verification of two potential glycosylation sites in human ncu-g1

liza hamdi mohamed hassan

Norwegian University of Life Sciences
Department of chemistry, BIOTECHNOLOGY AND FODD SCIENCE
Master Thesis 60 credits 2011

Acknowledgements

This thesis is a requirement for the Master of Science (M.Sc.) program in Microbiology at the Department of Chemistry, Biotechnology and Food science, Norwegian University of Life Sciences (UMB) with Professor Ragnar Flengsrud. The work in this project was conducted at the Department of Molecular Biosciences, University of Oslo under the supervision of Professor Winnie Eskild.

I would like to thank my supervisor Professor Winnie Eskild for her guidance, understanding, inspiration, encouragement, continuous help, and insightful discussions during this work.

I would like to thank my supervisor Professor Ragnar Flengsrud for his guidance, valuable advice, and help during this work.

Special thanks to my collegue Martine for the nice immunofluorescence work we did together, also thanks for the tips I got from you during my writing.

Great thanks to all members of the Eskild group for the nice time I spent with them, and the great help I got from each one of them. The time I spent with you will be in my heart for a long time. I would also like to specially thank Hilde Letnes for her continuous technical help during the lab work.

Kind thanks to my parents and siblings in Norway and Sudan, you gave me the encouragement I needed, and you were always there for me. Special thanks to my fiancé Hamdi Taweel for his continuous help, advice, encouragement and support. Looking forward to meet you soon enshalla. A very delicate thank to my fantastic friends in Sudan, you helped me lot thanks for your advice.

I really appreciate all your contributions to this work. Thank you all so much.

Ås, October 2011

_____

Liza Hamdi Mohamed Hassan

Abstract

Glycosylation is an important postranslational modification that greatly affects protein function. N-glycosylation is one class of this modification that has been characterized widely for some proteins. This glycosylation type involves the addition of glycans to asparagine amino acids. The most common protein acceptor motif for N-glycosylation is N-X-S/T. NCU-G1 is a highly glycosylated novel protein. According to the mentioned consensus motif, hNCU-G1 contains seven predicted glycosylation sites.

The aim of this study was to investigate some of the potential glycosylation sites on the hNCU-G1 amino acid sequence, and to study the effect of these sites on the protein molecular weight, expression, and subcellular localization.

Although prediction programs gave variable results concerning potential glycosylation sites on the hNCU-G1 protein sequence, they agreed on four of these sites. We changed asparagine amino acid of two of these potential sites to alanine using a site directed mutagenesis kit. These sites were located at positions 65 and 230 of the hNCU-G1 amino acid sequence.

We transiently transfected Hela cells with the wild type hNCU-G1 and the two N-glycosylation mutants (65mut and 230mut), then analyzed the expressed proteins by western immunoblotting and confocal microscopy. Both methods showed lower expression of the 65mut, as opposed to a very high expression of the 230mut. Two glycoforms of NCU-G1 proteins with the molecular weights of 62 and 75 kDa were detected. In addition, the 230mut was found to lower the apparent molecular weight of the 75 kDa glycoform by around 3 kDa. hNCU-G1 was found to localize in the cytoplasm, and it was also detected in the nuclei of the 230mut.

In conclusion, we hypothesize that glycosylation at position 65 of the hNCU-G1 amino acid sequence possibly affects its stability, folding, or antigenicity, and that glycosylation at position 230 possibly affects its antigenicity, turnover or conformation. In addition, the 230mut glycosylation site influenced the nuclear import, although this effect is not well understood. Finally, we expect that hNCU-G1 has more than one glycoform, with molecular weights that are cell and organelle specific. To this end, we recommend further characterization of more single and multiple glycosylation mutants. It would be of great value to study the exact mechanisms by which the glycosylation mutants exert their effects.

Abbreviations

A_{595}	Absorbance at 595 nm
Amp	Ampicillin
BSA	Bovine serum albumin
C-terminus	Carboxyl terminus
CNX	Calnexin
CRT	Calreticulin
DMEM	Dulbecco's modified Eagle's agar
DOL	Dolicol
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
ERAD	ER-associated degradation
Glc	Glucose
GalNAc	N-Acetylgalactosamine
GlcNAc	N-Acetylglucosamine
hr	Hour
HRP	Horse radish peroxidase
Kb	Kilobase
KDa	Kilodaltons
Man	Mannose
min	Minute
MW	Molecular weight
N	Asparagine
N-terminus	Amino terminus
NH_2	Amino group
NLS	Nuclear localization signal
nm	Nanometer
OH	Hydroxyl group

OST	Oligosaccharyltransferase
P	Phosphate
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PTM	Postranslational modification
PVDF	Polyvinylidene fluoride
QC	Quality control
RPM	Round per minute
RT	Room temperature
S	Serine
SDS PAGE	‘‘Sodium dodecyl sulfate’’- polyacrylamide gel electrophoresis
Sec	Second
STD	Standard
T	Threonine
TAE	Tris-Acetate-EDTA-buffer
V	Volt

Contents

1 Introduction

1.1 Objectives

The objectives of this study were to verify some of the predicted potential glycosylation sites on the human NCU-G1 (hNCU-G1) protein, and determine the effect of these glycosylation sites on the molecular weight, expression, and subcellular localization of this protein.

The hNCU-G1 protein coding sequence was obtained from a construct that has been made previously by our group. This construct is hNCU-G1 inserted into pcDNA 3.1 (+)/ myc-His A plasmid vector. Two of the predicted potential glycosylation sites were mutagenized, and the expressed protein analyzed by western immunoblotting.

In addition, subcellular localization of the wild type hNCU-G1 protein versus mutagenized types was compared using confocal microscope.

1.2 Postranslational modifications (PTMs)

Nearly all the proteins studied to date undergo chemical modifications, during or after polypeptide chain synthesis. These modifications take place in the endoplasmic reticulum (ER) or in other compartments of the living cell [1].

PTMs usually occur as covalent linkages or proteolytic cleavages, the latter being an irreversible modification. Covalent attachments can be removed enzymatically, and phosphorylation which is an example of a covalent linkage can be cleaved by a phosphatase enzyme. Removal of amino acids from the amino terminus is also considered as PTM [1].

The addition of molecules to a polypeptide chain occurs by an enzymatic attachment of either a functional group or unit. Two typical examples are methylation which is the addition of a methyl group to a polypeptide chain, and ubiquitination, which is the covalent linkage to the protein ubiquitin [1]. PTMS also can occur by a non-enzymatic addition such as glycation, which is the postranslational addition of any of the reducing sugars -for example glucose- to proteins. This process has been shown to increase in diabetes [2].

These modifications affect greatly the structure and function of the resulting protein. Therefore, the characterization of potential PTMs of a specific protein may help in understanding the biological mechanisms in which the protein participates.

1.3 Protein glycosylation

Protein glycosylation is the process by which an oligosaccharide is covalently linked to certain amino acids [3]. More than half of the proteins analyzed to date were found to be glycosylated. This modification mainly occurs during the translation step [4]. Previously it was believed that glycosylation biosynthesis occurs only via ER-Golgi pathway. But currently, it is well known that this process can take place in other subcellular compartments. This change in understanding came about after glycoproteins that originate from the cytoplasm or plasma membrane were recognized [5].

1.3.1 Glycosylation functions

Most proteins do not function as they should without being fingerprinted with one of the PTMs. Glycosylation represents one way of controlling protein function, as it has been found to play a major role in the following protein behaviors or functions:

- Protein folding: glycosylation affects the folding of certain proteins, since non-glycosylated versions were found to be improperly folded and subsequently accumulated in the ER and are degraded by the host ER-associated degradation (ERAD) system [6]. For example in some N-glycosylated proteins (a type of glycosylation described later), the monoglycosylated polypeptide is captured until correct folding occurs in the ER by calnexin (CNX) and calreticulin (CRT), the lectin-like chaperones [7, 8]. Many researchers ruled out this function for specific proteins [3, 9-13]. A research group went into even more details and identified that the conserved N-linked triose core ($ManGlcNA_2c$) is the glycan part responsible for folding and stability of the human immune cell receptor cluster of differentiation 2 protein (hCD2ad). But, this may not be the case for all glycoproteins [14].
- Localization and trafficking: for certain proteins, transport is controlled by glycosylation [15-18]. Doucette et al. related this function to N-glycosylation when they studied cellular distribution of the human folate receptor (FR), which is an

intracellular receptor. They determined the cellular localization of this protein before and after point mutation of one of the potential N-glycosylation sites. The change of an asparagine amino acid at position 201 of the human FR to aspartate led to the loss of glycosylation in this site. As a result, the mutated receptor can access the cell surface better than the wild type. They concluded that the loss of this glycosylation activates the trafficking [15].

- Protein stability: glycosylation confers stability for some proteins, as non-glycosylated forms can be easily degraded by proteases for instance. A good example of this is the highly glycosylated lysosomal membrane proteins Lamp-1 and Lamp-2. These proteins were found to be lost after a few hours of deglycosylation with endoglycosidase H (endo H), which is an enzyme that removes N-linked glycans from completely folded proteins. This loss was prevented using protease inhibitors, reflecting the importance of N-glycosylation in stabilizing these proteins that reside in a highly proteolytic environment [19].
- Glycosylation also plays a role in other important protein functions such as: cell-cell interactions, antigenicity [20], biological activity and half-life [21], and protein solubility [22].

1.3.2 Glycosylation types

According to the specific linkage between an amino acid and an oligosaccharide, protein glycosylation can be classified into four main branches [23, 24]:

- *N*-linked glycosylation: where a sugar is attached to the amino group (NH_2) of a specific asparagine (N) amino acid.
- *O*-linked glycosylation: where a sugar is added to the hydroxyl group (OH) of certain serine or threonine (S/T) amino acids.
- *C*-mannosylation: where an α-mannopyranosyl is linked to the indole C2 of tryptophan.
- Glycophosphatidylinositol (GPI): anchor attachments, where glycophosphatidyl-inositol groups are added near the C-terminus of a protein. This addition helps to attach the protein to the cell membrane at which the protein is residing.

N and O glycosylation comprise the two major types of glycosylation [25, 26]. For this reason, they need further elaboration, with more emphasis on N-glycosylation.

1.4 N-glycosylation

N-glycosylation as mentioned above is the addition of preformed oligosaccharide to the asparagine amino acids of the nascent polypeptide chains of secreted or membrane bound proteins [27]. Two main reactants determine the final outcome: those are the oligosaccharide donor and the protein acceptor. This process always occurs co-translationally [26, 28, 29]. One exception is the human blood coagulation factor VII. This glycoprotein has two glycosylation sites, with one site glycosylated co-translationally and the other one postranslationally [30]. In unusual conditions, N-glycosylation has also been observed to occur after the translation step. One such instance was the study of the hepatitis C virus envelope protein E1 using a truncated donor (mannosylphosphoryl dolichol deficient cells) in which the donor and acceptor were both involved in postranslational N-glycosylation [31].

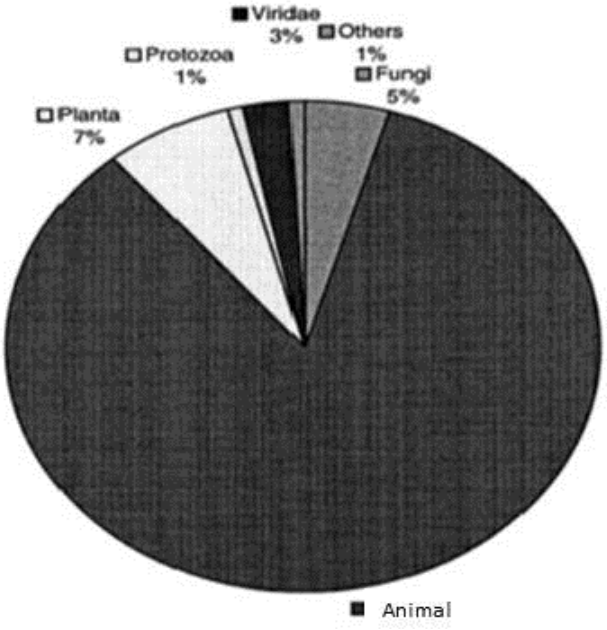


Figure 1-1 Origin of well-characterized glycoproteins, modified from the SWISS-PROT database [4].

As illustrated in figure 1-1, animal glycoproteins represent 83% of the total proteins confirmed as glycosylated. N-glycosylation is the most abundant type of oligosaccharide-protein linkages. More than 85% of the analyzed glycoproteins are either N-glycosylated or both N- and O-glycosylated [4].

A lot of studies have been conducted to characterize N-glycosylation, such as a study on human blood serum [23], human plasma [32], blood serum and cellular samples [33], and cell surface glycoproteins [34].

1.4.1 Sequence recognition motif

The sequence motif N-X-S/T, where X is any amino acid except proline, has been described as a prerequisite for N-glycosylation [35]. Some studies have proposed that threonine is more abundant at the third position of this motif than serine [3, 26, 36]. According to the SWISS-PROT database, two thirds of proteins have the N-X-S/T motif and so they are possibly N-glycoproteins [4].

At a lesser extent, cysteine or valine may form an acceptor motif when these amino acids occupy the second position after asparagine (N-X-C/V) [29]. Glycine has been described to occur at the first position after N in a non-consensus site, i.e. N-G [36].

Cui J et al. conducted a study on some eukaryotes and viruses, in which they concluded that recognition motif density is positively affected by: (a) adenine-thiamine (AT) content: where some viruses even change their AT content in order to gain more recognition motifs and so become more pathogenic, (b) N-glycan dependent quality control (QC) folding system : eukaryotes having N-glycan dependent (QC) folding system have an abundance of recognition motifs in secreted and membrane proteins, (c) N-glycan length: indirect effect, eukaryotes with longer glycan chains tend to have an N-glycan dependent (QC) folding system [37].

Although many features have been mentioned to characterize the area around N amino acid being modified, the sequence motifs described above are still not sufficient to act as glycosylation prerequisites [29]. Factors other than sequence motif may have the ability to direct glycosylation towards specific asparagine residues.

Interestingly, a non-consensus N-glycosylation site was found in position 162 of the constant domain of IgG1and2 (C_H1). The amino-acid sequence surrounding this N- glycosylation is: T-V-S-W-(N)-S-G-A-L, adding a new acceptor site (N-X-G), which was not previously presented as a glycosylation acceptor motif [38].

1.4.2 N-glycosylation biosynthesis

The attachment of a glycan to asparagine amino acids is a complex process that takes place at specific sites in the ER and Golgi apparatus, utilizing a number of molecules and enzymes [6, 11, 20, 26, 29]. Below is a brief summary of the main steps and reactants involved:

1. In the cytoplasmic part of the ER membrane, and with the help of GlcNAc-1 phosphotransferase enzyme, acetylglucosamine phosphate (GlcNAc-p) is transferred

from UDP-GlcNAc to the membrane bound polyisoprenol lipid dolicol phosphate (Dol-p). This is followed by the addition of a second GlcNAc-p, and then five mannose sugars are transferred from GDP-Man. Each added molecule is attached with the help of a specific glycosyltransferase enzyme. The resulting precursor molecule crosses the ER membrane bilayer, yielding a glycan that is exposed to the ER lumen. Here, four mannoses are accepted from DOL-p-Man. Finally three glucoses are transferred from DOL-p-Glc. This process results in a $Glc_3Man_9GlcNAc_2$-p-p-Dol formation (figure 1-2).

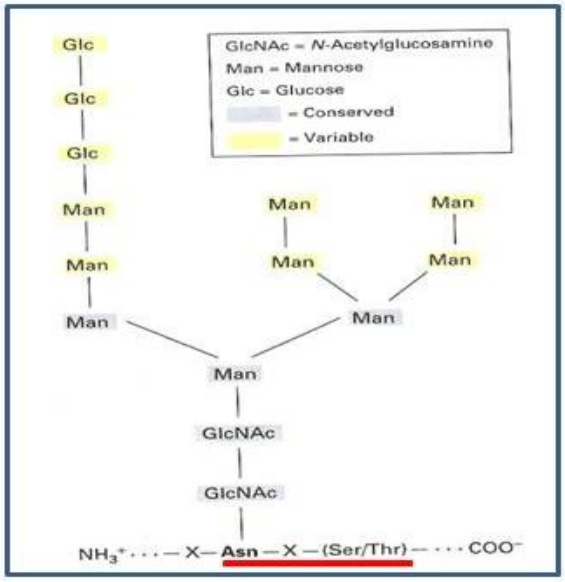


Figure 1-2 N-glycosylated protein, 14 residues attached to Asn amino acid. Consensus motif underlined with red line, figure modified from [20].

2. In the rough ER, preformed oligosaccharide precursors are covalently added to certain asparagine residues of growing polypeptide chains. This reaction is catalyzed by the oligosaccharyltransferase enzyme (OST).
3. In the ER and Golgi apparatus, these oligosaccharides are further modified, but 5 of the components remain conserved in all N-linked oligosaccharides as shown by the components highlighted in gray in figure 1-2.
4. Glucose and mannose residues are removed by glycosidase and α-mannosidase enzymes respectively, giving $Man_5GlcNAc_2Asn$.
5. For proteins that have an N-glycan dependent quality control folding system, the monoglycosylated polypeptide binds to lectin-like chaperones, and stays in the ER until complete folding occurs. It is then secreted and transported to the Golgi apparatus.

6. In the medial Golgi, diversification of glycans starts producing high mannose, complex or hybrid glycans. Finally, further sugars are added in the trans-Golgi.

1.4.3 Oligosaccharyltransferase (OST)

Oligosaccharyltransferase (OST) is the enzyme responsible for linking glycans to the asparagine amino acids of growing polypeptide chains. OST recognizes and acts on the consensus motif N-X-S/T [39]. This ER enzyme is a multi-subunit enzyme. Studies in yeast identified nine subunits, where each one assists somehow in the N-glycosylation process [40]. Earlier it was proposed that Ost1p subunit identifies N sites having the recognition motif N-X-S/T that is going to be glycosylated [41, 42]. Later, the same group found that it is a specific region in the Stt3p subunit that recognizes the consensus site and not Ostp1, and that the Ost1p luminal domain reacts with other subunits [43-45].

Ost3p and Ost6p subunits yeast paralogues genes are needed for proper glycosylation at some glycosylation sites [46].

1.5 O-glycosylation

O-glycosylation is the postranslational addition of oligosaccharides to a serine or threonine amino acid. This process takes place in the Cis-golgi apparatus [47]. There are many types of O-glycosylation; the main one being the mucin-type O-glycosylation, in which an N-acetylgalactoseamine (GalNAc) is attached to the hydroxyl group of serine or threonine amino acids [25].

Assembly starts by the transfer of GalNAc from UDP-GalNAc to S/T amino acids in the presence of the polypeptide-N-acetylgalactoseaminyltransferase (ppGalNAcT) enzyme, the basic step for all O-GalNAc glycan production. This is followed by the addition of further sugars, with the help of various enzymes, to give different O-GalNAc glycans. In comparison to N-glycan modification, no lipid precursors are needed and no enzymes are required for further processing. For the synthesis of complex O-GalNAc glycans, other enzymes are required [48]. No consensus sequences have been recognized as of yet to act as acceptor sites for O-glycosylation [49].

1.6 Glycosylation prediction

It has been of great importance to create bioinformatics programs to predict PTMs such as glycosylation, partly due to the key role PTMs play in protein function as described above, but also due to the fact that it is expensive and labor demanding to identify these modified sites experimentally. Accordingly, in the glycosylation field, researchers have constructed programs based on sequence motifs mentioned previously and the amino acids surrounding these motifs. Such programs function by using different mathematic rules. They also differ in specificity, sensitivity and accuracy. Although these programs offer potential assistance to scientists, biases cannot be excluded as some unglycosylated sites can be misinterpreted by the prediction as glycosylated, and vice versa.

Caragea et al. used ensembles of Support Vector Machines to identify glycosylated sites. In this method, they depend on information about the amino acids being glycosylated, as well as the surrounding sequence. They named the web server of their method EnsembleGly [50]. Using random forest algorithm and pairwise model, Hamby S E and Hirst J D made the GPP program (http://comp.chem.nottingham.ac.uk/glyco/). They gathered statistical informations from a frequency analysis performed using another glycosylation prediction programs. These informations concern the probability of finding certain amino acids in specific positions rather than others e.g. at position -6 Aspartate is more represented and at position -5 Methionine as well more abundant. GPP is 90.8% and 92.0% accurate to predict O and N glycosylated sites respectively. [51]. NetNglyc (http://www.cbs.dtu.dk/services/NetNGlyc/) and NetOglyc (http://www.cbs.dtu.dk/services/NetOGlyc/) are other examples of prediction program. As mentioned in the article about NetOglyc, this server was built depending on sequence context, protein secondary structure and surface accessibility. Regarding sequence, they abstracted a general charge dependent role: (a) negatively charged amino acids, especially glutamic acid are more abundant, (b) while positively charged amino acids are less frequent at position -1 and +3, they are favored at position +1 [49].

1.7 Human kidney predominant protein (hNCU-G1)

NCU-G1 is a novel protein, first discovered in a study of the complement component number 3 (C3) in mouse embryonic carcinoma (EC) cells [52], then later found among 12 unknown proteins in a study of human placental proteins [53].

At the mRNA level, NCU-G1 has been detected in most human tissues, being most abundant in the kidney especially in cortex, hence its initial name. It is highly expressed also in the prostate and liver, less in the placenta, ovary and adrenal glands [54].

The human NCU-G1 (hNCU-G1) protein consists of 406 amino acids, and has 6 exons [54]. In comparison to its human counterpart, mouse and rat NCU-G1 have an open reading frame (ORF) that is two amino acids shorter [52]. hNCU-G1 has a leucine and lysine after amino acids 52 and 139 respectively [54]. Bioinformatics analysis predicted that its isoelectric point is 6.1 and its molecular weight (MW) is 43.8 kDa (http://www.justbio.com/). According to the NCBI database and the Ensembl database (http://www.ensembl.org/index.html), hNCU-G1 is located in chromosome 1, as shown in figure 1-3, which is modified from the UCSC database (http://genome.ucsc.edu/index.html?org=Human&db=hg19&hgsid=203729833). The mouse orthologue is located in chromosome 3 [55].

A BLAST comparison analysis of NCU-G1 protein against other proteins shows no paralogues for NCU-G1.

Figure 1-3 The red line indicates the hNCU-G1 gene location in chromosome 1. Figure modified from UCSC.

Conflicting data has been presented concerning NCU-G1 subcellular localization. Research groups followed the prediction that the NCU-G1 protein has the same last amino acids as lysosomal proteins. It contains a transmembrane segment and a single tyrosine-based localization motif at its C-terminus as shown in figure 1-4. Subsequently, after a series of experiments they approved that NCU-1 is a lysosomal protein [53, 55, 56]. In contrast to these findings the Eskild group proposes that NCU-G1 is a nuclear protein. This was based on

bioinformatics predictions that NCU-G1 contains four nuclear receptor boxes (LXXLL). Using a polyclonal antibody, the Eskild group verified NCU-G1 nuclear localization [54].

