Synthesis of 3-azido-2,4,6-tri-O-(4-methoxybenzyl)-3-deoxy-galactopyranoside

Towards making 1,3-linked macrocyclic galactose ligands for galectin-3 for thermodynamic and inhibitory studies

2019-06-10

Esbjörn Reistad

Supervisor: Ulf Nilsson

Co-supervisor: Alexander Dahlqvist,

Examinor: Ulf Ellervik

Abstract

Galectin-3 is a protein that is involved in many physiological processes such as metastasis, T-cell regulation etc. which makes it a good drug target for the medicinal chemist. In a paper made by Doak ete al (2015) macrocyclization emerged as an interesting strategy for new drug design as macrocycles remained orally active in chemical space previously thought of as "unsuitable" as per the classical rules defined by Lipinsky. Previous work done in the Nilsson group at Lund university has explored the binding event of monosaccharide galectin-3 inhibitors based on a galactose scaffold. Based on X-ray crystallography obtained from previous work we predicted that a macrocyclic structure would retain much of the same configuration in the binding pocket as the monosaccharides. By comparing the open and cyclised compounds we expect to obtain valuable thermodynamic data of the binding event to aid future drug design. With this goal in mind 3-azido-2,4,6-tri-O-(4-methoxybenzyl)-3-deoxy-galactopyranoside was synthesized and a strategy to obtain the macrocycle has been proposed in this paper.

Introduction

The Galectin family

In the search for new therapeutics the galectin family has arisen as an interesting family of proteins involved in a wide variety of physiological functions for the pharmaceutical chemist to target. The galectin family is a protein family whose members all have carbohydrate recognizing domains (CRD:s) specific to β-galactosides but with different structures. The galectin members present in humans can be subdivided into three groups:

- The dimer or prototypical group containing galectins-1,2,7,10 and 13. The members of this group have one CRD but may act both as a monomer or dimer in different cases (often concentration dependant).
- The tandem repeat group containing galectins-4,8,9 and 12. The members of this group have two different CRD:s linked together by a peptide linker.
- The chimera group with galectin-3 as its sole member. Galectin-3 distinguishes itself from the other galectins by having a unique N-terminal extension. Galectin-3 can oligomerize into oligomers when binding multivalent glycans (Leffler et al, 2004, Yang et al, 2008).

Galectin-3

As previously stated, galectin-3 is the sole chimeric galectin in humans and it is almost ubiquitously present in humans both intra- and extracellularly as well as on cell surfaces. The exact mechanism for the migration over the cell membrane is not known but is thought to be non-classical meaning it does not go via the Golgi-apparatus (Leffler et al, 2004). Being expressed in such different biological zones it is no wonder that galectin-3 is thought to be involved in many different processes such as nuclear splicing of pre-mRNA, T-cell receptor regulation, apoptosis (both pro and anti), angiogenesis and cell adhesion in the extracellular matrix. (Johannes et al, 2018, Liu et al, 2002).

At an organismal level it is difficult to link galectin-3 to specific physiological phenomena as there is substantial redundancy in its functions as demonstrated in knock-out experiments on mice (Leffler et al, 2004). However, there are clear links between the expression and regulation of galectin-3 and physiological conditions such as inflammation, cancer growth and subsequent metathesis which makes galectin-3 an interesting drug target (Liu et al, 2002, Johannes et al, 2018).

Pharmaceutical background

In the process of developing new pharmaceuticals both the API (active pharmaceutical ingredient) *and* formulation must be considered. Rather intuitively, drugs that are orally available offer many advantages over drugs that have to be administered through injection (so called parental drugs): they are easy to administer with less discomfort for the patient which increases the patients likelihood to follow a treatment plan; the injection itself comes with a number of risks such as infection etc. For an API to be orally active it has to exist within a certain chemical/polarity space; it has to be polar enough to dissolve in the gastro-intestinal tract (GIT) but not so polar that it cannot pass through the cell membrane (unless actively transported across). In lieu of this background an empirical set of guidelines were developed named "Lipinski's rule of five" which states that an orally active API should:

- Have a molecular weight of less than five hundred Da.
- Should not have more than five groups able to participate as donors in hydrogen bonding.
- Should not have more than ten groups able to participate as hydrogen bond acceptors.
- Not have a log P value of more than +5, where P is defined as the partition coefficient of the API between a water- and 1-octanol phase at equilibrium.

It is sometimes useful to supplement this set of rules with information about the number of rotatable bonds as it turns out that, generally, the number of flexible bonds is negatively correlated with oral bioavailability. While useful, Lipinski's rule of five is by no mean a rule but rather a guideline. However, the space set by these restrictions is rather narrow and so efforts have been made to widen the scope for what may be considered "orally available drugs" (Doak et al, 2016, Patrick, 2017).

Trends in macrocycles

In a recent analysis of 475 drugs in different clinical stages trends for drug molecules extending outside the chemical space, as defined by Lipinski's rules, were explored. What the researchers noticed was a higher ratio of macrocyclic drugs to non-macrocyclic drugs in this extended chemical space. Furthermore, many of the drugs in this extended space were still orally available (Doak et al, 2016).

The study provided evidence that generally macrocyclic and non-macrocyclic drugs tended to take on different conformational shapes with non-macrocycles taking on more rod-like conformations while macrocycles took on either disc-like or spherical conformations. The shape of the drug molecule was linked to the shape/structure of its target where the disc or spherical shape of the macrocycles seemed to be preferred for flat or grooved targets. Perhaps surprisingly, the authors conclude that the main driving force behind the accumulation of macrocycles in the extended chemical space is not due to a rigid structure but instead due to a match of geometry. Rigidification due to macrocyclization does however offer other

advantages such as improved pharmacokinetic, increased oral bioavailability and sometimes better affinity/selectivity (Doak et al, 2016). In the interest of our project; lectin binding sites are generally shallow grooves (Zetterberg et al, 2018).

Aim and previous work

Previous work in our research group (to be published) has been done on naphtoylamide extensions (figure 1) on the anomeric position of galactose as well as on sulfonphenyltriazole extensions (figure 1) on the three-position of galactose – both yielding test molecules with good affinity and selectivity for galectin-3. Furthermore, Verteramo et al (2019) recently published a comprehensive analysis of the binding thermodynamics for the binding event of galectin-3, offering a set of methods which supplement one another to produce a detailed picture of the binding event. In lieu of their work and the interesting properties of macrocycles proposed by Doak et al (2016); we wanted to examine whether macrocyclization could provide us with a new approach to galectin-3 inhibitors. The project aims to evaluate the thermodynamic properties of the binding event for both open and cyclised compounds; the contribution of entropy and of enthalpy of the binding event using isothermal calorimetry (ITC) and NMR, as well as binding affinity and selectivity over other galectins using competitive fluorescent polarization assay. By dissecting the different contributions of entropy and enthalpy to the overall binding event it is possible to provide a clearer rationale for structure-based ligand design.

Macrocycle design

Previous work done by the Nilsson group has determined the crystal structure for compounds **1a** and **1b** (figure 2) binding to the CRD of galectin-3. The plan was to combine the two non-galactose moieties on the same galactose scaffold to make compound **ER3-13** which could be cyclised into compound **ER3-15**. The crystal structures of compounds **1a** and **1b** were used as input into computational modulations using Schrödinger software to test the feasibility of such a macrocycle as well as to estimate the optimal length of the carbon link between the sulphonyl phenyl and naphtoyl moieties (details of the simulation of the snapshot are provided in the supporting information). The model used estimated that linkers of 3,4 or 5 carbons in length would not displace neither the sulphonyl nor the naphtoyl moieties substantially compared to the structures obtained for compounds **1a** and **1b**. An even number of carbons in the link is preferable as it allows for symmetrical synthesis of the pre-linked allylic groups so for initial testing a four-carbon link was chosen for further studies.

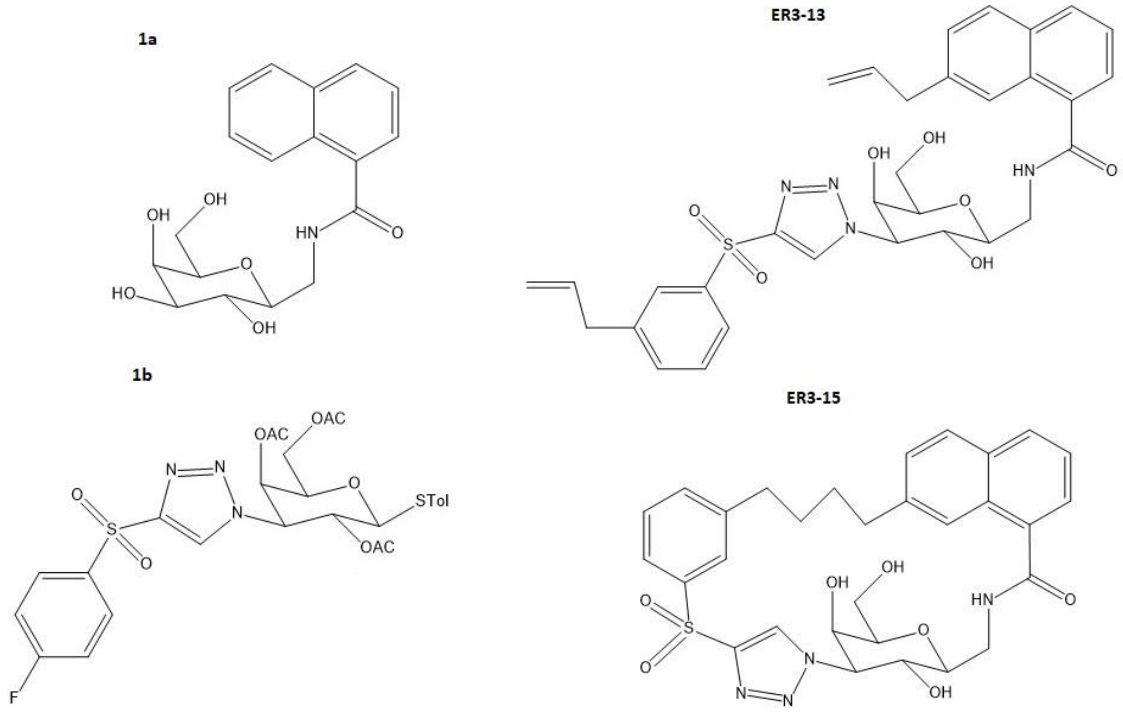


Figure 1. Structure of 1C-(1-naphthamidomethyl)-β-D-galactopyranoside (**1a**); 4-metylphenyl 2,4,6-tri-*O*-acetyl-3-deoxy-3-(4-(4-fluorophenylsulfonyl)-1*H*-triazol-1-yl)-1-thio-β-D-galactopyranoside (**1b**); 3-deoxy-1C-((7-allyl-1-naphtamido)-metyl)-3-(4-(3-allylphenylsulfonyl)-1*H*-triazol-1-yl)-β-D-galactopyranoside (**ER3-13**) and the macrocyclic compound (**ER3-15**).

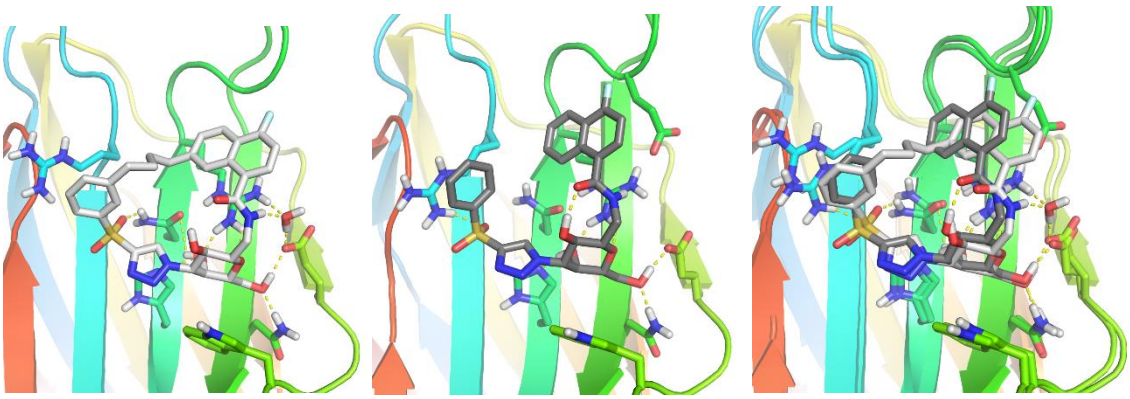
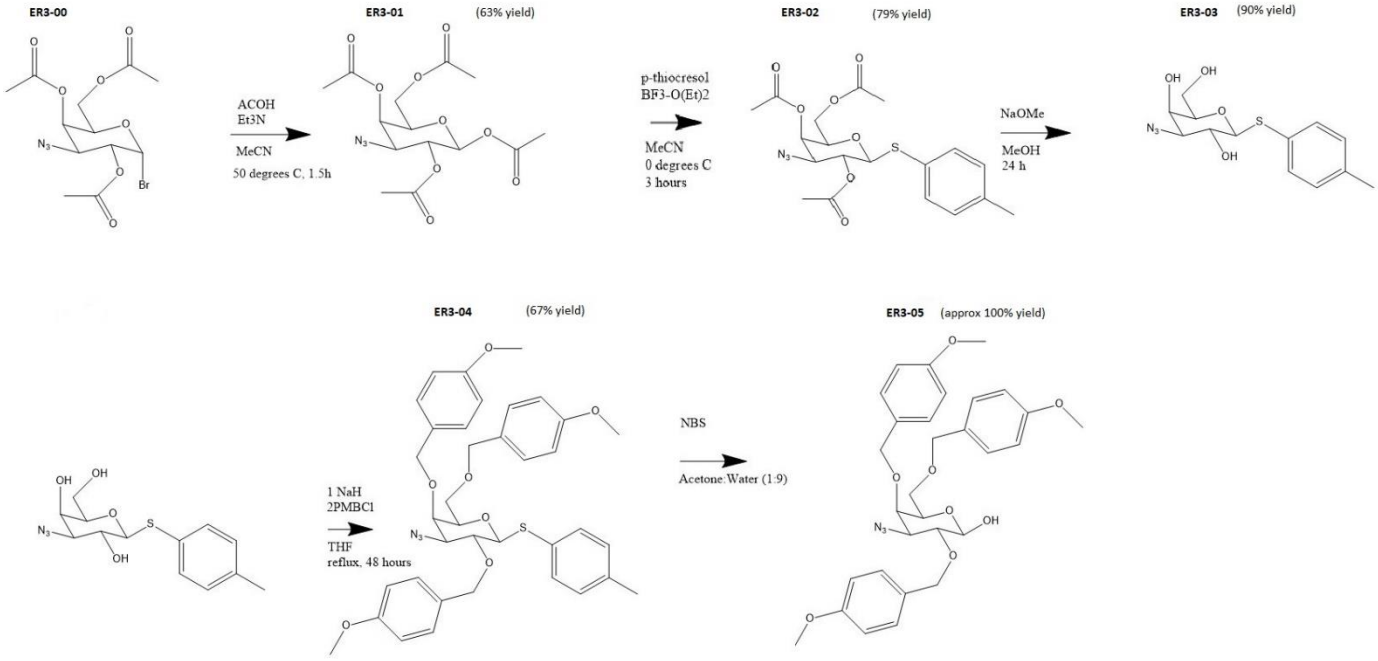


Figure 2. On the left is the simulated fit for compound **ER3-15** in the galectin pocket. The middle figure shows the fit of compound **ER3-14**. The right figure shows both simulated fits superimposed.

