Tutor

Dr. Miquel Pons Vallès
Departament de Química Inorgànica i Orgànica. Secció de Química Orgànica

Grau de Química

Treball Final de Grau

GlcNAcylation of the Unique domain of Lyn A protein of the Src family

GlcNAcilació del domini únic de la proteïna Lyn A de la família Src

Núria Górriz López

June 2018

Universitat de Barcelona

B:KC Barcelona Knowledge Campus
Campus d'Excel·lència Internacional

Defiende tu derecho a pensar, porque incluso
pensar de manera errónea es mejor que no pensar

Hipatia de Alejandría

Agrair a tots els companys de laboratori, l'Hèctor, Ferran, Laareb i Guille, per l'ajuda durant tots aquests mesos. Agrair també al Dr. Miquel Pons i al Dr. João Teixeira per trobar temps per seure amb mi a resoldre dubtes i, explicar-me tot allò que no entenia.

Gràcies a la meva família i amics pel suport.

REPORT

CONTENTS

1. SUMMARY

Lyn is a protein member of the Src family of non-receptor tyrosine kinases (SFK). These proteins are involved in cell signalling pathways related to cell growth, migration, invasion, and survival. Lyn mediates the induction of apoptosis. Lyn tyrosine kinase exists as two isoforms, Lyn A and Lyn B of 56 and 53 kDa, respectively. These isoforms differ by a 21-aa insert found in the N-terminal unique domain of Lyn A. In this research, the whole protein Lyn A will not be studied, the focus will be on the Unique and SH3 domain.

O-GlcNAcylation is a post-translational modification that involves the attachment of single O-linked N-acetylglucosamine (O-GlcNAc) moieties to Ser and Thr residues of cytoplasmic, nuclear and mitochondrial proteins, by an enzymatic reaction with OGT enzyme.

Phosphorylation can also target Ser and Thr residues, then, between O-GlcNAcylation and phosphorylation exists and intensive and complex interplay. O-GlcNAcylation has been proposed to function as a nutrient and stress sensor that regulates cellular processes.

In this research, the O-GlcNAcylation of Lyn A will be studied by performing an enzymatic linked assay of NADH/NAD^+. To do so, $USH3^{LynA}$ protein and OGT enzyme will be expressed by *E. coli* and purified.

Keywords: Lyn, SKF, O-GlcNAcylation, phosphorylation, OGT, NADH.

2. RESUM

Lyn és una proteïna membre de la família Src de quinases de tirosines no receptores. Aquestes proteïnes estan implicades en vies de senyalització cel·lular relacionades amb el creixement cel·lular, la migració, la invasió i la supervivència. Lyn regula la inducció de l'apoptosi. La proteïna Lyn existeix en dues isoformes, Lyn A i Lyn B de 56 i 53 kDa, respectivament. Aquestes isoformes es diferencien per una inserció 21-aa trobada en el domini únic N-terminal de Lyn A. En aquest estudi, no es treballarà amb la proteïna Lyn A completa, sinó amb el domini Únic i el domini SH3.

La O-GlcNAcilació és una modificació post-translacional que implica la unió de N-acetilglucosamina (O-GlcNAc) a residus de Ser i Thr de proteïnes citoplasmàtiques, nuclears i mitocondrials, mitjançant una reacció enzimàtica amb l'enzim OGT.

La fosforilació també pot atacar els residus Ser i Thr, per tant, entre l'O-GlcNAcilació i la fosforilació existeix una interacció complexa. Es creu que l'O-GlcNAcilació funciona com un sensor de nutrients i estrès, que regula els processos cel·lulars.

En aquest treball, l'O-GlcNAcilació de Lyn A serà estudiada mitjançant un assaig enzimàtic acoblat amb NADH / NAD^+. Per tal de fer-ho, la proteïna $USH3^{LynA}$ i l'enzim OGT seran expressats per *E. Coli* i posteriorment purificats.

Paraules clau: Lyn, SFK, O-GlcNacilació, fosforilació, OGT, NADH.

3. INTRODUCTION

Proteins are macromolecules formed by one or more chains of amino acids that have different conformations. Proteins that lack a static three-dimensional structure are known as Intrinsically Disordered Proteins (IDPs) and they fluctuate through a range of conformations.[1] Some proteins that have well structured domains also contain disordered sequences, known as Intrinsically Disordered Regions (IDRs). These proteins are very common in eukaryotic cells because they have a role in the regulation of signalling pathways and crucial cellular processes.[1] For this reason, mutations on IDRs are associated with diseases, as perturbed cellular signalling may lead to pathological conditions such as cancer. [2]

3.1. LYN PROTEIN

The family of non-receptor tyrosine kinases known as Src family (SFK) is formed by nine proteins: Src, Fyn, Yes, Yrk, Fgr, Hck, Lyn, Blk, and Lck. These proteins are involved in cell signalling pathways related to cell growth, migration, invasion, and survival.[3] SFK family shares a common domain structure with three folded domains; SH1 contains the catalytic centre, SH2 and SH3 are regulatory domains and an N-terminal IDR that includes the SH4 domain and Unique domain (UD), which is the one that presents lower sequence homology among SFK members. However, the UD of each protein of the SFK family is well conserved between different organisms, implying a more specific role than that of a simple spacer.[4] The UDs of Fyn, Lyn, and Src mediate the induction of apoptosis.[3]

Lyn tyrosine kinase exists as two isoforms, Lyn A and Lyn B of 56 and 53 kDa, respectively. These isoforms differ by a 21-aa insert found in the N-terminal unique domain of Lyn A,[5] as shown in **Figure 1**. In this research the protein Lyn A will be studied with Unique and SH3 domain (sequence in **Appendix 1**).

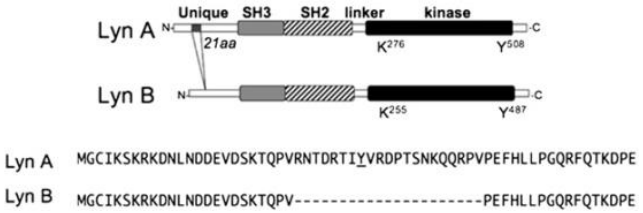


Figure 1. Schematic representation of Lyn isoforms with the sequence of 21-aa insert in Lyn A *(extracted image from Alvarez-Errico et al, ref 5).*

3.2. O-GlcNAcylation

Glycosylation is the site-specific enzymatic addition of saccharides to proteins and lipids. It has many functions in a cell: allowing the correct folding of proteins, conferring stability, allowing cell–cell adhesion and modulating intracellular signaling pathways.[6] There are many types of glycosylation, two general classes of protein-bound glycans are represented by N- and O-linked glycans. In N-linked glycosylation, the carbohydrate is bonded to a nitrogen of asparagine or arginine side chains. In O-linked glycosylation, the carbohydrate is bonded to the hydroxyl oxygen of serine, threonine, tyrosine, hydroxylysine or hydroxyproline side chains of nuclear and cytoplasmic proteins.[6]

O-GlcNAcylation is a post-translational modification that involves the attachment of single O-linked N-acetylglucosamine (O-GlcNAc) moieties to Ser and Thr residues of cytoplasmic, nuclear and mitochondrial proteins. O-GlcNAcylation has been proposed to function as a nutrient and stress sensor that regulates cellular processes that range from transcription and translation to signal transduction and metabolism. Disruption of O-GlcNAc homeostasis has been involved in the pathogenesis of many human diseases, such as cancer, diabetes and neurodegeneration. [7]

Phosphorylation can also target Ser and Thr residues, and hence there is an intensive and complex interplay between O-GlcNAcylation and phosphorylation as shown in **Figure 2**. O-GlcNAcylation plays a role in modulating protein function by affecting protein phosphorylation.[8] This reciprocal occupancy may produce different activities or alter the stability in a target protein.