Functions of NCU-G1 have not yet been well established. However, the Eskild group suggests two preliminary roles for this protein. Firstly, it can act as a transcription factor from the CRBP1 (cellular retinol binding protein 1) promoter through binding of the FP1 (footprint 1) DNA element. This activity requires intact exon 2 and 4. Secondly, it is also described as a nuclear receptor co-activator of PPARα (peroxisome proliferators-activated receptor α), by activating transcription from the acyl CoA-oxidase promoter, a target for the mentioned receptor. This occurs in a ligand dependent manner, which is necessary for conformational change of the AF-2 domain of the receptor. This process requires a NR1 box (nuclear receptor) [54].

Mouse NCU-G1 has nine N-glycosylation sites as shown in figure 1-4 [55], and reviewed by us using the prediction programs mentioned in (section 1.6). According to the consensus motif of N-glycosylation (N-X-S/T), the human orthologue has seven N-glycosylation sites.

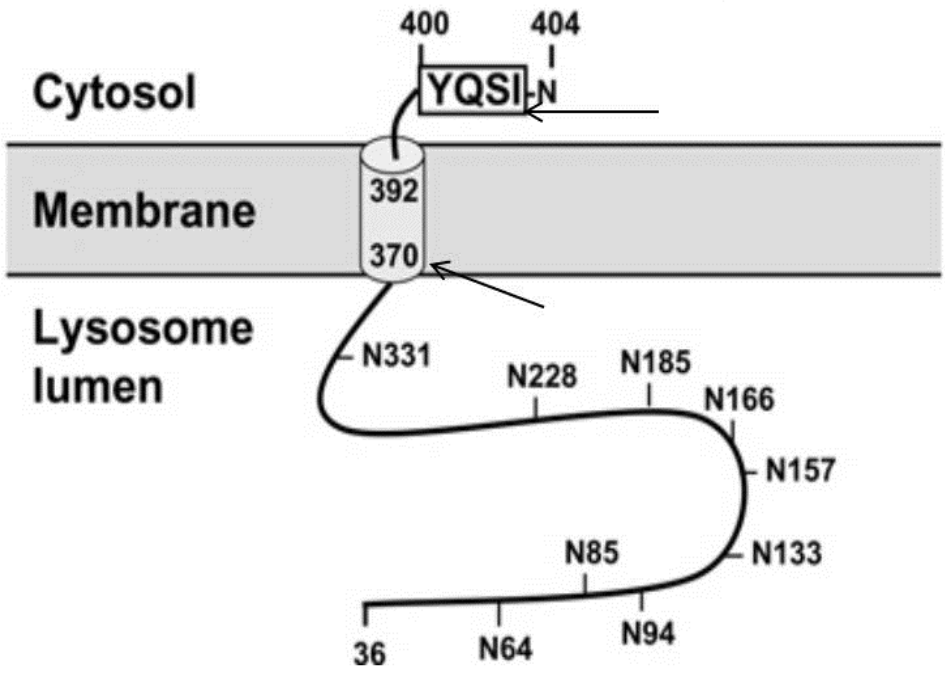


Figure 1-4 The mouse NCU-G1 predicted structure after cleaving of residue 1-35 (signal peptide). This figure shows nine glycosylation sites distributed along the sequence (indicated by a number beside N), a Tyrosine-based sorting signal (upper arrow), and a transmembrane segment (lower arrow). Figure modified from [55].

2 Materials

2.1 Chemicals

Chemicals	Range of applicability	Supplier
Agarose LE	Agarose gel electrophoresis	SeaKem
Ampicillin	LB-amp plates	Saveen Werner
Bactoagar	LB-amp plates	Sigma
Calcium chloride monohydrate	TB buffer	Sigma
DMSO	Competent cells production	Biowhittaker
GeneRuler I Kb DNA ladder	Agarose gel electrophoresis	Fermentas
Dulbecco's modified Eagle medium	Cell culture medium	GIBCO
EDTA	Lysis buffer (RIPA)	Sigma
EDTA (disodium)	Lysis buffer (RIPA)	AppliChem
Ethanol	Ethanol precipitation	Arcus
Ethedium bromide	Agarose gel electrophoresis	Biowhittaker
Fetal bovine serum	Cell culture medium preparation	Sigma
Glacial acetic acid	TAE buffer	MERCK
L-Glutamin	Cell culture medium additive	Biowhittaker
Gluteraldehyde (25 %)	X-gal staining	MERCK
Glycine	Blotting buffer	MERCK
Magnesium chloride	SOB medium, X- gal solution	Sigma
Magnesium sulfate	Competenet cells production , β-galactosidase measurement	Sigma
Manganese chloride	TB buffer	Sigma
β-mercaptoethanol	β-galactosidase measurement	Sigma
Methanol	Blotting buffer	Sigma
O-nitophenyl-β-D-galactosidase (ONPG)	β-galactosidase measurement	GIBCO
Opti-MEM	Transfection medium	Lonza

Pencillin/Streptomycin (P/S) 5000 U	Cell culture medium additive	Biowhittaker
Pipes	TB buffer	Sigma
PageRuler prestained protein ladder	SDS gel electrophoresis	Fermentas
Potassium chloride (KCl)	various	MERCK
Potassium dihydrogen phosphate	PBS	MERCK
Potassium hexacyano-ferrate(III)	X-gal staining	MERCK
Potassium hexacyano-ferrate(II) trihydrate	X-gal staining	MERCK
Reporter lysis buffer (5X)	β-galactosidase measurement	Promega
Sodium carbonate	β-galactosidase measurement	MERCK
Sodium chloride	various	Sigma
Sodium dihydrogenphosphate	Phosphate buffer, Z buffer	MERCK
di sodium hydrogen phosphate	Phosphate buffer, Z buffer	MERCK
di-sodium hydrogenphosphate-dihydrate	PBS	MERCK
Tris-Base	Blotting buffer, TAE buffer	MERCK
Triton® X-100	Lysis buffer (RIPA)	Sigma
Trypsin EDTA	Cell culture medium additive	Sigma
Tryptone	LB medium/plates, SOB medium	Oxoid
Tween 20	Western immunoblotting	Sigma
X-Gal	X-gal staining	Fermentas
XT MOPS Running buffer (20x)	Western immunoblotting	Nupage
XT Reducing agent (10x)	Western immunoblotting	Nupage
XT Sample buffer (4x)	Western immunoblotting	Nupage
Yeast extract	LB medium/plates, SOB medium	Oxoid

2.2 Other reagents

2.2.1 Antibodies

Antibody	Dilution	Supplier
Anti-mouse IgG HRP-linked antibody	1: 3000	Cell Signaling
HRP-Goat Anti-Rabbit IgG	1:4000	Invitrogen
Mouse anti-myc-tag monoclonal antibody	1:2000	Stressgen
Mouse anti-β actin monoclonal antibody	1:3000	Santa Cruz
Rabbit anti-NCU-G1 antibody	1:1000	Gift from Lubke T

2.2.2 Enzymes

Enzyme	Buffer	Supplier
DyNAzyme™ II DNA Polymerase	DyNAzyme™ Buffer	Finnzymes
KpnI	KpnI buffer	Fermentas
PNGase F	Denaturing buffer+ G7 reaction buffer+ NP-40	New England Biolabs
SmaI	Tango buffer	Fermentas
SspI	Buffer G	Fermentas
T4 DNA Ligase	Ligase buffer	New England Biolabs
XhoI	Buffer R	Fermentas
XmaI	Buffer 4 and BSA	New England Biolabs

2.2.3 Kits

Kit name	Supplier
BioRad protein assay (Bradford)	BioRad
Criterion XT Bis-Tris Gel, 10 %	BioRad
ECL plus Western Blotting Detection System	Amersham
FuGENE® 6 Transfection Reagent	Roche

GFX PCR DNA and Gel Band Purification Kit	GE Healthcare
NuPAGE® Novex® Tris-Acetate Mini Gels, 4-12 %	Invitrogen
Plasmid DNA Purification	Macherey Nagel
QuikChange® Lightning Site-Directed Mutagenesis Kit	Stratagene
QuikChange® Site-Directed Mutagenesis Kit	Stratagene
Wizard® Plus SV Minipreps DNA Purification System	Promega
X-tremeGENE 9 DNA Transfection Reagent	Roche

2.3 Equipments

Laboratory equipment	**Supplier**
Adjustable Reciprocating Orbital Shaker (AROS 160)	Thermolyne
Avanti J-25 centrifuge	Bechman
Biofuge 13	Heraeus
Blotting paper (3MM)	Whatman
Cell scraper	TPP
Counting chamber	Bürker
C24 incubator shaker	New Brunswick scientific
Electrophoresis power supply (EPS 600)	Pharmacia Biotech
Flat bottom microtest III plates (96 wells)	ArcSoft
Image station 4000R pro	Kodak
MicOcular 3.0 MP	Becton Dickinson
Model 25 incubator shaker	New Brunswick scientific
Nanodrop 2000	Thermo Scientific
Petri dish	Sarstedt
Tissue culture flask (75 cm^2)	Sarstedt
Tissue culture plate (6 well) flat bottom with lid	Sarstedt
Water bath	Grant
Western blotting apparatus	BioRad

3 Methods

3.1 Glycosylation prediction

Most of the programs mentioned in the introduction chapter (section 1.5) were used to predict NCU-G1 glycosylation sites from the amino acid sequence. The easiest programs and which gave the same results for NCU-G1 glycosylation sites compared to previous studies are NetNglyc and NetOglyc for N and O glycosylation prediction respectively.

3.2 Molecular biological techniques

3.2.1 DNA measurement using Nanodrop

DNA concentration was measured using The NanoDrop 2000 Spectrophotometer:

- 2 µl of sample was applied to the instrument and analyzed.
- 260/280 ratio was checked to assess the purity of DNA, a ratio of ~1.8 is generally considered as "pure".
- 260/230 ratio was also checked as a secondary measure of purity, pure DNA ratio should be within the range of 1.8-2.2.

3.2.2 Polymerase Chain Reaction (PCR)

PCR is a powerful technique by which millions of DNA copies may be produced from a DNA molecule. Primers are oligonucleotides that are designed to bind and flank the region going to be amplified. These primers are mixed in excess amount with the DNA template. Polymerase enzyme is added to synthesize the complementary DNA strand from nucleotides. A buffer is added to adjust the reaction condition. Finally, a cycling program is setup to obtain three main phases: denaturation (separation of DNA double strands), annealing (primers binding) and extension, as shown in (figure 3-1). The number of cycles used determines the amount of DNA copies produced and can be calculated from 2^n, where n is the number of cycles [57].

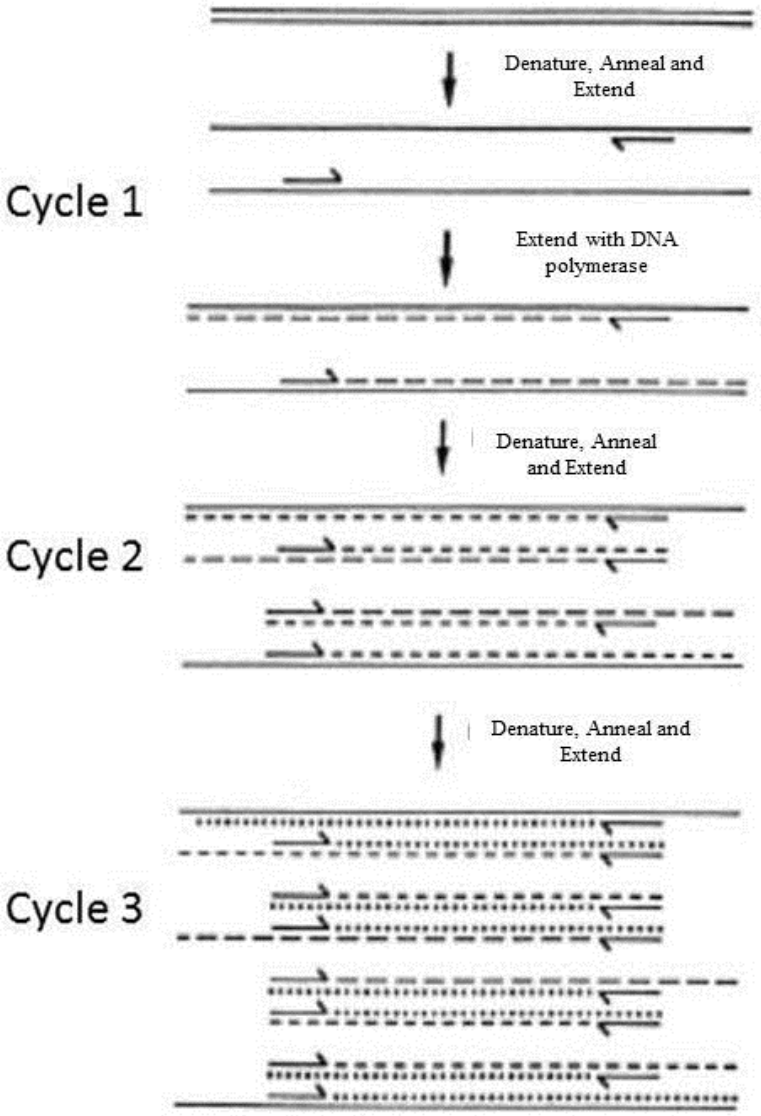


Figure 3-1 Schematic representation of the PCR technique, showing the 3 main phases: denaturation, annealing, and elongation. Figure modified from [57].

Procedure:

- On a 1.5 ml tube pre-inserted on ice mix the following components, (note: calculate for the number of PCR reactions required, then aliquot into 0.2 ml tubes):
  - 5 µl buffer.
  - 2 µl forward primer (5 pmol/µl).
  - 2 µl reverse primer (5 pmol/µl).
  - 1 µl dNTPs (10mM).
  - 1 µl template.
  - 1 µl polymerase enzyme.
  - Fill to 50 µl with nuclease free water.

- PCR cycling:

Temprature	Time	Cycles
94	4 min	1
94	30 sec	20
65	30 sec	
72	2 min	
72	7 min	1

Table 3-1 PCR cycling for NCU-G1

3.2.3 Restriction analysis of DNA

To treat DNA with more than one enzyme we need to use more than one buffer. As each enzyme has maximum digestion activity in its specific buffer we should remove the old buffer. Here we used ethanol precipitation technique to clear old buffer before digestion with new enzyme. After ethanol precipitation dried pellet was dissolved well in the new enzyme buffer prior to enzyme addition.

Kpn I digestion (Fermentas):

- Take x µg DNA.
- KpnI buffer to the final concentration of 1x.
- 20u Kpn I enzyme.
- Incubate for 2 hr in 37^{o}C water-bath.

XhoI digestion (Fermentas):

- Take x µg DNA.
- R buffer to the final concentration of 1x.
- 20u XhoI enzyme.
- Incubate for 2 hr in 37^{o}C water-bath.

SmaI digestion (Fermentas):

- Take x µg from the DNA.
- Tango buffer to the final concentration of 1x.
- 20u SmaI enzyme.
- Incubate for 2 hr in 30^{o}C water-bath.

XmaI digestion (NEB):

- Take x µg DNA.
- Buffer 4 to the final concentration of 1x.
- BSA to the final concentration of 1x.
- 20u XmaI enzyme.
- Incubate for 2 hr in 37^{o}C water-bath.

SspI digestion (Fermentas):

- Take x µg DNA.
- G buffer to the final concentration of 1x.
- 20u SspI enzyme.
- Incubate for 2 hr in 37^{o}C water-bath.

Ethanol precipitation:

- Take x µl from sample treated with the first enzyme.
- 0.1x µl 3M Na acetate PH 5.0.
- 2.5x µl room-temperature (RT) ethanol.
- Incubate at -20 overnight or at -80 for 1 hr.
- Centrifugate at 4^{o}C for 30 min at 13000 RPM.
- Discard supernatant and add 500 µl 70% cold ethanol.
- Centrifugate at 4^{o}C for 10 min at 13000 RPM.
- Remove supernatant and dry pellet.

Notes:

- Total digestion volume should be adjusted so that DNA concentration should not exceed 0.33g/µl. Also glycerol (a component in enzyme storage buffer) should not exceed 5%.
- Enzyme buffer with the less salt concentration should be use before enzyme buffer with the higher salt concentration.

3.2.4 Agarose gel electrophoresis

A method used for separation and identification of DNA fragments according to size. DNA as a negatively charged molecule moves through agarose matrix from cathode (black) to anode (red). Shorter molecules move faster and migrate further in the gel than longer ones [58].

Procedure:

- Preparation of 1% agarose gel, 50 ml:
  - Mix 0.5 g agarose with 50 ml 1x TAE buffer.
  - Boil up the mixture for 7 min and cool to 60^{o}C.
  - Add 0.2 µg/ml ethidium bromide.
  - Place the comb in the gel casting tray, pour the gel and allow polymerize for 30 min.
  - Carefully pull out the comb and place the gel in the electrophoresis chamber. Then add sufficient 1x TAE buffer to cover the depth of ~ 1cm.
- Preparation of sample:
  - Add loading buffer to the final concentration of 1x to the sample and then load to the gel wells.
  - Load x µl 1 Kb DNA marker to a well to enable band size comparison.
- Gel run:
  - Adjust volt (V) to 80 V.
  - Let gel run until the bromophenol blue dye has migrated two-third the length of the gel (~ 90 min).

3.2.5 Isolation and elution of DNA from agarose gel

Here we used Purification of DNA from gel bands protocol of the "GFX PCR DNA and Gel Band Purification Kit". This is a rapid technique in which gel fragment is dissolved in capture buffer at 60^{o}C, and applied to spin-columns that have great affinity to DNA. DNA is then washed with a washing buffer to remove contaminants. Purified DNA is then eluted in low salt concentration buffer.

3.2.6 Ligation of DNA fragments

Ligase is an enzyme that catalyzes the covalent formation of phosphodiester bonds between phosphate groups and hydroxyl groups. This process requires the hydrolysis of ATP or other energy molecule e.g. NAD^{+}/NADH. According to this principle ligase joins 5' phosphate group of DNA with the 3' hydroxyl group.

The most commonly used DNA ligase is purified from T4 bacteriophage and so called T4 DNA ligase.

Here we used rapid ligation kit from New England Biolabs (NEB).

Procedure:

- Mix vector DNA with the insert DNA in the following ratio:

	1:5	1:10
Vector	20 ng	20ng
Insert	20*y*5 ng	20*y*10 ng (y= insert basepairs/vector basebairs)

- Add T4 DNA ligase buffer to the final concentration of 1x.
- Add 1 µl T4 DNA ligase enzyme.
- Incubate overnight at RT.
- Ligation mixture can then be used in transformation reactions or stored at -20 for further usage.

3.2.7 Site directed mutagenesis

QuikChange® Lightning Site-Directed Mutagenesis Kit was used to generate substitutions at various nucleotide positions on hNCU-G1 sequence.

The principle of this method depends on generation of mutation on a DNA fragment pre-inserted in a plasmid vector. This is done by PCR amplification of the DNA using mutagenic primers containing the desired mutation. The reaction is catalyzed by the highly effective *PfuUltra* HF DNA polymerase. This is followed by removal of the precursor DNA using Dpn I endonuclease, and transformation into XL10-Gold ultra-competent cells.

Blue white screening system was used to determine mutagenesis efficiency. This method depends on that lacZ is the coding gene for β-galactosidase enzyme which has the ability to breakdown media 5-bromo-4-chloro-3-indolyl- β-D- galactoside (X-gal) yielding blue coloured colonies. When this gene is interrupted with DNA insertion, β-galactosidase will not form and so media X-gal will remain intact and the resulting colonies will appear white instead [59].

The pWhitescript™ plasmid was used for mutagenesis control. This vector contains the stop codon (TAA) at position nine of the amino acid sequence of β-galactosidase. This codon is changed to (CAA) when using control primers supplied with the kit, giving blue colonies on LB plates containing X-gal and IPTG. For transformation control, pcDNA vector containing an insert was used.

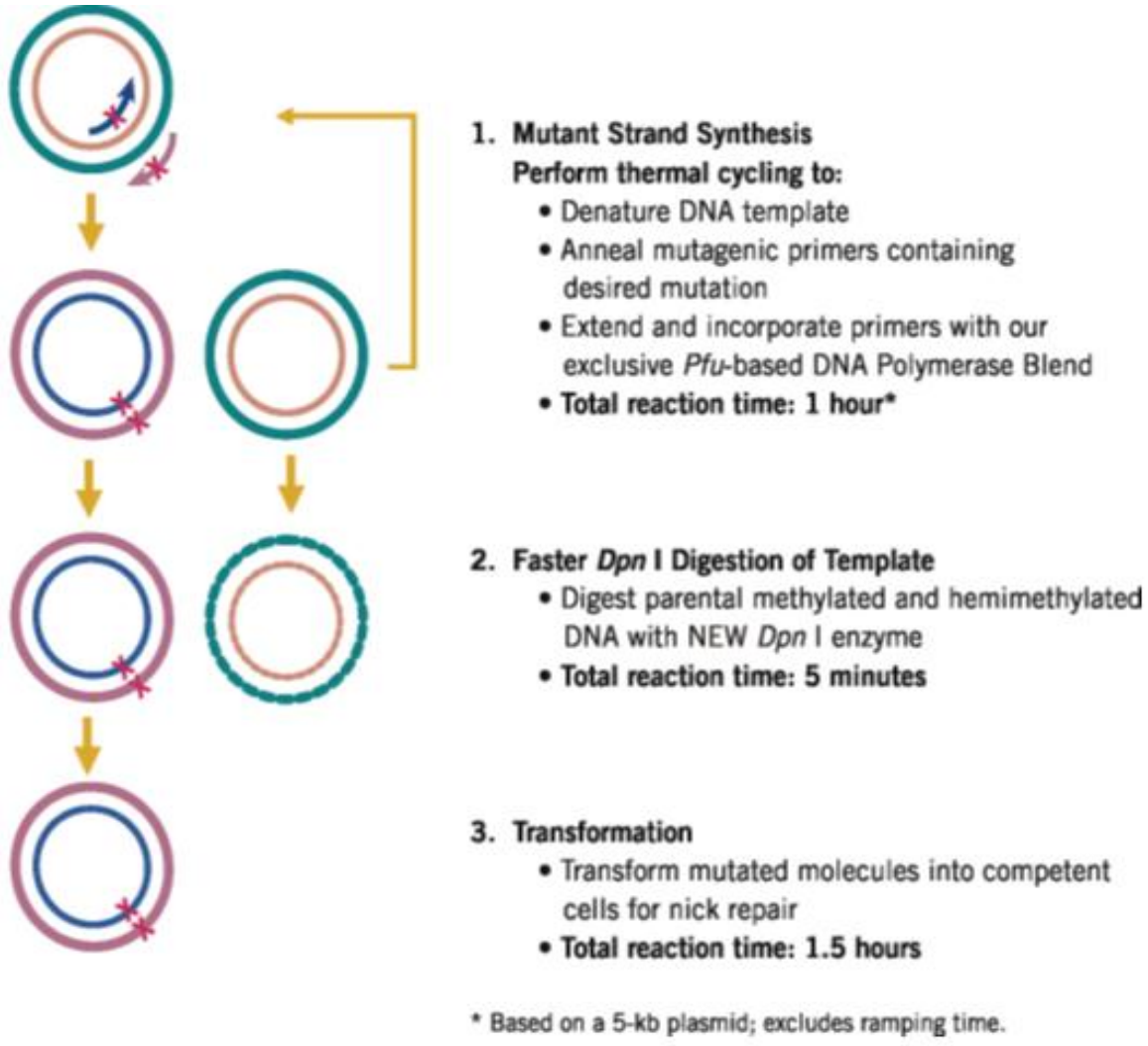


Figure 3-2 Fast guide for QuikChange® mutagenesis procedure, modified from the kit manual.

