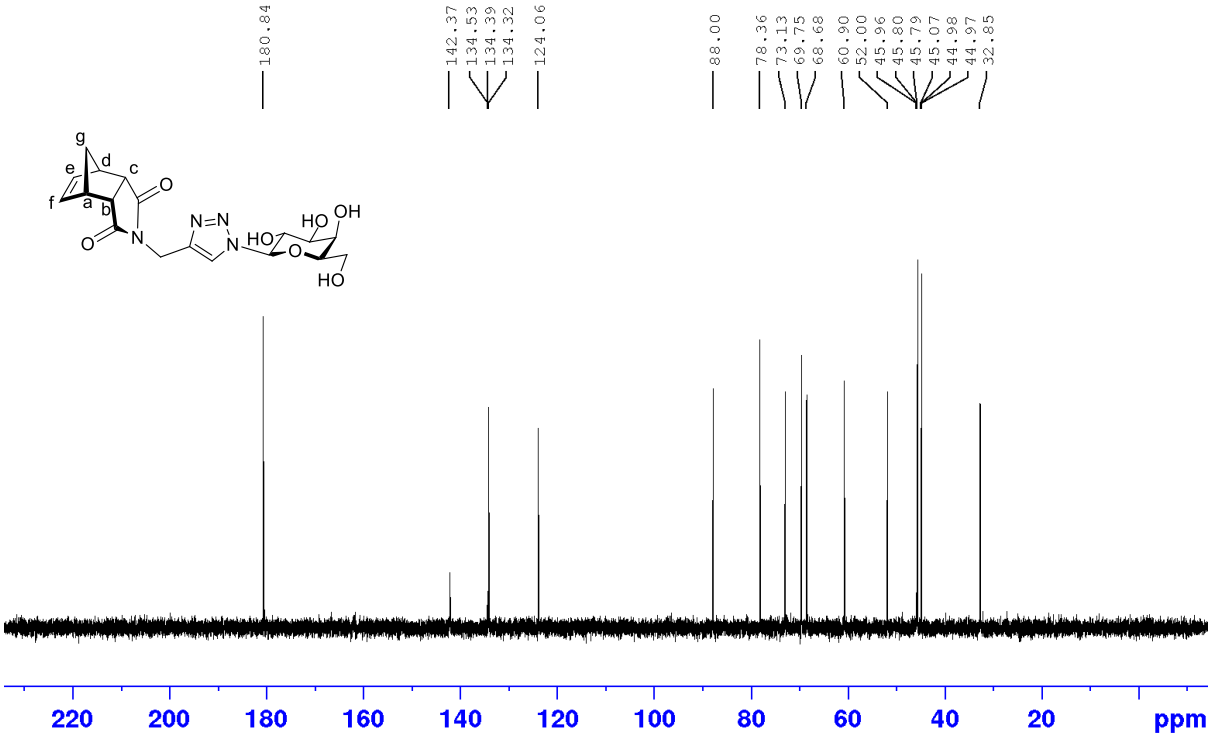

*[13]C NMR spectrum of bicylco[2.2.1]cis-hept-5-ene-2,3-endo-2,3-dicarboxamide, N-(β-D-galactopyranosyl-1,2,3-triazol-4-ylmethyl)* **21**



*[1]H NMR spectrum N-(β-D-galactopyranosyl-1,2,3-triazol-4-ylmethyl)biocylco[2.2.1]cis-hept-5-ene-2,3-endo-dicarboximide* **22**




*[13]C NMR spectrum of N-(β-D-galactopyranosyl-1,2,3-triazol-4-ylmethyl)biocylco[2.2.1]cis-hept-5-ene-2,3-endo-dicarboximide* **22**



