UNIVERSITA' DELLA CALABRIA

Dipartimento di Chimca e Tecnologie Chimiche

Dottorato di Ricerca in

Organic Matreials of Pharmacological Interest (OMPI)

XXVII Ciclo (CHIM/01)

ADVANCED MASS SPECTROMETRY-BASED STRATEGIES FOR THE ISOLATION AND CHARACTERIZATION OF G PROTEIN-COUPLED ESTROGEN RECEPTOR 1 (GPER)

Coordinatore:

Ch.mo Prof. Bartolo GABRIELE

Supervisori:

Ch.mo Prof. Giovanni SINDONA

Ch.ma Prof.ssa Anna NAPOLI

Co-supervisore:

Ch.mo Prof. Kenneth J SHEA (UCI)

Dottorando:

Dott. Hariprasad THANGAVEL

A.A. 2011 - 2014

To my parents,

Malathi and Thangavel,

and my beloved brother,

for their endless support.

ABSTRACT

Estrogen signaling plays a vital role in breast, ovarian and endometrial cancers. The actions of estrogen are mainly mediated by classical estrogen receptors, ERα and ERβ that belongs to the nuclear receptor superfamily. In recent years, a class of membrane-associated estrogen receptors are found to mimic the functions of classical ERs, including genomic as well as non-genomic signaling. These non-genomic signaling events include pathways that are usually thought of as arising from transmembrane growth factor receptors and G protein-coupled receptors (GPCRs). GPCRs belong to a superfamily of cell surface signaling proteins. GPCRs represent the most significant family of validated pharmacological targets in medical biology. A member of the GPCR family, named GPER, mediates rapid biological responses to estrogen in diverse normal and cancer cells, as well as transformed cell types. The identification and characterization of GPER will lead to understand the mechanisms underlying complex biological pathways and identify potentially new drug targets.

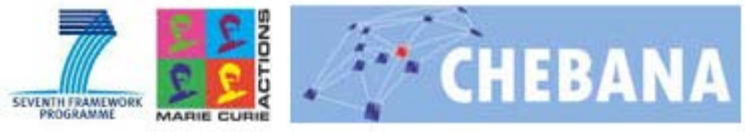
Here, we proposed a novel gel-free method to isolate and enrich GPER from crude lysate using home-made hydroxyapatite column (HTP). The HTP eluate was subjected to cellulose acetate (CA) filteration, followed by on-membrane protein digestion with different proteases and analyzed by MALDI MS. GPER was identified by peptide mass fingerprinting (PMF) after intensive data analysis. Sequence analysis reports 3 potential N-glycosylation in GPER. We manually validated 2 out of 3 glycosylation sites in GPER from the obtained MS/MS data and also validated the glycan moieties predicted by Glycomod. This approach is the first of its kind to identify GPER and characterize post-translational modifications (PTMs) by MS-based proteomic analysis. The proposed method is simple, robust and unique with great reproducibility. Finally, we designed and synthesized polymer nanoparticles (NPs) in an effort to capture GPER with high affinity and selectivity from crude lysate. PNIPAm-based NPs were synthesized by a free radical precipitation polymerization method with no control

over the functional monomer sequence. The NP binding affinity was evaluated against both truncated-GPER (short peptide epitopes) and GPER (whole protein). As the NPs were designed with complementary functionality against the peptides/protein, the NPs-peptide/protein binding will be through multipoint interactions. The initial qualitative results obtained by immunoblotting analysis revealed interesting hints on GPER's competitive affinity towards NPs when probed against multiple antibodies. We anticipate to use this strategy as a sample purification step prior to MS-based proteomic analysis.

Key words: GPR30/GPER, breast cancer, MALDI MS, *N*-glycosylation, synthetic polymer nanoparticles.

ACKNOWLEDGEMENTS

This thesis has been written as a part of my three year PhD at Department of Chemistry and Chemical Technologies in the University of Calabria, Italy, between March 2012 and February 2015. My research project was funded by the European commission under grant agreement number 264772 (ITN-CHEBANA).



First and foremost, I praise God, the almighty for providing me this opportunity and granting me the capability to proceed successfully. This thesis appears in its current form due to the assistance, guidance and encouragement of several people. I would therefore like to offer my sincere thanks to all of them.

I would like to express my deep and sincere gratitude to Prof. Giovanni Sindona, my esteemed supervisor and the director of the department of Chemistry and Chemical Technologies, for accepting me as a PhD student, and for his warm encouragement and continuous support throughout my PhD. I could not have imagined having a better supervisor for my PhD study.

I am deeply grateful to my cosupervisor, Prof.ssa Anna Napoli, whose support and motivation from the initial to the final level enabled me to develop an understanding of the subject. Her understanding and thoughtful guidance have provided a good basis for this thesis.

I would like to thank Dott.ssa Donatella Aiello for her advises and friendly assistance with various problems time to time. I greatly appreciate her time in teaching me the knowhow about the MALDI instrument during my initial days.

I express my sincere thanks to Prof. Marcello Maggiolini for his critical suggestions, and Dott.ssa Assunta Pisano for providing me the samples for my research work.

Next, I cordially thank Prof. Kenneth J Shea, for accepting me as a visiting scholar and giving me the opportunity to work in SHEA lab (UCI). I really had a valuable and enjoyable experience working with his team. His excellent guidance, encouragement and motivation have been of great value in this thesis.

I especially want to express my huge thanks to Dr Keiichi Yoshimatsu, Post Doctoral Research Fellow in SHEA lab, who gave his continuous help and support by answering all of my questions related to the research work. He was my colleague and my best “buddy” at UCI.

I would like to extend my thanks to Prof. Jennifer Prescher for understanding my research needs and allowing me work in her lab, utilizing her analytical instruments. I share my thanks to Lidia Nazarova, graduate student in Prescher lab (UCI), for helping me with me all my research needs throughout my stay.

My warm and special thanks are due to Dr Yi Ge, my MSc course coordinator (Cranfield University), for giving me confidence and support on all aspects of my career after MSc. For me, he is more than a course tutor, without his personal advice this PhD would not have been possible.

I would like to thank Prof. Bartolo Gabriele, coordinator of OMPI curriculum and Prof. Roberto Bartolino, director of ‘Bernardino Telesio’ Doctoral school of science and technology, UNICAL.

I cannot finish without thanking my family. I warmly thank and appreciate my parents and my brother. They were always supporting me and encouraging me with their best wishes.

Finally, I would like to thank my dearest friend, Dhanya Dhanyalayam. She was always there cheering me up and stood by me through the good times and bad.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

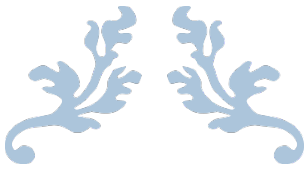
LIST OF ACRONYMS & ABBREVIATIONS

2DE	Two-dimensional gel electrophoresis
7TMRs	7-transmembrane receptors
AAc	Acrylic acid
APS	Ammonium per sulfate
ARs	Androgen receptors
BIS	Bis-acrylamide
CA	Cellulose acetate
cAMP	cyclic-adenosine monophosphate
cGMP-PDE	cyclic guanosine monophosphate phosphodiesterases
CHCA	α-Cyani-4-hydroxycinnamic acid
CID	Collision induced dissociation
CL2	Cytoplasmic loop - 2
CL3	Cytoplasmic loop – 3
CTGF	Connective tissue fgrowth factor
Cys	Cysteine
DDT	Dichlorodiphenyltrichloroethane
DHB	2,5-Dihydroxybenzoic acid
D_H	hydrodynamic diameter
DLS	Dynamic light scattering
DMEM	Dulbecco's modified eagle medium
DMF	N,N-Dimethylformamide
DNA	Deoxyribonucleic acid
DNS-Cl	Dansyl chloride
E1	Estrone
E2	Estradiol (17β-estradiol)
E3	Estriol

EDT	Ethanedithiol
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
EGTA	Ethylene glycol tetraacetic acid
EL1	Endoplasmic loop - 1
EL2	Endoplasmic loop - 2
EL3	Endoplasmic loop - 3
ELISA	Enzyme-linked immuno sorbent assay
ERs	Estrogen receptors
ERK	Extracellular-signal-regulated kinase
FBS	Fetal bovine serum
FMOC-CL	Fluorenylmethyloxycarbonyl chloride
GDI	Guanine nucleotide dissociation inhibitor
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
GLcNAc	*N*-acetylgucoseamine
GPCR	G protein-coupled receptor
GPER	G protein-coupled estrogen receptor
GPR30	G protein-coupled receptor 30
GRK	G protein-coupled receptor kinase
GTP	Guanosine triphosphate
GTs	Glycosyltransferases
HB-EGF	Heparin-bound-epidermal growth factor
Hex	Hexose
HexNAc	*N*-acetylhexoseamine
hGPER	human G protein-coupled estrogen receptor
HPLC	High performance liquid chromatography
HTP	Hydroxyapatite
IgG	Immunoglobulin G

IP3	Inositol triphosphate
IR	Infrared
LCST	Lowest critical solution temperature
m/z	mass-to-charge ratio
MALDI	Matrix-assisted laser desorption/ionization
Man	Mannose
MAPK	Mitogen-activated protein kinase
MC	Missed cleavage
Met	Methionine
MIP	Molecularly imprinted polymer
MMP	Matrix metalloproteinase
mRNA	messenger Ribonucleic acid
MS	Mass spectrometry
MWCO	Molecular weight cut-off
NCBI	National center for biotechnology information
Nd:YAG	Neodymium doped: yttrium aluminium garnet
NMR	Nuclear magnetic resonance
NPs	Synthetic polymer nanoparticles
OHT	4-hydroxytamoxifen
P1	Peptide 1
P2	Peptide 2
P3	Peptide 3
P4	Peptide 4
P5	Peptide 5
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PEG	Poly ethylene glycol
pI	Isoelectric point
PI3K	Phosphoinositide 3-kinase

PLC	Phospholipase C
PMSF	Phenylmethanesulfonyl fluoride
PNGase	Peptide-*N*-glycosidase
PNIPAm	Poly(N-isopropylacrylamide)
ppm	parts per million
PTM	Post-translational modification
RGS	Regulators of G protein
RPMI	Rosewell park memorial institute medium
SDS	Sodium dodecyl sulfate
SERDs	Selective estrogen receptor down regulators
SERMs	Selective estrogen receptor modulators
SPE	Solid-phase extraction
SPPS	Solid phase peptide synthesis
TBAm	N-tert butylacrylamide
TFA	Trifluoroacetic acid
TIS	Thioanisole
TM1	Transmembrane - 1
TM3	Transmembrane - 3
TM4	Transmembrane - 4
TM7	Transmembrane - 7
TOF	Time-of-flight
UV	Ultra violet
WB	Western blot analysis



CHAPTER 1

GENERAL INTRODUCTION

1.1 Receptors and ligands: an overview

Cell membranes in eukaryotes are naturally equipped with thousands of receptors, of many different kinds. Eukaryotic cells also encase their cell organelles like nucleus, ribosomes, endoplasmic reticulum, Golgi apparatus, mitochondria, lysosome with internal membranes that play host to a bunch of intracellular receptors. In general, receptors are nothing but protein molecules, ingrained in either the cell membrane or the cytoplasm of a cell. As the name denotes, receptors are macromolecular structures that receive information.[1] More specifically, receptors enable cells to sense stimuli or physical changes in the internal or external environment, so that the cells can adjust to new situations. Based on their physical presence, receptors can be easily put into two broad categories, cell surface receptors and internal receptors (cytoplasmic and nuclear receptors). Receptors that are found on the membrane of internal cell organelles are also categorized under intracellular receptors, and interestingly, they share functional similarities with cell surface receptors.

Receptor proteins recognize and respond to endogenous chemical signals. The chemical signals can act either at the plasma membrane or within the cytoplasm (or nucleus) of the target cell.[2] The signaling molecules that bind to the receptors are referred to as ligands and can be endogenous or exogenous in origin. A ligand can be any small molecule such as light-sensitive compound, odorant molecule, hormone, pheromone, growth factor, cytokine, neurotransmitter, toxin, pharmaceutical drug, or peptide (small protein).[3][4] Irrespective of the nature of initiating signal, the cellular responses are determined by the presence of receptors that specifically binds the signaling molecules.[2] Each receptor is unique and assigned to activate a specific cellular biochemical pathway when triggered. Every single receptor will just tie to ligands of a specific structure. Receptor-ligand interaction can be compared to a lock and key system, where a lock will just accept a specifically fashioned key **(Fig. 1.1)**.[5] On binding its corresponding receptor, the ligand initiates or inhibits the receptor's designated biochemical pathway.

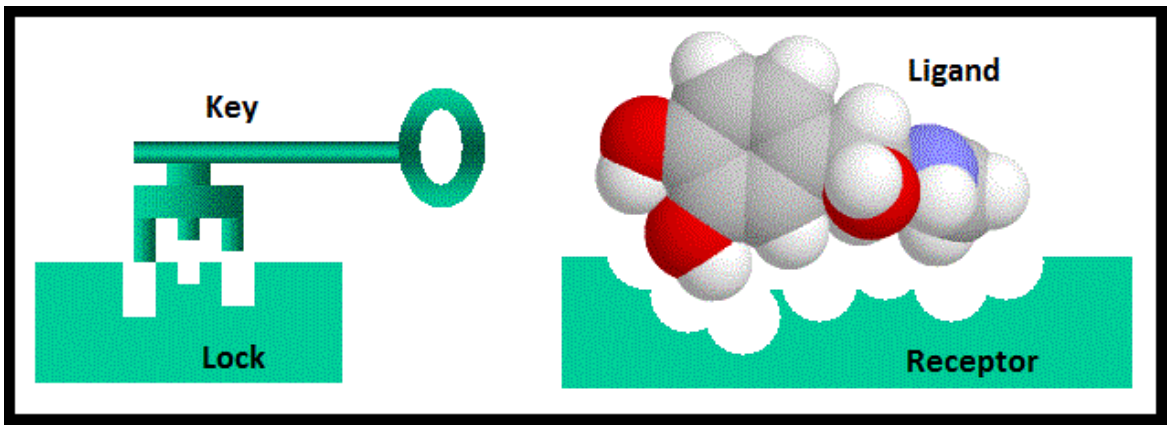


Figure 1.1 Lock and key model.

1.1.1 Cell surface receptors

Human cells are constantly communicating with each other and the outside world through the specialized integral membrane proteins that are collectively known as cell surface receptors (membrane receptors, transmembrane receptors). Cell surface receptors bind to an external ligand molecule and perform signal transduction, converting an extracellular signal into an intracellular signal. By doing so, the cell surface receptors play a unique and significant role in cellular communications and signal transduction. Ligands that interact with cell surface receptors are mostly impermeant signal molecules that can't enter the cell. Every cell surface receptor has three main components: an *N*-terminal ligand binding domain (extracellular domain), a hydrophobic membrane-spanning region, and a *C*-terminal cytoplasmic domain (intracellular domain) inside the cell. The extracellular domain usually includes the binding site for the ligand, while the intracellular domain activates a series of intracellular signaling events once the ligand binds. The size and extent of each of these domains vary extensively, depending on the type of receptor. So far, a wide range of these receptors have been identified and studied.[2] They are grouped into three main classes of receptors, namely: ligand-gated ion channel receptors, enzyme-coupled receptors, and G protein-coupled receptors. The names of these receptor classes are defined by the mechanism used to transform external signals into internal ones - via ion channel opening, enzyme activation, or protein action,

respectively. Because cell surface receptors interact with signal molecules or ligands externally and permit them to affect cell function without actually entering the cell.[6]

Ligand-gated ion channel receptors **(Fig. 1.2)** are also known as ionotropic receptors. These receptors bind a ligand and open a channel that allows the flow of specific types of ions such as Na^+, K^+, Ca^+ or Cl^- across the cell membrane, which changes the membrane potential, causing an electric current.[7] To form a channel, this type of cell surface receptor have an extensive membrane-spanning region. In order to interact with the phospholipid fatty acid tails that form the crux of the cell membrane, many of the amino acids in the membrane-spanning region are hydrophobic in nature. In contrast, the amino acids that line up on the inside of the channel are hydrophilic to allow the passage of water or ions. These receptors are responsible for the rapid transmission of signals across synapses in the nervous system. Good examples of such receptors are the neurotransmitter receptors.[2] Although the ligand-gated ion channel receptors are found mainly in the nervous system and other electrically excitable cells such as muscle cells, the other two types of cell surface receptors are found particularly in every cell type of the body.[7]

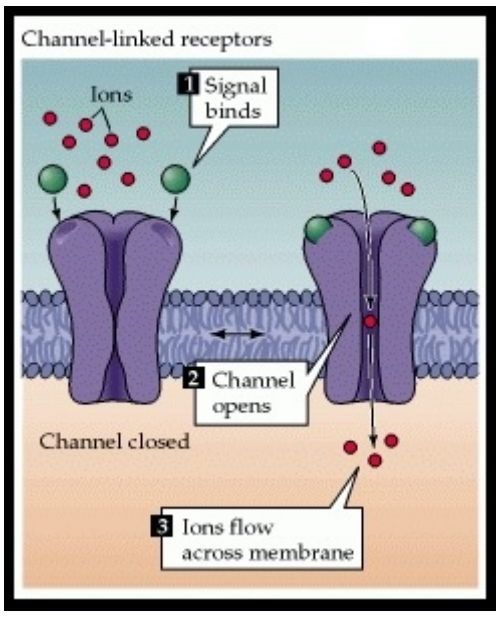


Figure 1.2 Ligand-gated ion channel receptors.[2]

Enzyme-coupled receptors **(Fig. 1.3)** are cell surface receptors, composed of an extracellular domain containing the ligand binding site and an intracellular domain, often associated with an enzyme. In some cases, the intracellular domain of such receptor itself is an enzyme whose catalytic activity is regulated by the binding of an extracellular chemical signal. As of 2009, only six types of such receptors are known and they are receptor tyrosine kinases, tyrosine kinase associated receptors, receptor-like tyrosine phosphatases, receptor serine/threonine kinases, receptor guanylyl cyclases, and histidine kinase associated receptors.[8] The great majority of them are protein kinases, often tyrosine kinases, which phosphorylate intracellular target proteins, thereby changing the physiological function of the target cells.[2]

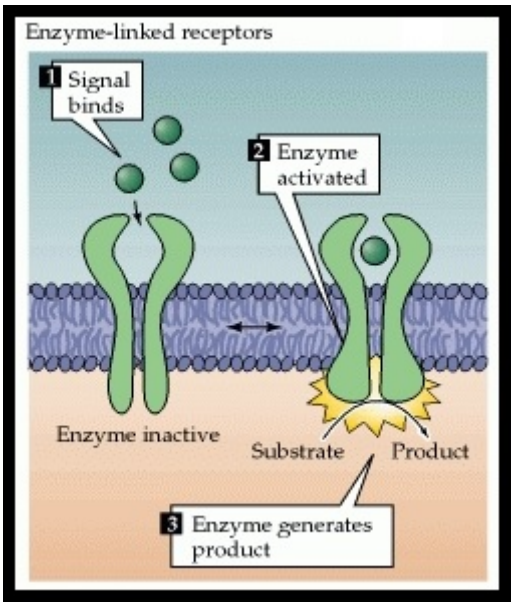


Figure 1.3 Enzyme-coupled receptors.[2]

The enzyme-coupled receptors normally have large extracellular and intracellular domains, but the membrane spanning-region consists of a single alpha-helical region of the peptide strand.[9] On binding their ligands externally, the receptors undergo conformational change that activates the enzyme, which then turn on a variety of intracellular signaling pathways. They are discovered through their role in responses to extracellular signal proteins that regulates the growth, proliferation, differentiation and

survival of cells in animal tissues. Disorders of cell growth, proliferation, differentiation, survival and migration are fundamental to cancer, and abnormalities in signaling via enzyme-coupled receptors have a major role in the development of this class of diseases.[8]