Protocol:

- Mutagenic primers construction:
  - Both of the mutagenic primers must contain the mutation and bind to the same sequence on the two strands of the plasmid.
  - Primers length should be between 25 and 45 bases.
  - Melting temperature (Tm) should be $\geq 78^{o}C$.

$$Tm=81.5+0.41(GC\%)-675/N - \%mismatch$$

Above is the formula used for estimating the *T*m of primers. Where N is the primer length in bases, values for GC% and % mismatch are whole numbers.

- PCR reactions were prepared and cycled at the same run because the control vector (pWhitescript) is 4.5 kb, and the template vector (pBlueskript II SK (+) 3kb and the insert 1.2) is 4.2 kb.
  - Reaction preparation:

Components	**Control/µl**	**Sample/µl**
10x Reaction buffer	5	5
Quick solution reagent	1.5	1.5
dNTPs mix	1	1
PfuUltra HF DNA polymerase	1	1
Double-distilled water (ddH_2O) to a final volume of 50 µl	34	37
(125 ng) Oligonucleotide primer #1/ 2 (100 ng)	1.25 each	
(125 ng) Mutagenic forward and reverse primers (50 ng)		2.5 each
(25 ng) of pWhitescript 4.5-kb control plasmid (5 ng/µl)	5	
(50 ng) Template DNA (25 ng/µl)		2

Table 3-2 Mutagenesis reaction preparation.

  - PCR cycling:

Temperature	**Time**	**Cycles**
95	30 sec	1
95	30 sec	16
55	1 min	
68	2.5 min	
68	7 min	1

Table 3-3 Mutagenesis PCR cycling

- Amplification Products were then digested with *Dpn* I, and checked by electrophoresis.
- PCR Products were then transformed into XL10-Blue Supercompetent cells, S.O.C broth was used for this purpose.

3.2.8 Sequencing

Samples were sequenced using the Sanger dideoxy termination sequencing method. In this method DNA sample is incubated with polymerase enzyme, primers and dNTPs. Four polymerase reactions are carried out. Each reaction also contains one of the four dideoxy NTPs. When a dideoxyNTP is added, chain extension terminates because ddNTP nucleotides lack 3' hydroxyl groups, thus each reaction results in fragments terminating at that base. The four reactions result in four clusters of fragments with lengths representing the positions of each of the four respective bases. These fragments are then separated by denaturing acrylamide gel [60]. Sequencing was mainly performed at the ABI lab in Oslo University (http://www.mn.uio.no/bio/forskning/om/infrastruktur/abi-lab/). NCU-G1 N- and C -termini sequencing was done by the help of the GATC Biotech in Germany (http://www.gatc-biotech.com/en/index.html).

3.3 Prokaryotic methods

3.3.1 Production of competent DH5α cells (E. coli)

- Mix old competent cells with 100 ml SOB-medium, incubate at 37^{o}C/ 200-250 RPM for around 3 hours (hr).

  Note: all steps should be done as sterile as possible.
- Measure OD_{600} and calculate how much of the culture is needed to give OD_{600} = 0.05 in 250 ml SOB-medium. Incubate 250 ml culture (SOB-medium added to the culture to give final OD_{600} = 0.05), and let it stay overnight (16-18 hr) at 18^{o}C.
- Measure OD_{600}, it should be 0.3-0.6.
- Incubate cells 10 min in ice.
- Centrifugate for 10 min, 2500x g, then discard supernatant.
- Resuspend pellet in 80 ml TB-buffer (note: this is the total volume, therefore if cells were divided to four tubes, each tube should receive 20 ml. Put them all in two tubes).
- Incubate 10 min in ice, then centrifugate as in point 5.
- Resuspend pellet in 20 ml TB-buffer (this is for total volume, so for two tubes 10 ml in each).
- Add 350 µl DMSO to each tube, mix well and let it stay for 5 min in ice.
- Add 350 µl DMSO, mix well and incubate 10 min in ice.

- Aliquot cells each 200 µl in cold Eppendorf tube, and shock freeze cells in liquid Nitrogen.

The newly produced competent cells should be controlled:

- Cultivate cells in LB-plates containing amp. Cells should not grow in the presence of amp.
- Competence level should be determined:
  - Transform cells with plasmid containing amp resistance gene (1, 0.1, 0.01 ng), and inoculate them in LB-plate containing amp.
  - Count produced colonies.
  - Calculate how many cells should be produced if transformed with 1 µg. This should be between 10^6-10^9, and it reflects competency.

3.3.2 Transformation

DNA can be introduced into bacterial cells by a method called transformation. There are several methods by which bacteria can take up DNA, as examples electroporation and heat shock. Certain bacteria are able to take up external DNA, these are called competent cells. To do transformation, DNA should be inserted in a plasmid, the resulting recombinant plasmid can be taken up by competent cells under specific conditions. This plasmid is then replicated using the bacterial expression system. High copy numbers of the plasmid including the previously inserted DNA may be produced. As not all cells can succeed to take up the plasmid, a selective marker should be included in the plasmid e.g. antibiotic resistant gene. This marker will apply positive selection pressure when including the specific antibiotic in the growth medium [61].

Procedure:

- Thaw the competent cells on ice.
- In Eppendorf tube take:
  - 2 µl ligation mixture.
  - 50 µl competent cells.
- Incubate on ice 30 min..
- Heat shock cells in 42^{o}C for 30 sec.
- Return cells to ice and leave for 2 min.

- Add 450 µl LB medium without antibiotic, incubate in 37^{O} C water-bath for 30-60 min. This can be used to:
  - Inoculate LB plates for colonies selection.
  - Or transfer cells to LB-medium containing amp for DNA amplification.
- Incubate overnight at 37^{o}C.

3.3.3 Isolation of plasmid DNA from bacterial culture

Many methods are available to isolate DNA plasmid from bacterial culture. Here we used two kits. For transformation product that has been amplified in a large medium volume we used NucleoBond® Xtra Midi/ Maxi protocol of plasmid DNA purification kit from Macherey Nagel. This method uses NaOH / SDS to lyse bacterial cells, the DNA is then captured using specially designed column filters, after a proper washing step the plasmid DNA is then eluted, precipitated, and dissolved in a suitable buffer for further use.
For small culture medium volume we used centrifugation protocol of miniprep kit from Promega.

3.4 Eukaryotic cell techniques

All treatments of cells were done in sterile hood, with sterile equipment and solutions. The cells which will be used here are Hela cells.

3.4.1 Growing of cells

All solutions were pre-warmed to 37^{o}C in water-bath, and then the outer walls of all bottles were sprayed with 70% ethanol before transfer to the cell hood. Cells were subcultured when they were about 80% confluent. Part of the cell suspension was transferred to new $75cm^2$ flasks from Sarstedt, containing 20-25 ml medium.
Cells were incubated at 37^{o}C in a highly humidified atmosphere with 5% CO_2. CO_2 is used to keep the medium pH within the physiological pH 7.0-7.5. CO_2 and H_2O form equilibrium with HCO_3^-. This equilibrium regulates the pH.

$$CO_2+H_2O \leftrightarrow HCO_3^-+H^+$$

3.4.2 Cell subculturing

- Remove old medium and wash the cells with 10 ml PBS (37^{o}C), then remove PBS.
- Add 2.5 ml Trypsin, and leave at RT for 1 min, remove most of the trypsin, and incubate for 2 min at 37^{o}C.
- Shake the bottle well until cells float.
- Resuspend the cells in 10 ml of media.
- Transfer the required volume to new culture flasks containing 20-25 ml fresh medium e.g. 1:5 (2 ml) 1:10 (1 ml). Return flasks to the incubator for a time that is required to achieve 80% confluency before next subculturing.

3.4.3 Cells counting

Before experiment, cells were counted using a Bürker chamber. A drop of well mixed cells from cells resuspened in 10 ml media (section 3.4.2 point 4) is applied to the chamber and covered with a cover-glass. Under microscope a person will see a typical picture to (figure3-3). 2x12 squares were counted and total cells were estimated from the following formula:

Average of diagonals x 2.1 /100 = X x 10^{6} cells /ml

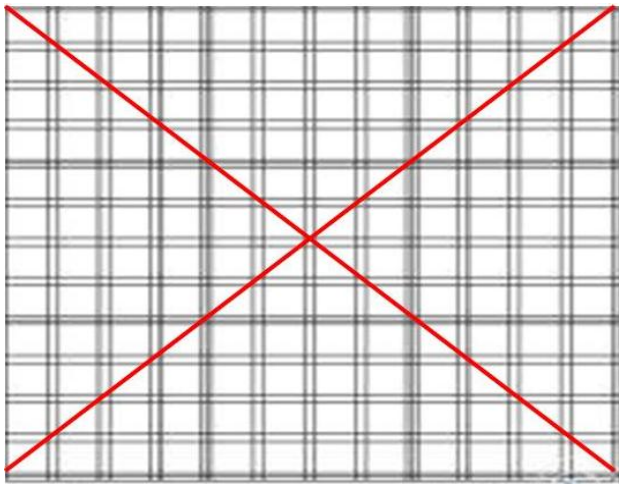


Figure 3-3 Bürker counting chamber. Figure modified from Wenk Lab Tech.

3.4.4 Transfection

Transfection is the technique by which a plasmid is taken up by cultured eukaryotic cells. A plasmid is prepared to carry an origin of replication extracted from mammalian virus.
The gene of interest is inserted into the plasmid. When this plasmid enters the cell, the virus origin of replication facilitates its replication producing multiple copies of the plasmid and the previously inserted DNA fragment. This process can be stable or transient depending on the integration between the plasmid and the transfected cell genome. Here we used transient transfection. Numerous methods available to transfect cells depending on specific general principles, most of them apply the same general notion of transformation technique mentioned in (section 3.3.2). Here we used cationic lipid transfection technique; lipid is prepared to be cationic for two reasons, to allow binding to negatively charge nucleic acids, and to facilitate entry of the lipid DNA complex through negatively hydrophobic cell membrane of eukaryotic cell [62, 63]. Here we used X-tremeGENE 9 kit from Roche. Hela cells were transiently transfected with X-tremeGENE reagent for 24 hr.
Procedure:

- One day before transfection Hela cells were subcultured into 6 well plates.
- The next day cells were transfected using 3:1 ratio (3 µl of transfection reagent to 1 µg of DNA) all steps mentioned by the producer were followed exactly.
- After 24 hr cells were treated for protein expression analysis.

Transfection control:
To verify transfection of Hela cells with X-tremeGENE 9, and for optimization of cell density that will be used for western immunoblotting, cells were transfected with a vector containing lacZ gene. LacZ expression was measured (β-galactosidase assay***)***. LacZ function was analyzed (x-gal assay). How lacZ gene reacts with x-gal was described (mutation section: 3.2.7).

3.4.5 Harvesting of cells for β-galactosidase analysis

- Remove all media.
- Wash cells carefully with 1x PBS.
- Remove the PBS (all of it).
- Add 300µl 1x lysis buffer to each well.
- Let stay at 4^{o}C for 10 min.

- Scrape into Eppendorf tubes.
- Centrifugate at 6000 RPM for 4 min.
- Transfer supernatant to new tube and store at -20 until measurement time.

3.4.6 β-galactosidase assay

- Mix 50µl cell extract with 150µl Z-buffer.
- Start reaction with 50µl ONPG-reaction buffer.
- Incubate at 30^{o}C for 10-60 min.
- Stop reaction with 100 µl 1M Na_2CO_3 when yellow color appears, this color is stable at 4^{o}C.
- Calculate β-galactosidase production per hr per well from the following formula:

$$\frac{A_{240\ nm}}{\text{Reaction time (hr) x Reaction volume per well}}$$

3.4.7 X-gal staining

- Remove media from cells 24 hr after transfection, then wash twice with 1x PBS.
- Fix cells by incubation with 2 ml per well of freshly made gluteraldehyde solution, incubate 15 min at RT.
- Remove gluteraldehyde, and wash twice with 1x PBS.
- Add 1 ml x-gal solution for each well, and incubate at 30^{o}C for 4 hr.
- Remove x-gal solution, and take picture of cells. Cells that are transfected and express lacZ gene will appear blue/green under microscope.

3.5 Protein analysis

Proteins were extracted from Hela cells, and measured. Proteins were then separated by western immunoblotting technique before and after the treatment with PNGase F enzyme.

3.5.1 Protein extraction

- Wash cells twice with 1x PBS (2 ml each well).
- Add 300µL RIPA lysis buffer, and 3µl of (100x) protease inhibitor. Incubate 5-10 min at RT.

- Scrape cells and transfer the lysate to an Eppendorf tube. Mix well up and down using small syringe.
- Centrifugate cells at 4^{o}C, 6000 RPM, for 5 min.
- Transfer the supernatant to a new tube.

3.5.2 Protein measurement

We used Bradford protein assay from BioRad to measure total protein. The principle of this method depends on the addition of the red acidic dye Coomassie brilliant blue to the proteins, this dye binds basic and aromatic amino acids mainly arginine yielding blue color, which can be measured spectrophotometrically. A set of serial dilutions of bovine serum albumin (BSA) standards were measured. A standard curve was constructed from concentrations against optical densities. Samples concentrations were determinated from the standard curve.

Procedure:

- Dilute standards of serum albumin in water (5, 10, 15, 20 ng/ml), in addition to the zero point (the blank).
- Dilute samples (1/400-1/800).
- Add 100 µl dye reagent to 400µl of the standards, samples or water. Mix well immediately.
- Read within 30 min with spectrophotometer against 595 nm.

3.5.3 Western immunoblotting

Western immunoblotting was used to determine the MW of proteins. Proteins were first separated according to their size by electrophoresis (SDS PAGE), then transferred to a Polyvinylidene Fluoride (PVDF) membrane by electroblotting, where proteins were detected with antibody against the desired protein. The primary antibody binds the epitope of the desired protein. The secondary antibody binds the FC region of the primary antibody, and is bound to a peroxidase enzyme. This enzyme breaks down H_2O_2 producing an end product that can be visualized with different techniques.

SDS Electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) is a method for protein separation according to their size. SDS denatures proteins, so they return to their basic

primary structure. SDS as an ionic detergent covers proteins with a negative charge. So, proteins separation using this method will depend only on size, and small proteins will move faster in the gel than large ones.

Here we used NuPAGE® Novex® Tris-Acetate Mini Gels (4-12% gradient) from Invitrogen, and Criterion XT precast gels 10% from BioRad.

Protocol:

- Add XT sample buffer 4x and XT reducing agent 20/10x to 1x concentration to the desired protein amount.
- Heat sample at 95^{o}C for 5 min.
- Add 1x XT MOPS running buffer to the electrophoresis tank.
- Load samples and standard to the gel.
- Adjust the run at 200 V, 165-175 mA/gel for around 60 min.

Wet blot

- Remove the gel from the glass plates and incubate in cold blotting buffer for 15 min to remove SDS.
- Wet two pieces of fiber pads, two pieces of Whatman papers with cold blotting buffer.
- Soak a PVDF membrane in methanol, then wash well with cold blotting buffer.
- Assemble the transfer sandwich as follow:

White part of the blotting cassette

Fiber pad

Whatman paper

PVDF membrane

Gel

Whatman paper

Fiber pad

Black part of the blotting cassette

- Remove any air bubbles between the sandwich layers.
- Put the sandwich in the electroblotting tank from positive to negative (the black cassette to the black site of the tank, and the white cassette to the light site of the tank).
- Connect the power supply, and electrophoretically transfer proteins from the gel to the PVDF membrane at 20 V overnight at 4^{o}C. Stir the buffer by a magnet-mixer.

Detection

- Blocking: Block non-specific binding by incubating the membrane for 1 hr at RT with agitation in Blocking Buffer (5% dry milk PBS/T).
- Primary antibody: Remove blocking solution. Incubate the membrane for 2 hr with a primary antibody diluted in blocking buffer with slow agitation.

  Primary antibodies used:

  - Mouse anti-β-Actin antibody. Dilution 1: 3000.
  - Mouse anti-myc antibody. Dilution 1: 2000.
  - Rabbit anti-NCU-G1 antibody. Dilution 1: 1000.
- Washing: Rinse the membrane with PBS/T with rapid agitation, wash 3 times 5 min, then 3 times 15 min.
- Secondary antibody: Incubate the membrane for 2 hr at RT with secondary HRP-linked antibody diluted in blocking buffer.

  Secondary antibodies used:

  - Anti-mouse IgG antibody. Dilution 1:3000.
  - Anti-rabbit IgG antibody. Dilution 1: 4000.
- Washing: Rinse membrane with 1x PBS/T, 3 times 5 min and 2 times 15 min. Then wash 1 time for 15 min with 1x PBS and keep in 1x PBS until visualization.

Visualization

Here we used ECL plus Western Blotting Detection system from Amersham. The secondary antibody used here is conjugated to horseradish peroxidase enzyme (HRP). This enzyme breaks down peroxide producing O_2, the later oxidizes the kit substrate (lumigen) producing chemilluminescence.

- Mix detection solutions A and B (40:1), protect from light if reagents will not be use immediately.
- Drain the excess wash buffer from the washed membrane.
- Put the membrane with the protein side up on a glass plate.
- Pipette the mixed detection reagent on to the membrane, and incubate for 5 min.
- Drain off excess detection reagent.
- Take picture with the ‘‘Kodak Image Station 4000 R pro’’ camera and ‘‘Carestream Molecular Imaging Software’’.

3.5.4 Membrane stripping

The purpose of this technique is to remove antibodies which are bound to the proteins in the membrane and treat with other antibodies. Membrane is treated with 50 ml stripping-buffer at 50^{o}C for 30 min. Then washed well with 1x PBS/T (2x5 min, then 2x15 min). Now the membrane is ready for blocking and incubation with new antibodies.

After stripping the membrane was blocked and treated with Anti-β-Actin antibody, then with Anti-IgG antibody, following the same steps described in (section 3.5.3).

3.5.5 Deglycosylation with PNGase F enzyme

PNGase F (Peptide N-glycosidase F) is an endoglycosidase enzyme which cleaves the linkage between asparagine and N-acetylglucosamines. Treatment with PNGase enables researchers to study the N-linked oligosaccharides of proteins.

Protocol:

- Denature proteins by adding 10x denaturing buffer to the final concentration of 1x, and incubate 10 min at 100^{o}C.
- Add 10x G7 reaction buffer to the final concentration of 1x.
- Add 10% NP-40 to the final concentration of 1%.
- Add 1 µl PNGase F enzyme (500.000 u/ml).
- Incubate 1 hr at 37^{o}C water-bath.

3.6 Immunostaining work

This part was done by my colleague Martine. Hela cells were grown on cover-glasses, and subsequently transfected with a myc-tagged wild type NCU-G1 and the two different myc-tagged mutants constructs. Cells were then fixed with paraformaldehyde, and permeabilized with Triton X-100. The myc-tagged proteins were captured with an anti-myc primary antibody. An Alexa Fluor 488 (green)-coupled secondary antibody was used. Nuclear DNA was stained with Hoechst. Confocal microscopy was used to detect immunofluorescence.

4 Results

4.1 Prediction of hNCU-G1 potential glycosylation sites

Literature presented NCU-G1 as a highly glycosylated protein. N-X-S/T is the main acceptor motif for the oligosaccharide to form N-glycosylation linkage (section 1.4.1.). Accordingly, hNCU-G1 amino acid sequence contains seven N-glycosylation acceptor sites as illustrated in figure 4-1.

MRGSVECTWGWGHCAPSPLLLWTLLLFAAPFGLLGEKTRQVSLEVIPNWLGPLQNL
LHIRAVGT**NST**LHYVWSSLGPLAVVMVATNTPHSTLSV**NWS**LLLSPEPDGGLMVLP
KDSIQFSSALVFTRLLEFDST**NVS**DTAAKPLGRPYPPYSLADFSWN**NIT**DSLDPATLS
ATFQGHPMNDPTRTFA**NGS**LAFRVQAFSRSSRPAQPPRLLHTADTCQLEVALIGASP
RG**NRS**LFGLEVATLGQGPDCPSMQEQHSIDDEYAPAVFQLDQLLWGSLPSGFAQWR
PVAYSQKPGGRESALPCQASPLHPALAYSLPQSPIVRAFFGSQNNFCAF**NLT**FGAST
GPGYWDQHYLSWSMLLGVGFPPVDGLSPLVLGIMAVALGAPGLMLLGGGLVLLLH
HKKYSEYQSIN

Figure 4-1 The predicted N-glycosylation sites on hNCU-G1 amino acid sequence, having the sequence motif that marked as follow **N**-**X**-**T**/**S**

NetNGlyc 1.0 Server was the main prediction program used to scan the hNCU-G1 protein sequence for potential N-glycosylation sites. This program found four potential sites as indicated by red crosses in table 4-1.

Interestingly, most of the web sites that offer N-glycosylation prediction use NetNGlyc e.g. Gene Infinity (http://www.geneinfinity.org/sp/sp_proteinptmodifs.html).

Ensembles of Support Vector Machine Classifiers (EnsembleGly) was used as a secondary check for the potential N-glycosylation sites detected by NetNGlyc. In addition to the four sites recognized by NetNGlyc, two more sites characterized as potential N-glycosylation sites by EnsembleGly. Those sites at positions 187 and 333 of the hNCU-G1 amino acid sequence.

The Uniprot protein knowledgebase (uniprotKB), a protein annotation database which extracts information from literature, adds position 187 to the potential N-glycosylation sites recognized by NetNGlyc.

Amino acid position	Sequence motif	Potentiality
65	N-S-T	++
95	N-W-S	+
134	N-V-S	++
159	N-I-T	++
187	N-G-S	+
230	N-R-S	++
333	N-L-T	-

Table 4-1 The potential N-glycosylation sites on hNCU-G1 amino acid sequence, predicted by NetNglyc. Red crosses indicate the potential sites.