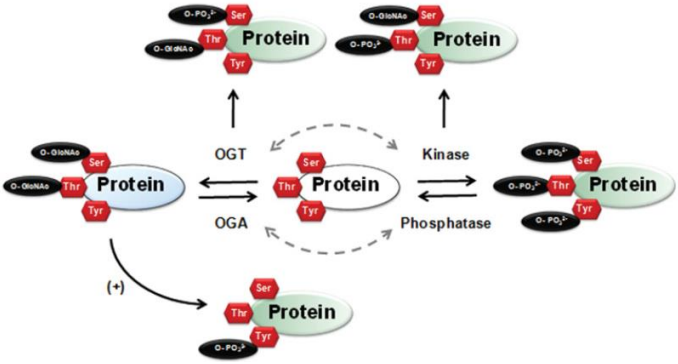


Figure 2. Interplay between O-GlcNAcylation and phosphorylation of proteins (extracted image from Lima et al, ref 6).

3.2.1. OGT enzyme

O-GlcNAcylation is controlled by two enzymes, O-Linked N-Acetylglucosamine Transferase (OGT) and O-GlcNAc hydrolase (OGA). OGT catalyses the addition of O-GlcNAc to the hydroxy group of Ser and Thr residues of a target protein using UDP-GlcNAc as substrate. OGA catalyses the hydrolytic cleavage of O-GlcNAc from post-translationally-modified proteins.[6 7]

The OGT used in this project is a sequence of the whole protein (sequence in **Appendix 1**) that contains an initial tail of histidine (Histag), a small spacer and tetratricopeptide repeat (TPR). The TPR motif is a protein-protein interaction module that mediates specific interactions with a partner protein. It appears in multiple copies in many functional proteins.[9] Some functional proteins with the TPR motif are involved in neurogenesis, cell cycle control, transcription, and peroxisomal transport. Therefore, it is likely that the TPR motif of OGT plays a role targeting the enzyme by protein-protein interactions.[10]

3.3. NADH/NAD$^+$ - LINKED ASSAY

In this project, the O-GlcNAcylation of the USH3^{LynA} protein will be studied. To see if the reaction takes place, the analysis will be performed indirectly based on the NADH/NAD$^+$ linked assay.[11] This assay consists of three coupled reactions shown in **Figure 3**. In it, a series of alcohols react with UDP-glucuronate, for this research USH3^{LynA} and UDP-GlcNAc are used instead.

If the reaction within USH3^{LynA} and UDP-GlcNAc takes place, the formation of UDP will be observed by the decrease of NADH. UDP with phosphoenolpyruvate and pyruvate kinase generates pyruvate, which reacts with NADH and lactate dehydrogenase.

$$\text{R–OH} + \text{UDP-glucuronate} \xrightarrow[\text{transferase}]{\text{UDP-glucuronyl-}} \text{R–O–glucuronide} + \text{UDP} \quad (1)$$

$$\text{UDP} + \text{phosphoenolpyruvate} \xrightarrow[\text{kinase}]{\text{Pyruvate}} \text{UTP} + \text{pyruvate} \quad (2)$$

$$\text{Pyruvate} + \text{NADH} + \text{H}^+ \xrightarrow[\text{dehydrogenase}]{\text{Lactate}} \text{lactate} + \text{NAD}^+ \quad (3)$$

Figure 3. NADH/NAD$^+$ linked assay *(extracted image from Mulder et al, ref. 11)*.

The consumption of NADH is studied with an UV-VIS spectrophotometer through absorbance measurements. As shown in **Figure 4**, NADH has two absorption bands whereas NAD$^+$ has only one. This phenomenon can be explained by the analysis of the conjugated double bonds, which absorb in the UV-VIS region, of each compound.

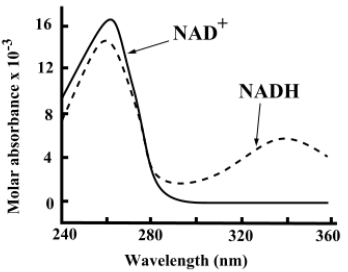


Figure 4. NADH and NAD$^+$ spectra *(By Cronholm144 [Public domain], from Wikimedia Commons)*.

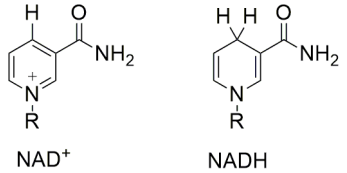
UV-VIS spectrometry provides information about compounds with conjugated double bonds. Visible light and ultraviolet light have the suitable energy to cause an electronic transition on those compounds that have π electrons. There are two only possible transitions.

The transition that requires less energy is the promotion of a non-bonding electron (n) to an antibonding molecular orbital (π*). This transition is called n → π*.

The transition of greater energy is the promotion of an electron on a π molecular orbital to antibonding molecular orbital. This transition is called π → π*.

NAD^+ has a heterocyclic aromatic ring, as shown in **Figure 5**, so the electronic transition corresponds to $\pi \rightarrow \pi^*$. There is only one kind of transition, and the absorption is at low wavelength (λ=260nm).

NADH has a heterocyclic non-aromatic ring. It has a pair of non-bonding electrons in the nitrogen of the heterocycle, that can suffer the transition $n \rightarrow \pi^*$. This transition requires less energy, so it appears at higher wavelength (λ=340nm). Equally to NAD^+, NADH has π electrons that undergo the same transition $\pi \rightarrow \pi^*$ at λ=260nm.



NAD^+ NADH

Figure 5. NAD^+ and NADH simplified structure.

4. OBJECTIVES

The aim of this research is to study if the O-GlcNAcylation of the unique domain of Lyn A takes place. To answer this question, a set of specific objectives are defined below:

- Express and purify $USH3^{LynA}$ protein.
- Express and purify OGT enzyme.
- Perform an enzymatic assay linked to the conversion of NADH to NAD^+, to see if the O-GlcNAcylation of the unique domain of Lyn A takes place.
- Quantify the O-GlcNAcylation sites in Lyn A.

5. EXPERIMENTAL SECTION

5.1. MATERIALS AND INSTRUMENTAL

5.1.1. Materials

The reagents used for the reaction performed in the experimental section are listed below:

- Uridine 5'-diphospho-N-acetylglucosamine sodium salt (UDP-GlcNAc) is purchased from Sigma-Aldrich.
- Uridine 5'-diphosphate disodium salt hydrate (UDP) is purchased from Sigma-Aldrich.
- Phospho(enol)pyruvic acid monopotassium salt (PEP) is purchased from Sigma-Aldrich.
- Pyruvate Kinase from rabbit muscle type III (P9136-1kU) is purchased from Sigma-Aldrich.
- Sodium pyruvate is purchased from Sigma-Aldrich.
- β-Nicotinamide adenine dinucleotide, reduced dipotassium salt (NADH) is purchased from Sigma-Aldrich.
- L-Lactic Dehydrogenase from bovine heart (L3916-2.5mL) is purchased from Sigma-Aldrich.

All the buffers used during the experimental section are described in **Appendix 2**.

5.1.2. Instrumental

The instrumental material used during this project is described in **Table 1**.

Instrument	Model	Software
Sonicator	KIKA LABORTECHNIK U200Scontrol	-
Centrifuge	Allegra 25R	-
Centrifuge	Avanti J-25	-
Nickel Cartridge Affinity Column	HisTrapTM HP 1mL	-
PD10	GE Healthcare	-
FPLC	AKTA FPLC W FRAC-950	Unicorn 5.20
Vivaspin centrifugal concentrator	VS2041 Sartorius Stedim	-
UV-VIS Spectrophotometer	SHIMADZU UVmini-1240	UV Probe 2.42

Table 1. Instruments used for purification and experiments.