G protein-coupled receptors (GPCRs) **(Fig. 1.4)** are the largest of all the cell surface receptors. GPCRs bind a ligand and activate a membrane-bound, trimeric GTP-binding protein (G protein). The activated G protein then interacts with either an ion channel (effector) or an enzyme in the cell membrane, initiating a sequence of other effects. All GPCRs share the structural feature of crossing the cell membrane seven times, but each receptor has its own specific extracellular domain containing the ligand binding site and intracellular domain with G protein binding site.[9] GPCRs are also referred to as 7-transmembrane receptors (7TMRs), heptahelical receptors, serpentine receptors or metabotropic receptors. Metabotropic receptors do not form an ion channel passage, instead, they are indirectly linked with ion channels on the cell membrane through signal transduction mechanisms.[10] Heterotrimeric G proteins have three subunits: α, β, and γ. When a ligand binds to a G protein-coupled receptor in the cell membrane, a guanosine diphosphate (GDP) molecule associated with the α subunit is exchanged for guanosine triphosphate (GTP). The β and γ subunits dissociate from the α subunit, and a cellular response is triggered either by the α subunit or the dissociated β-γ complex. Hydrolysis of GTP to GDP terminates the signal.[9] Cell signaling using GPCRs occurs as a cyclic series of events. These receptors mediate responses involving hormones, local mediators and neurotransmitters.[2] Because of their involvement in wide range of cellular processes, GPCRs are typically an appealing target for the development of drugs to treat a number of diseases.[11] Hundreds of different GPCRs have been identified so far. Some of the well know examples include the β-adrenergic receptor, metabotropic glutamate receptors, receptors for odorants in the olfactory system, and many types of receptors for peptide hormones.[11]

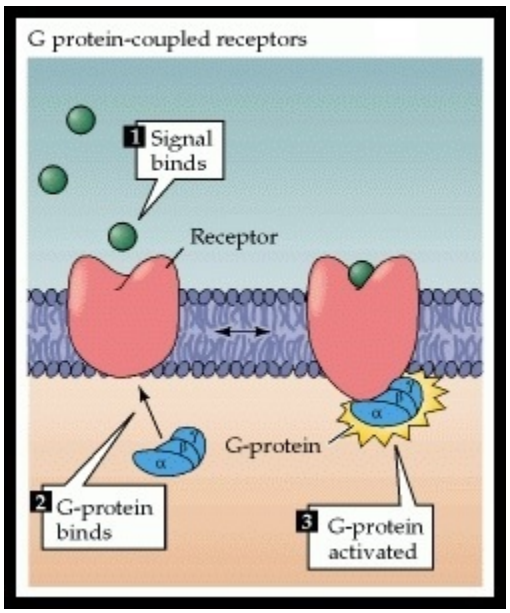


Figure 1.4 G protein-coupled receptors.[2]

1.1.2 Internal receptors

Internal receptors **(Fig. 1.5)**, also known as intracellular receptors, are found in the cytoplasm or nucleus of the cell and are normally activated by cell-permeant, hydrophobic or lipophilic ligand molecules that can pass through the cell membrane.[2]

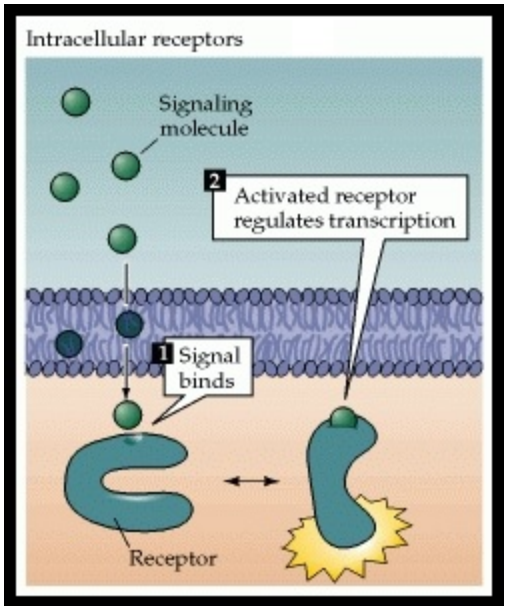


Figure 1.5 Internal receptors.[2]

To initiate signal transduction, these ligands must passively diffuse through cell membrane. On entering the cell, many of these molecules bind to proteins that act as regulators of mRNA synthesis to mediate gene expression. Gene expression is the cellular process of transforming the information in a cell's DNA into a sequence of amino acids that ultimately forms a protein. When the ligand binds to the internal receptor, a conformational change exposes a DNA-binding site on the protein. The ligand-receptor complex moves into the nucleus, binds to specific regulatory regions of the chromosomal DNA, and promotes the initiation of transcription **(Fig. 1.6)**.[12] As the ligand-receptor complex makes it all the way to the nucleus of the cell, these receptors are often called nuclear receptors.[13] Some intracellular receptors are located primarily in the cytoplasm, while others are in the nucleus. In either case, once these receptors are activated they can affect gene expression by altering DNA transcription. Internal receptors can directly influence gene expression without having to pass the signal on to other receptors or messengers. Intracellular receptors are used widely by some classic hormones such as thyroid and steroid hormones.[14][15]

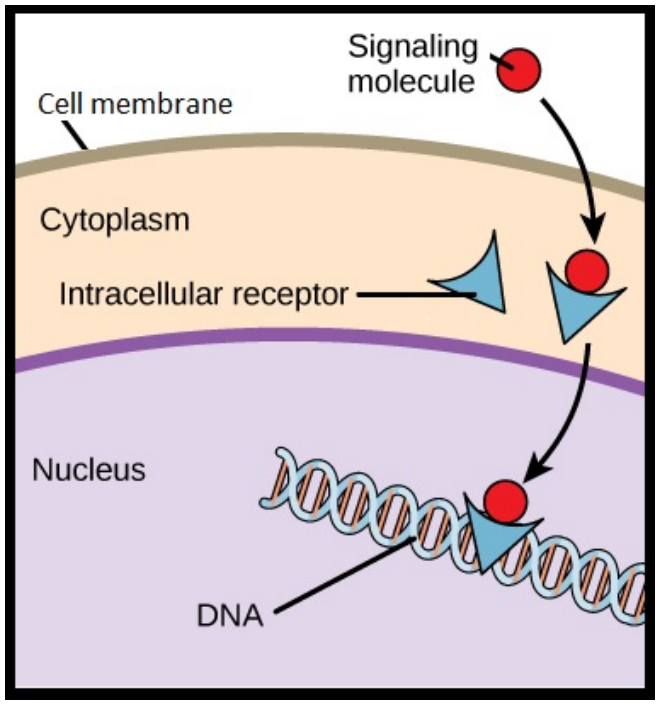


Figure 1.6 Signal transduction through internal receptors.[9]

Among all the above discussed receptors, GPCRs are the most abundant class of receptors in the human body.[16] They play a crucial role in an incredible range of functions in humans**.** More than one-half of all prescribed drugs achieve their effects by binding to GPCRs. However, only a small portion of GPCRs have been investigated as drug targets leaving a wide area to explore and understand.

1.2 G protein-coupled receptors

G Protein-coupled receptors constitute by far the largest and most distinct superfamily of cell membrane signaling proteins in eukaryotes, with their unique seven-transmembrane-helix structure. They transduce extracellular signals as exerted by a hormone or neurotransmitter to an intracellular effector pathway through the activation of heterotrimeric G proteins.[17] In human, nearly 800 different genes code for GPCRs, which account for ~4% of the entire protein-coding genome.[18] GPCRs are virtually expressed in all types of tissues in the body.[19] GPCRs involvement in numerous physiological processes and diseases including tumor growth and metastasis have been well documented in many scientific reports over the years. GPCRs have become drug targets for several life-threatening diseases. They are often expressed in low levels and in specific cell types, which contributes to the fact that they are the most important family of protein receptors serving as targets in drug discovery. An increased understanding of these receptors has significantly affected modern medicine.[20] Presently, one-quarter of the top 100 best-selling drugs are targeted mostly to GPCRs that bind amines. In 2012, the Nobel Prize in Chemistry was jointly awarded to Robert Lefkowitz and Brian Kobilka for their groundbreaking research work which gave the first insight on how GPCRs function.[21] Moreover, there have been at least seven other Nobel Prizes awarded for some aspect of G protein-mediated signaling in the past.

1.2.1 Structure of GPCRs

GPCRs consist of a single, serpentine-like polypeptide chain of variable length (from 300 to 1000 amino acids) that is folded into a globular structure and embedded in

the cell membrane.[22] Seven segments of this molecule span the entire width of the membrane explaining why GPCRs are sometimes called 7-transmembrane receptors (7TMRs). The intervening portions that connect the seven membrane spanning α-helices loop both inside and outside of the cell forming three intracellular and three extracellular loops **(Fig. 1.7)**. The extracellular amino terminal segment and cytoplasmic carboxyl terminal segment are attached to the TM1 domain and TM7 domain, respectively. Both termini are highly variable in length, and the amino-termini can comprise different functional domains each of which is able to provide specific properties to the relevant receptor.[23] Some GPCRs bear amine-linked glycosylation sites near their amino terminal segment. The three extracellular loops (EL1, EL2 and EL3) are considered to play an important role in structure stabilization and ligand binding, whereas, the cytoplasmic loops (CL2 and CL3) are mainly engaged in G protein recognition and activation.[22]

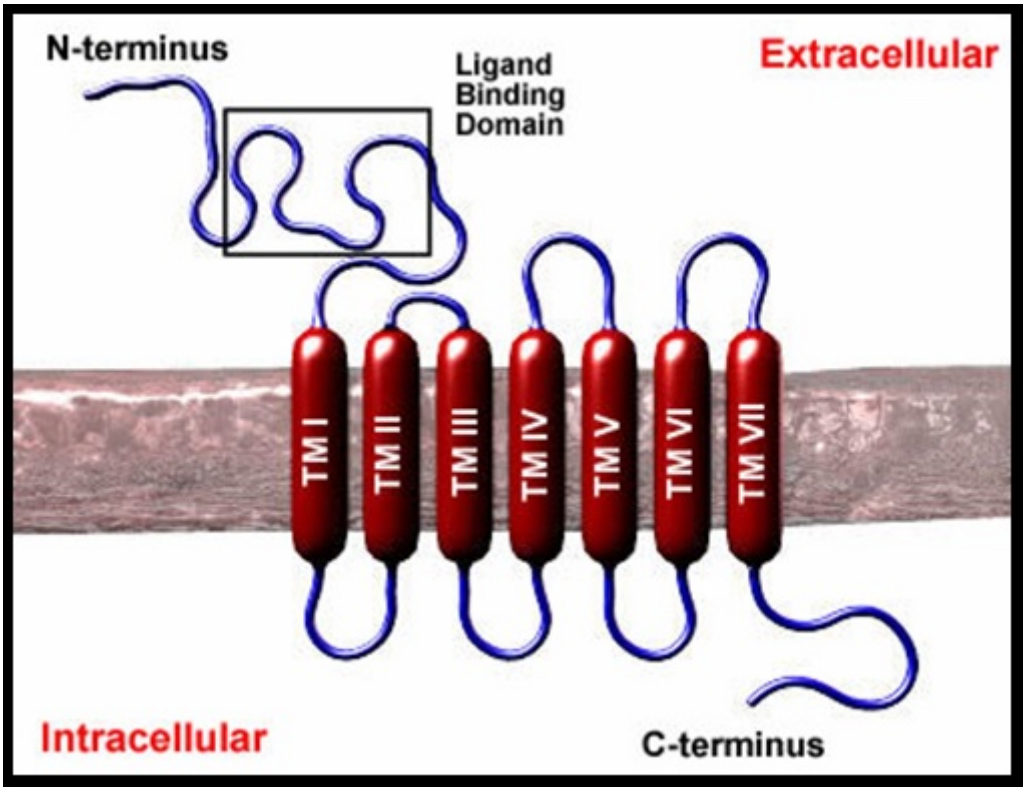


Figure 1.7 Structural representation of GPCR.

1.2.2 G proteins: types and functions

G proteins are specialized proteins with the ability to bind the nucleotides GTP and GDP. The G protein acts as a molecular switch by binding either GTP (active/on) or GDP (inactive/off). Some G proteins, such as the signaling protein Ras, are small protein with a single subunit. The G proteins activated by GPCRs are trimeric in structure consisting of an α-, β- and γ- subunit. There are more than 20 different α-subunits, 6 different β-subunits and 12 different γ-subunits, creating a large number of theoretical combinations.[24] However, only a small number of combinations form biological complexes. In humans, 16 Gα genes encode 23 known Gα isoforms.[24][25] Based on sequence similarities, Gα proteins are grouped into 4 classes including $G\alpha_{(S)}$, $G\alpha_{(i/o)}$, $G\alpha_{(q/11)}$ and $G\alpha_{(12/13)}$.[26] The Gα subunit binds to GDP when inactive. The tightly associated Gβγ complex functions as a single unit and facilitates the association of Gα to the cytoplasmic part of the GPCR. Moreover, it inhibits the release of GDP from Gα and acts as a guanine nucleotide dissociation inhibitor (GDI).[24] Each Gα and Gβγ subunits activate and regulate specific pathways that are shown in **Table 1.1**. The βγ subunits of G protein can also act as second messenger molecules, although their actions are not completely characterized.

Table 1.1 G protein types and associated functions.

Type	Pathways and functions	References
$G\alpha_{(S)}$	Activates Ca^{2+} channels, stimulates adenylyl cyclase pathway and cyclic adenosine monophosphate (cAMP) production	[16] [27]
$G\alpha_{(i/o)}$	Activates K^{+} channels, inhibits Ca^{2+} channels, inhibits adenylyl cyclase and cAMP production	[16] [28]
$G\alpha_{(q/11)}$	Stimulates phospholipase C (PLC) pathway	[16] [29]
$G\alpha_{(12/13)}$	- Diverse ion transporter interactions - Regulates G protein RhoA and stimulates PDZ-Rho guanine nucleotide exchange factors (PDZ-RhoGEF)	[16] [24]

Gβγ complex	- Inhibits the release of GDP from Gα and acts as GDI - Regulates Ca^{2+} and K^{+} channels - Regulates kinase and small G protein including Phosphoinositide 3-kinase-γ (PI3Kγ) - Various other regulation pathways have been considered for downstream activation of Gβγ	[24]

1.2.3 Classification of GPCRs

There are many different approaches for classifying the GPCRs. Both physiological and structural features have been used to classify GPCRs. The most commonly used system of classification is that implemented in the GPCRDB database, which divide GPCRs into six classes (Class A-F).[30] This A-F system is designed for both vertebrate and invertebrate GPCRs. Class A contain rhodopsin-like and biogenic amine receptors, with over 80% of all GPCRs in humans; Class B: Secretin-like; Class C: Metabotropic glutamate receptors; Class D: Pheromone receptors; Class E: cAMP receptors; and the much smaller Class F contain Frizzled/smoothened receptors. Here, Classes A, B, C and F are found in mammalian species while Class D receptors are found only in fungi and Class E are exclusive to *Dictyostelium*.[31] The above six classes are further divided into sub-divisions and sub-sub-divisions based on the function of a GPCR and its specific ligand.[32] As some classes of the A-F system do not exist in human, an alternative classification system called GARFS has been proposed for classifying mammalian GPCRs. In GARFS system, the receptors are grouped into five major classes based on phylogenetic analyses and named *Glutamate* (G, with 15 members), *Rhodopsin* (R, with 701 members), *Adhesion* (A, with 30 members), *Frizzled/Taste2* (F, with 24 members) and *Secretin* (S, with 15 members). Only a few human receptors (nearly 23 protein sequences) could not be designated to any of the above five classes and these were thus categorized as “Other 7TMRs”. It is, however fairly straight forward to place most of these “other” receptors into any of the main classes or groups using sequence similarity only.[33]

1.2.4 GPCR activation and signaling

Generally, all GPCRs have three characteristic domains: a signal recognizing domain (extracellular), a signal transmission domain (transmembrane), and a signal response and amplification domain (intracellular).[34] GPCRs receive a wide range of ligands such as lipid analogues, amino acid derivatives, small peptides, as well as stimuli from light (photons), taste, odor (pheromones). The ligand is docked in a binding pocket that is usually present on the extracellular side.[35] The mechanism by which GPCRs transmit extracellular signals through the cell membrane to intracellular responses is mediated by heterotrimeric G proteins. Since GPCRs do not have intrinsic enzymatic activity, binding of a ligand to the external domain of GPCR triggers a conformational change in the receptor, specifically in an ionic interchange between the TM3 and TM4 domains, which leads to receptor activation. Thereby, transducing the ligand's message mechanically to the G protein which is closely associated to the intracellular or cytoplasmic side of the receptor and leads to different downstream signaling events.[36] Specific G proteins bind to specific GPCRs[37], it is hard to determine these pairings based on primary amino acid sequence. The interaction appears to depend on the whole tertiary structure of the GPCR.[35]

On receiving a signal, G protein becomes active, detaches from the GPCR and binds to an enzymatic effector protein lodged in the membrane. Activation of a single G protein can affect the production of hundreds or even thousands of second messenger molecules. The G proteins function as amplifiers, inducing the effectors to produce cascades of secondary messenger molecules that activate other enzymes, creating a diverse range of physiological responses.[34] Effector/second messenger systems include retinal cyclic guanosine monophosphate phosphodiesterases (cGMP-PDE), ion channels (potassium, calcium), and several phospholipases and adenylyl cyclase subtypes. A list of cellular activities controlled by the effector/second messenger systems are shown in **Table 1.2**.[35]

Table 1.2 Second messengers and their cellular activities.

Effector/second messenger system	Cellular activities
cGMP-PDE	- Conversion of light signal into electrical nerve activity in rod cells - Color vision in cone cells
Phospholipases	- Autocrine and paracrine regulation - Protein kinase C activation - Ion channel conductance - Neurotransmitter release - Smooth muscle contraction - Platelet activating factor synthesis
Adenylyl cyclases	- Gene transcription - Mitogenesis - Metabolism - Growth factor

Most G proteins involved in GPCR signaling are heterotrimeric with α, β, and γ subunits. When a ligand activates the GPCR, it induces a conformational change allowing the receptor to act as GEF that exchanges GDP for GTP on the Gα subunit. GTP binding promotes the dissociation of Gα from Gβγ, and then, the free GTP-bound Gα subunit and Gβγ heterodimer can activate various effector proteins, thus propagating an intracellular signaling cascade **(Fig. 1.8)**.[38] The signaling continues until the G proteins are inactivated by a mechanism dependent on the intrinsic GTPase activity of the Gα subunit, which is facilitated by the direct binding of regulators of G protein signaling (RGS) to activated GTP-bound Gα.[39] In simple, the intrinsic GTPase activity of the Gα subunit inactivates Gα by hydrolyzing GTP back to GDP. The inactivated GDP-bound Gα subunit reforms and inactivates G protein with Gβγ complex turning off other downstream events.[40] A single activated GPCR may activate multiple G proteins, and each G protein may activate numerous effector proteins, resulting in a considerable amplification of the signal.[41][42]

Upon prolong stimulation however, the receptors eventually inactivate even if their activating ligands remain bound. In this case, a G protein-coupled receptor kinase (GRK) phosphorylates the cytosolic portions of the activated receptor.[43] Once the receptor is phosphorylated in this way, it binds with high affinity to β-arrestin protein, which inactivates the receptor by preventing its interaction with G proteins and decreasing its response to ligands or agonists (desensitization).[42] β-arrestins also act as adaptor proteins and recruit the phosphorylated receptors to clathrin-coated pits from where the receptors are endocytosed and afterwards they can either be degraded in lysosomes or activate new signaling pathways.[44]

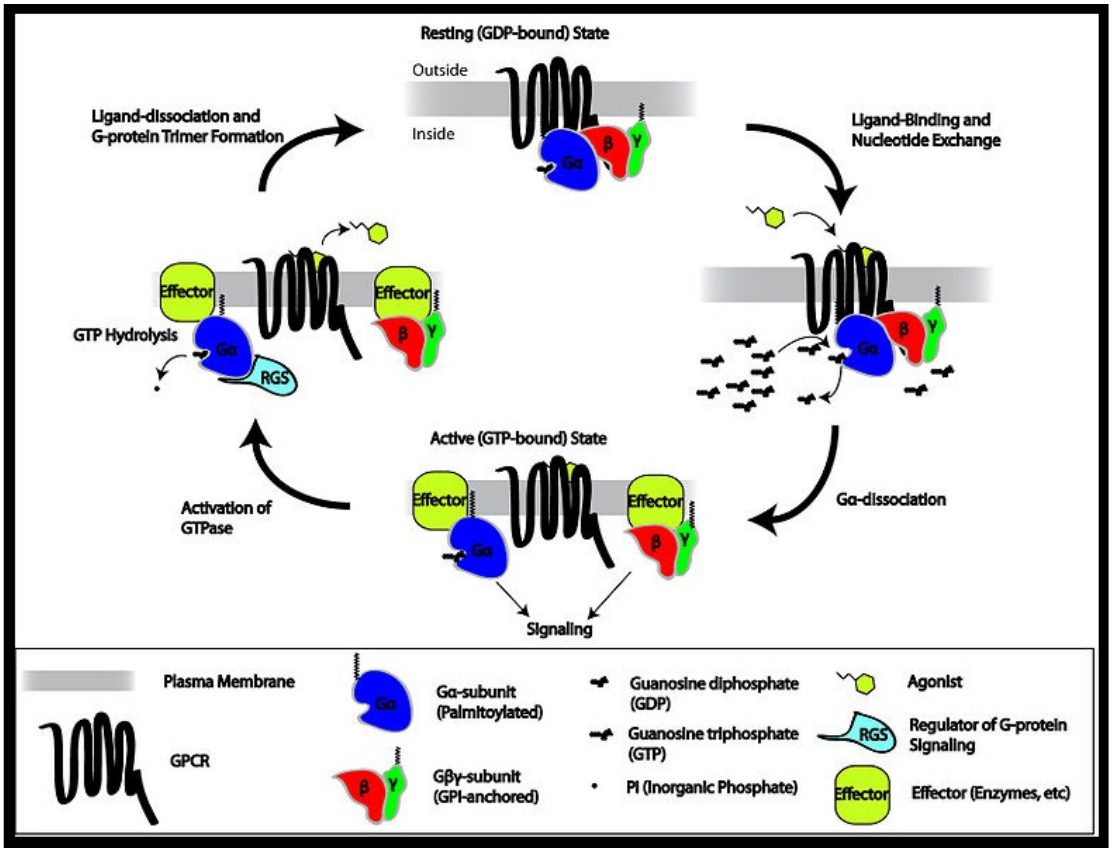


Figure 1.8 GPCR activation/deactivation cycle.

