Ryan McIntire and Dr. Gabriel Cook, Cook Research Group, Department of Chemistry
Oklahoma State University

# Expression and Isolation of the Membrane Proteoglycan Syndecan-1 from *E. coli*

## Abstract

Syndecan-1 is a membrane proteoglycan, which is a protein with glycosaminoglycans, or long carbohydrate chains, attached to its extracellular domain. The syndecan family is composed of various proteoglycans that each aid in cell-to-cell communication, and cell signaling pathways. Syndecan-1 plays a role in the initiation and progression of many different types of cancer. The overall goal of our research is to express and isolate the human syndecan-1 protein in bacterial cells, and then proceed with glycosylation of the protein. In order to have effective expression and cleavage of syndecan-1, a $His_6$-TrpΔLE fusion partner was attached to the syndecan-1 protein, and the sequence of syndecan-1 was altered to replace any cysteine and methionine residues with serine residues. The path to achieving this involves a growth of cells with emphasis on the expression of syndecan-1, a chemical cleavage of the TrpΔLE fusion partner from the syndecan-1 protein, purification by size-exclusion chromatography, and each step monitored by SDS-PAGE to show expression and purity. The syndecan-1 protein in its altered form weighs 31.903kDa, and the TrpΔLE weighs approximately 14kDa. We found that we were successful in cleavage and expression via SDS-PAGE, and got separation of protein through size-exclusion chromatography, but have not expressed a large amount of protein from growths, nor have we gotten proof of our protein expression by mass spectrometry. In the future, we plan on making slight adjustments in protocol and doing more growths to create more protein to work with, and find a reason that we are unable to identify the protein by mass spectrometry.

## Introduction

Syndecan-1 can serve as a signal for tumors and as a marker for multiple myeloma by its up-regulation within plasma cells of the bone marrow [1][2]. Even though it has been found to be in correlation with multiple myeloma, there is not much understanding of it in ways to be treated [1][2]. Myeloma is one cancer in which syndecan-1 enhances cell growth, but the presence of syndecan-1 in lung cancer and mesothelioma have actually been correlated with favorable results [3]. The seemingly contradictory relations of syndecan-1 from various studies illustrates exactly why the knowledge of the structure of syndecan-1 may be an important step in the treatment of cancers. The extracellular domain of syndecan-1 contains glycosaminoglycan groups such as heparan sulfate that are able to bind ligands. Examples of these ligands are growth factors, chemokines, or other cell stimulating agents, which all result in syndecan-1 upregulation. This syndecan-1 upregulation is observed in correlation with many cancers [3].

Cancer Significance: The binding of growth factors or chemokines to the glycosaminoglycan chains associated with the extracellular domain of syndecan-1 serves to sequester these substrates in close proximity to their respective receptors. This relative closeness of growth

factors to their receptor counterparts allows for their eventual binding [3]. Once a growth factor is bound to its designated receptor on the surface of the cell, it initiates a cascade of intracellular enzymes and reactions related to the extracellular-signal-regulated kinase (ERK) pathway, or more specifically, the mitogen activated protein kinase (MAPK) pathway. These cascades ultimately result in the production of matrix metalloproteases, which serve to cleave a portion of the syndecan-1 extracellular domain, with the location varying by the specific metalloprotease [3]. This cleaved portion of the extracellular domain still contains glycosaminoglycan chains capable of binding substrates, so its journey through the blood may trigger more intracellular cascades within cells of a different area of the body. In this way, syndecan-1 is able to propagate the effects of cell proliferation, as well as serve as a marker for possible cancer if the blood contains an abnormally large concentration of shed syndecan-1.

Research: The goal of our project is to overexpress the syndecan-1 protein in the hopes of being able to isolate and work with a large amount of it. By overexpressing syndecan-1, we can characterize its structure, function, dynamics, and interactions in cell-like conditions. Once overexpression and isolation have been proven, the next step is to glycosylate the protein to see how it may affect its structure, and therefore, its function. Because syndecan-1 is a membrane protein, its overexpression in bacteria may cause instability of the membrane, so in order to alleviate this problem, a fusion partner named TrpΔLE was included in front of the syndecan-1 sequence in the transformed plasmid. This partner causes the protein to aggregate into inclusion bodies within the cytosol, and thus prevents syndecan-1 from overburdening the cell membrane as it would ordinarily if overexpressed. This TrpΔLE fusion partner also contains a $His_6$-tag for the utilization of nickel affinity chromatography if the protein should need further purification.