5.2. USH3^{LynA} PROTEIN

The USH3^{LynA} protein is prepared from a precursor construct that contains the following components:

- His6-tag (histidine tail) for purification purposes.

- Glutathione S-transferase (GST) protein to ensure solubility.

- TEV (Tobacco Etch Virus) protease cleavage site.

- The SH4, Unique and SH3 domain. The SH4 domain contains a Cys to Ser mutation at position 3 to facilitate purification and prevent unwanted oxidation.

His-tag---GST---TEV---SH4---Unique---SH3

The coding DNA had been previously obtained by complete gene synthesis optimized for bacterial codon usage. This expression can be carried out with standard *E. coli* strains.

All the protocols followed for the expression and purification of proteins have been provided by the Biomolecular NMR group, University of Barcelona.[12]

5.2.1. Expression

The protein expression begins with a glycerol stock of the bacteria *E. coli* that has been transformed with a plasmid. This plasmid contains the gen to produce the protein USH3^{LynA} and

has resistance to the antibiotic Kanamycin. These stocks have been stored in aliquots at a -80ºC of temperature. To work in sterile conditions all the procedure is done in a laminar flow cabinet.

The expression culture is prepared in a 3L Erlenmeyer with 20g of Luria Broth, which is the most widely used rich medium for the growth of bacteria, and 1L of MiliQ water. Once prepared, the medium must be sterilized by heating to 121ºC at a pressure of 15psi during 15-30min in an autoclave. Through this technique, the denaturation of cellular components, including possible spores, takes place.

Simultaneously, the preculture is prepared in a sterile 50mL falcon. It is added 20mL of a solution of Luria Broth, previously autoclaved, 20µL of Kanamycin and 20µL of the glycerol stock. Kanamycin is added to eliminate all the bacteria that are in the culture apart from those that have the plasmid with resistance to the antibiotic. This means that only the bacteria that can produce the protein $USH3^{LynA}$ will grow. The solution is divided in two falcon tubes of 50mL in order to have a better oxygenation and avoid a poor growth of the bacteria. These tubes are left overnight shacking at 220rpm and 37ºC. One preculture falcon tube must be prepared for each litre of expression culture.

The following day, the falcon tubes are stored on ice until they are used. Then, 0.5mL of Kanamycin and the preculture of one of the falcon tubes are added to the expression culture. The culture is incubated at 37ºC and 130rpm during 4h approximately, until the optical density at 600nm (OD_{600nm}) reaches 0.6-0.8.

Then, the expression of the protein is induced by adding Isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 1mM. The expression culture is left overnight incubating at 25ºC and 130rpm.

The next day, the protein has been expressed so it must be collected. To do so, the solution is centrifuged at 4000rpm and 4ºC during 20min. This way, the cells sediment and the supernatant is removed. The remained pellet, that contains the cells and the protein, is resuspended in a falcon tube of 50mL with 20mL of the corresponding Lysis Buffer. The tube is stored at -80ºC until the purification.

5.2.2. Purification

The protein must be extracted from the bacterial cells and purified. The resuspended pellet is thawed with water at 4ºC for 20min. Once thawed, Dithiothreitol (DTT) is added to a final

concentration of 1mM. DTT is a reducing agent added to avoid intermolecular and intramolecular formation of disulphide bonds between cysteine residues of the protein.

The pellet is sonicated during 1min30s within intervals of 10s ON/OFF. This cycle is repeated two or three times, until the pellet becomes a liquid phase. By sonication it is applied sound energy to agitate particles in the pellet, so the cell disruption takes place and the protein is released from inside the cell. DNAase is added to a final concentration of 1mM and the falcon is left in ice for 10min. DNA is a large molecule, when it is released in the medium it causes some viscosity. DNAase cleaves DNA in smaller fragments so the solution becomes less viscous.

The sample is centrifuged at 20.000rpm for 20min at 4ºC. The supernatant is transferred in situ to a 50mL falcon tube, discarding the last drops to avoid mixing the sample with the resuspended membrane debris, that could interfere with the Nickel Cartridge affinity column.

The protein is purified with the Nickel Cartridge affinity column by the method of immobilized metal ion affinity chromatography (IMAC). The column is packed with Ni Sepharose High Performance affinity resin. This resin consists of highly cross-linked agarose to which a chelating group has been coupled. The chelating group is charged with Nickel, which selectively retains proteins with exposed histidine groups, which have imidazole groups on its lateral chain.[13]

The Support Buffer is added to the sample to equilibrate the pH to 8, due to the decrease of pH when cell lysis takes place. If the pH was lower than 7.5, the His-tag would be protonated and the pair of non-binding electrons would be lost, so would not bind the Nickel. The sample is loaded into the column and the protein is retained. To elute any proteins that could be weakly bounded to the column the Lysis Buffer is used, with a concentration of 10mM of Imidazole. Finally, the $USH3^{LynA}$ protein is eluted with the Elution Buffer, which has a concentration of 400mM of Imidazole. Imidazole competes with His-tag for the interaction with Nickel.

PD10 Desalting column, a size-exclusion column that allow for almost a complete buffer exchange, is used to remove the salt of the eluted sample, and the buffer is changed to the Lysis Buffer with 10mM of Imidazole. The sample is subjected to the TEV assisted proteolysis by adding DTT, EDTA and TEV protease at ratio 1:50 at 4ºC overnight agitating in the wheel mildly, to separate $USH3^{LynA}$ from GST.

The following day, the sample is centrifuged at 4000rpm for 15min at 4ºC. A second Nickel Cartridge affinity column is used to purify the protein. In this case, the $USH3^{LynA}$ protein will not be retained, since the GST, construct fragment containing the histidine tag, has been split.

Once the separation has been carried out, the sample is introduced in the FPLC size exclusion chromatography to finalise isolating the protein, where the molecules are separated by their size. Column is packed with porous particles. Those proteins bigger than the porous will not enter in any porous and will be the firsts to elute. Those proteins smaller than the porous will take a longer time in the column, going through those porous that can fit. Finally, the chromatogram will show the proteins, in descending size order.

The sample is eluted with Buffer F and it is collected in fractions. Only those fractions that contain the protein would be collected and then concentrated using a Vivaspin centrifugal concentrator to the desired concentration, that is measured by UV spectrophotometer.

5.3. OGT ENZYME

The OGT enzyme is prepared from a precursor construct that contains 4.5 TPR and a N-terminal His-tag. The plasmid induces resistance to kanamycin.

5.3.1. Expression

The protein expression starts with a glycerol stock of the bacteria *E. coli* that has been transformed with a plasmid. This plasmid contains the gen to produce the enzyme OGT and has resistance to the antibiotic Kanamycin. These stocks have been stored in aliquots at a temperature of -80ºC. To work in sterile conditions the whole procedure is done in a laminar flow cabinet.

The expression culture is prepared in a 3L Erlenmeyer with 1L of Terrific Broth medium. Terrific Broth is a culture medium that consists on 24g of Yeast Extract, 20g of Tryptone, and 4mL of Glycerol. Once the medium is prepared, it must be sterilized. Terrific Broth is used for the expression of OGT because it is a richer medium than Luria Broth so the cells have a better growth.