## 3- Additional Figures Biological Evaluation

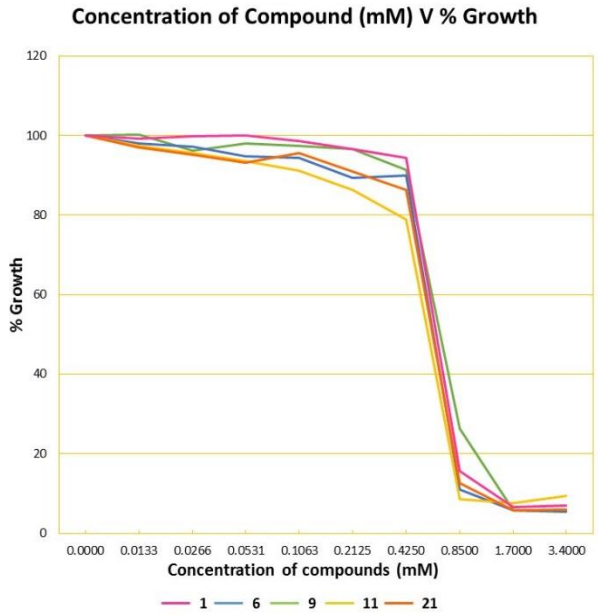



**Figure SI.1:** Toxicity of glycoconjugates against *Candida albicans*.

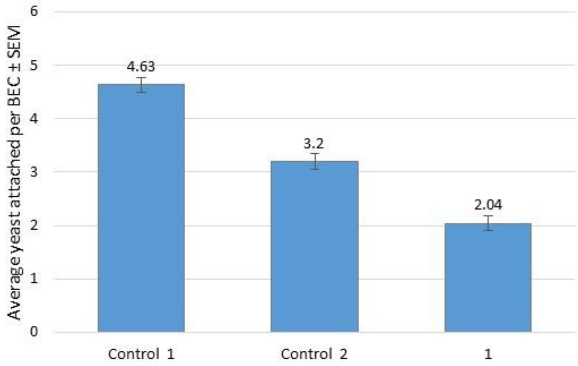



**Figure SI.2:** Displacement assay of compound **1** at 138 µM.

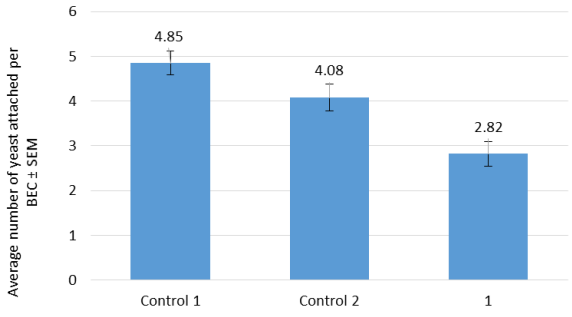



**Figure SI.3:** Displacement assay of compound **1** at 13.8 mM.

## 3. Additional Figures Computational Analysis

Initial pre-screening of the conformational space was achieved by means of the Conformer-Rotamer Ensemble Sampling Tool (crest) that utilizes the GFN2-xTB method. The method is designed around a semiempirical tight-binding quantum model to facilitate efficient and robust screening of the conformational space of large molecular systems. This procedure runs through iterative metadynamics sampling and subsequent pre-optimizations steps of selected MTD snapshots, narrowing the number of potential conformers down to 150 and 35 unique conformers for compound **1** and **21**, respectively. Starting from chemically reasonable starting geometries of **1** and **21**, the default conformer search algorithm implemented was utilised to automatically generate an ensemble of low-energy conformers and rotamers. By this technique, the initially >2000 structures generated by the independent MTD runs were narrowed down to 150 and 35 unique conformers for compounds **1** (Figure SI.6) and **21** (Figure SI.8), respectively, based on the GFN2-xTB energy criterion. These final structures were subsequently subjected to more accurate density functional theory (DFT) geometry optimisations (SMD-B3PW91/6-31G**) and single-point energies (SMD-B3PW91/D3/6-31G**), which allowed for their ranking in terms of their relative enthalpies (Table SI.1 and SI.2). The geometry of one of the representative low-energy conformers is stabilised by several hydrogen bonding interactions. One of the two side-arms is slightly folded due to the presence of a hydrogen bond between a galactosyl OH group and the amide carbonyl group attached to the central benzene ring. This "semi-open" structural motif is also featured in the other two low-energy conformers. Alternatively, hydrogen bonds can be formed between the two terminal galactosyl units, furnishing conformers in which these units are grouped together to maximise their contacts giving rise to asymmetrically globular structural motifs (Figure SI.4). Inclusion of dispersion corrections (D3) during geometry optimisations yielded geometries that were more densely packed, an apparent artefact of the presence of the atom pair-wise dispersion corrections which in the present case seem to maximise the contacts between individual parts of the molecules (Figure SI.5). These artefacts may stem from the lack of explicit solvent molecules in the calculations, resulting in overestimation of intramolecular dispersion.

The overall structures of the majority of low-energy conformers of compound **21** are best described as basket-shaped, stabilised through intramolecular hydrogen bonding between the galatosyl units and triazolyl linkers of adjacent branches.