**Methods**

Plasmid

The pHLV plasmid was used to insert our TrpΔLE and syndecan-1 proteins. The specific insertion of syndecan-1 is between the HindIII and BamHI restriction endonucleases. This modified plasmid was then transformed into *E. coli* cells and grown in order to set up a permanent cell stock in the -80 degree Celsius freezer.

Expression

*E. coli* cells containing the aforementioned transformed plasmid are grown in a 37 degree Celsius room starting in a 50mL tube containing 5mL of Luria Bertani (LB) medium, 10µL of cell stock, and 5µL of carbenicillin. This starter culture is grown for 2 hours before transferring 400µL from it to each of two 250mL flasks containing 100mL of LB along with 100µL of carbenicillin. These 250mL flasks are grown for 14-16 hours and then 50mL is transferred to each of four 2L flasks containing 950mL of LB along with 1mL of carbenicillin. These flasks are grown and monitored until they reach an optical density of 0.6 at 600nm and then they are induced with 1 mL of isopropyl β-D-1 thiogalactopyranoside (IPTG). Following induction, the

cells are grown for 4 hours and then pelleted via centrifugation and these pellets are stored at -80 degrees Celsius. Carbenicillin is added at each transfer step to maintain purity due to the fact the plasmid we transformed into the bacteria containing our TrpΔLE-Syndecan-1 gene also contains the gene for carbenicillin resistance.

Lysis

Resuspension 1 buffer is added to the frozen pellets, and this is sonicated to break open the cells. The resulting solution is centrifuged and the supernatant is removed. Then after the addition of resuspension 2 buffer to the pellet, we sonicate again, and that is followed by another centrifugation. Following centrifugation after resuspension 2 buffer, protein is dissolved for at least eight hours in binding buffer, sonicated, centrifuged, and then dialyzed in 5 liters of water for at least eight hours before placing the resulting precipitate on the lyophilizer.

Cleavage

Chemical cleavage of syndecan-1 from TrpΔLE is achieved using cyanogen bromide, formic acid, and water. This is all added to a tube containing the protein and repeatedly inverted on a rocker without light exposure for three to three and a half hours. Following the completion of the reaction, sodium hydroxide is added to neutralize the solution. We then transfer the solution to dialysis tubing and place it in 4L of water followed by two water exchanges, one at 30 minutes, and then another at one hour and 30 minutes. Once the dialysis is complete, we lyophilize our protein for use in the next steps.

Purification via Size-Exclusion Chromatography

Fast protein liquid chromatography (FPLC) is used to separate syndecan-1 from the TrpΔLE protein. Using the lyophilized sample following cleavage, 25mg of protein is dissolved in 500µL of 10% SDS. Once the protein is dissolved, 1.5mL of water is added to the sample and then it is centrifuged to remove any non-dissolving protein. The solution is then taken up by a syringe and injected with a filter tip onto a 2mL loop of the FPLC machine. The machine then pumps FPLC SDS buffer pH 8.2 through the loop and onto the column. As protein runs through the column, a light of wavelength 280nm provides a chromatogram of peaks indicating the passage of protein through the column and into the fraction collection tubes. By analyzing the resulting chromatogram from a sample run via FPLC, we can locate which fraction tubes contain our protein.

Mass Spectrometry Preparation

After isolating the syndecan-1 protein in fraction tubes via FPLC, we transfer the solution to dialysis tubing in water for precipitation. Then, the post-dialysis solution is centrifuged, and the resulting pellet is lyophilized before being ready for analysis by mass spectrometry. MALDI-

TOF is the typical mass spectrometry method used, and it requires the protein sample to be dissolved in an organic solvent, such as acetonitrile or tetrafluorethylene.

Electrophoresis

We utilized SDS-PAGE for quantification of each step of the process including expression, lysis, cleavage, FPLC, and mass spectrometry. This consists of running a current through a gel that was loaded with our desired samples and a molecular weight standard.

**Results**

The growth protocol has been successful based off gel runs with our pre- and post-induction samples, which showed an obvious change in band thickness and color around the 46kDa range with syndecan-1 being ~32kDa and the TrpΔLE at 14kDa. The exact molecular weight of our mutated syndecan-1 protein is 31,903Da. Following the cleavage step, the amount of protein obtained has ranged from 22-40mg. FPLC chromatograms have shown effective separation of various proteins by way of multiple peaks and their distance apart from one another. To improve separation of peaks in FPLC, we decided to run a double column setup, and the chromatogram relating to this can be seen in Figure 3. An SDS-PAGE of a post-FPLC dialysis sample revealed a thick protein band in the ~64kDa region. The below figures are relating to various steps of the protocol, with Figure 1 showing cleavage results, Figure 2 the single column FPLC chromatogram, Figure 3 the double column FPLC chromatogram, and Figure 4 the mass spectrometry results.