Simultaneously, the preculture is prepared in a sterile 50mL falcon tube. It is added 20mL of a solution of Luria Broth, previously autoclaved, 20µL of Kanamycin and 20µL of the glycerol stock. Kanamycin is added to eliminate all the bacteria that are in the culture apart from those that have the plasmid with resistance to the antibiotic. This means that only the bacteria that can produce the OGT protein will grow. The solution is divided in two falcon tubes of 50mL to have a better oxygenation and avoid a poor growth of the bacteria. These falcon tubes are left overnight

shacking at 220rpm and 37ºC. One preculture tube must be prepared for each litre of expression culture.

The following day, falcons are stored on ice until they are used. 100mL of Phosphate Buffer and 0.5mL of Kanamycin are added to the expression culture then, the preculture of the falcon is also added. The culture is incubated at 37ºC and 130rpm during 5h approximately, until the optical density at 600nm (OD_{600nm}) reaches 1,2-1,4.

Then, the expression of the protein is induced by adding Isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 0.1mM. The expression culture is left overnight incubating at 16ºC and 130rpm.

The next day, the protein has been expressed so it must be collected. To do so, the solution is centrifuged at 4000rpm and 4ºC during 20min. This way, the cells sediment and the supernatant is removed. The remained pellet, that contains the cells and the protein, is resuspended in a falcon of 50mL with 20mL of the corresponding Lysis Buffer. The falcon is stored at -80ºC until the purification.

5.3.2. Purification

The protein must be extracted from the bacterial cells and purified. The resuspended pellet is thawed with water at 4ºC for 20min. Once thawed, Dithiothreitol (DTT) is added to a final concentration of 1mM.

The pellet is sonicated for 1min30s within intervals of 10s ON/OFF. This cycle is repeated two or three times, until the pellet becomes a liquid phase. By sonication it is applied sound energy to agitate particles in the pellet so the cell disruption takes place and the protein is released from inside the cell. DNAase is added to a final concentration of 1mM and the falcon is left in ice for 10min.

The sample is centrifuged at 20.000rpm for 20min at 4ºC. The supernatant is transferred in situ to a 50mL falcon, discarding the last drops to avoid mixing the sample with the resuspended membrane debris, that could interfere with the Nickel Cartridge affinity column.

The protein is purified with the Nickel Cartridge affinity column since the histidine tag is retained in the column. Support Buffer is added to the sample to equilibrate the pH to 8, due to the decrease of pH when cell lysis takes place. If the pH was lower than 7.5, the His-tag would be protonated and would not bind the column. The sample is load into the column and the protein

is retained. To elute any proteins that could be weakly bounded to the column the Lysis Buffer is used, with a concentration of 10mM of Imidazole. Finally, the OGT protein is eluted with the Ellution Buffer, which has a concentration of 400mM of Imidazole.

Once the separation has been carried out, the sample is introduced in the FPLC size exclusion chromatography to finish isolating the protein. The sample is eluted with Buffer F and it is collected in fractions. Only those fractions that contain the protein would be collected and then concentrated using a Vivaspin centrifugal concentrator to the desired concentration, which is measured by UV spectrophotometer.

5.4. NADH/NAD+ LINKED CONTROL ASSAYS

As mentioned before, the study of the O-GlcNAcytalion of $USH3^{LynA}$ will be measured by coupling it with the conversion of NADH to NAD^+ by a set of three reactions. Previously to the experiments, it must be verified that all the reagents are in optimal conditions. For this reason, the following assays had been done.

All the experiments have been performed in a Buffer of 75mM TrisBase, 50mM MgCl, 2mM DTT and pH=7,3. The assays are carried out in a quartz spectrophotometer cuvette with a light path of 10mm. The total volume of each reaction is 100µL.

5.4.1. Lactate dehydrogenase assay

This assay corresponds to the third reaction of the set of coupling reactions. Lactate Dehydrogenase (LDH) enzyme is tested by the decrease of the absorbance at 340nm corresponding to the consumption of NADH.

Pyruvate	NADH	$\xrightarrow{LDH}$	NAD^+	Lactate
100µM	60µM			

Table 2. Experimental conditions.

Once the spectra at time zero is registered, 2µL of LDH 10 U/mL are added and the absorbance is measured over the time.

5.4.2. Pyruvate kinase assay

This assay corresponds to two consecutives reactions, where the enzyme Pyruvate Kinase (PK) is tested. If the enzyme PK works, pyruvate will be generated and the conversion of NADH will be observed.

UDP	PEP	$\overset{PK}{\longrightarrow}$	UTP	Pyruvate
100µM	100µM			
Pyruvate	**NADH**	$\overset{LDH}{\longrightarrow}$	**NAD^+**	**Lactate**
-	100µM	0.4U/mL		

Table 3. Experimental conditions.

Once the spectra at time zero is registered, 3µL of PK 13mg/mL are added and the absorbance is measured over the time.

5.5. USH3^{LYNA} ASSAY

Once the NADH/NAD+ linked control assays have been done correctly, and all the reagents have been verified, the reaction between LynA and UDPGlcNAc can be studied.

Many reactions are done, varying the concentration of reagents, to find the optimal conditions of the experiment. Then, more experiments are done to quantify the O-GlcNAcylation sites in LynA.

Simultaneously to each experiment, a control assay has been done to ensure that there are not any secondary reactions. The control has the same reagents as the studied reaction, except for OGT enzyme.

Observing the results of these experiments, it is decided to do another control assay to check if there is any secondary reaction between OGT and UDPGlcNAc. This control has all reagents apart from the substrate LynA.

6. RESULTS AND DISCUSSION

6.1. PROTEIN EXPRESSION AND PURIFICATION

During the purification of both proteins, USH3^{LynA} and OGT, a 10μL sample has been collected on each step of the purification to analyse it in a polyacrylamide gel electrophoresis (SDS-PAGE). Those samples are mixed with 2μL of blue gel loading dye (6x) to charge the protein with negative charge proportionally to their molecular weight.

Acrylamide gels are prepared following the established protocols.[12] Depending on the gel density the samples will run faster or slower. For those proteins with low molecular weight, such as USH3^{LynA}, a denser gel (15%) must be used. For those proteins with higher molecular weight, such as OGT, a lighter gel (12%) must be used.

Proteins on the gel are separated based on their molecular weight, small proteins run faster while large proteins run slower.

6.1.1. USH3^{LYNA} PROTEIN

As described above, purification of USH3^{LynA} consists in two Nickel affinity columns separated by TEV cleavage. In this purification 8 samples are collected, as shown in **Figure 6**. Each sample and its acronym are explained below.

Cell lysis (CL) is collected from the sample before loading it on the Nickel affinity column, it contains the whole cell expression profile. The flow through (FT) of the column contains the whole cell expression profile except for the protein of interest, which is retained on the column. The wash step (W) should contain the whole cell expression profile except for the protein of interest, confirming the effectiveness of the wash. The elution (E1) contains only the desired protein. The sample taken after PD10 desalting column (PD) is used to evaluate the loss of protein. After centrifugation of TEV reaction (TEV) three bands are observed, GST, TEV and the protein. On the second Nickel affinity column the flow through contains the protein (LynA). The elution (E2) contains the GST, TEV and some of the protein-GST that has not been cleaved. A sample of one-

week long LynA purified (LynA-1W) shows the degradation of the protein. Last sample is the marker (M).

