Rita Alexandra Lucas Santana

Degree in Biotechnology

EGFR Glycosylation in Cancer

Dissertation to obtain a Master's degree in Biotechnology

Supervisor: Patrícia Gomes-Alves, PhD, Senior scientist, IBET/ITQB-NOVA
Co-supervisor: Lara Marcos da Silva, PhD, Scientist, IBET/ITQB-NOVA

Jury:

President: Prof Doctor Rui Manuel Freitas Oliveira
Arguer: Prof Doctor Maria Angelina de Sá Palma

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

September 2018

Rita Alexandra Lucas Santana

Degree in Biotechnology

EGFR Glycosylation in Cancer

Dissertation to obtain a Master's degree in Biotechnology,

by Universidade Nova de Lisboa, Faculdade de Ciências e Tecnologia

Supervisor: Patrícia Gomes-Alves, PhD, Senior scientist, IBET/ITQB-NOVA
Co-supervisor: Lara Marcos da Silva, PhD, Scientist, IBET/ITQB-NOVA

Jury:

President: Prof Doctor Rui Manuel Freitas Oliveira
Arguer: Prof Doctor Maria Angelina de Sá Palma

September
2018

EGFR Glycosylation in Cancer

Acknowledgments

I would like to thank all the people that directly or indirectly supported and contributed to the work developed and presented in this MSc thesis, particularly:

To Professor Paula Alves and Professor Manuel Carrondo for giving me the opportunity to do my master thesis at the Animal Cell Technology Unit at ITQB/IBET, for offering the best conditions to work and for being an example of rigor, leadership and professionalism.

To my supervisors Dr. Patrícia Alves and Dr. Lara Silva for the opportunity to work in this challenging project. To Dr. Patrícia Alves for all the guidance, knowledge, helpful discussions, suggestions and encouragement during this thesis and specially to Dr. Lara Silva for the scientific and technical guidance, precious help, productive discussions, encouragement, constant support and for being a strong example of a scientist. I learned a lot this last year, and I am thankful for that.

To the UniMS members for all that I have learned and the willingness to help.

To the 3D Advanced cell models group with whom I have always learned something new and to all Animal Cell Technology Unit colleagues for the good working environment and all the support during this year.

To all the professors that lectured in the MSc in Biotechnology at FCT-UNL for the valuable knowledge they share with their students.

A todos os meus amigos e colegas de mestrado que acreditam em mim e nas minhas capacidades e que me apoiaram e ajudaram sempre, em especial à Patrícia, à Sónia, ao Pedro, ao Nuno, à Ana Teresa, à Inês e ao Tiago. Aprendi sempre algo mais com vocês. Obrigada pela vossa amizade!

Ao Ivo pela companhia, incentivo, compreensão, carinho e por seres um exemplo para mim. Obrigada por tudo!

À minha família pelo apoio incondicional, por acreditarem em mim em todos os momentos e ajudarem em tudo para que eu consiga atingir os meus objetivos. Sem vocês não era possível.

Abstract

Cancer is a leading cause of mortality around the world. Head and neck (HNC) and esophageal cancers (EC) are the sixth and eighth most common cancers worldwide, respectively. A common feature in both diseases is the overexpression of epidermal growth factor receptor (EGFR).

EGFR a transmembrane glycoprotein and cell-surface receptor is mainly involved in regulation of cell proliferation, survival, differentiation and its deregulation is associated with cancer. EGFR is highly glycosylated, and it is known that aberrant glycosylation interferes with its functions.

Several EGFR-targeted therapies have been already developed, though showing some undesirable secondary effects, such as skin toxicity, as normal keratinocytes also express EGFR. Identification of a specific target for tumor cells would be essential to minimize those effects. Since EGFR is glycosylated and aberrant glycosylation is a common feature in cancer, identifying a distinct EGFR glycosylation pattern between cancer and normal cells is a promising approach to improve specificity of targeted therapies.

In this work, primary human keratinocytes, cell lines and patient-derived tissues representative of HNC and EC were used to implement an optimized strategy to obtain EGFR enriched protein fractions from which EGFR glycosylation pattern could be defined. This thesis reports the results attained for the optimization steps on the different methodologies explored for protein extraction both from cells and tissues, EGFR detection and quantification, EGFR enrichment approaches and glycoprofiling methods (lectin ELISA and LC-MS analysis). Although promising strategies are proposed and validated to extract EGFR from different biological sources in adequate conditions for further mass spectrometry (MS) analysis, some difficulties were found in glycans detection by MS. Further work needs to be done to surpass this method sensitivity problem, as EGFR amounts recovered from enrichment protocols are extremely low, adding to the fact that when dealing with patient-derived samples, specimens are limited in size and availability.

Key words: Head and neck cancer; Esophageal cancer; Epidermal growth factor receptor (EGFR); N-Glycosylation; Protein enrichment methodologies; Mass spectrometry analysis.

Resumo

O cancro é das principais causas de mortalidade no mundo. Cancros de cabeça e pescoço (HNC) e de esófago (EC) são o sexto e oitavo mais frequentes no mundo, respetivamente. Tendo em comum a sobre-expressão do recetor do fator de crescimento epidérmico (EGFR).

O EGFR é um recetor transmembranar importante na regulação da proliferação, sobrevivência e diferenciação celular, estando desregulado em muitos cancros. O EGFR é altamente glicosilado e sabe-se que alterações na glicosilação interferem com as suas funções.

Existem atualmente terapias anti-EGFR, porém mostram efeitos secundários indesejáveis, como toxicidade cutânea, pois EGFR também é expresso em queratinócitos. A identificação de alvos específicos em células tumorais seria essencial para minimizar esses efeitos. Dado que EGFR é glicosilado e alterações da glicosilação são comuns em cancro, identificar um padrão de glicosilação de EGFR distinto entre células tumorais e normais poderá melhorar a especificidade das terapias dirigidas.

Neste trabalho, queratinócitos humanos, linhas celulares e tecidos derivados de pacientes de HNC e EC foram usados para otimizar e implementar uma estratégia para obtenção de frações proteicas enriquecidas em EGFR, a partir das quais o padrão de glicosilação de EGFR poderia ser definido. Esta tese relata os resultados obtidos na otimização das diferentes metodologias exploradas para extração proteica a partir de células em cultura e tecidos, para deteção, quantificação e enriquecimento em EGFR e para identificação de perfis de glicoformas de EGFR (ELISA com lectinas e LC-MS). Embora sejam propostas e validadas estratégias promissoras para extração de EGFR de diferentes fontes biológicas em condições adequadas para posterior análise por espectrometria de massa (MS), foram encontradas dificuldades na deteção de glicanos por MS. Assim, serão necessários mais estudos para aumentar a sensibilidade do método, pois as quantidades de EGFR recuperadas são extremamente baixas e as amostras de pacientes são limitadas em tamanho e disponibilidade.

Palavras-chave: Cancro de cabeça e pescoço; Cancro de esôfago; Recetor do fator de crescimento epidérmico (EGFR); N-glicosilação; Metodologias de enriquecimento de proteínas; Análise de espectrometria de massa.

List of contents

List of Figures

List of Tables

Abbreviations

AAL	*Aleuria aurantia* lectin
ACN	Acetonitrile
AFP	Alpha-fetoprotein
ANXA2	Annexin A2
AmBic	Ammonium bicarbonate
AS ODNs	Antisense oligodeoxynucleotides
Asn	Asparagine
Asp	Aspartic acid
ATP	Adenosine 5' triphosphate
BCA	Bicinchoninic acid assay
BSA	Bovine serum albumin
CBMs	Carbohydrate-binding molecules
CR	Cysteine rich
DAD1	Dolichyl-diphosphooligosaccharide
DNA	Deoxyribonucleic acid
DMEM	Dulbecco's Modified Eagle Medium
Dol-P	Dolichol phosphate
DTT	Dithiothreitol
EBV	Epstein-Barr virus
EC	Esophageal cancer
ECL	Enhanced chemiluminescence
ECL	*Erythrina cristagalli* lectin
EDTA	Ethylenediamine tetraacetic acid
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
ELISA	Enzyme-linked immunosorbent assays
ER	Endoplasmic reticulum
Erk MAPK	Extracellular signal-regulated kinase, mitogen-activated protein kinase
ESI	Electrospray ionization
FA	Formic acid
FBS	Fetal Bovine Serum
Fuc	Fucose
GAG	Glycosaminoglycan
Gal	Galactose
GalNAc	N-Acetyl-D-galactosamine
GBPs	Glycan-binding proteins
Glc	Glucose
GlcA	Glucuronic acid
GlcNAc	N-Acetyl-D-glucosamine
GNL	*Galanthus nivalis* lectin
GPI	Glycosylphosphatidylinositol
GRPs	Glycan-recognizing probes
HEPES	Hydroxyethyl piperazineethanesulfonic acid
HNC	Head and neck cancer
HPLC	High-pressure liquid chromatography
HPV	Human papillomavirus
H&E	Hematoxylin and eosin
H_2SO_4	Sulfuric acid
IAA	Iodoacetamida
IHC	Immunohistochemical
IP	Immunoaffinity purification
KCl	Potassium chloride
KH_2PO_4	Potassium phosphate monobasic
LacNAc	N-acetyllactosamine
LDS	Lithium dodecyl sulfate
LLO	Lipid-linked oligosaccharide
mAbs	Monoclonal antibodies
Man	Mannose

MagT1	Magnesium Transporter 1
MALDI	Matrix-assisted laser desorption/ionization
MES	2-(N-morpholino)ethanesulfonic acid
MGAT5	N-acetylglucosaminyltransferase V
mRNA	Menssager RiboNucleic Acid
MS	Mass spectrometry
NaCl	Sodium chloride
Na_2CO_3	Sodium carbonate
Na_2HPO_4	Sodium hydrogen phosphate
NeuAc	Acetylneuraminic acid
NH_4OH	Ammonium hydroxide
OST	Oligosaccharyltransferase
PBS	Phosphate-buffered saline
PHA-E	*Phaseolus vulgaris* erythroagglutinin
PI	protease inhibitor
PKC	Protein kinase C
PNA	Peanut agglutinin
PNGase	Peptide-N-glycosidase
PTB	Phospho-tyrosine binding
PVDF	Polyvinylidene difluoride
RB	Retinoblastoma
RNAi	RiboNucleic Acid interference
ROR1	Receptor tyrosine kinase-like orphan receptor 1
RPMI	Roswell Park Memorial Institute medium
RT	Room temperature
SDS-PAGE	Sodium Dodecyl Sulphate - Polyacrylamide Gel Electrophoresis
Ser	Serine
SH_2	Src homology 2
SLe	Sialyl Lewis
SNA	*Sambucus nigra* agglutinin
STAT	Signal transducer and activator of transcription
TBS-T	Tris-buffered saline - tween
TGF-α	Transforming growth factor alpha
Thr	Threonine
TKIs	Tyrosine kinase inhibitors
TMB	Tetramethylbenzidine
TNM	Tumor, lymph-node, metastasis
TNS	Trypsin Neutralization Solution
TOF	Time-of-flight
UV	Ultraviolet
WB	Western blot
WFL	*Wisteria floribunda* lectin
WGA	Wheat germ agglutinin
Xyl	Xylose

Units

h	Hour
kDa	Kilodalton
mg	Milligram
min	Minute
mL	Milliliter
mM	Millimolar
ng	Nanogram
nm	Nanometer
nM	Nanomolar
rpm	Rotation *per* minute
V	Volt
µg	Microgram
µL	Microliter
ºC	Celsius degrees

1. Introduction

1.1 Cancer

Cancer is one of the leading causes of mortality around the world, with an estimated 12.7 million new cancer cases and 7.6 million deaths annually [1]. The type of cancer which is more frequently diagnosed in females is breast cancer and in males is lung cancer, but other types like colorectal, stomach, liver or head and neck are also common among the population [1]. The increasing cancer incidence in the population is due to ageing, increase of the population and adoption of a bad lifestyle that includes physical inactivity, smoking, drinking and poor nutritional diets [1]. Some cancer cases could be prevented by the application of programs that would promote healthier lifestyles, like anti-tobacco campaigns, vaccination, physical activity and healthier diets. Death rates could also be decreased with earlier detection of cancer [1, 2]. However, determination of the cancer stage is important to cancer surveillance and control [3]. Cancer stage is determined by clinical examination, imaging, cytology of lymph nodes and histopathology after surgery. This evaluation allows a TNM classification by determining the size and extent of the tumor (T), presence of lymph-node metastasis (N) and distant metastasis (M) [3-5]. In this system, numbers (0 to 4) are assigned to these parameters according to the severity of the examined case [3]. Other less detailed cancer stage classification is based on location of the cancer cells. Stage 0 is identified when the tumor is found only in one part of the organism, stage I, II or III is considered when cancer has extended beyond the limits of the organ of origin to surrounding organs or tissues and there is involvement of regional lymph nodes, and stage IV is declared when the cancer has spread to distant parts of the body. In both classifications the higher the number, the more severe the situation is [5].

In general, cancer is a genetic disease that occurs when the DNA information is altered with an accumulation of mutations, leading to abnormal patterns of gene expression that can change fundamental biological processes of the cells [6].

The damage in the genome can be the result of endogenous processes, including errors during the replication of DNA, intrinsic chemical instability of some DNA bases and/or attack by free radicals generated during metabolism. It can also be the result of interactions with exogenous agents, such as UV radiation, ionizing radiation and chemical carcinogens [7].

Cells have mechanisms to repair the damage on DNA, like p53, retinoblastoma (RB) or cell-cycle checkpoints. However when mutations affect genes responsible for genome maintenance and normal homeostatic mechanisms that control cell proliferation and death, these cells will acquire more mutations, evolving eventually to malignancy [6, 7].

The hallmarks of cancer or the changes that occur in cancer cells were described by Hanahan and Weinberg [8] and are:

a) Growth factor independence or self-sufficiency by expressing growth factors to stimulate the receptors or altering the number, structure or function of the growth factor receptors;

b) Insensitivity to anti-growth signals;

c) Avoidance of programmed cell death (apoptosis);

d) Ability to recruit a blood supply by upregulating production of pro-angiogenic proteins and by downregulating production of anti-angiogenic proteins;

e) Immortalization by reactivation of telomerase;

f) Ability to invade adjacent normal tissues and metastasize to distant sites by expression of molecules that allow cells to digest the extracellular matrix, migrate and go into the blood or lymphatic vessels, reaching new locals in the organism;

g) Reprogrammed energy metabolism, by altering the glycolysis that obtains pyruvate, which is usually oxidized in the mitochondria, to glycolysis with generation of lactate;

h) Evasion of immune destruction, by recruiting immunosuppressive components of the immune system [6, 8].

The treatments applied to prolong cancer patients' survival vary according with the type of cancer, the stage level at diagnosis and with the patient's response to treatment [9]. Normally these may include:

a) Debulking surgery;
b) Radiation therapy (causes DNA damage in cancer cells causing cell death);
c) Chemotherapeutic regimens (kills tumor cells or slows their growth);
d) Immunotherapy (uses monoclonal antibodies or other components to boost patient's immune response against the tumor);
e) Targeted therapy (a treatment that targets the alterations of cancer cells and prevents their proliferation, helps the immune system to destroy these cells, stops cell signaling, delivers cell-killing substances, among others);
f) Hormone therapy (applied in cancers that need hormones to grow, or to prevent or reduce some symptoms);
g) Stem cells transplants (the patient receives healthy blood-forming stem cells to help the recovery after a treatment with high doses of radiation or chemotherapy);
h) Personalized medicine (treatments selected based on the genetic knowledge of the patient's disease) [10].

Personalized medicine is still not applied to the general community, due to high costs and developing time, so patients receive a standard treatment for their type of cancer. However, investigation on targeted therapies is evolving, with studies focusing on specific alterations on cancer cells in different types of cancer. So, there is a clinical necessity to have a comprehensive characterization of cancer types, to know the differences between normal and cancer cells, in order to establish new targeted therapies with minimal adverse effects [11].

1.1.1. Head and Neck cancer

Head and neck cancer (HNC) includes epithelial malignancies of mucosal linings of upper aerodigestive tract that affects several organs, including the paranasal sinuses, nasal cavity, oral cavity, nasopharynx, oropharynx, hypopharynx and larynx (Figure 1.1). HNC can also affect the salivary glands but it is unusual in comparison with other cases [4, 12-14].

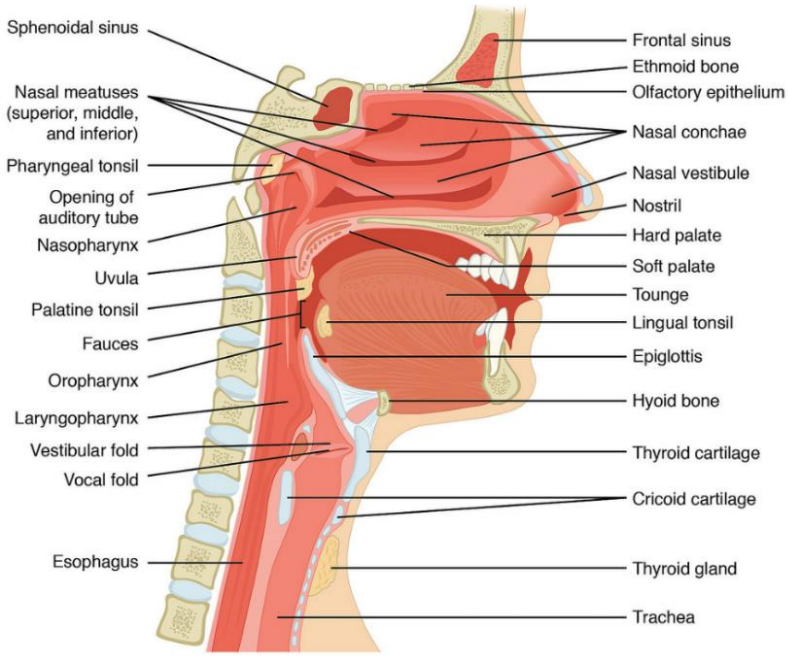


Figure 1.1 - Head and neck anatomy [15].

Most of HNC cases are squamous cell carcinomas, affecting epithelial cells, but also a variety of mesenchymal derived sarcomas and carcinomas, affecting salivary glands (adenomas/ adenocarcinomas), muscular structures and lymphoid aggregates, can be found [16, 17].

The incidence of this type of cancer in Portugal is approximately 2500 new cases *per* year [18] and worldwide is about 600 000 new cases *per* year, being the sixth leading cancer affecting the population in the world [4] and the ninth in Portugal [19]. This cancer affects more men than women and the incidence increases with age [13]. The 5-year survival rate is about 60%, but this number depends on the cancer stage at diagnosis.

The main exogenous risk factors for this type of cancer are tobacco use and alcohol consumption, and it is believed that there is a synergistic effect between them [12]. Human papillomavirus (HPV) infections are also a particular risk factor for cancers with origin in the oropharynx. Other risk factors are the use of betel quid [20]; consumption of certain preserved or salted foods [21]; poor oral hygiene and missing teeth [22]; occupational exposure to wood dust, formaldehyde, asbestos, synthetic fibers [21]; and Epstein-Barr virus (EBV) infection [23].

Endogenous risk factors include inherited disorders, like Li-Fraumeni or Fanconi anemia. Gastroesophageal reflux disease has also been associated as a potential contributing factor [4].

Depending of the organs affected, symptoms manifestation can be different. The main symptoms for cancer affecting the oral cavity are irritation, pain and difficulty on eating. Malignancies in the pharynx can lead to painful and difficulties in swallowing leading to weight loss. Patients with lesions in the larynx present hoarseness, while patients with lesions in the nasal cavity may complain of nose bleeding and obstruction. Symptoms of malignancies from salivary glands can present as swelling under the chin, paralysis of the face muscles and continuous pain [13].

The most common treatments for patients with HNC are surgical excision, radiotherapy and chemotherapy, although targeted therapies can also be applied. The application of single or combined treatments depend on the location of the tumor, its stage, size, among other characteristics.

Furthermore, the organs affected in HNC are essential for mastication, swallowing, breathing, communication and since they are the most visible portions of the body, the quality of life of the patients can be affected. So, treatments need to take this into account, having the objective to cure the malignancy while trying not changing the functionality and the appearance of the organs. Therefore, a non-surgical therapy is preferred, and the application of induction chemotherapy followed by radiotherapy can have the equivalent survival rate of surgery, but with better preservation of the organ and its function [13,14].

Genetic alterations in HNC cells lead to modifications that have been described by Hanahan and Weinberg, including self-sufficiency in growth signals, with changes in the epidermal growth factor receptor (EGFR) pathways. In this type of cancer, EGFR is overexpressed and can present point mutations on its sequence. Due to the central role of this receptor in head and neck cancer, it is an obvious target for therapies [4,12]. These therapies are applied together with radiotherapy or chemotherapy and can improve the survival of patients in the majority of cases [14].

1.1.2. Esophageal cancer

Esophageal cancer (EC) is a disease that affects the cells in the esophagus tissue. The esophagus is an organ that serves as a conduit to the gastrointestinal tract for food, extending from the larynx to the stomach. This organ is constituted by several layers of tissue, which are the mucous membrane, muscle and connective tissue (Figure 1.2) [24].

The incidence of EC is increasing, with about 1000 new cases in Portugal [19] and approximately 500 000 new cases and 338 000 deaths estimated worldwide, being the eighth most common cancer in the world [24] and the tenth in Portugal [19]. It affects about three

times more males than females and corresponds to 4% of male deaths caused by cancer. The median age of EC patients is about 67 years. The 5-year survival rate for this type of cancer is around 20% depending on the cancer stage upon diagnosis and on other characteristics of the patient [2, 25].

There are two forms of EC, depending on the type of cells that become malignant. The squamous cell carcinoma, or epidermoid carcinoma, which is formed in squamous lining cells of the esophagus and is mostly found in the upper and middle part of esophagus. The adenocarcinoma affects the glandular or secretory cells in the lining of the esophagus and these are usually found in the lower part of the esophagus, near the stomach [24].

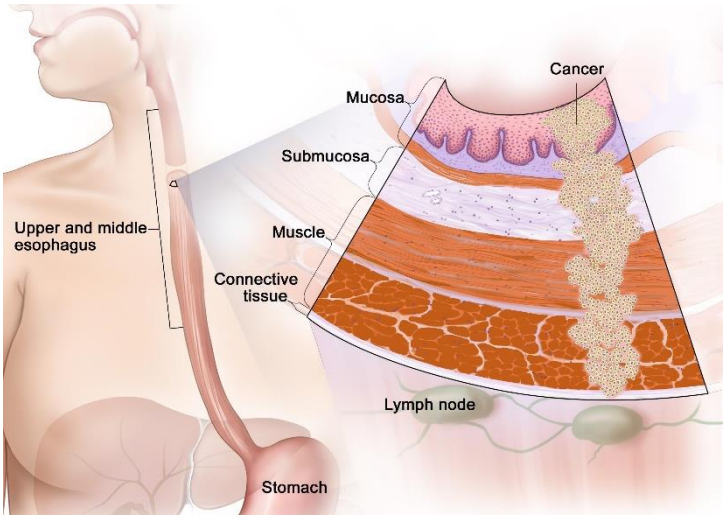


Figure 1.2 – Scheme of location of a stage II squamous cell esophageal cancer, placed in upper and middle esophagus, affecting the several layers of mucosa, muscle and connective tissue [171]

The risk factors for development of esophagus squamous cell carcinoma are the tobacco and alcohol use and human papillomavirus (HPV), Epstein–Barr virus (EBV) and polyoma viruses [26]. The risk factor associated to adenocarcinoma is a condition named Barret esophagus, in which the squamous epithelium of the lower part of the esophagus is replaced by columnar epithelium (metaplasia) due to gastroesophageal reflux [27]. Other risk factors for both cancer types are ageing, bad nutrition and genetic predisposition [2, 28].

The signs and symptoms for EC can be pain or difficulty in swallowing, weight loss, pain behind the breastbone, hoarseness, cough, indigestion and heartburn.

The treatment plan will depend of the stage and tumor grade [28]. The standard treatment options are surgery, radiation therapy, chemotherapy, chemoradiation therapy, laser therapy and electrocoagulation. These treatments may cause negative side effects and during the treatments patients will have special nutritional needs, since the affected organ belongs to the gastrointestinal tract [29].

Like in HNC, the use of molecular targeting agents such as antibodies and Tyrosine Kinase inhibitors (TKIs), among others is also in place [24]. A potential target for these type of

therapies is the EGFR, since it is overexpressed in esophageal cancer cells and its location on the cell membrane makes it an interesting target [30].

1.2. Epidermal growth factor receptor

EGFR is a transmembrane glycoprotein and a cell-surface receptor that belongs to the ErbB family of receptor tyrosine kinases. It participates mainly in the regulation of cell proliferation, survival, differentiation, through different signaling pathways [31]. EGFR is present on all epithelial and stromal cells as well some glial and smooth muscle cells [32].

- **Structure of EGFR**

EGFR gene is located across 188.3 kb on chromosome 7p11.2 [33] and encodes to a type 1 protein of 1210 amino acids. After cleavage of the N-terminal sequence, the protein with 1186 amino acid residues is inserted into the cell membrane [32, 34]. Several splice variants that encode for different protein isoforms have also been described [32, 35]

EGFR has a molecular mass of ~170 kDa [31, 34] and a structure divided into extracellular portion, or ectodomain, transmembrane domain and intracellular portion, that includes a juxtamembrane sequence, a tyrosine kinase domain and carboxyl-terminal domain (Figure 1.3) [32].

The ectodomain of EGFR has 622 amino acids and it is divided into four domains (I, II, III and IV) [32]. Domain I and III, also termed L1 and L2, are members of the leucine rich repeat family and are responsible for ligand binding. However, the domain III alone contributes to most of the binding energy, with an affinity to epidermal growth factor (EGF) of approximately 400 nM, while the domain I is involved in a secondary interaction. Domain II and IV, also termed CR1 and CR2, respectively, are cysteine-rich domains, with 22 and 20 cysteines, respectively, containing multiple small disulfide-bonded modules [31, 36]. Domain II participates in homo and heterodimerization of the receptor with ErbB family members [37].

The transmembrane domain of EGFR has 23 amino acids and it is constituted by a hydrophobic single pass membrane structure that anchors EGFR to the membrane, playing a role in its dimerization [38, 39].

The intracellular portion of EGFR has 542 amino acids. The juxtamembrane sequence comprises ~50 amino acids and is involved in EGFR dimer stabilization and signaling [34]. It has a site for feedback attenuation by protein kinase C (PKC) and Erk MAP kinases (extracellular signal-regulated kinase, mitogen-activated protein kinase) and a motif within this domain that possibly links to heterotrimeric G proteins [32, 40]. Next to this sequence there is a tyrosine kinase domain with ~250 amino acids followed by a 229-amino acid long carboxyl-terminal tail, containing five phosphorylation sites with tyrosine residues. When the receptor is

activated, C-terminal tail residues are phosphorylated, providing docking sites for signaling proteins for subsequent activation of signaling pathways [32, 36, 39],

- **Activation of EGFR**

EGFR exists at the cell surface as monomers or dimers. Before ligand binding, domains II and IV present a tethered conformation with disulfide bonds that inhibit dimerization, but with the binding of the ligand, EGFR achieves an active conformation [34, 39, 41]. EGFR can be activated by various ligands, including EGF, transforming growth factor alpha (TGF-α), amphiregulin, betacellulin, epiregulin, heparin-binding EGF, neuregulin 2β and epigen [31, 42, 43].

When a ligand binds to the extracellular part of EGFR, the receptor is activated and undergoes a series of changes. These alterations start in the membrane with the conversion of the monomeric form of the receptor to the dimeric form (Figure 1.3). This dimer formation can occur with another EGFR or with others receptors from the ErbB family [39]. This binding induces EGFR internalization and trafficking to early endosomes [44].

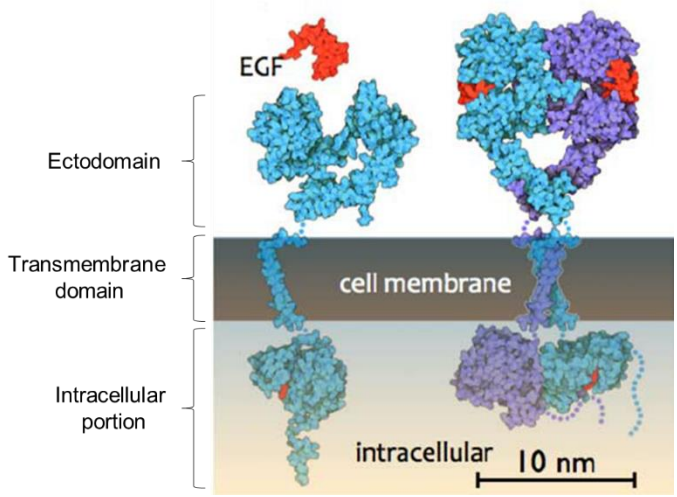


Figure 1.3 - Schematic representation of epidermal growth factor receptor in monomeric and dimeric form, after binding of the ligand *EGF* (adapted from [172]).

The dimerization process leads also to the activation of the kinase domain, increasing the affinity for adenosine 5' triphosphate (ATP) binding due to the associated conformational change [34], resulting in phosphorylation of its own tyrosine residues on the C-terminal domain [45, 46]. Then, the phosphorylated tyrosine kinase residues function as docking sites for target molecules, such as signal transducers or activators of intracellular substrates, and initiate a signal transduction cascade leading to cellular events such as DNA synthesis and cell division [36, 47],

Moreover, the molecules that interact with the binding sites can be several cytosolic proteins containing Src homology 2 (SH_2) domains or phospho-tyrosine binding (PTB) motifs [48], resulting in the activation of signaling pathways, including MAPK, PKC, signal transducer and activator of transcription ($STAT_3$), among others [49, 50]. This activation of EGFR leads to

multiple cell responses, like gene expression, cellular growth, proliferation, survival, differentiation, migration, angiogenesis and inhibition of apoptosis (Figure 1.4) [31, 34, 47].