NetOGlyc 3.1 Server (the twin for NetNGlyc database) together with EnsembleGly conclude that hNCU-G1 amino acid sequence is not modified with O-glycosylation.

We chose to study and characterize two of the potential N-glycosylation sites shared by NetNGlyc, EnsembleGly, and other databases. Those sites at positions 65 and 230 of the hNCU-G1 amino acid sequence.

4.2 Mutation of hNCU-G1 inserted in pcDNA3.1(+)/myc-His A vector

The starting material was hNCU-G1 protein coding sequence pre-inserted in a pcDNA 3.1 (+)/ myc-His A expression vector. We used QuikChange® Site-Directed Mutagenesis Kit to carry out nucleotide substitutions as follow:

Mutation position on hNCU-G1 amino acid sequence	Nucleotide triplet change	Nucleotide substitution
65 N ⟶ A	AAT ⟶ GCT	AA ⟶ GC
230 N ⟶ A	AAC ⟶ GCC	AA ⟶ GC

Table 4-2 Mutation construction. At amino acid level Asparagine amino acid (N) was changed to Alanine (A), and at nucleotide level AA was changed to GC on both mutations.

We designed primers and did the first mutation at position 65 of the hNCU-G1 amino acid sequence. The control vector got the mutation perfectly i.e. according to the producer more than 80% of the colonies were positive. On the other hand, no colonies were seen in the test sample, indicating unsuccessful mutation. This may be because pcDNA as a large vector (5.5 kb) is not easy to mutate. But also due to the fact that the kit we used has limited efficiency

towards large constructs, as it is difficult to mutate constructs more than 8 kb in size. Here our vector carrying the insert is 6.7 kb that is somewhat near the kit efficiency borderline.

4.3 Vector construction (1)

To facilitate the mutation, pBluescript II SK (+) plasmid was chosen because of its relative small size (3kb), and it contains amp resistance gene.

NCU-G1 was excised from pcDNA 3.1 (+)/ myc-His A vector using KpnI and XhoI restriction enzymes. These are the same enzymes by which NCU-G1 was previously inserted into this vector. pBluescript II SK (+) vector was cut by the same enzymes. Agarose gel picture in figure 4-2 shows the result of this cutting. The lower two bands in lanes 2 and 3 are NCU-G1 bands (Insert) 1.2 kb. And the bands on lane 4 and 5 are pBluescript bands (vector).

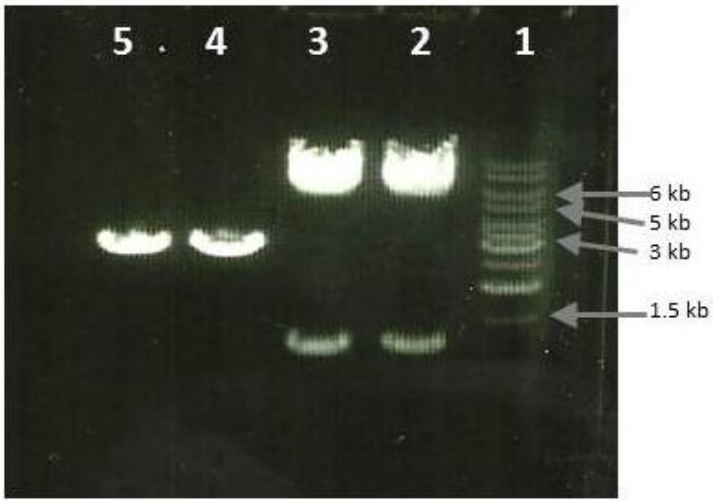


Figure 4-2 Agarose gel electrophoresis of KpnI and XhoI cutting of pcDNA 3.1 (+)/ myc His A containing NCU-G1, and pBluescript II SK (+) vectors. Lane 1, 1 kb DNA marker. Lanes 2 and 3 show the cutting result of pcDNA–NCU-G1 construct. These give two bands: around 5.5 kb is the pcDNA band, and around 1.2 kb is the NCU-G1 band. Lanes 4 and 5 are pBluescript cutting outcome, giving one band around 3 kb.

Vector and insert bands were excised from the gel and purified using the GE purification kit (section 3.2.5). Resulting DNA was then measured to calculate the volume we need from the vector and insert to perform ligation as described in (section 3.2.6). Ligation mixture was transformed into *E.coli* bacterial cells and plated on LB plates containing amp. Theoretically, only bacteria that carry the plasmid should grow in the presence of amp. For confirmation, PCR reaction for NCU-G1 for each individual colony was carried out. Three colonies gave positive PCR reactions (figure 4-3). We amplified these colonies on 5 ml LB medium containing amp, and then purified them using the Miniprep kit.

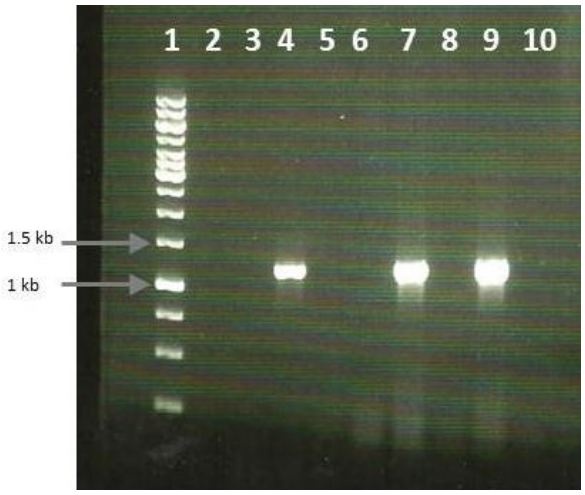


Figure 4-3 Agarose gel electrophoresis of NCU-G1 PCR products. Check for positive ligation of NCU-G1 to pBluescript vector. Lane 1, 1kb DNA ladder. NCU-G1 is 1.2 kb. Lanes 4, 7, and 9 gave positive results.

A secondary step of quality control was done. Small amount of the newly designated constructs were cut with KpnI and XhoI enzymes to check if the inserted fragment can be excised back from the vector. At the same time, it is also another way to confirm the ligation. An example of this cutting is indicated in figure 4-4. This figure shows that the NCU-G1 fragment 1.2 kb can be cut from the pBluescript vector 3 kb.

Removal of NCU-G1 from pcDNA 3.1 (+)/ myc His A and ligation into pBluescript II SK (+) vector is presented in the schematic draw (figure 4-10 A).

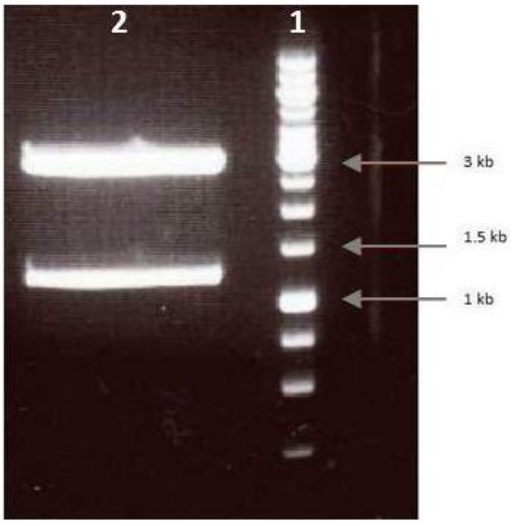


Figure 4-4 Removal of NCU-G1 from pBluescript II SK (+) using KpnI and XhoI restriction enzymes. Lane 1, 1kb DNA ladder. Lane 2, pBluescript vector 3 kb, NCU-G1 around 1.2 kb. Fractionated in agarose gel.

4.4 Mutation of hNCU-G1 inserted in pBluescript II SK (+)

For the newly constructed vector i.e. NCU-G1 inserted in pBluescript II SK (+), we did two separate single mutations typical as in table 4-2. For these mutations QuikChange® Lightning Site-Directed Mutagenesis Kit was used (section 3.2.7). This kit offers more potential materials, for example a highly specific polymerase enzyme to carry on the mutation PCR reaction. In addition, this kit also has the ability to mutate plasmids of up to 14 kb in size. Mutation products (control and samples) were spread on amp containing LB plates onto which X-gal and IPTG have been pre-spread.

Identity	**Growth**	**Color**	**Explanation**
Control vector	Present	Blue colonies	Stop codon is substituted and lacZ gene is intact
Mutated sample	Present	White colonies	lacZ gene is interrupted with an insert
Un-mutated sample	Absent	No	It should be digested with DpnI enzyme
E.coli lacking plasmid	Absent	No	The plates contain amp antibiotic

Table 4-3 Expected results for mutation supplemented with the logical reasons.

Results typically followed the expectations in table 4-3. A very high number of colonies was detected on both control and sample plates. More than 99% of the control colonies were blue. All colonies on the sample plate were white, indicating successful mutation (table 4-3).

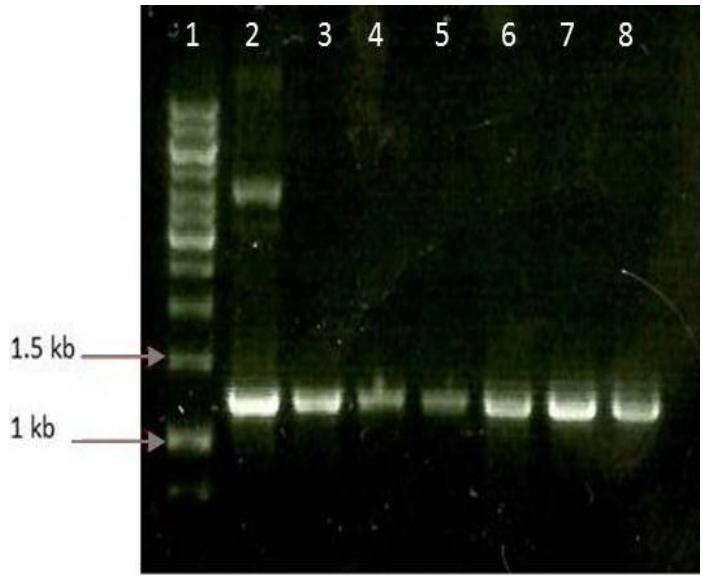


Figure 4-5 Agarose gel electrophoresis of NCU-G1 PCR test of mutations. Lane 1, 1 kb DNA marker. Lane 2, posotive control. The rest lanes are mutation samples. NCU-G1 is 1.2 kb.

A number of sample colonies were PCR tested for NCU-G1. All tested colonies gave positive result (figure 4-5). To confirm the mutation, PCR tested colonies were sequenced for NCU-G1. Both mutations were detected perfectly, whereas the rest of NCU-G1 sequence is not interrupted (appendix 7.5.1). Mutation is presented in figure 4-10 B.

4.5 Vector construction (2)

After mutation confirmation, mutated NCU-G1 samples were excised from pBluescript II SK (+) using KpnI and XhoI enzymes. pcDNA 3.1 (+)/ myc-His A vector was cut using the same enzymes. An example of this cutting is illustrated in figure 4-6. Both mutated NCU-G1 open reading frame sequences were then ligated into the pcDNA vector.

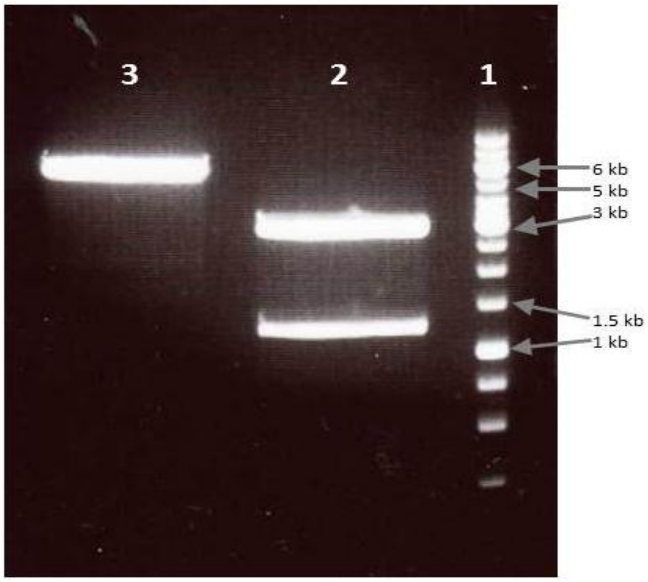


Figure 4-6 Cutting of pBluescript vector carrying mutated NCU-G1, and pcDNA myc His A vector with KpnI and XhoI enzymes. Fractionated on agarose gel. Lane 1, 1 kb DNA marker. Lane 2, pBluescript around 3 kb and NCU-G1 around 1.2 kb. Lane 3, pcDNA vector around 5.5 kb.

Along the way of cloning mutated NCU-G1 to pcDNA vector many control steps were done. For example to confirm vector identity i.e. to ensure that we are using pcDNA vector, vector was cut using two restriction enzymes. Figure 4-7 illustrates cutting of the pcDNA 3.1 (+)/ myc-His A (5.5 kb) vector with KpnI and SmaI enzymes, these enzymes cut at nucleotide positions 912 and 2137 on the pcDNA vector respectively.

The pcDNA 3.1 (+)/ myc-His vector comes in three versions A, B, or C. All versions have general main features e.g. they all have the same antibiotic resistance gene (amp resistant).

Minor but critical differences exist in their multiple cloning sites. So, it is of great importance to clone the insert to the same original vector version that it was taken from. Doing this will keep the gene in frame with the C-terminal myc epitope and the polyhistidine tag. We did sequencing of KpnI and XhoI restriction sites to ensure ligation to the right vector version. The sequencing results came up with a successful ligation of each of the mutated hNCU-G1 genes to the pcDNA3.1 (+)/myc-His A vector (appendix 7.5.2).

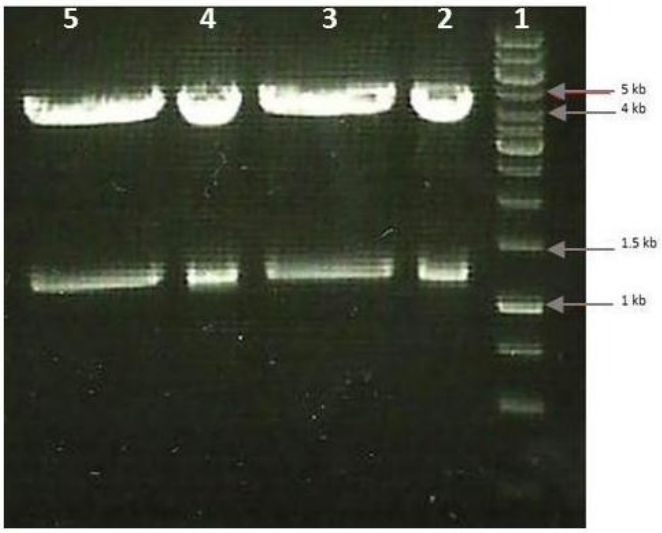


Figure 4-7 Agarose gel electrophoresis shows cutting of pcDNA vector with pcDNA one cutters: KpnI and SmaI restriction enzymes. Resulting fragments are around 4.3 and 1.2 kb.

Ligation is a technique that depends on most of the steps which precede it. Many control steps were done to optimize ligation. These include production of DNA material that is highly purified with the GE kit, taking care to keep ethanol (a component in the washing solution) concentration as low as possible. This was done by increasing sample drying time. Using positive ligation control was a wonderful tip to follow.

We did the same control steps as in section 4.3. Firstly, ligation result colonies were checked by NCU-G1 PCR using three different primer pairs. One primer set gives a product around 200 base pairs (bps), another set gives around 350 bps (figure 4-8 A). The third pair flanks NCU-G1 and so amplifies the whole region about 1200 bps (figure 4-8 B). Secondly, recombinant vector was cut with KpnI and XhoI restriction enzymes (figure 4-9). Reinsertion of mutated NCU-G1 into pcDNA 3.1 (+)/ myc-His A vector is drawn schematically in (figure 4-10 C).

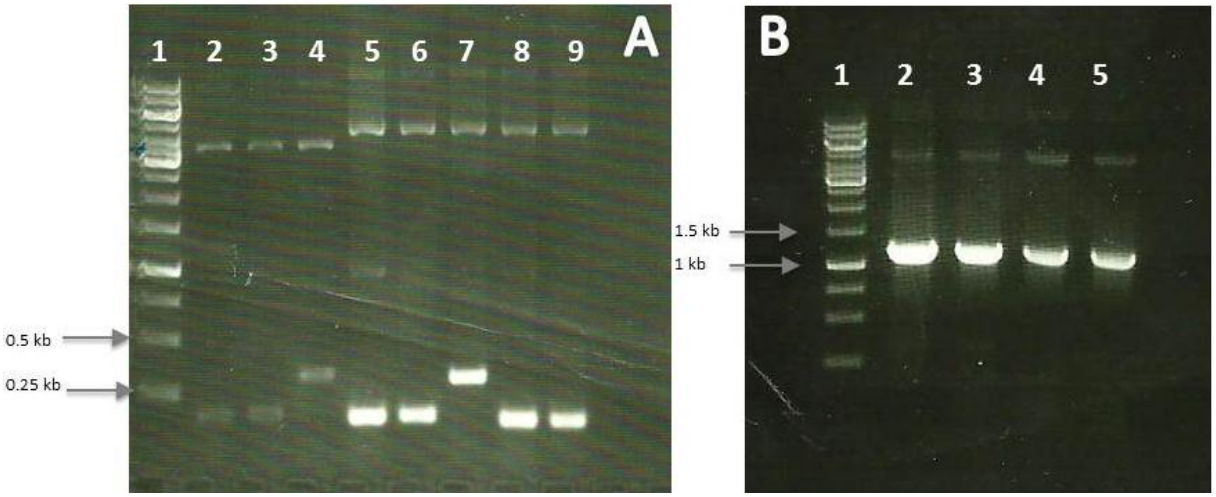


Figure 4-8 Agarose gel electrophoresis of NCU-G1 PCR products using different primer sets. **A**: Two different primer sets. Lane 1 is the DNA marker. Lanes 2, 3, 5, 6, 8, and 9 are amplification products with primers that give a product around 200 bp, lanes 4 and 7 with primers that give around 350 bps. **B**: lane 1, 1 kb DNA marker. The remaining 4 lanes, PCR products with primers that amplify the whole NCU-G1 region (1200 bps).

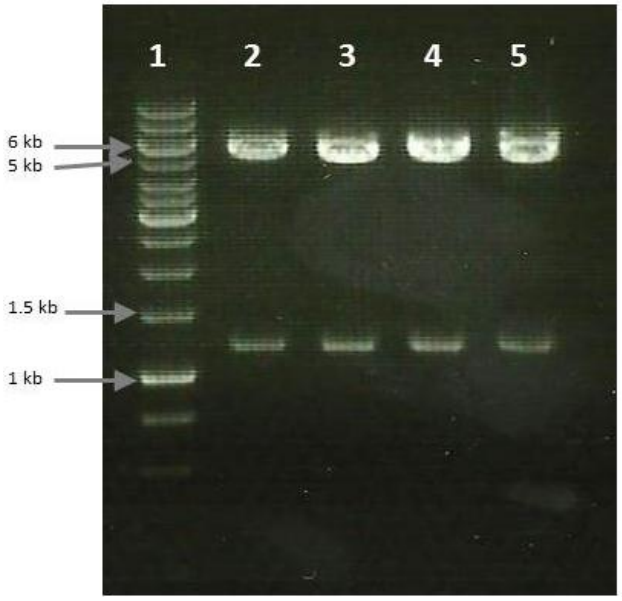


Figure 4-9 KpnI and XhoI cutting of pcDNA 3.1 (+)/ myc His A that contains mutated NCU-G1. Lane 1, 1 kb DNA marker. The remaining four lanes contain upper bands around 5.5 kb are pcDNA vector, and lower bands around 1.2 kb are NCU-G1 bands.

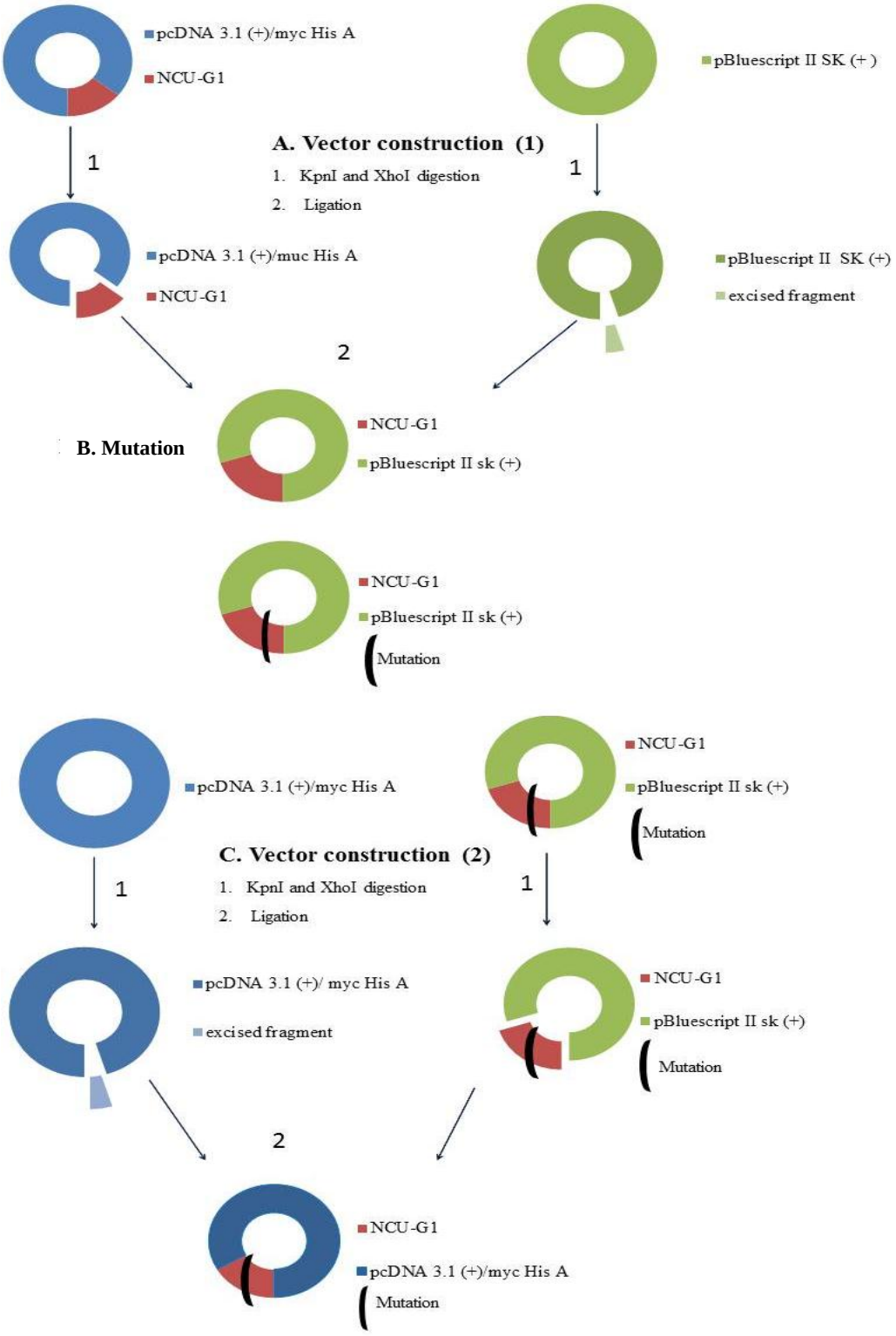


Figure 4-9 Schematic presentation of: **A**. Removal of NCU-G1 from pcDNA 3.1 (+)/ myc-His A vector, and ligation into pBlueskript SK II (+). **B**. Mutation. **C**. Return of mutated NCU-G1 to pcDNA 3.1 (+)/ myc-His A.