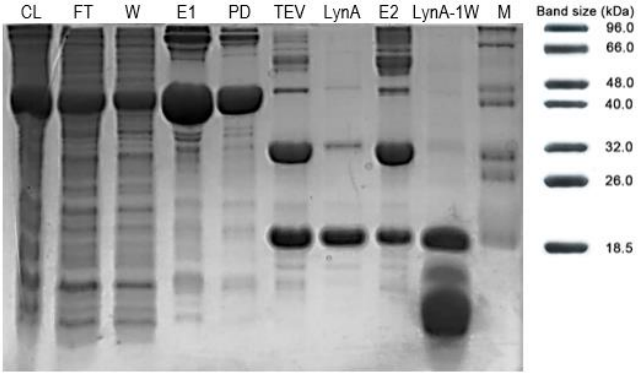


Figure 6. USH3^{LynA} purification SDS gel.

The purification of the protein has been done correctly, molecular weight of USH3^{LynA} is 14.34kDa and it matches with the marker. After the two purifications with the Nickel column the sample (LynA) is almost purified, although still contains some GST and the protein-GST. Some protein has been lost during the purification process. The SDS gel must be done once the FPLC purification has finished, otherwise the protein degrades spontaneously, as shown in sample of one-week long Lyn A.

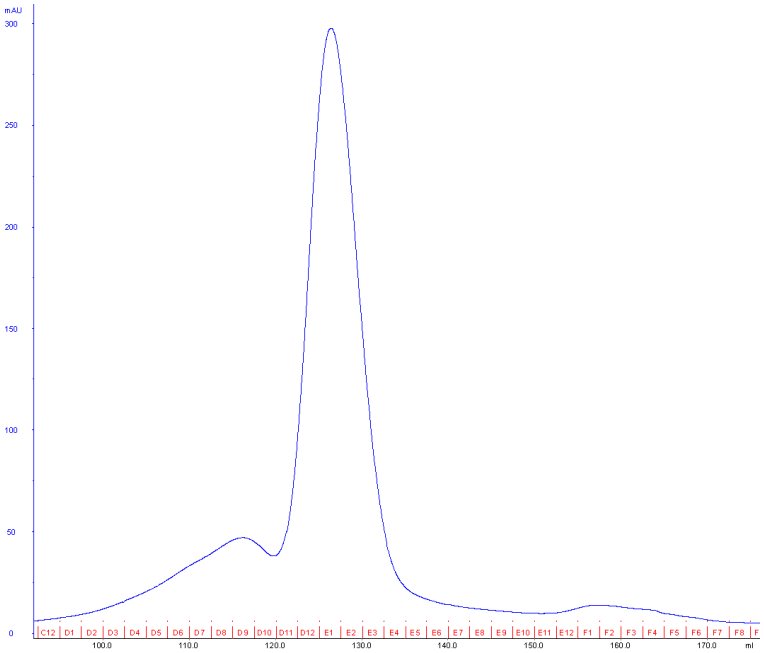


Figure 7. USH3^{LynA} FPLC profile.

After the second Nickel purification the sample is loaded in the FPLC to accomplish a better purification. **Figure 7** depicts the profile of the purification. The main peak, of 300mAu approximately, corresponds to USH3^{LynA}.

6.1.2. OGT ENZYME

As already explained, purification of OGT consists only of just one Nickel affinity column followed by FPLC. That is, the His tag is retained in the final product. While this purification 4 samples have been collected. Each sample and its acronym are explained below.

Cell lysis (CL), flow through (FT) and wash step (W) are the same as explained for USH3^{LynA} above. The elution (E) should contain only the desired protein, although it is observed that it is not completely pure. After the purification with FPLC the sample is obtained completely purified as seen in the fifth sample (OGT). Last sample corresponds to marker (M).

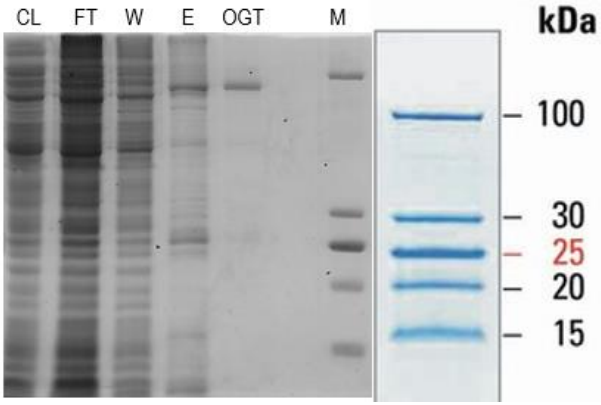


Figure 8. OGT purification SDS gel.

The SDS gel shows that the purification of OGT with Nickel column has been done correctly although there are some impurities in the sample (E), but the molecular weight of OGT, 83kDa, corresponds with the marker. After FPLC purification the protein is obtained without impurities (OGT). The profile of OGT purification in FPLC is shown in **Figure 9**.

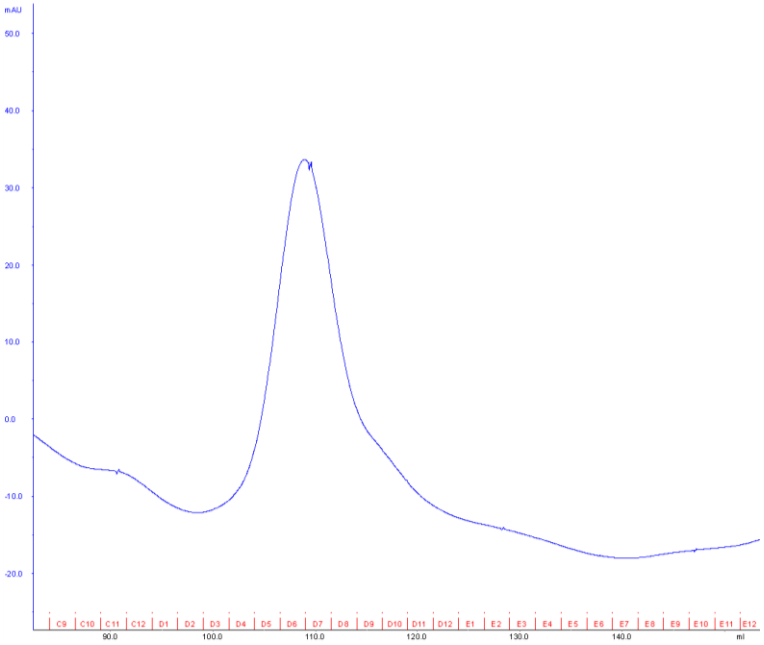


Figure 9. OGT FPLC profile

6.2. NADH/NAD+ LINKED CONTROL ASSAYS

6.2.1. LACTATE DEHYDROGENASE ASSAY

As explained above, in this assay the LDH enzyme is tested. The assay results are shown in **Figure 10**.

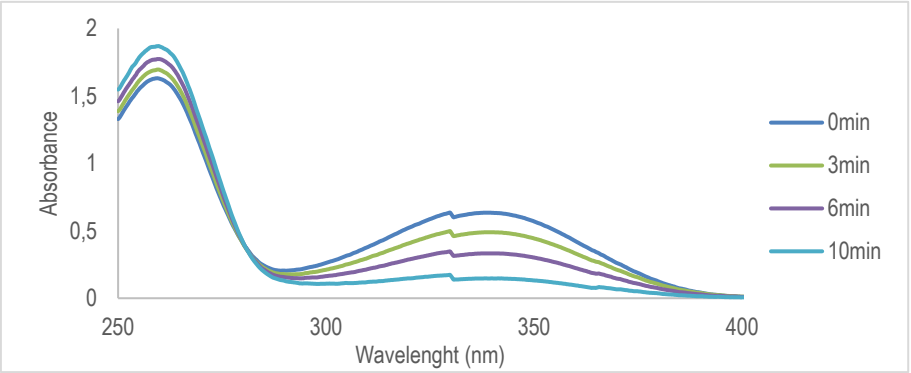


Figure 10. Lactate Dehydrogenase assay.