In absence of ligand, the receptor is internalized with an half-life of 30 min, being rapidly recycled back to the cell surface leading to a EGFR distribution of 80 to 90% on the cell surface [31].

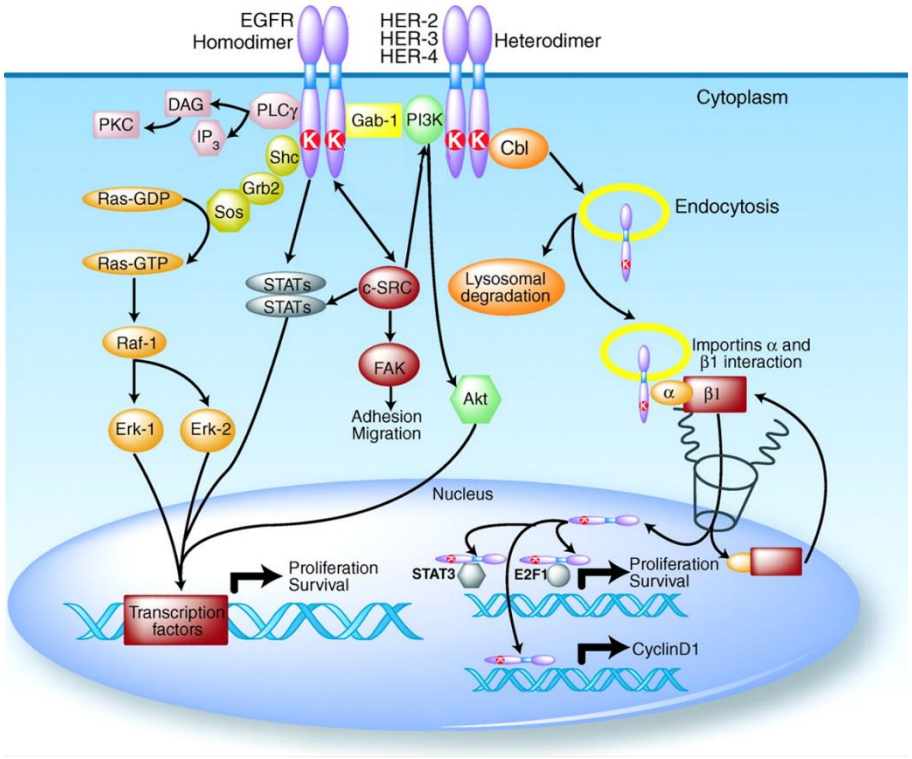


Figure 1.4 – Scheme of the signaling pathways of EGFR, including MAPK, PKC and STAT, leading to activation of transcription factors, promoting multiple cell responses (adapted from [51]).

1.2.1. EGFR in cancer

EGFR was the first receptor directly associated with cancer, since its activity contributes to proliferation which can provide an advantage to tumor cells survival [31].

Overexpression of EGFR has been observed in various types of cancer, for example, in head and neck cancer EGFR is overexpressed in 80-100% of the cases, being also overexpressed around 22-75% in colon and 35-70% in ovarian carcinoma [31] Furthermore, it has been found that signaling deregulation is due to aberrant expression of EGFR. Nearly 50% of grade IV gliomas have amplified EGFR genes, which are correlated with the structural rearrangement of the gene, resulting in in-frame deletions that preserve the reading frame of the protein's mRNA [52]. Hypoxic microenvironment of tumors can also induce overexpression of EGFR by increasing the translation of EGFR mRNA due to gene amplification, which can result in high levels of autocrine signaling [31, 53]. Aberrant gene expression is well correlated with increased cell growth, proliferation, invasion, differentiation and metastasis [54].

The metabolic half-life of EGFR in cancer cells is 20 h, much superior to the 30 min of normal cells. This means that one receptor will cycle through the endocytic pathways many

times during its lifetime which can affect the global pharmacokinetics of antibody drugs that bind to the receptor, due to antibody internalization and subsequent degradation [31].

It has also been observed in tumors several EGFR mutants, with deletions, regions of sequence duplication or defective kinase regulatory signals [34, 36]. The most common and best characterized mutant is EGFRvIII (or ΔEGFR), found in up to 40% of HNC cases [55] and also found in EC cases [56], in which the sequence spanning residues [6-273] is deleted from the extracellular portion, namely domain I and II. This mutant receptor with a molecular mass of approximately 145 kDa is able to dimerize even in the absence of ligand binding activity, its kinase is constitutively active due to self-dimerization and has a defective downregulation behavior [9, 34-36, 52, 57].

The frequent alterations of EGFR in several types of cancer makes it an evident target to new options of treatments or therapies for these malignancies.

1.3. Cancer therapies targeting EGFR

Several strategies targeting EGFR have been developed and shown remarkable results in various human malignancies. Those include the use of monoclonal antibodies (mAbs), small-molecule TKIs, immunoconjugates, antisense oligodeoxynucleotides (AS ODNs), RNAi, among others [31].

Functionally, mAbs bind to the extracellular domain of EGFR and compete with endogenous ligands, blocking the ligand-binding region [58] and preventing EGFR tyrosine kinase activation. These mAbs are highly selective, recognizing EGFR exclusively. Examples of anti-EGFR mAbs that are used in cancer treatment are cetuximab and panitumumab [31]. Cetuximab is a chimeric human murine immunoglobulin G1 mAb used in combination with chemotherapy and radiotherapy in patients with metastatic colorectal or head and neck cancer. It binds to EGFR and to EGFRvIII, with higher affinity compared to EGF or TGF-α. Consequently, this leads to a reduction of EGFR dependent downstream signaling pathways due to degradation of the receptor without phosphorylation and activation [31, 59]. Panitumumab is a high affinity human mAb that blocks ligand-binding and induces EGFR internalization. However, the receptor can be recycled back to the membrane [31], [60]. These antibodies, besides binding to cancer cells overexpressing EGFR, they also affect epithelial tissue cells that express EGFR, like skin and mucosa. This leads to numerous side effects, such as acneiform rash and other skin problems, diarrhea, hypomagnesemia, nausea, vomiting and, rarely, interstitial lung diseases [11, 61].

The TKIs are molecules (e.g. Gefitinib, Erlotinib, Lapatinib and Canertinib) that compete reversibly with ATP-biding pocket in the intracellular catalytic domain inhibiting EGFR autophosphorylation and downstream signaling. This prevents tumor cell proliferation, angiogenesis, metastasis and increases apoptosis [31, 62, 63]. The side effects of these molecules are the same of the antibodies but the patients also present loss of appetite, fatigue, conjunctivitis, elevated liver chemistries, among others [11].

Antibody based immunoconjugates act as a prodrug by releasing the drug into tumor cells after internalization. They improve the therapeutic window of chemotherapy or reduce the drug inactivation, by altering their *in vivo* distribution by conjugation to tumor-targeting mAbs. An example of this strategy is the conjugate composed by an anti-EGFR monoclonal antibody (EQ75) linked to an anti-mitotic drug Adriamycin (ADR) [31, 64, 65].

AS ODNs are a potential anti-cancer therapy in the sense these decrease the EGFR expression and regulate cell proliferation. There are some difficulties in the clinical development of these therapies due to the lack of effective delivery systems and the inability of applying the desired bio-activity of these AS ODNs [31, 66].

Other agents like affibodies, nanobodies, peptides, among others, have a structural similarity to EGF or have high binding affinity to EGFR, interfering with the mechanism of EGF binding to the receptor. They can be used for delivery of therapeutic agents to cancer cells due to its characteristics, but they are still in preliminary tests [31, 67, 68].

All these targeted therapies offer additional treatment options as they can be better tolerated than the normal treatments like chemotherapy and, in some cases, can prolong survival in patients. However, they continue to be associated with several adverse effects. So, more studies and a better characterization of the potential targets in cancer cells are needed [11].

1.4. Glycosylation

Glycosylation is a posttranslational modification in which glycans are covalently linked to molecules like proteins and lipids, via an enzymatic process.

Glycans, also called saccharides, carbohydrates or sugar chains, are based in the general formula $C_x(H_2O)_n$. The simplest form is the monosaccharide, which can connect to another monosaccharide through a glycosidic linkage, which forms between the anomeric carbon of one monosaccharide and a hydroxyl group of another. Joined monosaccharides form oligosaccharides (if composed by less than 20 monosaccharides) or polysaccharides, that can be a linear or a branched polymer [69]. Each monosaccharide presents several hydroxyl groups in its structure that can originate a glycosidic linkage. Thus, different types of oligosaccharides can be formed depending on what hydroxyl group participates in that linkage.

The most common monosaccharides found are D-Glucose (Glc), N-Acetyl-D-glucosamine (GlcNAc), D-Galactose (Gal), N-Acetyl-D-galactosamine (GalNAc), D-Mannose (Man), D-Xylose (Xyl), D-Glucuronic acid (GlcA), L-Fucose (Fuc) and N-Acetylneuraminic acid (NeuAc), being the latter the most common form of sialic acid (Figure 5) [69].

The glycosylation process originates different glycoconjugates depending of the modified molecule (Figure 1.5). Such example are glycoproteins that carry glycans covalently attached to their polypeptide backbone, generally via N-glycosylation, if the sugar chain is linked to an

asparagine residue on the protein; or O-glycosylation, if the sugar chain is linked to a hydroxyl group of a serine (Ser) or threonine (Thr) residues [70].

Other type of glycoconjugates include proteoglycans that are constituted by glycosaminoglycan (GAG) chains attached to a core protein through a xylose linked to the hydroxyl group of a serine residue [70].

Glycosylphosphatidylinositol (GPI) anchor is composed by a glycan structure that connects the carboxyl terminus of a membrane protein and the lipid bilayer membrane [70].

Glycosphingolipids are other type of glycoconjugates, in which a glycan is usually attached via glucose or galactose to the terminal hydroxyl group of ceramide [70].

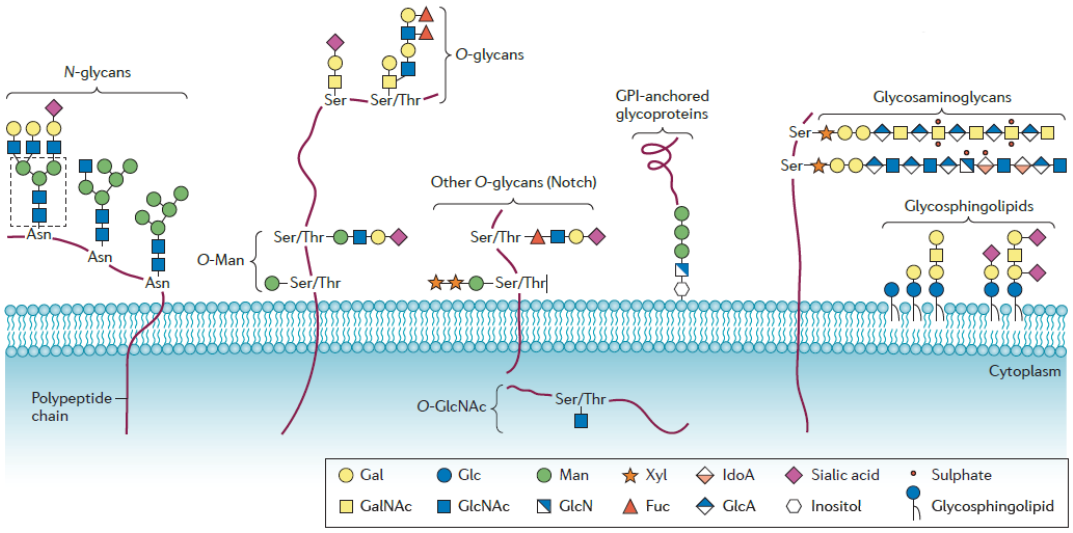


Figure 1.5 – Structure of the different glycoconjugates. Including N-Glycans, O-Glycans, GPI anchored glycoproteins, Proteoglycans/Glycosaminoglycans and Glycosphingolipids. (Adapted from [71])

These alterations on proteins and lipids add more diversity to these molecules and more associated functions.

1.4.1. Functions of Glycans

In nature, glycans are very abundant and have a large diversity, presenting different biological roles in the living organisms.

The biological functions of glycans can be divided in categories, including structural and modulatory properties, specific recognition by other molecules and molecular imitation of host glycans [72].

Depending on the different tissues, development time or environmental contexts, the same glycan can have different functions [73], such as:

Structural and modulatory functions:

Glycans have many protective, stabilizing, organizational and barrier functions, being the latter observed on the glycocalyx that covers eukaryotic cells. Globally, these constituents are important for maintenance of the structure, porosity and integrity of the tissue.

In glycoproteins, the glycans on the external position can protect the outer layers from recognition by proteases [74], block antibody binding [75] and, in some cases, protect against microbial attachment [76, 77].

Glycans are also important for proper folding of newly synthesized polypeptides in the endoplasmic reticulum (ER) and for protein solubility and conformation maintenance, avoiding protein degradation in proteasomes [74].

Glycosylation can also interfere with protein-protein interactions, for example controlling the binding capacity of growth factor receptors with growth factor proteins that are synthesized in the same cell, avoiding early interactions [72].

Glycosylation is also a mechanism to generate additional functional diversity than the obtained from gene products, for example influencing the interactions of receptor and ligand [73].

Additionally, glycans can also act as a storage depot for biologically important molecules, such as growth factors, water, ions or immune regulatory proteins [72].

Specific intrinsic recognition functions:

Glycans have central roles in cell-cell recognition and cell-matrix interactions. Glycan-binding proteins (GBPs) and glycans on cell surfaces can interact with the molecules in the matrix or with the glycans in the same cell surface [78], mediating cell trafficking [79], angiogenesis [80], adhesion and signaling events [81]. Some glycans can also act in a way they mask other molecules making them unrecognizable for external ligands, for example avoiding T-cells recognition [72, 82].

Specific extrinsic recognition functions:

Glycans have also a role in cell-microbe interactions. Various viruses, bacteria, parasites and some toxins evolved to recognize, with specificity, a sequence of glycans of the host to successfully invade the organism [83].

Glycan sequences in glycoconjugates, such as secreted mucins, act as a trap for microorganisms and parasites since the pathogen interacts with the soluble mucin and is removed without interacting with the target cells [76].

In cases of symbioses, the interactions between the host and the commensal bacteria are mediated by specific glycan recognition [84].

Also, foreign glycans or glycan patterns from microorganisms can be recognized by innate immune cells and are detected by specific receptors of the host [72, 73, 85].

Molecular imitation of host glycans function:

Moreover, when a pathogen invades a multicellular organism, it acquires glycan structures that are identical to those on host cell surfaces [84]. This is a strategy to evade host immune responses, since it can block the recognition of its antigenic epitopes [72, 73].

1.4.2. Glycosylation process

In eukaryotic cells, most of cell-surface and secreted proteins undergo post-translational modifications, including glycosylation reactions. Proteins are translocated into the ER, where they are folded, modified and subjected to quality control mechanisms. Then they pass through an intermediate compartment, through Golgi apparatus and are distributed to the various destinations [86].

Glycan structures are secondary gene products, meaning that they are not encoded directly in the genome [87]. But about 1-2% of the genome is dedicated to produce enzymes, such as glycosyltransferases, glycosidases and transporters, forming a complex glycosylation machinery present in various cellular compartments [82].

In glycosylation reactions, activated forms of monosaccharides are used as donors for glycosyltransferases. These donors are synthesized within the cytoplasmic or nuclear compartment from precursors of endogenous or exogenous origin. Due to their negative charge, they are actively transported through the membrane of ER or Golgi, by transporters that deliver nucleotide sugars into the lumen of these organelles with the simultaneous exiting of nucleoside monophosphates [86, 88].

Glycosyltransferases are a large family of enzymes responsible for the assembling of monosaccharides, in a sequential way, into linear and branched chains. They are very specific for both donor and acceptor substrates [89].

Some glycosylation enzymes of ER are soluble proteins, that can be involved in quality control and different types of glycosylation [86].

Most of glycosylation enzymes that act on Golgi apparatus are membrane proteins with an amino-terminal cytoplasmic tail, a transmembrane region and a carboxyl-terminal region. This region contains a membrane proximal region and a catalytic domain, placed in the Golgi lumen, participating in the synthesis of the glycan chains on proteins during their secretory pathway.

Other enzymes that participate in glycosylation are the glycosidases, which remove monosaccharides and are involved in the degradation of glycans. The glycans can also be modified by other types of enzymes, like sulfotransferases, phosphotransferases, acetyltransferases, methyltransferases, among others [89].

Protein glycosylation occurs in the secretory pathway, with some differences depending on the type of glycosylation. It is an ordered and sequential process involving glycosyltransferase reactions. For this process to occur, the proteins (or acceptors), the sugar

donors and enzymes must be in the same compartment of the cell. Glycosyltransferases are distributed along the Golgi compartments by the order of acting in this process, maintaining their location [86].

Each type of glycosylation have different mechanisms for initiation: transfering a large preassembled percursor (N-glycosylation and modification with GPI anchors) or with addition of a single monosaccharide (O-glycans, glycosaminoglycans and glycosphingolipids). The trimming and elongation proceed in the ER-Golgi pathway with distint enzymes involved in these modifications [86].

1.4.3. Major types of protein Glycosylation

The most common protein glycosylation processes are the O-glycosylation and N-glycosylation. These modifications are found in about half of eukaryotic proteins, in secreted and transmembrane proteins [90].

1.4.3.1. O-Glycosylation

O-linked glycosylation is initiated with the addition of a glycan molecule to an oxygen atom of hydroxyl group of exposed Ser or Thr residues in proteins [90].

There are several types of O-glycans, including O-GalNAc (the most common), O-fucose, O-glucose, O-mannose and O-GlcNAc.

O-GalNAc glycosylation occurs in Golgi with the addition of a GalNAc residue to proteins that are completely folded and assembled, followed by an enzymatic elongation by specific glycosyltransferases that extend the glycan structure [91]. O-glycosylated proteins can present four major core structures, that can be extended in linear or branched chains, reaching more than 20 sugar residues, being then modified by sialylation, fucosylation, methylation or acetylation [92].

In O-glycosylation with fucose and glucose, these sugars are added between conserved cysteines of substrates, such as EGF-like repeats of cell surface and secreted proteins [93]. The O-mannose glycosylation is found in several proteins, including cadherins and α-dystroglycan protein, being important to establish interactions with various matrix proteins [94]. These processes take place in ER with actuation of different glycosyltransferases [91].

O-GlcNAc glycosylation occurs in Ser or Thr residues of nuclear, mitochondrial and cytoplasmic proteins. This modification is not static, with the O-GlcNAc being attached and removed several times in the life of the glycosylated protein [95]. O-GlcNAc is also found in extracellular proteins containing EGF-like repeats [96]. However, in this case, this modification occurs in ER where a different enzyme mediates the process and the modification can occur in a putative consensus sequence between the fifth and sixth conserved cysteines of an EGF-like repeat [94].

1.4.3.2. N-Glycosylation

N-glycans are covalently attached to the nitrogen atom of an asparagine (Asn) residue in secreted and membrane-bound proteins by an N-glycosidic bond. This type of glycosylation occurs within the consensus sequence motif Asn-X-Ser/Thr, in which "X" is any amino acid except proline. This sequence needs to be accessible to the ER lumen to be glycosylated. All N-glycans share a common core sequence: Manα1-3(Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ1-Asn-X-Ser/Thr [97].

There are three types of N-glycans: oligomannose or high mannose, with Man residues extending the core; complex, in which the branches are initiated with GlcNAc and the core is extended with multiple sugar types; and hybrid type in which Man extends Manα1-6 arm of the core and one or two GlcNAcs extend the Manα1-3 arm (Figure 1.6) [97].

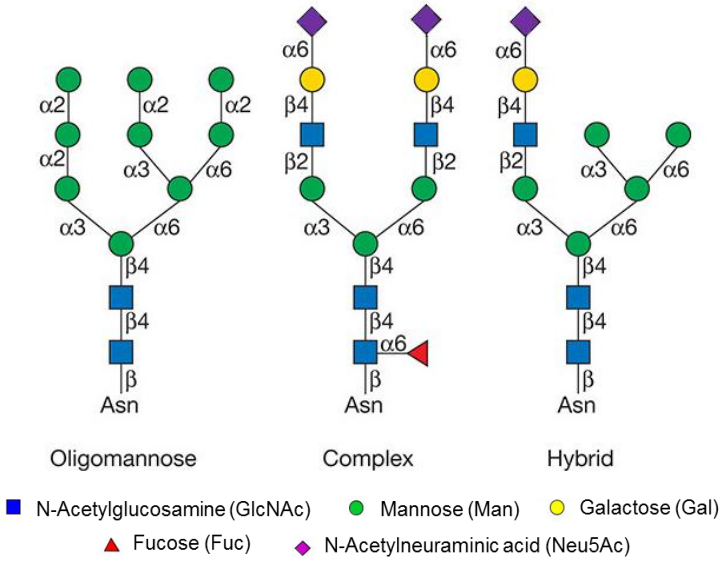


Figure 1.6 – Examples of structures for the different types of N-glycans. (Adapted from [97])

Synthesis of N-glycans

N-glycosylation occurs with a transfer of a N-glycan to an Asn-X-Ser/Thr sequence at the lumen side of ER membrane during or after the translocation of the protein substrate. The efficiency of the glycosylation can be affected, reduced or enhanced, depending of the "X" identity. The transfer of the N-glycan to this sequence does not always occur, due to the conformation of the glycoprotein during its folding. So, this sequence is a potential site for N-glycosylation, but not all of them present in the protein are mandatory N-glycosylated [97].

Synthesis of N-glycans happens in two phases within the compartments of endoplasmic reticulum (ER) and Golgi apparatus.

The first phase, highly conserved, takes place on the ER membrane with the involvement of the Dolichol phosphate (Dol-P). During the translocation of the protein into the ER, a precursor lipid-linked oligosaccharide (LLO) with 14 sugar units assembled on Dol-P is transferred to a selected Asn residue within the consensus sequence [97, 98]. The LLO

substrate is initially assembled on the cytoplasmic side of the ER membrane, using three carbohydrate building blocks (GlcNAc, Man and Glc) that enter the pathway as nucleotide activated sugars, serving as substrates for the transferases. The assembly pathway of the LLO is then terminated in the lumen of ER, requiring the translocation of lipid-linked biosynthetic intermediates across the membrane [99].

Then, the enzymatic complex oligosaccharyltransferase (OST) composed by a group of proteins including DAD1, N33/Tusc3 or MagT1/IAP, OST48, ribophorin I, ribophorin II, STT3A or STT3B, KCP2 and DC2 [99], will catalyze the transfer of the oligosaccharide from the lipid carrier to the amide group of selected asparagine residues in the Asn-X-Ser/Thr sequence of the polypeptide chains, forming an N-glycosidic linkage. This process can also affect the protein folding, contributing to the correct conformation of a functional protein [99].

The second phase begins in the lumen of the ER with glycosidases performing an initial trimming via hydrolysis to remove terminal glycans from the precursor. This process serves to monitor protein folding or to indicate a future protein degradation [100]. Next the glycoprotein is trafficked to the Golgi where different glycosyltransferases and glycosidases, distributed within the Golgi compartments, process the N-glycans into the various types of glycan structures (Figure 1.7) [101].

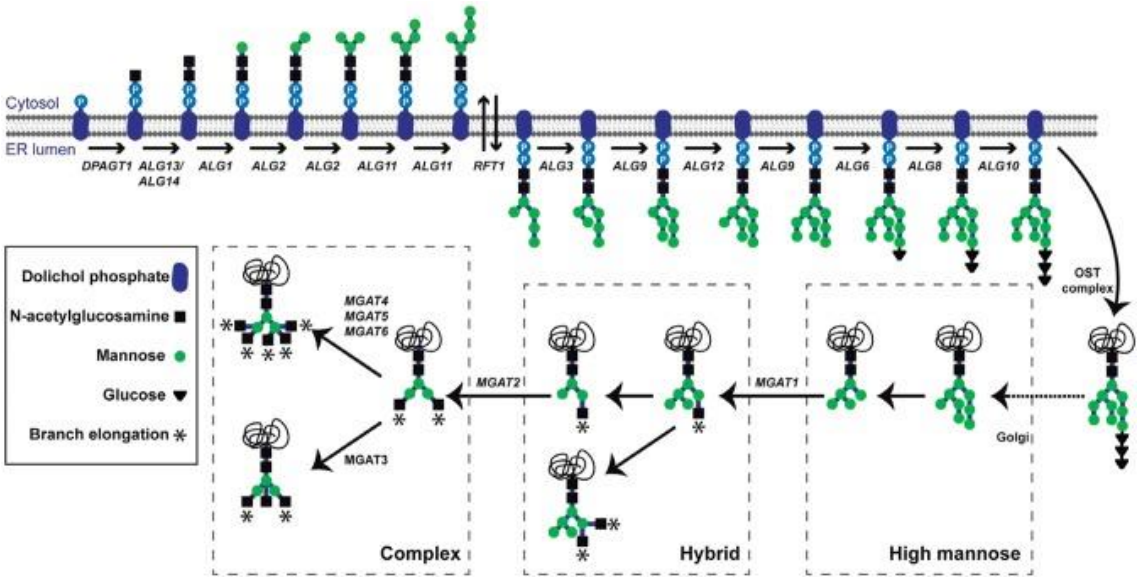


Figure 1.7 - N-glycosylation pathway. The precursor is assembled on Dolichol phosphate molecule (Dol-P) being this process initiated on cytosol, with actuation of Dolichyl-Phosphate N-Acetylglucosamine-phosphotransferase 1 (DPAGT1) and UDP-N-Acetylglucosaminyltransferase Subunits (ALG13, ALG14, ALG1, ALG2, ALG11), the glycan structure is translocated to endoplasmic reticulum (ER) lumen by RFT1, and there is the actuation of UDP-N-Acetylglucosaminyltransferase Subunits (ALG3, ALG9, ALG12, ALG6, ALG8 and ALG10). Then, it is attached to the protein with oligosaccharyltransferase (OST) complex intervention. Next, the glycoprotein transit to Golgi where glycosidases and glycosyltransferases act, including N-acetylglucosaminyltransferases (MGAT1, MGAT2, MGAT3, MGAT4, MGAT5 and MGAT6), originating different types of glycans structures [173].

After all these processes, N-glycans can also be subjected to capping of elongated branches. These capping reactions involve the addition of sialic acids, Fuc, Gal, GlcNAc and sulfate to complex N-glycans, in a way that facilitates the presentation of terminal sugars to lectins and antibodies [97].

N-glycosylation depends on the protein and on its N-glycosites, but also on the cell type, since glycosidases and glycosyltransferases are differently expressed and sensitive to the physiological state of the cell, which affects its localization and activity [97].

Glycoproteins can have different N-glycans on a particular site, which leads to glycan microheterogeneity at each site. Moreover, a given protein may have more than one consensus sequence motif in its structure that can be modified with different attached N-glycans, leading to a glycoprotein macroheterogeneity, constituting the glycoforms.

Variation of N-glycosylation can be due to the:

a) conformation of the protein that can affect substrate availability for glycosidases or glycosyltransferases acting in Golgi;
b) nucleotide sugar metabolism;
c) transport rate of the protein through the lumen of the ER and Golgi;
d) proximity of the Asn-X-Ser/Thr sequence to a transmembrane domain or to the next N-glycosite [99];
e) location of the glycosyltransferases within sub compartments of the Golgi that determines the enzymes that will act on the protein;
f) enzyme competition for the same acceptor and dependence of some enzymes of prior activity of other glycosyltransferases and glycosidases [97];
g) variations in micro-environmental stimuli, such as oxygen levels, and factors like age, gender and epigenetic background [82, 102].

1.4.4. Glycosylation in disease

The glycosylation process requires an involvement of many individual cellular components. This increases the probability for the occurrence of disease associated glycan alterations [82].

Glycosylation can be considered a key regulatory mechanism that controls various physiopathological processes. Defects in this mechanism are connected to several diseases, thus the glycome is expected to contain a remarkable amount of biological information.

Alterations in glycosylation are a hallmark of various diseases at several levels, including cardiovascular [103], dermatological [104], endocrinological [105], neurological [106] and oncological [107]. Knowing these alterations contribute to improve diagnosis and to develop therapies for these diseases [82, 108].

1.4.4.1. Glycosylation in cancer

Many physiological processes associated with cancer, including cell adhesion, proliferation, cellular signaling and immune response are influenced by glycans. Significant

alterations in the glycosylation machinery can cause or be the result of pathological events such as malignant transformation, tumor development, metastasis and invasion [54, 82, 98, 109].

These can be due to under- or overexpression of glycosidases and glycosyltransferases, variability on the abundance of sugar nucleotide donors as well as their acceptor substrates, changes in the tertiary conformation of the peptide backbone and altered localization of glycosyltransferases in the Golgi apparatus [71, 110, 111].

Increased sialylation is one common modification associated with cancer. It has an important role in cellular recognition, cell adhesion and signaling. Increased levels of sialylated glycans promote cell detachment from the tumor mass through electrostatic repulsion of negative charges, inhibiting and disrupting cell-cell adhesion [71].