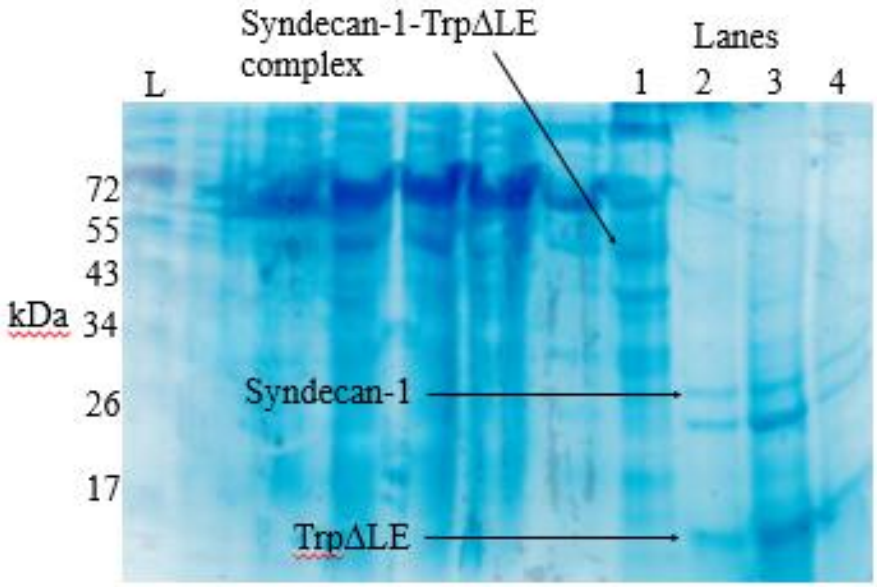


Figure 1: SDS-PAGE showing cleavage of syndecan-1

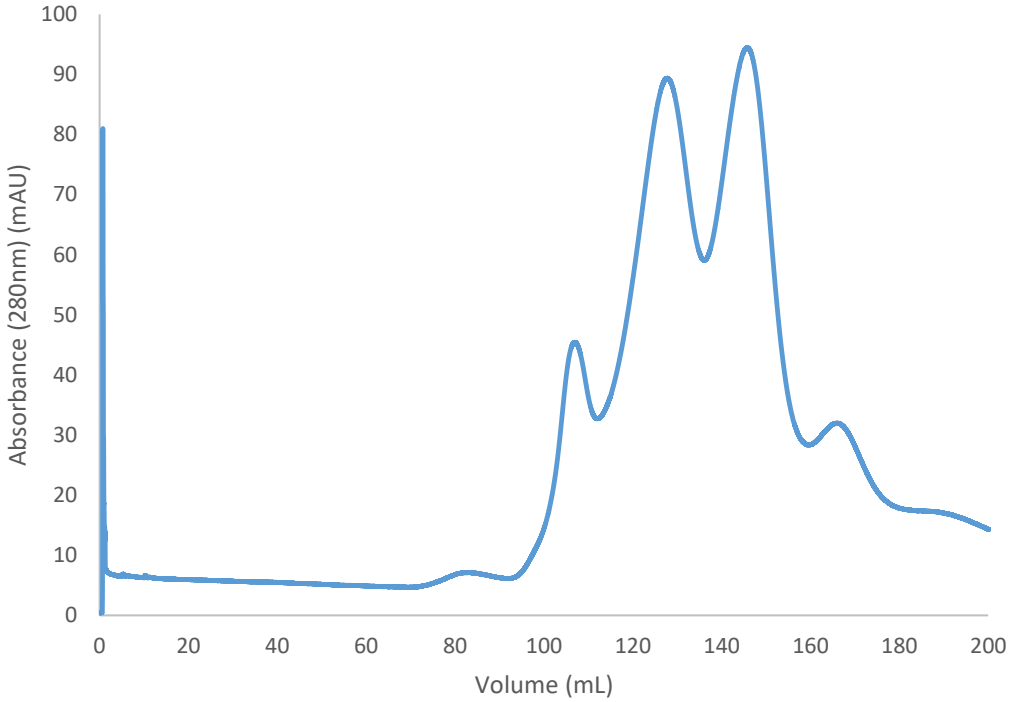


Figure 2: Single column FPLC of syndecan-1 with syndecan-1 being the second peak in the ~130mL range