4.6 Transfection optimization

Transfection is a technique by which the DNA can be introduced into eukaryotic cells. Here we transfected Hela cells with a vector containing lacZ gene (the coding gene for β-galactosidase enzyme). Cells were transfected with Fugene 6 and X-tremegene 9 transfection reagents using 3:1 (Transfection reagent: DNA) ratio, exactly 1 µg of DNA was added to the diluted transfection reagent, this contains 250 ng pCMV-lacZ vector and 750 ng pSG5 empty vector. Cells were harvested for β-galactosidase measurement. Very high transfection efficiency was achieved by Hela cells transfected with X-tremegene 9 transfection reagent compared to the same cells transfected with Fugene 6 transfection reagent (figure 4-11).

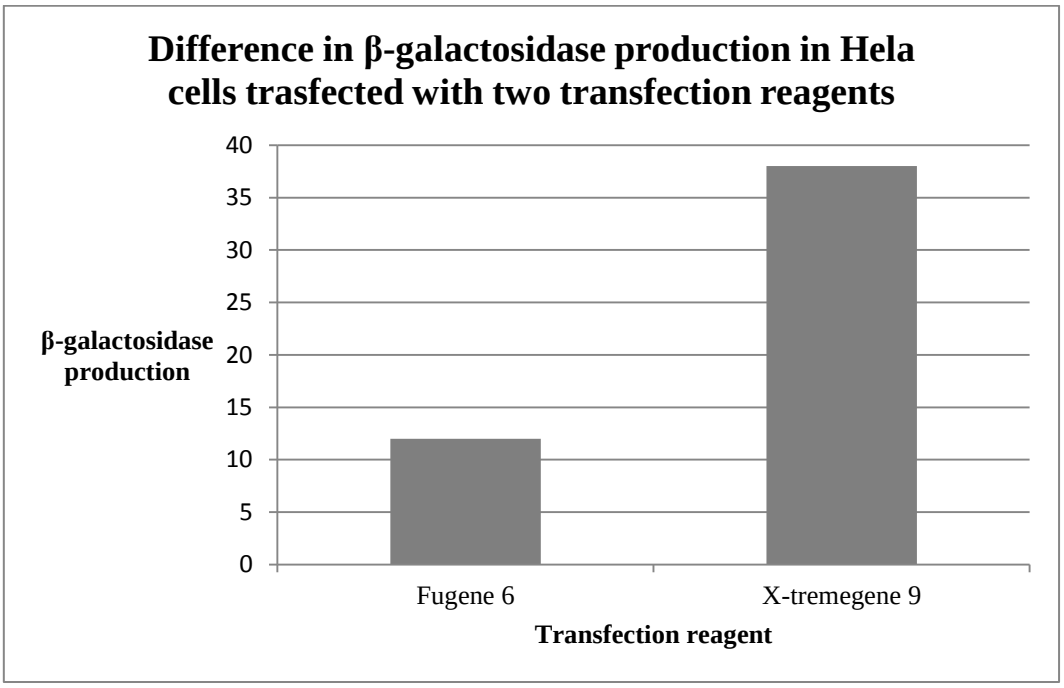


Figure 4-11 Measurement of β-galactosidase production in Hela cells transfected with Fugene 6 or X-tremegene 9 transfection reagents.

We proceeded with X-tremegene 9 transfection reagent. Hela cells were transfected with the same lacZ containing vector as above using 3:1, 3:2, or 6:1 ratios. Cells were fixated and stained with X-gal stain. β-galactosidase enzyme has the ability to break down x-gal producing blue colored cells. The highest transfection efficiency was obtained from cells transfected with the 3:1 ratio as shown in figure 4-12. The 3:1 ratio was used in the remaining experiments.

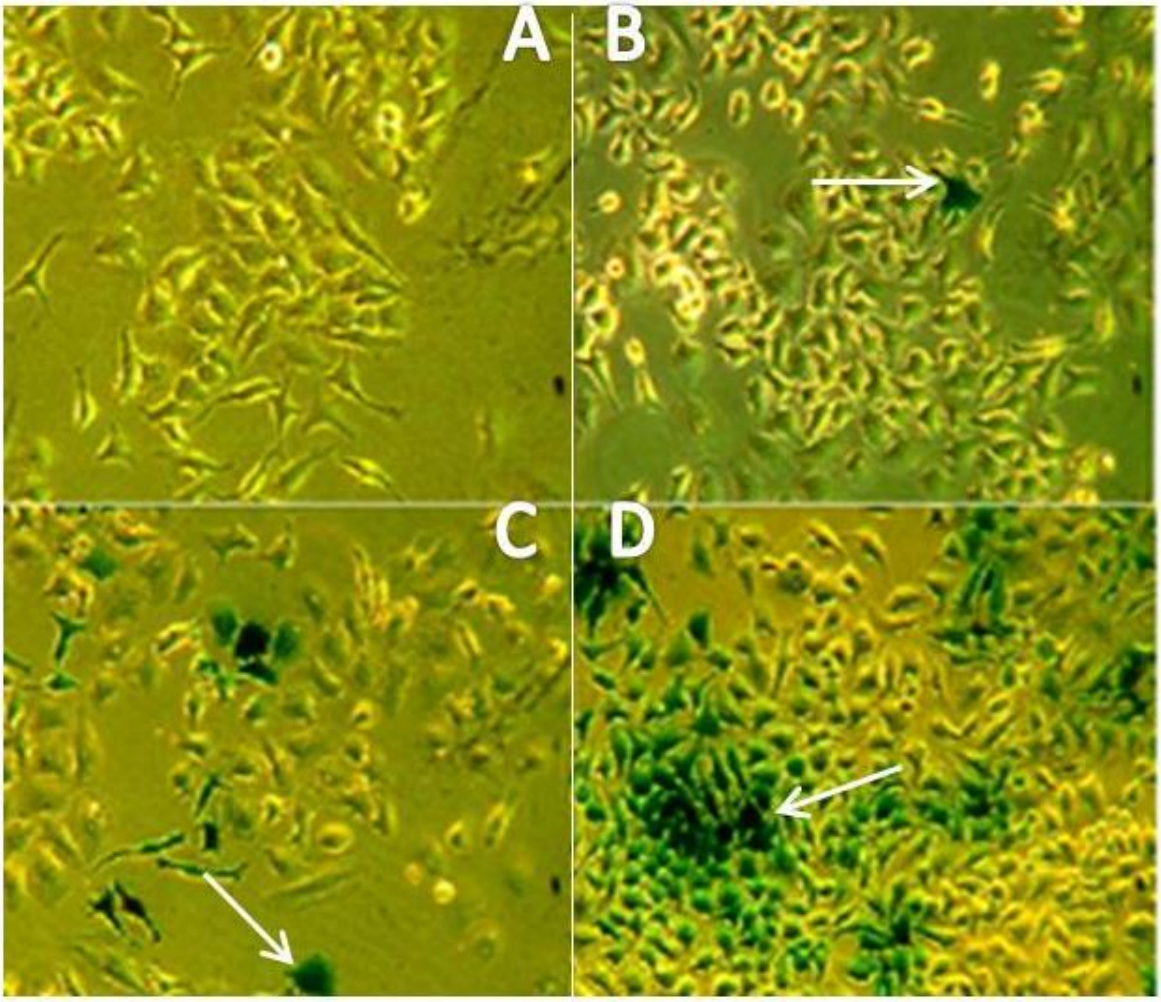


Figure 4-12 X-gal stain of Hela cells transfected with different X-tremegene 9: DNA ratios. White arrows are pointed toward some of the transfected cells. **A**. Non-transfected cells. **B**. Cells transfected with 3: 2 ratio. **C**. Cells transfected with 6:1 ratio. **D**. Cells transfected with 3: 1 ratio.

Cells were plated at different cell densities, to determine the best cell density. Cell densities used were 125.000, 150.000, 175.000, 200.000 cells/well. Cells were transfected with lacZ containing vector and harvested for β-galactosidase measurement. As shown in figure 4-13 lower cell densities yielded better transfection. We proceeded with the 150.000 cells/well to avoid using very low density, as the 125.000 and 150.000 cell densities gave almost identical results.

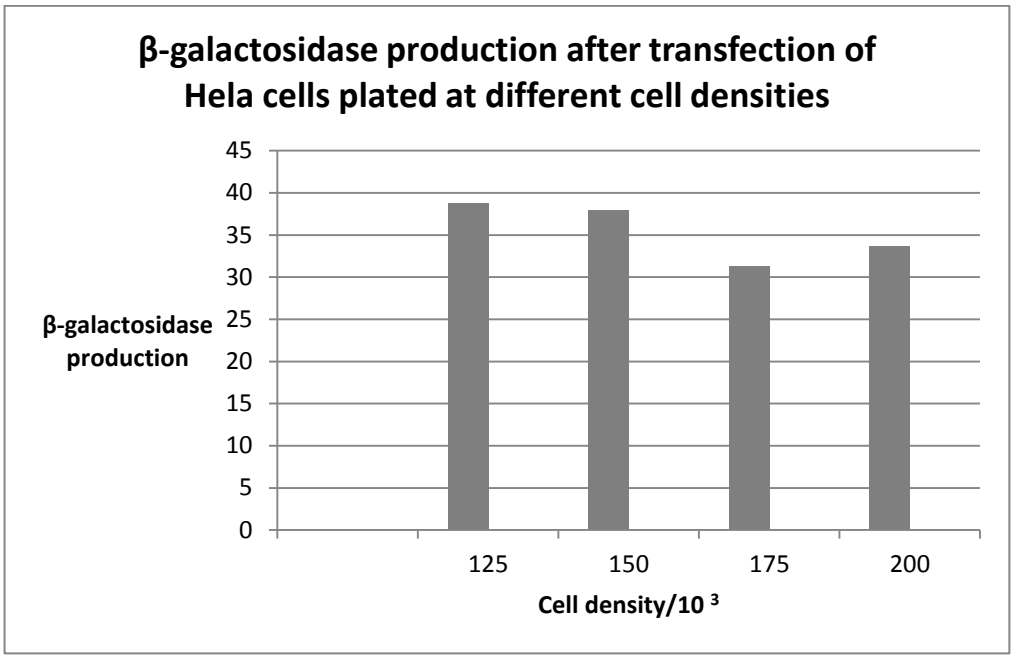


Figure 4-13 The effect of cell density on transfection.

4.7 The effect of glycosylation mutation on hNCU-G1 protein molecular weight and expression

Hela cells were transfected with the wild type and the two mutations of hNCU-G1 inserted in pcDNA 3.1 (+)/ myc-His A expression vector. DNA added was 375 ng from NCU-G1 containing vector and 625 ng pSG5 empty vector. Cells were harvested for total protein measurement 24 hr after transfection. These are (a) non-transfected cells (NT), (b) cells transfected with NCU-G1 construct of: wild type (WT), mutation 65 (65mut), and mutation 230 (230mut). 20 µg of the measured protein was loaded from each sample to an SDS PAGE (4-12% gradient gel). After blotting the NCU-G1 protein was captured using an antibody against the myc tag of the vector (anti-myc). Very wide bands were detected by Western analysis indicating glycosylation.

As a positive control, a well transfected with a lacZ containing vector was included in each experiment, DNA added was 250 ng pCMV-lacZ vector and 750 ng pSG5 empty vector. This was used to measure β-galactosidase production. β-galactosidase produced was between 30-40 O.D. unit per hr per well.

Another SDS PAGE (4-12% gradient gel) was loaded with 5 µg protein and captured with an anti-myc antibody in a western blotting experiment. Western analysis of this batch indicates

that, NT has a band at around 62 kDa, WT has a main band at around 75 kDa and has the same 62 kDa band of the NT, 65mut has the same two bands of the WT plus a band around 60 kDa, 230mut has a main band at around 72 and the same 62 and 60 kDa bands of the 65mut (figure 4-14). This reflects that disruption of the consensus sequence in position 230 of the hNCU-G1 amino acid sequence decreased the protein MW by about 3 kilodaltons, whereas 65mut had no effect on the hNCU-G1 protein MW.

Band intensities were variable between the 4 samples. To exclude inconsistent loading, all membranes were stripped after NCU-G1 detection and treated with an antibody against β-actin. β-actin band is around 44 kDa. All these experiments conclude that the 62 kDa band is denser in the three transfected samples compared to the NT, the 75 kDa band of 65mut is a bit weaker compared to the WT, the 230mut bands are very dense in comparison to all samples.

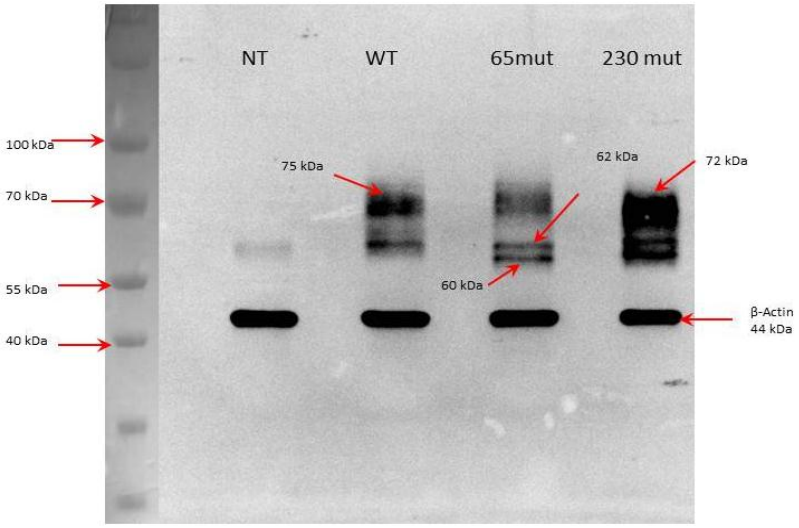


Figure 4-14 Western immunoblotting picture of NT, WT, 65mut, and 230mut protein samples. 5µg of the protein was loaded into a 4-12 % gradient polyacrylamide gel and separated by western technique using anti-myc antibody. The membrane was then stripped and treated with an antibody to detect β-actin (lower uniform bands at 44 kDa).

Using the Carestream program, we can take the same picture with different light effects, i.e. increase the background light, decrease bands sharpness and so on. This helped us to see the duplex bands of 60 and 62 kDa of both mutations a bit clearer (figure 4-15).

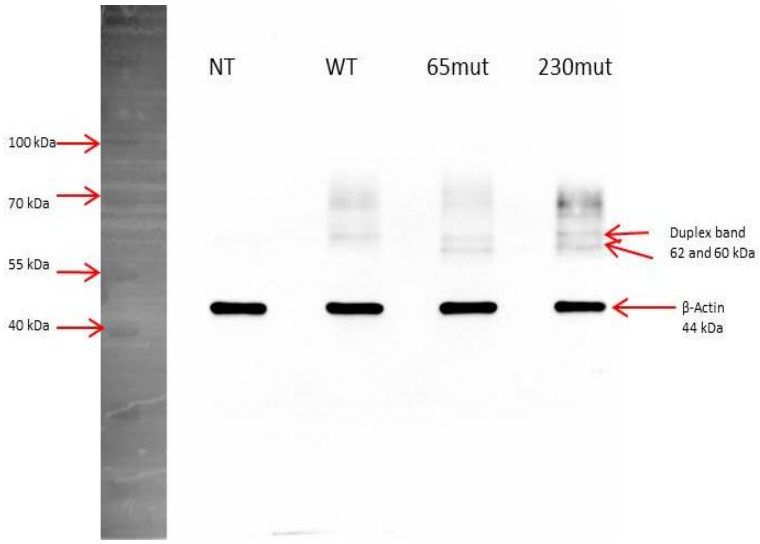


Figure 4-15 Western immunoblotting picture with decrease light. 5 µg NT, WT, 65mut and 230mut proteins was loaded into a 4-12 % polyacrylamide gel and separated by western technique using anti-myc antibody. This picture shows the duplex bands at around 60 and 62 kDa of the 65 and 230muts. The lowest bands are the β-actin bands around 44 kDa.

To increase bands resolution, 1 and 2 µg of the four samples were run at a lower volt 150 volt for longer time around 3 hr .This experiment gave the same result as previous (figure 4-16) .In addition, a 10% polyacrylamide gel was loaded with 5 µg from the four protein samples. The same bands patterns were detected as well (figure 4-17).

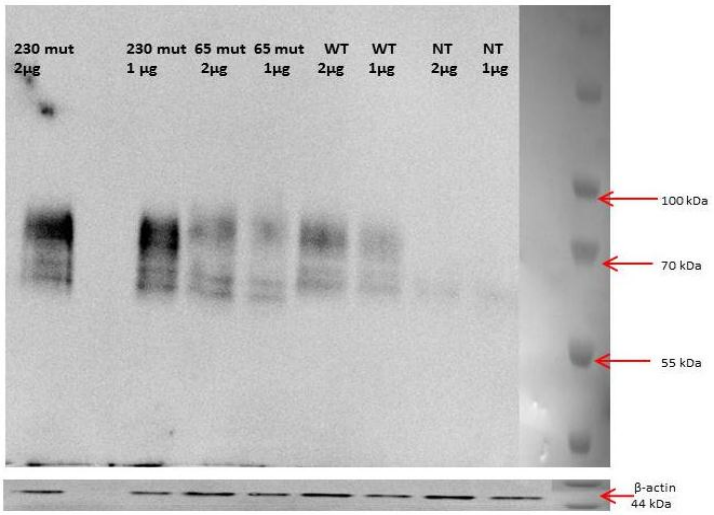


Figure 4-16 Western immunoblotting picture from a 4-12% polyacrylamide gel loaded with 1 and 2µg NT, WT, 65mut, and 230mut protein samples. This gel was run at 150 volt for around 3 hr. The proteins were captured with an anti-myc antibody. The membrane was then stripped and treated with an antibody to detect β-actin (lowest part of the picture).

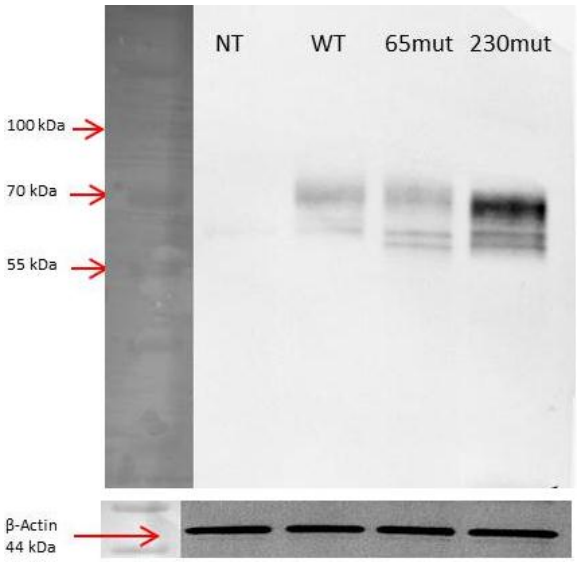


Figure 4-17 Western immunoblotting picture of 5 µg NT, WT, 65mut and 230mut proteins loaded into a 10% polyacrylamide gel. These proteins were captured with an anti-myc antibody. The membrane was then stripped and treated with an antibody to detect β-actin (lowest part of the picture).

For more confirmations, NCU-G1 protein was captured by an anti-NCU-G1 antibody in a western blotting analysis of the four samples loaded into a 10% polyacrylamide gel. This experiment yields four identical bands at around 80 kDa and four other identical bands at around 60 kDa (figure 4-18). This reflects that Hela cells produce NCU-G1, and the NCU-G1 captured by this antibody is only the endogenously produced one.

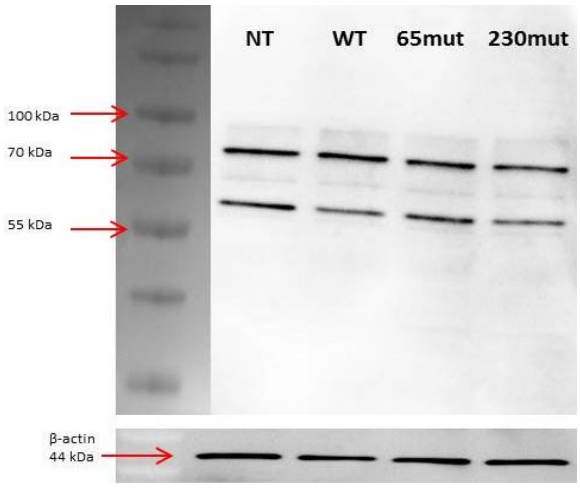


Figure 4-18 Western immunoblotting picture of 5 µg NT, WT, 65mut and 230mut proteins loaded into a 10% polyacrylamide gel. These proteins were captured with an anti-NCU-G1 antibody. The membrane was then stripped and treated with an antibody to detect β-actin (lowest part of the picture).

To verify the identity of the detected bands, the four protein samples were treated with PNGase F enzyme (Peptide N-glycosidase F). This enzyme removes glycans from N-

glycoproteins. Treatment of WT, 65mut, and 230mut protein samples with this enzyme lowers the MW to around 44 kDa which is the predicted MW for the unglycosylated NCU-G1 protein (http://www.justbio.com/). A faint band at 62 kDa was also detected in the PNGase treated transfected samples, which is weaker than the one in the non-treated samples. On the other hand, the NT 62 kDa band was not affected by the PNGase treatment (figure 4-19).