Modulation of compounds **ER3-13** and **ER3-15** showed that both compounds fit similarly to the tested compounds **1a** and **1b** in the galectin-3 CRD pocket. Figure 2 shows the simulated fits for compounds **ER3-13** and **ER3-15** in the galectin pocket which implies that the open and cyclized structures have very similar fits in the galectin-3 CRD pocket.

Results and discussion

The synthesis of **ER3-05** started with purifying **ER3-00** (3-azido-1-bromo-2,4,6-tri-O-acetyl-α-galactopyranoside) crude on flash column. The bromide in the α-position was substituted with an acetyl group using acetic acid and triethylamine which produced **ER3-01** (62.8% yield). The β-acetyl group was substituted for a tolylthio group by reacting **ER3-01** with p-thiocresol in the presence of borontriflouride diethyletherate to give **ER3-02** (79% yield). **ER3-02** was then deprotected with sodium methoxide forming **ER3-03**, (89% yield). **ER3-03** was then first treated sodium hydride to deprotonate the hydroxyl groups after which paramethoxybenzylchloride was added as the new protecting group producing **ER3-04**, (73% yield). Finally, the tolylthiol group was removed with NBS in a water/acetone solution giving **ER3-05** (approximately 99% yield, see scheme 1).



Scheme 1. Synthesis of **ER3-01** (1,2,4,6-Tetra-*O*-acetyl-3-azido-3-deoxy-β-D-galactopyranoside), **ER3-02** (4-Methylphenyl 2,4,6-tri-*O*-acetyl-3-azido-3-deoxy-1-thio-β-D-galactopyranoside), **ER3-03** (4-Methylphenyl 3-azido-3-deoxy-1-thio-β-D-galactopyranoside), **ER3-04** (4-Methylphenyl 3-azido-3-deoxy-2,4,6-tri-*O*-(4-methoxybenzyl)-1-thio-β-D-galactopyranoside) and **ER3-05** (3-Azido-3-deoxy-2,4,6-tri-*O*-(4-methoxybenzyl)-β-D-galactopyranoside).

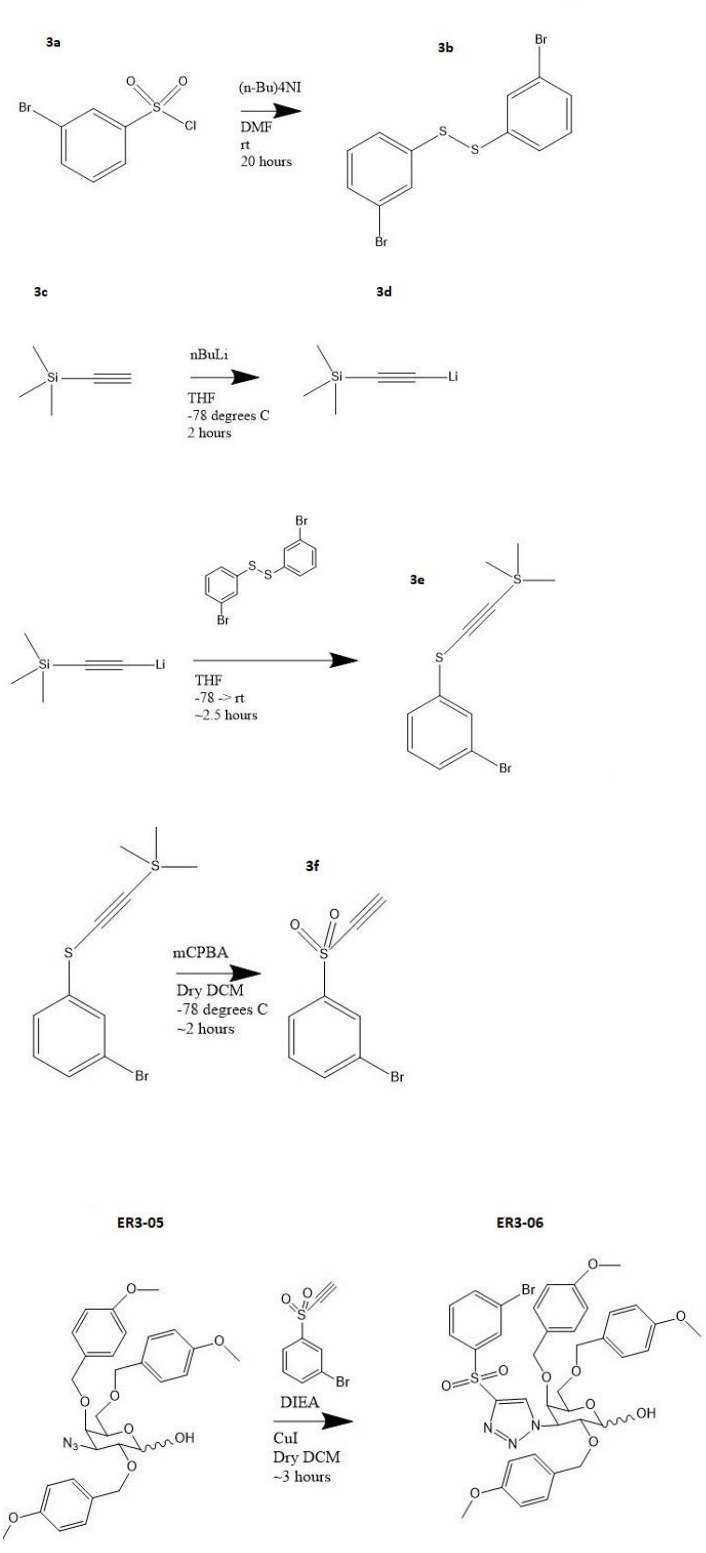
Proposed future work

It was difficult to analyse **ER3-05** partially due to the presence of the two different anomers and partially due to impurities. Better NMR data was collected on small scale, which showed that the correct compound had been synthesized. One easy way to check whether the correct compound has been made is to continue the synthesis on small scale and see if the correct product forms in the next step. The presence of the two anomers is not problematic as the oxidation by Dess-Martin periodane does not discriminate between the two anomers (see scheme 4).

We propose the following synthetic path from **ER3-05** in order to make the final compounds **ER3-14** and **ER3-15**. The proposed strategy is general and exact conditions must be experimentally determined.

Synthesis of 1-bromo-3-(ethynylsulfonyl)benzene and its cycloaddition with ER3-05.

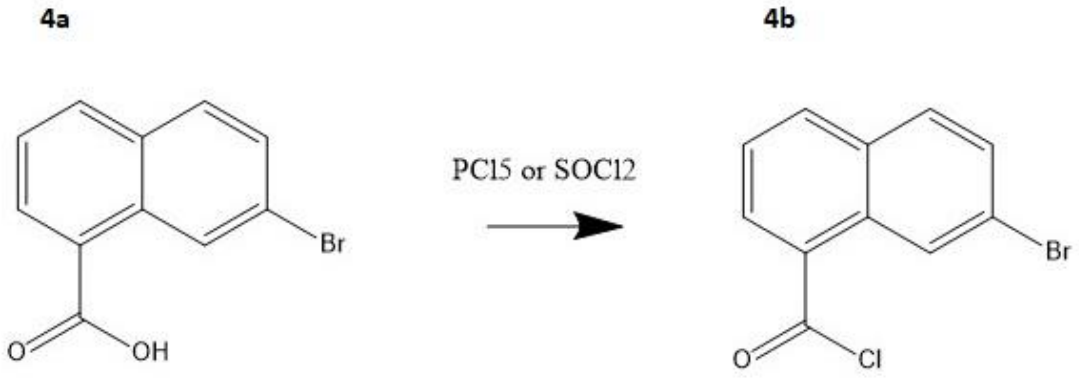
Mukul Mahanti (to be published) developed a pathway to making ethynylsulfonylbenzenes with good yield and we propose to follow the same pathway in order to make 1-bromo-3-(ethynylsulfonyl) benzene. The brominated disulphide required is however not commercially available but can be made in house following the findings of Zheng et al (2016, scheme 2).



Scheme 2. Proposed synthesis of 1-bromo-3-(ethynylsulfonyl)benzene and its cycloaddition to **ER3-05**. The formation of the disulphide was proposed by Zheng et al (2016) and the subsequent steps were developed by Mukul Mahanti (unpublished).

Making the naphthoic acid chloride group

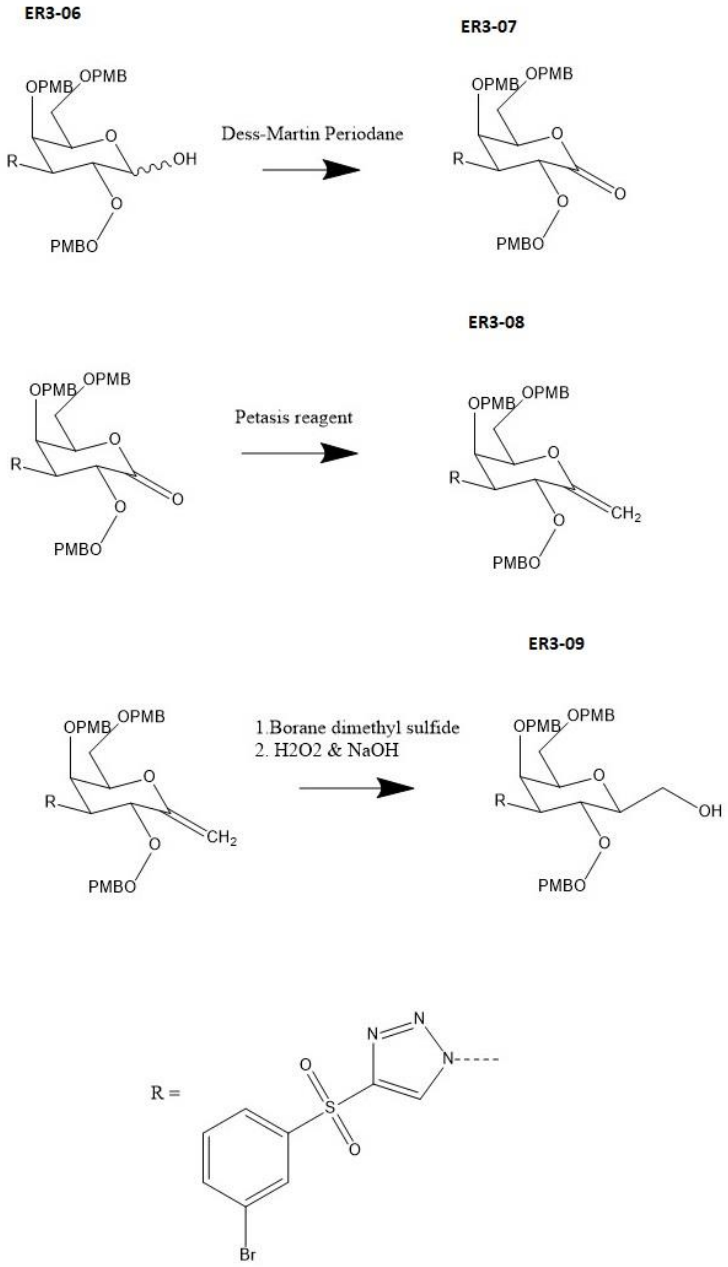
We propose to convert commercially available 7-bromo-1-naphthoic acid into 7-bromo-1-naphthoyl chloride using either $SOCl_2$ or PCl_5, (scheme 3).



Scheme 3. Possible strategy for formation of 7-bromo-1-naphthoic acid chloride from the corresponding carboxylic acid.

Preparing the galactose scaffold for amide coupling

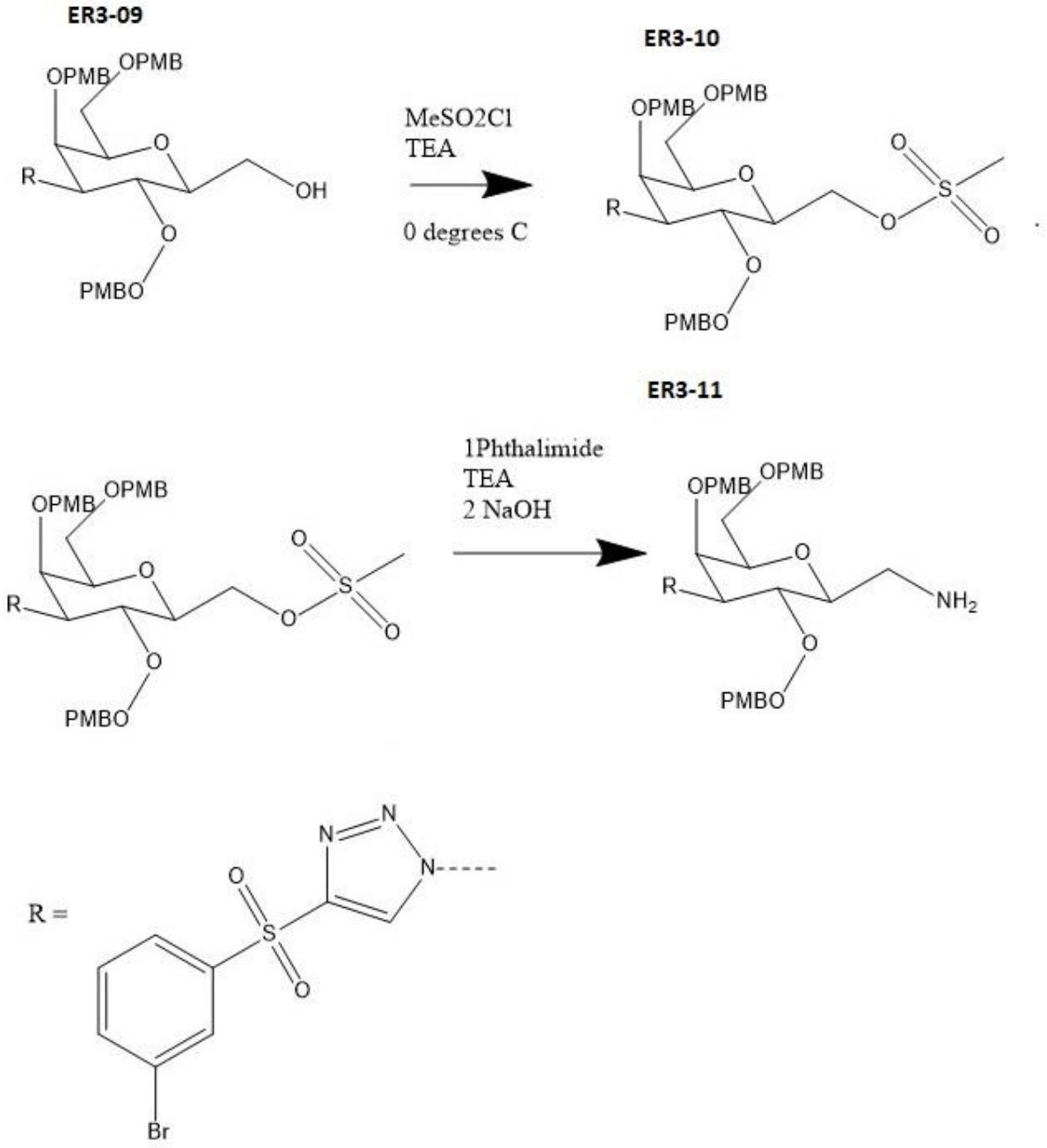
Following a hydroboration method developed (scheme 4) by Alexander Dahlqvist (Dahlqvist et al, 2019) where excellent diastereoselectivity was achieved, it is hoped that the same results can be achieved with para-methoxybenzyl protecting groups instead of benzyl groups used in the original experiment.