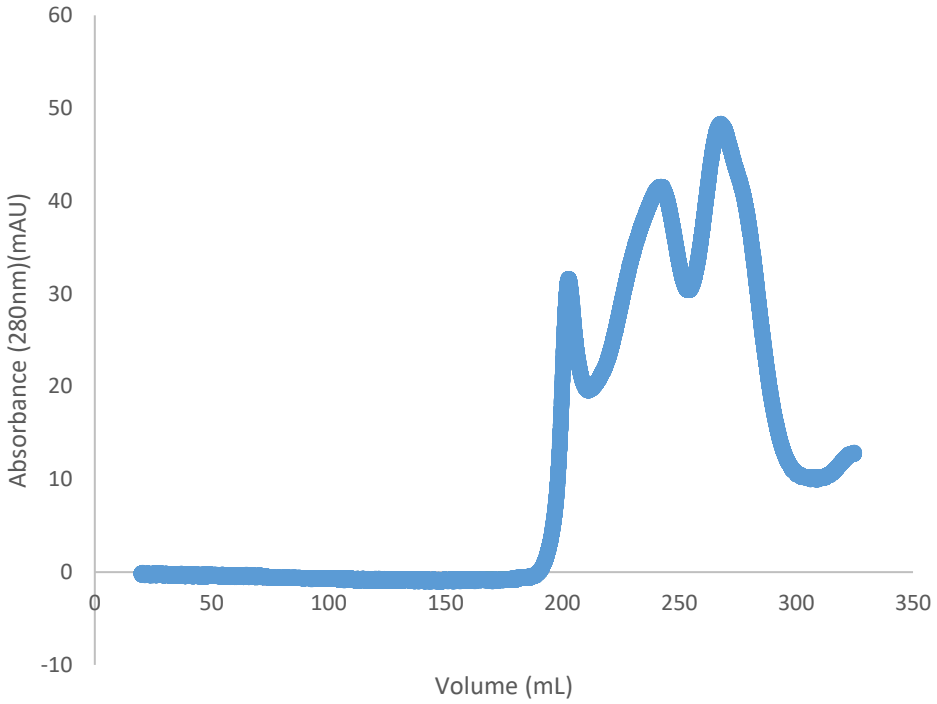


Figure 3: Double column FPLC of syndecan-1 with syndecan-1 being the second peak in the ~240mL range (truncated due to large file size)