The decrease in the absorbance at 340nm corresponds to the consumption of NADH. Therefore, the assay of LDH has been tested correctly, the enzyme has activity in this condition and it can be used in following reactions.

6.2.2. PYRUVATE KINASE ASSAY

In this assay, PK enzyme is tested by coupling it with reaction of NADH/NAD+. The assay results are shown in **Figure 11**.

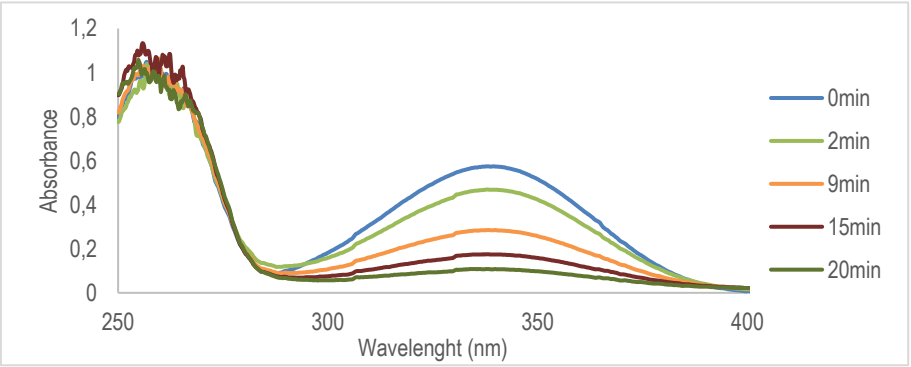


Figure 11. Pyruvate Kinase assay.

The decrease in the absorbance at 340nm corresponds to the consumption of NADH. Therefore, the assay of PK has been tested correctly, the enzyme has activity in this condition and it can be used in following reactions.

6.3. USH3^{LYNA} ASSAYS

As explained above, many experiments were done to find optimal conditions to study the O-GlcNAcylation of USH3^{LynA}. Finally, the first experiment with optimal results was performed following the conditions shown in **Table 4**.

USH3^{LynA}	**UDP-GlucNAc**	$\xrightarrow{OGT}$	**UDP**	**USH3^{LynA}-GlucNac**
60μM	60μM	2.2μM		
UDP	**PEP**	$\xrightarrow{PK}$	**UTP**	**Pyruvate**
-	120μM	0.13mg/mL		
Pyruvate	**NADH**	$\xrightarrow{LDH}$	**NAD$^+$**	**Lactate**
-	60μM	0.2U/mL		

Table 4. Experimental conditions.

The results obtained are shown in **Figure 12**. A reduction in absorbance at 340nm is observed along time, meaning that the O-GlcNAcylation of USH3^{LynA} is accomplished. This result lets us think, that at least, there is one glycosylation site, since all the NADH reagent has been consumed.

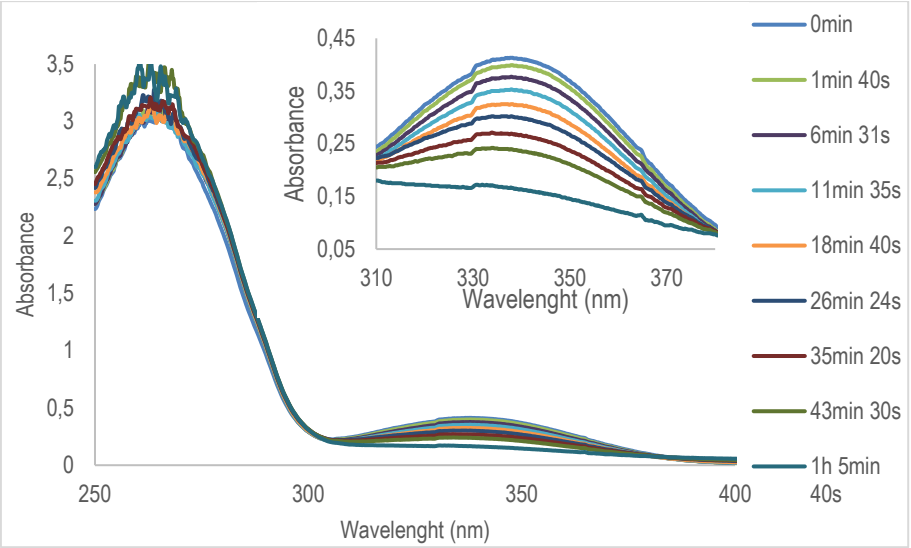


Figure 12. Experiment **1** of USH3^{LynA} assay.

After this experiment, a series of experiments have been performed. The conditions of each experiment are shown in **Table 5**, including the decrease on the absorbance at 340nm along time.

Entry	USH3^{LynA} (µM)	UDP-GlucNAc (µM)	OGT (µM)	NADH (µM)	Decrease on absorbance at 340nm (%)	Time
1	60	60	2.2	60	70	1h
2	40	80	3.0	80	50	2h
3	60	60	1.2	60	25	50min
4	60	60	0.6	60	30	1h
5	30	60	0.3	60	20	1h 50min
6	30	60	2.2	60	80	2h
7	30	90	2.2	90	80	2h 15min
8	-	60	2.2	60	60	1h

Table 5. Experimental conditions for each experiment and its result.

Results obtained in experiments **1** and **2** let us think that the O-GlcNAcylation of USH3^{LynA} is accomplished. On experiment **2** two equivalents of UDP-GlucNAc and two of NADH for each USH3^{LynA} we are added to study the stoichiometry of the reaction. Since the decrease of

absorbance in experiment **2** is of 50%, it is reasonable to say that there is only one site of O-GlcNAcylation in the protein USH3^{LynA}.

Once the O-GlcNAcylation sites have been quantified we tried to modify the kinetics of the reactions, by changing the proportions of OGT and UDP-GlucNAc with respect to USH3^{LynA} on experiments **3**, **4** and **5**.

On experiment **3** the reaction with one equivalent of UDP-GlucNAc and OGT in a proportion 1:50 was studied. The decrease on absorbance at 340nm was only of 25%. The control done simultaneously had a greater decrease on absorbance (30%) than the reaction, as shown in **Figure 13**, so no conclusions can be extracted from this experiment.

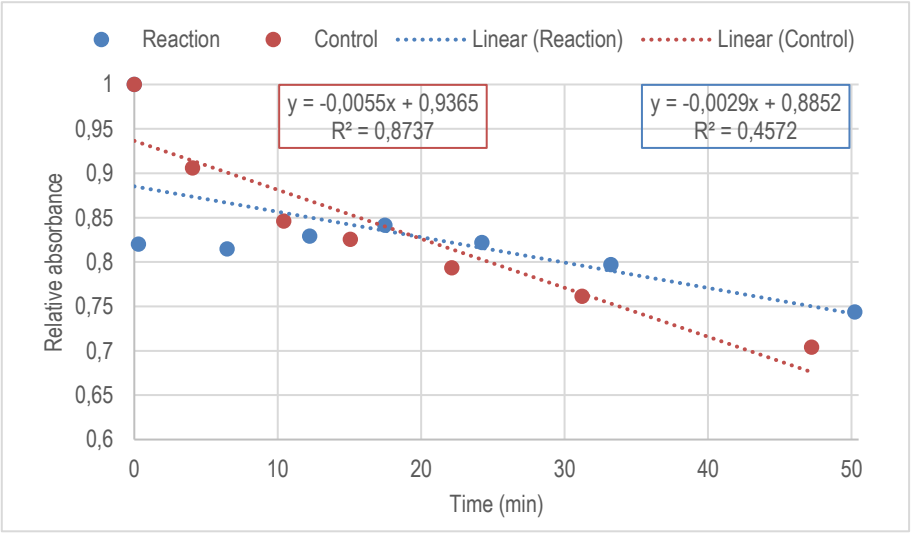


Figure 13. Relative absorbance along time of Experiment **3** and its respective Control.