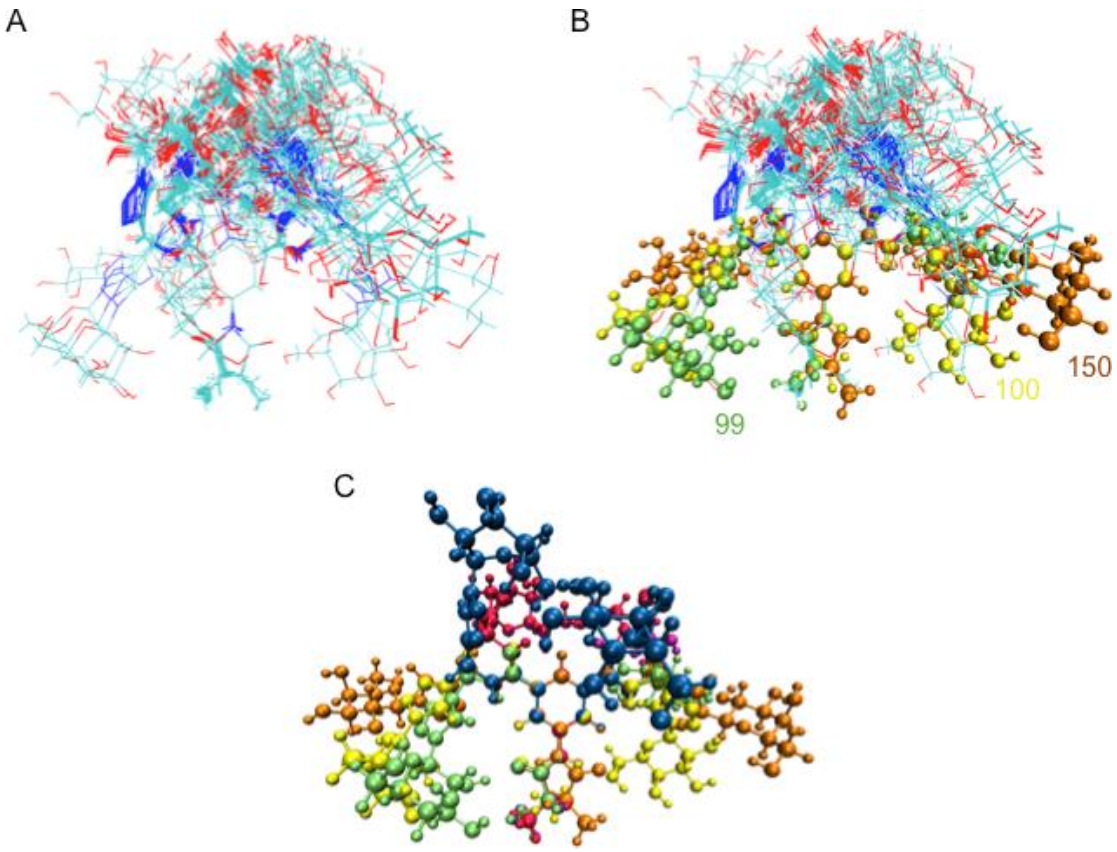



**Figure SI.4**. Overlay of optimised geometries highlighting selected conformers for compound **1** in colour and CPK representation along with line representations for remaining conformers: (A) all conformers (B) 3 lowest-energy conformers within a 2.5 kcal $mol^{-1}$ Free energy window along with all other conformers, (C) 6 low-energy conformers within a 3.0 kcal $mol^{-1}$ energy window measured by either $\mathbf{\Delta}G$ or $\mathbf{\Delta}H$. Numbers refer to the specified conformer.

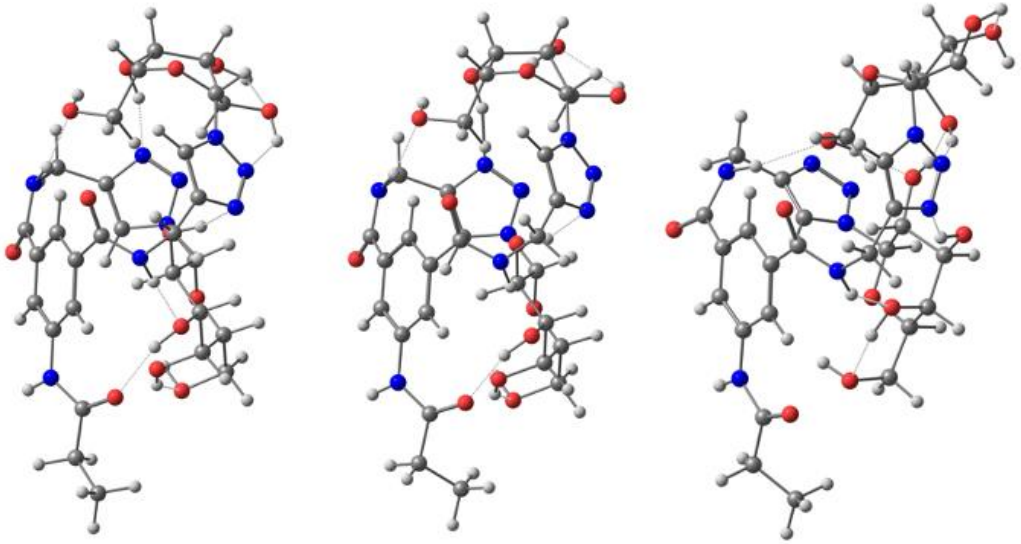
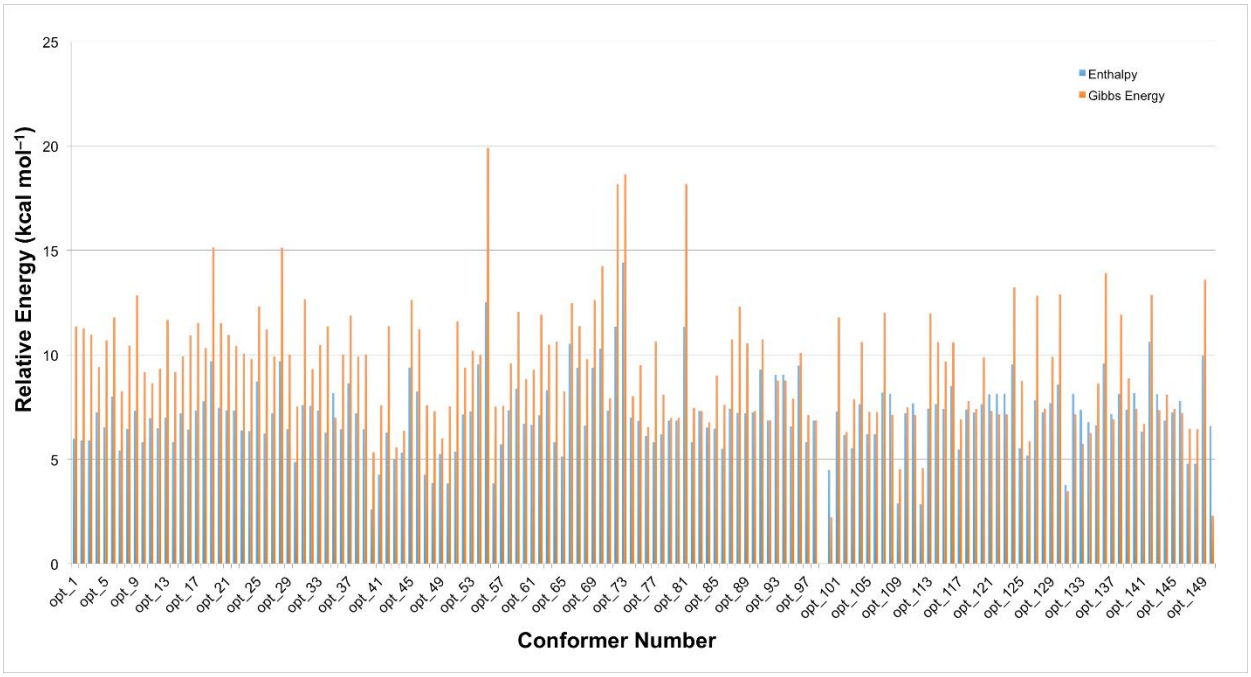