Scheme 4. Proposed synthesis of 1-hydroxy-3,5,7-*O*-(4-methoxybenzyl)-3-(4-(3-bromophenylsulfonyl)-1*H*-triazol-1-yl)-1,3-dideoxy-β-D-galactoheptulose according to the hydroboration method developed by Alexander Dahlqvist (Dahlqvist et al, 2019).

Installing the amine

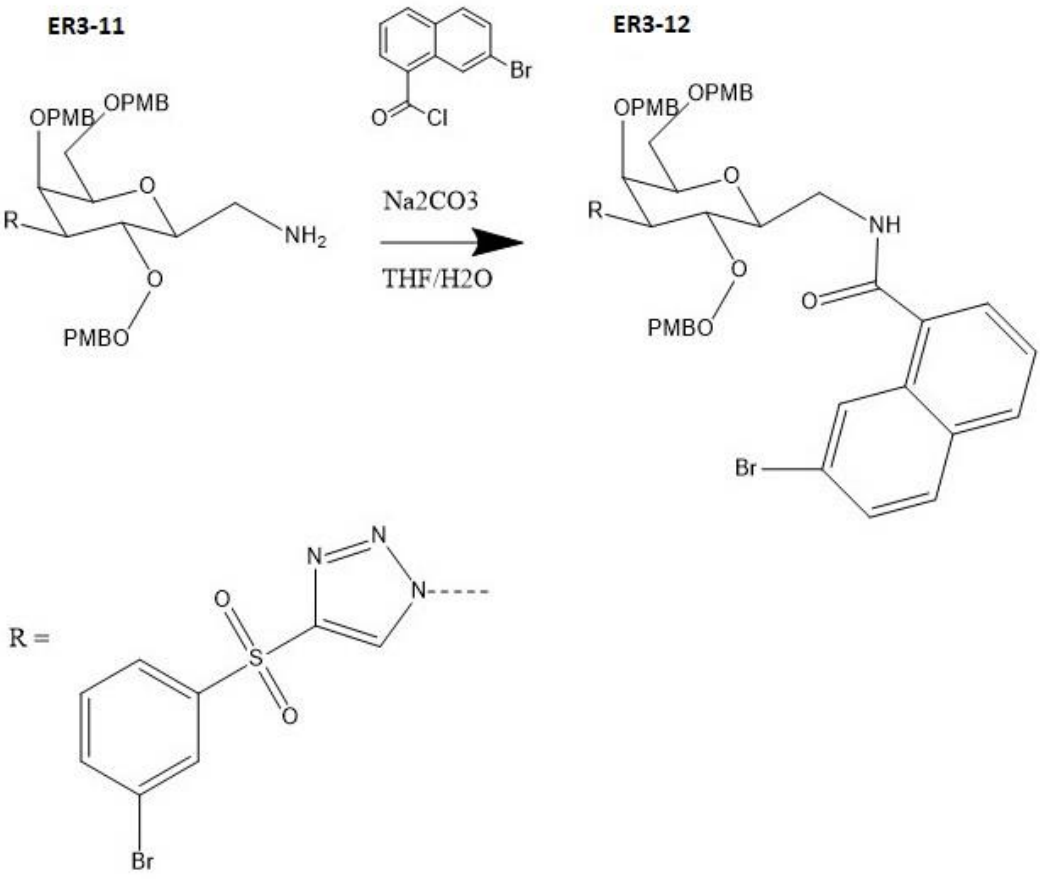
We propose to exchange the exo-hydroxy group of the **ER3-9** with an amine by first activating the hydroxy group by mesylation and then substituting it with an amine using Gabriel synthesis and hydrolysis, (scheme 5).



Scheme 5. Proposed strategy for installation of an exo-amine by mesylation and Gabriel synthesis with triethyl amine (TEA) as a hindered base.

Synthesis of the naphthoyl amide

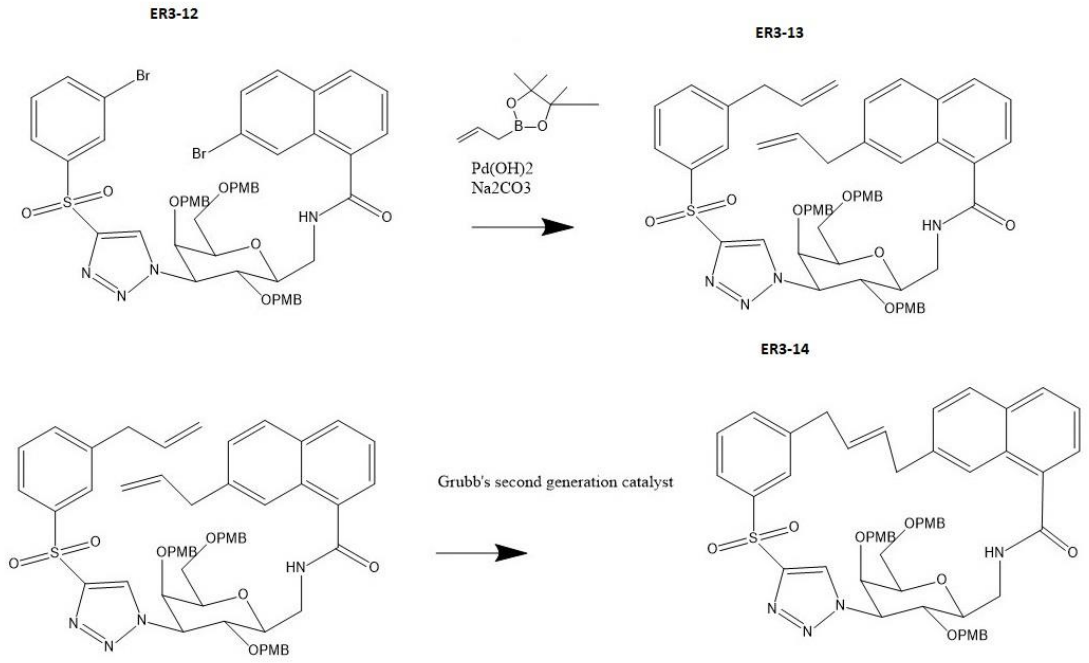
Following another manuscript produced by Alexander Dahlqvist (to be published) we propose to join the naphthoyl moiety by reacting the naphthoic acid chloride with **ER3-11** under mildly basic conditions (scheme 6).



Scheme 6. Proposed strategy for joining the naphthoyl moiety onto the galactoheptulose scaffold.

Synthesis of the open diene ER3-13 and ring-closing metathesis into the macrocycle ER3-15.

We propose to perform a Suzuki-Miyaura coupling of **ER3-12** with allylboronic acid pinacol ester, substituting the bromides on both the phenylsulfonyl- and naphthoyl moieties with allyl groups yielding the open compound **ER3-13**. By performing a ring-closing metathesis (RCM) on **ER3-13** with Grubb's second generation ruthenium catalyst we expect to cyclise the compound into **ER3-14** (Scheme 7).

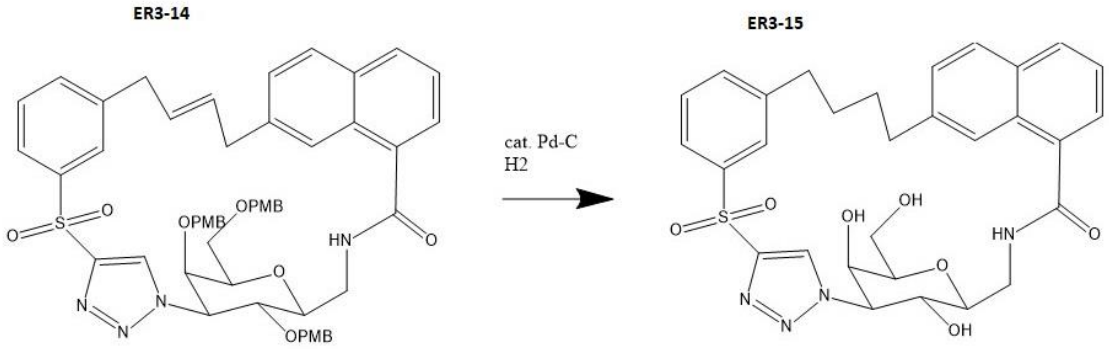


Scheme 7. Proposed reactions for of the allyl installation and ring-closing metathesis reactions.

Deprotection

The PMB groups can be deprotected using many different procedures including removal with 2,3-dichloro-5,6-dicyano-p-benzoquinone (DDQ); palladium catalysed hydrogenation; removal with oxalyl chloride or other methods (Ilangovan et al, 2015). For this project it would be preferable to deprotect the hydroxyl groups after cyclisation but it is possible that the structure will be too sterically congested for metathesis (while protected) due to the bulkiness of the protecting groups. Should cyclisation be possible with the protecting groups present then hydrogenation with palladium would be preferable as it would also hydrogenate the double bond on the linker which would be the prime testing target for further studies. However, should deprotection be necessary before cyclisation then the palladium catalysed deprotection is unsuitable as it would hydrogenate the allyl-groups thus oxalyl chloride would be preferable as Ilangovan et al (2015) were able to selectively deprotect the PMB groups in presence of double bonds. This approach would come with further risk as the catalyst might not tolerate the free hydroxyl functional groups. Using the more traditional DDQ deprotection

would risk oxidizing the double bond on the linker which, while it might prove to be interesting for further studies should the macrocycle be preferable to the open compound, would render the compound less suitable for comparison with the open compound (scheme 8).



Scheme 8. Schematic picture of the PMB deprotection using palladium catalysis.

Conclusions

In this project **ER3-05** (3-azido-3-deoxy-2,4,6-tri-O-(4-methoxybenzyl)-β-D-galactopyranoside) was successfully synthesized with good yields along the entire synthetic pathway thus far. However, purification of **ER3-04** leaves some to be desired as it was necessary to purify it by flash column twice in each attempt. Finding a better solvent system or changing the silica properties might be worth looking into to improve the yields further.

The open and macrocyclic compounds proposed in this study remains interesting for thermodynamic studies. Further understanding of the binding interaction and dissecting the enthalpy/entropy contributions of the binding event for galectin-3 would greatly enhance the prospect of a rational target-based design for future projects. Furthermore, it would be very beneficial to test the viability of the macrocyclic and open compounds in lieu of the work and propositions of Doak et al (2016) and see whether the macrocyclic compound offers advantages in terms of cell permeability and oral availability. While this project was not able to produce the final compounds in the set timeframe there is currently nothing that suggests that such a task should be impossible.

Acknowledgement

I would like to thank my supervisor Professor Ulf Nilsson for all your help and for giving me the opportunity to work on this project. A special thanks to my co-supervisor Dr Alexander Dahlqvist for your guidance and insight and helping me develop as a synthetic organic chemist. I would like to thank the entire Nilsson and Ellervik groups as well as Sofia Essén for their support and interest in this project. All the interesting talks, discussions, wonder and

woe experienced together has been great both in and out of lab gear. And finally I'd like to thank Maria Levin and Katarina Fredriksson for keeping the building together.

Experimental section

Instruments and specifications

All flash chromatography was performed with silica gel of the following specifications: a pore size of 60Å; a mesh particle size of 230-400Å and a particle size of 43-60μm.

The TLC:s were performed on "TLC Silica gel 60" on aluminium sheets supplied by Merck. TLC plates were visualized using a charring solution (7% sulfuric acid in ethanol), UV light (254 nm) or both.

NMR was perfomed on a Bruker advance II spectrometer (^{1}H NMR 400 MHz, ^{13}C 100 MHz) and the chemical shifts are given in units of ppm.

All chemicals were purchased from Sigma-Aldrich and used without further purification except for the starting material which was purchased from Aptuit chemicals.

Molecular dynamics simulations

Molecular dynamics simulations were performed with the OPLS3 force field in Desmond (Schrödinger Release 2018-4: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2017. Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY, 2017) using default settings except for the length of the simulation and the use of light harmonic constraints (1 kcal/mol/Å) on all stranded backbone atoms and on the galactose O4 atom. Compound 2b was positioned to replace lactose in the binding site of galectin-3 (pdb id 1KJL) with the galactose ring of 2b in an orientation identical to that of galactose ring in 1KJL. This complex was subjected to a 200 ns molecular dynamics simulation. Molecular images were generated using PyMOL v1.7 (Schrodinger LLC).

Attempted synthesis of 1-bromo-3-(ethynylsulfonyl)benzene

Synthesis of 1-bromo-3-(ethynylsulfonyl)benzene was attempted in two separate ways, neither of which succeeded. The first method (method 1) attempted was a Friedel-Craft alkynylation-like attempt using Aluminium(III) chloride and TMSacetylene and the second method was an attempt at a organo-lithium addition of lithium acetylene to the sulfonylchloride. Several attempts were made with each method.