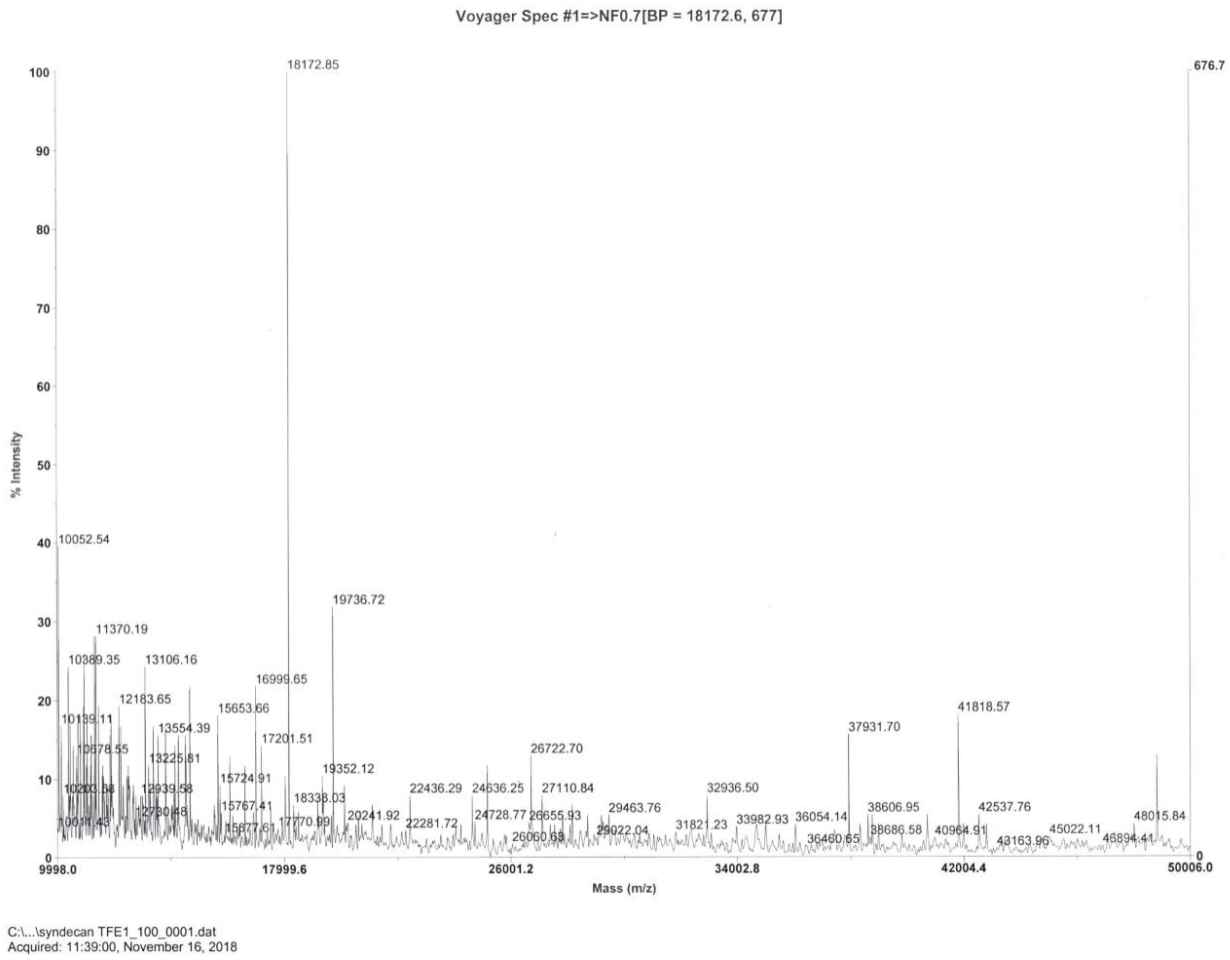


**Figure 4:** Mass spectrometry chromatogram of post-FPLC syndecan-1 sample

## Discussion

While we have evidence of expression of the syndecan-1 protein with the TrpΔLE after growth, it has not yielded a large enough amount to enable testing of different methods in terms of FPLC within the same growth. We had decided to run a gel of the resuspension samples, as outlined in the methods section, and the lanes containing the resuspension 1 and 2 supernatant samples show thick bands in the 46kDa region, which may mean that some of our protein is still soluble in the weak detergent. The second gel of the lysis samples did not show the same results as the previous one, due to a lack of band in the 46kDa region, so the lysis that was performed on the first gel may have been from errors made. As can be seen in Figure 1, there is still a lot of noise present in our FPLC-isolated fraction, and no indication of the expected 31,903Da syndecan-1 protein. The SDS-PAGE of the post-FPLC dialysis sample, which yielded a band in the 64kDa region, is likely representative of our syndecan-1 in the form of a dimer. Following cleavage, the syndecan-1 protein should be separate from the TrpΔLE, which would make it around 32kDa in mass. As a way to improve the separation of our protein peaks in FPLC, the protocol was altered to incorporate the utilization of a second column. This involves using

600mL of buffer instead of 300mL, and doubles the running time, but the extended journey of the protein through a second column proved to space out the peaks on the chromatogram, and thus separated the fraction tubes containing different proteins. Even though cysteine and methionine residues have been removed from the sequence to avoid protein misfolding by way of disulfide bonds, the protein may still have hydrophobic interactions with other syndecan-1 proteins, which can cause dimerization.

Moving forward, we intend to add salt to our mass spectrometry-prepared protein sample to see if this can interrupt possible hydrophobic interactions, and leave us with syndecan-1 monomers. We also plan to try different types of mass spectrometry such as ElectroSpray ionization, and even different organic solvents for MALDI-TOF. If we are able to verify its structure and identity through mass spectrometry, we will then attempt to characterize its structure via nuclear magnetic resonance spectroscopy (NMR). Once we have accomplished this, the goal is to glycosylate the protein to see how its dynamics and function may be affected. With the use of a lipid environment-mimicking nanodisc, we would be able to view the protein and its protein-protein interactions in conditions similar to a real cell.

**References**

[1] Indatuximab Ravtansine (BT062) In Combination With Lenalidomide and Low-Dose Dexamethasone In Patients With Relapsed and/Or Refractory Multiple Myeloma: Clinical Activity In Len/Dex-Refractory Patients

[2] Rawstron AC (May 2006). *"Immunophenotyping of plasma cells"*. Current Protocols in Cytometry. Chapter 6: Unit6.23. doi:10.1002/0471142956.cy0623s36. ISBN 0-471-14295-6. PMID 18770841

[3] Tünde Szatmári, Rita Ötvös, Anders Hjerpe, and Katalin Dobra, “Syndecan-1 in Cancer: Implications for Cell Signaling, Differentiation, and Prognostication,” Disease Markers, vol. 2015, Article ID 796052, 13 pages, 2015. doi:10.1155/2015/796052