Other alterations found in cancer patients are increased levels of Sialyl Lewis X (SLe^x) and Sialyl Lewis A (SLe^a) structures, presenting Nau5acα2-3Galβ1-3(Fucα1-3)GlcNAc-R and Nau5acα2-3Galβ1-3(Fucα1-4)GlcNAc-R structures respectively. These antigens are present in the terminal non-reducing ends of β1,6-branching of N-linked chains or of O-linked chains. When they are displayed on the surface of tumor cells, contribute to cancer cell dissemination and invasion [71, 82, 110],

In transformed cells, the size of N-glycans increases due to increased activity of enzymes, like those encoded by *MGAT4* and *MGAT5* genes. These enzymes catalyze GlcNAc branching of N-glycans, increasing the number of N-acetyllactosamine (LacNAc) units that can be sialylated and fucosylated, consequently enhancing tumor progression [111].

Other tumor specific alterations are the onco-fetal antigens that are expressed in tumor cells and embryonic tissues [111] and are important biomarkers for different types of cancer [112, 113]. These alterations can be glycan epitopes, such as SLe^a antigen (CA19-9), or glycoproteins such as 5T4 [114], receptor tyrosine kinase-like orphan receptor 1 (ROR1) [115], alpha-fetoprotein (AFP) [116], annexin A2 (ANXA2), carcinoembryonic antigen (CEA), and Mucin 16 (CA-125) [112].

All these glycosylation alterations are associated with tumor development steps, such as proliferation, invasion, metastasis and angiogenesis. They are an important target for cancer diagnosis, a cancer biomarker for prognosis, monitoring and therapy [82, 98].

1.4.5. Glycosylation of EGFR

EGFR is a glycoprotein highly N-glycosylated and this modification contributes with 40 kDa to the 170 kDa of mature EGFR [117]. EGFR is also O-glycosylated, presenting structures with O-GalNAc [118] and O-GlcNAc [119].

EGFR has up to 11 potential N-linked canonical glycosylation sites in its extracellular portion. The identification of these sites was obtained by analysis of three different cell lines, including human epidermoid carcinoma cell line (A431) [117], chinese hamster ovary (CHO)-derived cell lines [120, 121] and insect cell line (Sf9) [41]. These cell's characteristics allied to the presentation of the studied protein, being in full-length or only the extracellular domain, lead

to different results, which together allows the characterization of the glycosites of EGFR [122]. In these results, the potential glycosylated Asn identified include Asn104, Asn151, Asn172, Asn328, Asn337, Asn389, Asn420, Asn504, Asn544, Asn579, Asn599 [41, 117, 120-122], One noncanonical site, Asn32, that belongs to Asn-Asn-Cys sequence, was also identified [121]. All the canonical glycosites show to be glycosylated in different ways, mass spectrometry studies revealed that eight sites are fully glycosylated, two are not glycosylated and one is partially glycosylated, but these results depend on cell type [122].

Using inhibitors of N-glycosylation, such as tunicamycin, site mutagenesis or genetically manipulating specific glycosyltransferases, it has been possible to understand the importance of N-linked glycosylation to EGFR [123]. It has been shown that glycosylation is important for EGFR translocation and maturation, for the conformational arrangement of the extracellular region and its membrane interactions and for the proper folding of the receptor to generate an active conformation that allows ligand binding. It also plays a role in receptor self-association, receptor activation, stability of the protein and endocytosis [36, 117, 124, 125].

Previous studies have showed that ligand binding, dimerization and tyrosine phosphorylation can be reduced by both α2,3 and α2,6 sialylation and fucosylation of EGFR [123, 126]. However, it has been reported that core fucosylation is necessary for ligand binding and intracellular signaling of EGFR, since the biological function of this receptor decreases in its absence [123, 127, 128]. EGFR activity and signaling pathways can also be enhanced by β1,6-branched glycans and their biosynthetic enzyme N-acetylglucosaminyltransferase V (MGAT5) [123, 129].

As referred above, the glycosylation of EGFR also depends on the cell and tissue type, so it can exhibit different glycosylation patterns depending on the conditions found on the cell [117].

1.4.5.1. Aberrant glycosylation of EGFR in cancer

Aberrant glycosylation is a hallmark of cancer. Since EGFR function depends on proper glycosylation, it is expected that deregulation of EGFR activity in malignancy to be related with abnormalities in its glycosylation [54, 125]. All the alterations in glycosylation that promote ligand binding, dimerization and activation of this receptor are expected to be found in cancer cells, such as the increase of core fucosylation and branched glycans.

Core fucosylation and overexpression of enzymes such as MGAT5, affect EGFR by increasing the dimerization and phosphorylation. This promotes EGFR-mediated signaling that is associated with tumor cells progression and invasion [71, 126].

In some studies, with lung cancer cells, increased sialylation and α1,3-linked fucosylation lead to a decrease in the dimerization and phosphorylation of this receptor, affecting the metastatic ability of cancer cells. So, inhibiting sialylation and fucosylation may contribute to more aggressive cancer cells [126]. However, studies with ovarian cancer cells demonstrated that sialylation of EGFR by overexpressed α2,3-sialyltransferase type I (ST3Gal-I) or α2,6-

sialyltransferase type I (ST6Gal-I) was positively associated with its function, enhancing its activation, leading to migration and invasiveness of cancer cells [130, 131].

Glycosylation of EGFR with Lewis^y (Le^y), with Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ1-R structure, was observed in cancer cells and the increase of this modification is able to promote cell migration, stabilize EGFR upon activation and regulate its signaling pathways, leading to more aggressive tumors [132].

The overexpression of EGFR in several cancer types is associated with the overexpression of some glycosidases and glycosyltransferases that interact with EGFR, participating in its glycosylation process. An example of those enzymes is the sialidase NEU3, that promotes EGFR phosphorylation and activation of downstream signaling [133]. Another example is the ribophorin II, a part of OST complex, that participates in EGFR glycosylation transferring high mannose oligosaccharides to the asparagine in the consensus motif. This enzyme promotes cancer cells proliferation through mediating EGFR glycosylation, affecting its location, expression and the cellular functions [125]. The enzyme β1,4-N-acetylgalactosaminyltransferase III (B4GALNT3), that forms GalNAcβ1-4GlcNAc (LacdiNAc), has been reported to be highly expressed in some cancer cases and participates in EGFR glycosylation, affecting its phosphorylation and downstream signaling, leading to migration and invasion of cancer cells [134].

EGFR activity can also be affected by modification of its O-glycans by N-acetylgalactosaminyl transferases (GALNTs), increasing phosphorylation and enhancing the invasive potential of cancer cells [135, 136].

All this information was obtained by analyzing different types of cancer cells. These cells can present different EGFR glycoforms that may exhibit distinct activities. Nevertheless, these overexpressed enzymes that participate in EGFR glycosylation, may constitute a potential therapeutic target, together with the aberrant EGFR glycosylation present in certain cancer cells.

1.5. Glycomics and Glycoproteomics

Glycomics describes the complete repertoire of glycans that are produced by a cell or tissue under specific conditions, time, location and environment. It is influenced by factors like the transcriptome, the proteome, environmental nutrients, pH, among others [137].

Glycoproteomics ascertains which sites are glycosylated on a glycoprotein and includes also the identification and quantification of each glycan structure at each site on the different glycoforms present in the cell [137, 138]. To obtain this information, various techniques have been developed and applied in parallel: the “bottom up” technique, analyzing glycan structures that were obtained from one cell, and “top down” approach, taking advantage on the analysis of expressed glycans of the global tissue [137].

1.5.1. Glycan analysis

In glycan analysis, different molecules termed glycan-recognizing probes (GRPs) can be used, such as antibodies, lectins, microbial adhesins, viral agglutinins and other proteins with capacity to bind to glycans. These molecules have specificities for certain glycan structures, binding with high affinity and allowing its identification among a variety of glycan structures [139].

There are different methodologies in which lectins, antibodies and other carbohydrate-binding molecules (CBMs) can be applied to the identification of glycan structures existing on glycoconjugates. These molecules can also be used in glycan purification by isolating glycoconjugates that express specific glycan' determinants. Some examples are represented in Figure 1.8 [139].

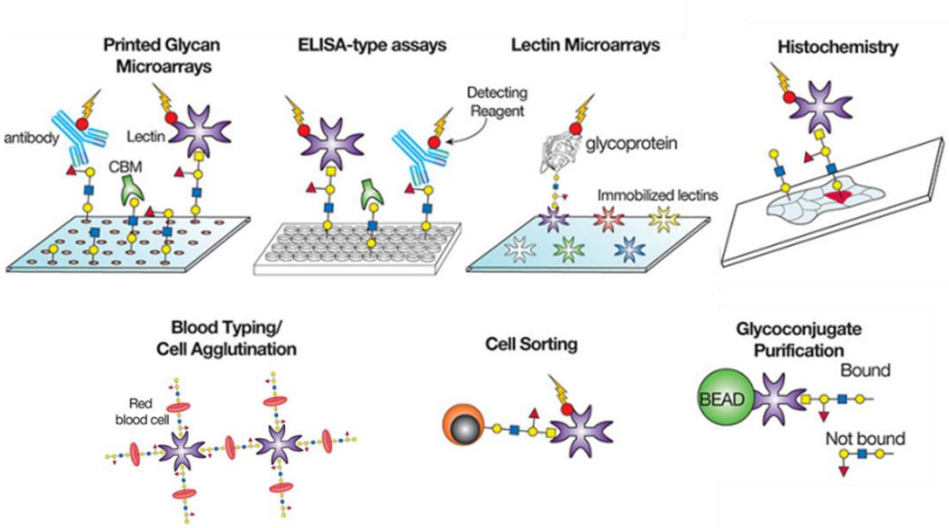


Figure 1.8 - Different applications of lectins, carbohydrate-binding molecules (CBMs) and antibodies for identification of glycan structures (adapted from [139])

Characterization of cell-surface glycoconjugates can use antibodies, lectins and other CBMs in different approaches, such as microarrays. Here, glycans can be printed on the slide and detected by biotinylated GRPs or instead with different immobilized lectins or antibodies printed on a slide detecting the glycosylation of cells and glycoconjugates, revealing the differences in protein glycosylation between samples, and contributing to the characterization of the glycan structures [140]. Another technique comprises enzyme-linked immunosorbent assays (ELISA), in which glycans can be immobilized on a plate and detected with GRPs or be captured by immobilized GRPs. Additional approaches involve histochemical techniques, in which fixed tissue or cells are incubated with biotinylated or peroxidase-labelled lectins or antibodies and visualized with secondary reagents. A different strategy is cell agglutination and precipitation, in which these molecules are added to a solution and if they recognize certain glycan composition of a target cell, they will bind to that target promoting agglutination and precipitation. Other method is flow cytometry with cell sorting, in which cells are incubated with

fluorescent lectins or antibodies, allowing the correlation of the degree of fluorescence with the number of binding sites and identification of the cells by their fluorescence.

Glycan purification using antibodies or lectins include affinity chromatography, or immunoprecipitation, which can be used to isolate glycoconjugates expressing specific glycan determinants, thus separating glycoconjugates from non-glycosylated material [139].

1.5.1.1. Mass spectrometry

Mass spectrometry (MS) technique, besides allowing the identification of different types of molecules such as peptides, proteins, metabolites and lipids, it has been also optimized for analysis of the glycans present in a diverse biological material. This technique can be applied with high throughput and high-sensitivity, requiring only a small sample [138], and allowing to define:

a) structural features of glycans including the degree of heterogeneity and type of glycosylation;

b) sites of glycosylation and identity of the protein carrier;

c) glycan branching, the number and lengths of antennae;

d) glycans composition and substitution with Fuc, Sia or other capping groups like sulfate, phosphate or acetyl esters;

e) the complete sequences of individual glycans [141].

Before proceeding with free-glycan analysis by MS, the glycans need to be released from glycoconjugates. For N-glycosylated proteins, N-glycans can be released using N-glycosidases such as peptide-N-glycosidase F (PNGase F) and peptide-N-glycosidase A (PNGase A) or by chemical approaches like acid hydrolysis [137]. For O-glycosylated proteins, O-glycans can be released by β-elimination with sodium hydroxide, since O-linkage is labile under alkaline conditions [92]. Then, glycans need to be separated and isolated from the mixture, normally by chromatographic separations, such as size exclusion chromatography, strong or weak anion exchange chromatography and some forms of reverse-phase high-pressure liquid chromatography (HPLC) [141].

The glycans are then analyzed by a mass spectrometer that is composed by an ion source, a mass analyzer and a detector. The ion source ionizes molecules converting them into the gaseous phase. Then, a mass analyzer separates ions and measures the mass-to-charge ratio (m/z) of ionized analytes. The detector registers the number of ions at each m/z value [142-144]. It provides masses of ionizable glycans and their fragments, generating a mass spectrum (i.e. intensity vs m/z graph) [141].

There are different types of ionization used in MS methods, such as matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) [143]. In MALDI-MS, the sample is dried on a plate in the presence of a light-absorbing matrix that absorbs energy from laser pulses. The matrix forms crystals containing trapped sample molecules and the laser actuation causes the desorption of matrix and ions from the sample into the gas phase [145]. ESI-MS

consists of a system coupled to liquid-based separation tools, like liquid chromatography, where a stream of liquid that contains the sample enters the source through a capillary interface, stripping the sample molecules from the solvent, producing intact gas-phase ions, obtaining multiply charged species [141, 142, 146].

MALDI-MS and ESI-MS provide structural information of masses of intact molecules (molecular ions). However, the ionization process is usually not sufficiently energetic to fragment ions. To overcome this, these instruments have analyzers such as quadrupole acceleration time-of-flight (Q-TOF) or TOF arranged in tandem, allowing the detection of the fragment ions by the last analyzer, after molecular ions undergo collisions with an inert gas in chambers placed between the analyzers [141, 145].

Based on the experimentally determined mass, a suggested list of glycan composition can be acquired using glycoinformatics tools like GlycoMod, that suggest possible glycan compositions from the experimental mass values [147].

2. Aim of the thesis

Anticancer therapeutics directed against EGFR although effective are associated with a high prevalence of dermatologic side effects, most commonly skin rash, due to the high expression of this receptor in keratinocytes.

The main aim of this thesis was the development of a strategy for the identification of EGFR glycosylation patterns. Discovery of these profiles would allow the comparison between cancer, normal tissue and human keratinocytes, enabling the definition of cancer specific glycan signatures and identification of possible targets to improve specificity of targeted therapies. Due to the scarcity of cancer specimens in satisfactory conditions and low quantity of tissue available for research purposes different cell lines representative of HNC and EC cancer types were explored before analyzing primary cells/tissues.

The work developed during this thesis was divided in two main parts:

- The first goal was to optimize protein extraction and enrichment methods to guarantee the presence of EGFR protein from which glycans could be analyzed.
- The second goal was the optimization of sample preparation steps before LC-MS analysis and consequent identification of glycans (and proteins) to establish EGFR glycan profiles.

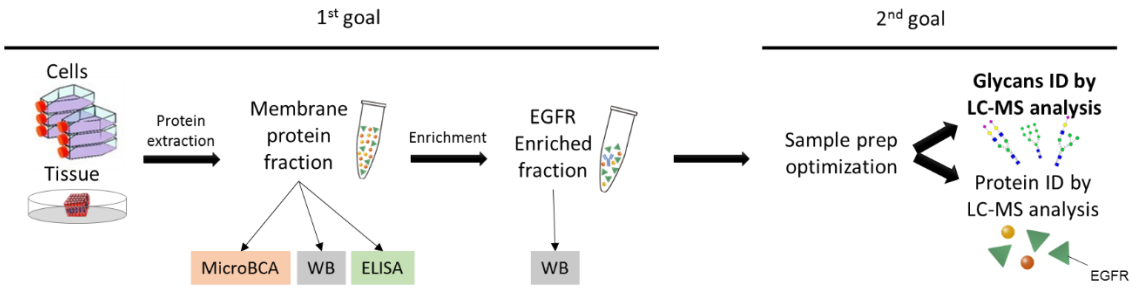


Figure 2.1 – Schematic representation of the work strategy implemented to ultimately establish EGFR glycan profiles from human cells and tissues.

3. Materials and methods

3.1. Biological material and culture media

3.1.1. Cell lines culture

Different cell lines representative of HNC, EC and normal keratinocytes were cultured as described below.

FaDu cell line, derived from human pharynx squamous cell carcinoma, was cultured in DMEM (Dulbecco's Modified Eagle Medium) (Merck), supplemented with 10% of Fetal Bovine Serum (FBS) (Gibco), with a seeding of 22000 cell/cm^2.

SNU-1076 cell line, derived from human larynx cancer cell line, was cultured in growth media composed by RPMI 1640 (Glibco) supplemented with 10% of FBS and 2 mM of L-Glutamine, with a seeding of 14000 cell/cm^2.

SCC-4 cell line, derived from human tongue squamous cell carcinoma, was cultured in 45% of DMEM/F12 (Gibco), 45% of DMEM, 10% of FBS and 400 ng/mL of hydrocortisone (Sigma), with a seeding of 11000 cell/cm^2.

Kyse-520 cell line, derived from human esophageal squamous cell carcinoma, was cultured in RPMI 1640, supplemented with 10% of FBS and 2 mM of L-Glutamine, with a seeding of 10000 cell/cm^2.

Kyse-30 cell line, derived from human esophageal squamous cell carcinoma, was cultured in 44.25% of RPMI 1640, 44.25% of Ham's F12 (Pan Biotech), supplemented with 10% of FBS, 2 mM of L-Glutamine and 1 mM of Sodium Pyruvate, with a seeding of 8000 cell/cm^2.

Kyse-450, derived from human esophageal squamous cell carcinoma, was cultured in 45% of RPMI 1640, 45% of Ham's F12, supplemented with 10% of FBS, with a seeding of 40000 cell/cm^2.

NHEK cells, derived from normal human epidermal keratinocytes (adult – pooled donors), was cultured in PromoCell Keratinocyte Growth Medium 2 (Serum-free), composed by Basal Medium, 0.06 mM of $CaCl_2$ and supplement mix (PromoCell), with a seeding of 5000 cell/cm^2.

All cell lines were cultured at 37ºC and 5% CO_2 in adherent conditions in T-flasks.

Cells were split twice a week, after reaching 80 to 95% confluence. For cell sub-culturing, all cell lines except NHEK were washed with PBS (Phosphate-buffered saline) and then trypsinized using 0.05% trypsin-EDTA (Ethylenediamine tetraacetic acid). For NHEK, detachment was performed using a detach kit where cells were washed with HEPES-buffered balanced Salt Solution (30 mM HEPES (hydroxyethyl piperazineethanesulfonic acid), D-Glucose, NaCl, KCl, Na-Phosphate and Phenol Red), trypsinized with Trypsin/EDTA (0.04%/0.03%) solution and neutralized with Trypsin Neutralization Solution (0.05% Trypsin inhibitor from soybean and 0,1% Bovine Serum Albumin). Viable cell concentration was assessed using Trypan blue exclusion test explained in 3.4.1.

3.1.2. Human tissue samples

Tumor tissues and their corresponding adjacent non-tumorigenic tissues were collected from patients with Head and Neck or Esophageal carcinomas at Instituto Português de Oncologia de Lisboa Francisco Gentil (IPOLFG) under the scope of iNOVA4Health program. Samples were embedded in OCT matrix and conserved at -80ªC. All cases are described in table 3.1.

Table 3.1 – Identification of collected cases from Head and Neck and Esophagus cancer patients. Number of the case, tumor and adjacent non-tumorigenic tissue weight and organ of origin are indicated.

Case	Weight (mg)		Organ
	Tumor	Normal	
1	175.6	147.5	Tongue
2	194.2	109.6	Tongue
3	281.8	151.1	Tongue
4	660.1	242.3	Tongue
5	179.8	54.8	Buccal floor/tongue
6	31.2	21.7	Buccal floor/tongue
7	117.6	233.9	Buccal floor/tongue
8	213.3	125.8	Buccal floor/tongue
9	47.2	24.7	Pyriform sinus
10	176.5	14.9	Pyriform sinus
11	106.9	48.8	Oropharynx
12	95.9	38.9	Supraglottic larynx
13	187	-	Maxillary sinus
14	57.00	100.10	Esophagus
15	153.40	71.20	Esophagus
16	129.40	112.10	Esophagus

3.2. Protein extraction

Protein extraction was performed using the Mem-PER™ Plus Membrane Protein Extraction Kit (Thermo Scientific) according to the manufacturer's instructions with some exceptions, as described below.

3.2.1. Protein extraction from cell lines and membrane fraction preparation

Cells were first washed with PBS and then 8 mL/T150 flask of complete growth media were added to scrape adherent cells from the plastic surface. The harvested cell suspension was centrifuged at 300 x g for 5 min at room temperature (RT) and the supernatant removed.

Cell pellet was then washed with Cell Wash Solution (3 mL *per* $5x10^6$ cells), centrifuged at 300 x g for 5 min at RT and supernatant discarded. Cell pellet was ressuspended again in half the volume of Cell Wash Solution, transferred to a new tube and centrifuged at 300 x g for 5 min at RT. Supernatant was discarded. Next the cell pellet was ressuspended in Permeabilization buffer (750 µL *per* $5x10^6$ cells) with 1X protease inhibitor (PI) (complete, EDTA-free (Roche) and incubated for 10 min at 4ºC with constant mixing. Permeabilized cells were then centrifuged at 16000 x g for 15 min at 4ºC and the supernatant containing the cytosolic proteins was transferred to a new tube and stored at -20ºC. The cell pellet was then ressuspended in Solubilization buffer (500 µL *per* $5x10^6$ cells) with 1X PI and incubated for 30 min at 4ºC with constant mixing. Cells were centrifuged at 16000 x g for 15 min at 4ºC and the supernatant containing solubilized membrane and membrane-associated proteins was collected - membrane fraction 1 – and stored at -20ºC.

To maximize membrane protein extraction, an additional step with Solubilization buffer was added. Hence the cell pellet was again ressuspended in half the volume of Solubilization buffer with 1X PI and incubated for 30 min at 4ºC with constant mixing. Cells were centrifuged at 16000 x g for 15 min at 4ºC and the supernatant containing the solubilized membrane and membrane-associated proteins was collected - membrane fraction 2 – and stored at -20ºC. Both membrane fractions were later pooled together and stored at -20ºC.

3.2.2. Protein extraction from human tissues and membrane fraction preparation

All tissue fragments were separated from the matrix in a petri dish using razors and tweezers. After measuring their weight each fragment was washed with Cell Wash Solution (4 mL *per* 20-40 mg of tissue) with a brief vortex mixing and cut into small pieces with a razor in a petri dish.

All pieces were divided by tubes containing ceramic beads (Tissue homogenizing CKMix beads (Bertin Corp.)), with about 50 mg of tissue *per* tube, and Permeabilization buffer (1 mL *per* 20-40 mg of tissue) with 1X PI was added.

Tissues were homogenized using different methods. The first used a Minilys homogenizer (Bertin Instruments), where tissues were homogenized during 4 cycles of 30 sec at 5000 rpm, with a 30 sec pause on ice between cycles. Then the homogenate was recovered and an equal volume of Permeabilization buffer with 1X PI was added.

Esophagus samples were homogenized during 3 cycles of 30 sec at 6000 rpm, with a 30 sec pause on ice between cycles, in a Precellys instrument and the homogenate was recovered. Due to the presence of tissue that was not totally homogenized in the tubes, an additional step of homogenization with Permeabilization buffer was performed during 4 cycles of 30 sec at 6000 rpm, with a 30 sec pause on ice between cycles. Then this homogenate was recovered and mixed with the previously obtained.

The homogenates were incubated for 10 min at 4ºC with constant mixing, followed by a centrifugation at 16000 x g for 15 min at 4ºC. The supernatant, containing the cytosolic fraction, was collected and stored at -20ºC. The pellet was ressuspended in Solubilization buffer (500 µL *per* 20-40 mg of tissue) with 1X PI, incubated for 1 h at 4ºC with constant mixing and centrifuged at 16000 x g for 15 min at 4ºC. The supernatant containing membrane protein (membrane fraction 1) was saved and stored at -20ºC. This step was repeated using half the volume of solubilization buffer to increase membrane protein extraction and the supernatant (membrane fraction 2) was stored at -20ºC. Both membrane fractions were later pooled together and stored at -20ºC.

3.3. EGFR Immunoaffinity

EGFR was purified by immunoaffinity using cetuximab-coated magnetic beads. For that, 5 mg (~$3.3x10^8$) of Dynabeads M-270 Epoxy beads (Invitrogen) were resuspended in 0.1 M Sodium Phosphate buffer pH 7.4 and incubated for 10 min with tilting and rotation. Tubes were placed in a DynaMag™- 5 magnet (Life Technologies) and supernatant discarded. The beads were resuspended again in 0.1 M Sodium Phosphate buffer pH 7.4, vortexed for 30 sec and supernatant discarded.

Beads were incubated in a solution with 0.1 M Sodium Phosphate buffer pH 7.4, 40 µg of cetuximab (provided by Merck) in PBS and 3M Ammonium sulfate (ratio 1:1:1) ($1.1x10^9$ beads/mL) for 16 to 24 h at 37ºC, at 900 rpm in a ThermoMixer (Eppendorf). The tube was placed in the magnet and supernatant collected. The coated beads were washed twice with 1 mL PBS pH 7.4 - 0.1% BSA (bovine serum albumin) and twice with 1 mL PBS pH 7.4 by pipetting up and down, followed by washes in PBS pH 7.4 -0.1% Tween-20 and PBS pH 7.4 with rotation for 10 min.

Beads' saturation with cetuximab was evaluated by SDS-PAGE (Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis) and Coomassie Blue staining (3.4.5).

Cetuximab bound to the beads was de-glycosylated with 5U PNGase F (from *Elizabethkingia miricola*) in 25 mM Ammonium bicarbonate (AmBic) buffer at 37ºC at 300 rpm overnight in a ThermoMixer. At the end, the supernatant was discarded and the beads washed 4 times with 1 mL PBS pH 7.4.

To pull-down EGFR, coated beads were incubated for 1 h at 4ºC with membrane protein extracts from cells or tissues. Whenever possible, these incubations were repeated to promote EGFR saturation on the beads. This was verified by SDS-PAGE and Western blot analysis.

Next, the coated beads were extensively washed by pipetting up and down as it follows: 5 times with 1 mL PBS pH 7.4 - 0.1% Tween-20; 5 times with 1 mL PBS pH 7.4; 3 times with 1 mL PBS pH 7.4 - 1 M NaCl; 3 times with 1 mL PBS pH 7.4; 3 times with 1 mL PBS pH 7.4 - 1 M NaCl; 3 times with 1 mL PBS pH 7.4, 3 times with 1mL PBS pH 7.4 - 1 M Urea by tilting and rotation for 5 min and 3 times with 1 mL PBS pH 7.4. Afterwards, beads were incubated with 150 µL 0.1 M Citric acid pH 2.2, at 37ºC for 5min at 300 rpm. The supernatant was collected and immediately neutralized with 1 M Tris pH 8.5. Beads were then incubated with 0.5 M NH_4OH, 0.5 M EDTA pH 10.8, at RT for 5 min and shaking.

After that, tubes were changed and EGFR was eluted by incubating the beads 3 times with 0.1% RapiGest™ SF (Waters), diluted in 50 mM AmBic buffer at 100ºC for 10min at 500 rpm and once with 2X LDS sample buffer at 70ºC for 10 min at 500 rpm. Finally, beads were ressuspended in 2X LDS sample buffer.

All washes were concentrated using a Speed-Vac concentrator (Thermo Electron Corporation) in manual run mode, with a temperature of 35ºC and vacuum pressure of 10 vac. EGFR expression was evaluated in the concentrated washes and elutions by SDS-PAGE and Western blot analysis.

3.4. Analytical methods

3.4.1. Determination of cell concentration and viability

Cell concentration and viability were determined using the trypan blue exclusion assay. Trypan blue is a cell-impermeable dye, staining in blue non-viable cells that are membrane damaged while viable cells appear colorless. Cells were incubated with a 0.1% Trypan blue solution and viable and non-viable cells were counted in a Fuchs-Rosenthal hemocytometer chamber under a light field microscope.

3.4.2. Total protein quantification

Total protein quantification from cytosolic and membrane fractions from cell lines and tissue fragments was performed using the Micro BCA™ (Bicinchoninic acid assay) Protein assay kit (Thermo Fisher Scientific), according to manufacturer's protocol. BSA was used for the calibration curve. The assay took place in a clear 96-well plate (Falcon) and the absorbance was measured at 562 nm on a Tecan Infinite 200 Pro NanoQuant equipment.

3.4.3. EGFR quantification by ELISA

EGFR quantification was performed in flat-bottomed MaxiSorp 96 well plates (Sigma-Aldrich). Plates were coated with 1 µg/mL of cetuximab diluted in a carbonate-bicarbonate buffer pH 9.6 (0.035 M Na_2CO_3, 0.015 M $NaHCO_3$) overnight at 4ºC and blocked with PLI-P

buffer pH 7.4 (0.5 M NaCl, 0.003 M KCl, 0.0015 M KH_2PO_4, 0.0065 M Na_2HPO_4,2H2O, 1% BSA, 1% Triton X-100) for 1 h at RT. Next, wells were incubated with protein extracts from an initial concentration ranging between 250 to 50 µg/mL serially diluted in PLI-P buffer for 2h at RT. The standard curve was established with a recombinant EGFR protein (produced in Hi5 cells) serially diluted from an initial concentration of 30 ng/mL and incubated in the same conditions. Binding of EGFR was detected by incubating with primary antibody Matuzumab-biotinylated (1 µg/mL) for 1 h. Subsequently, wells were incubated with 1 µg/mL of Streptavidin-HRP (Life Technologies) for 1 h. Between all incubation steps several washes were performed with PBS-T (1xPBS, 0.05% Tween-20). Plates were developed with TMB+ (Dako) and the reaction stopped with 0.5 M H_2SO_4. Absorbance was measured at 450 nm on Tecan Infinite 200 Pro NanoQuant instrument. As controls, protein extracts were substituted by PLI-P buffer and wells with the maximum concentration of protein extract were also tested without Matuzumab-biotinylated antibody.