Method 1

In the first flask, 0.0929 g of Aluminium(III) chloride (0.70 mmol) was dissolved in 2 ml dry DCM after which 0.06 ml (0.4 mmol) of 3-bromo benzoyl chloride was added under N2 and left to stir for 1 hour. In a separate flask 0.0978 g Bis(trimethylsilyl)acetylene (0.57 mmol) was dissolved in 1 ml dry DCM and cooled to zero degrees centigrade under N2. The warm liquid was then slowly transferred to the cold solution with an additional 2ml dry DCM to clean the first flask. The mixture was kept on ice for 30 minutes after which it was allowed to warm up to room temperature for an additional 48 hours. After the 48 hours no reaction had occurred (controlled by TLC) so an additional 3 ml dry DCM was added together with

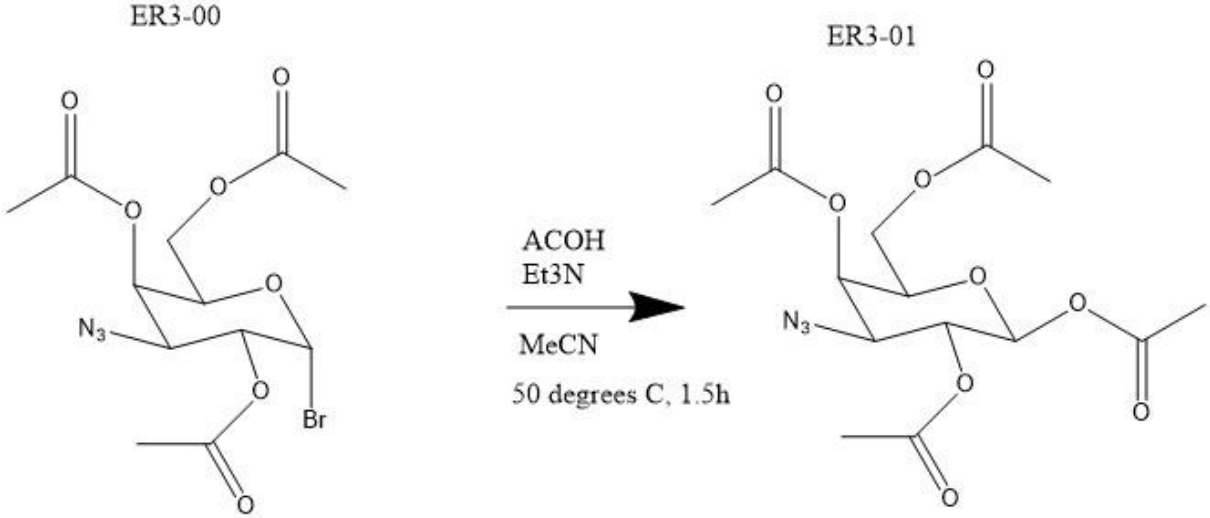
0.0686g Bis(trimethylsilyl)acetylene (0.40 mmol) and 0.0549 g Aluminium(III) chloride; the mixture was heated to 30 degrees centigrade and put under reflux. After 72 hours only some minor amount product had formed and the experiment was terminated.

Method 2

0.221 g Cerium(III) chloride heptahydrate (0.58 mmol) was dried under vacuum for 1 hour at 120 degrees centigrade and 1.5 hour at 140 degrees centigrade. The dried cerium(III) chloride was cooled on ice, dissolved in 2.1 ml dry THF and warmed to room temperature. In a separate flask 0.9 bis(trimethylsilyl)acetylene was cooled to -78 degrees centigrade after which 0.28 ml n-Butyl lithium (0.7 mmol, 2,5 M in hexanes) was added slowly. After one hour the cerium solution was cooled to -78 degrees centigrade and the second solution was added slowly with an additional 1 ml THF used to clean the flask. 0.09 ml of 3-bromo benzoyl chloride (0.58 mmol) was then dissolved in a separate flask at -78 degrees centigrade and after 30 min of stirring the sulfonyl chloride solution was added dropwise to the cerium solution with an additional 1 ml of THF to clean the flask. With this addition came a colour change from a slightly yellow solution to a reddish one and then to a purple one. The reaction mixture was left in a dry ice bath over the night and the solution was -9 degrees centigrade the day after. TLC showed a minor amount of product that was also unreasonably polar. The reaction was neutralized with HCl. The aqueous phase was washed three times with 15 ml DCM; the organic phases were pooled and washed twice with 15 ml water; twice with 15 ml brine; dried with anhydrous sodium sulphate; filtered with dcm and reduced under vacuum. NMR and LC-MS on the crude were inconclusive and suggested that the compound was impure so attempts were made to purify the compound. No suitable eluent system was found as the compound stuck on silica and only moved slightly in a 77% DCM, 20% MeOH and 3% water eluent. The compound was then crystalized and analysed with NMR and LC-MS, neither of which could confirm the compound. The water phases from the experiment were also pooled but no presence of the product could be found in the combined phases.

Detailed synthesis.

ER3-01 (1,2,4,6-Tetra-*O*-acetyl-3-azido-3-deoxy-β-D-galactopyranoside)



Scheme 9. Schematics for the synthesis of compound **ER3-01**.

Starting with purifying **ER3-0** (3-azido-1-bromo-2,4,6-tri-O-acetyl-α-galactopyranoside) on a flash column using 1:1 ethyl acetate:heptane as the eluent system, the resulting oil weighing 5.731g (14.6 mmol). The oil was dissolved in acetonitrile (73 ml) in a nitrogen atmosphere to which acetic acid (30.5 ml) was added. When the oil was fully dissolved triethylamine (15.5 ml) was added slowly and the solution was heated to fifty degrees centigrade for 1.5 hours. Upon completion the solvent was evaporated and the remaining oil was first dissolved in ethyl acetate (50 ml) which was then washed with saturated sodium bicarbonate solution (50 ml); saturated ammonium chloride solution (50 ml), brine (50 ml); dried with anhydrous sodium sulphate, filtered and evaporated. The crude was purified with flash chromatography using an ethyl acetate:heptane gradient of ratios of 1:2; 1:1 and 2:1 which gave **ER3-01 (**3.42 g, 63% yield). The product was a white powder (scheme 9).

NMR for ER3-01 (1,2,4,6-Tetra-*O*-acetyl-3-azido-3-deoxy-β-D-galactopyranoside)

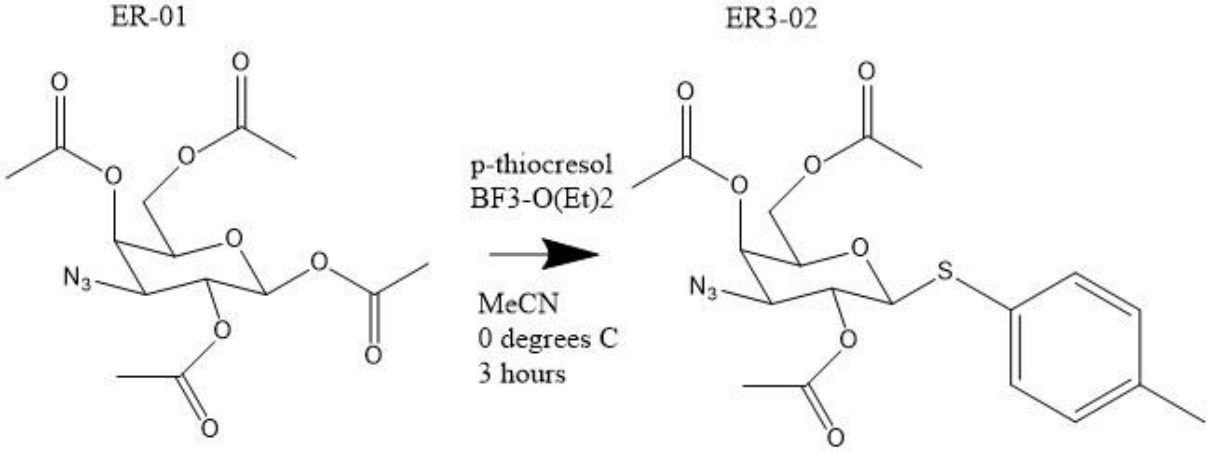
In deuterated chloroform

^{1}H: 5.701 (d, J=8.22 Hz, 1H, H-1), 5.46 (dd, J=0.89 Hz, 3.42 Hz, 1H, H-4), 5.28 (dd, J=8.23 Hz, 10,61 Hz, 1H, H-2), 4.178-4.136 (m, 1H, H-5), 4.081-3.995 (m, 2H, H-6), 3.67 (dd, J= 3.40 Hz, 10,60 Hz, 1H, H-3), 2.179 (s, 3H, acetyl protons), 2.124 (s, 3H, acetyl protons), 2.118 (s, 3H, acetyl protons), 2.056 (s, 3H, acetyl protons).

^{13}C: 170.53, 170.01, 169.38, 169.20, 92.35, 77.48, 76.84, 72.84, 68.84, 67.49, 61.80, 61.37, 20.96, 20.808, 20.744.

Structure confirmed by COSY and the proton spectrum is aligned with previous results obtained by Verča Chadimová.

ER3-02 (4-Methylphenyl 2,4,6-tri-*O*-acetyl-3-azido-3-deoxy-1-thio-β-D-galactopyranoside)



Scheme 10. Schematics for the synthesis of compound **ER3-02**.

Compound **ER3-01** (2.953 g, 7.91 mmol) was dissolved in dry acetonitrile (15 ml) and cooled to zero degrees centigrade in a nitrogen atmosphere. p-Thiocresol (1.429 g, 11.1 mmol) was added to the solution and after it had dissolved boron trifluoride diethyl etherate (2.75 ml, 11.1 mmol) was added slowly and the solution acquired a faint brown colour. After thirty minutes the solution had started to solidify so it was partially removed from the ice bath so that the phase was completely liquid. After three hours and fifteen minutes the reaction was

complete and the solution was quenched slowly with saturated sodium bicarbonate solution (36 ml) and the mixture was allowed to warm to room temperature. The aqueous phase was extracted four times with 30 ml DCM after which the organic phases were pooled; washed three times with 30 ml saturated sodium bicarbonate solution; washed three times with 30 ml brine; dried over anhydrous sodium sulphate; filtered through a stinted glass and reduced under vacuum. The crude was purified with flash chromatography using a gradient eluent system of 1:5, 1:3, 1:2, 1:1 and 2:1 ethyl acetate:heptante yielding **ER3-02** (2.74 g, 79% yield) and an unidentified minor side product presumed to be the alpha anomer (scheme 10).

NMR for ER3-02 (4-Methylphenyl 2,4,6-tri-*O*-acetyl-3-azido-3-deoxy-1-thio-β-D-galactopyranoside)

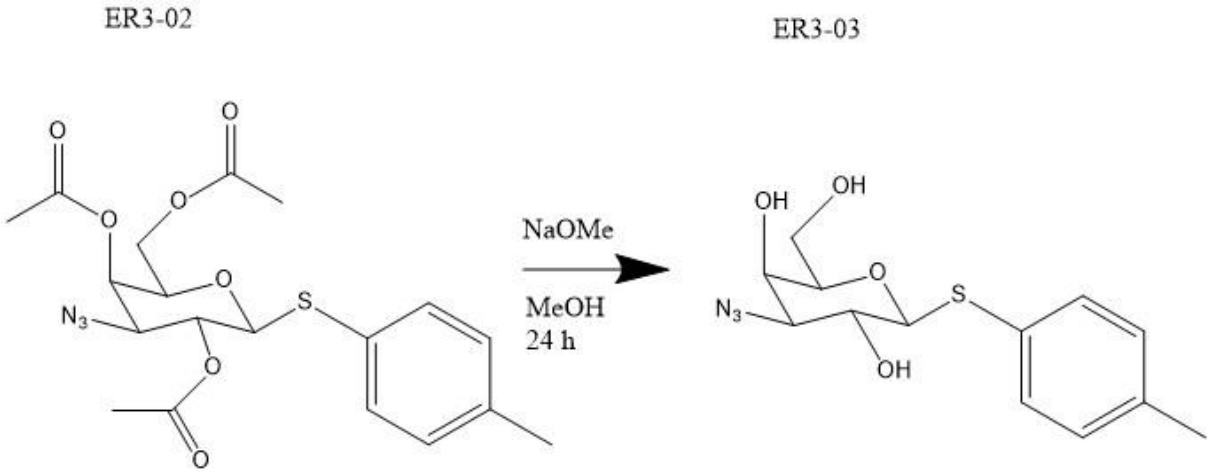
In deuterated chloroform

^{1}H: 7.41 (d, J=8.18 Hz, 2H, meta thiol protons), 7.12 (d, J=7.89 Hz, 2H, ortho thiol protons), 5.43 (dd, J=0.87 Hz, 3.27 Hz, 1H, H-4), 5.19 (T, J=9.98 Hz, 1H, H-2), 4.62 (d, J=9.89 Hz, 1H, H-1), 4.11 (d, J=6.54 Hz, 2H, H-6), 3.86 (td, J=1.00 Hz, 6.46 Hz, 1H, H-5), 3.62 (dd, J=3.33 Hz, 10.09 Hz, 1H, H-3), 2.34 (s, 3H, CH3-stol), 2.178 (s, 3H, acetyl protons), 2.137 (s, 3H, acetyl protons), 2.053 (s, 3H, acetyl protons).

^{13}C: 133.12, 129.67, 87.32, 77.34, 77.02, 76.70, 75.21, 68.40, 67.79, 62.92, 61.81, 21.17, 20.91, 20.70, 20.62.

Assignment was done with COSY.

ER3-03 (4-Methylphenyl 3-azido-3-deoxy-1-thio-β-D-galactopyranoside)



Scheme 11. Schematics for the synthesis of compound **ER3-03**.

Compound **ER3-02** (2.714 g, 6.27 mmol) was put in dry methanol (26 ml) which created a suspension. After ten minutes of stirring sodium methoxide in methanol (26 ml, 26 mmol) was added and the solution acquired a faint greenish colour. During a five hour span an additional sodium methoxide solution (24 ml, 24 mmol) was added after which the reaction

was left overnight. The mixture was quenched with DOWEX until the pH of the solution was between six and seven after which the mixture was filtered with methanol through a stinted glass. The crude was evaporated and purified on a flash column using a gradiented eluent going from 1:1 to 2:1 and finally 3:1 ethyl acetate:heptane yielding **ER3-03** (1.754g, 90% yield). The resulting product was a white substance akin to cotton candy (scheme 11).

NMR for ER3-03 (4-Methylphenyl 3-azido-3-deoxy-1-thio-β-D-galactopyranoside)

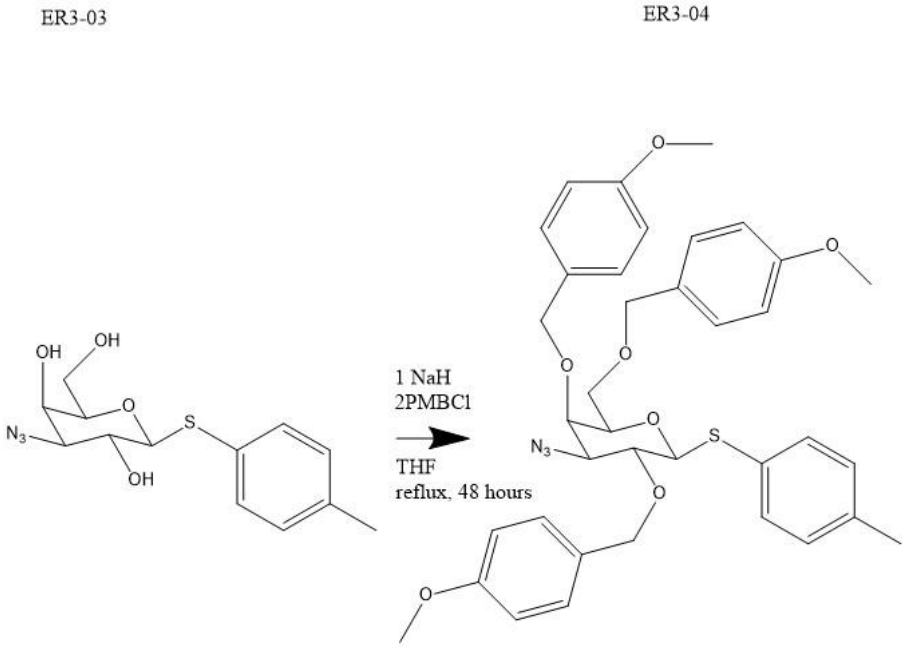
In deuterated methanol

^{1}H: 7.45 (d, J=8.16 Hz, 2H, meta thiol protons), 7.12 (d, J=7.88 Hz, 2H, ortho thiol protons), 4.59 (d, J=9.59 Hz, 1H, H-1), 3.96 (dd, J=0.66, 3.00, 1-H, H-4), 3.78 (t, J=9.71 Hz, 1H, H-2), 3.69 (qd, J= 6.73 Hz, 11.40 Hz, 2H, H-6), 3.562-3.529 (m, 1H, H-5), 3.37 (dd, J=3.09 Hz, 9.87 Hz, 1H, H-3), 2.31 (s, 3H, CH3-stol).