**Figure SI.5**. Three examples of closely packed conformations of compound **1** obtained from geometry optimisations at the SMD-B3PW91-D3/6-31G** level of theory.

**Table SI.1**. Relative energies (in kcal $mol^{-1}$) of selected low-energy conformers for compound **1** calculated at different levels of theory. Thermal and entropic corrections to the electronic energy were taken from the B3PW91/6-31G** frequency calculations.

| | SMD-B3PW91-D3/6-31G** | | DSD-PBEP86-D3/def2-TZVPP | | DLPNO-CCSD(T)/def2-TZVPP | |
|---|---|---|---|---|---|---|
| Conformer | **Δ***H* | **Δ***G* | **Δ***H* | **Δ***G* | **Δ***H* | **Δ***G* |
| 99 | 3.6 | 0.9 | 3.6 | 0.9 | 3.2 | 0.5 |
| 100 | 11.3 | 6.2 | 9.4 | 4.4 | 8.9 | 3.9 |
| 150 | 16.9 | 9.8 | 19.8 | 12.8 | 19.1 | 12.1 |
| 109 | 2.7 | 1.6 | 3.3 | 2.2 | 3.1 | 2.0 |
| 112 | 2.7 | 1.7 | 3.2 | 2.2 | 3.1 | 2.1 |
| 40 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |



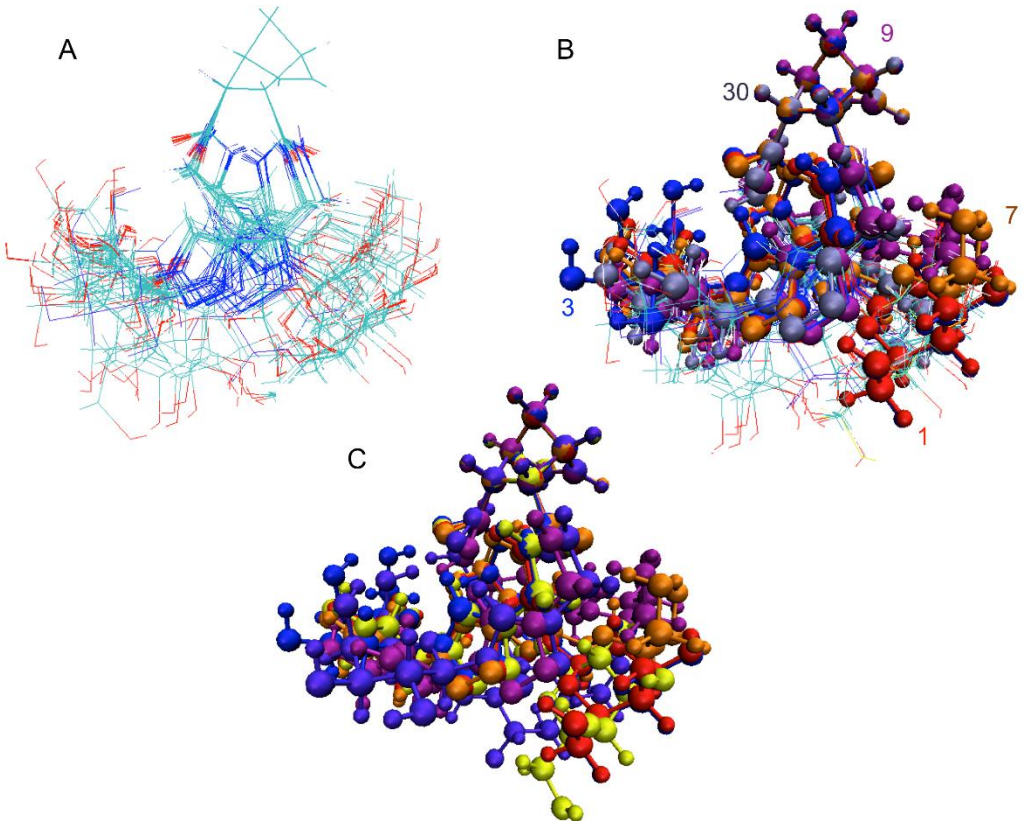




**Figure SI.7**. Overlay of optimised geometries of compound **21** highlighting selected conformers in colour and CPK representation along with line representations for remaining conformers: (A) all conformers (B) 5 lowest-energy conformers within a 2.5 kcal mol$^{-1}$ Free energy window along with all other conformers, (C) low-energy conformers measured by either **Δ***G* or **Δ***H*. Numbers refer to the specified conformer.