3.4.4. EGFR glycoprofiling by ELISA

The analysis of glycans structures present on EGFR was performed in similar conditions as described above. In brief, after coating with cetuximab and blocking, wells were incubated with membrane protein extracts from cell lines serially diluted in PLI-P buffer from an initial concentration of 50 ng/mL for 2 h at RT. Following washing, wells were incubated with 1 µg/mL of biotinylated lectins (Table 3.2) for 1 h at RT. The rest of the protocol was performed as described in the preceding section. Negative control wells included substitution of protein extracts by PLI-P buffer. Protein extracts were also tested at the maximum with 1 µg/mL of Matuzumab-biotinylated antibody as a control for biotin-streptavidin interaction.

Table 3.2 – Lectins used in this study and their specificity.

Lectin	Supplier	Specificity
WGA (*Triticum vulgaris*)	Galab	GlcNAc Sialic acid
PHA-E (*Phaseolus vulgaris* erythroagglutinin)	Vector Labs	Gal(β4)GlcNAc(β2)Man(α6) (GlcNAc(β4)) (GlcNAc(β4)Man(α3)) Man(β4)
WFL (*Wisteria floribunda*)	Vector Labs	N-GalNAc (LAcdiNAc)
SNA (*Sambucus nigra*)	Galab	Sialic Acid-(α-2,6) -Gal/GalNAc Sialyl Tn
PNA (*Arachis hypogaea*)	Galab	Gal-(β1,3)-GalNAc (T antigen)
ECL (*Erythrina cristagalli*)	Galab	Gal-(β1,4)-GalNAc Gal-(β1,4)-GlcNAc
AAL (*Aleuria aurantia*)	Galab	Fuc-(α-1,6)-GlcNAc Fuc-(α-1,2)-Gal-(β1,4)-(Fucα-1,3/4)Gal β1-4LacNAc
GNL (*Galanthus nivalis*)	Vector labs	Man-(α1,3)Man (Man-(α1,6)Man)

3.4.5. SDS-PAGE

Samples were prepared with 1X NuPAGE LDS sample buffer (Novex) and 1X NuPAGE sample reducing agent (Novex), followed by denaturation at 70ºC for 10 minutes. Samples were subjected to electrophoresis on pre-casted gels NuPAGE 4-12% Bis-Tris (ThermoFischer Scientific) while using 1X NuPAGE MES Running buffer (Novex). The molecular marker used was SeeBlue®Plus 2 prestained standard (ThermoFischer Scientific).

For protein extract samples, electrophoresis was performed at constant voltage of 200 V for 40 min. For analysis of concentrated washes and elutions samples obtained from immunoaffinity purification procedure, electrophoresis was performed at 80 V for 2 h and at 100 V for 40 min. For samples containing high concentration of citric acid, in order to be further prepared and analyzed by MS, electrophoresis was performed at 80 V for 1 h 20 min and at 100 V for 1 h 10 min. In these two cases the tank was maintained on ice to avoid overheating.

Gels were stained either with Coomassie Blue (InstantBlue, Expedeon) for 1 h under gentle agitation, or using SilverQuest™ Silver Staining Kit according to manufacturer's protocol.

3.4.6. Western blot

Gels were blotted to polyvinylidene difluoride (PVDF) membranes (iBlot® Transfer Stacks) (ThermoFischer Scientific) using the program P3 for 7 min in iBlot® Transfer Device (ThermoFischer Scientific). Transfer efficiency was assessed by Coomassie Blue staining the gels. Membranes were blocked in 5% of albumin fraction V (Merck) diluted in tris buffered saline pH 8.0 (Merck) with 0.1% tween 20 (Merck) (TBS-T) for 1 h at RT, under gentle agitation. Membranes were then incubated with a rabbit monoclonal antibody – anti-EGFR (clone 15F8) (4405S, Cell Signaling Technology) diluted 1:2000 in blocking solution, for 2 h or overnight at 4ºC under gentle agitation. After washing thrice for 10 min with TBS-T, membranes were incubated with horseradish peroxidase-conjugated goat anti-rabbit antibody (A9169, Sigma) diluted 1:80000 in TBS-T for 1 h at RT, under gentle agitation. Following washing, blots were developed with the enhanced chemiluminescence detection system (ECL) prime western blotting (Amersham Biosciences) and the digital image acquired using ChemiDoc™ XRS+ (Bio-Rad).

3.5. Mass spectrometry analysis

The protein and glycan content present in membrane protein extracts or in elution samples with 0.1% RapiGest™ SF and citric acid were analyzed by mass spectrometry at the UniMS-ITQB/iBET facility.

Samples run in SDS-PAGE gels were stained as described above and the area between the molecular markers 198 and 98 kDa was cut. Gel pieces were firstly destained with 50 mM

AmBic in 50% (v/v) acetonitrile (ACN) for 15 min with agitation, and then destaining solution was discarded. This step was repeated until the gel pieces were transparent.

Next, samples were dehydrated, by incubating with 100% ACN for 15 min and then supernatant was discarded. Samples were after reduced with 10 mM DTT (Dithiothreitol) in 50 mM AmBic for 45 min at 56 ºC and alkylated with 55 mM iodoacetamide (IAA) in 50 mM AmBic for 30 min at RT, in the dark. Supernatant was discarded.

The gel pieces containing the samples were again dehydrated, by incubating with ACN for 15 min, liquid was removed and the gel pieces were dried for 10 min at 37ºC. An *in gel* digestion with 2.5 units of PNGase F (Sigma) in 20 mM AmBic was performed overnight at 37ºC with agitation. Glycans were obtained in the supernatant and further extracted from the gel by incubating the gel pieces with water LC-MS grade 3 times for 30 min in an ultrasound bath. The sample obtained from the previous digestion was then dried on Speed-Vac and ressuspended in 0.1% of formic acid (FA).

After that, gel pieces were dehydrated again (incubated twice for 15 min with ACN) and dried for 10 min at 37 ºC. A second digestion with 10 ng/mL of trypsin/lys-C (Promega) in 50 mM AmBic was performed overnight at 37ºC with agitation. The digestion was stopped with 5% of FA and proteins were obtained in the supernatant by two alternated incubations of the gel pieces with ACN and water LC-MS grade for 10 min in an ultrasound bath. The sample obtained from the digestion was then dried on Speed-Vac and ressuspended in 0.1% FA.

In-solution digestion was also tested. Briefly, after the samples being reduced with 10 mM DTT for 40 min at 56 ºC and alkylated with 20 mM IAA for 30 min in the dark at RT, PNGase F digestion was performed overnight (500 units/mL in 5 mM potassium phosphate) to obtain the glycans and proteins were digested overnight with trypsin/lys-C (0.1 µg/µL in 0.001% HCL).

The glycans and tryptic peptides obtained were analyzed by NanoLC MSMS using an ekspert™ NanoLC 425 cHiPLC® system coupled with a TripleTOF® 6600 with a NanoSpray® III source (Sciex).

Peptides were separated through reversed-phase chromatography (RP-LC) in a trap-and-elute mode. Trapping was performed at 2 µL/min with 100% A (0.1% formic acid in water, Fisher Chemicals, Geel, Belgium), for 10 min, on a Nano cHiPLC Trap column (Sciex 200 µm x 0.5 mm, ChromXP C18-CL, 3 µm, 120 Å). Separation was performed at 300 nL/min, on a Nano cHiPLC column (Sciex 75 µm x 15 cm, ChromXP C18-CL, 3 µm, 120 Å). The gradient was as follows: 0-1 min, 5% B (0.1% formic acid in acetonitrile, Fisher Chemicals); 1-46 min, 5-35% B; 46-48 min, 35-80% B; 48-54 min, 80% B; 54-57 min, 80-5% B; 57-75 min, 5% B.

Peptides were sprayed into the MS through an uncoated fused-silica PicoTip™ emitter (360 µm O.D., 20 µm I.D., 10 ± 1.0 µm tip I.D., New Objective). The source parameters were set as follows: 12 GS1, 0 GS2, 30 CUR, 2.5 keV ISVF and 100 ^{0}C IHT. An information dependent acquisition (IDA) method was set with a TOF-MS survey scan of 400-2000 m/z for 250 msec. The 50 most intense precursors were selected for subsequent fragmentation and the MS/MS were acquired in high sensitivity mode (150-1800 m/z for 40 msec each). The selection criteria

for parent ions included a charge state between +2 and +5 and counts above a minimum threshold of 125 counts per second. Ions were excluded from further MSMS analysis for 12 s. Fragmentation was performed using rolling collision energy with a collision energy spread of 5.

The obtained spectra were processed and analyzed using ProteinPilot™ software, with the Paragon search engine (version 5.0, Sciex). The following search parameters were set: search against Homo sapiens from Uniprot/SwissProt database (release 2015_05); Iodoacetamide, as Cys alkylation; Tryspsin, as digestion; TripleTOF 6600, as the Instrument; ID focus as biological modifications and Amino acid substitutions; search effort as thorough; and a FDR analysis. Only the proteins with Unused Protein Score above 1.3 and 95% confidence were considered.

Glycans were separated through graphitic carbon column in a trap-and-elute mode. Trapping was performed at 2 µl/min with 100% A (0.1% formic acid in water, Fisher Chemicals, Geel, Belgium), for 10 min, on a Nano cHiPLC Graphitic Carbon trap column (200 µm x 0.5 mm, 3 µm, 250 Å). Separation was performed at 300 nL/min, on a Nano cHiPLC Graphitic Carbon analytical column (75 µm x 150 mm, 3 µm, 250 Å). The gradient was as follows: 0-1 min, 5% B (0.1% formic acid in acetonitrile, Fisher Chemicals); 1-40 min, 5-40% B; 40-42 min, 40-90% B; 42-56 min, 90% B; 56-58 min, 90-5% B; 58-72 min, 5% B.

Glycans were sprayed into the MS through an uncoated fused-silica PicoTip™ emitter (360 µm O.D., 20 µm I.D., 10 ± 1.0 µm tip I.D., New Objective). The source parameters were set as follows: 15 GS1, 0 GS2, 30 CUR, 2.5 keV ISVF and 100 ^{0}C IHT. An information dependent acquisition (IDA) method was set with a TOF-MS survey scan of 300-1800 m/z for 250 msec. The 30 most intense precursors were selected for subsequent fragmentation and the MS/MS were acquired in high sensitivity mode (150-1800 m/z for 80 msec each). The selection criteria for parent ions included a charge state between +2 and +5 and counts above a minimum threshold of 125 counts per second. Ions were excluded from further MSMS analysis for 6 s. Fragmentation was performed using rolling collision energy with a collision energy spread of 5.

The obtained spectra were filtered for the presence of glycans MSMS marker ions using PeakView software version 2.2 (Sciex) and glycans identification was performed using SimGlycan software version 5.80 or GlycoMod from Expasy.

Several proteins (recombinant EGFR (Abcam, HEK293 derived) and IgA (Sigma, human colostrum derived)) were tested for the implementation and optimization of sample preparation and LC-MS procedures according to the described above.

4. Results and discussion

4.1. Cell culture and EGFR expression

Cell lines representative of HNC and EC were expanded in vitro to obtain cells from where EGFR could be extracted. In parallel human keratinocytes were also cultured.

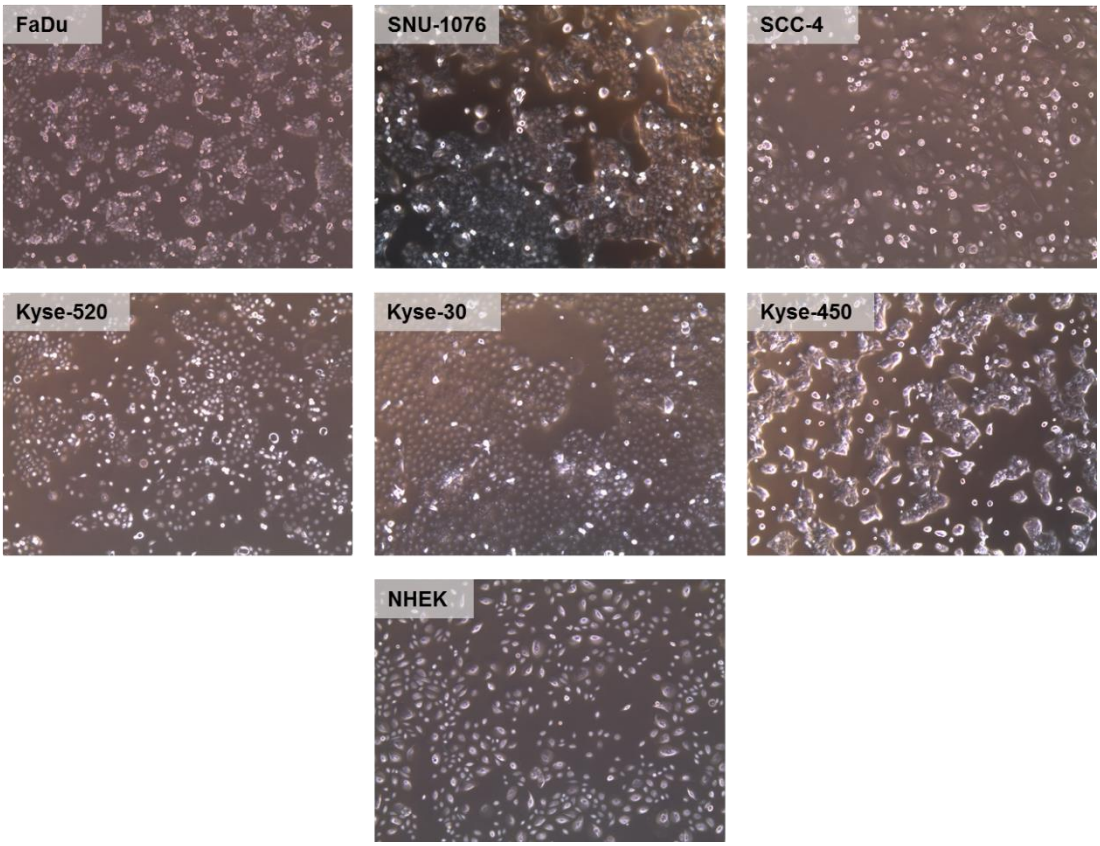


Figure 4.1 – Cell morphology of cultured cells. Light field microscopy images from FaDu, SNU 1076, SCC-4, Kyse 520, Kyse 30, Kyse 450 and NHEK cell lines in adherent culture. Magnification 50x.

Figure 4.1. shows microscopy images of the different cultured cells. FaDu cell line, derived from human pharynx squamous cell carcinoma, SNU-1076 cell line, derived from human larynx cancer, and SCC-4 cell line, derived from human tongue squamous cell carcinoma, are representative of head and neck cancer. Kyse-520, Kyse-30 and Kyse-450 cell lines, derived from human esophageal squamous cell carcinomas, are representative of esophageal cancer. NHEK, derived from normal human epidermal keratinocytes, is representative of normal cells that proficient in EGFR.

All cell cultures were performed in adherence, under rigorous conditions of asepsis and controlled factors such as temperature (at 37ºC), pH (near 7.4), osmolality and gas concentration of 5% carbon dioxide [148]. Despite the same culture conditions used, there were differences in cellular growth between cells.

These cells present different forms of proliferation, forming diverse structures. FaDu and Kyse-450 cell lines grow radially and form groups of cells, until forming a cellular sheet. SNU-

1076, Kyse-520, Kyse-30 and NHEK have a radial uniform growth, leaving some hollow spaces. SCC-4 presents a radial and also axial growth, forming a layer of cells when high confluency is reached.

All cancer cell lines reached 90% confluency in 3 days, presenting a faster cellular growth when comparing with NHEK cells, that reached 90% confluency in 7 days. Cancer cell lines are expected to present a faster growth, since they derive from more proliferative cells, whereas NHEK are normal cells with primary origin and are not immortalized. Thus, these cells have a limited number of passages possible leading to limitations in the number of cells attained and consequent obtained protein amount.

Protein extracts were produced from all cultured cells using Mem-PER™ Plus Membrane Protein Extraction Kit. The optimized methodology is schematized in Figure 4.2.

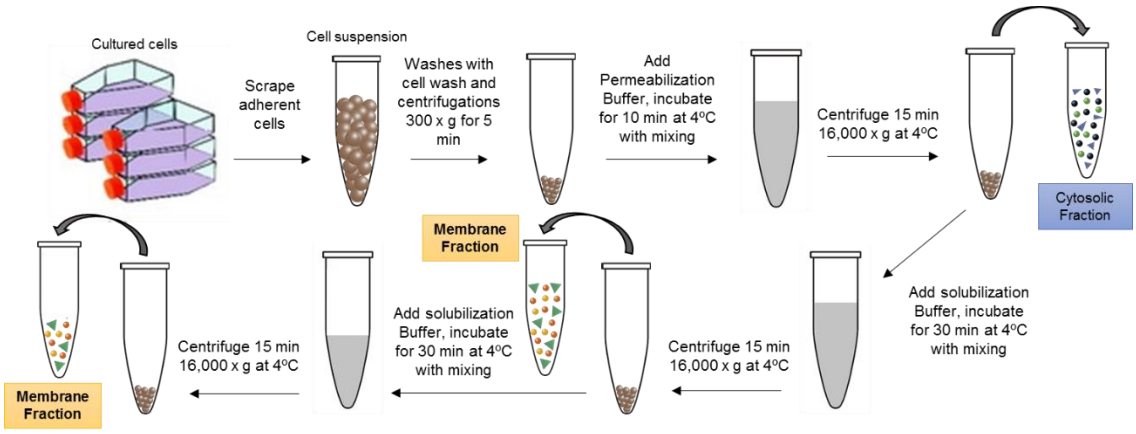


Figure 4.2 - Schematic representation of the procedure implemented for protein extraction from cultured cells and membrane fraction preparation.

This procedure uses two solutions with different proprietary detergent concentrations (and other components) to obtain cytosolic and membrane fractions. The first solution applied was the permeabilization buffer, that provokes damages in the lipid bilayer of cell membrane, allowing the extraction of soluble cytosolic proteins. Proceeding with a pellet a second solution used was the solubilization buffer, that solubilizes the membrane, extracting integral membrane proteins and membrane-associated proteins, in which EGFR is included. Hence, we chose this kit in order to obtain enriched fractions in EGFR, namely within the membrane fraction.

Further optimization of this procedure, aiming at obtaining the highest amount of EGFR in the membrane fraction, led us to an additional incubation with the solubilization buffer. We confirm by Western Blot (data not shown) that this additional step increased the amount of EGFR extracted, so it was added to the extraction protocol.

Table 4.1 indicates for each cultured cell the amount of total protein obtained in pooled cytosolic and membrane fractions for all the protein extractions performed. The values were normalized considering the number of cells used in each cell line.

Table 4.1 – Quantities of total protein for each cultured cell. The amount of total protein was determined by MicroBCA.

Cultured Cells		Cytosolic fraction		Membrane fraction	
		Total protein (µg)	Total protein (µg/10^6 cells)	Total protein (µg)	Total protein (µg/10^6 cells)
HNC	FaDu	47372.59	147.72	20192.88	62.97
HNC	SNU-1076	11411.55	20.97	18668.71	34.31
HNC	SCC-4	2875.14	19.25	5181.29	34.70
EC	Kyse-520	1196.21	24.59	1880.08	38.65
EC	Kyse-30	7509.86	120.74	10136.80	162.97
EC	Kyse-450	10795.14	75.60	11443.09	80.13
NHEK		474.19	25.32	2709.48	144.66

It is noticeable that the quantity of total protein obtained *per* million of cells differs between cells. In this way, SNU-1076 and SCC-4 cell lines generated the lowest protein yields of both cytosolic and membrane fractions, while FaDu and Kyse-30 produced the highest amounts of protein in the cytosolic and membrane fractions, respectively.

The diverse yields obtained from protein extraction can result from the different characteristics of the cells. Indeed, the protocol used has been tested with other cell lines, including C6, NIH-3T3 and HeLa, and these produced different amounts of protein for the same number of cells [149].

The presence of EGFR in the cytosolic and membrane fractions pool of each cultured cell type was evaluated by Western Blot (Figure 4.3). The amount of loaded protein depended on the highest volume of sample that could be applied on the gel due to their protein concentration.

Detection of EGFR was performed using a monoclonal antibody (clone 15F8) that recognizes residues near the C-terminal region of EGFR.

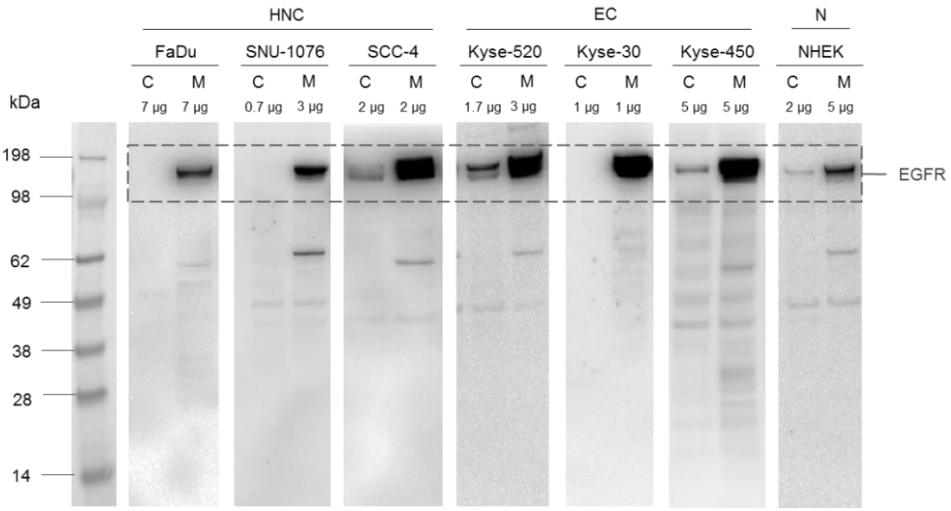


Figure 4.3 - Western blot analysis of EGFR in cytosolic and membrane fraction pools of cultured cells. The samples were loaded with different amounts of protein in Novex™ NuPAGE™ 4 – 12% Bis-Tris gel and

transferred to a PVDF membrane. The primary antibody used was a rabbit anti-EGFR (clone 15F8) monoclonal antibody, diluted 1:2000, and the secondary antibody was horseradish peroxidase-conjugated goat anti-rabbit antibody, diluted 1:80000.

Figure 4.3 shows that EGFR is present in the membrane fractions of all cells under study, but it was also detected in cytosolic fractions of SCC-4, Kyse-520, Kyse-450 and NHEK. The cytosolic fractions from FaDu, SNU-1076 and Kyse-30 may also contain EGFR present in solution that was not detected due to the amounts loaded. The presence of EGFR in cytosolic fraction means that membranes were also partly solubilized by permeabilization buffer. This can happen due to long incubation time or strong agitation. However, there had to be a compromise, because the main objective is to obtain an enriched membrane fraction with a minimum amount of cytosolic proteins present. Nevertheless, the extracted EGFR was mostly in membrane fractions all cells.

In order to better visualize the differences in EGFR band present in membrane fractions of all cells under study, an additional WB was performed (Figure 4.4).

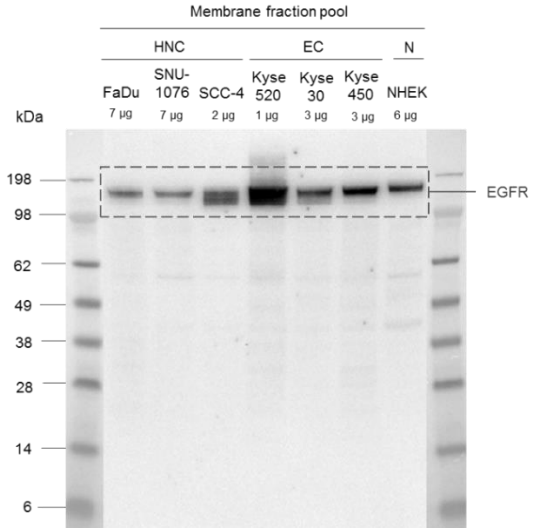


Figure 4.4 – Western blot analysis of EGFR in the membrane fraction pool of all cells. The samples were loaded with different amounts of protein in Novex™ NuPAGE™ 4 – 12% Bis-Tris gel and transferred to a PVDF membrane. The primary antibody used was a rabbit anti-EGFR (clone 15F8) monoclonal antibody, diluted 1:2000, and the secondary antibody was horseradish peroxidase-conjugated goat anti-rabbit antibody, diluted 1:80000. This image was obtained with a development of 5 seconds.

Figure 4.4 shows that FaDu, SNU-1076 and NHEK present only one band that is expected to correspond to the 170 kDa full-length EGFR [35]. In SCC-4, Kyse-520, Kyse-30 and Kyse-450 two bands can be observed, one corresponding to full-length EGFR and the other that may correspond to the mutant EGFRvIII, with 145 kDa [35]. This mutant is generally found together with full-length EGFR and it is present in HNC and EC cases [55, 56], but not in normal cells [150], which is consistent with its absence in NHEK.

This EGFR variant can be recognized by the antibody used in the WB analysis, since this antibody recognizes residues that are also present in this mutant [35].

Differences in the intensity of the lower band are also observed between cells, indicating a differential expression of this mutant in the different cell lines. Indeed, it has been reported

that changes in the expression level of EGFRvIII can be related with cancer cells phenotype and development stage [55].

Quantification of EGFR in the membrane fraction was assessed by a sandwich type ELISA, using two different antibodies (cetuximab and matuzumab), recognizing distinct epitopes on domain III in the extracellular part of EGFR [151, 152].

Of notice is that these antibodies bind not only to the full-length protein but can also bind to other variants that contain domain III, such as EGFRvIII [153].

For the calibration curve we used a recombinant EGFR, comprising the amino acids 25 to 642 from the ectodomain, produced in Hi5 insect cells. Thus, it is expected to present less elongated N-glycans, such as oligomannose and paucimannose structures, and lack complex-type sugars present in mammalian cells [154-156]. Nevertheless, this does not interfere with antibody binding [152].

Table 4.2 shows the amount of EGFR obtained in membrane fractions pools for all the protein extractions performed.

Table 4.2 – Quantities of EGFR obtained in membrane protein extractions for each cultured cell. The amount of EGFR was established by ELISA.

Cultured Cells		Membrane fraction		Ratio
		Total EGFR (ng)	Total EGFR (ng/10^6 cells)	EGFR/Total Protein (ng/µg)
HNC	FaDu	12108.03	37.76	0.60
	SNU-1076	11255.61	20.68	0.60
	SCC-4	21054.25	140.99	4.06
EC	Kyse-520	6710.71	137.97	3.57
	Kyse-30	14193.75	228.20	1.40
	Kyse-450	16672.10	116.75	1.46
NHEK		1882.93	100.53	0.69

We have observed differences between cells under study regarding the quantity of EGFR obtained *per* million of cells (Table 4.2). Indeed, FaDu and SNU-1076 cell lines produced the lowest amounts of EGFR, while Kyse-30 originated the highest quantities.

When we analyzed the ratio of EGFR amount to total membrane protein we noted that SCC-4 and Kyse-520 presented the highest ratios, in comparison to the other cell lines. In contrast, FaDu and SNU-1076 cell lines showed the lowest ratios, very similar to what is observed with NHEK.

These EGFR expression differences between these cultured cells were expected, since EGFR expression is also variable within tissues [150, 157]. Also, the protein extraction yields obtained were different, which means that the total EGFR expressed by the cells might not be entirely represented in the obtained membrane fractions.

4.2. Implementation of a lectin ELISA for EGFR glycoprofiling

EGFR is a glycoprotein highly N-glycosylated with up to 11 potential N-linked canonical glycosylation sites in its extracellular portion and is also O-glycosylated [118, 119, 125]. In cancer, aberrant EGFR glycosylation has been found and shown to affect its functions [54], [125]. Thus, EGFR in cancer is expected to show aberrant structures, different from normal cells. By comparing these structures between cancer and normal cells, a possible target can be discovered to be applied in targeted therapies towards cancer cells and not normal tissue.

For that, an initial characterization of EGFR glycosylation was performed by sandwich ELISA using lectins that recognize specific determinants on glycan structures, schematized in Figure 4.5.