^{13}C: 137.20, 131.65, 130.41, 129.17, 89.83, 79.48, 68.06, 68.04, 67.16, 61.02, 19.68.

Assignment was done by COSY..

ER3-04 (4-Methylphenyl 3-azido-3-deoxy-2,4,6-tri-*O*-(4-methoxybenzyl)-1-thio-β-D-galactopyranoside)



Scheme 12. Schematics for the synthesis of compound **ER3-04**.

Compound **ER3-03** (1.284 g, 4.12 mmol) was dissolved in dry THF (24 ml) after which NaH (0.6412g, 60% in oil, 16.5 mmol) was added in portions and the mixture was left to stir at room temperature for twenty minutes. After the starting material had fully dissolved PMBCl (2.8 ml, 20.6 mmol) was added and the vessel was heated to reflux. After 48 hours the reaction was quenched with saturated NH4Cl solution (60 ml) and diluted with ethyl acetate which invoked a colour change of the mixture from brown to red. The phases were separated and the aqueous phase was extracted three times with ethyl acetate which were pooled and washed twice with water; twice with brine; dried over anhydrous sodium sulphate; filtered through a stinted glass with ethyl acetate and reduced under vacuum. The resulting crude was difficult to purify and required two separate flash columns. The first flash column utilized a

gradient eluent system of 1:7, 1:5, 1:3 and 1:1 ethyl acetate:heptane. The second flash column was made in 2:3 ethyl acetate:heptane which yielded **ER3-04** (1.869 g, 67% yield), (scheme 12).

NMR for ER3-04 (4-Methylphenyl 3-azido-3-deoxy-2,4,6-tri-*O*-(4-methoxybenzyl)-1-thio-β-D-galactopyranoside)

In deuterated chloroform

^{1}H: 7.41 (d, J = 8.15 Hz, 2H, meta thiol protons), 7.38 (d, J = 8.62 Hz, 2H, benzyl), 7.23 (d, J = 8.62 Hz, 2H, benzyl), 7.20 (d, J = 8.57 Hz, 2H, benzyl), 7.00 (d, J = 8.04 Hz, 2H, ortho thiol protons), 4.85 (d, J = 9.60 Hz, 1H, benzylic-a proton), 4.78 (d, J = 10.87 Hz, 1H, benzylic-b proton), 4.62 (d, J = 9.58 Hz, 1H, benzylic-a proton), 4.56 (d, J = 9.53 Hz, 1H, H-1), 4.48 (d, J = 10.88 Hz, 1H, benzylicb proton), 4.37 (q, J = 11.30 Hz, 2H, benzylic-c protons), 3.85 (d, J = 2.85 Hz, 1H, H-4), 3.824-3.795 (m, 14H, H-2; H-5; H-6; 9 methoxy protons), 3.48 (dd, J = 2.89, 9.68, 1H, H-3), 2.30 (s, 3H, CH3-stol).

^{13}C: 159.52, 159.37, 159.25, 137.42, 132.03, 130.39, 130.29, 130.07, 129.83, 129.68, 129.66, 129.59, 113.98, 113.86, 113.84, 113.63, 88.5 4, 77.63, 77.35, 77.03, 76.71, 74.79, 74.56, 74.48, 73.22, 68.02, 67.19, 55.31, 55.29, 55.28, 21.12.

Assignment was made with COSY, HMQC and HRMS performed by Sofia Essén.

ER3-05 (3-Azido-3-deoxy-2,4,6-tri-*O*-(4-methoxybenzyl)-β-D-galactopyranoside)

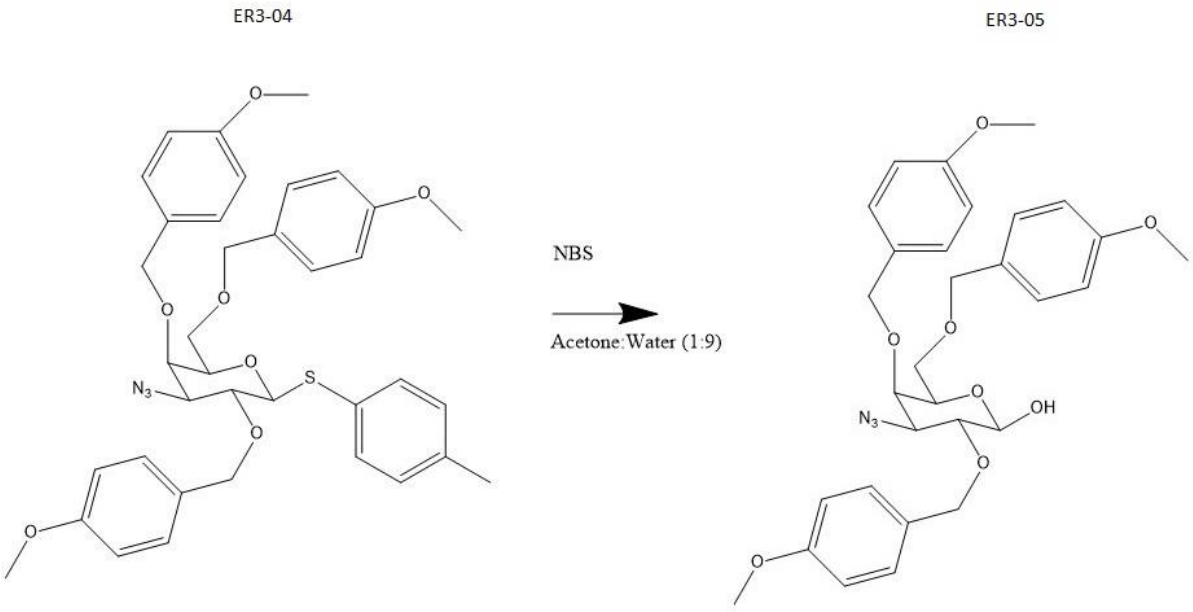


Figure 14. Schematics for the synthesis of compound **ER3-05**.

Compound **ER3-04** (1.72 g, 2.56 mmol) was dissolved in an acetone:water mixture (9:1, 40 ml) after which NBS (2.54 g, 14.1 mmol) was added portion wise. At each addition the mixture turned reddish which subsided slower and slower by each addition (probable formation of Br_2). After three hours the reaction was quenched with water. The mixture was washed three times with ethyl acetate; the organic phases were pooled; washed twice with saturated sodium bicarbonate solution; washed twice with brine; dried over anhydrous sodium

sulphate; filtered with additional ethyl acetate through a stinted glass and reduced under vacuum. The crude was purified on a flash column using a gradient system of 2:3; 1:1 and 3:2 (ethyl acetate:heptane) yielding **ER3-05** (1.51g, unknown yield). Attempts were made at drying the product which had formed a thick viscous oil but to no avail. Pending HRMS data will check the purity of the compound.

NMR for ER3-05 (3-azido-2,4,6-tri-O-(4-methoxybenzyl)-3-deoxy-α,β-galactopyranoside).

NMR data is currently too heavily obscured to confirm the structure.

References

Dahlqvist, A., Peterson, K., Mandal, S., Håkansson, M., Logan, D., Zetterberg, F., Leffler, H., Nilsson, U. Aromatic C-galactosides 2-O-substitutetd with aryl-aminopyrimidines interact with two unique sub-sites for high-affinity and high-selectivity inhibition of galectin-3. (manuscript).

Dahlqvist, A., Furevi, A., Warlin, N., Leffler, H., Nilsson, U. (2019). Stereo- and regioselective hydroboration of 1-exo-methylene pyranoses: discovery of aryltriazolylmethyl C-galactopyranosides as selective galectin-1 inhibitors. *Beilstein journal of organic chemistry.* Volume 15, pp 1046-1060.

Doak, B., Zheng, J., Dobritzsch, D., Kihlberg, J. (2016). How Beyond Rule of 5 Drugs and Clinical Candidates Bind to Their Targets. *Journal of medicinal chemistry*. Volume 59 (6), pp 2312-2327.

Ilangovan , A., Anandhan , K., Kaushik, M. (2015). Facile and selective deprotection of PMB ethers and esters using oxalyl chloride. *Tetrahedron Letters.* Volume 56 (9), pp 1080-1084.

Johannes, L., Jacob, R., Leffler, H. (2018). Galectins at a glance. *Journal of Cell Science*. Volume 131 (9), p1-9.

Leffler, H., Carlsson, S., Hedlund, M., Qian, Y., Poirier, F. (2004). Introduction to galectins. *Glycoconjugate Journal*. Volume 19 (7-9), pp 433-440.

Liu, F., Patterson, R., Wang, J. (2002). Intracellular functions of galectins. *Biochimica et Biophysica Acta (BBA).* Volume 1572 (2-3), pp 263-273

Patrick, G. (2017). *An introduction to Medicinal Chemistry.* 6th edition. Oxford: Oxford university press, pp 162-165.

Verteramo, M., Stenström, O., Ignjatovic, M., Caldararu, O., Olsson, M., Manzoni, F., Leffler, H., Oksanen, E., Logan, D., Nilsson, U., Ryde, U., Akke, M. (2019). Interplay between Conformational Entropy and Solvation Entropy in Protein−Ligand Binding. *Journal Of The American Chemical Society.* Volume 141 (5), pp 2012-2026.

Yang, R., Rabinovich, G., Liu, F. (2008). Galectins: structure, function and therapeutic potential. *Expert Reviews in Molecular Medicine.* Volume 10 (17

Zetterberg, F., Peterson, K., Johnsson, R., Brimert, T., Håkansson, M., Logan, D., Leffler, H., Nilsson, U. (2018). Monosaccharide Derivatives with Low-Nanomolar Lectin Affinity and High Selectivity Based on Combined Fluorine–Amide, Phenyl–Arginine, Sulfur–π, and Halogen Bond Interactions. *ChemMedChem.* Volume 13 (2), pp 133-137.

Zheng, Y., Qing, FL., Huang, Y., Xu, XH. (2016). Tunable and Practical Synthesis of Thiosulfonates and Disulfides from Sulfonyl Chlorides in the Presence of Tetrabutylammonium Iodide. *Advanced synthesis and catalysis.* Volume 358 (21), pp 3477-3481.

Supporting Information

Experiment and NMR data

2019-05-08 **HRMS Result for ER3-04**

MS methods:
ESI+: capillary voltage 3 kV, cone voltage 20 V,
Ext 4, Source Temp 120, Des Temp 300, Cone gas 50, Des gas 400. , continuum resolution mode, m/z 100-1200, manual lock mass correction by Leucine Enkephalin (m/z 556.2771),
Project: CAS_2019 Sample list: HRMS_1904
MS Methods: DirectInfusion, DirectInfusion_ESI_man

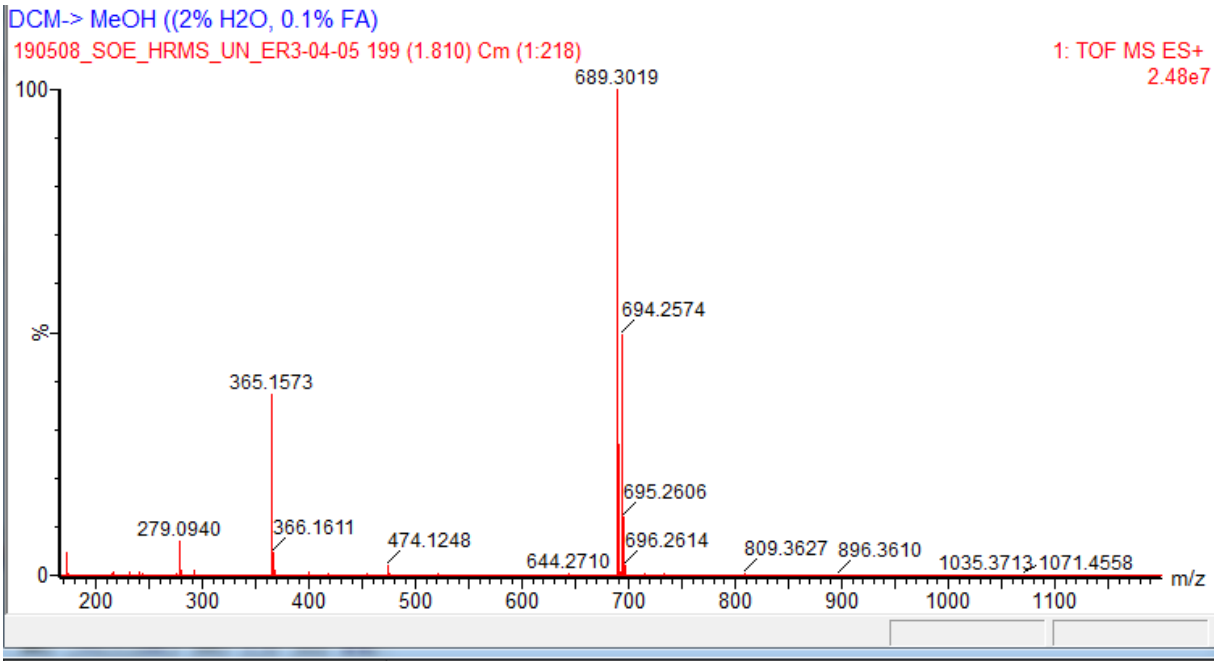
Dissolved in DCM and diluted in MeOH (2% water + 0.1% formic acid).

RESULT SUMMARY, see attachment for details:

The identity was confirmed within 2 mDa (3 ppm) .