**Table SI.2**. Relative energies (in kcal mol$^{-1}$) of selected low-energy conformers for compound **21** obtained at different levels of theory. Thermal and entropic corrections to the electronic energy were taken from the B3PW91/6-31G** frequency calculations.

| | SMD-B3PW91-D3/6-31G** | | DSD-PBEP86-D3/def2-TZVPP | | DLPNO-CCSD(T)/def2-TZVPP | |
|---|---|---|---|---|---|---|
| Conformer | **Δ***H* | **Δ***G* | **Δ***H* | **Δ***G* | **Δ***H* | **Δ***G* |
| 1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 3 | 4.5 | 3.8 | 6.8 | 6.1 | 6.4 | 5.7 |
| 7 | 4.8 | 3.6 | 6.9 | 5.8 | 6.4 | 5.2 |
| 9 | 5.6 | 3.4 | 10.7 | 8.5 | 9.7 | 7.6 |
| 12 | 4.8 | 4.5 | 7.1 | 6.9 | 6.7 | 6.4 |
| 30 | 13.0 | 9.2 | 20.5 | 16.6 | 19.5 | 15.6 |
| 31 | 11.5 | 10.7 | 17.9 | 17.1 | 16.4 | 15.5 |

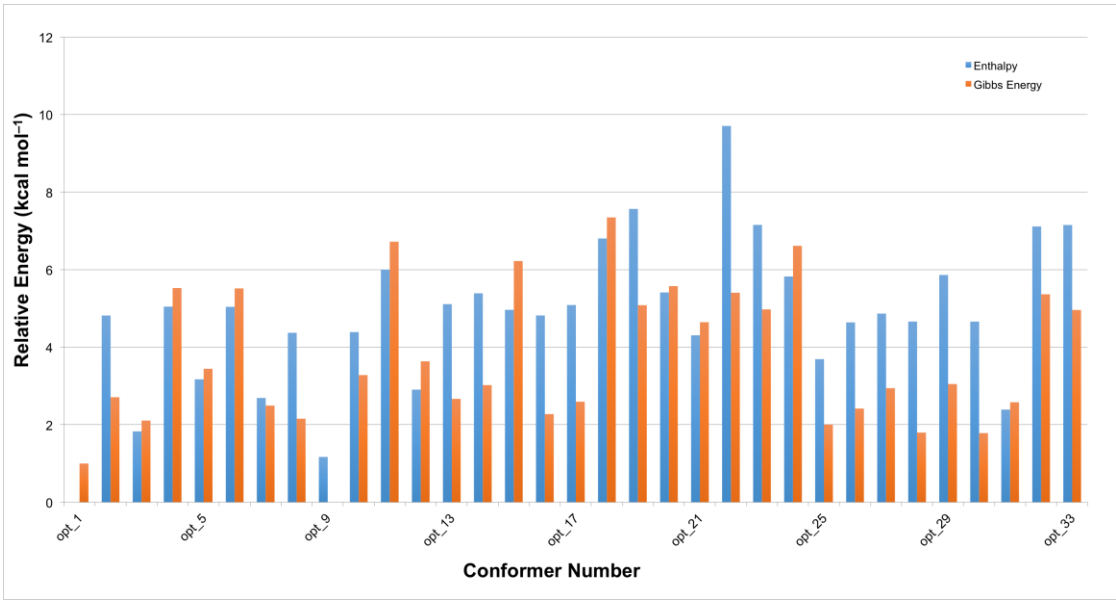



**Figure SI.8**. Relative energies of DFT-optimised (B3PW91/6-31G**) conformers of compound **21**.

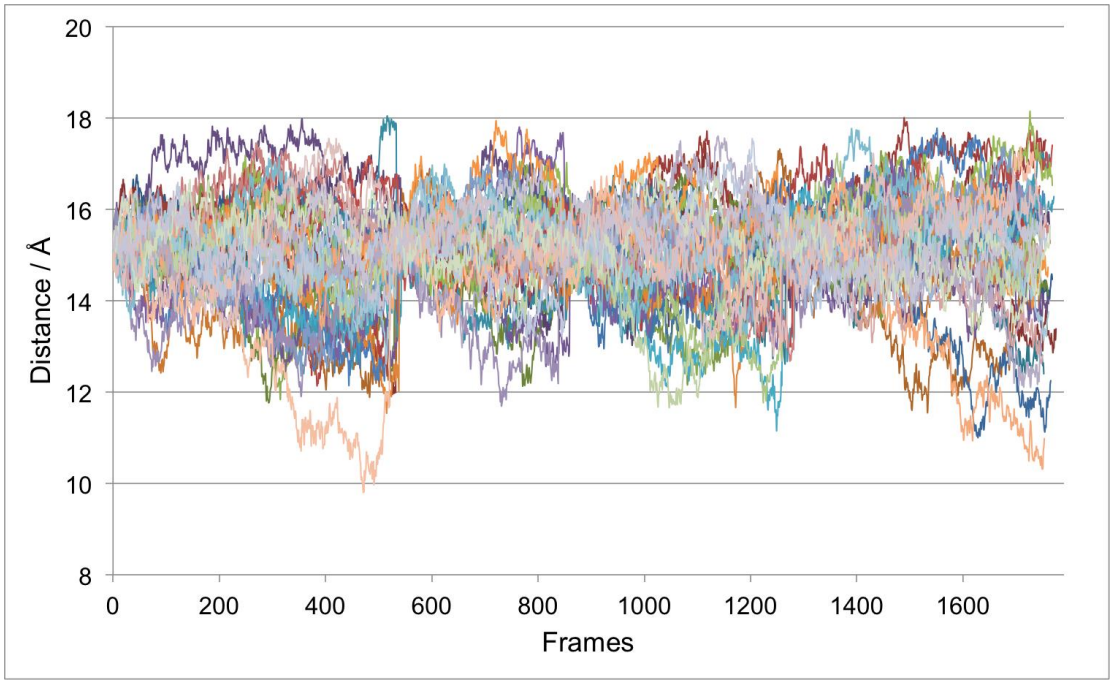