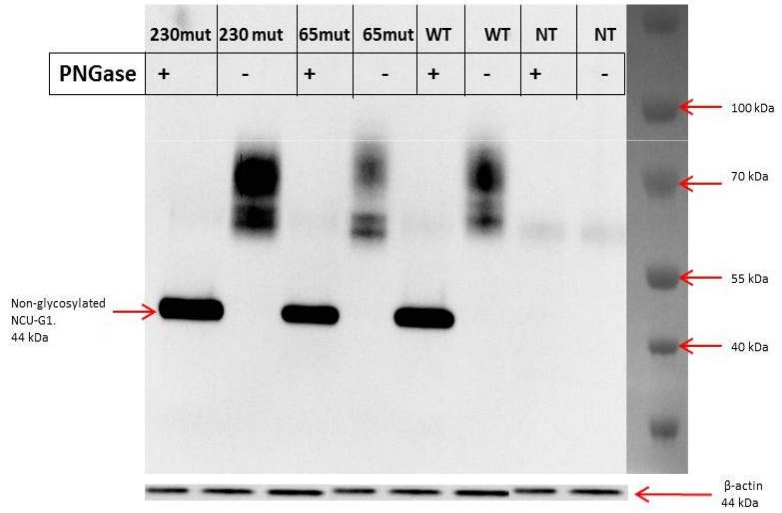


Figure 4-19 Western immunoblotting picture of PNGase treated and non-treated NT, WT, 65mut, and 230mut protein samples. 5µg of the protein was loaded into a 4-12 % gradient polyacrylamide gel and separated by western technique using anti-myc antibody. The membrane was then stripped and treated with an antibody to detect β-actin (lower part of this figure).

4.8 The effect of glycosylation mutation on hNCU-G1 subcellular localization

Hela cells were transiently transfected with hNCU-G1-myc-tagged of WT, 65mut and 230mut. After fixation the expressed protein was captured with an anti-myc antibody. This was detected with an anti-IgG carrying a fluorescent probe and visualized using confocal microscopy. The wild type was found to be diffusely localized in the cytoplasm including the cytoplasmic membrane, and some accumulated in a specific cytoplasmic area beside the nucleus (figure 4-20, WT). The 65mut was detected with a bit lower expression compared to the WT, and with a very low expression in the cytoplasmic membrane (figure 4-20, 65mut). On the other hand, the 230mut was highly expressed in all cytoplasmic areas (figure 4-20, 230mut a and b), and it was also detected in the nucleus (figure 4-20, 230 b). A summary of these results is presented in table (5-1).

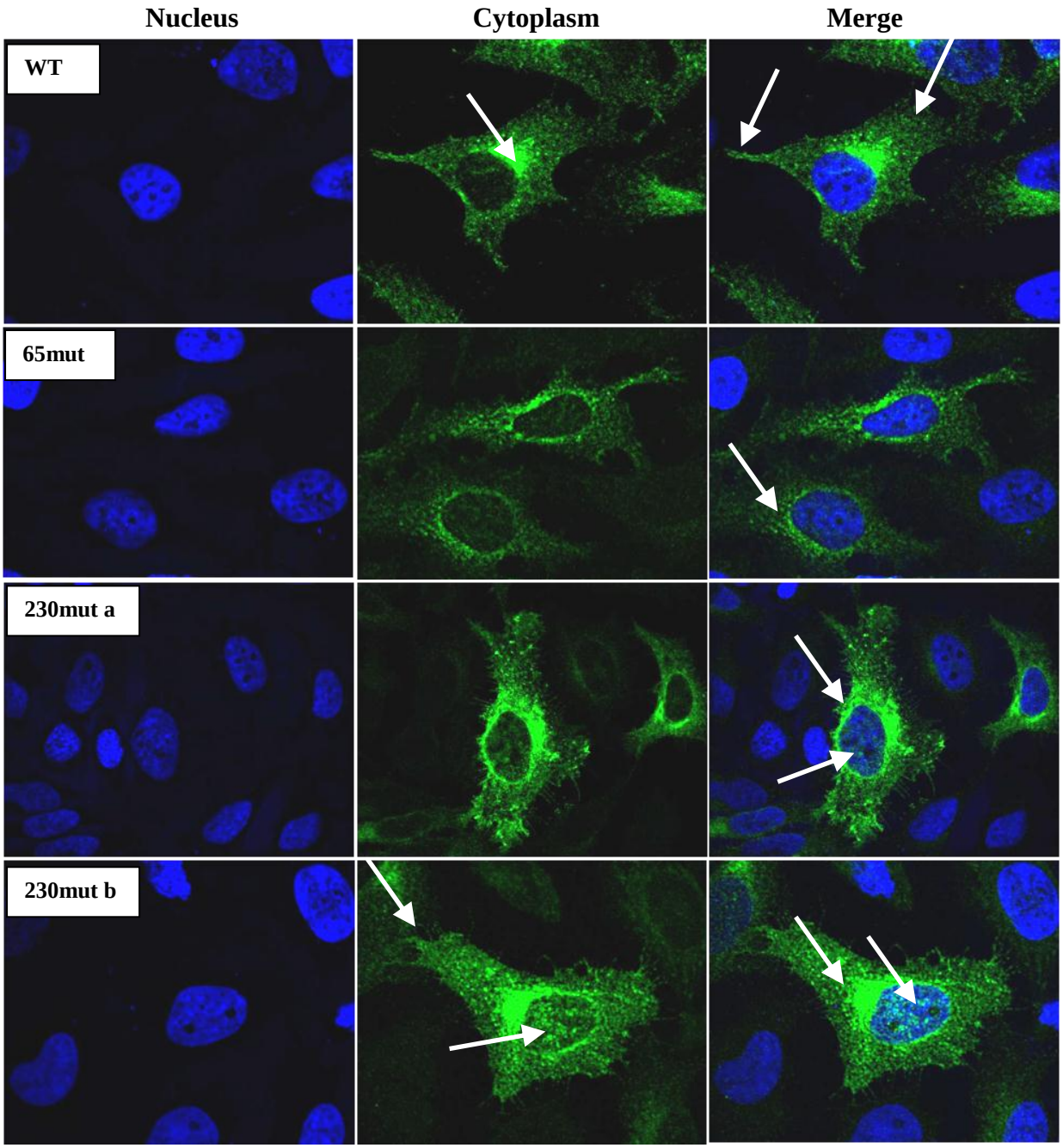


Figure 4-20 Subcellular localization of the WT, 65mut, and 230mut of the hNCU-G1 protein in Hela cells. Hela cells were transfected with hNCU-G1-myc-tagged of WT, 65mut and 230mut. NCU-G1 was captured with an anti-myc antibody and an Alexa Fluor 488 (green)-coupled IgG antibody. Nuclei were stained with Hoecst (blue). Immuoflouresce was detected using confocal microscope. White arrows are pointed toward some of the hNCU-G1 localization areas for more explanation. Columns from left to right: nucleus, cytoplasm, and merge. Rows from top to base: WT, 65mut, 230mut a, 230mut b.

5 Discussion

Postranslational modifications are among the most important events that greatly influence proteins in many aspects. PTMs have been studied widely, but the effect of one such modification on certain protein may not be the same for another protein. These modifications have been found to affect proteins structure, folding and conformation, stability, trafficking, immunogenicity, half-life, and other effects [18, 64, 65].

In the present project, we studied one of these modifications and its effect on the hNCU-G1 protein. NCU-G1 is a highly glycosylated protein. Glycosylation is one of the most important PTMs, that regulates protein activities and functions. Glycosylation is classified to many sub-types, of these N- and O-linked glycosylation are the main types. While N-glycosylation is the co-translational addition of glycans to specific asparagine residues on nascent polypeptide chains, O-glycosylation is the postranslational addition of glycans to serine or threonine amino acids [23, 24]. This process is not a random process i.e. not all asparagine amino acids will become N-glycosylated for example. Certain factors are directing this process towards specific asparagine, or serine/threonine residues in N- or O-glycosylation respectively. Although the exact prerequisites have not been fully identified yet, N-X-S/T motif has been described as the main glycans acceptor motif in N-glycosylation [35]. On the other hand, no acceptor motif has been identified yet for O-glycosylation. Based on the N-linked glycosylation consensus motif hNCU-G1 protein contains seven N-glycosylation predicted sites of which four have been characterized as potential (table 4-1). In contrast, the mouse NCU-G1 orthologue protein has nine predicted N-glycosylation sites. Although the seven glycosylation sites of hNCU-G1 are the same of that of mouse NCU-G1, they differ a bit in their position, this is because the human NCU-G1 has a leucine and lysine amino acid inserted in positions 52 and 139 respectively. There are more two sites at positions 85 and 166 of the mouse NCU-G1, the acceptor motifs for these sites are N-T-T and N-L-S respectively, compared to N-T-P and T-L-S in the corresponding positions on the hNCU-G1 protein sequence. According to NetOglyc and EnsemblGly databases, hNCU-G1 is devoid of O-glycosylation.

5.1 Mutation of two of the potential glycosylation sites on the hNCU-G1 amino acid sequence

Two of the predicted potential glycosylation sites on the hNCU-G1 protein sequence were separately mutagenized to disrupt the N-glycosylation acceptor motif at these sites. This involved the change of asparagine amino acid to alanine amino acid on both sites (table 4-2). The mutation was carried out on a pBluescript II SK (+) vector because it was difficult to mutagenize hNCUG-1 in the pcDNA3.1 (+)/ myc-His A expression vector. This is partly because pcDNA is a relatively large vector (5.5 kb) compared to the 3 kb size of the pBluescript vector, but it may also be because the kit we used to mutate our target in pcDNA vector has lower efficiency towards large constructs. With pBluescript the mutation was achieved sucessfully, and we verified it by sequencing through the whole NCU-G1 nucleotide sequence, and also by PCR test (figure 4-5). The two Mutated hNCU-G1 versions 65mut and 230mut were cut and ligated again to the original pcDNA vector. This was confirmed by N- and C-termini sequencing.

5.2 Effect of glycosylation mutants on hNCU-G1 protein expression as detected by western analysis

The effect of specific glycosylation sites on protein expression have been studied for several proteins. Glycosylation is known to affect protein folding and stability and so protein expression. Some N-glycoproteins have an N-glycan dependent quality control folding system, these proteins do not leave the ER before complete folding occurs, which is enhanced by glycosylation for them. These proteins if not glycosylated at these specific sites they will be degraded in the ER. This will lead to lower or even no expression [6]. As an example for that, mutation of N-glycosylated sites of the Gn membrane proteins of the Hantaan virus gave different results regarding protein folding. Firstly, single mutations at site 134 or 347 of the Gn protein led to retention of the protein in the ER. Secondly, while single mutations at positions 235 or 399 of this protein didn't affect the folding, the double mutant led to mis-folding [12]. Another example is on the effect of glycosylation on protein stability, prevention of glycosylation of the Kv1.4 potassium channel glycoprotein decreased it is stability [64].

In this study, we transiently expressed the wild type hNCU-G1 and the two single mutants 65mut and 230 mut in Hela cells. The expressed protein was captured by an antibody against

the myc tag of the pcDNA expression vector in a western analysis. We noticed lower expression of the 65mut compared to the WT, as clearly shown in all figures in section 4.7. This is may be because glycosylation at position 65 of the human NCU-G1 protein sequence affects its folding, and loss of this glycosylation site leads to partial effect on the protein folding and so retention in the ER. It could be also that this site protects the protein from degradation, or positively influences the protein antigenicity and so binding to the antibody, since glycosylation has been described to influence the protein antigenicity [20]. Another possibility is that this mutation increased our target protein cycling.

In contrast, mutation at position 230 of the hNCU-G1 protein sequence apparently increases its expression (all figures in section 4.7). Glycosylation at this site may have an effect on the protein conformation, and loss of this glycosylation makes the antibody binding site in the protein more accessible by the antibody. Especially, because the small size myc-tag may be folded within the protein and evade the detection. Another possibility is that this mutation slows down the protein turnover. It could also be that this glycosylated site decreases the NCU-G1 antigenicity, and loss of this modification may makes the protein more antigenic, thus binds to the anti-myc antibody more efficiently. The difference in immunogenicity between the normal and cancerous Heat shock protein gp96 was found to be due to differences in the monosaccharide composition between the two protein versions, and samples from more potential cellular transformations were found to have decreasing amounts of many monosaccharides [66]. Thus, loss of some monosaccharaides may render NCU-G1 more antigenic.

5.3 Effect of glycosylation mutants on hNCU-G1 protein molecular weight as detected by western analysis

Glycosylation is the addition of oligosaccharides to specific amino acids, because of that it can increase the protein MW. Indeed, this depends on the oligosaccharide composition to be added. Loss of some glycosylated sites has been found to decrease the protein MW [65, 67-69]. For highly glycosylated proteins, the oligosaccharide addition may increase the protein MW by more than 30%, as in the case of NCU-G1 [55]. A change as small as 1 kDa can exist between a glycoprotein and its glyosylation mutated version. For instance, oligosaccharides

attached to asparagines 467 and 496 in the Alzaheimer's disease amyloid precursor protein (APP) each one weigh 1 kDa [68].

In the present work, when we detected the NCU-G1 myc tagged with an anti-myc antibody (figure 4-14), we found a major band at around 75 kDa in both WT and 65mut.This result shows that mutation of the hNCU-G1 at position 65 of the amino acid sequence didn't give noticeable difference in the protein MW. On the contrary, the corresponding band of 230mut was found to be 3 kDa lower than the wild type. This result was found also in some glycosylation sites of other proteins as mentioned above.

A band at around 60 kDa was detected only in the two mutated versions. This band is likely to be the result of slower glycosylation/partial degradation of the two mutated versions, or the formation of a less glycosylated form as a result of mutation. Mutation is a kind of disturbance to the normal protein behavior i.e. structure, function, activity and so on. So, it is expected that the resulting protein can differ from the wild type. In congenital disorder of glycosylation (CDG) a group of diseases characterized by glycosylation defects, which are caused by mutation in some types, these defects may lead to serious complications, some of them are fatal [70-72].

A band at around 62 kDa was detected in all samples, being denser in the transfected samples. We assume that the 62 kDa band in the non-transfected well is a result of binding of the anti-myc to the endogenous myc protein. Supporting this hypothesis, the myc protein has been detected at this size in yeast [73], and breast cancer [74]. In addition, because the detected 62 kDa band from the transfected samples was denser compared to the non-transfected sample, we expect that the hNCU-G1 protein has two glycoforms, and the band we detect at 62 kDa is a merged version of one of the isoforms and the endogenous myc protein. This could be tested using a different tag.

For more characterization, the four hNCU-G1 protein samples were captured with an anti-NCU-G1 antibody (figure 4-18). Two bands of 62 and 80 kDa were detected in each sample. Regarding bands molecular sizes and intensities, these bands are identical in the 4 samples, meaning that Hela cells can produce NCU-G1. It also reflects that these cells produce two forms of our protein. This could be a function of the cells or the protein itself. The detection of the two bands by the anti-NCU-G1 and anti-myc antibodies, confirms that Hela cells express two isoforms of the hNCU-G1 protein. However, the MW of the endogenous NCU-

G1 differs from the transfected one. Schieweck et al. found two isoforms of 80 and 70 kda when they expressed NCU-G1 in HT1080 cells, but they found a 75 kDa band of NCU-G1 from a lysosomal fraction [55]. This raises the possibility that NCU-G1 has more than one glycoform, the difference between them might be cell and organelle specific.

5.4 Deglycosylation of the hNCU-G1

To verify the identity of the detected bands, the four samples were treated with PNGase F enzyme, an enzyme that removes N-linked glycans. The deglycosylated proteins were captured with an anti-myc antibody (figure 4-19). Treatment of these samples with the deglycosylase enzyme removed all the previously detected bands on the WT, 65mut and 230mut, and yielded a new band at around 44 kDa. A light band at 62 kDa can be seen in these three samples. Also, the NT band was not sensitive to the PNGase treatment. The 44 kDa band is the predicted MW for the unmodified NCU-G1 protein (http://www.justbio.com/). This reflects that all the previously detected bands were NCU-G1 glycoforms, and proves our assumption that Hela cells produce two hNCU-G1 glycoforms. In addition, these data confirm the prediction that NCU-G1 is an N-glycoprotein and it is not O-glycosylated.

The presence of two glycoforms for the NCU-G1 protein can be explained by many theories. One is that four of the acceptor motifs for N-glycosylation on the hNCU-G1 protein contain serine in the third position (table 4-1), this may lead to inefficient glycosylation of these sites. Concerning this assumption, the Antithrombin III (AIII) was found to have two glycoforms, resulting from partial glycosylation of one of its N-glycosylated sites due to the presence of serine amino acid at the third position [75]. It has been shown that T amino acid is more likely to be glycosylated than S amino acid [76]. Another explanation could be that NCU-G1 is distributed in many subcellular compartments and that each isoform comes from specific site (s). This is the case for the protein Nucleolin which was found with the MW of 105 kDa [77], and later two glycoforms were detected with 105 and 113 kDa, with a form being localized in the nucleus (105 kda) and the other one in the cytoplasm (113 kDa) [78]. Some proteins tend to have more than one glycoform because they originate from more than one tissue [79, 80]. We exclude this possibility as we expressed our protein in only one kind of cells.

5.5 Effects of glycosylation mutants on the hNCU-G1 subcellular localization

We studied the subcellular localization of the hNCU-G1 myc-tagged protein using confocal microscopy. We saw noticeable differences between the wild type hNCU-G1 and the two mutations we made (see figure 4-20, and table 5-1). These results show that hNCU-G1 is basically localizes in the cytoplasm including the cytoplasmic membrane of Hela cells, with some aggregation at one cytoplasmic area beside the nucleus. The same kind of cytoplasmic diffusion of NCU-G1 was observed by Schieweck et al. After further analysis they conclude that NCU-G1 basically localize in the lysosomes and late endosomes [55]. Supporting this result, NCU-G1 was found to have two nuclear export signals [54].

In addition, we noticed that while the 65mut decreased hNCU-G1 total expression, the 230mut increased the expression. The decrease of hNCU-G1 expression in the cytoplasmic membrane of the 65mut is probably due to the total lower expression caused by this mutation. These results match the results we got by the western analysis.

In addition, in the 230mut, hNCU-G1 was also detected in the nucleus. Indicating that this glycosylation site causes partial decrease in the hNCU-G1 nuclear sorting. Small proteins with MW less than 50 kDa can passively diffuse to the nuclei [81]. Most other nuclear proteins contain nuclear localization signal (NLS) to facilitate their active entry to the nuclei. Some nuclear glycoproteins do not have NLS. However, they enter the nuclei by a sugar dependent nuclear import mechanism, although it is a less efficient mechanism in comparison to the NLS dependent nuclear import [82]. NCU-G1 does not have NLS [54], but the amount observed in the nuclei of 230mut could be transferred from the cytoplasm upon this glycosylation site blockage using the sugar dependent import mechanism. If this hypothesis is true, it will suggest that glycosylation of our target protein in position 230 prevents its nuclear import using the sugar dependent import mechanism. Although the trafficking of some other proteins is found to be affected by glycosylation [18, 21], this is not the case for all proteins [67]. In addition, we cannot exclude that WT NCU-G1 can be present in the nucleus. However, we succeeded in its detection only in the 230mut because of the relative high protein expression resulted from this mutation. This agrees with the results from the Eskild group [54].

The variables	WT	65mut	230mut
Total expression	Present	Lower	Very high
Cytoplasmic expression	Present	Lower	Very high
Aggregation in a cytoplasmic area beside the nucleus	Present	Present	Very high related to the expression
Expression in the cytoplasmic membrane	Present	Almost absent	Higher
Nuclear expression	Absent	Absent	Present but very low

Table 5-1 Summary of the subcellular localization of the hNCU-G1: the WT and the two mutants 65mut and 230mut.

6 Conclusion

In this study we showed that hNCU-G1 has two glycoforms when expressed in Hela cells. These isoforms have a MW of around 75 and 62 kDa. Disruption of the N-glycosylation acceptor motif at amino acid 65 showed no effect on the apparent MW. However, disruption at position 230 decreased band 75 to 72 kDa. Both mutations produced a third band at around 60 kDa, which might be a result of the protein degradation or slower glycosylation. In addition, the 65mut lowered the total protein expression, while the 230mut increased it.

hNCU-G1 was found to be localized in the cytoplasm. But in the case of the 230mut it has been detected also in the nucleus.

Further studies should be carried out to enable better characterization. This should include:

- Construction of single and multiple mutations of all of the potential glycosylation sites of the hNCU-G1 protein sequence, and study of their effect (s) on the protein.
- Identification of the exact causative agents for the results we obtained here. For example, why do we get low expression as a result of a mutation at position 65 of our target protein?
- Expression of hNCU-G1 in cells other than Hela. And a western analysis ^{o}f this protein extracted from different cellular organelles. This will enable identification and characterization of NCU-G1 glycoforms.
- Study of the subcellular localization of the fully glycosylated and glycosylation mutants of the hNCU-G1 in more details. This may involve the use of co-localization markers to determine the exact location. Also, we can use Leptomycin B (LMB) to inhibit nuclear export and therefore study the nuclear localization of NCU-G1 in depth.

7 Appendix

7.1 Solutions

7.1.1 β-galactosidase assay

ONPG-reaction buffer

4mg/ml ONPG

ONPG is diluted in phosphate buffer, 100 mM, pH 7.0. Prepared right before use.

Phosphate-buffer

100 mM Na_2HPO_4/NaH_2PO_4

dH_2O

Mix 100 ml of Na_2HPO_4 with 100 ml of NaH_2PO_4 to give pH 7.0.

Z-buffer

60 mM Na_2HPO_4

40 mM NaH_2PO_4

10 mM KCl

1 mM $MgSO_4$

50 mM β-mercaptoethanol

dH_2O

Mix all components except β-mercaptoethanol, adjust pH to 7.0, store at $4^{O}C$. add β-mercaptoethanol to the volume that will be used.

7.1.2 Cell culture media

Medium for Hela cells

500 ml Dulbecco's modified Eagles medium (DMEM) 4.5 g/l glucose

25 ml fetal calf serum

5 ml Pen/Strep (5000 units/ml)

5 ml L-Gluamine

Sterilize filtrate all additives

7.1.3 Competent cell production

SOB-medium

10 g tryptone

2.5 g yeast extract

0.29 g NaCl

625 µl 2M KCl

MilliQ-H_2O to 490 ml

Autoclave, and keep at 4^oC

Right before use, add the following filter-sterilized solutions:

5 ml 1M $MgCl_2$

5 ml 1M $MgSO_4$

TB-buffer

1.5 g pipes

1.1 g $CaCl_2$ x H_2O

9.3 g KCl

5.3 g $MnCl_2$

465 ml milliQ H_2O

Mix all components except $MnCl_2$, adjust the pH to 6.7 with KOH. Then add $MnCl_2$ and filtersterilize the solution.