Formula	Ion.	Found m/z	Predicted m/z	mDa*	ppm*	Comment
$C_{37}H_{41}N_3O_7S$	$+H^+$	694.2574	294.2563	1.1	1.6	And ammonium adduct at m/z 689

2019-05-08
Sofia Essen
Lunds universitet, CAS



Single Mass Analysis

Tolerance = 2.0 mDa / DBE: min = -1.5, max = 100.0

Element prediction: Off

Number of isotope peaks used for i-FIT = 3

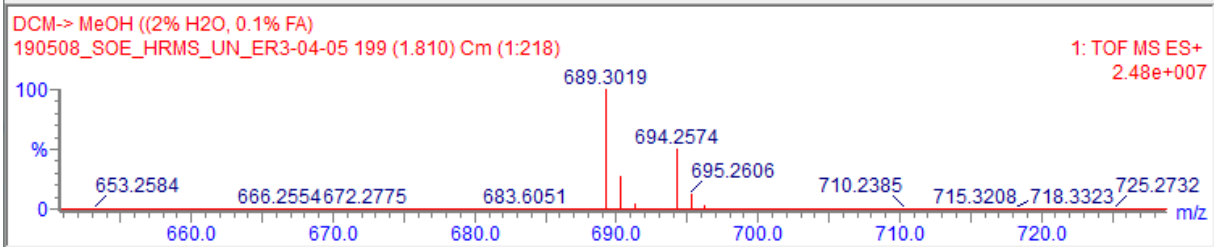
Monoisotopic Mass, Even Electron Ions

2604 formula(e) evaluated with 8 results within limits (up to 50 closest results for each mass)

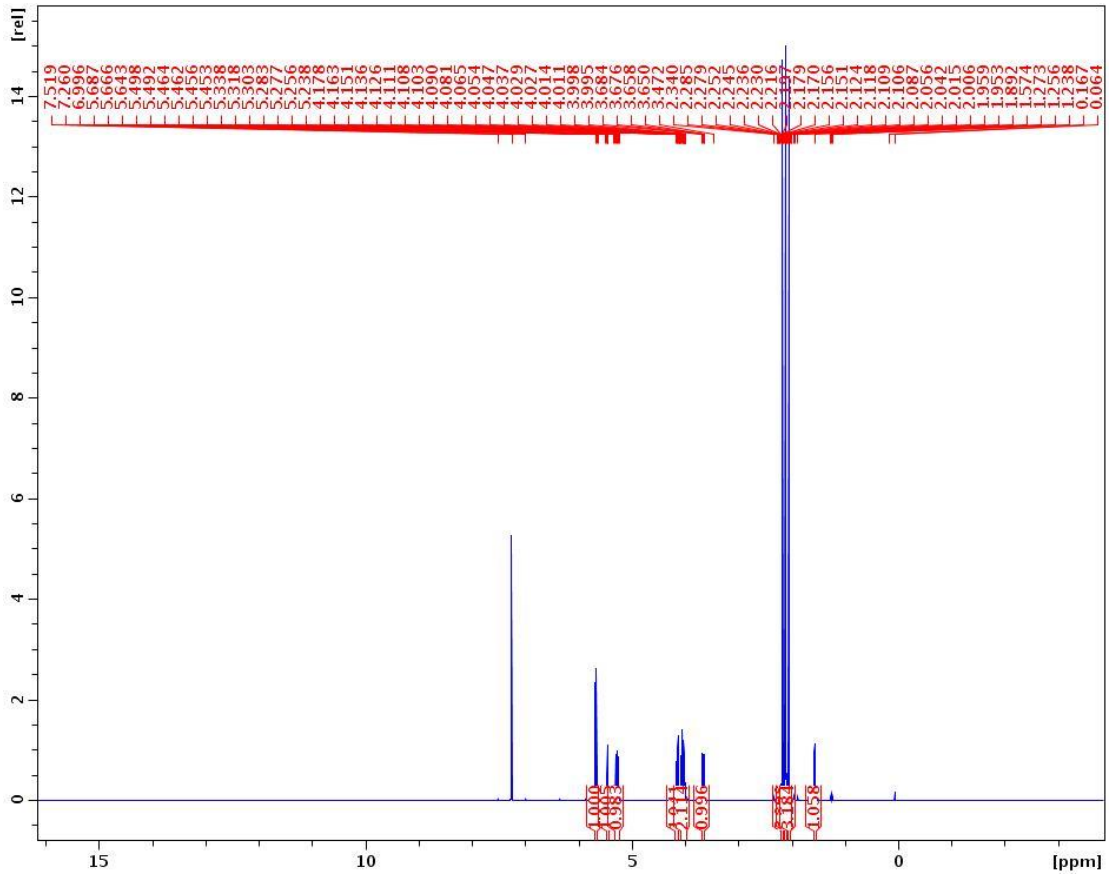
Elements Used:

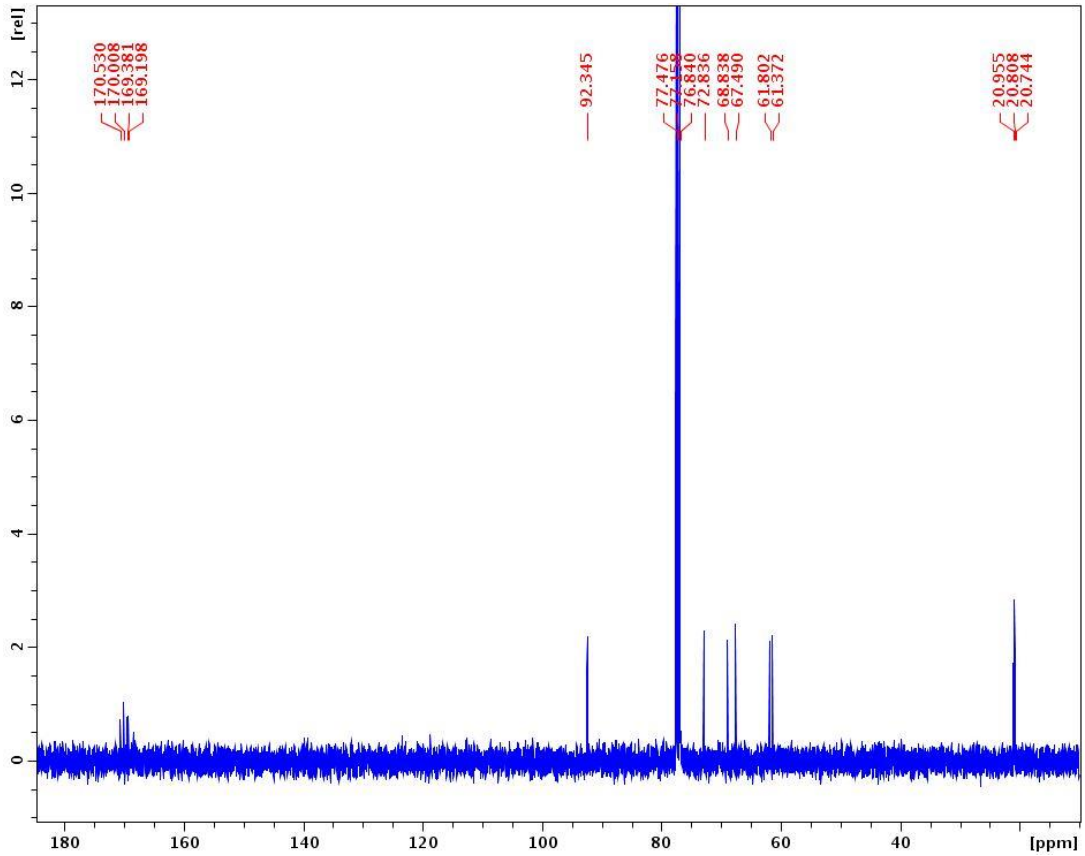
C: 0-50 H: 0-100 N: 0-5 O: 1-10 Na: 0-1 S: 0-2

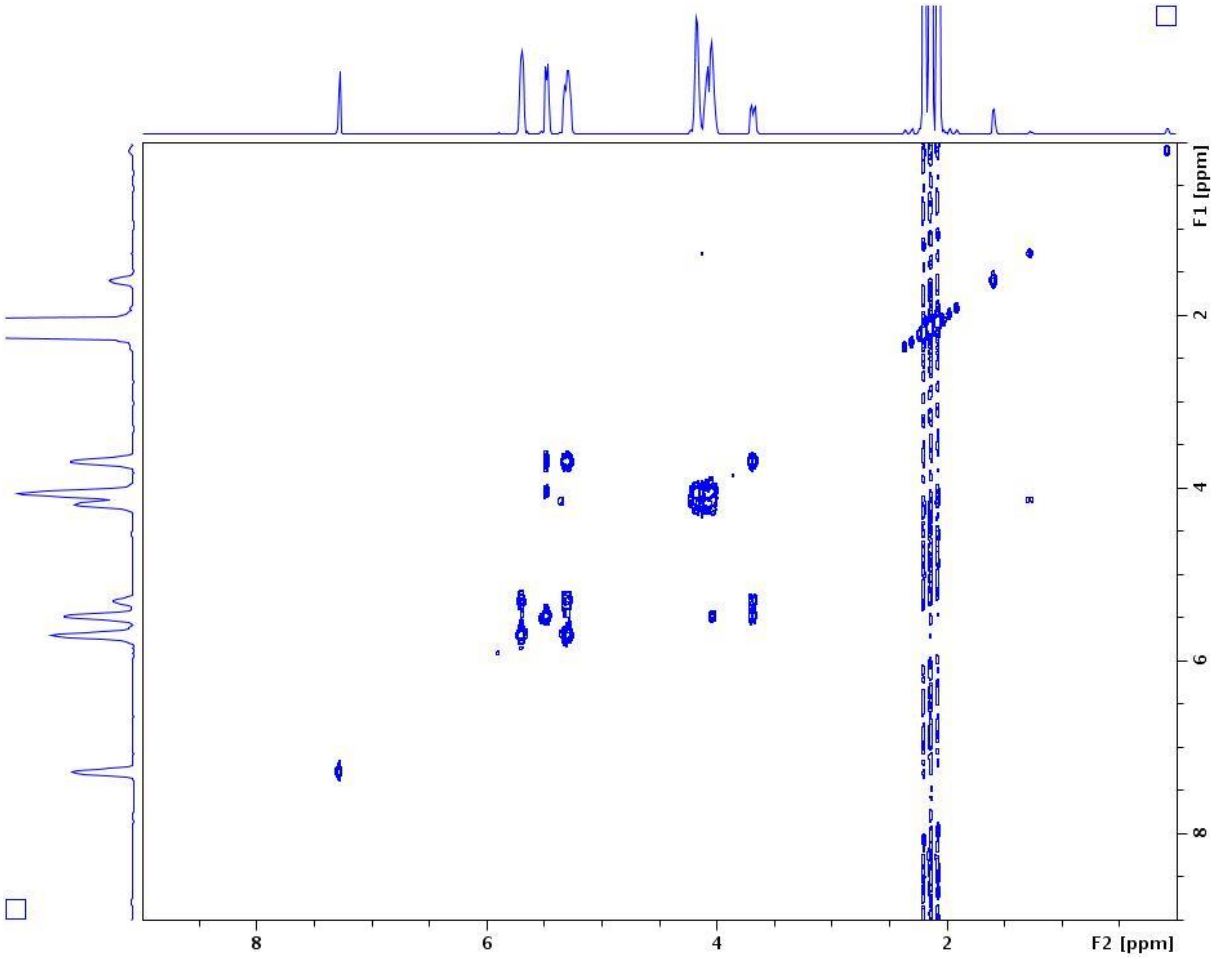
Mass	Calc. Mass	mDa	PPM	DBE	Formula	i-FIT	i-FI...	Fit Conf %	C	H	N	O	Na	S
694.2574	694.2569	0.5	0.7	27.5	C45 H37 N O5 Na	731.5	2.590	7.50	45	37	1	5	1	
	694.2580	-0.6	-0.9	12.5	C31 H44 N5 O9 S2	733.3	4.341	1.30	31	44	5	9		2
	694.2583	-0.9	-1.3	32.5	C46 H33 N5 O Na	732.2	3.249	3.88	46	33	5	1	1	
	694.2563	1.1	1.6	18.5	C37 H41 N3 O7 Na S	730.3	1.395	24.79	37	41	3	7	1	1
	694.2562	1.2	1.7	25.5	C43 H40 N3 O2 S2	733.4	4.425	1.20	43	40	3	2		2
	694.2587	-1.3	-1.9	21.5	C39 H40 N3 O7 S	729.4	0.514	59.82	39	40	3	7		1
	694.2556	1.8	2.6	9.5	C29 H45 N5 O9 Na S2	734.4	5.499	0.41	29	45	5	9	1	2
	694.2593	-1.9	-2.7	30.5	C47 H36 N O5	733.4	4.514	1.10	47	36	1	5		