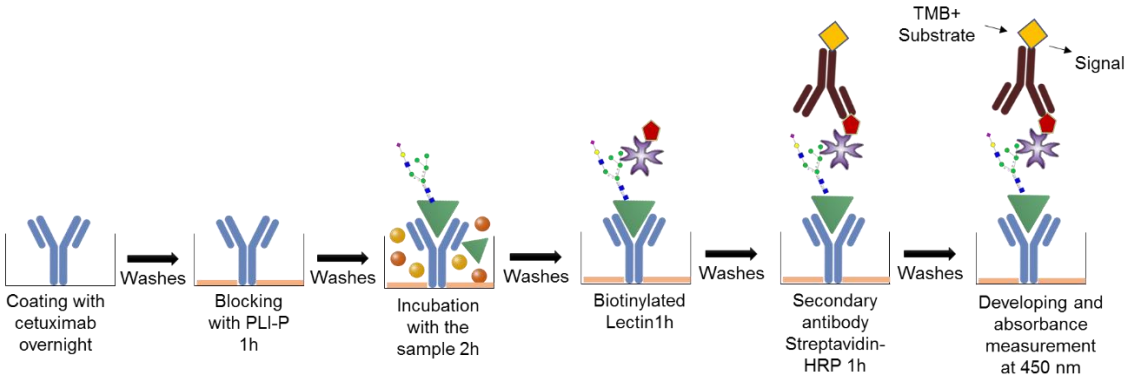
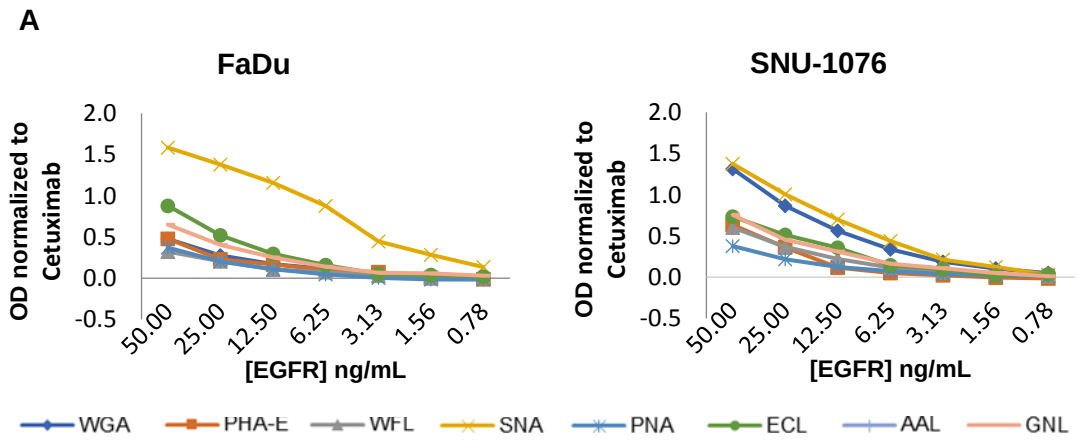
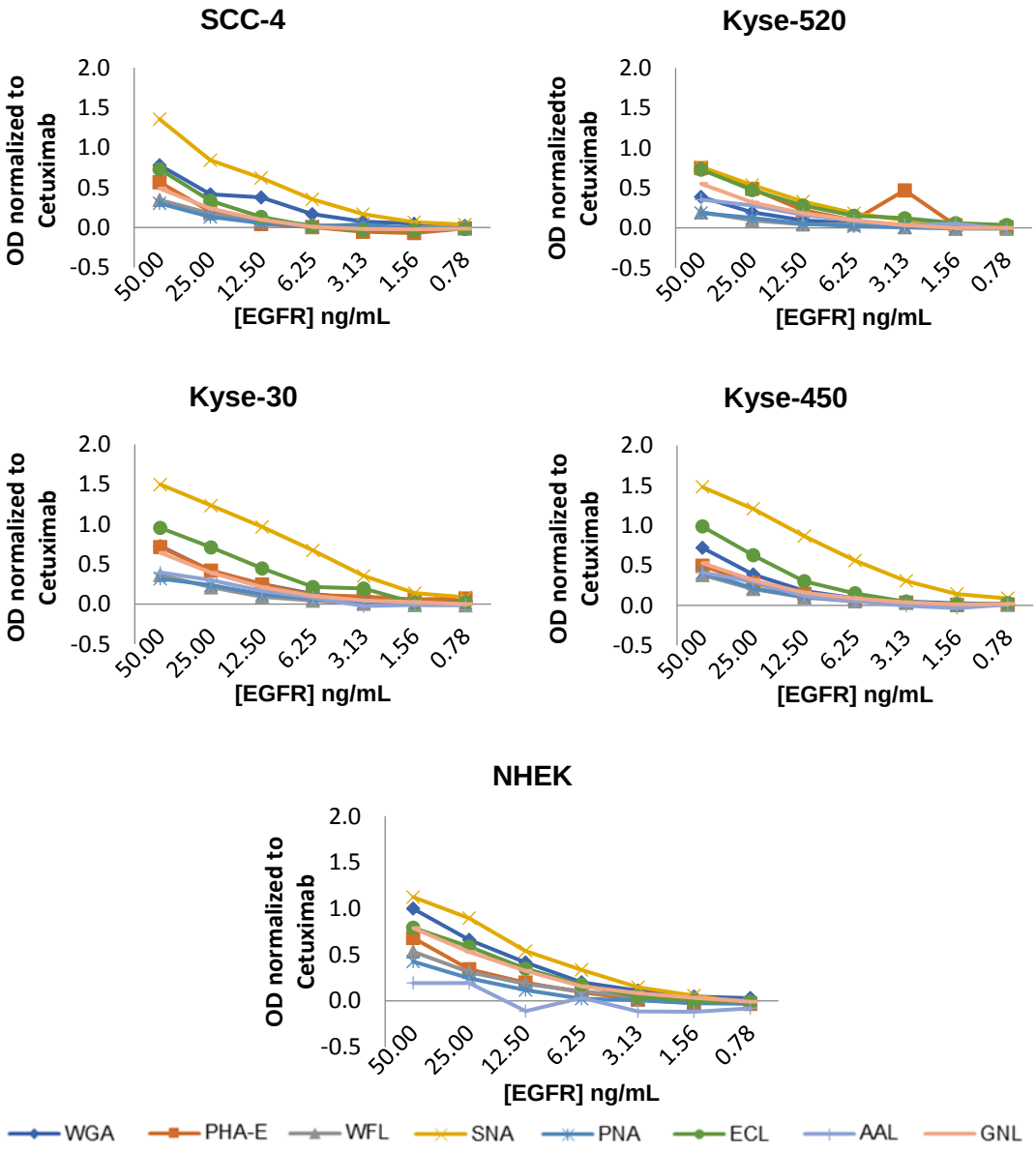


Figure 4.5 - Schematic representation of the lectin ELISA implemented for EGFR glycoprofiling.

In this assay, the membrane fraction of each cultured cells was incubated in a plate previously coated with cetuximab. This antibody allows the capture of EGFR to be further probed with biotinylated lectins. Since cetuximab is also glycosylated, the background of this assay can sometimes be very high. Thereat, the absorbance for EGFR had to be normalized for cetuximab related background.

The results of lectin glycoprofiling are presented in Figure 4.6, in which a correspondent obtained signal of each lectin is represented for each cell under study.





B

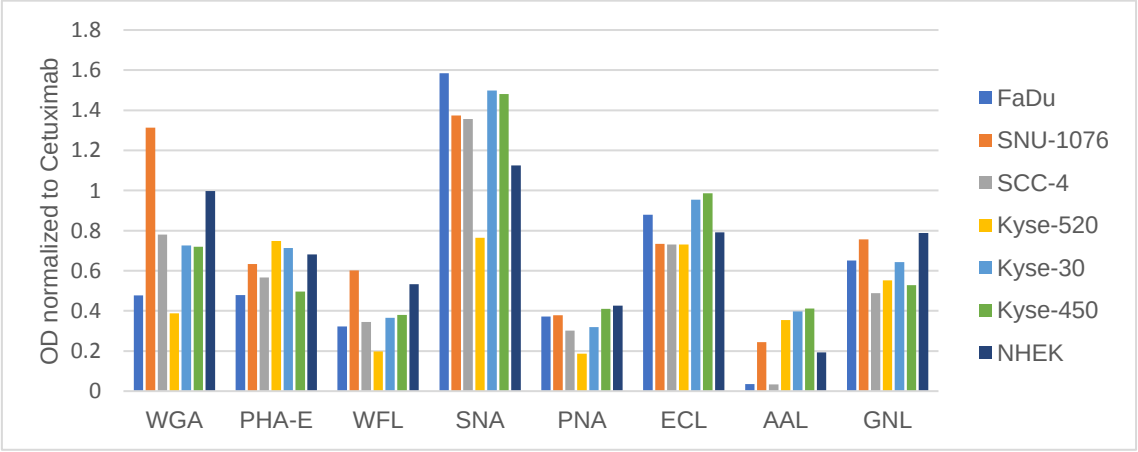


Figure 4.6 – Glycoprofiling of EGFR in membrane fractions from cells by sandwich ELISA**.** **A**) The glycan structures of EGFR were characterized by WGA, PHA-E, WFL, SNA, PNA, ECL, AAL and GNL lectins. Membrane fractions of FaDu, SNU-1076, SCC-4, Kyse-520, Kyse-30, Kyse-450 and NHEK cell lines were incubated with serial dilutions of EGFR. Absorbance was measured at 450 nm. **B**) Summary of the normalized OD values obtained at 50 ng/mL of EGFR for all the cells and lectins tested.

Lectin glycoprofiling (Figure 4.6) shows that SNA lectin corresponding signal was the highest in all cells, except for Kyse-520. This lectin recognizes sialic acid α2-6-linked to terminal Gal or GalNAc [158], which indicates that the analyzed EGFR has this structure in complex or hybrid type of N-glycans. This lectin also recognizes sialyl-Tn (Neu5Ac-(α2,6)-GalNAc-O-Ser/Thr) antigen, which means that this structure can also be present on EGFR [159]. A study by Wu and colleagues have previously identified EGFR as a carrier of sialyl-Tn in cell lysates from tissue and cell lines of hepatocellular carcinomas [160].

ECL also presented a generally elevated reactivity, particularly with EC cell lines (Kyse-30 and Kyse-450). This lectin reacts primarily with LacNAc structures (Gal-(β1,4)-GlcNAc) is also reactive to lesser degree with GalNAc and weaker still with Gal [161]. The elevated reactivity indicates the existence of these structures on EGFR. This lectin does not tolerate sialylation, which could also contribute for differences in ECL-reactivity.

The signal of WGA lectin was also broadly high, especially in NHEK and SNU-1076 cell line. This lectin recognizes GlcNAc-(β1,4)-GlcNAc and sialic acid structures [162, 163]. The GlcNAc recognition is expected, since all N-glycans contain this structure, although WGA has a higher affinity to hybrid-type than to complex-type or high-mannose type N-glycans [97].

The PHA-E and GNL lectins signal was not so intense. PHA-E has a specificity towards biantennary galactosylated complex-type N-glycans with bisecting GlcNAc [164], while GNL binds preferentially Man-(α1,3)-Man in high-mannose type N-glycans [165].

In general, WFL, PNA and AAL lectins showed the lowest reactivities in all cell lines. WFL recognizes GalNAc-(β1,4)-GlcNAc (LacdiNAc) or Gal-(β1,3/6)-GalNAc structures, PNA recognizes essentially T antigen (Gal-(β1,3)-GalNAc-O-Ser/Thr) [166] and AAL has a broad specificity to fucosylated glycans. These results mean that these structures are present in low quantities. Regarding PNA reactivity, it can be prevented by the existence of sialic acid, which can possibly explain the low affinity observed. If this was to be true, neuraminidase treatment would allow PNA binding.

When comparing cancer cell lines with NHEK we observe that the most prominent difference is the higher levels of sialic acid structures in cancer cell lines.

In a study of colorectal cancer, using both cell lines and tissue, it was found in EGFR an increase in α2,6-sialylation [54, 98], which is consistent with our results with SNA lectin. Another study using cell lines and *in vivo* models of ovarian cancer demonstrated that increased sialylation of EGFR led to its activation and to a more aggressive tumor phenotype [130, 131]. In our study, we also observed an increased sialylation of EGFR in cancer cell lines derived from HNC and EC, thus showing that this effect on EGFR is common to several cancer types.

Despite the obtained lectin glycoprofiling results, a more thorough characterization of EGFR glycosylation was required. For that, purification or enrichment of EGFR was necessary in order to perform MS analysis.

4.3. Implementation of EGFR immunoaffinity purification / enrichment

In order to perform EGFR immunoaffinity purification/enrichment (IP) we developed a method that is schematized in Figure 4.7.

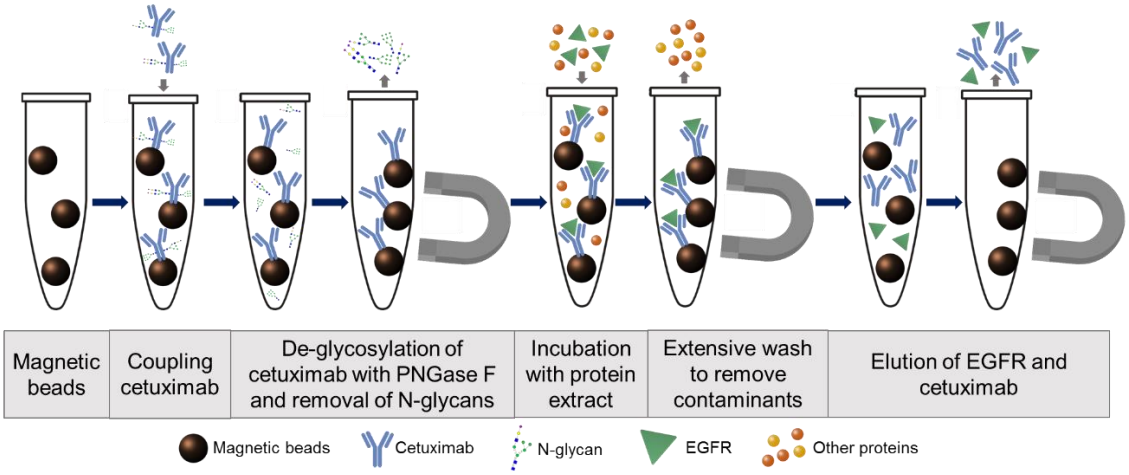


Figure 4.7 – Schematic representation of EGFR immunoaffinity purification/enrichment procedure implemented.

Our procedure used epoxy magnetic beads covalently linked with cetuximab. Saturation of the beads with the antibody was assessed by analyzing the resulting unbound fraction by SDS-PAGE followed by InstantBlue™ Coomassie protein staining (data not shown). The next step included de-glycosylation of cetuximab with PNGase F to avoid that glycans from the antibody interfere with MS analysis of EGFR specific glycan profile at the end of the process. PNGase F removes N-glycans by cleaving the N-glycosidic bond between the Asn residue and GlcNAc, converting the Asn into Asp through deamination [168].

Then, beads were consecutively incubated several times with each membrane fraction in order to saturate the antibody-coated beads with EGFR. It was noted that EGFR binding to cetuximab seemed to vary with concentration of protein extract despite the same quantity of EGFR, i.e. in more diluted samples EGFR has possibly fewer interactions with the antibody, resulting in less EGFR bound to the coated beads.

In order to remove contaminant proteins, we performed extensive washes with different solutions, including PBS-T, PBS, PBS 1 M NaCl and PBS 1 M Urea. However, we noted loss of beads attached to the tips.

EGFR elution was achieved with citric acid pH 2.2 and RapiGest™ SF. The low pH of citric acid disrupts the EGFR and antibody bond [169], though not all EGFR is eluted with this solution. Thus, incubations with RapiGest™ SF surfactant at high temperature were necessary to elute all EGFR. However, the covalent bond between the antibody and the beads is broken at high temperature, resulting in contamination of cetuximab in these elution fractions. Due to this, the initial step of cetuximab de-glycosylation is of major importance.

During the establishment of this method several optimizations were performed to maximize final EGFR yield. For that, we increased EGFR quantity *per* incubation with membrane fraction and performed several incubations to promote saturation of the beads. To minimize loss of beads during the washes, we used *LoBind* tubes, which also reduced binding of proteins to the walls of the tubes.

To assess the amount of EGFR lost during washes, fractions were concentrated 4 times on a Speed-Vac, to facilitate detection in further assays.

The assessment of the immunoprecipitation procedures was performed by ELISA (data not shown). However, results were inconclusive, so we investigated if it was related with solution composition affecting absorbance readings (Figure 4.8).

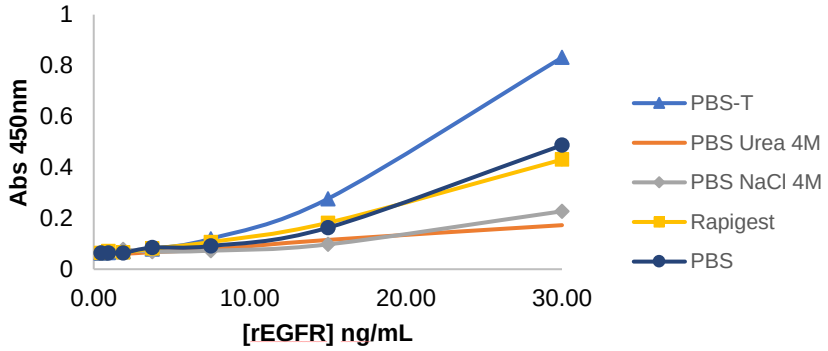


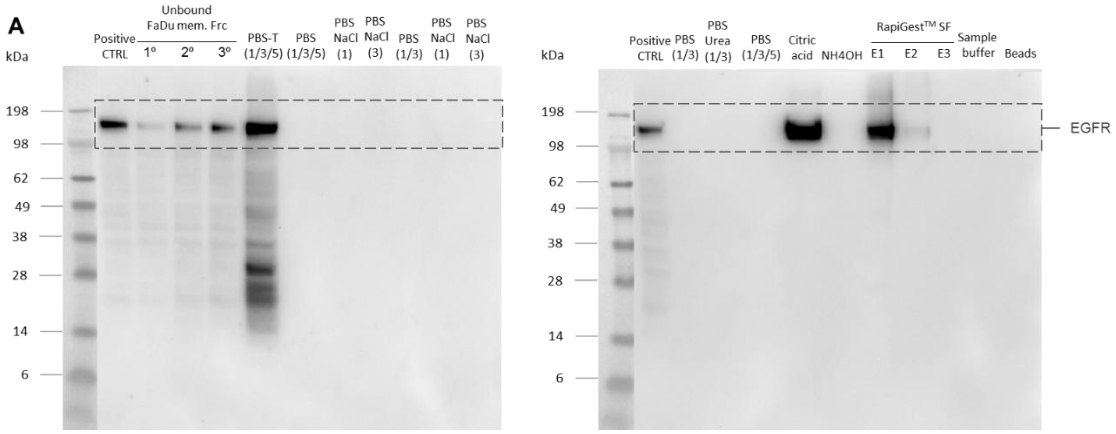
Figure 4.8 – Effect of different solutions used in immunoprecipitation procedure in EGFR quantification by ELISA**.** Recombinant EGFR (rEGFR), produced in Hi5 cells, was diluted in different solutions and incubated on a plate in the same conditions. The reference value was the dilution in PBS (which is the regular solvent used in ELISA quantifications). rEGFR was also diluted in PBS-T, PBS with 4M urea and 4M NaCl (that corresponds to the concentration of the samples after being concentrated on a Speed-Vac) and in RapiGest™ SF solution 0.1%. The absorbance was measured at 450 nm.

For that, we diluted a recombinant EGFR protein in different solutions used during immunoprecipitation (Figure 4.8). It can be observed that PBS-T presents higher absorbance values ([EGFR]>10 ng/mL), which will overestimate the calculated concentration, while PBS with NaCl or urea present lower absorbances, which suggests that this high salt concentration can interfere with EGFR binding. Dilution in RapiGest™ SF showed similar results to dilution in PBS.

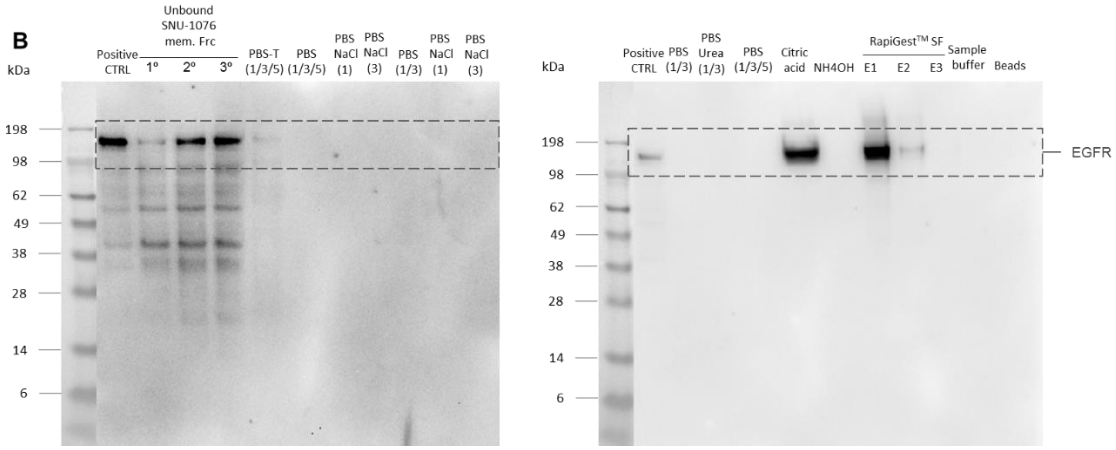
The effect of citric acid solution on quantification by ELISA was also assessed and results were similar to the effect of high salt (data not shown).

Although quantification by ELISA should be very sensitive, these results showing buffer interference led us to assess the presence of EGFR in the different fractions from IP procedure by WB (Figure 4.9).

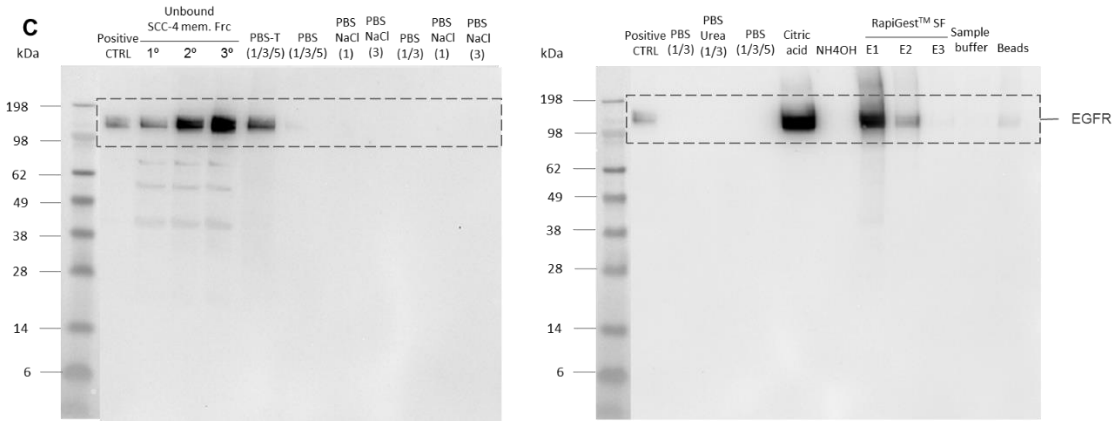
FaDu



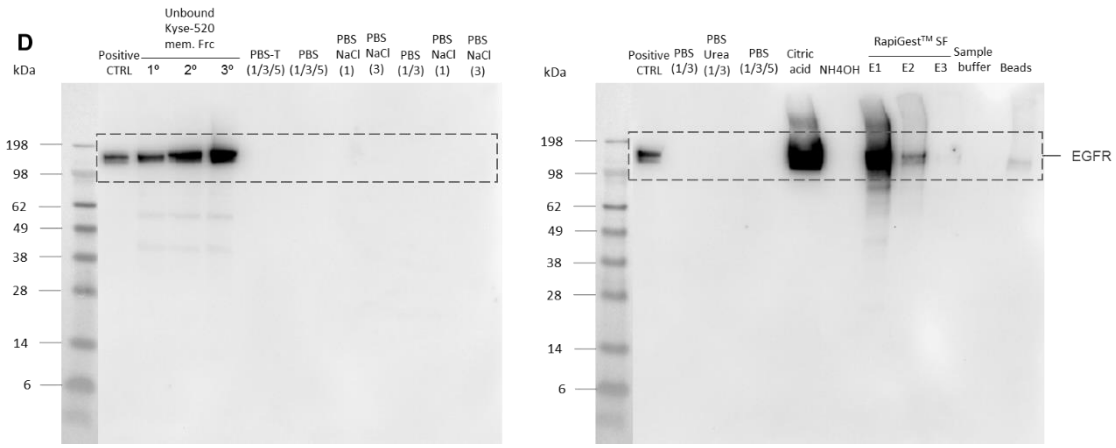
SNU-1076



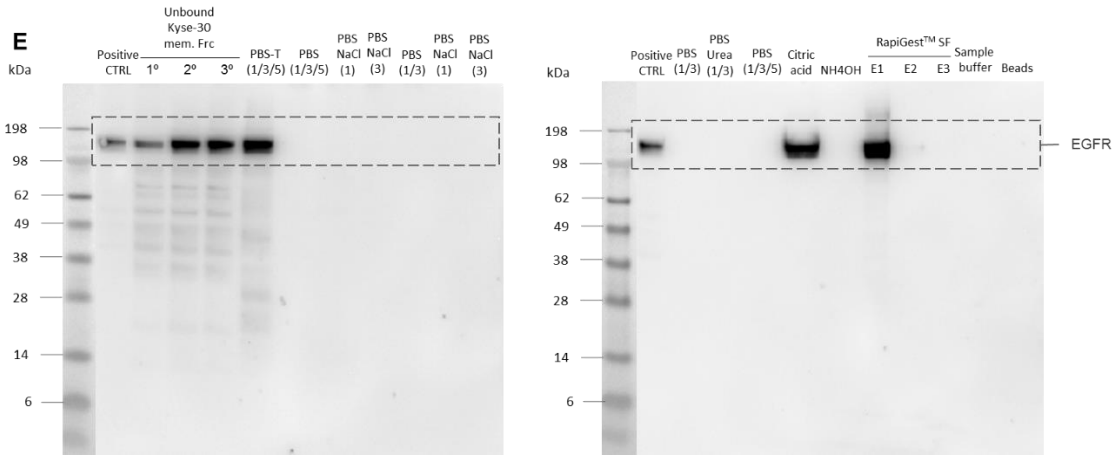
SCC-4



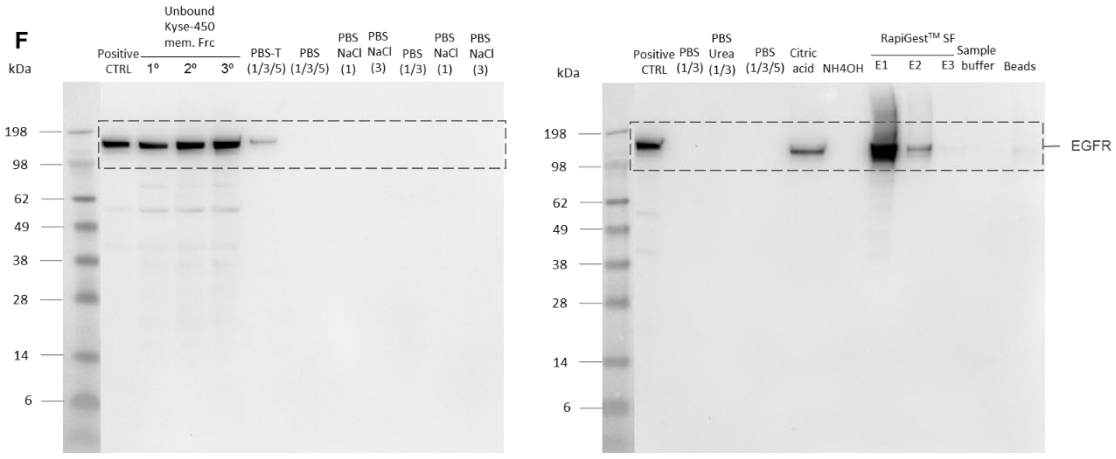
Kyse-520



Kyse-30



Kyse-450



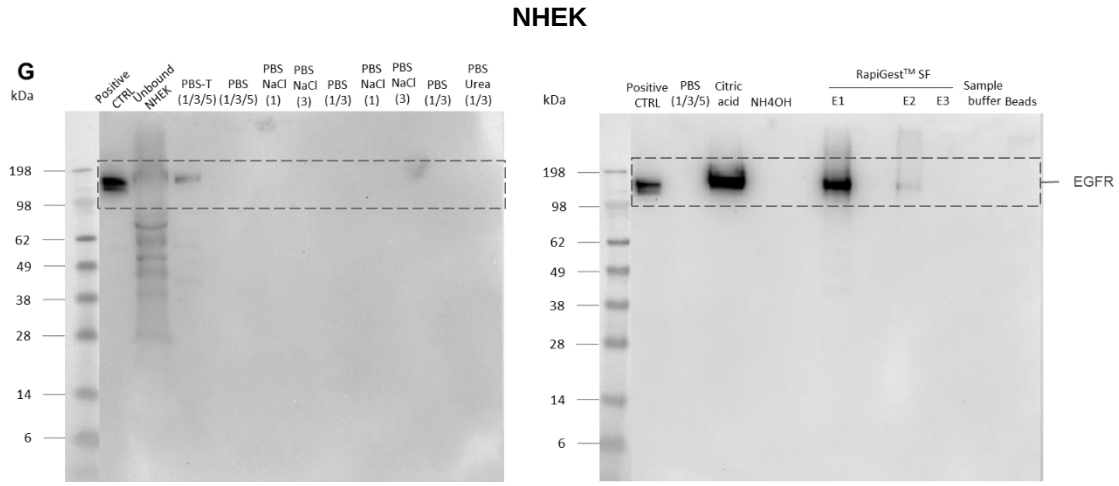


Figure 4.9 – Western blot analysis of EGFR present in immunoprecipitated fractions obtained from cell lines representative of HNC (FaDu, SNU-1076, SCC-4), EC (Kyse-520/-30/-450) and adult skin cells (NHEK). The samples were loaded in Novex™ NuPAGE™ 4 – 12% Bis-Tris gel and transferred to a PVDF membrane. The primary antibody used was a rabbit anti-EGFR (clone 15F8) monoclonal antibody, diluted 1:2000, and the secondary antibody was horseradish peroxidase-conjugated goat anti-rabbit antibody, diluted 1:80000. The numbers 1/3/5 are the identification of the washes loaded in each well. (**A**) Results of FaDu cell line. Positive control was 5 µg of FaDu membrane fraction. (**B**) Results of SNU-1076 cell line. Positive control was 2 µg of SNU-1076 membrane fraction. (**C**) Results of SCC-4 cell line. Positive control was 1 µg of SCC-4 membrane fraction. (**D**) Results of Kyse-520 cell line. Positive control was 0.4 µg of Kyse-520 membrane fraction. (**E**) Results of Kyse-30 cell line. Positive control was 1 µg of Kyse-30 membrane fraction. (**F**) Results of Kyse-450 cell line. Positive control was 3 µg of Kyse-450 membrane fraction. (**G**) Results of NHEK cell line. Positive control was 0.4 µg of Kyse-520 membrane fraction. Development time of 1 min in results of cancer cell lines and 2 min in results of NHEK.