On experiment **4** and **5** OGT was added in a proportion of 1:100, and there were one and two equivalents of UDP-GlucNAc, respectively. On these experiments, the decrease on absorbance was of 30% and 20% each to each. In contrast to experiment **3**, in experiment **4** and **5**, the control taken had a lower decrease on absorbance than the reaction. Even though, reactions are much slower than what was expected as seen in experiments **1** and **2**.

Since results obtained with low concentrations of OGT had not been satisfactory, it was decided to do experiments **6** and **7** with the same concentration as in experiment **1**, where satisfying results were obtained.

On experiment **6** two equivalents of UDP-GlucNAc and NADH were added, and it took place the consumption of the 80% of NADH, so it can be thought that there are two O-GlcNAcylation sites.

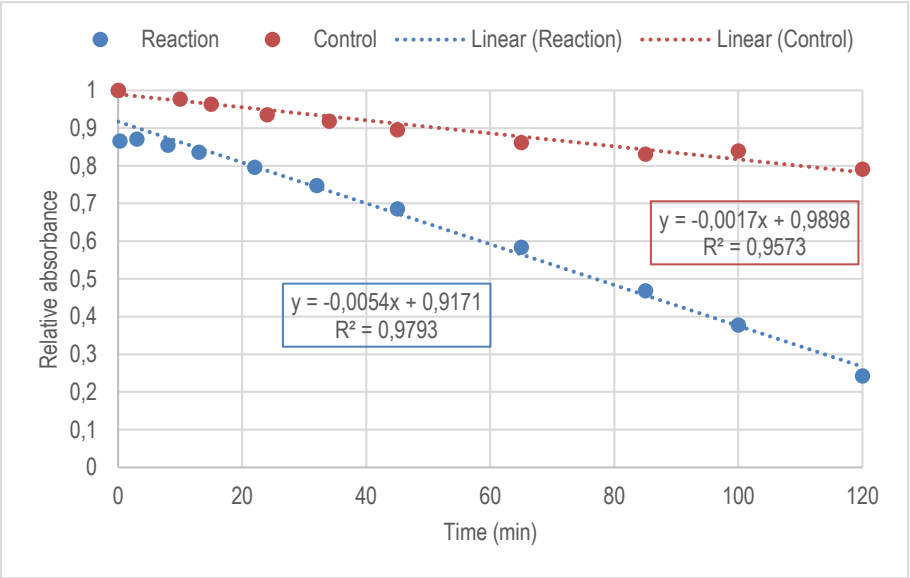


Figure 14. Relative absorbance along time of Experiment **6** and its respective Control.

Therefore, on experiment **7** three equivalents of UDP-GlucNAc and NADH were added, to study if there is a third O-GlcNAcylation site. The same way as in experiment **6**, the consumption of 80% of NADH took place as if there were 3 O-GlcNAcylation sites.

These results from experiments **6** and **7**, were found suspicious, since previously only one O-GlcNAcylation sites had been found. The USH3^{LynA} protein used for those experiments was frozen at -80ºC for a week, and it was thawed to do these experiments. When the SDS gel was done, it resulted in the protein being degraded.

Finally, a last experiment was done, **8**, to check if there was a secondary reaction between OGT and UDP-GlucNAc. The results are shown in **Figure 15**.

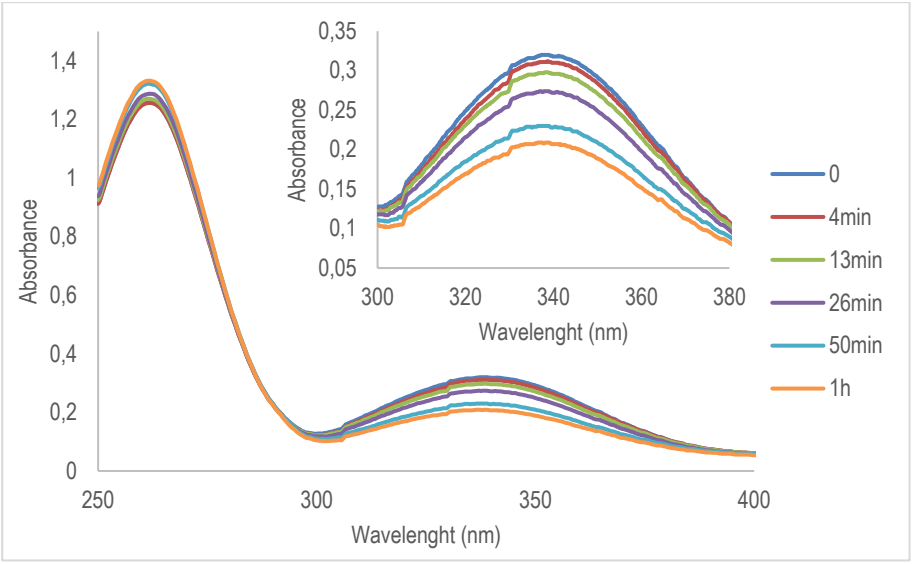


Figure 15. Experiment **8** of OGT.

It can be observed that there is a reaction between OGT and UDP-GlucNAc since there is a decrease of the absrobance of NADH at 340nm. Then, it must be considered whether the decrease on the other experiments is caused by the O-GlcNAcylation of USH3^{LynA} or OGT.

Since all experiments had different conditions, the decrease of relative absorbance of each experiment against the OGT control is compared.

Experiments **6** and **7** have the same decrease along time than the OGT control, as shown in **Figure 16**. Then, it is thought that the decrease of absorbance observed on these experiments is caused by OGT O-GlcNAcylation, so degraded protein does not react.

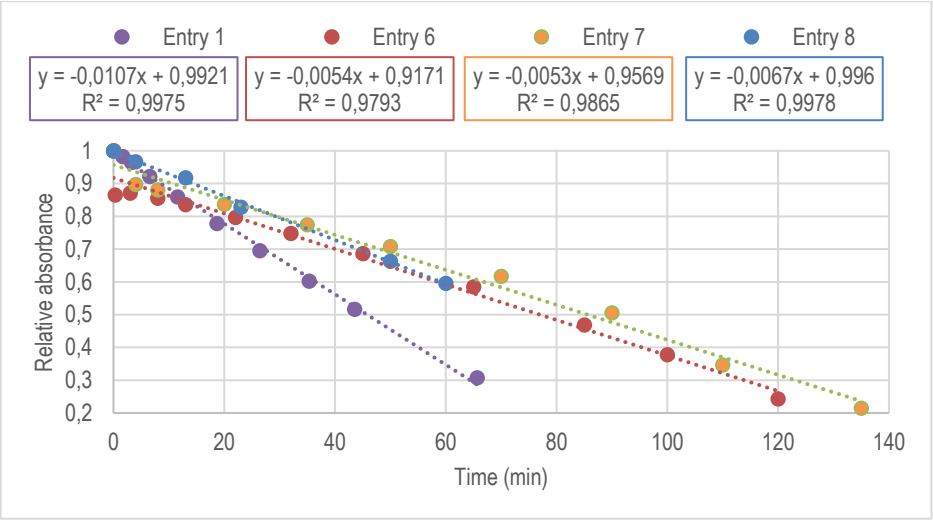


Figure 16. Relative absorbance along time of Experiment **1**, **6**, **7** and **8**.