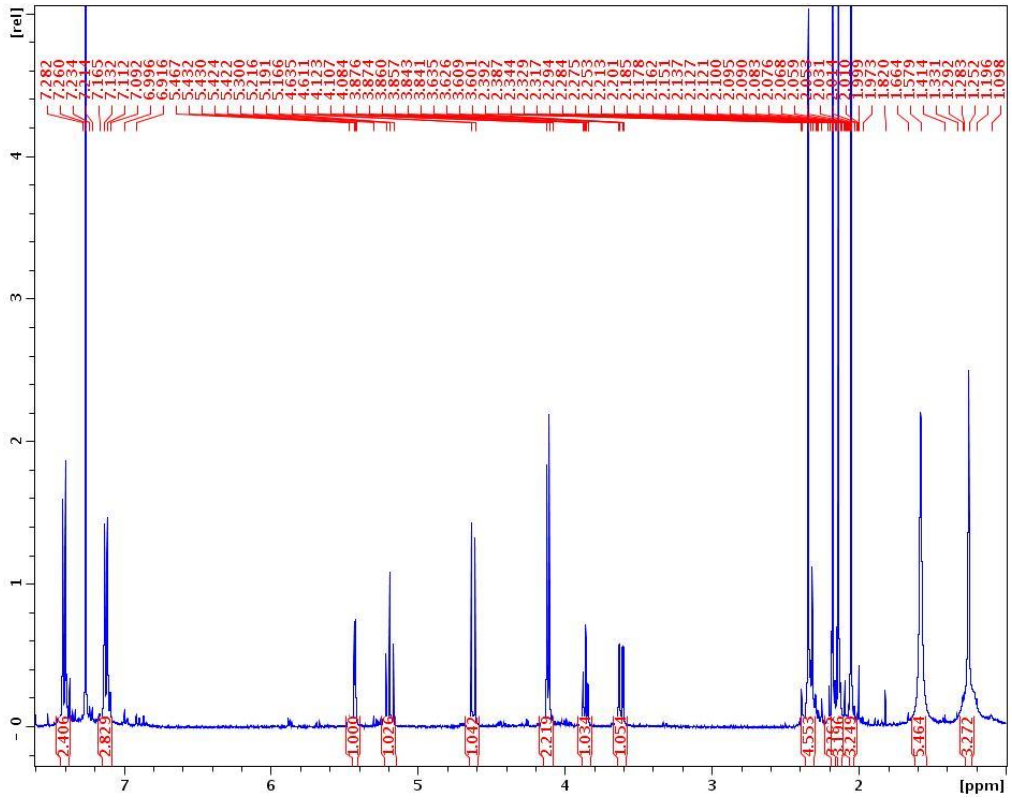
ER3-01

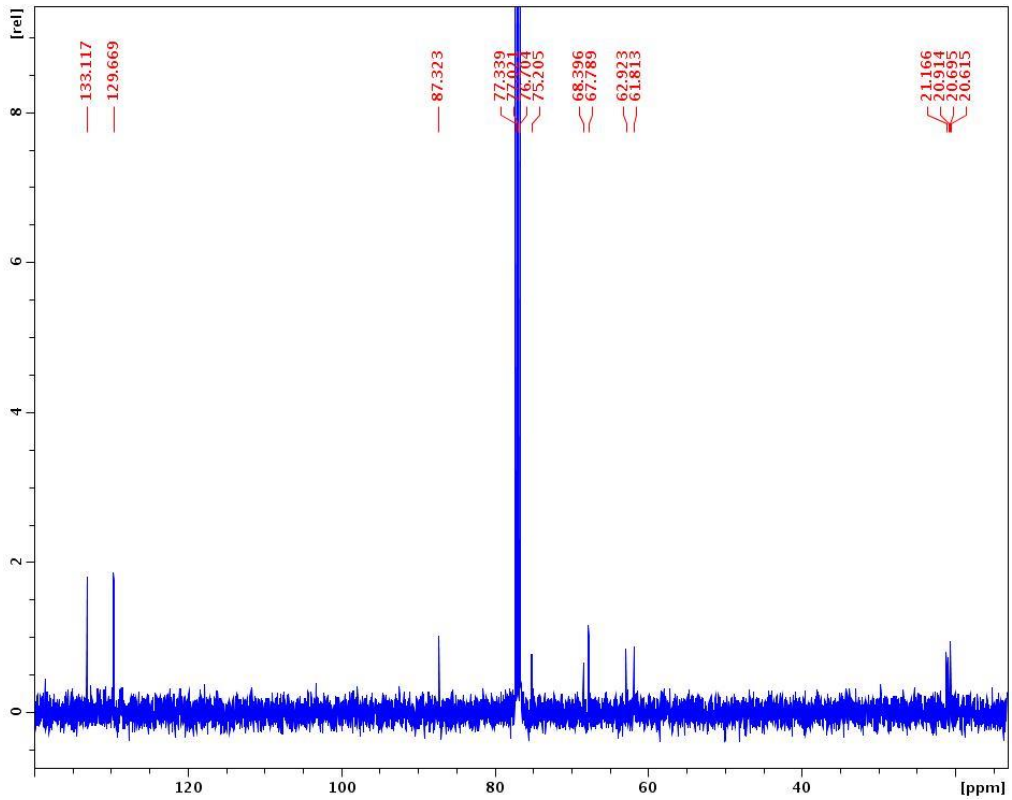

[rel]
7.519
7.260
6.596
5.687
5.666
5.643
5.498
5.492
5.464
5.462
5.456
5.453
5.338
5.318
5.303
5.283
5.277
5.256
5.238
5.178
4.163
4.151
4.136
4.126
4.111
4.108
4.103
4.090
4.081
4.065
4.054
4.047
4.037
4.029
4.027
4.014
4.011
3.998
3.995
3.684
3.676
3.658
3.650
3.472
2.340
2.285
2.279
2.252
2.245
2.236
2.230
2.216
2.197
2.179
2.170
2.156
2.151
2.124
2.118
2.109
2.106
2.087
2.056
2.042
2.015
2.006
1.959
1.953
1.892
1.574
1.273
1.256
1.238
0.167
0.064
1.000
1.085
0.983
2.111
0.996
1.058
[ppm]

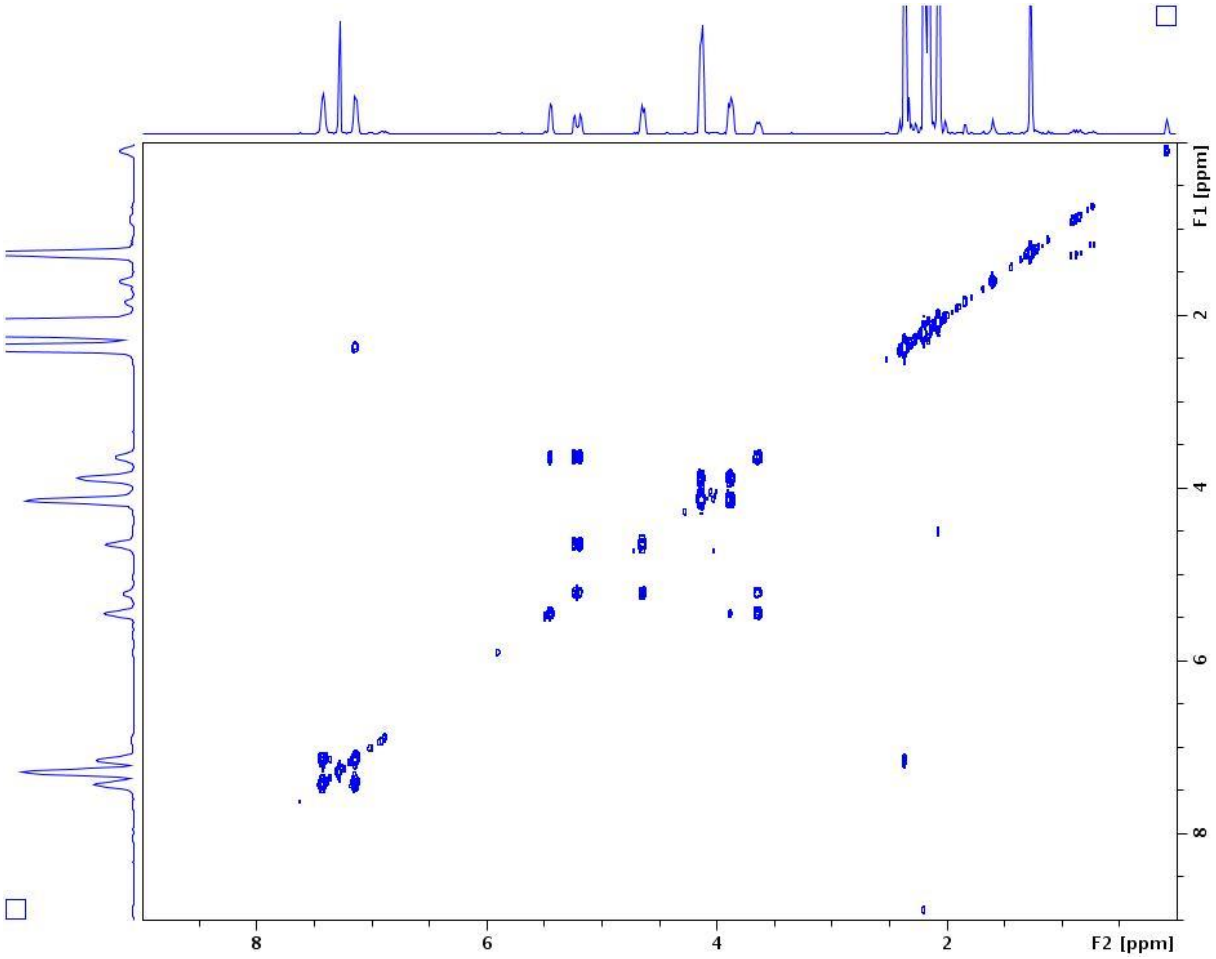

[rel]
170.530
170.008
169.381
169.198
92.345
77.476
77.158
76.840
72.836
68.838
67.490
61.802
61.372
20.955
20.808
20.744
[ppm]


F1 [ppm]
2
4
6
8
8
6
4
2
F2 [ppm]

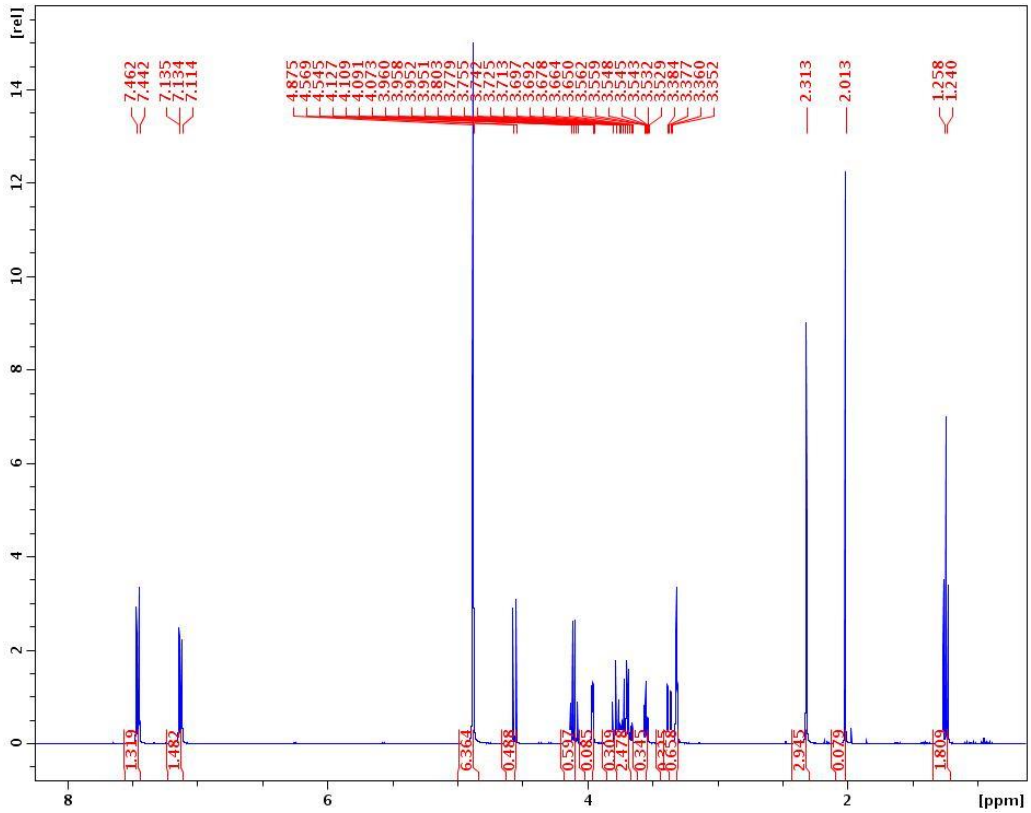
ER3-02

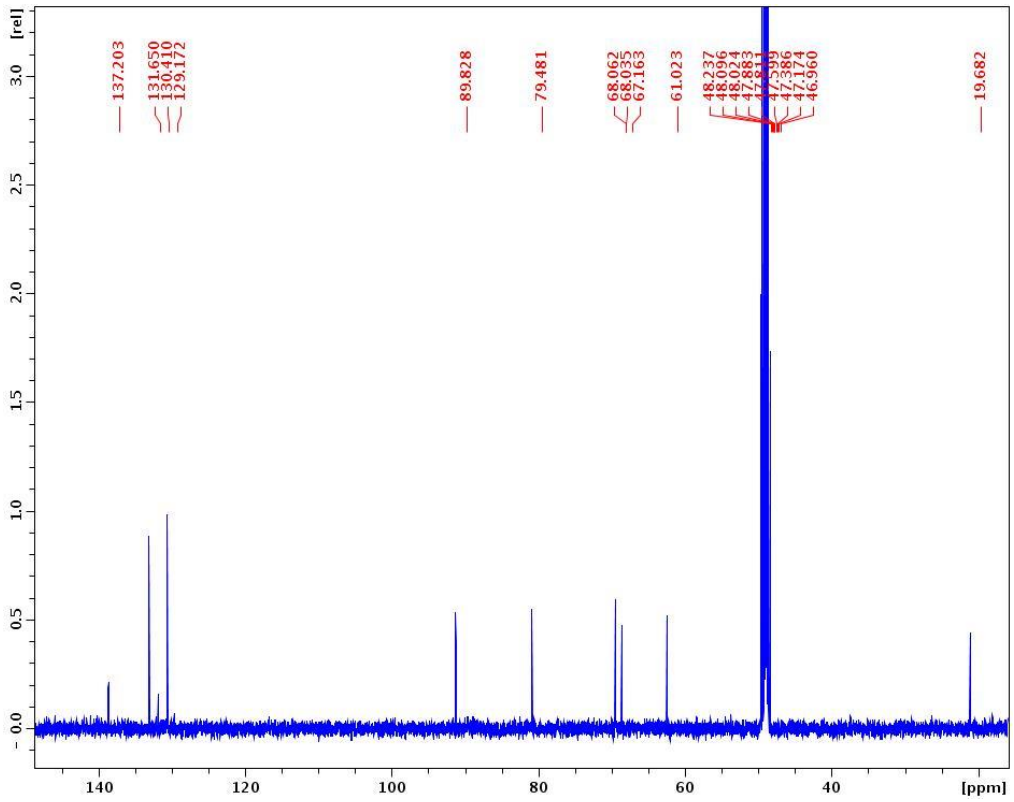

[rel]
[ppm]
7.282
7.260
7.234
7.211
7.165
7.132
7.102
7.082
6.996
6.916
5.467
5.432
5.430
5.424
5.422
5.300
5.216
5.191
5.166
4.635
4.611
4.123
4.107
4.084
3.876
3.874
3.860
3.857
3.843
3.841
3.635
3.626
3.609
3.601
3.392
3.387
3.344
3.329
3.317
3.294
3.284
3.275
3.253
3.233
3.201
3.185
3.178
3.162
3.151
3.137
3.127
3.121
3.100
3.095
3.090
3.083
3.076
3.068
3.059
3.053
3.031
3.011
3.010
1.999
1.973
1.820
1.664
1.579
1.414
1.331
1.292
1.283
1.252
1.196
1.098
2.406
2.829
1.000
1.022
1.042
2.219
1.034
1.054
4.553
4.167
3.249
3.224
5.464
3.272

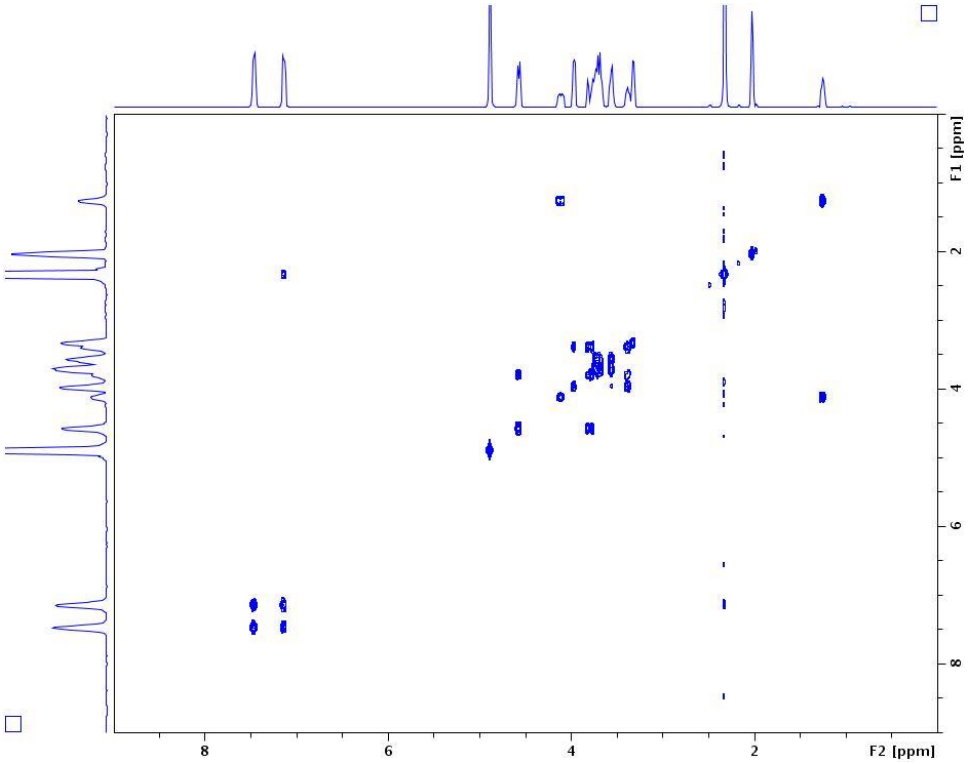

[rel]
[ppm]
133.117
129.669
87.323
77.339
77.021
76.704
75.205
68.396
67.789
62.923
61.813
21.166
20.914
20.695
20.615