In these cases, the IP was performed with all parameters optimized, with triplicates, performing three sequential incubations with membrane fraction of each cultured cell. IP using protein extracts from cancer cell lines was performed with the same amount of EGFR (4500 ng) while the IP with NHEK was performed with lesser amount of EGFR (1832.93 ng), due to the protein extracts limitation in this case.

WB analysis (Figure 4.9) of the sequential unbound fractions of all cancer cell lines was performed to evaluate if the EGFR was being bound to the coated beads. In these samples, a band correspondent to EGFR is observed in consecutive unbound fractions with an increasing intensity, due to consecutive incubations with EGFR-containing protein extracts. These results demonstrate that, although EGFR signal is detected in the unbound fractions, the antibody on the beads is not completely saturated in the first incubation, since an increasing of the signal is observed in the next two incubations. The incubation time might be short to allow the interaction of cetuximab with all EGFR present in solution. However, extending the incubation time could lead to unspecific interactions with other proteins. In NHEK results, a very tenuous band of EGFR on unbound fraction is observed, which means that the antibody-coated beads might not be completely saturated with EGFR, since the amount of EGFR used was smaller.

The initial washes with PBS-T revealed that part of EGFR is coming off in all cells under study, except for Kyse-520. This is probably EGFR that did not bind to cetuximab and was only adsorbed to the beads. EGFR was not detected in the remaining washes.

Figure 4.9 also shows that EGFR was essentially eluted with citric acid and RapiGest™ SF solutions. However, EGFR was not completely eluted with citric acid and RapiGestTM SF at

high temperature eluted the majority of the remaining EGFR. Nevertheless, we observed that in SCC-4, Kyse-520 and Kyse-450 cell lines some EGFR was still bound to the beads.

It is clear that there are differences between cell lines regarding the amount of EGFR obtained in the citric acid and RapiGest™ SF elutions, which could be related with structural differences in EGFR that could interfere with cetuximab binding.

To assess the protein content that is removed in the several solutions applied during the IP procedure, the fractions from an IP with 600 µg of NHEK protein extract were applied in SDS-PAGE and stained with silver (Figure 4.10).

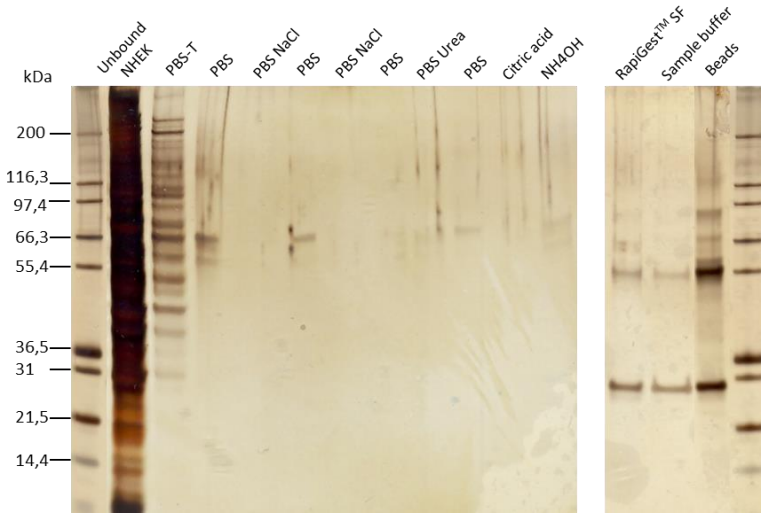


Figure 4.10 – Representative silver staining of all fractions from an IP with 600 µg of NHEK protein extract. The samples were loaded in Novex™ NuPAGE™ 4 – 12% Bis-Tris gel and stained with SilverQuest™ Silver Staining Kit.

The results obtained showed that a higher number of proteins were present in the unbound fraction. The first and second wash solutions used (PBS-T and PBS) also contained proteins, and most of the contaminant proteins were removed with these solutions. In the next wash solutions some bands can be observed, which justifies the necessity of several washes to be performed.

In the elution fractions with RapiGest™ SF, sample buffer and on the beads is possible to observe the presence of cetuximab, with the two bands corresponding to the heavy and light chain of the antibody. The use of high temperature in RapiGest™ SF and sample buffer steps disrupt the covalent link between the antibody and beads. The same was not observed on samples of citric acid elution.

Since bands corresponding to contaminant proteins are visible on elution solutions (where EGFR is observed, by WB analysis), this developed methodology does not allow the obtention of an EGFR purified fraction but an enriched fraction.

To verify the robustness of our methodology, we repeated our developed IP protocol using a different batch of a SCC-4 membrane extract (Figure 4.11).

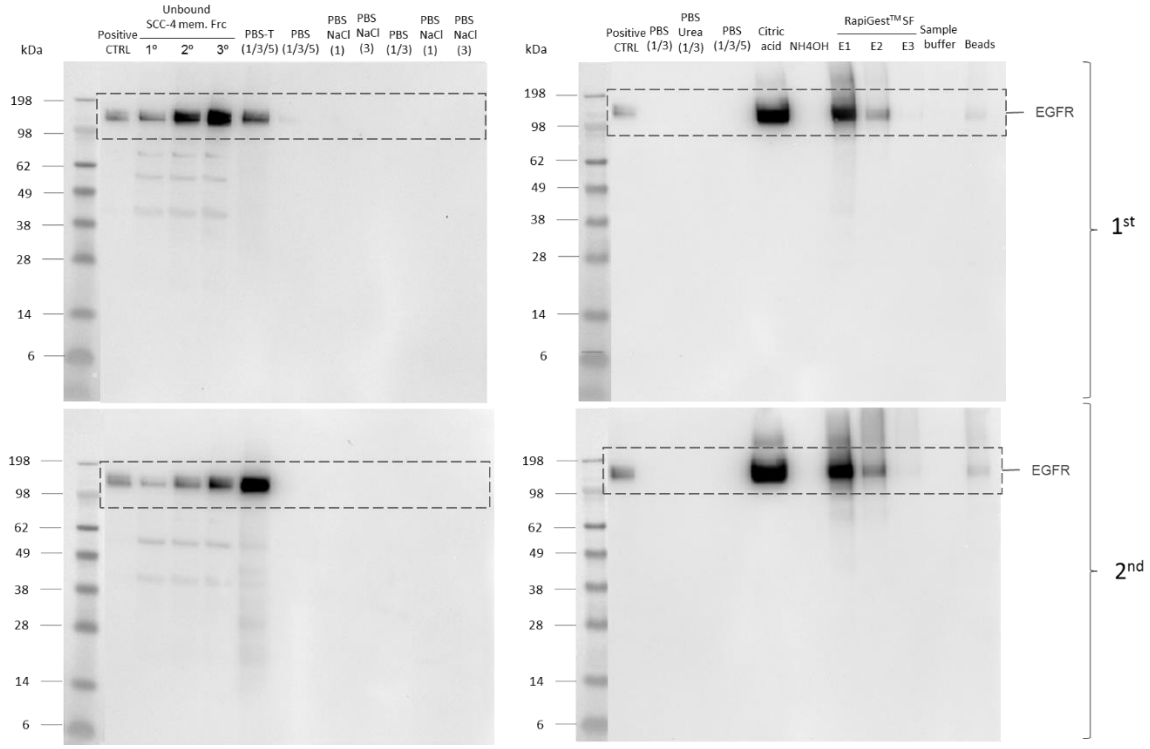


Figure 4.11 – Western blot analysis of EGFR present in immunoprecipitated fractions from independent IPs with different batches of SCC-4 membrane protein extracts. The samples were loaded in Novex™ NuPAGE™ 4 – 12% Bis-Tris gel and transferred to a PVDF membrane. The primary antibody used was a rabbit anti-EGFR (clone 15F8) monoclonal antibody, diluted 1:2000, and the secondary antibody was horseradish peroxidase-conjugated goat anti-rabbit antibody, diluted 1:80000. The numbers 1/3/5 are the identification of the washes loaded in each well. Positive control was 1 µg of SCC-4 membrane fraction. Development time of 1 min

Results obtained with the second batch of SCC-4 membrane fractions reveled a very similar profile to the one obtained with the first batch. This demonstrates the reproducibility and robustness of our methodology and suggests that differences in the IP profiles between cell lines are likely due to their own biologic characteristics and not related to the method itself.

Following EGFR immunoaffinity enrichment step, the elution fractions were analyzed by MS methods in order to assess sample purity and identify EGFR glycan structures.

4.4. Mass spectrometry analysis

To evaluate the protein content in both membrane fractions and elution samples, we performed LC-MS protein analysis. Proteins in the samples were reduced with DTT and alkylated with IAA (details in Material and Methods section) to maintain protein linearity and facilitate the proteases/glycosidase access to the respective cleaving sites. Samples were then digested with trypsin and Lys-C after the removal of the glycans with PNGase F (as described in Material and Methods Section 3.5). The digestion with trypsin cleaves the peptide chain at the carboxyl side of lysine and arginine residues, while Lys-C completes the digestion, cleaving peptides on the C-terminal side of lysine residues.

Results obtained with the analysis of Kyse-520 membrane fraction identified EGFR as the third protein in the rank list, however amongst a total of approximately 450 identified proteins. Table 4.3 shows the top 10 identified proteins in this analysis.

Table 4.3 – Identification of the top 10 proteins. The number of identified peptides, the protein score and sequence coverage are indicated.

N	Protein	Peptides (95%)	ProtScore	% COV (95)
1	Glucose-regulated protein	31	60.05	49.1
2	Endoplasmin	31	59.47	45.0
3	Epidermal growth factor receptor	27	55.04	29.9
4	Heat shock protein	35	52.18	57.1
5	Actin	61	51.27	82.4
6	Vimentin	25	49.87	62.2
7	Transferrin receptor protein	23	42.87	35.8
8	Protein disulfide-isomerase A3	25	40.66	51.5
9	Calnexin	32	40.37	48.0
10	ATP synthase subunit beta	21	40.2	54.1

EGFR was identified with 27 peptides (≥ 95% confidence) and a protein score of 55.04 (scores > 2 represent proteins identified with > 99% confidence). The results also show 29.9% of sequence coverage (percentage of amino acids in a protein sequence identified with 95% confidence).

Results obtained for the IP elution samples (citric acid or RapiGest^{TM} SF) from the different cells under study, identified EGFR again in the top ranked proteins (higher protein scores). However, this time within a smaller number of total proteins, ranging between 7 and 89 identified proteins. These results indicate that EGFR is present after the IP process, but still co-eluting with some proteins (which is also observed in Figure 4.10), though much less than the ones identified in the direct analysis of the membrane fraction. It was also observed that most of those identified proteins are not described as being N-glycosylated (e.g. actin [170]) or seem to be present in small amounts, so their presence should not compromise the identification of EGFR specific glycans.

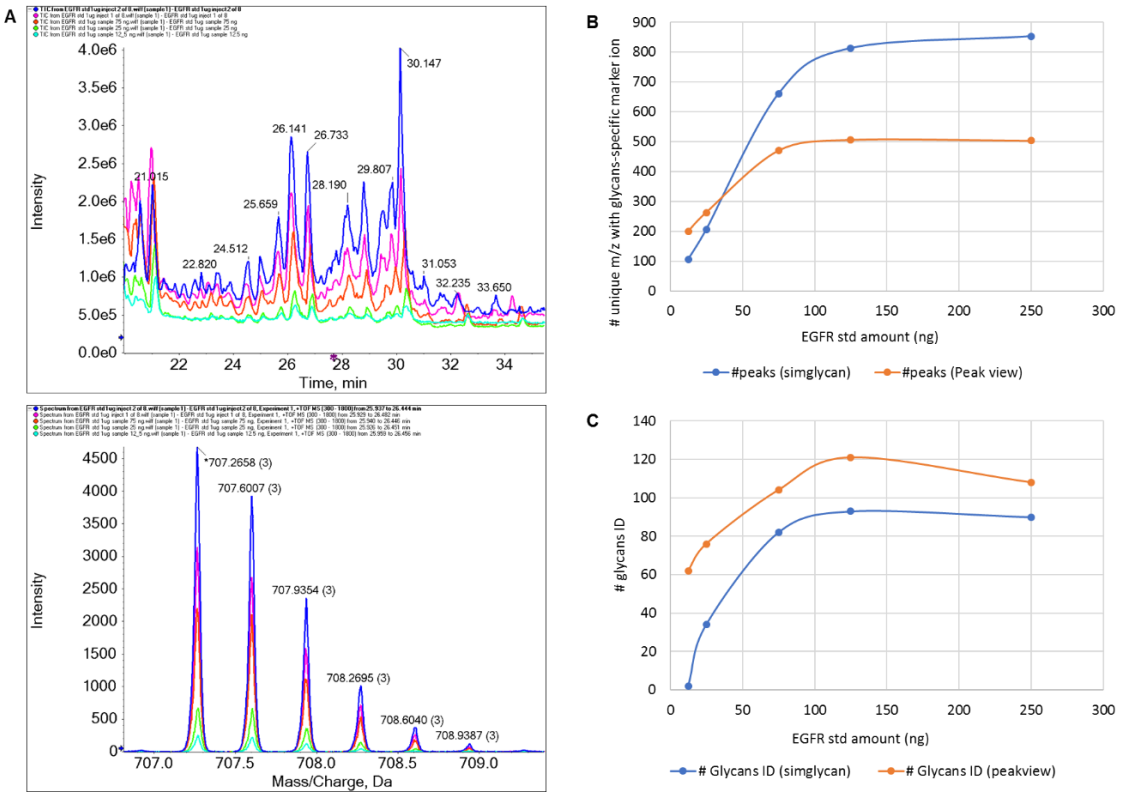
Before analyzing IP-derived samples by MS, we sought to establish the proper sample preparation and the minimum amount of EGFR necessary to allow glycan identification.

For that, we tested different quantities of a recombinant EGFR protein produced in mammalian cells (HEK293). This protein is constituted by amino acids 25 to 645 of full-length EGFR which correspond to the ectodomain. Its peptide sequence has ~70 kDa but it can reach 110 – 115 kDa due to glycosylation. This recombinant protein was chosen instead of the recombinant EGFR from Hi5 origin (used in the ELISA assays) since proteins produced in mammalian cells have far more complex glycosylation than insect cells. Indeed, insect cells are able to initiate N-glycosylation process in the ER as mammalian cells, but the processing of

oligosaccharides in the Golgi apparatus is different from what occurs in mammalian cells due to the absence or low activity of glycosyltransferases enzymes, resulting in proteins with limited glycosylation [154-156]. Given this, as the main objective was to optimize the process to identify the glycosylation pattern of human EGFR present in our samples, the closer our test/control protein is from real samples the more reliable and transferable our protocol should be. However, we cannot predict how well recombinant EGFR glycosylation levels will correlate with the levels present in the different endogenous human EGFR to be analyzed.

Furthermore, during the optimization of this step both *in solution* and *in gel* digestions were tested. However, the best results were obtained with samples processed in gel. Although some proteins may be lost in this procedure, it allows to concentrate and clean the samples from components in solution, such as salts and detergents that could interfere with the digestions and also with the MS. Here, proteins are immobilized on gel and stained (with MS-compatible staining protocols). Then, the band/piece of interest is cut and further processed to release glycans and peptides form the gel, as described in the Material and Methods Section 3.5. It was also tested a purification/enrichment of the glycans using graphite microcolumns, but some glycans were lost in this procedure, so the microcolumns were not used in this developed methodology.

Results obtained from MS analysis of different amounts (250, 125, 75, 25, and 12.5 ng) of recombinant EGFR released glycans can be observed in Figure 4.12.



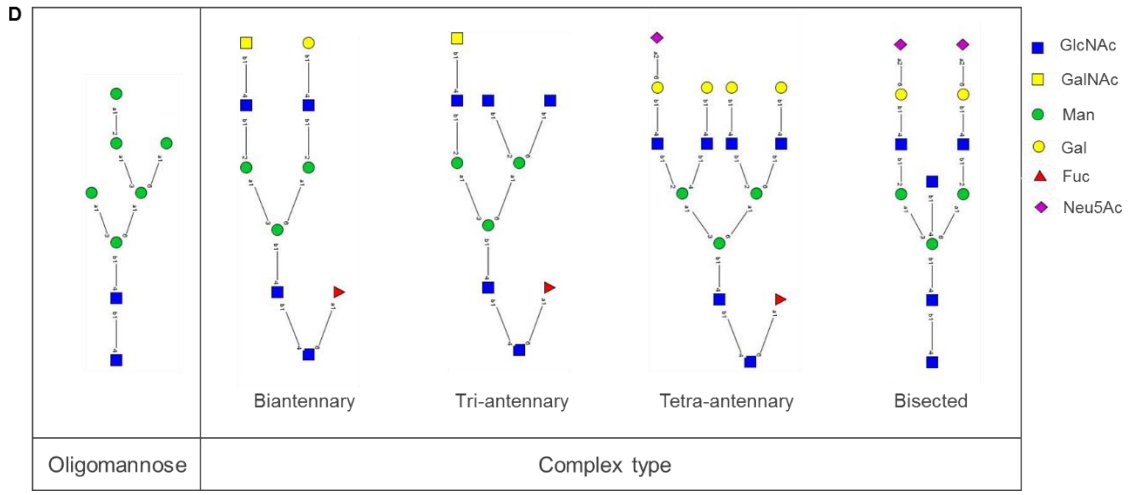


Figure 4.12 – (**A**) Total ion chromatogram (top panel) and average mass spectra of TIC peak at 26.14 min retention time (bottom panel). Different amounts of EGFR were injected on-column: 250 ng (dark blue), 125 ng (pink), 75 ng (red), 25 ng (green), 12.5 ng (light blue). (**B**) Number of unique m/z ions with the glycan specific marker ion 204.08 m/z and (**C**) number of glycan identifications, determined by SimGlycan software (blue) and PeekView with GlycoMod (orange). (**D**) Examples of glycan structures identified, using SimGlycan program.

MS analysis of recombinant EGFR showed that intensities of the peaks obtained in the total ion chromatograms and mass spectra correlate with the amount of EGFR used. The higher the amount of EGFR injected the higher the intensity of the peaks observed (Figure 4.12 (A))

The number of peaks and of glycan identification was tested with different softwares (Figure 4.12 (B) and (C)). SimGlycan identifies glycan structures by matching MS/MS data with theoretical fragmentation. On the other hand, PeakView+GlycoMod predicts possible glycan structures based on experimentally determined masses (MS). Furthermore, these results show that a plateau was reached, both for the number of peaks and glycan structures identified, when glycans corresponding to 100 ng of EGFR were injected. Higher amounts/volume of glycans injected resulted in approximately the same number of identifications, allowing us to define that 100 ng of recombinant EGFR is the minimum amount of this protein that is needed to perform a robust glycan profile characterization.

The N-glycan structures identified on recombinant EGFR were mostly of the complex type, including biantennary, tri-antennary, tetra-antennary and bisected structures, and oligomannose structures were also identified (Figure 4.12 D), which is consistent with structures found in mammalian cells.

The described procedure was also performed with IgA, starting with 250 ng of pure protein (data not shown). Here, 22 glycan structures were identified using GlycoMod, while SimGlycan software identified 26 glycan structures. These included oligomannose, hybrid and complex type of N-glycans. These results indicate that our method is robust and suitable for different proteins (varying in size, complexity, number of glycosylation sites, biological origin, …).

Citric acid and RapiGest™ SF elutions were analyzed according to the optimized procedure described above. Briefly, these samples passed through a step of PNGase F digestion *in gel*, after reduction and alkylation of the proteins. The spectra obtained (data not shown) do not have any peak corresponding to glycan structures (confirmed by the absence of glycans marker ions, namely 204.08 $(HexNAc)^+$ and 366.12 $(Hex\text{-}HexNAc)^+$) not allowing any glycan identification in the samples.

These results can be explained by the small amount of EGFR obtained in the eluted fractions (that is observed when the samples are loaded and stained in gel or in WB analysis), leading to a small amount of glycans present in the sample that are below the detection limit of the current nanoLC-MS method.

To evaluate this hypothesis, a sample of membrane fraction of Kyse-520 cells with 1.5 µg of total protein was directly analyzed using the protocol described but without any step of EGFR enrichment. The results obtained allowed to identify 38 glycan structures by GlycoMod and 35 glycan structures by SimGlycan, mainly from complex and hybrid type, but oligomannose structures were also identified. Of course, with this approach one cannot assure the specificity of the glycan profile obtained, as several other glycoproteins besides EGFR are present in this fraction (~450 proteins were previously identified as referred above).

Again, these results are showing that samples resulting from the IP processing may not have the amount of EGFR and consequently of glycans needed (above the detection limit) to guarantee the identification of glycan structures by this methodology.

Due to these results, further optimizations of our protocols are required in order to increase EGFR quantities in the final fractions to enable definition of its glycosylation profile in the different cell lines and tissues under study, which is the main objective of this work.

4.5. Application of the developed method with human tissue

The ultimate goal of this project was the characterization of EGFR glycosylation in cancer and normal human tissue. To achieve that, we established a collaboration with IPOLFG that enabled us to access patients' derived samples and clinical data of relevance.

A collection of tissue samples collected at IPOLFG (project approved by the Ethical committee of IPOLFG) is presented in Table 4.4, including the organ of origin, weight, percentage of tumor and immunohistochemical (IHC) scores. Examples of IHC analysis performed at IPOLFG are represented in Figure 4.13.

Table 4.4 – Characteristics of the collected samples from tumor tissues and their corresponding adjacent non-tumorigenic tissues, including the organ of origin, weight, percentage of tumor and immunohistochemical (IHC) scores.

Case	Organ	Tumor				Normal		
		Weight (mg)	% tumor	IHC*	Cell line	Weight (mg)	IHC	Cell line
1	Tongue	175.6	50	3+, >75	SCC-4	147.5	3+, >75	NHEK
2		194.2	80	3+, >75		109.6	3+, >75	
3		281.8	70	3+, >75		151.1	3+, >75	
4		660.1	60	3+, >75		242.3	3+, >75	
5	Buccal floor/tongue	179.8	30	3+, >75		54.8	3+, <25	
6		31.2	<5	3+, >75		21.7	3+, >75	
7		117.6	80	3+, >75		233.9	3+, >75	
8		213.3	60	3+, >75		125.8	3+, >75	
9	Pyriform sinus	47.2	60	3+, 50-75	FaDu	24.7	3+, >75	
10		176.5	70	3+, >75		14.9	-	
11	Oropharynx	106.9	70	3+, >75		48.8	3+, >75	
12	Supraglottic larynx	95.9	70	3+, >75	SNU-1076	38.9	3+, >75	
13	Maxillary sinus	187	80	-		-	-	
14	Esophagus	57.00	60	2+, 25-50	Kyse-520, Kyse-30, Kyse450	100.10	2+, <25	
15		153.40	50	3+, >75		71.20	3+, >75	
16		129.40	60	3+, <25		112.10	3+, >75	

***immunohistochemical scores**: Intensity of EGFR signal in a given percentage of cells.
Intensity: Neg, no staining; 1+, weakly positive; 2+, moderately positive; 3+, strongly positive.

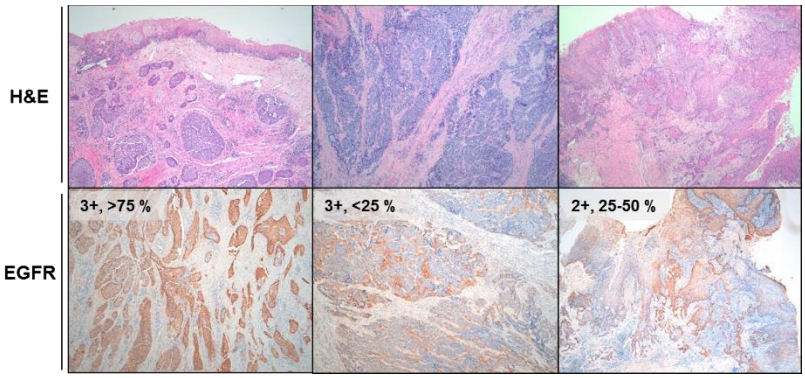


Figure 4.13 – Representative histochemical stainings of tumor tissue sections from HNC patients. The upper panel shows stainings by hematoxylin and eosin (H&E) while the lower panel depicts EGFR immunohistochemical stainings showing different intensities of EGFR expression and different percentages of EGFR positive cells.

Collected cases included samples of tumor tissues and their corresponding adjacent non-tumorigenic tissues, that we assumed as normal tissue. HNC cases comprised tissues mainly from the oral cavity (tongue and buccal floor), but also from the pharynx (pyriform sinus and oropharynx), the supraglottic larynx and maxillary sinus. These different origins give a representability of this type of cancer, and of the majority of the organs that can be affected by this disease. Only three cases were obtained from EC patients.

The weight of the specimens collected was highly variable, ranging between 31.2 mg and 660.1 mg for tumor tissues and 14.9 mg and 242.3 mg for normal samples (Table 4.4). During collection, cancer tissues were also analyzed to assess the percentage of tumor cells present. These values varied between <5% and 80%, but most of the cases presented values superior to 50%.

Immunohistochemical evaluation of EGFR was performed in both tumor and normal tissue. In the majority of tissue sections, more than 75% of the cells were stained with high intensity (3+), in both tumor and normal tissue (Table 4.4). Figure 4.13 shows representative images of EGFR stainings with different IHC scores. These results indicate that a large number of samples contain high expression levels of EGFR.

Visually, the samples were very heterogeneous with different morphologic structures, some presenting more blood, others more fat tissue and different consistencies (data not shown). These differences are expected with tumor samples, coming from different patients and from different organs.

After this qualitative analysis, tissue protein extraction was performed using Mem-PER™ Plus Membrane Protein Extraction Kit optimized protocol, schematized on Figure 4.14. The amount of protein obtained in tumor and normal samples is presented in Table 4.5.

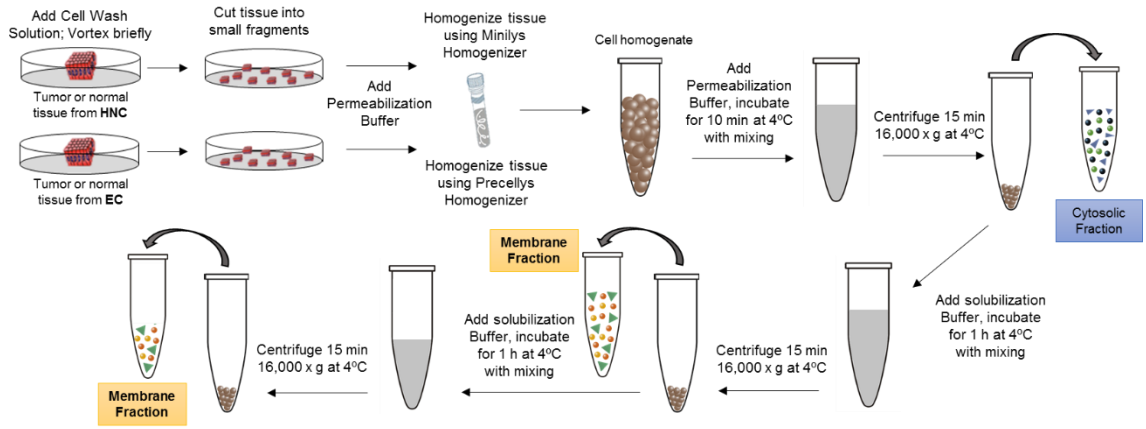


Figure 4.14 - Schematic representation of the procedure implemented for protein extraction from tissue and membrane fraction preparation.

Table 4.5 – Total protein amount obtained in protein extractions from tumor tissues and their corresponding adjacent non-tumorigenic tissues. The amount of total protein was determined by MicroBCA and normalized by weight. The organ of origin and weight of the samples is also showed.