Experiment **1**, which took place with protein not degradaded, had a greater decrease along time than the OGT control, so it is thought that both reactions, O-GlcNAcylation of OGT and USH3^{LynA}, were occurring simultaneously.

7. CONCLUSIONS

The study of the O-GlcNAcylation of the unique domain of LynA has led us to two main conclusions. Not only the O-GlcNAcylation of the unique domain of LynA takes place, but also the O-GlcNAcylation of OGT enzyme does.

OGT is a better substrate than LynA for O-GlcNAcylation. When there is a large concentration of the enzyme, both reactions take place simultaneously. At first, there is a fast decrease on the absorbance, meaning that LynA and OGT are reacting. Gradually, there is a lower decrease on the absorbance, which means that only OGT keeps reacting.

On the other hand, when there are lower concentrations of OGT, only a slow decrease on the absorbance can be observed, this is associated with the O-GlcNAcylation of OGT. As OGT is acting as a substrate of the reaction, it can not act simultaneously as an enzyme, so the O-GlcNAcylation of LynA does not occur.

It is not possible to quantitatively determine the precise split of LynA reaction from the OGT background reactions, and hence the O-GlcNAcylation sites in LynA had not been quantified.

8. REFERENCES AND NOTES

1. Wright, P. E. & Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* 16, 18–29 (**2015**).
2. Gsponer, J., Futschik, M. E., Teichmann, S. A. & Babu, M. M. Tight regulation of unstructured proteins: From transcript synthesis to protein degradation. *Science (80-.).* 322, 1365–1368 (**2008**).
3. Arbesú, M. *et al.* The Unique Domain Forms a Fuzzy Intramolecular Complex in Src Family Kinases. *Structure* 25, 630–640.e4 (**2017**).
4. Amata, I., Maffei, M. & Pons, M. Phosphorylation of unique domains of Src family kinases. *Front. Genet.* 5, 1–6 (**2014**).
5. Alvarez-Errico, D. *et al.* Functional Analysis of Lyn Kinase A and B Isoforms Reveals Redundant and Distinct Roles in Fc RI-Dependent Mast Cell Activation. *J. Immunol.* 184, 5000–5008 (**2010**).
6. Lima, V. V., Spitler, K., Choi, H., Webb, R. C. & Tostes, R. C. O-GlcNAcylation and oxidation of proteins: is signalling in the cardiovascular system becoming sweeter? *Clin. Sci.* 123, 473–486 (**2012**).
7. Yang, X. & Qian, K. Protein O-GlcNAcylation: Emerging mechanisms and functions. *Nat. Rev. Mol. Cell Biol.* 18, 452–465 (**2017**).
8. Wu, J. L. *et al.* Temporal regulation of Lsp1 O-GlcNAcylation and phosphorylation during apoptosis of activated B cells. *Nat. Commun.* 7, 1–11 (**2016**).
9. Blatch, G. L. & Lässle, M. The tetratricopeptide repeat: A structural motif mediating protein-protein interactions. *BioEssays* 21, 932–939 (**1999**).
10. Lubas, W. A., Frank, D. W., Hanover, J. A. & Krause, M. Cell Biology and Metabolism : O -Linked GlcNAc Transferase Is a Conserved Nucleocytoplasmic Protein Containing Tetratricopeptide Repeats O -Linked GlcNAc Transferase Is a Conserved Nucleocytoplasmic Protein Containing Tetratricopeptide Repeats *. 272, 9316–9324 (**1997**).
11. Mulder, G. J. & van Doorn, A. B. A rapid NAD+-linked assay for microsomal uridine diphosphate glucuronyltransferase of rat liver and some observations on substrate specificity of the enzyme. *Biochem. J.* 151, 131–40 (**1975**).
12. Protocols and guidelines, BioNMR Research Group, University of Barcelona
13. Öberg, K., Andersson, L. C., Lundqvist, J., Anderson, L. & Stålbrand, H. Effect of flow rate, molecular weight of (His)6-tagged proteins, and expression system on the performance of HisTrap HP columns. *Life Sci. News* 18, 3–5 (**2004**).

9. ACRONYMS

SFK	Src Family Kinases
Ser	Serine
Thr	Threonine
$USH3^{LynA}$	Unique domain and Src Homologous 3 domain of c-LynA
UTP	Uridine triphosphate
UDP	Uridine diphosphate
NADH	Nicotinamide adenine dinucleotide (reduced form)
NAD^+	Nicotinamide adenine dinucleotide (oxidized form)
TEV	Tobacco Etch Virus

APPENDICES

APPENDIX 1: PROTEIN SEQUENCES

USH3^{LynA} Sequence: contains the Unique domain and SH3 domain.

GA MGSIKSKGKD SLSDDGVDLK TQPVRNTERT IYVRDPTSNK QQRPVPESQL LPGQRFQTKD PEEQGDIVVA LYPYDGIHPD DLSFKKGEKM KVLEEHGEWW KAKSLLTKKE GFIPSNYVAK LNTLE

OGT Sequence: the sequence contains an initial histidine tag, a spacer and then the natural protein with 4.5 TPR and catalytic domain.

MGSSHHHHHHSSGLEVLFQGPAMCPTH ADSLNNLANIKREQG NIEEAVRLYRKALEVFPEF AAAHSNLASVLQQQGKLQEALMHYKE AIRISPTF ADAYSNMGNTLKEMQDVQGALQCYTRAIQINPAF ADA HSNLASIHKDSGNIPEAIASYRTALKLKPDF PDAYCNLAHCL QIV CDWTDYDERMKKLVSIVADQLEKNRLPSVHPHHSMLYPLSHGFRK AIAERHGNLCLDKINVLHKPPYEHPKDLKLSDGRLRVGYVSSDFG NHPTSHLMQSIPGMHNPDKFEVFCYALSPDDGTNFRVKVMAEAN HFIDLSQIPCNGKAADRIHQDGIHILVNMNGYTKGARNELFALRPA PIQAMWLGYPGTSGALFMDYIITDQETSPAEVAEQYSEKLAYMPH TFFIGDHANMFPHLKKKAVIDFKSNGHIYDNRIVLNGIDLKAFLDSL PDVKIVKMKCPDGGDNADSSNTALNMPVIPMNTIAEAVIEMINRGQ IQITINGFSISNGLATTQINNKAATGEEVPRTII VTTRSQYGLPEDAI VYCNFNQLYKIDPSTLQMWANILKRVPNSVLWLLRFPAVGEPNIQ QYAQNMGLPQNRIIFSPVAPKEEHVRRGQLADVCLDTPLCNGHTT GMDVLWAGTPMVTMPGETLASRVAASQLTCLGCLELIAKNRQEYE DIAVKLGTDLEYLKKVRGKVW KQRISSPLFNTKQYTMELERLYLQ MWEHYAAGNKPDHMIKPVE

APPENDIX 2: BUFFERS

- USH3^{LynA} Lysis Buffer pH=8

Reagent	Concentration
Trizma Base	50mM
Imidazole	5mM
Glycerol	2%
NaCl	500mM
DTT	1mM

- OGT Lysis Buffer pH=8

Reagent	Concentration
Trizma Base	50mM
Imidazole	10mM
Glycerol	2%
NaCl	500mM
DTT	1mM

- Phosphate Buffer

Reagent	Concentration
KH_2PO_4	0.17M
K_2HPO_4	0.72M

- Support Buffer pH=8: the corresponding Lysis Buffer with 0.5M Trizma Base.

- Elution Buffer pH=8: the corresponding Lysis Buffer with 0.4M Imidazole.

- Buffer F pH=7.3:

Reagent	Concentration
Trizma Base	75mM
DTT	2mM