1.2.5 Clinical impact of GPCRs

Through this sequence of events, GPCRs help regulate an incredible range of bodily functions, from sensation to growth to hormone responses. Hence, proper functioning of this 'molecular switching system' is essential to the health of every individual organism. When the system malfunctions, the results can lead to acute or chronic human diseases, a partial listing of which includes cardiovascular disease (β_1 - adrenergic receptor)[45]; asthma (β_2 - adrenergic receptor)[46]; endometrial, ovarian and breast cancer (membrane estrogen receptor)[47]; and strokes and cerebral hypoperfusion (A_{2a} - adenosine receptor)[48][49]. Other disease states directly linked to mutations in GPCRs include retinitis pigmentosa (rhodopsin), female infertility (follicle-stimulating hormone receptor), nephrogenic diabetes insipidus (vasopressin receptor), familial exudative vitreoretinopathy (frizzled receptors), and dominant and recessive obesity (melanocortin receptors)[50].

1.3 GPCRs and cancer

Miscommunication is the hallmark of cancer. Normally, our cells are in constant communication, deciding how to share resources, determining the best time to grow, and if necessary, the best time to quit. In contrast, cancer cells typically have corrupted these lines of communication, allowing them to grow without limits and greedily reserve resources for themselves. GPCRs are among the many different molecules of communication that are changed when a normal cell is transformed into a cancer cell. In multicellular organisms, GPCRs became indispensable to integrate and coordinate the function and proliferation of individual cell types.[51] As an aberration of the normal relationships that organize cells coexistence, tumors commonly deceive cell-cell communication in order to expand and spread in the body. GPCRs represent critical elements in this process too.[52] An increasing number of studies link aberrant GPCR expression and activation to numerous types of human malignancies.[52][53] For instance, several GPCRs are overexpressed in different tumors[53] and GPCR variants can lead to

increased cancer risk. Some of the GPCRs that are more frequently implicated in human cancer are listed in **Table 1.3**.[52]

Table 1.3 GPCRs associated with human cancers.

Human cancer	Receptor	Ligand	Process
Breast cancer	PAR1	Thrombin	Growth; metastasis; angiogenesis
	EP2; EP4	PGE2	Growth; metastasis; angiogenesis
	CXCR4	SDF1	Metastasis; angiogenesis
	GPR30/GPER	Estrogen	Growth? Hormone-therapy resistance
Colon cancer	EP2; EP4	PGE2	Growth; metastasis; angiogenesis
	LPA_1	LPA	Growth
	ET receptors	Endothelin-1	Survival
	PAR1	Thrombin	Growth; migration
	Frizzleds	Wnts	Growth
Head and neck cancer	CXCR2	IL8; GROα	Growth; metastasis; angiogenesis
	CXCR4	SDF1	Metastasis
	EP receptors	PGE2	Growth; angiogenesis; metastasis
	GRPR	GRP	Growth; survival
	PAR1	Thrombin	Metastasis; angiogenesis
Small-cell lung cancer	GRPR	GRP	Growth
	NMB-R	Neuromedin B	Growth
	CCK_1; CCK_2	CCK	Growth; survival
	CXCR4	SDF1	Growth; metastasis
Non-small-cell lung cancer	EP receptors	PGE2	Growth; metastasis; angiogenesis
	CXCR2	IL8; GROα	Growth; metastasis; angiogenesis
	CXCR4	SDF1	Migration; metastasis
	β1AR; β2AR	NNK	Growth?
Ovarian cancer	LPA_1-LPA_3	LPA	Growth; metastasis; angiogenesis
	CXCR2	GROα	Growth; angiogenesis
Pancreatic cancer	GRPR	GRP	Growth

	CCK_1; CCK_2	CCK	Growth
Parathyroid gland cancer	CASR	Calcium	Growth
Pituitary cancer	TSH receptor	TSH	Growth; survival
	ACTHR	ACTH	Growth
Prostate cancer	PAR1	Thrombin	Growth; invasion
	ET_A	Endothelin-1	Growth; survival; metastasis
	AT1	Angiotensin II	Growth
	EP2; EP4	PGE2	Growth; metastasis; angiogenesis
	LPA_1	LPA	Growth; invasion
	B1; B2	Bradykinin	Growth; survival; invasion
	GRPR	GRP	Growth; migration
Melanoma	MC1R	MSH	Sensitivity to UV-induced DNA damage
	CXCR2	IL8; GROα	Growth; metastasis; angiogenesis
	ET_B	Endothelin-1/3	Growth
Basal-cell carcinoma	Smoothened	Sonic hedgehog	Growth
Testicular cancer	LH receptor	LH	Growth
Thyroid cancer	TSH receptor	TSH	Growth

A very recent genomic characterization (1507 coding genes from 441 tumors) of somatic mutations with in the cancer genomes of multiple cancer types revealed an underestimated role for G protein signaling.[54] Moreover, emerging scientific reports indicate that GPCRs have a crucial but often not fully appreciated role in cancer progression and metastasis. Malignant cells often hijack the normal physiological functions of GPCRs to proliferate autonomously, evade the immune system, increase their nutrient and oxygen supply, invade their surrounding tissues and disseminate to other organs.[52] GPCRs are also the target of key inflammatory mediators, therefore providing a probable link between chronic inflammation and cancer.[52] In addition, GPCRs

have a central role in tumor-induced angiogenesis, and that tumor metastasis might involve the GPCR-guided migration of cancer cells to their target organs.[52] Abnormal expression of GPCRs and/or their ligands is directly observed in cancer cells of various origins that abuse GPCRs signaling to directly stimulate growth, induce angiogenesis, inhibit apoptosis, promote spreading and induce immune-tolerance.[52][53] Therefore, interfering with GPCRs and their downstream targets might provide an opportunity for the development of new, mechanism-based strategies for cancer diagnosis, prevention, and treatment.[52] Despite GPCRs represent one of the major pharmaceutical targets; it is surprising that the clinical practice of cancer treatment includes only a few drugs that act on GPCR-mediated signaling.[51]

1.3.1 Steroid receptors in cancer

Numerous hormone-activated receptors are overexpressed in hormone-dependent and hormone-independent tumors and trigger multiple transduction pathways, which mediate relevant biological effects in diverse cancer cells.[47] Aberrant signaling of steroid receptors play a role in several diseases, including hormone-dependent cancers such as breast, ovarian, endometrial and prostate cancer. Hormonal therapy is often the treatment of choice for breast and prostate cancers, as even in advanced cases the growth of cancer cells is still largely dependent on estrogens and androgens, respectively.[55][56] The steroid hormones activate their allied estrogen (ER) and androgen (AR) receptors, which are transcription factors of the nuclear hormone receptor family.[57][58] As treatment continues some patients develop hormone-refractory cancer lesions, which are characterized by their rapid growth and invasiveness.[58] The aberrant activity of GPCRs might contribute to this progression from hormone-dependent to hormone-independent tumors, and might therefore represent suitable targets for the treatment of hormone-insensitive breast and prostate cancers.[52]

Cancers of female reproductive organs such as breast, ovarian and endometrial cancer are often dependent on steroid hormone, estrogen. These cancers tend to display variable expression of estrogen receptors (ERs) as well as various growth factor receptors

including epidermal growth factor receptor (EGFR). In addition to the cancers of female reproductive system, estrogen and ERs are reported to have roles in colon and prostate cancer.[59][60]

1.3.2 Physiological importance of estrogen

All living organisms produce hormones which act as chemical messengers in transmitting signals. Estrogen is a steroid hormone comprising a group of chemically similar compounds **(Fig. 1.9)** which include estrone (E1), estradiol (E2), and estriol (E3), and is the primary sex hormone in women, essential to the menstrual cycle. Steroid hormones, including 17β-estradiol (E2), regulate a wide range of physiological processes involved in the development and maintenance of an array of tissue types in mammals. Estrogen plays a key role in the development and general function of reproductive organs in women and has also shown to play a role in inflammation[61], cardiovascular protection[62][63], neuroprotection[64], and maintenance of bone structure and strength[65]. Development of reproductive organs, regulation of estrus and menstrual cycling, and establishing pregnancies and maintaining pregnancy to term are the three classical roles of estrogen in female reproductive physiology.[66]

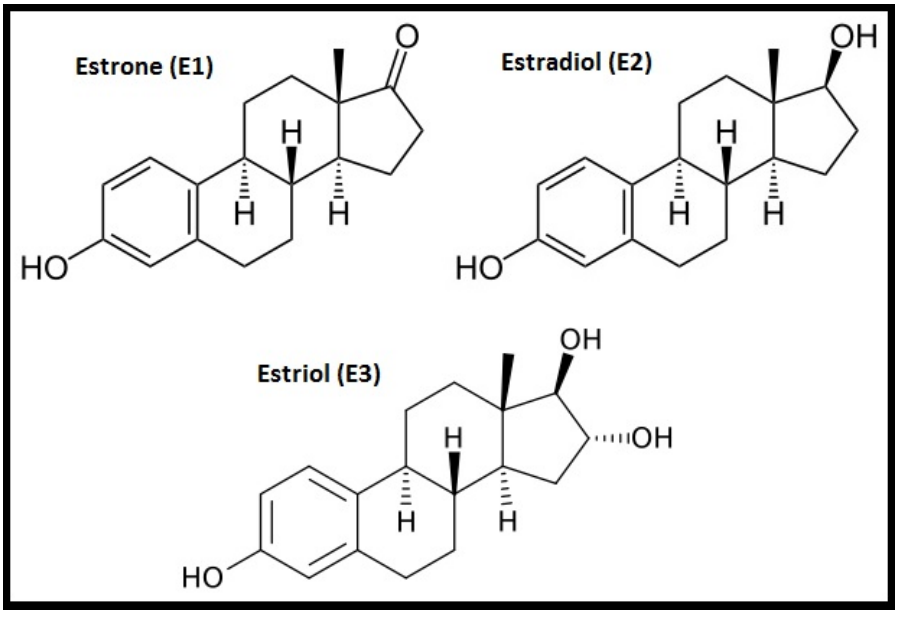


Figure 1.9 Chemical structure of estrone (E1), estradiol (E2) and estriol (E3).

The multiple biological actions elicited by estrogens are mainly mediated by the classical estrogen receptors, ERα and ERβ which belong to nuclear steroid hormone receptor superfamily.[67] ERα is critical to the development of mammary gland, as shown in mice lacking ERα, the branching ductal structures typically seen in mature mammary glands are not seen at the end of puberty, and rather, mice have the rudimentary, non-branching mammary glands that they are born with.[68] Both ERα and ERβ are involved in regulating the estrus cycle in mouse.[69] Estrogen is also critical to male fertility, shown by ERα knockout mice, which are infertile due to breakdown of testicular structures.[69]

E2 also plays a well-defined role in osteoporosis, wherein E2 normally acts to prevent osteoclast-mediated bone loss and promotes bone formation by osteoblasts; when E2 levels are deficient, such as following menopause, the rate of bone loss increases, resulting in osteoporosis.[70] In a study conducted to monitor the role of E2 in maintenance of bone density, a decrease in bone density is observed in patients with defective ERα or aromatase.[71]

Estrogen has been shown to play a role in neuroprotection in conditions like Parkinson's disease and Alzheimer's disease. In both conditions, postmenopausal hormone replacement therapy has been shown to be protective against disease onset, although E2 is not therapeutically useful after disease onset.[72][73] Estrogen can also protect against stroke and cardiovascular disease in premenopausal women.[74][75] E2 also induces rapid vasodilation via the release of nitric oxide and reduces the adhesion of inflammatory cells to atherosclerotic plaques.[76]

1.3.3 GPER as a GPCR for E2

It is universally appreciated that estrogen receptors, ERα and ERβ are primarily nuclear, and function as hormone-inducible transcription factors and induce estrogen-dependent gene transactivation.[77] However, the physical identity and nature of the receptor(s) that manifest pre-genomic estrogen have been a matter of healthy debate.[77] A number of recent reports has demonstrated the existence of membrane-associated

estrogen receptors that mimic the activity of the classical nuclear ERs.[78] These membrane-associated ERs trigger diverse cellular functions by activating both genomic (transcriptional) and non-genomic (rapid) signaling.[79] Non-genomic signaling is less well characterized and thought to involve a rapid mechanism and receptors located at the cell membrane. The rapid signaling events include pathways that involve a cross-talk between transmembrane growth factor receptors and G protein-coupled receptors.[79] The existence of G protein-mediated signaling by estrogen and localization of estrogen binding sites to membranes suggested the possibility of a 7-transmembrane G protein-coupled receptor family member, G protein-coupled receptor 30 (GPR30), being involved in certain aspects of estrogen function.[80][81] GPR30 has been implicated in mediating both rapid and transcriptional events in response to E2 under certain circumstances **(Fig. 1.10)**.[82] Several studies demonstrating estrogen pre-genomic signaling in GPR30-positive, ER-negative cells indicate that GPR30 can act as a stand-alone (independent) receptor.

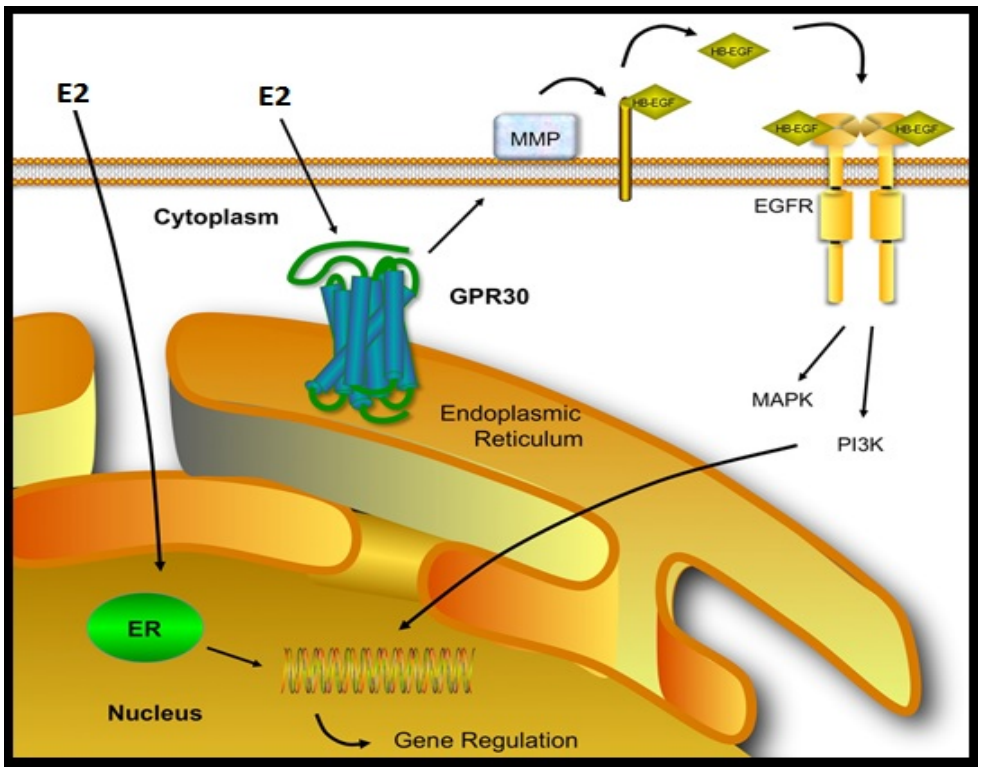


Figure 1.10 Estrogen signaling pathway.

GPR30 was first identified as an orphan 7TMR in different cells by multiple groups during 1996-1998.[83][84][85][86] Unlike ERs, which was isolated by a classical protein chemistry strategy well suited for a soluble receptor, GPR30 was discovered by molecular cloning approaches that have been widely successful for identifying a large number of GPCR. Since, its ligand was unknown at that time, it was named after its significant homology to GPCR superfamily. Based upon its structural homology to angiotensin II receptors, and other chemotactic peptide receptors, it was presumed that the ligand for GPR30 was a peptide.[87] Furthermore, this receptor was found to be associated with local ER expression in breast cancer cell lines.[85] Later in 2000, a study demonstrated that estrogen rapidly activate extracellular signal-regulated kinase (Erk)-1 and Erk-2 in two breast cancer cell lines, MCF-7 (ERα+/ERβ+/GPR30+) and SkBr3 (ERα-/ERβ-/GPR30+), with the cell line SkBr3 expressing non-ERs.[88] These findings demonstrated that estrogen might be a potential ligand for GPR30. This view was further confirmed by the observation that estrogen did not activate Erk-1/-2 in a breast cancer cell line MDA-MB-231 (ERα-/ERβ+/GPR30-) without GPR30 expression, whereas Erk-1/-2 was activated by estrogen after GPR30 transfection in the cells.[88] Therefore, GPR30 is necessary for the activation of Erk-1/-2 by E2. So far, GPR30 has been detected in numerous human tissues or cell lines, such as heart, uterus, placenta, prostate, subcutaneous adipose, visceral adipose, arteries and vessels.[89]

GPR30 is now widely recognized as a receptor for E2. This receptor is included in the official GPCR nomenclature and was designated G protein-coupled estrogen receptor-1 (GPER) by the *International Union of Basic and Clinical Pharmacology* in 2007.[90] As a GPCR, GPER has significantly different pharmacological properties and physiological roles than that of classical ERs. Rather than being a soluble receptor, GPER is a membrane receptor with seven transmembrane domains and is localized predominantly in the endoplasmic reticulum membrane.[91][92] Although, classic GPCRs are described as cell membrane receptor which binds its ligand at cell surface, it is becoming accepted that some GPCR may be functionally expressed at intracellular sites.[93] This is particularly true

of GPCRs with lipophilic or endogenously produced ligands. Estrogen is a cell permeable hormone, which suggests intracellular localization of GPER possible. However, the subcellular localization of GPER is still an object of controversy as this receptor is not truly intracellular and sometimes observed on the cell membrane.[94][95] GPER does not directly act as a transcription factor, but downstream signaling of GPER results in transcription of a variety of genes.[96][97] GPER also initiates a wide range of rapid signaling events, via adenylyl cyclase[98], transactivation of EGFR through the release of heparin-bound epidermal growth factor (HB-EGF)[88] and other pathways[99].