Case	Organ	Weight (mg)		Total protein (µg)				Total protein/Weight (µg/mg)			
				Tumor		Normal		Tumor		Normal	
		Tumor	Normal	Cytosolic	Membrane	Cytosolic	Membrane	Cytosolic	Membrane	Cytosolic	Membrane
1	Tongue	175.6	147.5	7117.87	2189.82	5565.58	895.12	40.53	12.47	37.73	6.07
2		194.2	109.6	4561.32	913.65	3696.37	850.15	23.49	4.70	33.73	7.76
3		281.8	151.1	9858.70	1288.83	5739.27	664.94	34.98	4.57	37.98	4.40
4		660.1	242.3	16964.50	4040.96	7847.70	1356.19	25.70	6.12	32.39	5.60
5	Buccal floor/tongue	179.8	54.8	6157.20	960.26	1398.65	139.46	34.24	5.34	25.52	2.54
6		31.2	21.7	761.96	245.26	348.45	42.89	24.42	7.86	16.06	1.98
7		117.6	233.9	3830.81	637.14	7524.43	979.88	32.57	5.42	32.17	4.19
8		213.3	125.8	6339.00	1404.26	3279.60	516.44	29.72	6.58	26.07	4.11
9	Pyriform sinus	47.2	24.7	1411.34	375.98	861.08	83.50	29.90	7.97	34.86	3.38
10		176.5	14.9	6091.74	1674.16	483.24	51.08	34.51	9.49	32.43	3.43
11	Oropharynx	106.9	48.8	2604.79	519.33	1383.35	109.76	24.37	4.86	28.35	2.25
12	Supraglottic larynx	95.9	38.9	3159.78	872.63	1498.96	209.85	32.95	9.10	38.53	5.39
13	Maxillary sinus	187	-	8008.76	3024.90	-	-	42.83	16.18	-	-
14	Esophagus	57.00	100.10	5133.02	1220.45	6784.69	1993.76	90.05	21.41	67.78	19.92
15		153.40	71.20	8740.92	3422.96	3065.00	924.38	56.98	22.31	43.05	12.98
16		129.40	112.10	6595.25	2666.32	5682.78	1940.57	50.97	20.61	50.69	17.31

Different homogenization procedures were performed for protein extraction. Cases from HNC patients were prepared with Minilys homogenizer, while EC samples were homogenized by Precellys instrument. Table 4.5 shows that the ratio of total protein *per* weight is higher when the Precellys instrument was used, when compared with Minilys. These differences can be related with the fact that Precellys is an advanced equipment as it reaches higher motion speed than Minilys, which can promote a higher extraction efficiency. Moreover, Precellys instrument was only tested with EC samples and we cannot exclude that this type of tissue might be better homogenized than the other samples or simply have a higher protein content.

During the homogenization process, differences between normal and tumor tissue from the same cases were observed. Normal tissue reveled to be more difficult to dissociate. The apparent existence of higher lipid content in normal tissue was also observed, as we observed the formation of a lipid layer on the top of the supernatants during the centrifugation steps.

In all cases, protein extraction was performed, with the cytosolic fractions having higher amounts of protein than membrane fractions, as expected, for both tumor and normal tissues. Variations between samples are observed, since the tissue pieces were highly heterogeneous presenting different types of cells, such as blood or adipose cells, which can influence the amount of extracted protein [157].

The amount of EGFR was quantified in all cytosolic and membrane fractions (Table 4.6).

Table 4.6 - Total EGFR amount obtained in protein extractions from tumor tissues and their corresponding adjacent non-tumorigenic tissues. The amount of EGFR was determined by ELISA and normalized by weight. The organ of origin and weight of the samples is also presented.

Case	Organ	Weight (mg)		Total EGFR (ng)				Total EGFR (ng)		Total EGFR/Weight (ng/mg)		Total EGFR/total protein (ng/µg)		Ratio Normalized by weight
				Tumor		Normal								Tumor/ Normal
		Tumor	Normal	Cyt.	Memb.	Cyt.	Memb.	Tumor	Normal	Tumor	Normal	Tumor	Normal	
1	Tongue	175.6	147.5	35.18	65.21	6.28	14.37	100.38	20.65	0.57	0.14	0.011	0.003	4.08
2		194.2	109.6	45.89	34.08	10.71	11.69	79.96	22.41	0.41	0.20	0.015	0.005	2.01
3		281.8	151.1	107.04	27.94	3.15	14.80	134.98	17.95	0.48	0.12	0.012	0.003	4.03
4		660.1	242.3	93.51	833.40	10.60	17.04	926.91	27.64	1.40	0.11	0.044	0.003	12.31
5	Buccal floor/tongue	179.8	54.8	172.39	22.64	0.43	2.83	195.03	3.26	1.08	0.06	0.027	0.002	18.23
6		31.2	21.7	5.82	4.41	0.44	1.73	10.23	2.17	0.33	0.10	0.010	0.006	3.28
7		117.6	233.9	49.32	48.85	6.00	10.84	98.17	16.84	0.83	0.07	0.022	0.002	11.59
8		213.3	125.8	43.56	263.44	7.11	2.80	307.00	9.91	1.44	0.08	0.040	0.003	18.26
9	Pyriform sinus	47.2	24.7	115.29	44.73	2.94	2.34	160.02	5.28	3.39	0.21	0.090	0.006	15.86
10		176.5	14.9	36.67	122.28	2.36	2.13	158.96	4.48	0.90	0.30	0.020	0.008	2.99
11	Oropharynx	106.9	48.8	87.38	35.71	0.77	2.48	123.09	3.25	1.15	0.07	0.039	0.002	17.29
12	Supraglottic larynx	95.9	38.9	101.94	31.35	12.54	15.98	133.29	28.51	1.39	0.73	0.033	0.017	1.90
13	Maxillary sinus	187	-	91.21	165.56	-	-	256.77	-	1.37	-	0.023	-	-
14	Esophagus	57.00	100.10	223.26	100.91	184.12	48.24	324.17	232.36	5.69	2.32	0.051	0.026	2.45
15		153.40	71.20	146.82	103.77	2.48	3.05	250.60	5.53	1.63	0.08	0.021	0.001	21.03
16		129.40	112.10	293.07	63.30	192.02	101.27	356.37	293.29	2.75	2.62	0.038	0.038	1.05

Results show that the total amount of EGFR obtained in the different cases was highly variable, despite most of the samples presented high EGFR expression when analyzed by IHC.

In HNC, quantities ranged between 10.23 ng and 926.91 ng in tumor tissues and 2.17 ng and 28.51 ng in normal tissues (Table 4.6). In EC, variation in EGFR amounts were not so remarkable, except for the normal tissue of case 15 where we only obtained 5.53 ng of EGFR (Table 4.6). Heterogeneity in the obtained quantities of EGFR may arise from the different characteristics of the samples but also from the efficiency of the homogenization procedure.

Regarding the amount of EGFR present per µg total protein extracted, some differences can also be observed in the tumor samples, with values varying from 0.010 to 0.090 ng EGFR /µg total protein, which indicates different levels of expression of EGFR in the tumor samples analyzed. However, normal tissue samples collected from HNC patients did not show such a high variability between them. Higher concentrations of EGFR/µg total protein were obtained in EC samples when compared with HNC samples. Once again, these differences can be due to an improvement in the extraction protocol with the use of Precellys equipment in the homogenization step, but also due to the biological differences between samples/tissue characteristics.

Overall, we observed that the amount of extracted EGFR normalized by protein content was higher in tumor than in normal tissues. These results were expected, since EGFR is overexpressed in both cancer types [31].

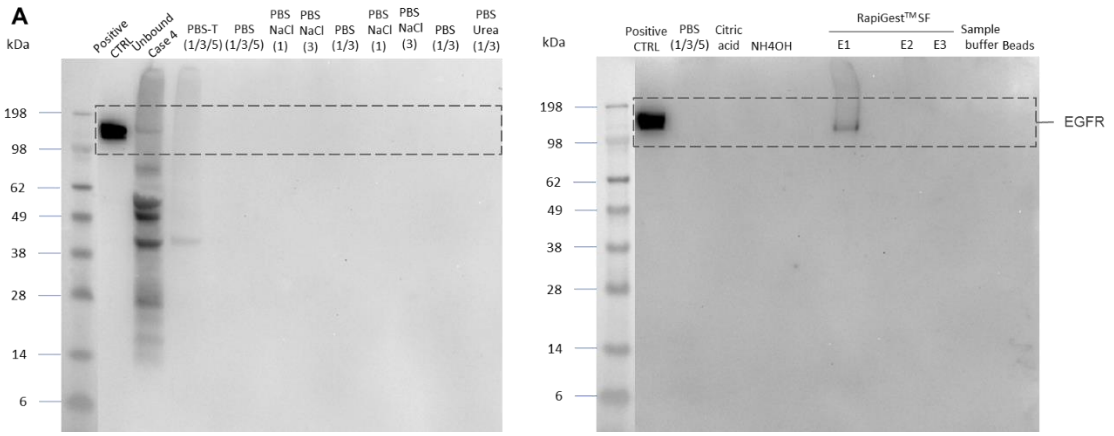
In some cases, cytosolic fractions presented higher quantities of EGFR than the corresponding membrane fractions (11/16 in tumors and 4/17 in normal tissue), especially in tumors. This may be explained by the fact that these tissue samples, that already present higher EGFR expression, were submitted to longer incubation times with permeabilization buffer during homogenization, which occurred because of difficulties in tissue disruption.

Considering that for a reliable MS glycan analysis the minimum amount of EGFR needed was 100 ng, we verified that the quantity of EGFR *per* sample before IP procedures was already below that limit, thus insufficient to proceed with individual IP in some of the cases. To overcome this limitation, for some of the cases we pooled samples according to their organ of origin, such as case 9, 10 and 11 of pharynx origin. We also tested individual cases, including case 4 from tongue, that had enough EGFR amount for an individual analysis, and case 12 from supraglottic larynx and 13 from maxillary sinus, that were the only cases representing the origin of those organs.

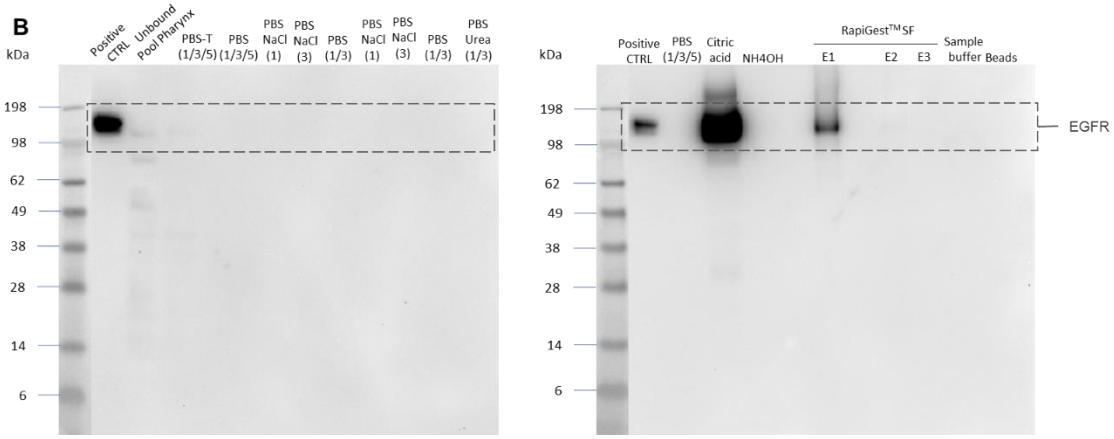
In contrast to what we have done with cell lines, tissue samples were incubated with both cytosolic and membrane fractions to increase the initial amount of EGFR.

The assessment of EGFR presence in the different fractions from the IP procedures was performed by WB analysis (Figure 4.14).

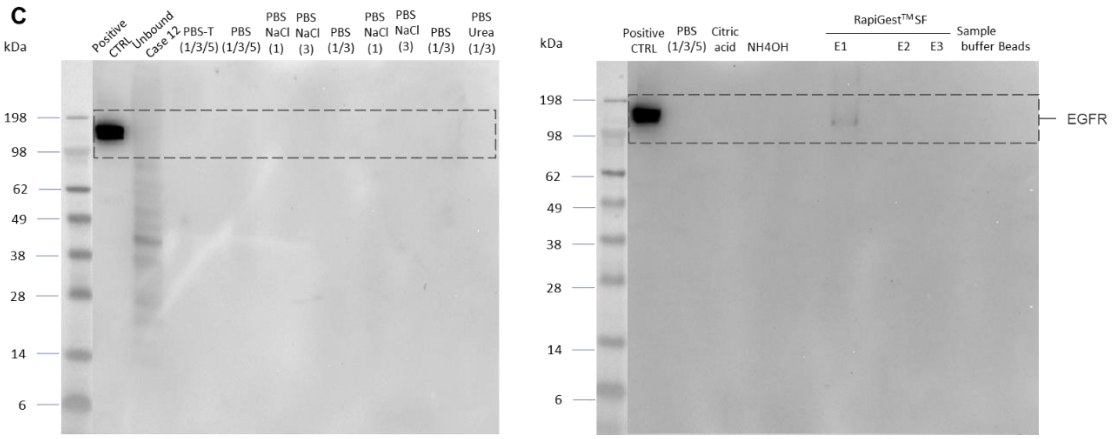
Case 4



Pool pharynx (Cases 9, 10 and 11)



Case 12



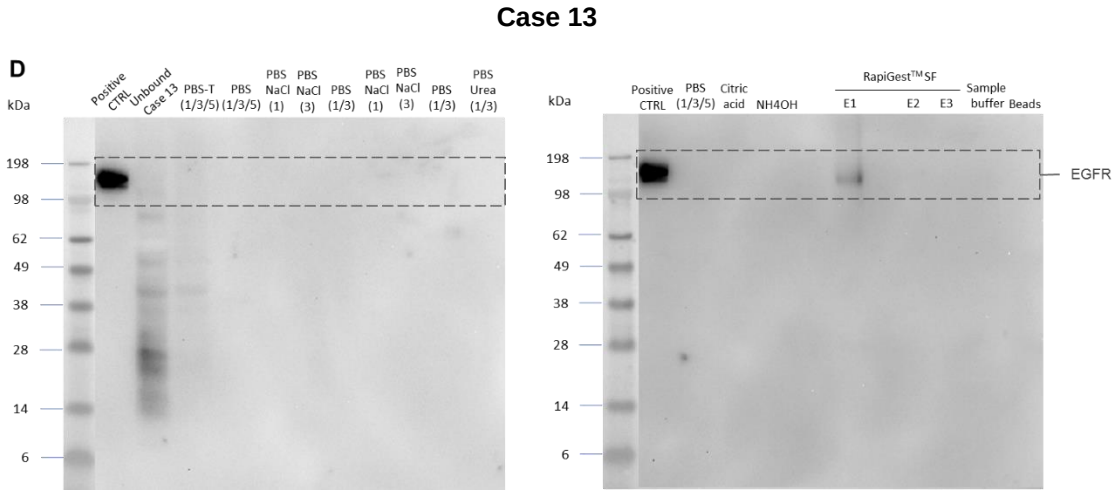


Figure 4.15 – Western blot analysis of EGFR present in immunoprecipitated fractions obtained from IPs with tumor samples. Samples were loaded in Novex™ NuPAGE™ 4 – 12% Bis-Tris gel and transferred to a PVDF membrane. The primary antibody used was a rabbit anti-EGFR (clone 15F8) monoclonal antibody, diluted 1:2000, and the secondary antibody was horseradish peroxidase-conjugated goat anti-rabbit antibody, diluted 1:80000. The numbers 1/3/5 are the identification of the washes loaded in each well. (**A**) Results of tumor case 4 (tongue). Development time of 7 min. (**B**) Results of pharynx pooled tumor cases (10, 11 from pyriform sinus and 12 from oropharynx). Development time of 2 min. (**C**) Results of tumor case 12 (supraglottic larynx). Development time of 6 min. (**D**) Results of tumor case 13 (maxillary sinus). Positive. Development time of 4 min. In all cases, 0.4 µg of Kyse-520 membrane fraction was used as positive control.

The IP results from case 4 showed a very faint band of EGFR present in the unbound fraction, suggesting saturation of cetuximab-coated beads with EGFR or that during incubations not all EGFR molecules in solution were able to interact and bind to cetuximab. This was not observed for the other cases, possibly because those had initial lower amounts of EGFR, meaning that all molecules were likely bound to the antibody.

Moreover, EGFR was not detected in the numerous washes in all of the cases tested (Figure 4.15).

The IP results from pharynx pooled cases (Figure 4.15) showed that EGFR was mainly eluted with citric acid solution and partly with RapiGest™ SF, while in cases 4, 12 and 13 EGFR was only detected at the RapiGest™ SF elution and at much lower quantities. These results are probably due to the initial lower amounts of incubated EGFR (cases 12 and 13), conversely to case 4. In this case, is likely due to few interaction events between cetuximab and EGFR. We cannot exclude that possible structural differences of EGFR between cases could affect its affinity to cetuximab.

Our developed IP methodology is yet to be applied to the remaining tumor samples from tongue/buccal floor, esophagus and all normal tissues samples.

5. Main achievements and conclusions

The main aim of this thesis was to develop a novel methodology capable of generating enriched EGFR fractions, which ultimate goal is the definition of EGFR glycosylation patterns to identify targets to be used in the development of novel cancer-specific therapies. During this work, we performed several steps that led to the following achievements and conclusions:

a) All cell lines and primary cells (NHEK) showed EGFR expression. Western blot analysis revealed that SCC-4, Kyse-520, Kyse-30 and Kyse-450 cancer cell lines also have an additional EGFR isoform, likely corresponding to EGFRvIII;

b) Cancer cell lines and primary cells were used to optimize a protocol to obtain EGFR-containing membrane fractions, with Kyse-30 and SNU-1076 showing the highest and lowest EGFR yields, respectively;

c) A lectin ELISA assay was developed to glycoprofile EGFR. The most striking result demonstrated that, generally, cancer cell lines have higher reactivities to α2,6-linked sialic acids than normal cells;

d) We developed a robust EGFR immunoaffinity procedure able to capture and elute this receptor from different cell types or cell batches;

e) We established that the minimum amount of recombinant EGFR (produced in HEK293) for a confident MS glycan analysis should be approximately 100 ng;

f) MS analysis of glycans from recombinant EGFR protein produced in HEK293 cells identified different N-glycan structures, mostly from complex type, including biantennary, tri-antennary, tetra-antennary and bisected structures, and oligomannose type;

g) The low amounts of EGFR present in the elution fractions from the IPs with cell lines and primary cells are most probably hindering glycan identification by MS;

h) Tumor tissues and their corresponding adjacent non-tumorigenic tissues from HNC and EC patients showed generally high EGFR expression levels by IHC analysis, although the quantities of extracted EGFR were highly variable between cases;

i) The developed EGFR immunoaffinity procedure was also applied to tissue samples from HNC patients. However, the quantities of eluted EGFR were much lower, especially for cases 4, 12 and 13, compared to what we obtained with cancer cell lines and primary cells, probably due to substantial differences in the initial quantities of EGFR present in protein extracts.

Overall, we developed a successful methodology for the production of EGFR enriched fractions that can be applied to diverse cell lines and tissue samples. However, we were still not able to define a glycosylation pattern for EGFR.

Future work will include further optimizations of the process, specially addressing LC-MS detection sensitivity, aiming at analyzing tissue derived enriched EGFR samples which ultimately will enable the definition of EGFR glycosylation patterns.

6. References

[1] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman, “Global cancer statistics,” 2011.

[2] G. Danaei, S. Vander Hoorn, A. D. Lopez, C. J. L. Murray, and M. Ezzati, “Causes of cancer in the world: comparative risk assessment of nine behavioural and environmental risk factors,” vol. 366, 2005.

[3] N. Denaro, E. G. Russi, and M. C. Merlano, “Pros and Cons of the New Edition of TNM Classification of Head and Neck Squamous Cell Carcinoma,” *Oncology*, pp. 1–9, 2018.

[4] C. R. Leemans, B. J. M. Braakhuis, and R. H. Brakenhoff, “The molecular biology of head and neck cancer,” *Nat. Publ. Gr.*, vol. 11, no. 1, pp. 9–22, 2010.

[5] C. K. Howlader N, Noone AM, Krapcho M, Miller D, Bishop K, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, “SEER Cancer Statistics Review,” *National Cancer Institute*, 2017. [Online]. Available: https://seer.cancer.gov/csr/1975_2014/.

[6] K. J.Harrington, “Biology of cancer,” *Medicine (Baltimore).*, vol. 39, no. 12, pp. 689–692, 2011.

[7] J. Bertram, “The molecular biology of cancer,” *Mol. Aspects Med.*, vol. 21, no. 6, pp. 167–223, 2001.

[8] D. Hanahan and R. A. Weinberg, “Hallmarks of Cancer : The Next Generation,” *Cell*, vol. 144, no. 5, pp. 646–674, 2011.

[9] P. Smilek *et al.*, “Epidermal growth factor receptor (EGFR) expression and mutations in the EGFR signaling pathway in correlation with anti-EGFR therapy in head and neck squamous cell carcinomas,” vol. 59, no. 5, pp. 508–515, 2012.

[10] “Types of cancer treatment,” *National cancer institute*, 2017. [Online]. Available: https://www.cancer.gov/about-cancer/treatment/types.

[11] D. E. Gerber, “Targeted Therapies: A New Generation of Cancer Treatments,” *Am. Fam. Physician*, vol. 77, no. 3, 2008.

[12] T. M. Rezende, M. de S. Preire, and O. L. Franco, “Head and Neck Cancer,” pp. 4914–4925, 2010.

[13] E. Crozier and B. D. Sumer, “Head and Neck Cancer,” *Med. Clin. NA*, vol. 94, no. 5, pp. 1031–1046, 2010.

[14] D. M. Cognetti, R. S. Weber, and S. Y. Lai, “Head and Neck Cancer: An Evolving Treatment Paradigm,” *Cancer Suppl.*, vol. 113, no. 7, 2008.

[15] OpenStax College, “Head and neck anatomy: sagittal illustration.” [Online]. Available: https://radiopaedia.org/cases/head-and-neck-anatomy-sagittal-illustration.

[16] R. F. Osborne and J. J. Brown, “Carcinoma of the oral pharynx: An analysis of subsite treatment heterogeneity,” *Surg. Oncol. Clin. N. Am.*, vol. 13, no. 1, pp. 71–80, 2004.

[17] D. M. Cohan, S. Popat, S. E. Kaplan, N. Rigual, T. Loree, and W. L. Hicks, “Oropharyngeal cancer: current understanding and management,” *Curr. Opin. Otolaryngol. Head Neck Surg.*, vol. 17, no. 2, pp. 88–94, 2009.

[18] Serviço Nacional de Saúde, “IPO Porto - Cancro da cabeça e pescoço,” 2017. [Online]. Available: https://www.sns.gov.pt/noticias/2017/07/25/ipo-porto-cancro-da-cabeca-e-pescoco/.

[19] “Portugal: Doenças Oncológicas em Números (2014),” 2014.

[20] P. Ho and Y. Ko, “The incidence of oropharyngeal cancer in Taiwan : an endemic betel quid chewing area,” *Oral Pathol Med*, vol. 31, no. 9, pp. 213–219, 2002.

[21] E. T. Chang and H. Adami, “The Enigmatic Epidemiology of Nasopharyngeal Carcinoma,” vol. 15, no. 10, pp. 1765–1778, 2006.

[22] N. Guha *et al.*, “Oral Health and Risk of Squamous Cell Carcinoma of the Head and Neck and Esophagus : Results of Two Multicentric Case-Control Studies,” *Am. J. Epidemiol.*, vol. 166, no. 10, pp. 1159–1173, 2007.

[23] Y.-C. Chien *et al.*, “SEROLOGIC MARKERS OF EPSTEIN – BARR VIRUS INFECTION AND NASOPHARYNGEAL CARCINOMA IN TAIWANESE MEN,” *N. Engl. J. Med.*, vol. 345, no. 26, pp. 1877–1882, 2001.

[24] L. Hong, S. Li, Y. Han, J. Du, H. Zhang, and J. Li, “Angiogenesis-related molecular targets in esophageal cancer,” pp. 637–644, 2011.

[25] R. Alteri *et al.*, “Cancer Facts & Figures 2018,” 2018.

[26] H. El-Zimaity *et al.*, “Risk factors for esophageal cancer: emphasis on infectious agents.,” *Ann. N. Y. Acad. Sci.*, 2018.

[27] J. LAGERGREN, R. BERGSTRÖM, A. LINDGREN, and O. NYRÉN, “SYMPTOMATIC GASTROESOPHAGEAL REFLUX AS A RISK FACTOR FOR ESOPHAGEAL ADENOCARCINOMA,” vol. 340, no. 11, pp. 825–831, 1999.

[28] M. J. D. Arnal, Á. F. Arenas, and Á. L. Arbeloa, “Esophageal cancer: Risk factors, screening and endoscopic treatment in Western and Eastern countries,” *World J. Gastroenterol.*, vol. 21, no. 26, pp. 7933–7943, 2015.

[29] H. Kato and M. Nakajima, “Treatments for esophageal cancer: A review,” *Gen. Thorac. Cardiovasc. Surg.*, vol. 61, no. 6, pp. 330–335, 2013.

[30] L. Hong, Y. Han, and L. Brain, “Epidermal growth factor receptor : an important target in esophageal cancer,” pp. 1179–1185, 2013.

[31] C. Yewale, D. Baradia, I. Vhora, S. Patil, and A. Misra, “Epidermal growth factor receptor targeting in cancer : A review of trends and strategies,” *Biomaterials*, vol. 34, no. 34, pp. 8690–8707, 2013.

[32] A. Wells, “EGF receptor,” *Int. J. Biochem. Cell Biol.*, vol. 31, pp. 637–643, 1999.

[33] Archieve Ensembl, “Gene: EGFR,” 2017. [Online]. Available: http://may2017.archive.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000146648;r=7:55019021-55211628.

[34] R. N. Jorissen, F. Walker, N. Pouliot, T. P. J. Garrett, C. W. Ward, and A. W. Burgess, “Epidermal growth factor receptor : mechanisms of activation and signalling,” *Exp. Cell Res.*, vol. 284, pp. 31–53, 2003.

[35] C. Abou-Fayçal, A. S. Hatat, S. Gazzeri, and B. Eymin, “Splice variants of the RTK family: Their role in tumour progression and response to targeted therapy,” *Int. J. Mol. Sci.*, vol. 18, no. 2, 2017.

[36] S. Bishayee, “Role of Conformational Alteration in the Epidermal Growth Factor Receptor (EGFR) Function,” *Biochem. Pharmacol.*, vol. 60, no. 0, pp. 1217–1223, 2000.

[37] R. Roskoski, “The ErbB/HER family of protein-tyrosine kinases and cancer,” *Pharmacol. Res.*, vol. 79, pp. 34–74, 2014.

[38] C. D. Carpenter *et al.*, “Structural analysis of the transmembrane domain of the epidermal growth factor receptor,” *J Biol Chem*, vol. 266, no. 9, pp. 5750–5755, 1991.

[39] P. Wee and Z. Wang, “Epidermal growth factor receptor cell proliferation signaling pathways,” *Cancers (Basel).*, vol. 9, no. 5, pp. 1–45, 2017.

[40] H. Sun *et al.*, “The juxtamembrane, cytosolic region of the epidermal growth factor receptor is involved in association with α-subunit of G(s),” *J. Biol. Chem.*, vol. 272, no. 9, pp. 5413–5420, 1997.

[41] K. M. Ferguson, M. B. Berger, J. M. Mendrola, H. S. Cho, D. J. Leahy, and M. A. Lemmon, “EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization,” *Mol. Cell*, vol. 11, no. 2, pp. 507–517, 2003.

[42] B. Singh, G. Carpenter, and R. J. Coffey, “EGF receptor ligands: recent advances,” *F1000Research*, vol. 5, no. 0, p. 2270, 2016.

[43] T. Sasaki, K. Hiroki, and Y. Yamashita, “The Role of Epidermal Growth Factor Receptor in Cancer Metastasis and Microenvironment,” *Biomed Res. Int.*, vol. 2013, 2013.

[44] K. Roepstorff *et al.*, “Differential effects of EGFR ligands on endocytic sorting of the receptor,” *Traffic*, vol. 10, no. 8, pp. 1115–1127, 2009.

[45] Y. Yarden and M. X. Sliwkowski, “Untangling the ErbB signalling network,” *Nat. Rev. Mol. Cell Biol.*, vol. 2, no. 2, pp. 127–137, 2001.

[46] X. Zhang, J. Gureasko, K. Shen, P. A. Cole, and J. Kuriyan, “An Allosteric Mechanism for Activation of the Kinase Domain of Epidermal Growth Factor Receptor,” *Cell*, vol. 125, no. 6, pp. 1137–1149, 2006.

[47] R. S. Herbst, “Review of epidermal growth factor receptor biology,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 59, no. 2 SUPPL., pp. 21–26, 2004.

[48] T. Pawson, “Specificity in Signal Transduction: From Phosphotyrosine-SH2 Domain Interactions to Complex Cellular Systems,” *Cell*, vol. 116, no. 2, pp. 191–203, 2004.

[49] I. Alroy and Y. Yarden, “The ErbB signaling network in embryogenesis and oncogenesis: Signal diversification through combinatorial ligand-receptor interactions,” *FEBS Lett.*, vol. 410, no. 1, pp. 83–86, 1997.

[50] K. Oda, Y. Matsuoka, A. Funahashi, and H. Kitano, “A comprehensive pathway map of epidermal growth factor receptor signaling,” *Mol. Syst. Biol.*, vol. 1, no. 1, pp. E1–E17, 2005.

[51] M. Scaltriti and J. Baselga, “The Epidermal Growth Factor Receptor Pathway : A Model for Targeted Therapy The Epidermal Growth Factor Receptor Pathway : A Model for T argeted Therapy,” *Clin Cancer Res*, vol. 12, no. 18, pp. 5268–5272, 2006.

[52] H. Fernandes, S. Cohen, and S. Bishayee, “Glycosylation-induced Conformational Modification Positively Regulates Receptor-Receptor Association,” vol. 276, no. 7, pp. 5375–5383, 2001.