**Figure SI.9**. Distance between anomeric centres along MTD trajectory for compound **1**. Each line represents the trajectory generated by a single walker (40 walkers in total).

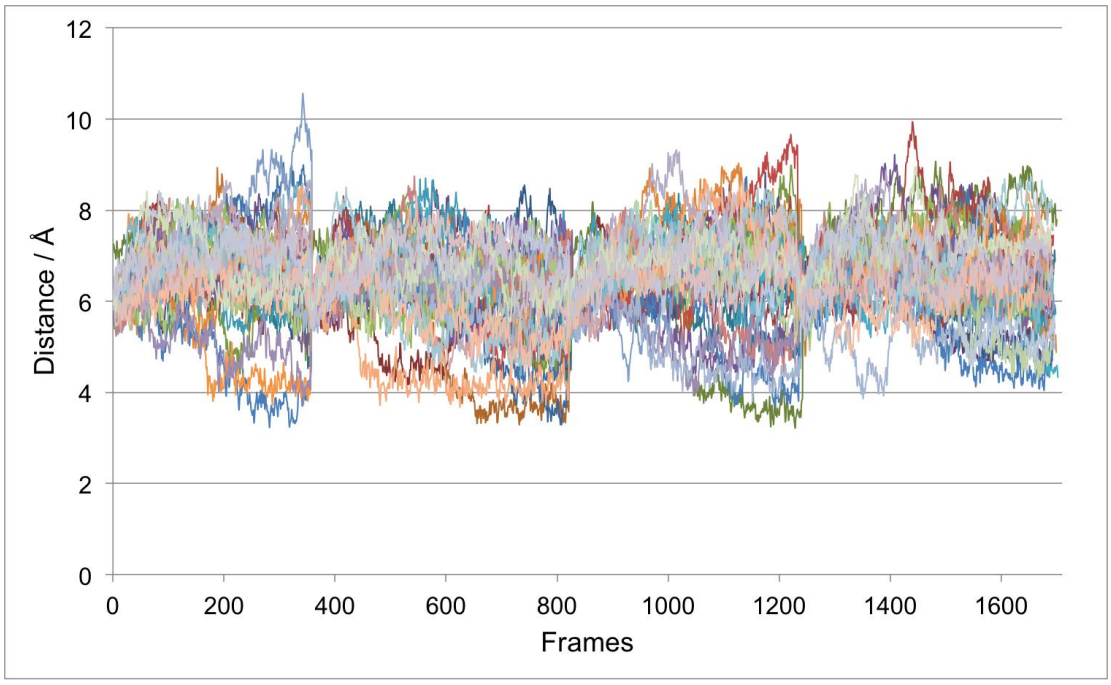



**Figure SI.10**. Distance between anomeric centres along MTD trajectory for compound **21**. Each line represents the trajectory genrated by a single walker (40 walkers in total).

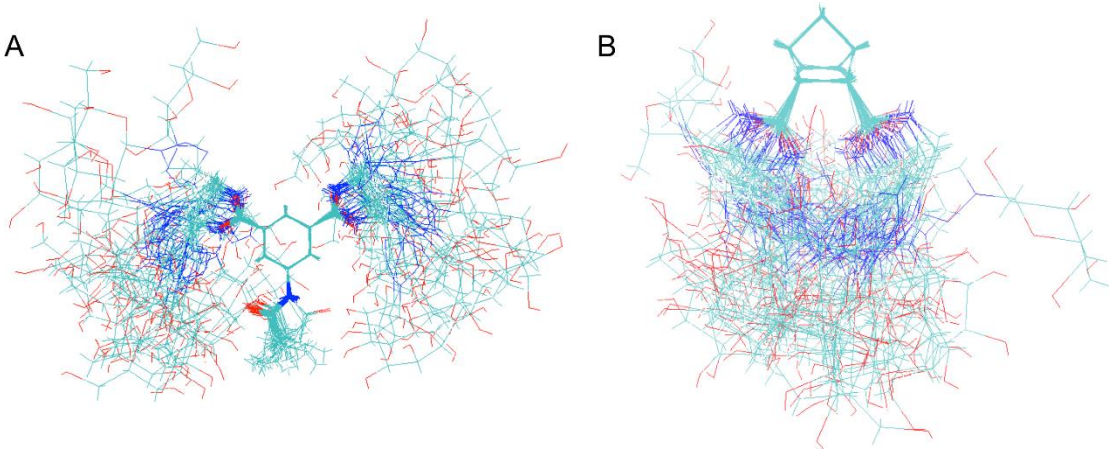



**Figure SI.11**. Representa tive accessible conformational space of (A) compound **1** and (B) compound **21**. Snapshots taken along the 40-walker MTD trajectory, indicating the different orientations of the terminal carbohydrate units linked to the core scaffold through triazolyl linkers. The atoms of the core region of each molecule were used to align all conformations in space in the same way using the VMD RMSD trajectory tool.