GPER has high affinity for E2, though not for other endogenous estrogens, such as E1 or E3.[94][100] 17α-estradiol and 17β-estradiol are two isomers of E2. Among the two, 17α-estradiol cannot bind GPER at all, neither can other steroid hormones, such as progesterone, testosterone, and glucocorticoid.[94] In addition, GPER can bind GPR30-specific compounds, G-1 (GPR30-selective agonist)[101]; G-15 (GPER-specific synthetic antagonist)[102]; and G-36 (GPER-specific synthetic antagonist)[103]. Selective estrogen receptor down regulators (SERDs), such as ICI 182,780[98][94]; and selective estrogen receptor modulators (SERMs), such as tamoxifen[98], reloxifene and 4-hydroxytamoxifen (OHT)[104][105][106] are also found to bind GPER and mimic E2 effects. Also, a variety of environmental estrogens, such as genistein, bisphenol A, zearalonone, nonylphenol, kepone, p,p'-DDT, 2,2',5',-PCB-4-OH and o,p'-DDE can bind GPER.[107] The affinities (Kd) of GPER to different ligands are shown in **Table 1.4.**

Table 1.4 Different ligands for GPER with tested affinity (Kd) values.

GPER Ligands	Affinity (Kd)
E2	2.7 nM [94], 6 nM[100]
E1	0.1% that of E2[94]
E3	0.1% that of E2[94]
G-1	11 nM[100]
G-15	20 nM[102]
ICI 182,780	~10% that of E2[94]

Tamoxifen	~10% that of E2[94]
Genistein	IC50 133 nM (~13% that of E2)[107]
Bisphenol A	2-3% that of E2[107]
Zearalonone	2-3% that of E2[107]
nonylphenol	2-3% that of E2[107]
kepone	0.25-1.3% that of E2[107]
p,p'-DDT	0.25-1.3% that of E2[107]
2,2',5',-PCB-4-OH	0.25-1.3% that of E2[107]
o,p'-DDE	0.25-1.3% that of E2[107]

1.3.4 GPER-mediated signaling

GPER is activated by E2, which also activates ERα and ERβ. As mentioned above, E2 initiates multiple intracellular signaling cascades. Although classical ERs have been demonstrated to be capable of mediating many of these responses, the signaling capabilities of GPER in response to estrogen have just begun to be described. GPER is capable of mediating both genomic and non-genomic responses induced by E2. Signaling pathways employed by GPER activation have not been fully elucidated yet. According to several published literatures, possible GPER-mediated signaling systems have been summarized in **Fig. 1.11**.[108]

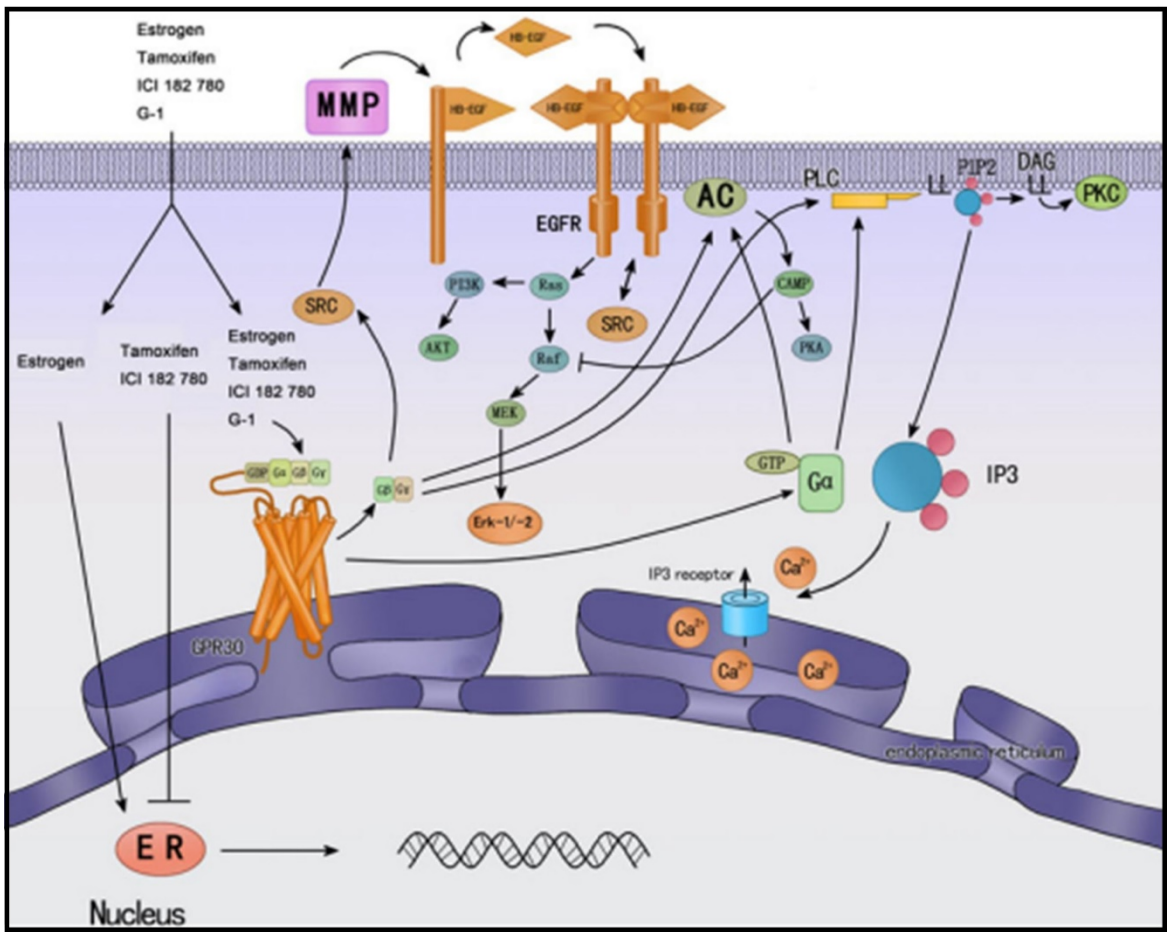


Figure 1.11 GPER-mediated genomic and non-genomic signaling.[108]

Some of the initial reports demonstrated that the GPER does indeed couple to G proteins in breast cancer cells.[94][88] In studies of the GPER-mediated signaling, much of the data have been obtained using the breast cancer cell lines, such as MCF-7 and SkBr3.[85][88] Briefly, E2 or other ligands with estrogenic properties may cross the cell membrane and bind to GPER, which is predominantly expressed on the membrane of endoplasmic reticulum, and activate heterotrimeric G proteins. The $G\alpha_s$ subunit in the activated trimeric G protein induces the activation of adenylyl cyclase, which results in the production of cAMP. On the other hand, the Gβγ subunits of the G protein activate Src tyrosine kinase, which binds to integrin α5β1 through an adaptor protein, Shc.[109] This complex then activates matrix metalloproteinase (MMP), and the activated MMP cleaves the pro-HB-EGF, releasing free HB-EGF into the extracellular space. The HB-EGF

transactivates the EGFR via an autocrine/paracrine mechanism, leading to multiple downstream events, including activation of PLC, PI3K, and mitogen-activated protein kinase (MAPK).[82] Activated PLC produces inositol triphosphate (IP3), which further binds to IP3 receptor and leads to intracellular calcium mobilization. The downstream signal of PI3K is Akt pathway. Main biological consequence of Akt activation is closely related to cancer cell growth; catalogued loosely into three aspects: survival, proliferation (increased cell number) and growth (increased cell size).[110] The activation of MAPK and PI3K further results in expression of transcription factors such as *c-fos*.[96] The activated EGFR can also induce extracellular-signal-regulated kinase (ERK) activation.[82][109] A recent study reported that the activation of ERK through GPER after E2 stimulation results in the secretion of connective tissue growth factor (CTGF) into the extracellular space, and that this secretion is involved in the proliferation of breast cancer cells.[106] In summary, the activation of GPER signaling cascades often leads to tumor promotion.

1.3.5 Clinical significance of GPER

Studies have demonstrated that GPER mediates rapid biological responses to estrogen in diverse normal, as well as transformed, cell types.[111][112] GPER gene expression has been spotted in at least four types of human cancer cell lines **(Table 1.5)**[89], including breast cancer[94][85][88][113][114][106], endometrial cancer[115][104][116], ovarian cancer[97][117][118] and thyroid cancer[105]. In human breast cancer, decreased GPER expression is observed on both mRNA[119] and protein levels[113] when compared to healthy tissues, and its expression level is positively correlated with ERα[119]. Whereas in human endometrial cancer, GPER expression is up-regulated on both mRNA and protein levels when compared to the healthy tissues.[115] A number of experimental evidence accumulates every year to prove that GPER is strongly associated with cancer proliferation, migration, invasion, metastasis, differentiation, prognosis, and drug resistance.

Table 1.5 Estrogen receptors expression in human cancer cell lines.

Human cancer cell lines	**ERα**	**ERβ**	**GPER**
Breast cancer cell lines			
MCF-7	+	+	+
SkBr3	-	-	+
MDA-MB-231	-	+	-
T47D	+	+	+
MDA-MB-468	-	+	+
Endometrial cancer cell lines			
KLE	-	-	+
RL95-2	+	+	+
Ishikawa	+	+	+
HEC-1A	-	+	+
Ovarian cancer cell line			
BG-1	+	+	+
Thyroid cancer cell line			
WRO	+	-	+

Since breast cancer cells proliferate in response to E2, E2 antagonists have been used for breast cancer therapy. However, relapse and metastasis have frequently been observed during therapy involving E2 antagonists, suggesting the possibility that a signal pathway in response to E2 other than the ERs may be present in breast cancer cells. It has been noted that GPER acts as a receptor in an alternative pathway of E2 activation.[120] Endocrine therapy is often the treatment of choice for breast cancer, including in advanced cases as long as they remain estrogen-dependent.[55] About two-thirds of all breast carcinomas express ERα, and yet, tamoxifen is used to treat ERα-positive tumors.[111] Tamoxifen is an antagonist of ERα in breast tissue via its active metabolite, 4-hydroxytamoxifen. But, this antagonist acts as GPER agonist, which could significantly

influence the outcome of treatment.[121] 25% of all ERα-positive breast cancer patients do not respond to tamoxifen therapy, instead, they develop hormone-refractory cancer lesions, which are characterized by their rapid growth and invasiveness.[111] Considering the expression and signaling profile of GPER in breast cancer cells, it is clear that this receptor constitutes a target for anti-carcinogenic drug design and emphasizes the importance of evaluating the level of GPER expression in an ERα-positive cancer before using tamoxifen in endocrine therapy.

Moreover, women treated with tamoxifen against breast cancer display an increased incidence of endometrial cancer.[122] In endometrial cancer, GPER is considered as a novel indicator of poor survival, as its high level expression is correlated with a more deteriorated cancer outcome.[123] Here again, GPER signaling is found to be involved in the development of endometrial carcinoma by promoting proliferation and enhancing invasion.[115] Another recent study, proved a similar role of GPER in ovarian carcinoma.[124] GPER was also involved in the stimulatory effects elicited by estrogen and ER antagonists in cancer-associated fibroblasts.[125] Together, these evidences support the hypothesis that GPER represents an estrogen-responsive receptor that is overexpressed and functionally relevant in high-risk breast, endometrial, and ovarian carcinomas.[111] But, the mechanism underlying the effect of GPER in estrogen-related cancer therapy is still unclear, and yet, there is no specific drug for blocking GPER action. It would be clearly important to clarify whether GPER is essential for certain cancer development and whether GPER is responsible for anti-estrogen therapy and chemotherapy resistance in these cancers.

Apart from cancer, reports have been published on other possible physiological roles of GPER in the nervous system as well as in reproduction, metabolism, bone, and in the cardiovascular and immune systems.[126] GPER has shown to play a role in insulin secretion[127], vascular and myocardial function[128], renal disease and proteinuria [129][130].

1.4 Rationale for project

Science is not about simply accepting or denying findings of others, but about understanding, integrating, and communicating findings to advance current knowledge to the greatest possible degree. In the field of GPER research, a number of questions particularly with regard to rapid and chronic actions of GPER activators, inhibitors, partial agonists, or genetic GPER deficiency as well as potential roles of GPER in disease are still open.[131]

Despite showing a broad clinical significance, GPER holds several confusions and challenges that remains unclear and are to be addressed. Though GPER is recognized as an ER, mediating non-genomic effects induced by E2, some groups raise the most controversial question concerning whether GPER is an ER at all.[132][133][134] The ultimate proof might have to come from a structural analysis of E2-bound GPER, but the first x-ray structures of any ligand-bound 7TM-GPCRs have only been solved very recently.[135] As a GPCR, GPER has very different sequence and structure than the classical ERs, and also, membrane proteins are difficult to crystallize in order to generate x-ray structures. The binding pocket for estrogen and estrogen analogues in GPER is not specifically known.

Other unresolved puzzles include: What is the physiologic function of GPER in normal tissues as well as disease states? What are the overlapping and distinct functions of GPER with respect to ERα and ERβ? Does it initiate mostly independent responses? Is GPER expressed in the same or different cells and tissues compared to ERα and ERβ? How does it go to the membrane and which membrane after all? What are its structural and functional relationships to its neighbors in the GPCR family? Will drugs that selectively target GPER versus ERα and ERβ and *vice versa* be superior to drugs currently available for treating cancer, cardiovascular, neurological, renal and immune disorders. Sex differences of ER and GPER expression and their genomic and non-genomic functions as well as post-translational and epigenetic modifications such as methylation of ER DNA which may significantly affect its function also needs to be addressed.

In order to address some of the afore-mentioned challenges and to further advance our understanding on GPER's mechanism of action, mass spectrometry-based proteomic approach was employed in the work presented in this thesis. This research was part of the project "Interaction of estrogen and estrogen receptors by MALDI-TOF/TOF" of the Initial Training Network - Chemical Bioanalysis (ITN-CHEBANA). The ITN is part of the Marie Curie Actions funded by the European commission. At the start of this research project, almost nothing was reported on the isolation and characterization of GPER by mass spectrometry, which was evident from the lack of published literature. By that time, most of the laboratories used molecular biological techniques for GPER research. Immunofluorescent and Western blot (WB) analyses were widely used to study expressed GPER. But today, we are first to report GPER identification by MALDI-TOF/TOF tandem mass spectrometry. For this thesis, we worked towards the development of a proteomics workflow for GPER investigation.

In **Chapter 2**, we discuss a bit about what proteomics is and how it could help in solving some of the puzzles associated with GPER, and move on to modern mass spectrometry-based tools to unravel complex physiological pathways.

Here, we demonstrate the hurdles in isolating membrane-bound GPER from crude lysate and propose a gel-free method using home-made hydroxyapatite (HTP) spin column to enrich and isolate GPER. During the study, we tested different proteolytic digestion conditions and made use of different proteases to pick the best one for GPER identification and characterization. The efficiency of the developed method for GPER isolation was verified by WB analysis with great reproducibility. This approach has proven to be successful as we were able to isolate and identify GPCRs including GPER by peptide mass fingerprinting (PMF).

In **Chapter 3**, we discuss on protein post-translational modifications (PTMs), and PTMs that are potentially found in GPER. For glycosylation study, we used the HTP method that we proposed in Chapter 2 for GPER isolation and carried out glycopeptide derivatization with dansyl chloride (DNS-Cl), followed by MS and MS/MS analyses. GPER

deglycosylation experiments were also performed to some extent and discussed in this chapter. From the obtained experimental data, we were able to validate 2 glycosylation sites and the predicted glycan structures, manually.

In **Chapter 4**, we introduce affinity capture-release strategy for GPER purification. This study revealed the possibility for developing synthetic antibodies for GPER. Here, we discuss about the design and synthesis of polymer nanoparticles to capture GPER with high affinity and selectivity among a mixture of proteins that are expressed in cancer cells. We evaluated the NPs-peptide/protein binding using HPLC and WB analyses. The initial results were interesting and we anticipate to use this strategy as a sample purification step prior to MS-based proteomic analysis.

Chapter 5 will bring the discussions together, summarizing the milestones achieved during the project. This chapter will also point some future directions that emerge from the results of this thesis, to achieve the long-term goal of studying molecular interactions of GPER with other receptors and ligands by tandem mass spectrometry.

CHAPTER 2

MASS SPECTROMETRY-BASED GPCR PROTEOMICS: Isolation and identification of GPER by Peptide Mass Fingerprinting

In this rapidly developing scientific era, proteomics represents a major promise in proteome-wide studies. The term proteome was first used by Marc Wilkins, in 1996[136], for the entire set of proteins expressed by a cell, tissue or organism at a given time point. Proteomics is the study of the proteome, and involves the large scale study of proteins, particularly their structures, biological functions and interactions with other proteins. According to the Oxford Dictionary of Biochemistry and Molecular Biology, proteome and proteomics are defined as "the complete expression profile of the proteins of an organism" and "the study of the proteome by the analysis of protein structure and composition", respectively.[137] The idea of identifying and analyzing all proteins encoded by a genome was proposed in mid 1990s[136] and the first proteome-scale analysis of a eukaryote was presented by Shevchenko and colleagues in 1996[138]. Since then, proteomics took off briskly and reached great heights. The challenge of proteomics is essentially based on archetypal 'Analytical Chemistry' strategies, such as separation, purification, and qualitative (identification) and quantitative analysis of proteinaceous analytes. Furthermore mass spectrometry (MS), one of the most powerful tool in modern analytical chemistry, is the core methodology in proteomics. Today, it is believed that the emergence of novel proteomics technologies will help researchers to better understand the mechanism underlying complex pathways including the ones that involve GPCRs. However, regardless of having many advances in technology and methodology, proteomics is still far from having reached the stage of productivity and utility that is necessary for it to be crucial to biological and biomedical research in the post-genome era. Areas requiring a prompt attention are related to sample preparation, separation technologies, quantitative methodologies and full exploitation of modern mass spectrometers.

Despite showing a broad pharmaceutic importance, the structural information available on different GPCRs' ligand-receptor complex with respect to molecular interaction is very limited and is mainly due to inadequate protein purification and hydrophobic nature of GPCRs. In addition, like many other membrane proteins, the

expression levels of GPCRs are very low in native systems. So far, structure-based drug design for GPCRs is mainly derived from computational structure prediction based on homology modeling combined with mutagenesis experiments. In the last decade, MS has proven to be the tool of interest for structural elucidation and for investigating protein-protein and ligand-receptor interactions. The use of selective ligands to covalently bind a receptor of interest, followed by MS analysis, could in principle precisely reveal amino acid residues that are part of the respective ligand-receptor binding sites. Moreover, MS techniques are well suited for the analysis of post-translational modifications but not frequently applied in the analysis of GPCRs. In the presented work, we intended to address some of the challenges related to GPER (refer section 1.5) by focusing on the analytical core of proteomics through the integration of an innovative separation and enrichment method for MS-based GPCR proteomics.

2.1 Mass spectrometry-based proteomics

The soul of any proteomics approaches is the ability to identify proteins in complex samples. Proteins can be identified based on structural and compositional features. The two main approaches to identify a protein are immunological approach and sequence-based approach. Immunological approaches identify a protein using specific antibodies.[139] The field of cellular biochemistry and molecular biology was dominated by antibody-based techniques such as WB analysis as well as protein array and enzyme-linked immunosorbent assay (ELISA) technologies, which requires certain prior knowledge about the proteins to be examined and are mostly restricted to a certain number of proteins to be analyzed. In sequence-based approach, proteins are identified based on the determination of their amino acid composition. Technological, methodological and computational limitations currently do not allow the determination of the sequence of a protein directly. Instead the sequence of a subset of its peptides is determined which is then used for protein identification.