[53] A. Franovic, L. Gunaratnam, K. Smith, I. Robert, D. Patten, and S. Lee, “Translational up-regulation of the EGFR by tumor hypoxia provides a nonmutational explanation for its overexpression in human cancer,” *Proc. Natl. Acad. Sci.*, vol. 104, no. 32, pp. 13092–13097, 2007.

[54] M. K. Sethi *et al.*, “In-depth N -glycome pro fi ling of paired colorectal cancer and non-tumorigenic tissues reveals cancer- , stage- and EGFR-speci fi c protein N-glycosylation,” *Glycobiology*, vol. 25, no. 10, pp. 1064–1078, 2015.

[55] J. C. Sok *et al.*, “Mutant epidermal growth factor receptor (EGFRvIII) contributes to head and neck cancer growth and resistance to EGFR targeting,” *Clin. Cancer Res.*, vol. 12, no. 17, pp. 5064–5073, 2006.

[56] X. Duan, S. Zhou, M. Zhang, P. Wang, J. Zhang, and J. Wang, “Clinical significance of EGFR and EGFRvIII expression in human esophageal carcinoma,” vol. 27, no. 3, pp. 6–11, 2011.

[57] H. K. Gan, A. N. Cvrljevic, and T. G. Johns, “The epidermal growth factor receptor variant III (EGFRvIII): where wild things are altered,” vol. 280, pp. 5350–5370, 2013.

[58] H. Ogiso *et al.*, “Crystal Structure of the Complex of Human Epidermal Growth Factor and Receptor Extracellular Domains,” vol. 110, pp. 775–787, 2002.

[59] M. Perez-torres, M. Guix, A. Gonzalez, and C. L. Arteaga, “Epidermal Growth Factor Receptor (EGFR) Antibody Down-regulates Mutant Receptors and Inhibits Tumors Expressing EGFR Mutations,” vol. 281, no. 52, pp. 40183–40192, 2006.

[60] X. Yang, X. Jia, J. R. F. Corvalan, P. Wang, and C. G. Davis, “Development of ABX-EGF, a fully human anti-EGF receptor monoclonal antibody, for cancer therapy,” vol. 38, pp. 17–23, 2001.

[61] J. Ocvirk and S. Cencelj, “Management of cutaneous side-effects of cetuximab therapy in patients with metastatic colorectal cancer,” pp. 453–459, 2010.

[62] S. M. Thomas and J. R. Grandis, “Pharmacokinetic and pharmacodynamic properties of EGFR inhibitors under clinical investigation,” *elsevier Heal.*, vol. 30, pp. 255–268, 2004.

[63] A. Arora and E. M. Scholar, “Role of Tyrosine Kinase Inhibitors in Cancer Therapy,” vol. 315, no. 3, pp. 971–979, 2005.

[64] C. Yewale, D. Baradia, I. Vhora, and A. Misra, “Proteins : emerging carrier for delivery of cancer therapeutics,” pp. 1–20, 2013.

[65] V. S. Goldmacher, W. A. Blättler, J. M. Lambert, and R. V. J. Chari, “Immunotoxins and Antibody-Drug Conjugates for Cancer Treatment,” *Biomed. Asp. Drug Target.*, pp. 291–292, 2002.

[66] T. Gazori, I. Haririan, S. Fouladdel, A. Namazi, and A. Nomani, “Inhibition of EGFR expression with chitosan / alginate nanoparticles encapsulating antisense oligonucleotides in T47D cell line using RT-PCR and immunocytochemistry,” *Carbohydr. Polym.*, vol. 80, no. 4, pp. 1042–1047, 2010.

[67] N. Calonghi *et al.*, “A new EGFR inhibitor induces apoptosis in colon cancer cells,” vol. 354, pp. 409–413, 2007.

[68] P.-À. Nygren, “Alternative binding proteins: Affibody binding proteins developed from a small three-helix bundle scaffold,” *FEBS J.*, vol. 275, pp. 2668–2676, 2008.

[69] P. H. Seeberger, “Monosaccharide Diversity,” in *Essentials of Glycobiology*, 3rd editio., et al. Varki A, Cummings RD, Esko JD, Ed. Cold Spring Harbor Laboratory Press, 2017.

[70] A. Varki and S. Kornfeld, “Historical Background and Overview,” in *Essentials of Glycobiology*, 3rd editio., A. Varki, R. Cummings, and J. Esko, Eds. Cold Spring Harbor Laboratory Press, 2017.

[71] S. S. Pinho and C. A. Reis, “Glycosylation in cancer : mechanisms and clinical implications,” *Nat. Publ. Gr.*, vol. 16, no. August, 2015.

[72] A. Varki and P. Gagneux, “Biological Functions of Glycans,” in *Essentials of Glycobiology*, 3rd editio., A. Varki, R. Cummings, and J. Esko, Eds. Cold Spring Harbor Laboratory Press, 2017.

[73] A. Varki, “Biological roles of glycans,” *Glycobiology*, vol. 27, no. 1, pp. 3–49, 2017.

[74] D. Russell, N. J. Oldham, and B. G. Davis, “Site-selective chemical protein glycosylation protects from autolysis and proteolytic degradation,” *Carbohydr. Res.*, vol. 344, no. 12, pp. 1508–1514, 2009.

[75] C.-C. Wang *et al.*, “Glycans on influenza hemagglutinin affect receptor binding and immune response,” *Proc. Natl. Acad. Sci.*, vol. 106, no. 43, pp. 18137–18142, 2009.

[76] M. E. V. Johansson, J. M. H. Larsson, and G. C. Hansson, “The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions,” *Proc. Natl. Acad. Sci.*, vol. 108, no. Supplement_1, pp. 4659–4665, 2011.

[77] P. Gagneux and A. Varki, “Evolutionary considerations in relating oligosaccharide diversity to biological function,” vol. 9, no. 8, pp. 747–755, 1999.

[78] B. E. Collins and J. C. Paulson, “Cell surface biology mediated by low affinity multivalent protein-glycan interactions,” *Curr. Opin. Chem. Biol.*, vol. 8, no. 6, pp. 617–625, 2004.

[79] R. Sackstein *et al.*, “Ex vivo glycan engineering of CD44 programs human multipotent mesenchymal stromal cell trafficking to bone,” *Nat. Med.*, vol. 14, no. 2, pp. 181–187, 2008.

[80] L. Xia *et al.*, “Defective angiogenesis and fatal embryonic hemorrhage in mice lacking core 1-derived O-glycans,” *J. Cell Biol.*, vol. 164, no. 3, pp. 451–459, 2004.

[81] Y.-Y. Zhao *et al.*, “Functional roles of N-glycans in cell signaling and cell adhesion in cancer,” *Cancer Sci.*, vol. 99, no. 7, pp. 1304–1310, 2008.

[82] A. Almeida and D. Kolarich, “The promise of protein glycosylation for personalised medicine,” *Biochim. Biophys. Acta*, vol. 1860, no. 8, pp. 1583–1595, 2016.

[83] A. Imberty and A. Varrot, “Microbial recognition of human cell surface glycoconjugates,” *Curr. Opin. Struct. Biol.*, vol. 18, no. 5, pp. 567–576, 2008.

[84] L. V. Hooper and J. I. Gordon, “Glycans as legislators of host-microbial interactions: Spanning the spectrum from symbiosis to pathogenicity,” *Glycobiology*, vol. 11, no. 2, pp. 1–10, 2001.

[85] L. G. Baum, O. B. Garner, K. Schaefer, and B. Lee, “Microbe-host interactions are positively and negatively regulated by galectin-glycan interactions,” *Front. Immunol.*, vol. 5, no. JUN, pp. 1–8, 2014.

[86] K. Colley, A. Varki, and T. Kinoshita, “Cellular Organization of Glycosylation,” in *Essentials of Glycobiology*, 3rd editio., et al. Varki A, Cummings RD, Esko JD, Ed. Cold Spring Harbor Laboratory Press, 2017.

[87] B. Henrissat, A. Surolia, and P. Stanley, “A Genomic View of Glycobiology,” in *Essentials of Glycobiology*, 3rd editio., A. Varki, R. Cummings, and J. Esko, Eds. Cold Spring Harbor Laboratory Press, 2017.

[88] H. Freeze, G. Hart, and R. Schnaar, “Glycosylation Precursors,” in *Essentials of Glycobiology*, 3rd editio., A. Varki, R. Cummings, and J. Esko, Eds. Cold Spring Harbor Laboratory Press, 2017.

[89] J. Rini and J. Esko, “Glycosyltransferases and Glycan-Processing Enzymes,” in *Essentials of Glycobiology*, 3rd editio., A. Varki, R. Cummings, and J. Esko, Eds. Cold Spring Harbor Laboratory Press, 2017.

[90] P. Van Den Steen, P. M. Rudd, R. A. Dwek, and G. Opdenakker, “Concepts and principles of O-linked glycosylation,” *Crit. Rev. Biochem. Mol. Biol.*, vol. 33, no. 3, pp. 151–208, 1998.

[91] E. P. Bennett, U. Mandel, H. Clausen, T. A. Gerken, T. A. Fritz, and L. A. Tabak, “Control of mucin-type O-glycosylation: A classification of the polypeptide GalNAc-transferase gene family,” *Glycobiology*, vol. 22, no. 6, pp. 736–756, 2012.

[92] I. Brockhausen and P. Stanley, “O-GalNAc Glycans,” in *Essentials of Glycobiology*, 3rd editio., A. Varki, R. Cummings, and J. Esko, Eds. Cold Spring Harbor Laboratory Press, 2017.

[93] D. J. Moloney, A. I. Lin, and R. S. Haltiwanger, “The O -Linked Fucose Glycosylation Pathway,” *J. Biol. Chem.*, vol. 272, no. 30, pp. 19046–19050, 1997.

[94] R. Haltiwanger, L. Wells, and H. Freeze, “Other Classes of Eukaryotic Glycans,” in *Essentials of Glycobiology*, 3rd editio., A. Varki, R. Cummings, and J. Esko, Eds. Cold Spring Harbor Laboratory Press, 2017.

[95] N. Zachara, Y. Akimoto, and G. Hart, “The O-GlcNAc Modification,” in *Essentials of*

Glycobiology, 3rd editio., A. Varki, R. Cummings, and J. Esko, Eds. Cold Spring Harbor Laboratory Press, 2017.

[96] A. Matsuura *et al.*, “O-linked N-acetylglucosamine is present on the extracellular domain of notch receptors,” *J. Biol. Chem.*, vol. 283, no. 51, pp. 35486–35495, 2008.

[97] P. Stanley, N. Taniguchi, and M. Aebi, “N-Glycans,” in *Essentials of Glycobiology*, 3rd editio., A. Varki, R. Cummings, and J. Esko, Eds. Cold Spring Harbor Laboratory Press, 2017.

[98] M. K. Sethi, W. S. Hancock, and S. Fanayan, “Identifying N - Glycan Biomarkers in Colorectal Cancer by Mass Spectrometry,” 2016.

[99] M. Aebi, “N-linked protein glycosylation in the ER,” *BBA - Mol. Cell Res.*, vol. 1833, no. 11, pp. 2430–2437, 2013.

[100] E. S. Trombetta, “The contribution of N-glycans and their processing in the endoplasmic reticulum to glycoprotein biosynthesis,” vol. 13, no. 9, pp. 77–91, 2003.

[101] P. Q. Control, E. Reticulum, C. Carcinoma, and A. Molecule, “Protein N -Glycosylation along the Secretory Pathway : Relationship to Organelle Topography and Function , Protein Quality Control , and Cell,” 2002.

[102] N. Ding *et al.*, “Human serum N-glycan profiles are age and sex dependent,” *Age Ageing*, vol. 40, no. 5, pp. 568–575, 2011.

[103] A. O. Akinkuolie, J. E. Buring, P. M. Ridker, and S. Mora, “A novel protein glycan biomarker and future cardiovascular disease events,” *J. Am. Heart Assoc.*, vol. 3, no. 5, pp. 1–12, 2014.

[104] K. Ley, “The role of selectins in inflammation and disease,” *Trends Mol. Med.*, vol. 9, no. 6, pp. 263–268, 2003.

[105] A. O. Akinkuolie, A. D. Pradhan, J. E. Buring, P. M. Ridker, and S. Mora, “Novel Protein Glycan Side-Chain Biomarker and Risk of Incident Type 2 Diabetes Mellitus,” *Arterioscler. Thromb. Vasc. Biol.*, vol. 35, no. 6, pp. 1544–1550, 2015.

[106] M. S. Freedman, J. Laks, N. Dotan, R. T. Altstock, A. Dukler, and C. J. M. Sindic, “Anti-α-glucose-based glycan IgM antibodies predict relapse activity in multiple sclerosis after the first neurological event,” *Mult. Scler.*, vol. 15, no. 4, pp. 422–430, 2009.

[107] M. L. A. de Leoz *et al.*, “High-Mannose Glycans are Elevated during Breast Cancer Progression,” *Mol. Cell. Proteomics*, vol. 10, no. 1, p. M110.002717, 2011.

[108] H. Freeze, T. Kinoshita, and A. Varki, “Glycans in Acquired Human Diseases,” in *Essentials of Glycobiology*, 3rd editio., A. Varki, R. Cummings, and J. Esko, Eds. NY: Cold Spring Harbor Laboratory Press, 2017.

[109] E. Rodríguez, S. T. T. Schetters, and Y. Van Kooyk, “The tumour glyco-code as a novel immune checkpoint for immunotherapy,” *Nat. Rev. Immunol.*, vol. 18, no. 3, pp. 204–211, 2018.

[110] F. Dall’Olio, N. Malagolini, M. Trinchera, and M. Chiricolo, “Mechanisms of cancer-associated glycosylation changes,” vol. 17, no. 2, pp. 670–699, 2012.

[111] A. Varki, R. Kannagi, and B. Toole, “Glycosylation Changes in Cancer,” in *Essentials of Glycobiology*, 3rd editio., A. Varki, R. Cummings, and J. Esko, Eds. NY: Cold Spring Harbor Laboratory Press, 2017.

[112] S. Cua *et al.*, “Targeting of embryonic annexin A2 expressed on ovarian and breast cancer by the novel monoclonal antibody 2448.,” *Oncotarget*, vol. 9, no. 17, pp. 13206–13221, 2018.

[113] P. M. Drake *et al.*, “Discovery Equation,” vol. 56, no. 2, pp. 223–236, 2010.

[114] O. J. McGinn *et al.*, “Targeting the 5T4 oncofetal glycoprotein with an antibody drug

conjugate (A1mcMMAF) improves survival in Patient-Derived xenograft models of acute lymphoblastic leukemia," *Haematologica*, vol. 102, no. 6, pp. 1075–1084, 2017.

[115] C. S. Leung, "Analysis of ROR1 Protein Expression in Mice with Reconstituted Human Immune System Components," *J. Immunol. Res.*, vol. 2018, pp. 1–10, 2018.

[116] X. Wang and Q. Wang, "Alpha-Fetoprotein and Hepatocellular Carcinoma Immunity," *Can. J. Gastroenterol. Hepatol.*, vol. 2018, 2018.

[117] Y. Zhen, R. M. Caprioli, and J. V. Staros, "Characterization of Glycosylation Sites of the Epidermal Growth Factor Receptor," *Biochemistry*, vol. 42, no. 18, pp. 5478–5492, 2003.

[118] M. Lin, P. C. H. Wu, and S. C. H. Juan, "C1GALT1 predicts poor prognosis and is a potential therapeutic target in head and neck cancer," *Oncogene*, 2018.

[119] S. R. Stateva and A. Villalobo, "O-GlcNAcylation of the human epidermal growth factor receptor," *Org. Biomol. Chem.*, vol. 13, no. 30, pp. 8196–8204, 2015.

[120] K. D. Smith, M. J. Davies, D. Bailey, D. V. Renouf, and E. F. Hounsell, "Analysis of the Glycosylation Patterns of the Extracellular Domain of the Epidermal Growth Factor Receptor Expressed in Chinese Hamster Ovary Fibroblasts," *Growth Factors*, vol. 13, no. 1–2, pp. 121–132, 1996.

[121] C. Sato, J.-H. Kim, Y. Abe, K. Saito, S. Yokoyama, and D. Kohda, "Characterization of the *N*-oligosaccharides attached to the atypical Asn-X-Cys sequence of recombinant human epidermal growth factor receptor," *J Biochem.*, vol. 127, no. 1, pp. 65–72, 2000.

[122] K. B. Whitson *et al.*, "Functional effects of glycosylation at Asn-579 of the epidermal growth factor receptor," *Biochemistry*, vol. 44, no. 45, pp. 14920–14931, 2005.

[123] I. Ferreira, M. Pucci, G. Venturi, N. Malagolini, and M. Chiricolo, "Glycosylation as a Main Regulator of Growth and Death Factor Receptors Signaling," 2018.

[124] K. Kaszuba, A. Or, R. Danne, T. Róg, K. Simons, and Ü. Coskun, "N -Glycosylation as determinant of epidermal growth factor receptor conformation in membranes," vol. 112, no. 14, pp. 2–7, 2015.

[125] H. Li, K. Al-Japairai, Y. Tao, and Z. Xiang, "RPN2 promotes colorectal cancer cell proliferation through modulating the glycosylation status of EGFR.," *Oncotarget*, vol. 8, no. 42, pp. 72633–72651, 2017.

[126] Y. Liu, H. Yen, C. Chen, C. Chen, P. Cheng, and Y. Juan, "Sialylation and fucosylation of epidermal growth factor receptor suppress its dimerization and activation in lung cancer cells," 2011.

[127] K. Matsumoto *et al.*, "N -Glycan fucosylation of epidermal growth factor receptor modulates receptor activity and sensitivity to epidermal growth factor receptor tyrosine kinase inhibitor," vol. 99, no. 8, pp. 1611–1617, 2008.

[128] X. Wang, J. Gu, H. Ihara, E. Miyoshi, K. Honke, and N. Taniguchi, "Core Fucosylation Regulates Epidermal Growth Factor Receptor-mediated Intracellular Signaling," *J. Biol. Chem.*, vol. 281, 2005.

[129] P. Guo, Q. Y. Wang, H. B. Guo, Z. H. Shen, and H. L. Chen, "N-acetylglucosaminyltransferase V modifies the signaling pathway of epidermal growth factor receptor," vol. 61, pp. 1795–1804, 2004.

[130] K.-C. Wen *et al.*, "α2,3-sialyltransferase type I regulates migration and peritoneal dissemination of ovarian cancer cells," *Oncotarget*, vol. 8, no. 17, 2017.

[131] C. M. Britain, A. T. Holdbrooks, J. C. Anderson, C. D. Willey, and S. L. Bellis, "Sialylation of EGFR by the ST6Gal-I sialyltransferase promotes EGFR activation and resistance to gefitinib-mediated cell death," *J. Ovarian Res.*, vol. 11, no. 1, pp. 1–11, 2018.

[132] W.-L. Lin, Y.-S. Lin, G.-Y. Shi, C.-F. Chang, and H.-L. Wu, "Lewisy promotes migration of oral cancer cells by glycosylation of epidermal growth factor receptor.," *PLoS One*,

vol. 10, no. 3, p. e0120162, 2015.

[133] K. Yamamoto *et al.*, “Potentiation of epidermal growth factor-mediated oncogenic transformation by sialidase NEU3 leading to src activation,” *PLoS One*, vol. 10, no. 3, pp. 1–17, 2015.

[134] M. Che *et al.*, “β1, 4-N-acetylgalactosaminyltransferase III modulates cancer stemness through EGFR signaling pathway in colon cancer cells.,” *Oncotarget*, vol. 5, no. 11, pp. 3673–84, 2014.

[135] M. C. Lin, M. J. Huang, C. H. Liu, T. L. Yang, and M. C. Huang, “GALNT2 enhances migration and invasion of oral squamous cell carcinoma by regulating EGFR glycosylation and activity,” *Oral Oncol.*, vol. 50, no. 5, pp. 478–484, 2014.

[136] S. Chugh, J. Meza, Y. M. Sheinin, M. P. Ponnusamy, and S. K. Batra, “Loss of N-acetylgalactosaminyltransferase 3 in poorly differentiated pancreatic cancer: Augmented aggressiveness and aberrant ErbB family glycosylation,” *Br. J. Cancer*, vol. 114, no. 12, pp. 1376–1386, 2016.

[137] P. Rudd, N. Karlsson, and K. Khoo, “Glycomics and Glycoproteomics,” in *Essentials of Glycobiology*, 3rd editio., A. Vark, R. Cummings, and J. Esko, Eds. Cold Spring Harbor Laboratory Press, 2017.

[138] S. M. Haslam, S. J. North, and A. Dell, “Mass spectrometric analysis of N- and O-glycosylation of tissues and cells,” *Curr. Opin. Struct. Biol.*, vol. 16, no. 5, pp. 584–591, 2006.

[139] R. Cummings, A. Darvill, and M. Etzler, “Glycan-Recognizing Probes as Tools,” in *Essentials of Glycobiology*, 3rd editio., A. Varki, R. Cummings, and J. Esko, Eds. Cold Spring Harbor Laboratory Press, 2017.

[140] D. F. Smith and R. D. Cummings, “Application of Microarrays for Deciphering the Structure and Function of the Human Glycome,” *Mol. Cell. Proteomics*, vol. 12, no. 4, pp. 902–912, 2013.

[141] B. Mulloy, A. Dell, and P. Stanley, “Structural Analysis of Glycans,” in *Essentials of Glycobiology*, 3rd editio., A. Varki, R. Cummings, and J. Esko, Eds. Cold Spring Harbor Laboratory Press, 2017.

[142] R. Aebersold and M. Mann, “Mass spectrometry-based proteomics,” vol. 422, no. March, 2003.

[143] B. Cañas, D. Lopez-Ferrer, A. Ramos-ferna, E. Camafeita, and E. Calvo, “Mass spectrometry technologies for proteomics,” vol. 4, no. 4, 2006.

[144] C. Lu, “Introduction to mass spectrometry.,” pp. 0–10, 2000.

[145] J. R. Yates, C. I. Ruse, and A. Nakorchevsky, “Proteomics by Mass Spectrometry: Approaches, Advances, and Applications,” *Annu. Rev. Biomed. Eng.*, vol. 11, no. 1, pp. 49–79, 2009.

[146] J. Jang-Lee *et al.*, “Glycomic Profiling of Cells and Tissues by Mass Spectrometry: Fingerprinting and Sequencing Methodologies,” *Methods Enzymol.*, vol. 415, no. 6, pp. 59–86, 2006.

[147] M. Campbell, K. Aoki-Kinoshita, and F. Lisacek, “Glycoinformatics,” in *Essentials of Glycobiology*, 3rd editio., A. Varki, R. Cummings, and J. Esko, Eds. Cold Spring Harbor Laboratory Press, 2017.

[148] R. I. Freshney, *Culture of Animal Cells: A Manual of Basic Technique and Specialized Applications*, Sixth Edit. John Wiley & Sons, Inc., 2011.

[149] M. W. Qoronfleh, B. Benton, R. Ignacio, and B. Kaboord, “Selective Enrichment of Membrane Proteins by Partition Phase Separation for Proteomic Studies,” *J. Biomed. Biotechnol.*, vol. 2003, no. 4, pp. 249–255, 2003.

[150] D. K. Moscatello *et al.*, “Frequent expression of a mutant epidermal growth factor receptor in multiple human tumors.,” *Cancer Res.*, vol. 55, no. 23, pp. 5536–5539, 1995.

[151] M. Voigt *et al.*, “Functional Dissection of the Epidermal Growth Factor Receptor Epitopes Targeted by Panitumumab and Cetuximab,” *Neoplasia*, vol. 14, no. 11, p. 1023–IN3, 2012.

[152] J. Schmiedel, A. Blaukat, S. Li, T. Knöchel, and K. M. Ferguson, “Matuzumab Binding to EGFR Prevents the Conformational Rearrangement Required for Dimerization,” *Cancer Cell*, vol. 13, no. 4, pp. 365–373, 2008.

[153] D. Patel *et al.*, “Monoclonal antibody cetuximab binds to and down-regulates constitutively activated epidermal growth factor receptor vIII on the cell surface,” *Anticancer Res.*, vol. 27, no. 5 A, pp. 3355–3366, 2007.

[154] P. Léo, A. L. L. Galesi, C. A. T. Suazo, and A. Moraes, “Animal Cells: Basic Concepts.,” in *Animal Cell Technology: From Biopharmaceuticals to Gene Therapy.*, L. Castilho, A. Moraes, E. Augusto, and M. Butler, Eds. New York: Taylor & Francis Group, 2008.

[155] D. L. Jarvis, Z. S. Kawar, and J. R. Hollister, “Engineering N-glycosylation pathways in the baculovirus-insect cell system,” *Curr. Opin. Biotechnol.*, pp. 528–533, 1998.

[156] C. F. Goochee, M. J. Gramer, D. C. Andersen, J. B. Bahr, and J. R. Rasmussen, “The oligosaccharides of glycoproteins: Bioprocess factors affecting oligosaccharide structure and their effect on glycoprotein properties,” *Bio/Technology*, vol. 9, no. 12, pp. 1347–1355, 1991.

[157] F. Pontén *et al.*, “A global view of protein expression in human cells, tissues, and organs,” *Mol. Syst. Biol.*, vol. 5, no. 337, pp. 1–9, 2009.

[158] N. Shibuya, I. J. Goldstein, W. F. Broekaert, M. Nsimba-Lubaki, B. Peeters, and W. J. Peumans, “The elderberry (Sambucus nigra L.) bark lectin recognizes the Neu5Ac(alpha 2-6)Gal/GalNAc sequence.,” *J. Biol. Chem.*, vol. 262, no. 4, pp. 1596–1601, 1987.

[159] G. Poiroux, A. Barre, E. J. M. van Damme, H. Benoist, and P. Rougé, “Plant lectins targeting O-glycans at the cell surface as tools for cancer diagnosis, prognosis and therapy,” *Int. J. Mol. Sci.*, vol. 18, no. 6, 2017.

[160] Y. M. Wu *et al.*, “Mucin glycosylating enzyme GALNT2 regulates the malignant character of hepatocellular carcinoma by modifying the EGF receptor,” *Cancer Res.*, vol. 71, no. 23, pp. 7270–7279, 2011.

[161] A. M. Wu, J. H. Wu, M. S. Tsai, Z. Yang, N. Sharon, and A. Herp, “Differential affinities of Erythrina cristagalli lectin (ECL) toward monosaccharides and polyvalent mammalian structural units,” *Glycoconj. J.*, vol. 24, no. 9, pp. 591–604, 2007.

[162] K. Yamamoto, T. Tsuji, I. Matsumoto, and T. Osawa, “Structural Requirements for the Binding of Oligosaccharides and Glycopeptides to Immobilized Wheat Germ Agglutinin,” *Biochemistry*, vol. 20, no. 20, pp. 5894–5899, 1981.

[163] T. Iskratsch, A. Braun, K. Paschinger, and I. B. H. Wilson, “Specificity analysis of lectins and antibodies using remodeled glycoproteins,” *Anal. Biochem.*, vol. 386, no. 2, pp. 133–146, 2009.

[164] E. D. Green and J. U. Baenziger, “Oligosaccharide specificities of Phaseolus vulgaris leukoagglutinating and erythroagglutinating phytohemagglutinins. Interactions with N-glycanase-released oligosaccharides.,” *J. Biol. Chem.*, vol. 262, no. 25, pp. 12018–12029, 1987.

[165] M. K. Hall, D. A. Weidner, Y. Zhu, S. Dayal, A. A. Whitman, and R. A. Schwalbe, “Predominant expression of hybrid N-glycans has distinct cellular roles relative to complex and oligomannose N-glycans,” *Int. J. Mol. Sci.*, vol. 17, no. 6, 2016.

[166] A. M. Wu and S. Sugii, “Coding and classification of d-galactose, N-acetyl-d-galactosamine, and β-d-Galp-[1→3(4)]-β-d-GlcpNAc, specificities of applied lectins,”

Carbohydr. Res., vol. 213, no. C, pp. 127–143, 1991.

[167] F. Fukumori *et al.*, “Cloning and expression of a functional fucose-specific lectin from an orange peel mushroom, Aleuria aurantia,” *FEBS Lett.*, vol. 250, no. 2, pp. 153–156, 1989.

[168] P. Kuhn, A. L. Tarentino, T. H. Plummer, and P. Van Roey, “Crystal Structure of Peptide-N4-(N-acetyl-β-d-glucosaminyl)asparagine Amidase F at 2.2-Å Resolution,” *Biochemistry*, vol. 33, no. 39, pp. 11699–11706, 1994.

[169] C. J. Janeway, P. Travers, M. Walport, and E. Al, “The interaction of the antibody molecule with specific antigen,” in *Immunobiology: The Immune System in Health and Disease.*, 5th editio., New York: Garland Science, 2001.

[170] J. R. Terman and A. Kashina, “Post-translational modification and Regulation of Actin,” *Curr. Opin. Cell Biol.*, vol. 25, no. 1, pp. 30–38, 2014.

[171] T. Winslow, “Esophageal cancer,” 2005. [Online]. Available: https://www.cancer.gov/types/esophageal/hp/esophageal-treatment-pdq.

[172] D. Peckys and N. Jonge, “Liquid Scanning Transmission Electron Microscopy: Imaging Protein Complexes in their Native Environment in Whole Eukaryotic Cells,” no. February, 2014.

[173] X. Varelas, M. P. Bouchie, and M. A. Kukuruzinska, “Protein N-glycosylation in oral cancer: Dysregulated cellular networks among DPAGT1, E-cadherin adhesion and canonical Wnt signaling,” *Glycobiology*, vol. 24, no. 7, pp. 579–591, 2014.