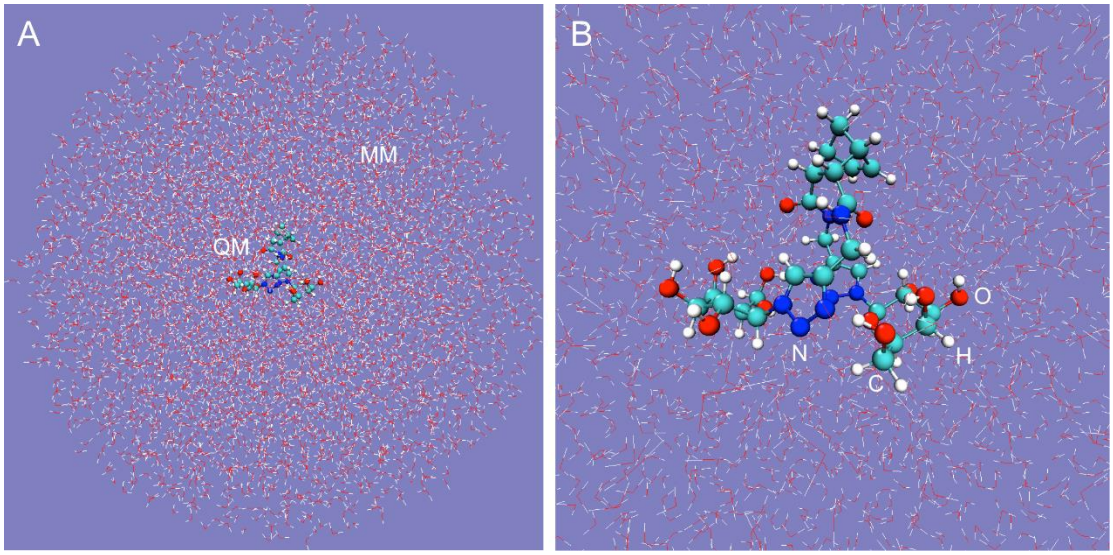



**Figure SI.12**. (A) Partitioning of QM and MM regions within the explicit solvation ("water droplet") model. (B) Close-up view of the QM region.

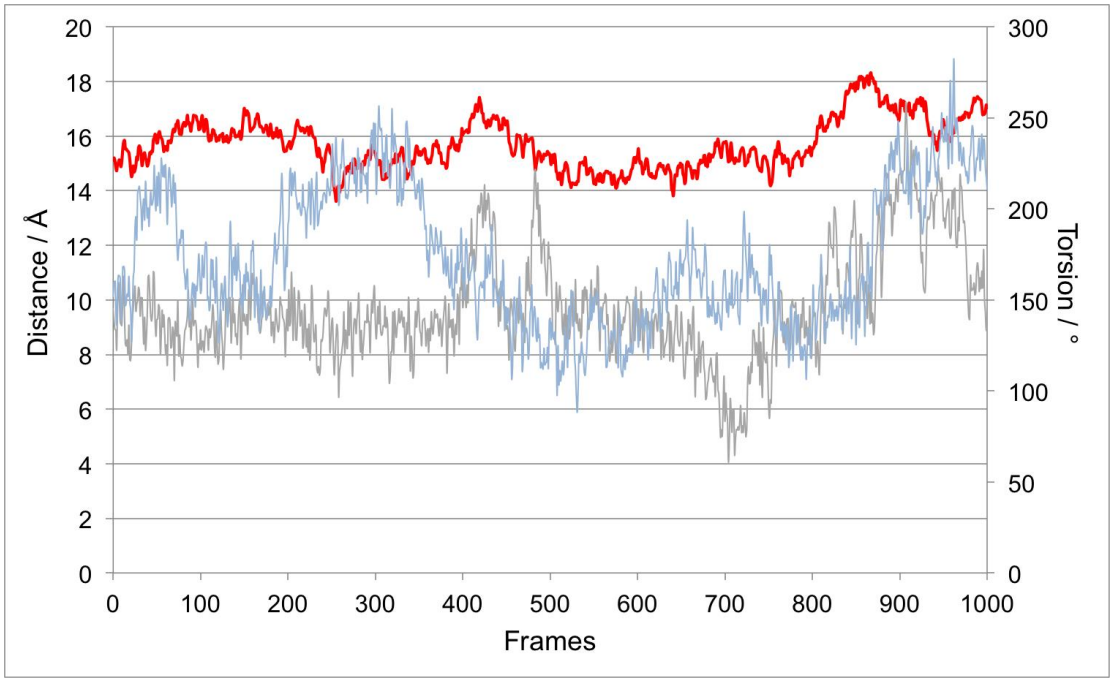



**Figure SI.13**. Distance between anomeric centres (red) and torsional angles $\varphi$ (grey) and $\vartheta$ (blue) along QM/MM trajectory (explicit solvation) for compound **1**.