Until the early 1990s, Edman degradation was the method of choice to directly determine the amino acid sequence of a peptide. This method, introduced by Per Edman

in 1950, is based on cyclic degradation of proteins with phenylisothiocyanate.[140] The detached amino acids are subsequently identified by ultra violet (UV) absorbance spectroscopy. Despite the automation of the process in the late sixties[141][142][143], with a cycle time of one hour per amino acid, determination by Edman degradation remains a slow and inefficient process. Later in mid 1990s, Edman degradation was replaced by mass spectrometry-based approaches to determine the amino acid sequences of peptides. In MS, the elemental composition of a sample is determined based on the mass-to-charge (*m/z*) ratio of ionized molecules or molecule fragments generated from the sample. The *m/z* ratio is measured based on the motion of the ions as they pass through an electromagnetic field.[144] The foundations of modern mass spectrometry were laid by Eugen Gold-stein and Wilhelm Wien in the late 19th century. The first fully functional mass spectrometers were built by Arthur Demster and Francis Aston in 1918 and 1919.[145] Large-scale and global proteomic studies of cellular biochemistry are enabled by the application of mass spectrometry, which allows the unbiased identification and characterization of hundreds to thousands of cellular proteins within one study.

Mass spectrometry provides highly sensitive and accurate mass determinations of biomolecules. It is a versatile tool in proteomics research with a wide range of applications, including protein identification, quality control of recombinant proteins and studies of PTMs. [146] MS-based proteomics approaches encompass several steps, including extraction, fractionation and/or enrichment of protein and/or proteolytically derived peptide mixtures; peptide separation and ionization; measurement and collection of mass spectra; and data analysis. The most prominent use of MS is in combination with two-dimensional gel electrophoresis (2-DE), in which masses of enzymatic digests of proteins are determined and peptide mass fingerprints are obtained for comparison with entries in available protein databases **(Fig. 2.1)**.[147][148][149] Various combinations of ionization techniques, mass analyzers and detectors may be used.

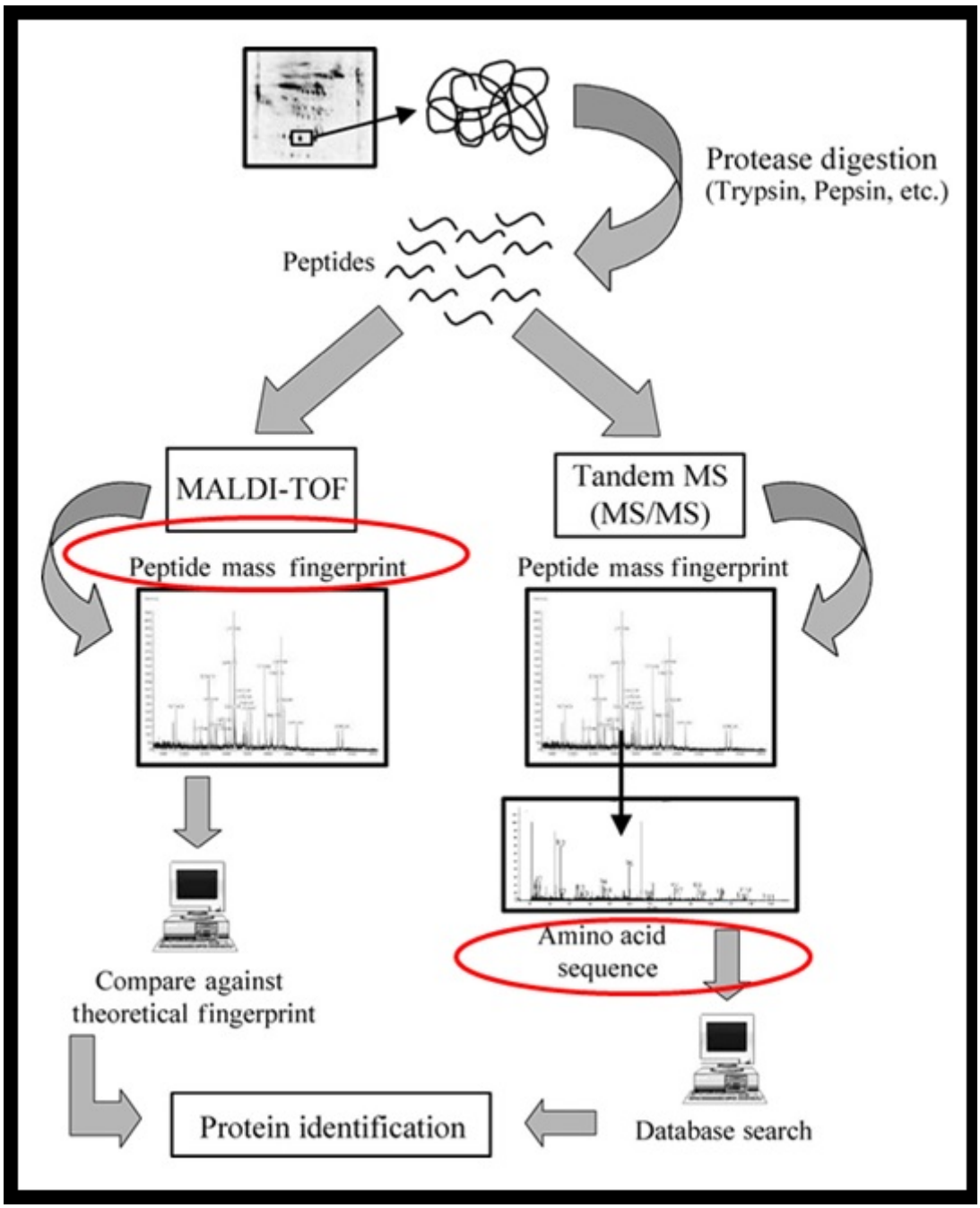


Figure 2.1 Classical proteomics workflow.

The ionization sources used for protein research are electrospray ionization (ESI)[150] and matrix-assisted laser desorption/ ionization (MALDI)[151] systems, since their relatively gentle modes of action permit analysis of large polypeptides **(Fig. 2.2)**. Both of these ionization types are classified as 'soft' techniques because there is little in-source

fragmentation of the ionized species, such that the mass of the intact molecular ion can be measured.[152][153] These ionization systems are combined with different mass analyzers in order to separate the ions formed, depending on the application. The most frequently used designs in proteomics are hybrid instruments that contain a combination of mass analyzers, such as: time of flight (TOF)[154][155], quadrupole (Q)[156], ion trap (LIT)[156], Fourier transform ion cyclotron resonance (FT-ICR)[157], and Orbitrap[158]. Tandem mass spectrometers can have more than one mass analyzer (for example, q-TOF and TOF-TOF) to take advantage of the strength of each.[159]

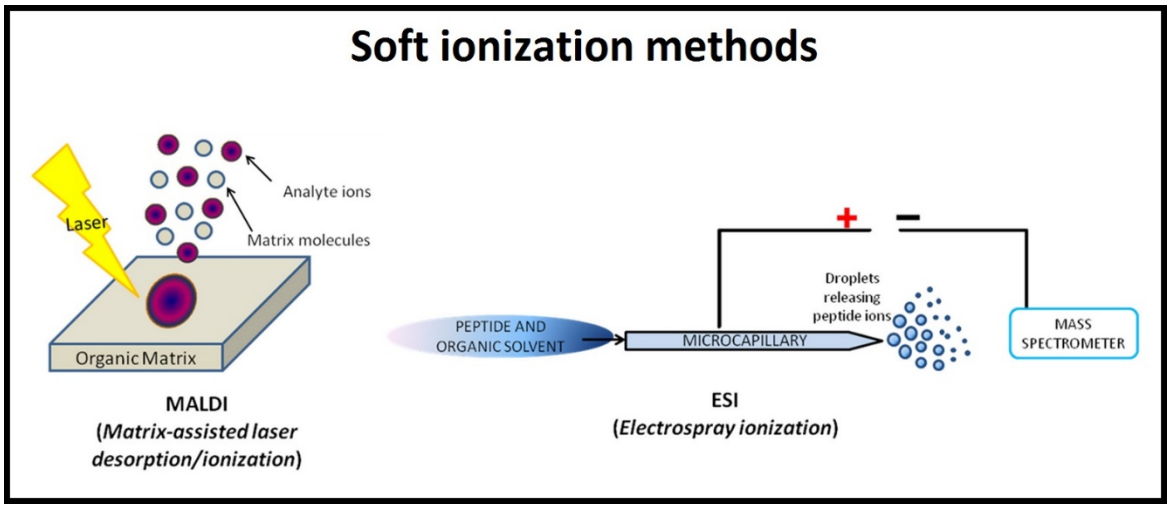


Figure 2.2 Schematic representation of MALDI and ESI.

MALDI is usually coupled to TOF analyzer that measure the mass of intact peptides, however, several other analyzers can also be coupled with MALD depending on the type of analysis. In the presented work, MALDI-TOF/TOF tandem mass spectrometer **(Fig. 2.3)** is used for GPER identification and characterization. MALDI sublimates and ionizes the analytes out of a dry, crystalline matrix via laser pulses. Generally, the matrix consists of crystallized molecules. Matrix should be of low molecular weight to allow facile vaporization. It is often acidic and acts as a proton source to encourage ionization. Matrix should have strong optical absorption either in ultra violet (UV) or in infra-red (IR) range and can be functionalized with any one of the following chemical agent: 3,5-dimethoxy-

4-hydroxyxinnamic acid (sinapinic acid), α-cyano-4-hydroxycinnamic acid (α-cyano or α-matrix), or 2,5-dihydroxybenzoic acid (DHB) **(Fig. 2.4)**.

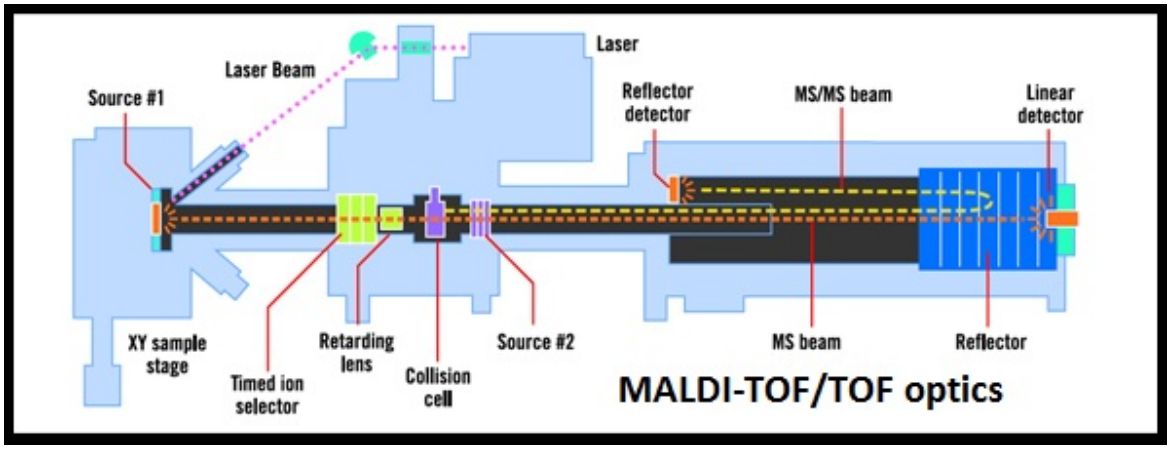


Figure 2.3 Schematic representation of MALDI-TOF/TOF optics.

The ionization is triggered by a laser beam. UV laser such as nitrogen laser (337 nm), frequency-tripled or -quadrupled neodymium-doped yttrium aluminium garnet (Nd:YAG) laser are used as ionization source. The advantages of MALDI-TOF MS are speed, sensitivity, ease of use and accuracy of the molecular weights obtained by the TOF detector.[160][161][162]

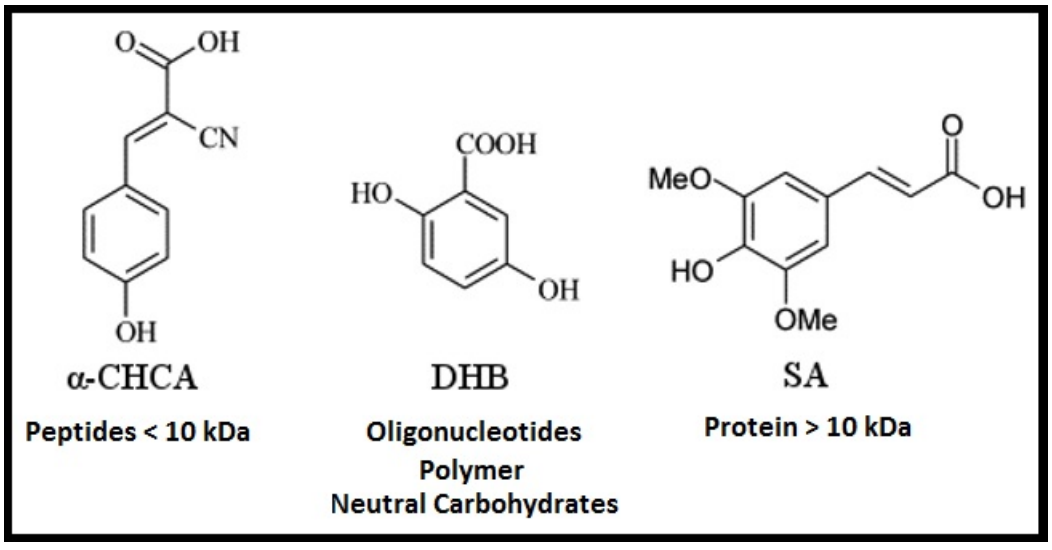


Figure 2.4 Chemical structures of different MALDI matrices.

2.2 Classical proteomics unfit for GPER analysis

GPER is a 42 kDa integral membrane protein that may contribute to normal estrogen physiology as well as pathophysiology. The GPER gene, well conserved in different species, is mapped to chromosome 7p22.3 in human.[114] There are four alternate transcriptional splicing variants with 2 or 3 exons encoding the same protein which is comprised of 375 amino acids.[163] GPER is highly conserved in mammals with 87% sequence identity between human and mouse.

Cell culture and cell lysis: Proteins prepared from different cancer cell lines were used in the presented research work. BG-1 (ovarian cancer cells), Ishikawa (endometrial cancer cells), MCF-7 (breast cancer cells) were maintained in DMEM without phenol-red supplemented with 10% fetal bovine serum (FBS). Whereas, SkBr3 (breast cancer cells) were maintained in RPMI 1640 without phenol-red supplemented with 10% FBS. HEK293 cells were overexpressed with hGPER gene and maintained in DMEM high glucose supplemented with 10% FBS. All cells were grown in 10 cm petri dishes, and then lysed in 500 µL lysis buffer containing 50 mM HEPES (pH 7.5), 150 mM NaCl, 1.5 mM $MgCl_2$, 1 mM EGTA, 10% glycerol, 1% Triton X-100, 1% SDS, and a mixture of protease inhibitors such as 1 mM aprotinin, 20 mM PMSF and 200 mM sodium orthovanadate. The protein concentration was determined using Bradford protein assay. The cell lysates were divided into small aliquots and stored at -20 °C for further use.

Protein depletion: The proteomics workflow was initiated by depleting high abundant proteins such as albumin & immunoglobulin G (IgG) using ProteoPrep Blue Albumin and IgG depletion medium from Sigma Aldrich. The buffers supplied by the manufacturer contain surfactants & salts that can cause signal suppression, therefore an alternative protocol was used. The cartridge was conditioned with 200 µl of 50 mM NH_4HCO_3 pH 8, three times giving 10 min incubation followed by centrifugation. 200 µl of total protein lysate (TPL) was loaded onto the column and incubated for 10 min at room temperature. After centrifugation at 3000 rpm for 1 min, the flow-through was loaded again and collected. The column was washed two times with 200 µl of 50 mM NH_4HCO_3 and the

relative flow-through fractions were collected and pooled. The bound proteins were eluted with 200 µl of $(NH_4)_2CO_3$ pH 10, two times giving 10 min incubation followed by centrifugation. The collected eluate fractions were pooled and subjected to chemical fractionation, where the depleted proteins were partitioned into soluble, acidic, neutral and alkaline proteins by chemical treatment.

Chemical fractionation: 100 µl of the depleted protein solution was precipitated with 800 µl $CHCl_3/CH_3OH$ (1:3, v/v). The pellet was partitioned with 200 µl of CH_3CN/NH_4HCO_3 (60:40, v/v), 200 µl of H_2O and 200 µl of TFA 0.1%/CH_3CN (90:10, v/v) systematically at room temperature under magnetic stirring. Each step was followed by centrifugation at 12000 rpm for 5 min. Supernatant samples were collected at each step after centrifugation.

Electrophoresis: The chemically fractionated samples were then subjected to sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) where the proteins were separated according to their electrophoretic mobility. 12.5% poly acrylamide gel was used to resolve the proteins. After electrophoresis, the gel was stained with a solution of Coomassie Brilliant Blue R 0.1% w/v in $H_2O/CH_3OH/CH_3COOH$ (9:9:2) and destained with destain solution made with 40% CH_3OH, 10% CH_3COOH to 1 L H_2O.

The bands appeared at around 42kDa were excised and processed for proteolytic digestion. On-gel digestion was performed overnight with trypsin (pH 8) at 37˚C. The tryptic peptides were concentrated and purified by eluting with C18 ZipTip, prior to MS analysis. α-CHCA (0.3% in TFA) was used as the matrix of choice in the presented work. 1 µL portion of sample-matrix solution was spotted on a MALDI target plate, dried at room temperature and directly analyzed MALDI MS. MS analyses were carried out on AB SCIEX TOF/TOF™ 5800 System equipped with a 1 kHz variable rate laser. MS data were acquired in reflectron positive-ion mode, at a laser repetition rate of 400 Hz with 4000 laser shots/spectrum (100 laser shots/sub-spectrum) with a mass accuracy of 50 ppm. Acquired data set was evaluated using Mascot search engine (www.matrixscience.com). Database searches were performed against Swiss-Prot & NCBI, with taxonomy restricted to *Homo*

sapiens and trypsin enzyme cleavage specificity with an initial mass tolerance of 50 ppm. Over all, this experiment was a letdown as the database search failed to fish out GPER.

In the successive experiments, protein depletion and chemical fractionation were replaced by ***solid-phase extraction (SPE)*** step. The TPL was fractionated using a reversed-phase C18 cartridge (6 ml, 1 g). *Conditionation*: 2 ml of CH_3CN/TFA 0.1% (50:50, v/v) followed by 2 ml of TFA 0.1% were drawn slowly through the column, avoiding column from drying. *Sample Adsorption*: 4 ml of sample (200 µl of total protein extract + 3.8 ml of TFA 0.1 %) was slowly drawn through the column. *Washing*: 3 ml of TFA 0.1% was drawn completely through the column. *Elution*: The analytes were gradient eluted by loading 1 ml of CH_3CN/TFA 0.1% subsequently, increasing the concentration of organic phase (from 10 to 100%). Samples were collected in 500 µl at every stage throughout the whole process followed by concentration in speed vac. The concentrated samples were then subjected to SDS-PAGE and MALDI MS analysis, as described in previous paragraph. Additionally, we also tried in solution digestion of SPE fractions with trypsin followed by sample clean-up and MALDI MS analysis. The database searches revealed no significant data in either case. The experiments were repeated multiple times to rectify unknown errors, if occurred. But the outcome was same every time.

It was learnt that classical proteomics was not suitable for GPER analysis by MALDI MS. Limitations with this approach include difficulties in separating very hydrophobic proteins like cell surface receptors, automation problems and the need for time-consuming optimization, depending on the type of proteins to be studied.[164] To apply MS techniques to GPER analysis, improved methods for membrane protein production and proteomics analysis are needed to overcome the major hurdles.

2.3 Method development for GPER isolation

As classical MS-based proteomics involves protein separation by gel-electrophoresis followed by proteolytic digestion and MS analysis **(Fig. 2.1)**, this approach may not be suitable for integral membrane proteins, like GPER. The extraction of

membrane proteins after on-gel digestion still remains a big challenge. To overcome the hurdles in characterizing GPER by MALDI MS, it is necessary to develop a novel gel-free proteomics method to isolate and enrich the protein of interest from crude lysate. Lack of published literature on GPER isolation motivated us to work towards the development of a proteomics work flow for investigating GPER by MS.