Store at 4^oC

Cold when use

7.1.4 LB agar and medium

LB agar with amp

500 ml LB medium

10 g bactoagar

Autoclave 121^{o}C for 20 min

After cooling add ampicillin 75µg/ml

LB medium

10 g trypone

5 g yeast extract

10 g NaCl

Complete to 1L with milliQ/dH_2O

Adjust PH to 7.2 with10M NaOH

Autoclave 121^{o}C for 20 min

7.1.5 Others

10x PBS

80 g NaCl

2 g KCl

7.4 g Na_2HPO_4 x $2H_2O$

2 g KH_2PO_4

MilliQ-water to 1 liter

50x TAE buffer

242 g Tris-base

57.1 ml Glacial acetic acid

37.2 g $Na_2EDTA*2H_2O$

dH_2O to 1 liter

7.1.6 Western immunoblotting

Lysis buffer (RIPA)

20mM Tris pH 7.5

150 mM NaCl

1 mM Na_2EDTA

1mM EDTA

1% Triton

2.5 mM Na pyrophosphate

1 mM β-glycerophosphate

1 mM Na_3VO4

1x PBS with Tween

50 ml 1x PBS

0.5 ml tween (0.1 %)

dH_2O to 500 ml

1x PBS –Tween with 5% milk powder (blocking buffer)

50 ml 1x PBS with Tween

2.5 g milk powder

Stripping buffer

121 g Tris-base (2M)

50 ml 0.5 M EDTA

28.5 ml Acetic acid

dH_2O to 500 ml

Transfer/blotting buffer

15 g Tris-base

72.5 g Glycine

1L Methanol

dH_2O to 5 liter

7.1.7 X-gal assay

Gluteraldehyde solution

0.25 % (v/v) gluteraldehyde

1x PBS

X-Gal solution

0.2% (v/v) X-gal

2mM $MgCl_2$

5mM $K_4Fe(CN)_6$*3H2O

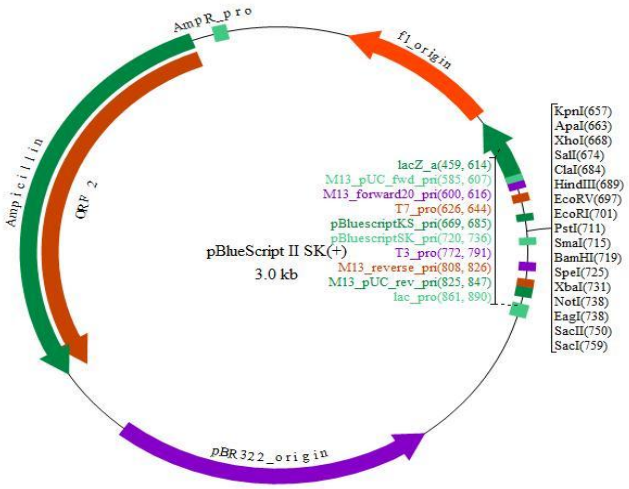
5mM $K_3Fe(CN)_6$

dH_2O

7.2 Plasmid maps

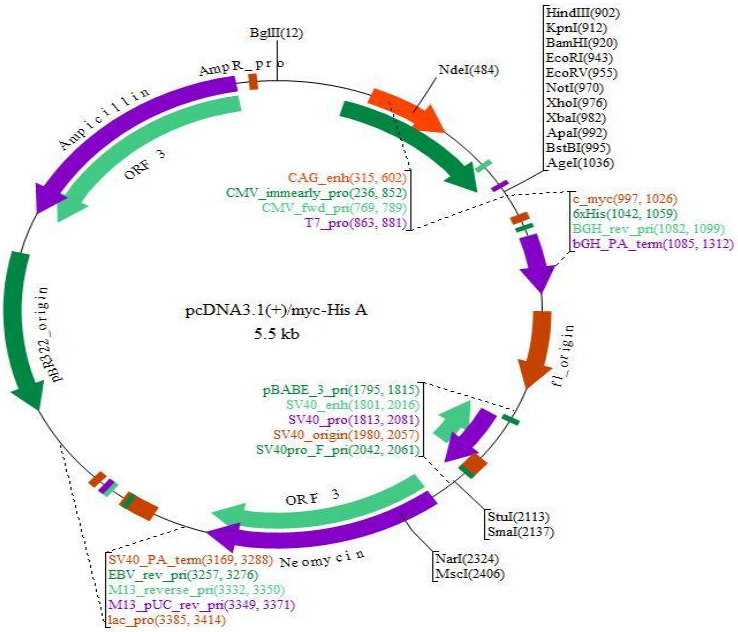
pBluescript II SK (+)

This is a vector with a relative small size that in this work will be used to mutate NCU-G1.



pcDNA 3.1 (+)/ myc-His A vector

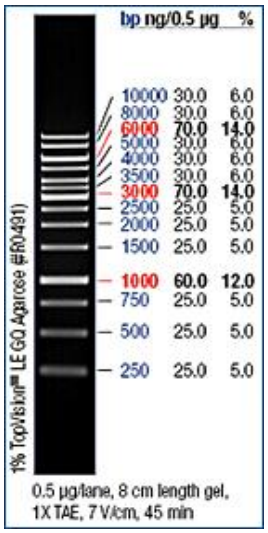
This is a mammalian expression vector that in this is study will be used to express NCU-G1.



7.3 Standards

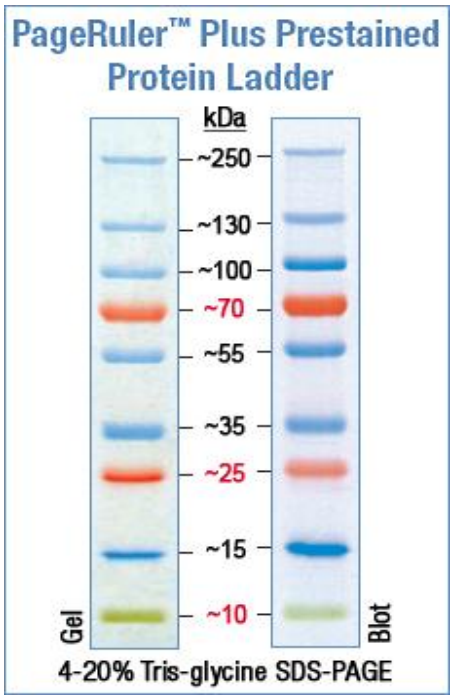
1 kb DNA ladder

This marker will be used to determine DNA size on gel electrophoresis.



Prestained protein ladder

This marker will be used to determine protein size in SDS PAGE.



7.4 Protein measurement

Absorbance at 595 nm of BSA protein standards with concentrations in the range 5-20 µg was measured with each protein measurement batch. Absorbance was measured twice and the mean then calculated for each standard (STD). Based on the means of the STDs a standard curve was constructed. Table 7-1 is an example of STDs readings from one batch, for which the STD curve is shown in figure 7-1.

STD concentration	**Mean of A_{595}**
5	0.109
10	0.31
15	0.488
20	0.714

Table 7-1 A_{595} of BSA protein standards in the concentration range 5-20 µg.

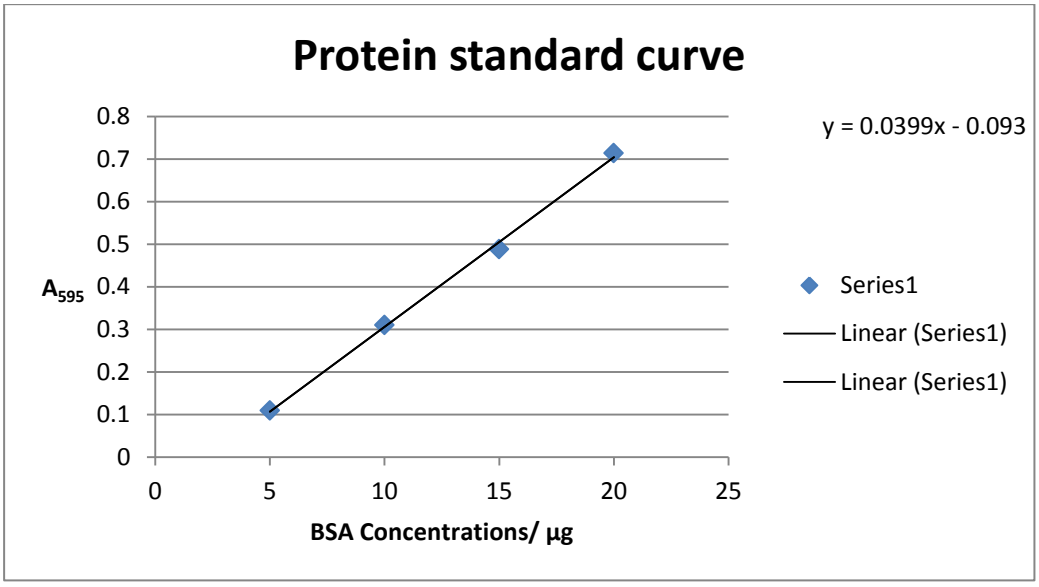


Figure 7-1 protein standard curve of serial BSA standards.

Cells protein harvested using RIPA lysis buffer was diluted 1:400 and 1:800 prior measuring of absorbance at 595 nm. Tests were measured twice and the mean then calculated. Samples concentrations were determinated from the STD curve, and then multiplied by the dilution factor. Table 7-2 shows the mean samples absorbance and calculated concentrations. These samples from the same batch of standards on table 7-1.

Sample	Mean A_{595} [a]	Sample concentration / µg [b]
Not transfected (NT)	0.28	3.7
Wild type (WT)	0.308	4
Mutation 65 (mut65)	0.211	3
Mutation 230 (mut230)	0.252	3.4

Table 7-2 Samples concentrations of the same batch for STDs in table 7-1.

[a] The mean A_{595} is calculated by [(the mean A_{595} for dilution 1: 400) + (the mean A_{595} for dilution 1:800 x 2)] /2.

[b] Sample concentration in µg is calculated by finding the concentration of the A_{595} from the STD curve then multiply by 400.

7.5 Sequences

7.5.1 Alignments of both mutated NCU-G1 versions against wild type NCU-G1

Here we used CLUSTALW (1.83) multiple sequence alignment. The Wild type NCU-G1 coding sequence (CDS) is taken from the NCBI database, with the accession number BC018757. The Stop codon TAA was taken out.

Alignment of mutation 65 amino acid against wild type NCU-G1

```
WT              ATGCGCGGCTCTGTGGAGTGCACCTGGGGTTGGGGGCACTGTGCCCCCAGCCCCCTGCTC
mut65           ATGCGCGGCTCTGTGGAGTGCACCTGGGGTTGGGGGCACTGTGCCCCCAGCCCCCTGCTC
                ************************************************************

WT              CTTTGGACTCTACTTCTGTTTGCAGCCCCATTTGGCCTGCTGGGGGAGAAGACCCGCCAG
mut65           CTTTGGACTCTACTTCTGTTTGCAGCCCCATTTGGCCTGCTGGGGGAGAAGACCCGCCAG
                ************************************************************

WT              GTGTCTCTGGAGGTCATCCCTAACTGGCTGGGCCCCCTGCAGAACCTGCTTCATATACGG
mut65           GTGTCTCTGGAGGTCATCCCTAACTGGCTGGGCCCCCTGCAGAACCTGCTTCATATACGG
                ************************************************************

WT              GCAGTGGGCACCAATTCCACACTGCACTATGTGTGGAGCAGCCTGGGGCCTCTGGCAGTG
mut65           GCAGTGGGCACCGCTTCCACACTGCACTATGTGTGGAGCAGCCTGGGGCCTCTGGCAGTG
                ************..************************************************

WT              GTAATGGTGGCCACCAACACCCCCCACAGCACCCTGAGCGTCAACTGGAGCCTCCTGCTA
mut65           GTAATGGTGGCCACCAACACCCCCCACAGCACCCTGAGCGTCAACTGGAGCCTCCTGCTA
                ************************************************************

WT              TCCCCTGAGCCCGATGGGGGCCTGATGGTGCTCCCTAAGGACAGCATTCAGTTTTCTTCT
mut65           TCCCCTGAGCCCGATGGGGGCCTGATGGTGCTCCCTAAGGACAGCATTCAGTTTTCTTCT
                ************************************************************
```

```
WT              GCCCTTGTTTTTACCAGGCTGCTTGAGTTTGACAGCACCAACGTGTCCGATACGGCAGCA
mut65           GCCCTTGTTTTTACCAGGCTGCTTGAGTTTGACAGCACCAACGTGTCCGATACGGCAGCA
                ************************************************************

WT              AAGCCTTTGGGAAGACCATATCCTCCATACTCCTTGGCCGATTTCTCTTGGAACAACATC
mut65           AAGCCTTTGGGAAGACCATATCCTCCATACTCCTTGGCCGATTTCTCTTGGAACAACATC
                ************************************************************

WT              ACTGATTCATTGGATCCTGCCACCCTGAGTGCCACATTTCAAGGCCACCCCATGAACGAC
mut65           ACTGATTCATTGGATCCTGCCACCCTGAGTGCCACATTTCAAGGCCACCCCATGAACGAC
                ************************************************************

WT              CCTACCAGGACTTTTGCCAATGGCAGCCTGGCCTTCAGGGTCCAGGCCTTTTCCAGGTCC
mut65           CCTACCAGGACTTTTGCCAATGGCAGCCTGGCCTTCAGGGTCCAGGCCTTTTCCAGGTCC
                ************************************************************

WT              AGCCGACCAGCCCAACCCCCTCGCCTCCTGCACACAGCAGACACCTGTCAGCTAGAGGTG
mut65           AGCCGACCAGCCCAACCCCCTCGCCTCCTGCACACAGCAGACACCTGTCAGCTAGAGGTG
                ************************************************************

WT              GCCCTGATTGGAGCCTCTCCCCGGGGAAACCGTTCCCTGTTTGGGCTGGAGGTAGCCACA
mut65           GCCCTGATTGGAGCCTCTCCCCGGGGAAACCGTTCCCTGTTTGGGCTGGAGGTAGCCACA
                ************************************************************

WT              TTGGGCCAGGGCCCTGACTGCCCCTCAATGCAGGAGCAGCACTCCATCGACGATGAATAT
mut65           TTGGGCCAGGGCCCTGACTGCCCCTCAATGCAGGAGCAGCACTCCATCGACGATGAATAT
                ************************************************************

WT              GCACCGGCCGTCTTCCAGTTGGACCAGCTACTGTGGGGCTCCCTCCCATCAGGCTTTGCA
mut65           GCACCGGCCGTCTTCCAGTTGGACCAGCTACTGTGGGGCTCCCTCCCATCAGGCTTTGCA
                ************************************************************

WT              CAGTGGCGACCAGTGGCTTACTCCCAGAAGCCGGGGGGCCGAGAATCAGCCCTGCCCTGC
mut65           CAGTGGCGACCAGTGGCTTACTCCCAGAAGCCGGGGGGCCGAGAATCAGCCCTGCCCTGC
                ************************************************************

WT              CAAGCTTCCCCTCTTCATCCTGCCTTAGCATACTCTCTTCCCCAGTCACCCATTGTCCGA
mut65           CAAGCTTCCCCTCTTCATCCTGCCTTAGCATACTCTCTTCCCCAGTCACCCATTGTCCGA
                ************************************************************

WT              GCCTTCTTTGGGTCCCAGAATAACTTCTGTGCCTTCAATCTGACGTTCGGGGCTTCCACA
mut65           GCCTTCTTTGGGTCCCAGAATAACTTCTGTGCCTTCAATCTGACGTTCGGGGCTTCCACA
                ************************************************************

WT              GGCCCTGGCTATTGGGACCAACACTACCTCAGCTGGTCGATGCTCCTGGGTGTGGGCTTC
mut65           GGCCCTGGCTATTGGGACCAACACTACCTCAGCTGGTCGATGCTCCTGGGTGTGGGCTTC
                ************************************************************

WT              CCTCCAGTGGACGGCTTGTCCCCACTAGTCCTGGGCATCATGGCAGTGGCCCTGGGTGCC
mut65           CCTCCAGTGGACGGCTTGTCCCCACTAGTCCTGGGCATCATGGCAGTGGCCCTGGGTGCC
                ************************************************************

WT              CCAGGGCTCATGCTGCTAGGGGGCGGCTTGGTTCTGCTGCTGCACCACAAGAAGTACTCA
mut65           CCAGGGCTCATGCTGCTAGGGGGCGGCTTGGTTCTGCTGCTGCACCACAAGAAGTACTCA
                ************************************************************

WT              GAGTACCAGTCCATAAATTAA
6mut5           GAGTACCAGTCCATAAAT---
                ******************
```

Alignment of mutation 230 amino acid against wild type NCU-G1

```
WT              ATGCGCGGCTCTGTGGAGTGCACCTGGGGTTGGGGGCACTGTGCCCCCAGCCCCCTGCTC
mut230          ATGCGCGGCTCTGTGGAGTGCACCTGGGGTTGGGGGCACTGTGCCCCCAGCCCCCTGCTC
                ************************************************************

WT              CTTTGGACTCTACTTCTGTTTGCAGCCCCATTTGGCCTGCTGGGGGAGAAGACCCGCCAG
mut230          CTTTGGACTCTACTTCTGTTTGCAGCCCCATTTGGCCTGCTGGGGGAGAAGACCCGCCAG
                ************************************************************

WT              GTGTCTCTGGAGGTCATCCCTAACTGGCTGGGCCCCCTGCAGAACCTGCTTCATATACGG
mut230          GTGTCTCTGGAGGTCATCCCTAACTGGCTGGGCCCCCTGCAGAACCTGCTTCATATACGG
                ************************************************************

WT              GCAGTGGGCACCAATTCCACACTGCACTATGTGTGGAGCAGCCTGGGGCCTCTGGCAGTG
mut230          GCAGTGGGCACCAATTCCACACTGCACTATGTGTGGAGCAGCCTGGGGCCTCTGGCAGTG
                ************************************************************

WT              GTAATGGTGGCCACCAACACCCCCCACAGCACCCTGAGCGTCAACTGGAGCCTCCTGCTA
mut230          GTAATGGTGGCCACCAACACCCCCCACAGCACCCTGAGCGTCAACTGGAGCCTCCTGCTA
                ************************************************************

WT              TCCCCTGAGCCCGATGGGGGCCTGATGGTGCTCCCTAAGGACAGCATTCAGTTTTCTTCT
mut230          TCCCCTGAGCCCGATGGGGGCCTGATGGTGCTCCCTAAGGACAGCATTCAGTTTTCTTCT
                ************************************************************

WT              GCCCTTGTTTTTACCAGGCTGCTTGAGTTTGACAGCACCAACGTGTCCGATACGGCAGCA
mut230          GCCCTTGTTTTTACCAGGCTGCTTGAGTTTGACAGCACCAACGTGTCCGATACGGCAGCA
                ************************************************************

WT              AAGCCTTTGGGAAGACCATATCCTCCATACTCCTTGGCCGATTTCTCTTGGAACAACATC
mut230          AAGCCTTTGGGAAGACCATATCCTCCATACTCCTTGGCCGATTTCTCTTGGAACAACATC
                ************************************************************

WT              ACTGATTCATTGGATCCTGCCACCCTGAGTGCCACATTTCAAGGCCACCCCATGAACGAC
mut230          ACTGATTCATTGGATCCTGCCACCCTGAGTGCCACATTTCAAGGCCACCCCATGAACGAC
                ************************************************************

WT              CCTACCAGGACTTTTGCCAATGGCAGCCTGGCCTTCAGGGTCCAGGCCTTTTCCAGGTCC
230mut          CCTACCAGGACTTTTGCCAATGGCAGCCTGGCCTTCAGGGTCCAGGCCTTTTCCAGGTCC
                ************************************************************

WT              AGCCGACCAGCCCAACCCCCTCGCCTCCTGCACACAGCAGACACCTGTCAGCTAGAGGTG
mut230          AGCCGACCAGCCCAACCCCCTCGCCTCCTGCACACAGCAGACACCTGTCAGCTAGAGGTG
                ************************************************************

WT              GCCCTGATTGGAGCCTCTCCCCGGGGAAACCGTTCCCTGTTTGGGCTGGAGGTAGCCACA
mut230          GCCCTGATTGGAGCCTCTCCCCGGGGAGCCCGTTCCCTGTTTGGGCTGGAGGTAGCCACA
                ****************************..********************************

WT              TTGGGCCAGGGCCCTGACTGCCCCTCAATGCAGGAGCAGCACTCCATCGACGATGAATAT
mut230          TTGGGCCAGGGCCCTGACTGCCCCTCAATGCAGGAGCAGCACTCCATCGACGATGAATAT
                ************************************************************

WT              GCACCGGCCGTCTTCCAGTTGGACCAGCTACTGTGGGGCTCCCTCCCATCAGGCTTTGCA
mut230          GCACCGGCCGTCTTCCAGTTGGACCAGCTACTGTGGGGCTCCCTCCCATCAGGCTTTGCA
                ************************************************************

WT              CAGTGGCGACCAGTGGCTTACTCCCAGAAGCCGGGGGGCCGAGAATCAGCCCTGCCCTGC
mut230          CAGTGGCGACCAGTGGCTTACTCCCAGAAGCCGGGGGGCCGAGAATCAGCCCTGCCCTGC
                ************************************************************

WT              CAAGCTTCCCCTCTTCATCCTGCCTTAGCATACTCTCTTCCCCAGTCACCCATTGTCCGA
```

```
mut230          CAAGCTTCCCCTCTTCATCCTGCCTTAGCATACTCTCTTCCCCAGTCACCCATTGTCCGA
                ************************************************************

WT              GCCTTCTTTGGGTCCCAGAATAACTTCTGTGCCTTCAATCTGACGTTCGGGGCTTCCACA
mut230          GCCTTCTTTGGGTCCCAGAATAACTTCTGTGCCTTCAATCTGACGTTCGGGGCTTCCACA
                ************************************************************

WT              GGCCCTGGCTATTGGGACCAACACTACCTCAGCTGGTCGATGCTCCTGGGTGTGGGCTTC
mut230          GGCCCTGGCTATTGGGACCAACACTACCTCAGCTGGTCGATGCTCCTGGGTGTGGGCTTC
                ************************************************************

WT              CCTCCAGTGGACGGCTTGTCCCCACTAGTCCTGGGCATCATGGCAGTGGCCCTGGGTGCC
mut230          CCTCCAGTGGACGGCTTGTCCCCACTAGTCCTGGGCATCATGGCAGTGGCCCTGGGTGCC
                ************************************************************

WT              CCAGGGCTCATGCTGCTAGGGGGCGGCTTGGTTCTGCTGCTGCACCACAAGAAGTACTCA
mut230          CCAGGGCTCATGCTGCTAGGGGGCGGCTTGGTTCTGCTGCTGCACCACAAGAAGTACTCA
                ************************************************************

WT              GAGTACCAGTCCATAAATTAA
mut230          GAGTACCAGTCCATAAAT---

                ******************
```

7.5.2 Aligment of the linkage between NCU-G1 and pcDNA 3.1 (+)/ myc-His A

N terminus:

```
wt    AAG CTT GGT ACC ATG CGC GGC TCT
      ||| ||| ||| ||| ||| ||| ||| |||
seq   AAG CTT GGT ACC ATG CGC GGC TCT
```

This nucleotide alignment shows the linkage between NCU-G1 N-terminus and pcDNA 3.1 (+)/ myc-His A vector. Kpn I cutting site on pcDNA vector is represented by blue colour. The start codon of NCU-G1 marked with red, and the rest of the begining nucleotides of NCUG1 marked with green colour.