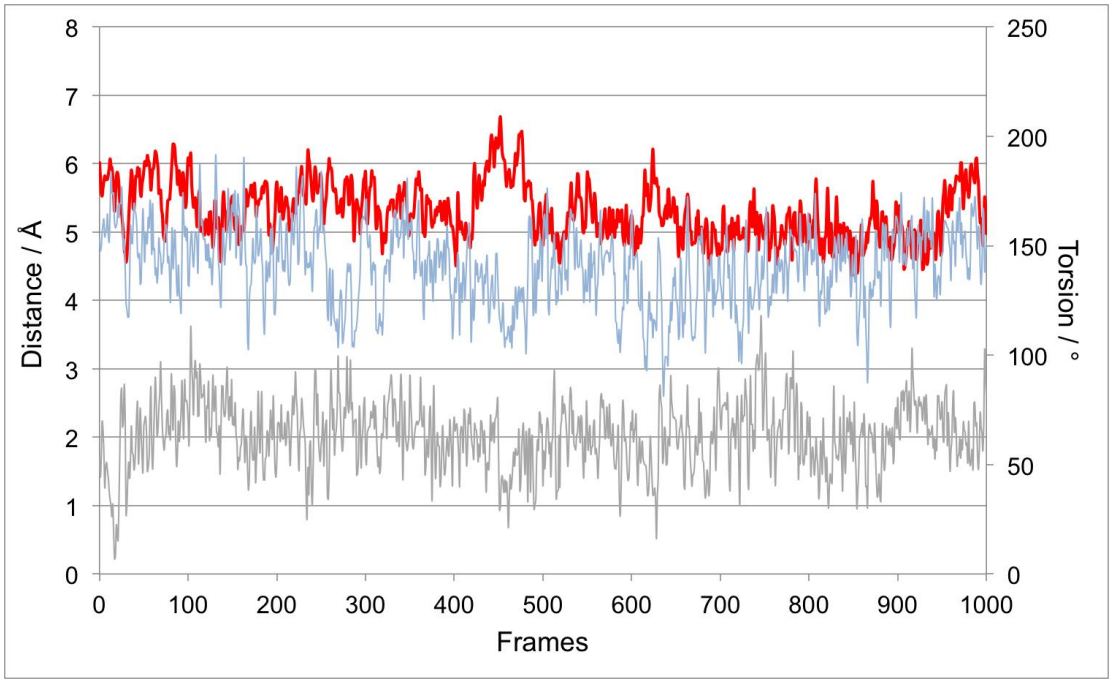



**Figure SI.14**. Distance between anomeric centres (red) and torsional angles $\varphi$ (grey) and $\vartheta$ (blue) along QM/MM trajectory (explicit solvation) for closed-form compound **21**.

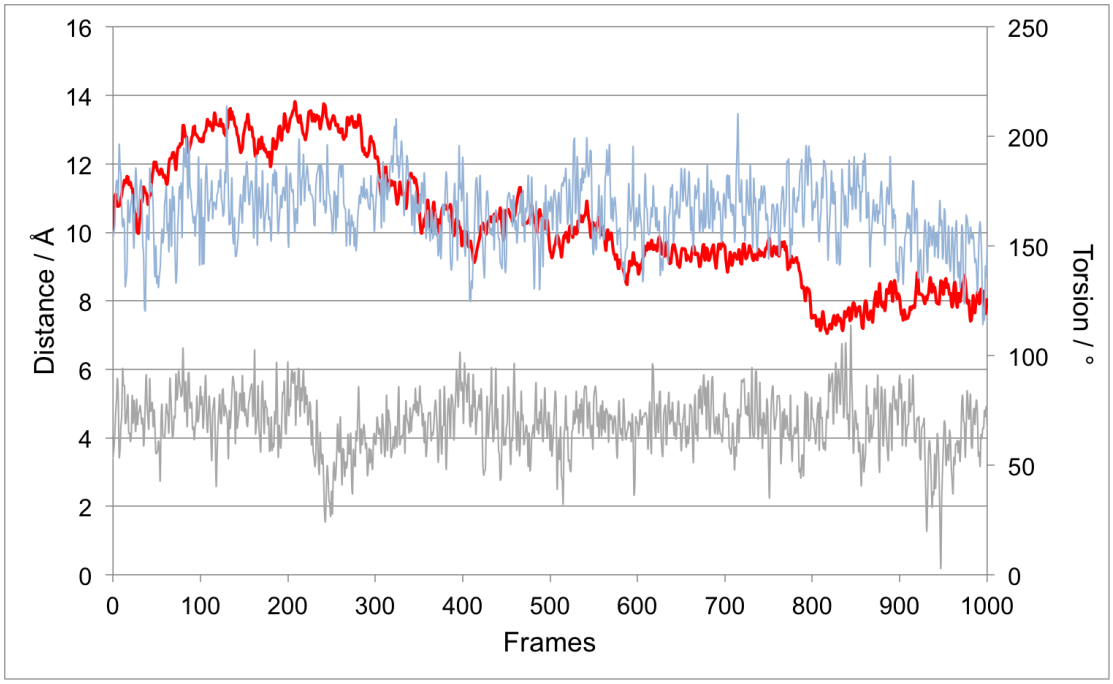



**Figure SI.15**. Distance between anomeric centres (red) and torsional angles $\varphi$ (grey) and $\vartheta$ (blue) along QM/MM trajectory (explicit solvation) for open-form compound **21**.