F1 [ppm]
2
4
6
8
8
6
4
2
F2 [ppm]

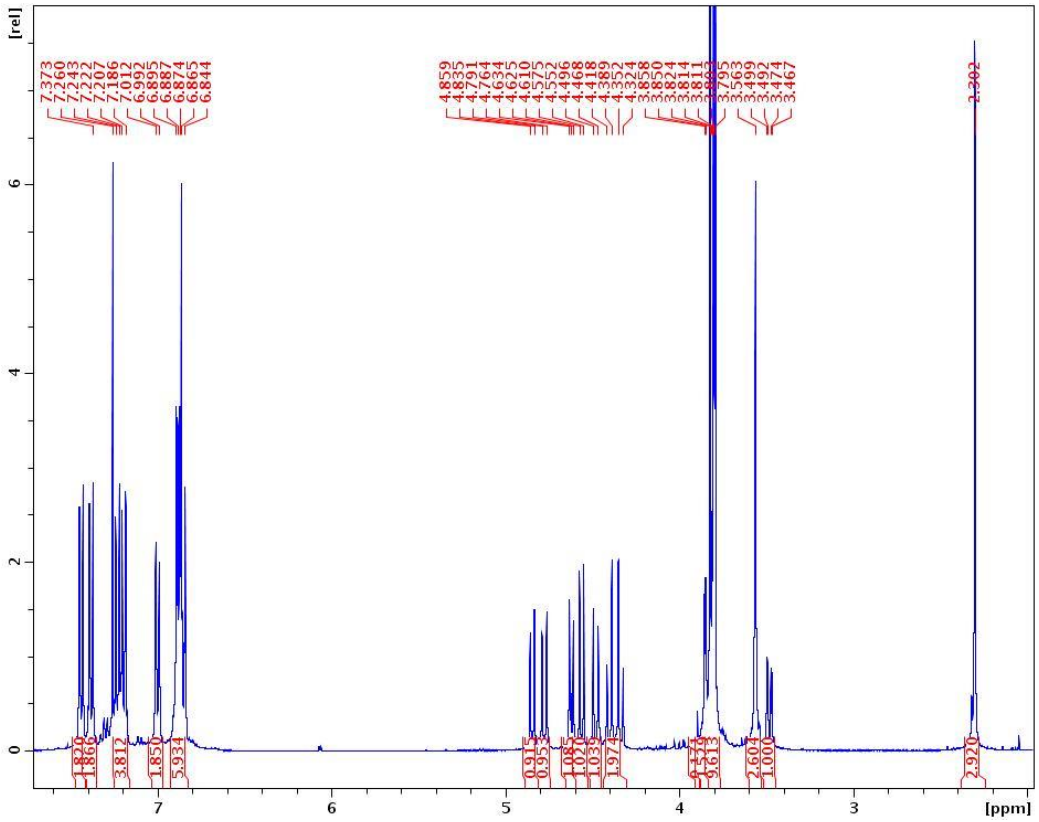
ER3-03

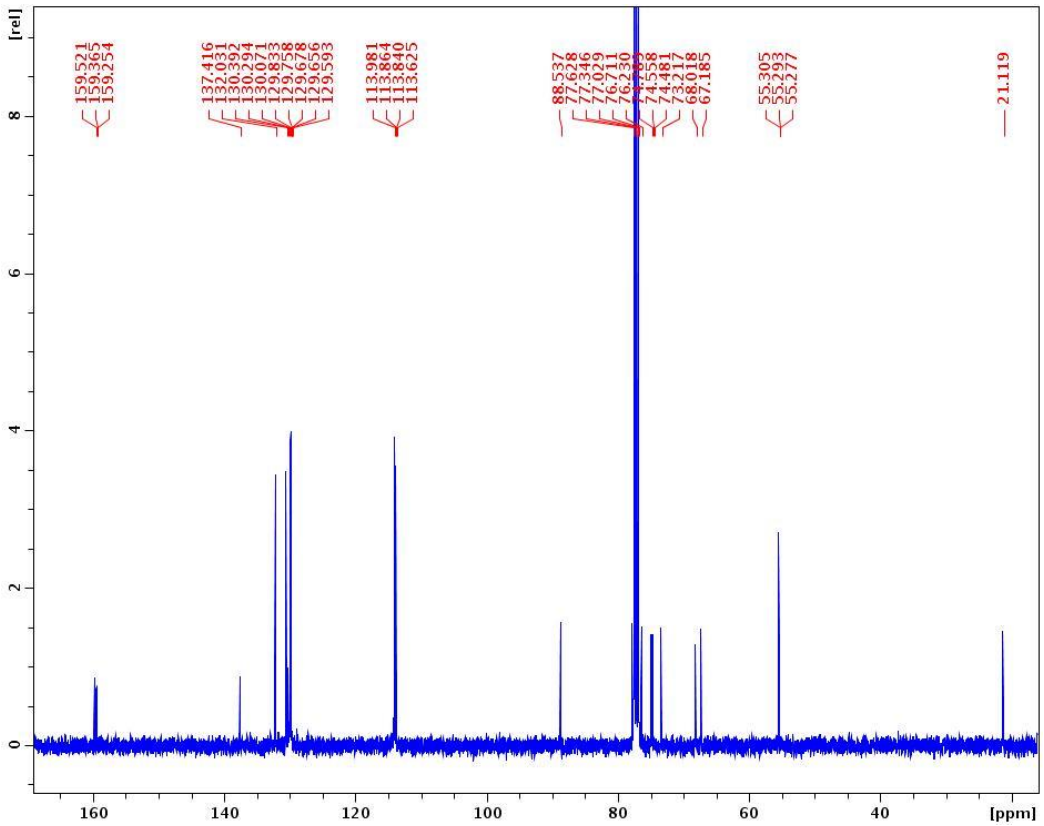

[rel]
7.462
7.442
7.135
7.134
7.114
4.875
4.569
4.535
4.177
4.169
4.091
4.073
3.960
3.958
3.952
3.951
3.803
3.779
3.755
3.742
3.725
3.713
3.697
3.678
3.664
3.650
3.562
3.559
3.548
3.545
3.543
3.532
3.529
3.384
3.377
3.360
3.332
2.313
2.013
1.258
1.240
1.319
1.482
6.364
0.488
0.597
0.085
0.309
2.478
0.335
0.245
0.638
2.945
0.079
1.809
[ppm]

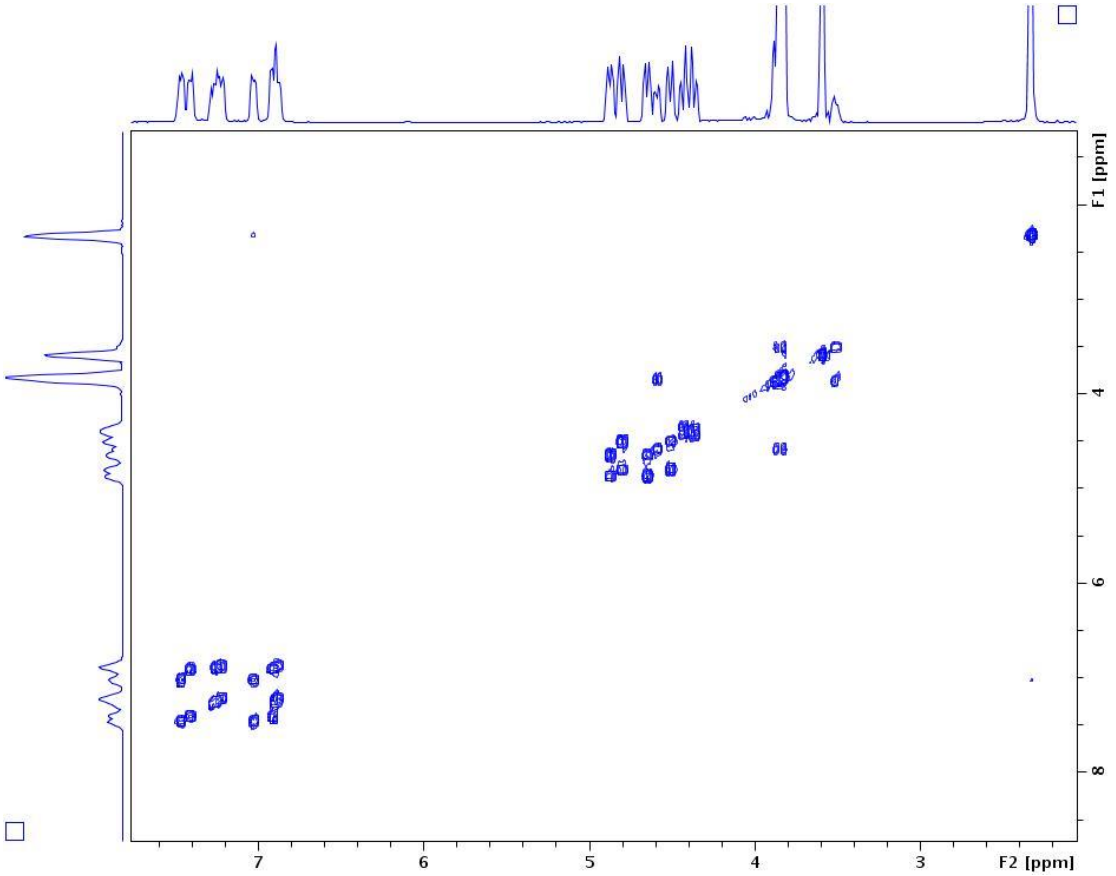

[rel]
137.203
131.650
130.410
130.171
129.172
89.828
79.481
68.062
68.035
67.163
61.023
48.237
48.096
48.024
47.883
47.811
47.599
47.386
47.174
46.960
19.682
[ppm]

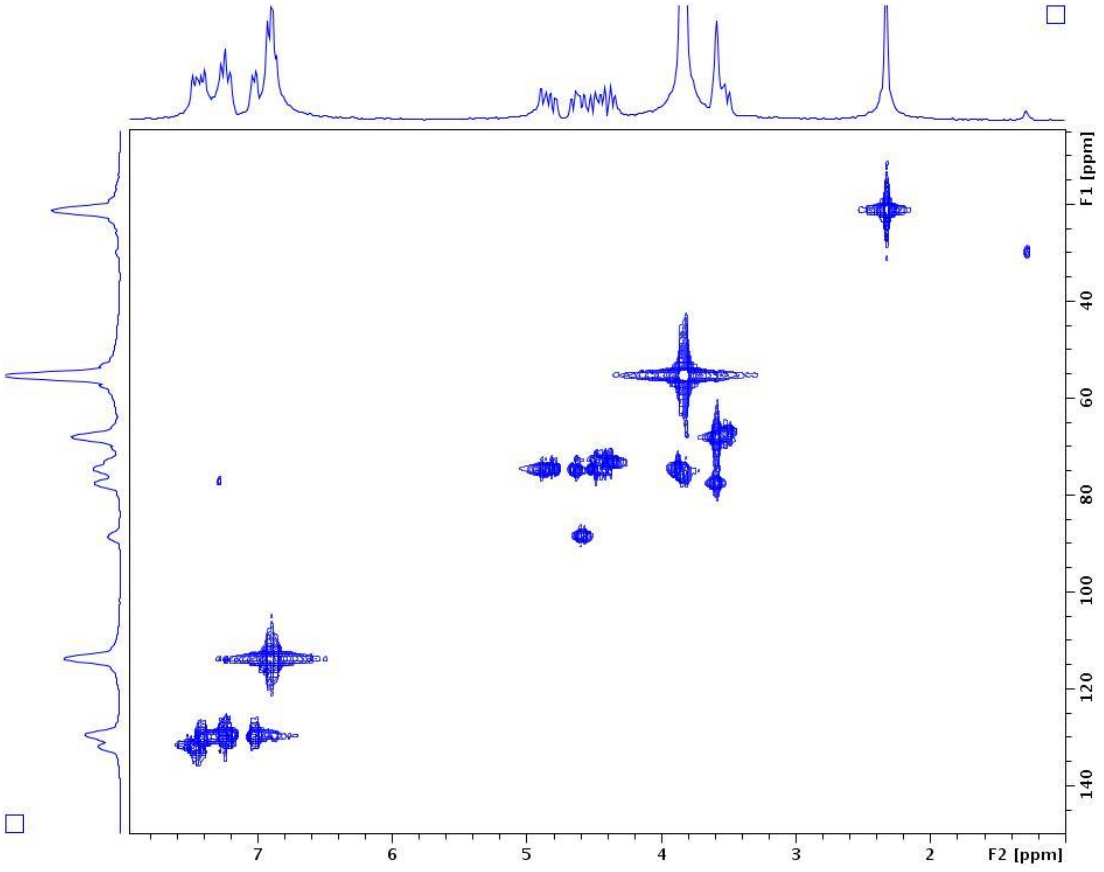

F1 [ppm]
2
4
6
8
F2 [ppm]
8
6
4
2

ER3-04


[rel]
7.373
7.260
7.243
7.222
7.207
7.186
7.012
6.992
6.895
6.887
6.874
6.865
6.844
4.859
4.835
4.791
4.764
4.694
4.625
4.610
4.575
4.552
4.496
4.468
4.418
4.389
4.352
4.324
3.858
3.850
3.824
3.814
3.811
3.795
3.763
3.493
3.492
3.474
3.467
2.302
1.826
1.866
3.817
1.850
5.934
0.915
0.953
1.085
1.076
1.035
1.974
6.137
4.529
9.613
2.604
1.000
2.920
[ppm]


[rel]
159.521
159.365
159.254
137.416
132.031
130.392
130.294
130.071
129.973
129.758
129.678
129.656
129.593
113.981
113.864
113.840
113.625
88.537
77.628
77.346
77.029
76.711
76.230
74.558
74.481
73.217
68.018
67.185
55.305
55.293
55.277
21.119
[ppm]


F1 [ppm]
4
6
8
7
6
5
4
3
F2 [ppm]


F1 [ppm]
40
60
80
100
120
140
7
6
5
4
3
2
F2 [ppm]