C terminus:

```
Wt   ATA AAT TAA CTC GAG TCT AGA GGG CCC TTC GAA CAA AAA CTC ATC TCA GAA
     ||| |||     ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
seq  ATA AAT --- CTC GAG TCT AGA GGG CCC TTC GAA CAA AAA CTC ATC TCA GAA

GAG GAT CTG AAT ATG CAT ACC GGT CAT CAT CAC CAT CAC CAT TGA GTT
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
GAG GAT CTG AAT ATG CAT ACC GGT CAT CAT CAC CAT CAC CAT TGA GTT
```

This nucleotide alignment shows the linking between NCU-G1 C-terminus and pcDNA 3.1 (+)/ myc His A vector. The last nucleotides on NCU-G1 marked with red (stop codon is

removed). XhoI cutting site on pcDNA vector is represented by blue color, nucleotides code for the myc epitope marked in green, and those code for the His epitope marked in yellow.

7.5.3 Primer sequences

Primer name	**Melting-temp**	**Sequence 5'-3'**
NCU-G1 sequencing primer nr 1 F	53.8	CCTCTGGCAGTGGTAATGGT
NCU-G1 sequencing primer nr 1 R	51.8	GGTGCTGTCAAACTCAAGCA
NCU-G1 sequencing primer nr 2 F	58.3	ACGGCTTGTCCCCACTAGTCC
NCU-G1 sequencing primer nr 2 R	57.9	TCCAGAGACACCTGGCGGGT
NCU-G1 sequencing primer nr 3 F	55.4	GGAGCAGCACTCCATCGAC
NCU-G1 sequencing primer nr 3 R	53.8	TCGGCCAAGGAGTATGGAGGA
NCU-G1 sequencing primer nr 4 F	62.4	CAGGTACCATGCGCGGCTCTGTG
NCU-G1 sequencing primer nr 4 R	60.1	CACTCGAGATTTATGGACTGGTACTCTGA
NCU-G1 sequencing primer nr 5 F	51.8	GCTTGAGTTTGACAGCACCA
NCU-G1 sequencing primer nr 5 R	51.8	TTGGCAAAAGTCCTGGTAGG
NCU-G1 mut 65 primer F	78.5	CGGGCAGTGGGCACCGCTTCCACACTGCACTA
NCU-G1 mut 65 primer R	78.5	TAGTGCAGTGTGGAAGCGGTGCCCACTGCCCG
NCU-G1 mut 230 primer F	78.07	GCCTCTCCCCGGGGAGCCCGTTCCCTGT
NCU-G1 mut 230 primer R	78.07	ACAGGGAACGGGCTCCCCGGGGAGAGGC

Note: For mutagenic primers, bases that carry the mutation are underlined.

References

1. Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S., and Brunak, S. *Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence.* Proteomics, 2004. **4**: 1633-49.
2. Vasan, S., Zhang, X., Kapurniotu, A., Bernhagen, J., Teichberg, S., Basgen, J., Wagle, D., Shih, D., Terlecky, I., Bucala, R., Cerami, A., Egan, J., and Ulrich, P. *An agent cleaving glucose-derived protein crosslinks in vitro and in vivo.* Nature, 1996. **382**: 275-8.
3. Shental-Bechor, D. and Levy, Y. *Folding of glycoproteins: toward understanding the biophysics of the glycosylation code.* Current opinion in structural biology, 2009. **19**: 524-33.
4. Apweiler, R., Hermjakob, H., and Sharon, N. *On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database.* Biochimica et Biophysica Acta, 1999. **1473**: 4-8.
5. Varki, A., Esko, J.D., and Colley, K.J. *Cellular Organization of Glycosylation*, in *Essentials of Glycobiology*. Varki, A., Cummings R.D., Esko J.D., Freeze H.H., Stanley P., Bertozzi C.R., Hart G.W., and Etzler M.E. 2009: Cold Spring Harbor (NY).
6. Ruddock, L.W. and Molinari, M. *N-glycan processing in ER quality control.* Journal of Cell Science, 2006. **119**: 4373-80.
7. Freeze, H.H., Esko, J.D., and Parodi, A.J. *Glycans in Glycoprotein Quality Control*, in *Essentials of Glycobiology*. Varki, A., Cummings R.D., Esko J.D., Freeze H.H., Stanley P., Bertozzi C.R., Hart G.W., and Etzler M.E. 2009: Cold Spring Harbor (NY).
8. Zhang, J.X., Braakman, I., Matlack, K.E.S., and Helenius, A. *Quality control in the secretory pathway: The role of calreticulin, calnexin and BiP in the retention of glycoproteins with C-terminal truncations.* Molecular biology of the cell, 1997. **8**: 1943-54.
9. Shental-Bechor, D. and Levy, Y. *Effect of glycosylation on protein folding: A close look at thermodynamic stabilization.* Proceedings of the National Academy of Sciences USA, 2008. **105**: 8256-61.
10. Zhou, F., Xu, W., Hong, M., Pan, Z., Sinko, P.J., Ma, J., and You, G. *The Role of N-Linked Glycosylation in Protein Folding, Membrane Targeting, and Substrate Binding of Human Organic Anion Transporter hOAT4.* Molecular Pharmacology, 2005. **67**: 868-76.
11. Hammond, C., Braakman, I., and Helenius, A. *Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control.* Proceedings of the National Academy of Sciences USA, 1994. **91**: 913-7.
12. Shi, X. and Elliott, R.M. *Analysis of N-linked glycosylation of hantaan virus glycoproteins and the role of oligosaccharide side chains in protein folding and intracellular trafficking.* Journal of virology, 2004. **78**: 5414-22.
13. Ermonval, M., Duvet, S., Zonneveld, D., Cacan, R., Buttin, G., and Braakman, I. *Truncated N-glycans affect protein folding in the ER of CHO-derived mutant cell lines without preventing calnexin binding.* Glycobiology, 2000. **10**: 77-87.
14. Hanson, S.R., Culyba, E.K., Hsu, T.-L., Wong, C.-H., Kelly, J.W., and Powers, E.T. *The core trisaccharide of an N-linked glycoprotein intrinsically accelerates folding and enhances stability.* Proceedings of the National Academy of Sciences USA, 2009. **106**: 3131-6.

15. Doucette, M.M. and Stevens, V.L. *Point Mutations Alter the Cellular Distribution of the Human Folate Receptor in Cultured Chinese Hamster Ovary Cells.* The Journal of Nutrition, 2004. **134**: 308-16.
16. Atiya-Nasagi, Y., Cohen, H., Medalia, O., Fukuda, M., and Sagi-Eisenberg, R. *O-glycosylation is essential for intracellular targeting of synaptotagmins I and II in non-neuronal specialized secretory cells.* Journal of Cell Science, 2005. **118**: 1363-72.
17. Li, H. and Pajor, A.M. *Mutagenesis of the N-glycosylation site of hNaSi-1 reduces transport activity.* American journal of physiology. Cell physiology, 2003. **285**: C1188-96.
18. Cai, G., Salonikidis, P.S., Fei, J., Schwarz, W., Schulein, R., Reutter, W., and Fan, H. *The role of N-glycosylation in the stability, trafficking and GABA-uptake of GABA-transporter 1. Terminal N-glycans facilitate efficient GABA-uptake activity of the GABA transporter.* The FEBS journal, 2005. **272**: 1625-38.
19. Kundra, R. and Kornfeld, S. *Asparagine-linked Oligosaccharides Protect Lamp-1 and Lamp-2 from Intracellular Proteolysis.* Journal of Biological Chemistry, 1999. **274**: 31039-46.
20. Lodish, H., Berk, A., Kaiser, C., Kriegar, M., Scott, M., Bretscher, A., Ploegh, H., and Matsudaira. *Moving proteins into membranes and organells*, in *Molecular cell biology*. 2008. 549-51.
21. Bieberich, E., Tencomnao, T., Kapitonov, D., and Yu, R.K. *Effect of N-Glycosylation on Turnover and Subcellular Distribution of N-Acetylgalactosaminyltransferase I and Sialyltransferase II in Neuroblastoma Cells.* Journal of Neurochemistry, 2000. **74**: 2359-64.
22. Tams, J.W., Vind, J., and Welinder, K.G. *Adapting protein solubility by glycosylation. N-glycosylation mutants of Coprinus cinereus peroxidase in salt and organic solutions.* Biochimica et Biophysica Acta, 1999. **1432**: 214-21.
23. Bunkenborg, J., Pilch, B.J., Podtelejnikov, A.V., and Wisniewski, J.R. *Screening for N-glycosylated proteins by liquid chromatography mass spectrometry.* Proteomics, 2004. **4**: 454-65.
24. Lee, A., Kolarich, D., Haynes, P.A., Jensen, P.H., Baker, M.S., and Packer, N.H. *Rat liver membrane glycoproteome: enrichment by phase partitioning and glycoprotein capture.* Journal of proteome research, 2009. **8**: 770-81.
25. Tian, E. and Ten Hagen, K.G. *Recent insights into the biological roles of mucin-type O-glycosylation.* Glycoconjugate journal, 2009. **26**: 325-34.
26. Kim, P.J., Lee, D.Y., and Jeong, H. *Centralized modularity of N-linked glycosylation pathways in mammalian cells.* PloS one, 2009. **4**: e7317.
27. Medzihradszky, K.F. *Characterization of protein N-glycosylation.* Methods in enzymology, 2005. **405**: 116-38.
28. Ruiz-Canada, C., Kelleher, D.J., and Gilmore, R. *Cotranslational and Posttranslational N-Glycosylation of Polypeptides by Distinct Mammalian OST Isoforms.* Cell, 2009. **136**: 272-83.
29. Stanley, P., Schachter, H., and Taniguchi, N. *N-Glycans*, in *Essentials of Glycobiology*. Varki, A., Cummings R.D., Esko J.D., Freeze H.H., Stanley P., Bertozzi C.R., Hart G.W., and Etzler M.E. 2009: Cold Spring Harbor (NY).
30. Bolt, G., Kristensen, C., and Steenstrup, T.D. *Posttranslational N-glycosylation takes place during the normal processing of human coagulation factor VII.* Glycobiology, 2005. **15**: 541-47.
31. Duvet, S., Op De Beeck, A., Cocquerel, L., Wychowski, C., Cacan, R., and Dubuisson, J. *Glycosylation of the hepatitis C virus envelope protein E1 occurs*

posttranslationally in a mannosylphosphoryldolichol-deficient CHO mutant cell line. Glycobiology, 2002. **12**: 95-101.
32. Liu, T., Qian, W.J., Gritsenko, M.A., Camp, D.G., 2nd, Monroe, M.E., Moore, R.J., and Smith, R.D. *Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry.* Journal of proteome research, 2005. **4**: 2070-80.
33. Zhang, H., Li, X.J., Martin, D.B., and Aebersold, R. *Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry.* Nature biotechnology, 2003. **21**: 660-6.
34. Wollscheid, B., Bausch-Fluck, D., Henderson, C., O'Brien, R., Bibel, M., Schiess, R., Aebersold, R., and Watts, J.D. *Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins.* Nature biotechnology, 2009. **27**: 378-86.
35. Gavel, Y. and von Heijne, G. *Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering.* Protein engineering, 1990. **3**: 433-42.
36. Zielinska, D.F., Gnad, F., Wisniewski, J.R., and Mann, M. *Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints.* Cell, 2010. **141**: 897-907.
37. Cui, J., Smith, T., Robbins, P.W., and Samuelson, J. *Darwinian selection for sites of Asn-linked glycosylation in phylogenetically disparate eukaryotes and viruses.* Proceedings of the National Academy of Sciences USA, 2009. **106**: 13421-26.
38. Valliere-Douglass, J.F., Kodama, P., Mujacic, M., Brady, L.J., Wang, W., Wallace, A., Yan, B., Reddy, P., Treuheit, M.J., and Balland, A. *Asparagine-linked Oligosaccharides Present on a Non-consensus Amino Acid Sequence in the CH1 Domain of Human Antibodies.* Journal of Biological Chemistry, 2009. **284**: 32493-506.
39. Breuer, W., Klein, R.A., Hardt, B., Bartoschek, A., and Bause, E. *Oligosaccharyltransferase is highly specific for the hydroxy amino acid in Asn-Xaa-Thr/Ser.* FEBS Letters, 2001. **501**: 106-10.
40. Knauer, R. and Lehle, L. *The oligosaccharyltransferase complex from yeast.* Biochimica et Biophysica Acta, 1999. **1426**: 259-73.
41. Yan, Q., Prestwich, G.D., and Lennarz, W.J. *The Ost1p Subunit of Yeast Oligosaccharyl Transferase Recognizes the Peptide Glycosylation Site Sequence, -Asn-X-Ser/Thr.* Journal of Biological Chemistry, 1999. **274**: 5021-5.
42. Yan, Q. and Lennarz, W.J. *Oligosaccharyltransferase: a complex multisubunit enzyme of the endoplasmic reticulum.* Biochemical and biophysical research communications, 1999. **266**: 684-9.
43. Yan, Q. and Lennarz, W.J. *Studies on the Function of Oligosaccharyl Transferase Subunits.* Journal of Biological Chemistry, 2002. **277**: 47692-700.
44. Yan, Q. and Lennarz, W.J. *Studies on the function of oligosaccharyl transferase subunits: a glycosylatable photoprobe binds to the luminal domain of Ost1p.* Proceedings of the National Academy of Sciences USA, 2002. **99**: 15994-9.
45. Li, G., Yan, Q., Nita-Lazar, A., Haltiwanger, R.S., and Lennarz, W.J. *Studies on the N-glycosylation of the subunits of oligosaccharyl transferase in Saccharomyces cerevisiae.* The Journal of biological chemistry, 2005. **280**: 1864-71.
46. Schulz, B.L. and Aebi, M. *Analysis of glycosylation site occupancy reveals a role for Ost3p and Ost6p in site-specific N-glycosylation efficiency.* Molecular & cellular proteomics : MCP, 2009. **8**: 357-64.

47. Roth, J., Wang, Y., Eckhardt, A.E., and Hill, R.L. *Subcellular localization of the UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-mediated O-glycosylation reaction in the submaxillary gland.* Proceedings of the National Academy of Sciences USA, 1994. **91**: 8935-9.
48. Brockhausen, I., Schachter, H., and Stanley, P. *O-GalNAc Glycans*, in *Essentials of Glycobiology*. Varki, A., Cummings R.D., Esko J.D., Freeze H.H., Stanley P., Bertozzi C.R., Hart G.W., and Etzler M.E. 2009: Cold Spring Harbor (NY).
49. Hansen, J.E., Lund, O., Tolstrup, N., Gooley, A.A., Williams, K.L., and Brunak, S. *NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility.* Glycoconjugate journal, 1998. **15**: 115-30.
50. Caragea, C., Sinapov, J., Silvescu, A., Dobbs, D., and Honavar, V. *Glycosylation site prediction using ensembles of Support Vector Machine classifiers.* BMC bioinformatics, 2007. **8**: 438.
51. Hamby, S. and Hirst, J. *Prediction of glycosylation sites using random forests.* BMC bioinformatics, 2008. **9**: 500.
52. Kawamura, T., Kuroda, N., Kimura, Y., Lazoura, E., Okada, N., and Okada, H. *cDNA of a Novel mRNA Expressed Predominantly in Mouse Kidney.* Biochemical Genetics, 2001. **39**: 33-42.
53. Schroder, B., Wrocklage, C., Pan, C., Jager, R., Kosters, B., Schafer, H., Elsasser, H.P., Mann, M., and Hasilik, A. *Integral and associated lysosomal membrane proteins.* Traffic, 2007. **8**: 1676-86.
54. Steffensen, K., Bouzga, M., Skjeldal, F., Kasi, C., Karahasan, A., Matre, V., Bakke, O., Guerin, S., and Eskild, W. *Human NCU-G1 can function as a transcription factor and as a nuclear receptor co-activator.* BMC Molecular Biology, 2007. **8**: 106.
55. Schieweck, O., Damme, M., Schröder, B., Hasilik, A., Schmidt, B., and Lübke, T. *NCU-G1 is a highly glycosylated integral membrane protein of the lysosome.* Biochemical Journal, 2009. **422**: 83-90.
56. Sardiello, M., Palmieri, M., di Ronza, A., Medina, D.L., Valenza, M., Gennarino, V.A., Di Malta, C., Donaudy, F., Embrione, V., Polishchuk, R.S., Banfi, S., Parenti, G., Cattaneo, E., and Ballabio, A. *A gene network regulating lysosomal biogenesis and function.* Science, 2009. **325**: 473-7.
57. Gibbs, R.A. *DNA amplification by the polymerase chain reaction.* Analytical chemistry, 1990. **62**: 1202-14.
58. Voytas, D. *Agarose gel electrophoresis.* Current Protocols in Protein Science. A.4F.1–A.4F.3., 2001.
59. Reddy, M. *Positive selection system for identification of recombinants using alpha-complementation plasmids.* BioTechniques, 2004. **37**: 948-52.
60. Sanger, F., Nicklen, S., and Coulson, A.R. *DNA sequencing with chain-terminating inhibitors.* Proceedings of the National Academy of Sciences USA, 1977. **74**: 5463-7.
61. Lodish, H., Lodish, H., Berk, A., Kaiser, C., Kriegar, M., Scott, M., Bretscher, A., Ploegh, H., and Matsudaira. *Molecular genetic techniques*, in *molecular cell biology*. 2008. 176-97.
62. Hawley-Nelson, P., Ciccarone, V., and Moore, M.L. *Transfection of cultured eukaryotic cells using cationic lipid reagents*, in *Current protocols in molecular biology* 2008. 81:9.4.1–9.4.17.
63. Ruozi, B., Forni, F., Battini, R., and Vandelli, M.A. *Cationic liposomes for gene transfection.* Journal of drug targeting, 2003. **11**: 407-14.
64. Watanabe, I., Zhu, J., Recio-Pinto, E., and Thornhill, W.B. *Glycosylation affects the protein stability and cell surface expression of Kv1.4 but Not Kv1.1 potassium*

channels. A pore region determinant dictates the effect of glycosylation on trafficking. The Journal of biological chemistry, 2004. **279**: 8879-85.
65. Xiao, Y.P., Morice, A.H., Compton, S.J., and Sadofsky, L. *N-linked glycosylation regulates human proteinase-activated receptor-1 cell surface expression and disarming via neutrophil proteinases and thermolysin.* The Journal of biological chemistry, 2011. **286**: 22991-3002.
66. Tiwari, R.K., Suriano, R., Ghosh, S.K., Ashok, B.T., Mittelman, A., Chen, Y.G., and Banerjee, A. *Differences in glycosylation patterns of heat shock protein, gp96: Implications for prostate cancer prevention.* Cancer research, 2005. **65**: 6466-75.
67. Kothari, H., Rao, L.V., and Pendurthi, U.R. *Glycosylation of tissue factor is not essential for its transport or functions.* Journal of thrombosis and haemostasis : JTH, 2011. **9**: 1511-20.
68. Yazaki, M., Tagawa, K., Maruyama, K., Sorimachi, H., Tsuchiya, T., Ishiura, S., and Suzuki, K. *Mutation of potential N-linked glycosylation sites in the Alzheimer's disease amyloid precursor protein (APP).* Neuroscience letters, 1996. **221**: 57-60.
69. Pelis, R.M., Suhre, W.M., and Wright, S.H. *Functional influence of N-glycosylation in OCT2-mediated tetraethylammonium transport.* American journal of physiology. Renal physiology, 2006. **290**: F1118-26.
70. Lecca, M.R., Maag, C., Berger, E.G., and Hennet, T. *Fibrotic response in fibroblasts from congenital disorders of glycosylation.* Journal of cellular and molecular medicine, 2011. **15**: 1788-96.
71. Dupre, T., Barnier, A., de Lonlay, P., Cormier-Daire, V., Durand, G., Codogno, P., and Seta, N. *Defect in N-glycosylation of proteins is tissue-dependent in congenital disorders of glycosylation Ia.* Glycobiology, 2000. **10**: 1277-81.
72. Eklund, E.A., Sun, L., Yang, S.P., Pasion, R.M., Thorland, E.C., and Freeze, H.H. *Congenital disorder of glycosylation Ic due to a de novo deletion and an hALG-6 mutation.* Biochemical and biophysical research communications, 2006. **339**: 755-60.
73. Miyamoto, C., Chizzonite, R., Crowl, R., Rupprecht, K., Kramer, R., Schaber, M., Kumar, G., Poonian, M., and Ju, G. *Molecular-Cloning and Regulated Expression of the Human C-Myc Gene in Escherichia-Coli and Saccharomyces-Cerevisiae - Comparison of the Protein Products.* Proceedings of the National Academy of Sciences USA, 1985. **82**: 7232-6.
74. Liao, D.J. and Dickson, R.B. *c-Myc in breast cancer.* Endocrine-Related Cancer, 2000. **7**: 143-64.
75. Picard, V., Ersdal-Badju, E., and Bock, S.C. *Partial glycosylation of antithrombin III asparagine-135 is caused by the serine in the third position of its N-glycosylation consensus sequence and is responsible for production of the beta-antithrombin III isoform with enhanced heparin affinity.* Biochemistry, 1995. **34**: 8433-40.
76. Nicolaes, G.A., Villoutreix, B.O., and Dahlback, B. *Partial glycosylation of Asn2181 in human factor V as a cause of molecular and functional heterogeneity. Modulation of glycosylation efficiency by mutagenesis of the consensus sequence for N-linked glycosylation.* Biochemistry, 1999. **38**: 13584-91.
77. Chen, C.M., Chiang, S.Y., and Yeh, N.H. *Increased stability of nucleolin in proliferating cells by inhibition of its self-cleaving activity.* The Journal of biological chemistry, 1991. **266**: 7754-8.
78. Carpentier, M., Morelle, W., Coddeville, B., Pons, A., Masson, M., Mazurier, J., and Legrand, D. *Nucleolin undergoes partial N- and O-glycosylations in the extranuclear cell compartment.* Biochemistry, 2005. **44**: 5804-15.

79. Fujimori-Arai, Y., Koyama, I., Hirano, K., and Komoda, T. *Intestinal alkaline phosphatase isoforms in rabbit tissues differ in glycosylation patterns.* Clinical biochemistry, 1997. **30**: 545-51.
80. Bosch, D., Smal, J., and Krebbers, E. *A Trout Growth-Hormone Is Expressed, Correctly Folded and Partially Glycosylated in the Leaves but Not the Seeds of Transgenic Plants.* Transgenic research, 1994. **3**: 304-10.
81. Macara, I.G. *Transport into and out of the nucleus.* Microbiology and Molecular Biology Reviews, 2001. **65**: 570-94.
82. Rondanino, C., Bousser, M.T., Monsigny, M., and Roche, A.C. *Sugar-dependent nuclear import of glycosylated proteins in living cells.* Glycobiology, 2003. **13**: 509-19.