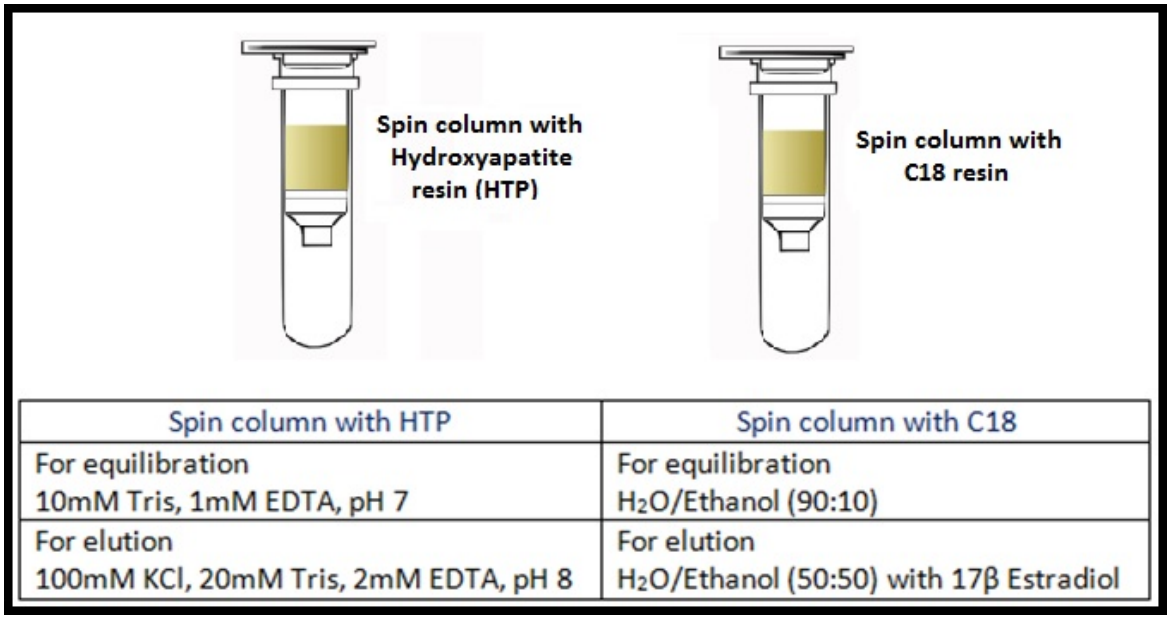
After days of literature hunt, we ended up in an interesting study dated back to 1980, where hydroxyapatite was used in partial purification and characterization of estrogen receptors.[165] Even before that in 1978, hydroxyapatite dissociation method was reported to be useful in the preparation of chromosomal proteins for characterization studies.[166] HTP is naturally occurring metal salt with the formula $Ca_5(PO_4)_3OH$. HTP crystallizes as a hexagonal close-packed structure orienting calcium ions at the surface in a triangle a few angstroms apart.[167] It is present in human bone and tooth enamel and forms biologically relevant interactions with proteins and phosphoproteins.[168] Over the years, HTP has enabled protein scientists to separate and purify proteins.[169][170] Recent studies demonstrated the use of HTP for single-step phosphopeptide enrichment from complex biological samples prior to MALDI analysis.[171][172][173] Also, the use of HTP column has been reported for the purification of human β_2-adrenergic receptor.[174]

From the existing literatures, it is very important to note that the hydroxyapatite was used to purify only classical ERs[165], not the membrane-ER, GPER, which was not even in existence by that time. So, we decided to perform a case study on SkBr3 human breast cancer cells and used home-made hydroxyapatite spin column **(Fig. 2.5)** for GPER enrichment and isolation from TPL. In comparison, we also used a column stacked with C18 resin **(Fig. 2.5)** to monitor the performance.

Spin column with hydroxyapatite (HTP) – 200 µl of TPL was diluted with 25 µl of equilibration buffer (10 mM TRIS, 1 mM EDTA, pH 7) and loaded onto HTP (100 g) spin column. After 30 min incubation at room temperature, the column was centrifuged at 8000 rpm for 2 min. The filtrate was collected and reloaded for second time to achieve maximum enrichment. The bound proteins were eluted with 200 µl of elution buffer (100

mM KCl, 40 mM TRIS, 2 mM EDTA, pH 8), three times giving 30 min incubation followed by centrifugation at 8000 rpm for 2 min. The volume of eluate was concentrated to 100 µl in speed vac.

Spin column with C18 – 200 µl of total protein lysate was diluted with 25 µl of equilibration solution (H_2O:C_2H_6O, 90:10) and loaded onto C18 (100 g) spin column. After 1 hr incubation at room temperature, the column was centrifuged at 8000 rpm for 2 min. The filtrate was collected and reloaded for second time to achieve maximum enrichment. The bound proteins were eluted with 200 µl of elution solution containing 1 µM 17β-estradiol in H_2O:C_2H_6O, 50:50 giving 5 hr incubation and centrifugation, then followed by two more times with 200 µl of H_2O:C_2H_6O, 50:50 giving 30 min incubation and centrifugation. The volume was concentrated to 100 µl in speed vac.



Spin column with HTP	Spin column with C18
For equilibration 10mM Tris, 1mM EDTA, pH 7	For equilibration H_2O/Ethanol (90:10)
For elution 100mM KCl, 20mM Tris, 2mM EDTA, pH 8	For elution H_2O/Ethanol (50:50) with 17β Estradiol

Figure 2.5 HTP and C18 spin column model.

The concentrated samples (C18 waste, C18 eluate, HTP waste and HTP eluate) from both the columns were subjected to SDS-PAGE for protein separation. The results were shown in **Fig. 2.6(a)**. On comparing the results, it was clear that C18 lost the battle against HTP as most of the proteins got flushed out in the waste fraction, while HTP

retained most of the proteins leaving no trace in the waste. Surprisingly, the eluate fractions from both the columns showed trace bands in the range of 42 kDa indicating the possibility for GPERs' presence. Later, the concentrated eluate from both C18 and HTP columns were proteolytically digested in solution using trypsin and analyzed by MALDI MS. The obtained spectra were submitted to MASCOT search engine. Unfortunately, no hits were found for GPER rather we noticed heavy traces of β-actin topping the list every time. It is crucial to note that the molecular weight of β-actin (42 kDa) is similar to that of GPER. β-actins are generally expressed in all eukaryotic cell types and they are involved in cell migration, structure and integrity. Simultaneously, efforts were made to detect the presence of GPER by protein immunoblotting technique.

***Western blot analysis*:** the C18 and HTP eluate fractions were resolved on a 10% polyacrylamide gel. After electrophoresis, the resolved proteins were transferred to a nitrocellulose membrane and probed overnight against GPR30 (N15) antibody (Santa Cruz Biotechnology) at 4°C. The developed blot was registered on a photographic film. The result turned out to be unexpected. HTP column emerged as the winner by showing a very clear and unique band at 42 kDa range, and confirmed the presence of GPER in samples processed through HTP column. Though 17β-estradiol was used in C18 elution solution, there was no significant outcome and this may be due to the poor protein retention as is evident from **Fig. 2.6(b)**.

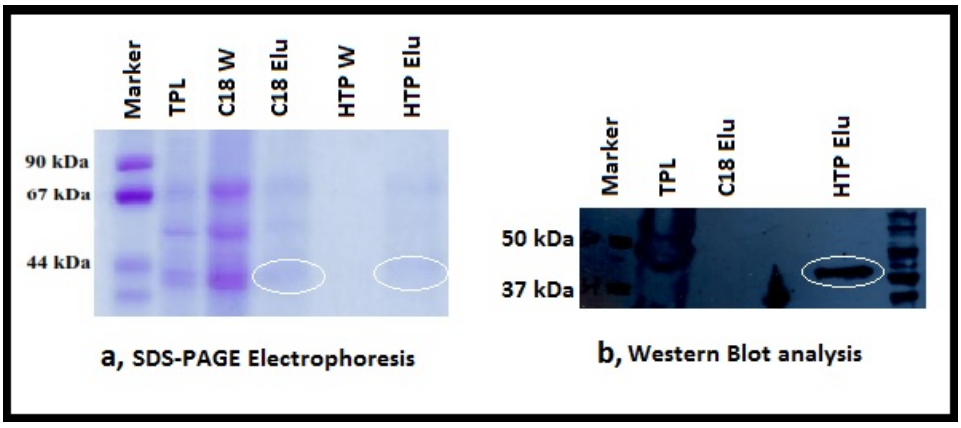


Figure 2.6 GPER isolation from SkBr3 lysate using HTP spin column.

Next, to check the consistency of GPER enrichment using HTP spin column, we repeated the experiment again. This time, we used TPL obtained from different cancer cells, such as BG-1, Ishikawa, SkBr3 and MCF-7. We processed the TPL through HTP spin column as descried earlier in this section and collected the eluate fractions. The eluate fractions were subjected to SDS-PAGE followed by WB analysis. The SDS-PAGE and WB results were shown in **Fig. 2.7(a)** and **Fig. 2.7(b)**, respectively. As expected, the immunoblot confirmed the presence of GPER when probed against GPR30 (N15) antibody.

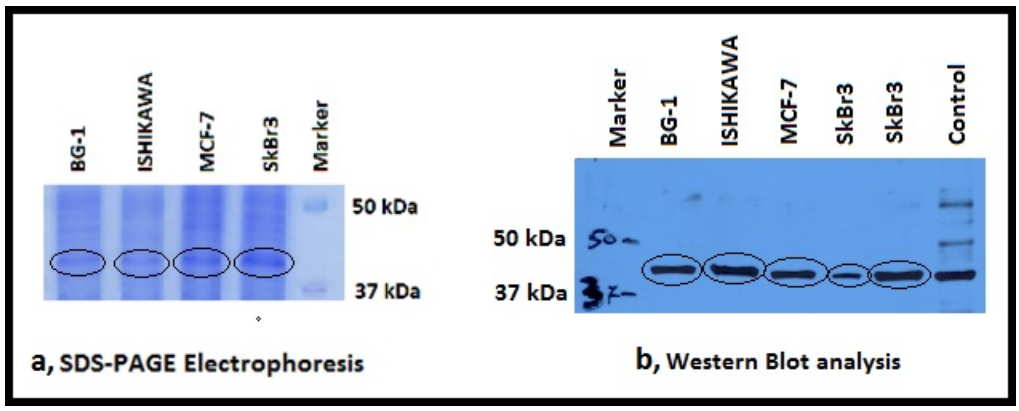


Figure 2.7 Consistency of HTP enrichment method.

Here, we made one successful step towards the development of proteomics workflow for GPER investigation. The proposed GPER enrichment method using HTP spin column was simple, robust and cost-effective, and mainly found useful to isolate GPER from complex protein mixture.

2.4 Identification of GPER by peptide mass fingerprinting

The next big step was to identify GPER by peptide mass fingerprinting. In order to achieve this, the known problem with β-actin needs to be rectified first and foremost. We postulated that the use of membrane filters would serve the purpose. Our idea was to trap the membrane protein, GPER on the filter membrane and wash out all other unwanted proteins including β-actin. Cellulose acetate (CA) membrane spin filters (0.22

µm) were used in the presented work **(Fig. 2.8)**. CA membranes get their name not from their structure, but from their materials of fabrication.[175] CA membranes have a very low binding affinity for most macromolecules and are especially recommended for applications requiring low protein binding.[176] They are naturally hydrophobic and have small amounts (less than 1%) of non-toxic wetting agents added during manufacture to ensure proper wetting of the membrane. If desired, these agents can be easily removed prior to use by filtering a small amount of warm purified water through the membrane or filter unit.

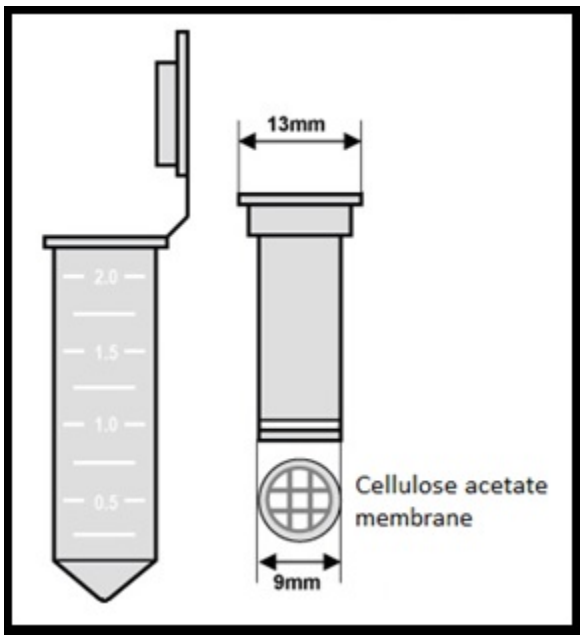


Figure 2.8 Cellulose acetate spin filter model.

Once the TPL was processed through HTP spin column, the collected eluate was filtered using CA membrane spin filter, centrifuging at 6000 rpm for 2 min. The flow-through fraction was re-filtered for the second time in the same spin filter, centrifuging at 6000 rpm for 2 min. Followed by this, the CA membrane was washed with d.H_2O and 50 mM NH_4HCO_3, three times each consecutively, centrifuging at 6000 rpm for 2 min each time. Then, the filter unit was cut-open to remove the CA membrane. The CA membrane, thought to hold GPERs was then immersed in a solution having protease enzyme for proteolysis. The proteolytic digestion was performed on CA membrane **(Fig. 2.9)** with

pepsin (pH 2) at 37°C for 2-3 hr. In parallel, we also proteolytically digested small aliquots of the flow-through, H_2O and NH_4HCO_3 fractions in solution with pepsin to compare the results. MALDI MS data were acquired from the peptic digests and evaluated using Mascot search engine. Database searches were performed against Swiss-Prot & NCBI, with taxonomy restricted to *Homo sapiens* and pepsin enzyme cleavage specificity with an initial mass tolerance of 50 ppm. The search results turned out to be a game changer and gave us new hope. The use of CA membrane filter before proteolytic digestion removed the much disturbing β-actin and other unwanted soluble proteins. The membrane retained mostly all the hydrophobic proteins that couldn't pass through it. Though the search didn't find the presence of GPER, it picked up traces of other GPCRs including GPR1, GPR6, GPR40, GPR112, GPR124.

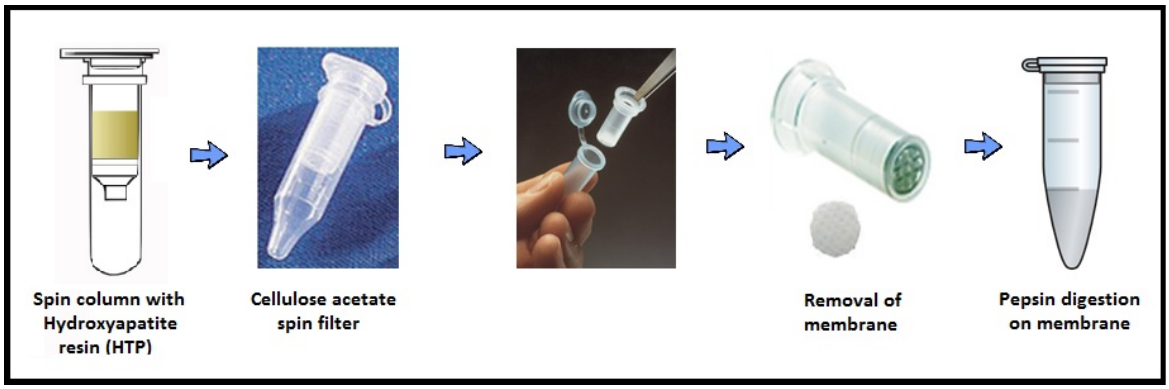


Figure 2.9 On-membrane digestion Schematic representation.

For example, in our very first experiment after integrating both HTP enrichment and on-membrane digestion steps to our workflow, we encountered GPR1 with 72% sequence coverage **(Fig. 2.10)**, topping the search list. As our research focus is mainly on identifying GPER, we repeated the experiments several time and continued our journey towards method development. At one point, we tried changing the protease from pepsin to α-chymotrypsin for on-membrane digestion.

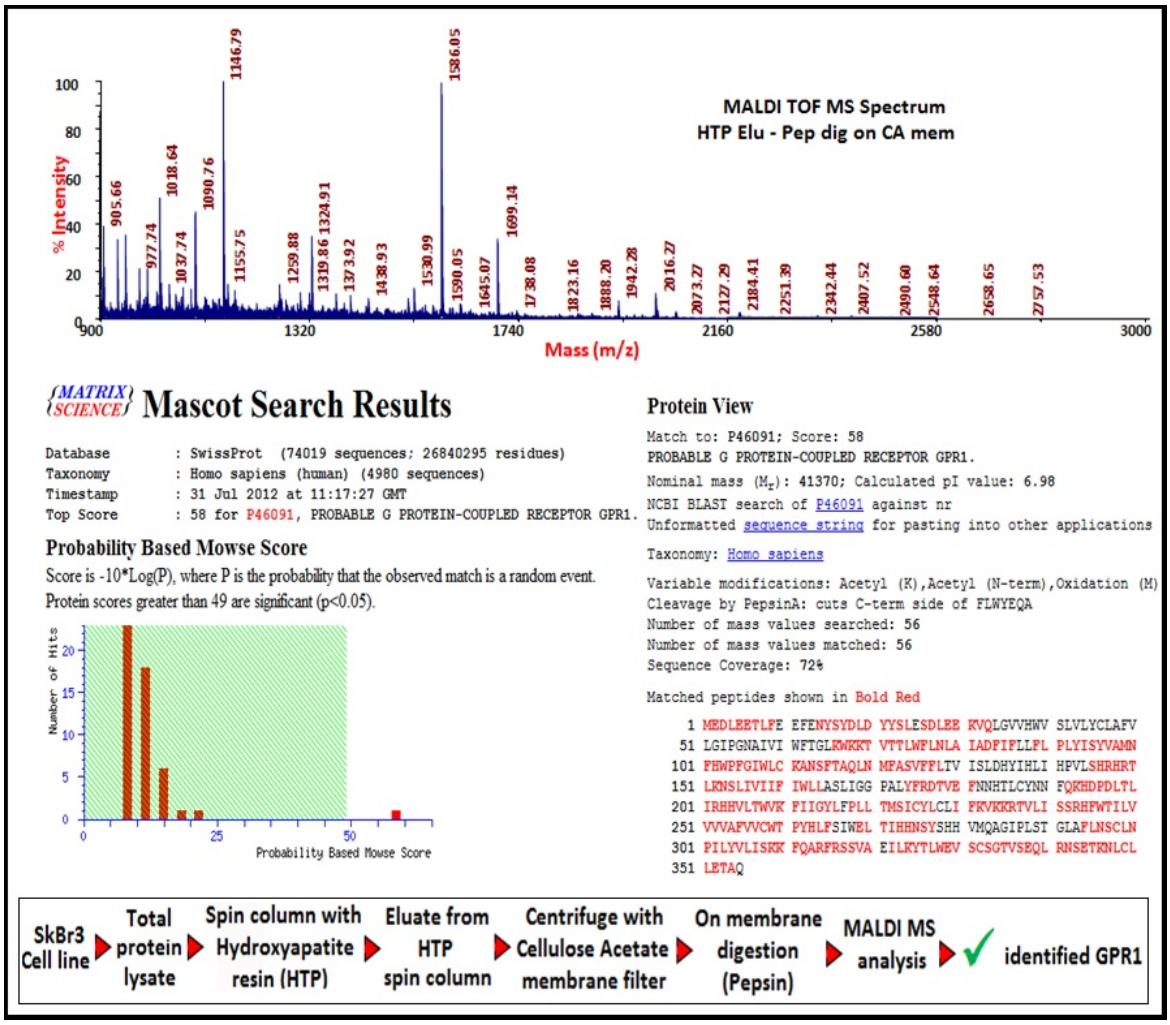


Figure 2.10 Peptide mass fingerprint of GPR1.

After passing the eluate through CA membrane, the membrane was washed with H_2O and 50 mM NH_4HCO_3 as described earlier and removed from the filter unit. The proteins on-membrane was then digested overnight with α-chymotrypsin (pH 7.8) at 30°C, followed by sample clean-up using C18 ZipTip and MALDI MS analysis. At first, we were not successful in fishing out GPER from the searches, but we didn't give up. We repeated the experiment several times, obtained a set of MS data by altering the acquisition parameters and performed intensive database search by fine tuning the search parameters. Finally, to our surprise, when using α-chymotrypsin for proteolysis, we were able to identify GPER **(Fig. 2.11)** by peptide mass fingerprinting with top score.

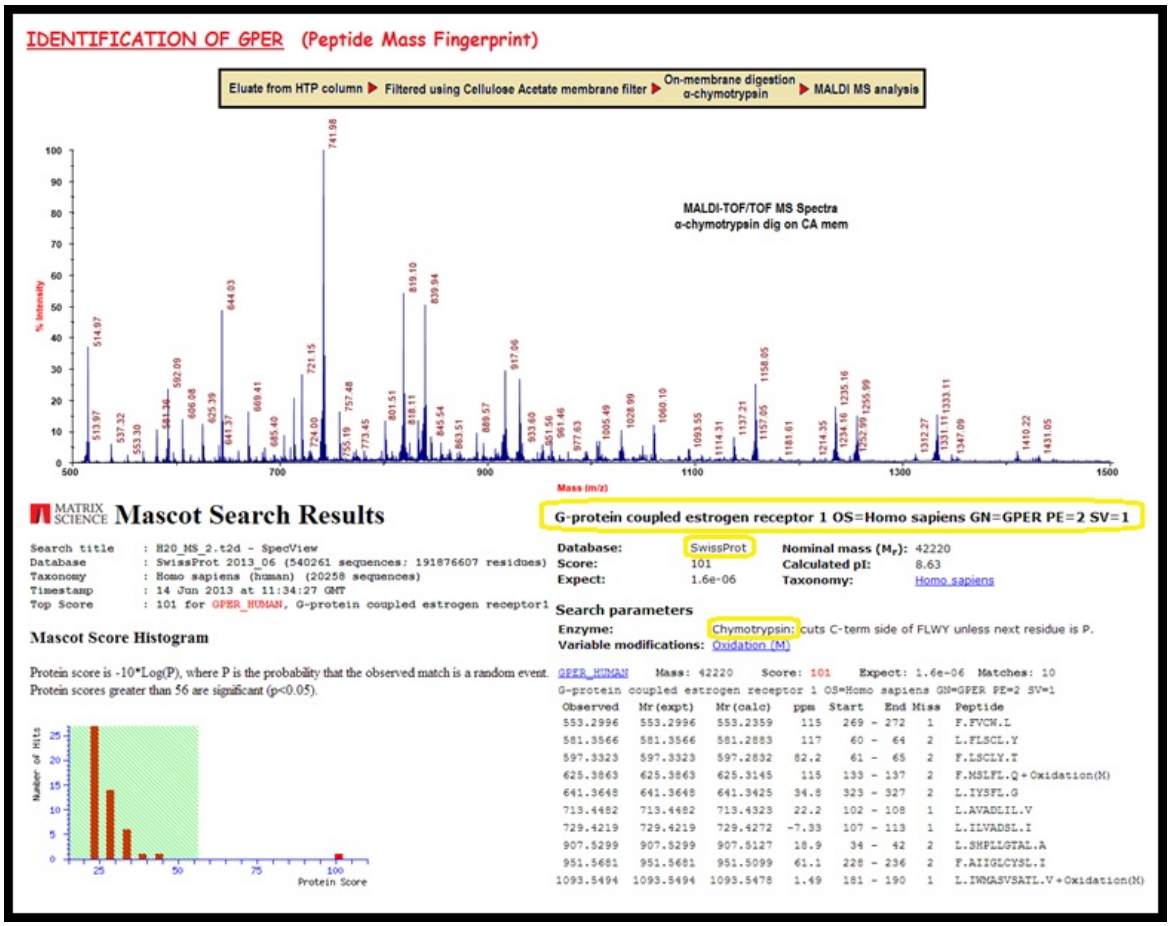


Figure 2.11 GPER identification by peptide mass fingerprinting (MALDI MS).

So far, this is the first study to report MS-based proteomics approach to isolate and identify GPER from total proteins extracted by lysing SkBr3 breast cancer cells. The use of CA membrane spin filters after enriching and isolating GPERs with HTP spin column made GPER identification easier and possible by MALDI MS. From our experiments while optimizing the workflow, we learnt that this protocol is not limited to GPER, but it can also be applied to study several other GPCRs by MALDI MS.

CHAPTER 3

GPER POST-TRANSLATIONAL MODIFICATIONS: Analysis of N-glycosylation in GPER by MALDI-TOF/TOF tandem Mass Spectrometry

Over the last few decades, researchers have discovered that the human proteome is incredibly complex and challenging than the human genome. Though it is estimated that the human genome comprises between 20,000 and 25,000 genes[177], the total number of proteins expressed in the human proteome is projected to be above 1 million[178]. These estimations demonstrate that single genes encode multiple proteins. Genomic recombination, transcription initiation at alternative promoters, differential transcription termination, and alternative splicing of the transcript are mechanisms that generate different mRNA transcripts from a single gene.[179] The increase in complexity from the genome level to the proteome is further facilitated by protein post-translational modifications (PTMs). PTMs are chemical modifications that play a key role in functional proteomics, because they regulate protein behavior such as activity, turnover, localization and molecular interactions.

3.1 Protein post-translational modifications

Post-translational modifications (PTMs), a step in protein biosynthesis that can actively play an essential role in regulating protein functions. Normally, proteins are created by ribosomes translating mRNA into polypeptide chains. These polypeptide chains undergo modifications such as folding, cutting and attaching other biochemical functional groups (e.g. acetate, phosphate, or carbohydrates) before maturing as a complete protein product.[180] Most common protein modifications **(Fig. 3.1)** include phosphorylation, glycosylation, ubiquitination, nitrosylation, methylation, acetylation, lipidation and proteolysis. The PTMs play a crucial role in generating the heterogeneity in proteins and also help in utilizing identical proteins for different cellular functions in different cell types. These PTMs regulate how a particular protein sequence should act in most of the eukaryotic organisms. Defects in PTMs have been linked to numerous developmental disorders and human diseases, highlighting the importance of PTMs in maintaining normal cellular states.[181]

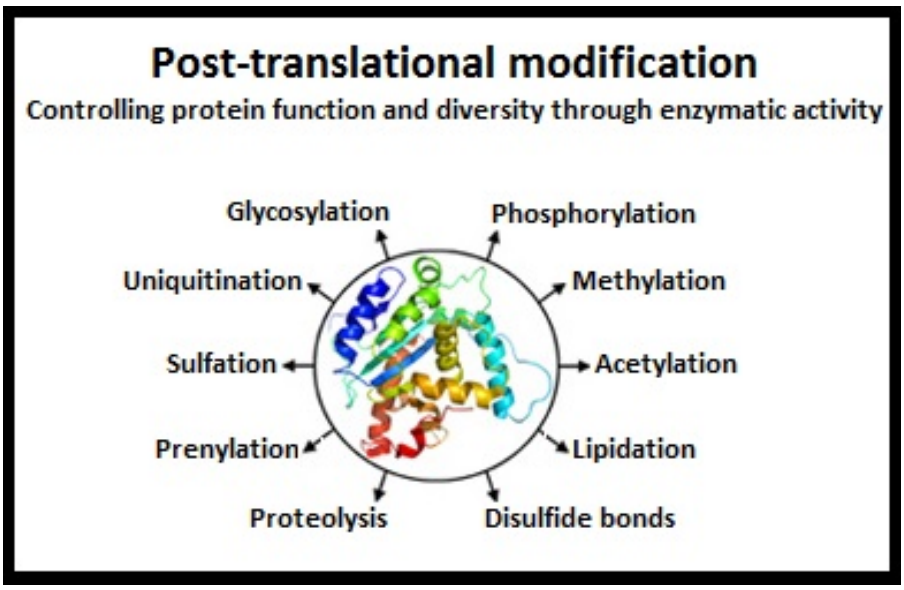


Figure 3.1 Common post-translational modifications.

PTMs occur at different amino acid side chains or peptide linkages and are often mediated by enzymatic activity. In fact, it is estimated that 5% of the proteome comprises enzymes that perform more than 200 types of PTMs.[182] These enzymes include kinases, phosphatases, transferases and ligases, which add or remove functional groups, proteins, lipids or carbohydrates to or from amino acid side chains, and proteases, which cleave peptide bonds to remove specific sequences or regulatory subunits.[182][183][184] PTMs can occur at any stage in the life cycle of a protein. For example, many proteins are modified shortly after translation is completed to mediate proper protein folding or stability or to direct the nascent protein to distinct cellular compartments such as nucleus and membrane. Some other modifications occur after folding and localization are completed to activate or inactivate catalytic activity or to influence the biological activity of the protein. Proteins are not limited to have single modification, often they are modified through a combination of post-cleavage and the addition of functional groups through a step-wise mechanism of protein maturation or activation. Some of the common types of PTMs studied in protein research are outlined in **Table 3.1**.

Table 3.1 Common types of post-translational modifications.

Protein post-translational modifications	
Glycosylation	Many proteins, particularly in eukaryotic cells, are modified by the addition of carbohydrates, a process called glycosylation. Glycosylation in proteins results in addition of a glycosyl group to either asparagine, hydroxylysine, serine, or threonine.[185]
Phosphorylation	The addition of a phosphate group, usually to serine, tyrosine, threonine or histidine.[186]
Ubiquitination	The attachment of ubiquitin to a substrate protein.[187]
Acetylation	The addition of an acetyl group, usually at the *N*-terminus of the protein.[188]
Alkylation	The addition of an alkyl group (e.g. methyl, ethyl).[189]
Methylation	The addition of a methyl group, usually at lysine or arginine residues. This is a type of alkylation.[190]
Sulfation	The addition of a sulfate group to a tyrosine.[191]
Nitrosylation	Specifically S-nitrosylation, involves the covalent incorporation of a nitric oxide moiety into thiol groups at cysteine residues, to form S-nitrosothiol.[192]
Lipidation	A method to target proteins to membranes in organelles (endoplasmic reticulum, Golgi apparatus, mitochondria), vesicles (endosomes, lysosomes) and the plasma membrane.[193]
Isoprenylation	The addition of an isoprenoid group (e.g. farnesol and geranylgeraniol). This is a type of lipidation.[194]
Palmitoylation	Specifically S-palmitoylation, the reversible addition of palmitate and other long-chain fatty acids to proteins at cysteine residues. This is a type of lipidation.[193]
Myristoylation	The irreversible attachment of a myristoyl group, derived from myristic acid by an amide bond to the α-amino group of an *N*-terminal glycine residue.[195]

Biotynylation	Acylation of conserved lysine residues with a biotin appendage.[196]
Glutamylation	Covalent linkage of glutamic acid residues to tubulin and some other proteins.[197]
Glycylation	Covalent linkage of one to more than 40 glycine residues to the tubulin *C*-terminal tail of the amino acid sequence.[198]
Phosphopantetheinylation	The addition of a 4'-phosphopantetheinyl moiety from coenzyme A, as in fatty acid, polyketide, non-ribosomal peptide and leucine biosynthesis.[199]
Selenation	The addition of selenate group to a cysteine. Selenates are analogous to sulfates and have similar chemistry.[200]
***C*-terminal amidation**	The addition of an amide group to the end of the polypeptide chain.[201]
Lipoylation	The attachment of a lipoate functionality.[202]
Proteolysis	A ubiquitous and irreversible PTM involving limited and highly specific hydrolysis of peptide and isopeptide bonds of a protein by a protease.[203]

Based on previous understanding and emerging data, it seems evident that PTMs are involved in regulating almost all cellular events, including gene expression, signal transduction, protein-protein interaction, cell-cell interaction, and communication between the intracellular and extracellular environment.[204] The analysis of proteins and their PTMs is especially important for the study of heart disease, cancer, neurodegenerative diseases and diabetes. Protein modifications that influence the activity of pivotal proteins can only be determined by studying the proteome.[205] Though the characterization of PTMs is well challenging, it cannot be left behind as it provides invaluable insight into the cellular functions underlying the processes that leads to subsequent disorders in living organism. Technically, the main challenges in studying post-translationally modified proteins are the development of specific detection and purification methods. Fortunately, these technical complexities are currently being

tackled with a variety of new proteomics technologies. Tandem mass spectrometric methods can be employed to localize the sites of post-translational modifications.[206][207]

3.2 Glycosylation in GPER

Post-translational modification through transfer of glycans or carbohydrates to proteins is a complex process requiring the concerted action of a series of glycosyltransferases (GTs), each catalyzing a specific step in the pathway. The conservation of this complicated process throughout evolution suggests that important functions are attached to protein glycosylation. Unfortunately, these functions remain poorly understood. Protein glycosylation is acknowledged as one of the major post-translational modifications, with significant effects on protein folding, conformation, distribution, stability and activity. It is well known that protein glycosylation plays a critical role in the regulation of protein structure[208], signal transduction[209], cell-cell and cell-environment interactions[210][211][212], immune responses[213][214], hormone action[215], cancer progression[216] and embryonic development[217][218]. Nearly half of all known proteins are potentially glycosylated[219] and this PTM is characterized by various glycosidic linkages **(Fig. 3.2)**, including *N*-linked glycosylation, *O*-linked glycosylation and *C*-linked mannosylation, glypiation (GPI anchor attachment) and phospho-glycosylation.[220] In cells, the most abundantly found protein glycosylation are *N*-linked and *O*-linked glycomodifications. *N*-linked glycosylation often occurs on a large variety of nascent proteins, whereas, *O*-linked monosaccharide modification of *N*-acetylglucosamine (GlcNAc) on serine, threonine or amino residues in close proximity to tyrosine phosphorylation sites is frequently observed in many cells. At these sites, glycosylation may contribute to the regulation of signaling pathways through a direct competition with serine and threonine phosphorylation or by indirectly disturbing the phosphorylation of tyrosine.[221]

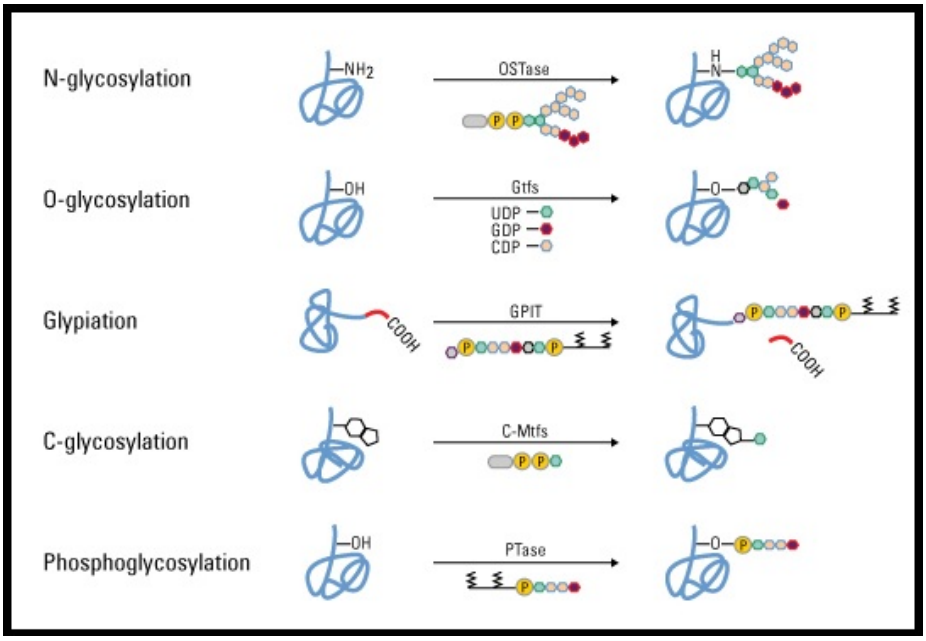


Figure 3.2 Different types of glycosidic linkages.[222]

In UniProt database (http://www.uniprot.org/uniprot/Q99527), GPER is reported to have three potential glycosylation sites, *N*-linked (GlcNAc) based on sequence analysis. The possible positions for the glycosylation sites are shown in **Table 3.2**. A number of different technologies are currently being used to study glycosylation, including mass spectrometry[223][224][225], nuclear magnetic resonance (NMR)[226] and liquid chromatography for 'glycan sequencing'[227][228]. Glycoproteins can be detected (glycan staining and visualization), purified (glycan cross-linking to agarose or magnetic resin) and analyzed by mass spectrometry-based proteomics. Over last decade, glycan microarrays are used for glycan-protein interaction profiling[229] and new techniques to fluorescently label glycans have enabled quantification of glycan species on the array[230]. In the presented work, we employed mass spectrometry-based approach to study GPER glycosylation.

Table 3.2 Glycosylation sites in GPER.

PTM	Position	Length	Description
Glycosylation	25	1	*N*-linked (GlcNAc)

Glycosylation	32	1	*N*-linked (GlcNAc)
Glycosylation	44	1	*N*-linked (GlcNAc)

3.3 Glycopeptide derivatization with dansyl chloride

Tandem mass spectrometry (MS/MS or MS^n) is a powerful tool for characterizing *N*-linked glycopeptide structure. Mass spectrometry-assisted glycopeptide analysis can provide information on both glycan structure heterogeneity and glycosylation site, enabling site-specific and/or protein-specific glycosylation analysis.[231] However, it is still difficult to obtain detailed structural information on the glycan moiety directly from glycopeptide ions. For structural analysis, collision-induced dissociation (CID) is used as the ion-activation method for fragmenting glycopeptide ions. Under CID conditions, cleavages of glycosidic bonds preferentially occur prior to peptide fragmentation. Thus, more harsh conditions for ion fragmentation or MS^3 measurements are required to obtain peptide sequence information.[232][233] The glycan structure can be simply deduced from CID spectra of glycopeptides; however, available structural information is limited. Dominant cleavages in glycosidic bonds provide only glycan compositional information. Moreover, monosaccharide rearrangements may occur in the dissociation of protonated glycopeptides.[234] Due to the inherent branched structures of *N*-linked glycans, glycan compositional information is not sufficient for characterizing the structure of glycan moiety in glycopeptides.

In the presented work, we used dansyl chloride (DNS-Cl) for glycopeptide derivatization and manually validated two glycosylation sites, reported in GPER. Glycopeptide derivatization using DNS-Cl have been reported successful in identifying glycosylation sites and characterizing the structure of glycan moieties.[235] Dansylation is a method for determining the *N*-terminal residue of a peptide. It reacts with the terminal primary amine. DNS-Cl is well known for its competency to improve ionization efficiency and peptide fragmentation.[236] Dansylation is expected to increase the peptide mass by

234 Da. It increases the *m/z* value of each individual peptide and enlarges the mass scale of the experiment. Hence, the increased ionization efficiency and the better signal-to-noise ratio thus achieved, allows the identification of dansylated peptides by MALDI, to a greater extent.[235]

MS analyses were performed using a 5800 MALDI-TOF/TOF analyzer equipped with a Nd:YAG (349 nm) laser, in reflectron positive-ion mode with a mass accuracy of 5 ppm. At least, 4000 laser shots were typically accumulated with a laser pulse rate of 400 Hz in the MS mode. In **Fig. 3.3**, the presence of possible glycans including hexose and GLcNAc were predicted in the MS spectrum obtained by shooting the pepsin-digested sample. The mono isotopic mass of hexose and GLcNAc are 162.05 Da and 203.07 Da, respectively.

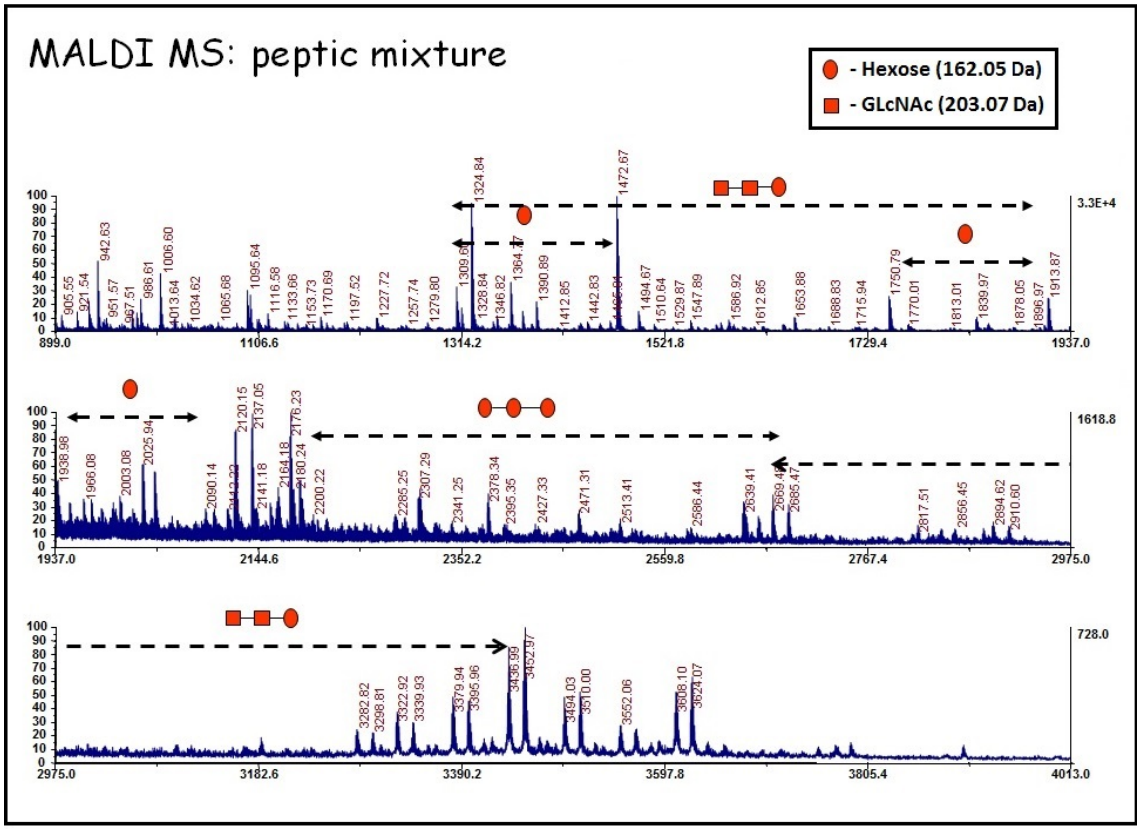


Figure 3.3 MALDI MS spectrum of peptic mixture (SkBr3 lysate).

A closer inspection of the characteristic pattern displayed in **Fig. 3.4** provided some information on the presence of a hexose sugar, as the difference between the ion peaks at *m/z* 1472.54 and *m/z* 1309.49 in the spectrum is about 162 Da. These MS spectra were acquired from the peptic digests before dansylation.

In the MS/MS mode spectra up to 5000 laser shots were acquired and averaged with a pulse rate of 1000 Hz. MS/MS experiments were performed at a collision energy of 1-2 kV, and ambient air was used as the collision gas with a medium pressure of 10^{-6} Torr. The MS/MS spectra of the precursor ion peaks at *m/z* 1472.54 and *m/z* 1309.49, before dansylation were shown in **Fig. 3.5**.

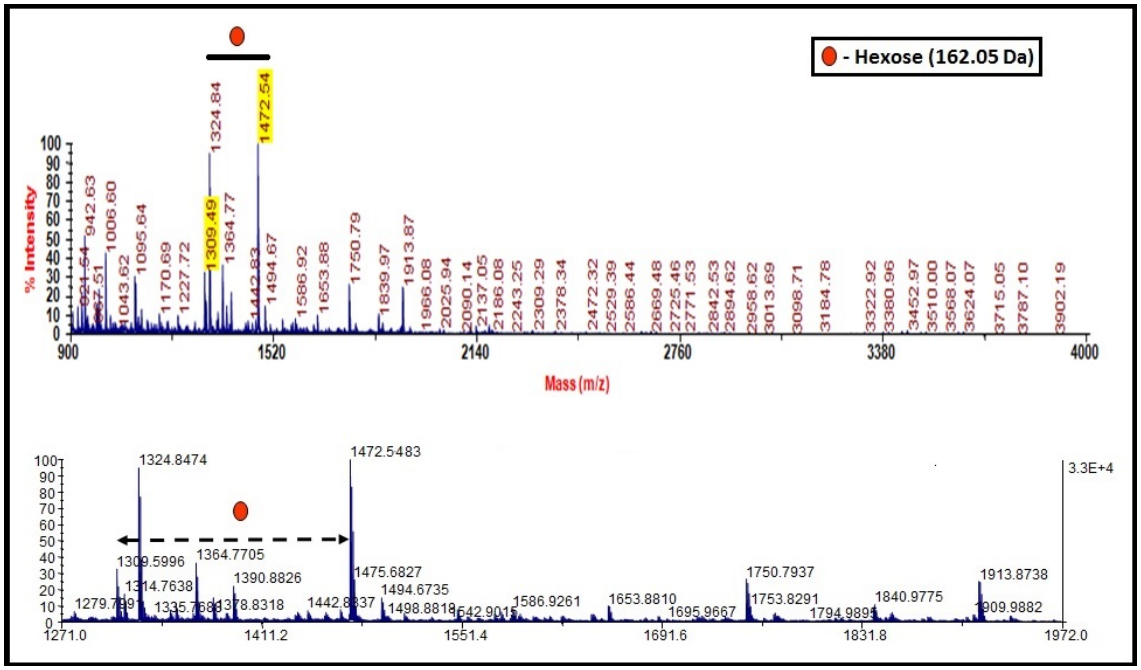


Figure 3.4 MALDI MS spectrum showing possible hexose sugar.

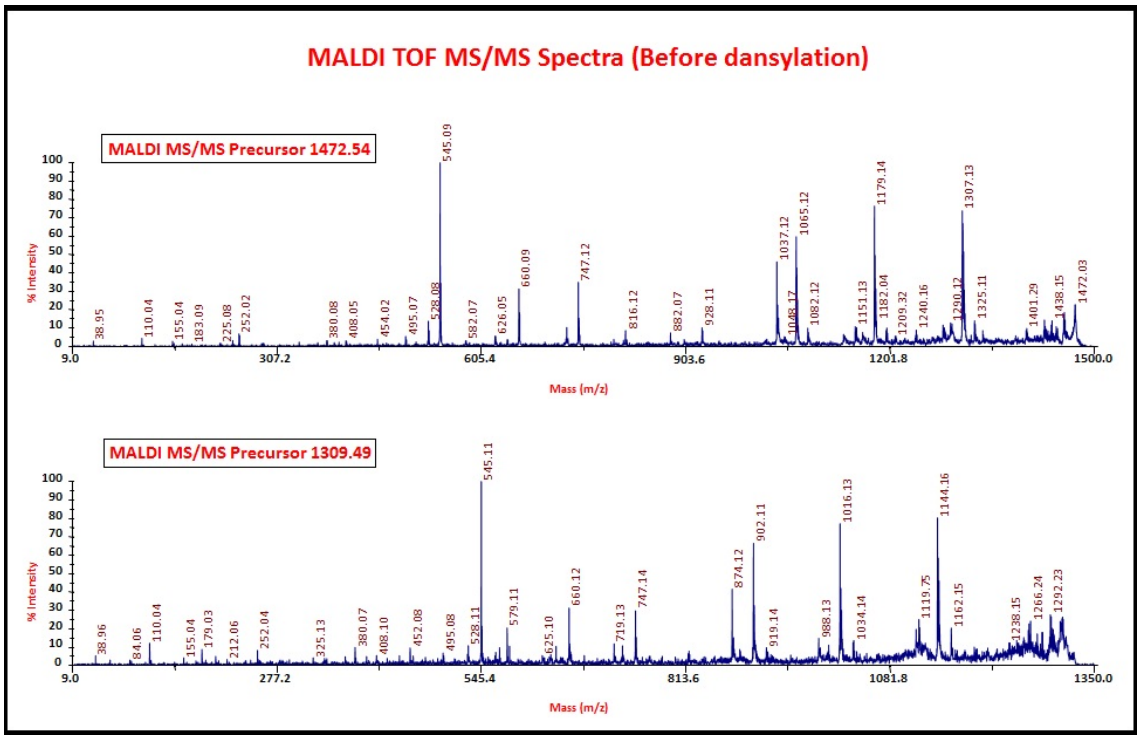


Figure 3.5 MS/MS spectra of precursor ions *m/z* 1472.54 and *m/z* 1309.49 before dansylation.

Peptide dansylation was performed by adding 16 µl DNS-Cl (0.22 mol/L, in CH_3COCH_3), 10 µl 0.1M $NaHCO_3$ and 5 µl Na_2CO_3 to 20 µl of sample (SkBr3 TPL - processed through HTP column and filtered using CA spin filter, followed by on-membrane digestion with suitable protease). For optimum results, pH 11 should be maintained throughout the reaction. The mixture should be wrapped in aluminium foil and kept under magnetic stirring for 3 hours at room temperature, followed by sample clean-up using C18 ZipTip and MALDI MS and MS/MS analysis. After dansylation, the mass spectrum showed more peaks than that of underivatized peptides. Peptide peaks appeared with one or two dansyl modifications, whereas, unmodified peptides were not observed showing that each pepsin-digested peptide has taken up at least one dansyl group. *N*-terminal dansylation increased desorption and ionization efficiency in α-CHCA matrix. The peptide mass should increase by 234 Da on dansylation, thereby increasing the *m/z* 1472.54 and *m/z* 1309.49 to *m/z* 1706.54 and *m/z* 1543.49, respectively. The MS/MS spectra of the

precursor ion peaks at the corresponding *m/z* 1706.54 and *m/z* 1543.49 were shown in **Fig. 3.6**.

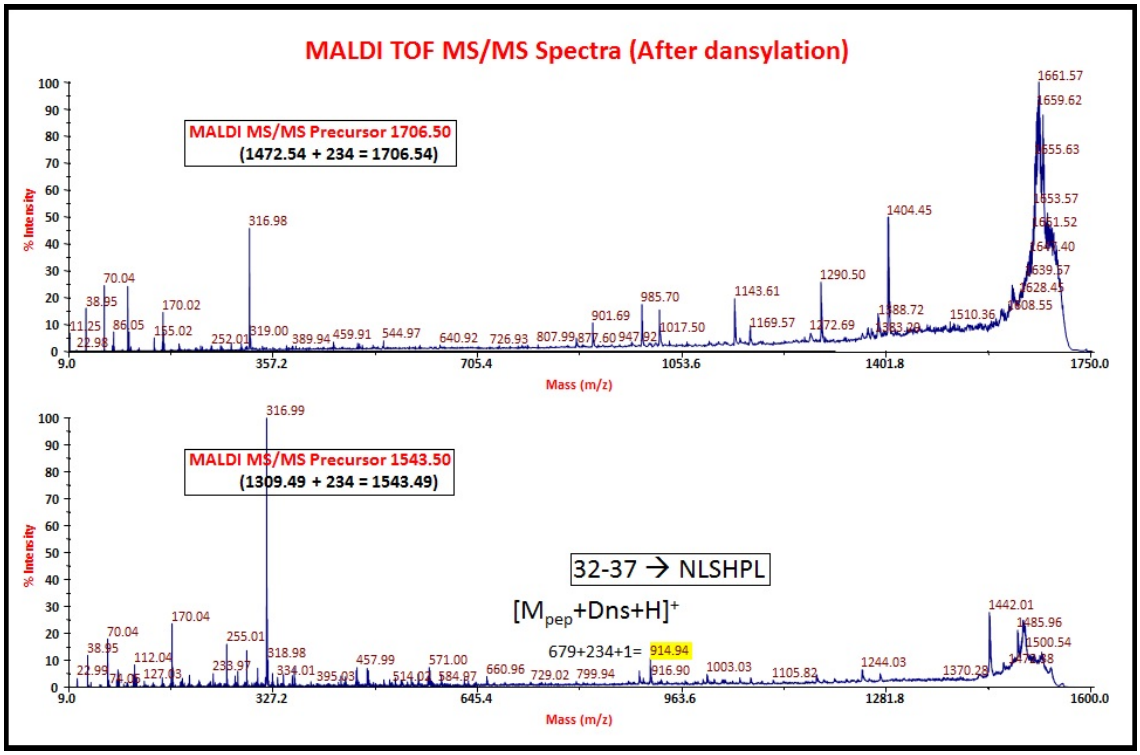


Figure 3.6 MS/MS spectra after dansylation.

The first *N*-glycosylation site at position no. 32 in GPER protein sequence was validated in **Fig. 3.6**. We used Glycomod program (http://web.expasy.org/glycomod/) to determine this. Glycomod is an online tool for predicting all possible composition of oligosaccharide structures in glycoproteins from their experimentally determined masses. From Glycomod, the structural information including possible glycosylation sites can be obtained by inputting data derived from preliminary evaluation of the most likely constituents present in the examined glycopeptides. The experimentally determined mass of glycopeptide can be used to find all possible compositions of the glycan moiety when glycosylation site and protein structure are known. The glycopeptide of *m/z* 679.37 with 1 missed cleavage (MC) from **Table 3.3** (generated from Glycomod) is more closely connected to the experimentally determined peak at *m/z* 914.94 in the MS/MS spectrum.

On subtracting the mass of DNS-Cl (234 Da) and H^+ (1 Da) ion from the experimentally observed peak at *m/z* 914.94, the glycopeptide mass is known to be *m/z* 679.94 which is corresponding to the peptide NLSHPL at position 32-37 **(Table 3.3)** in GPER protein.

Table 3.3 Possible glycopeptides predicted using Glycomod.

Position	MC	Peptide mass	Peptide
23-30	1	815.36609	APNTTSPE
24-30	0	744.32898	PNTTSPE
24-31	1	857.41304	PNTTSPEL
32-37	1	679.36530	NLSHPL
43-48	1	547.22378	ANGTGE
44-48	0	476.18667	NGTGE
44-49	1	589.27073	NGTGEL
80-87	1	903.54292	ILVVNISF
82-87	0	677.37480	VVNISF
82-89	1	962.51850	VVNISFRE
314-319	1	669.27920	FSNSCL
315-319	0	522.21079	SNSCL
315-322	1	846.39054	SNSCLNPL

We further exploited the MS/MS approach to validate the *N*-glycosylation site suggested by the Glycomod tool and the amino acid sequence of glycopeptide. **Fig. 3.7** and **Fig. 3.8** shows the high energy CID spectrum of the dansylated-glycopeptide with b- and y-ion coverage and MS/MS validation of structures, respectively. The inner structure in **Fig. 3.7** shows the cleavage pattern with respect to $[M_{pep}-16]^+$, $[M_{pep}+H]^+$, cleavage at the innermost *N*-acetylglucosamine residue $[M_{pep}+84]^+$ and Y-type cleavage of the chitobiose core $[M_{pep}+204]^+$ with the retention of the peptide moiety.

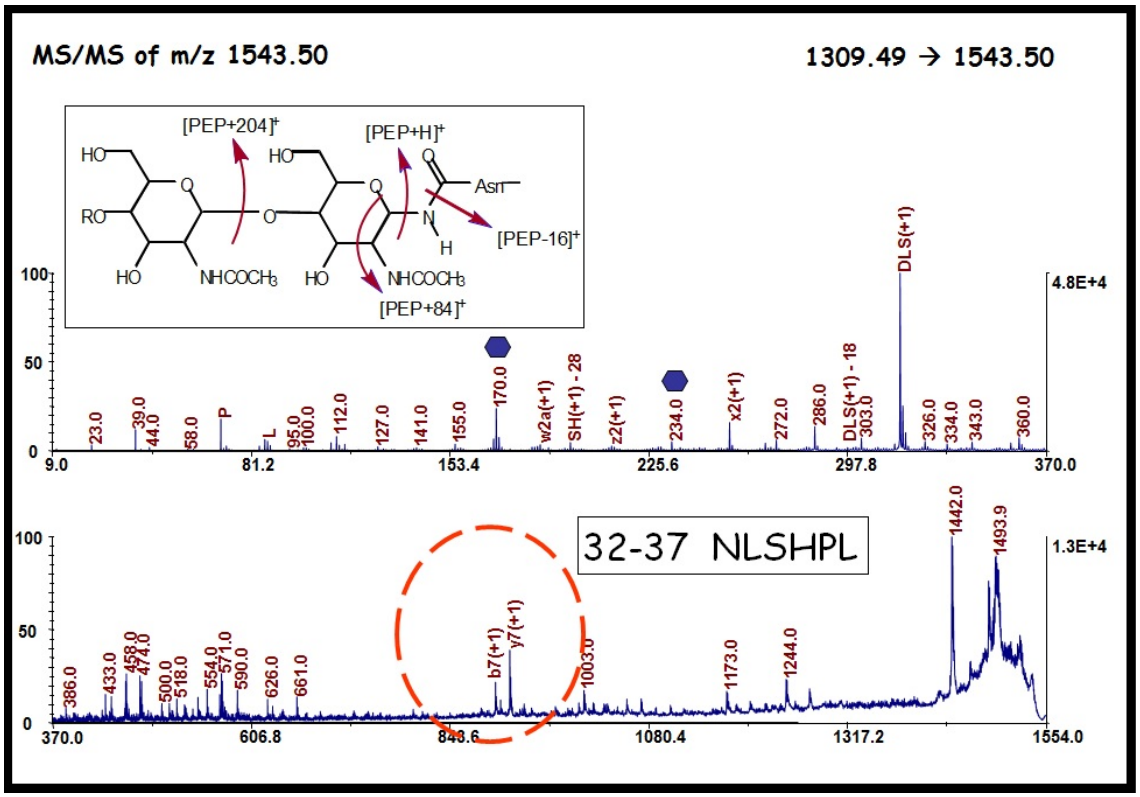


Figure 3.7 MS/MS spectrum of *m/z* 1543.50.

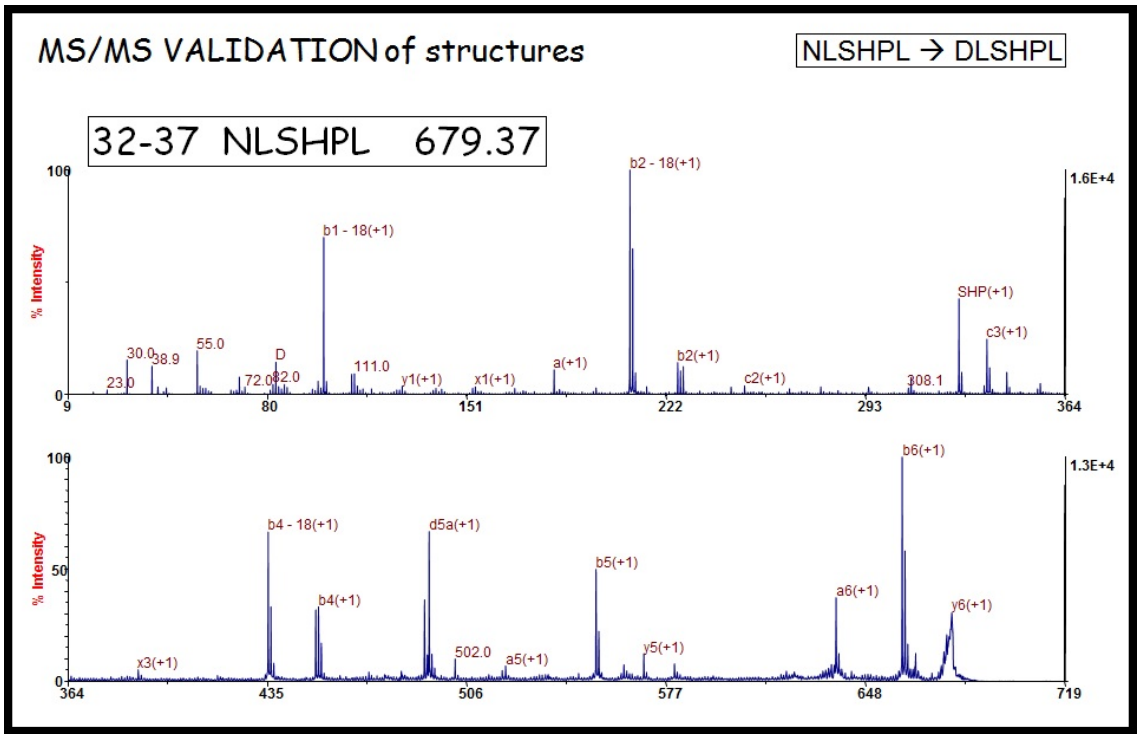


Figure 3.8 MS/MS spectrum of *m/z* 679.37.

Hybrid complex and high mannose structure type were predicted and assigned to the observed glycopeptide NLSHPL (MC 1) of *m/z* 679.37 **(Fig. 3.9)**. The composition of oligosaccharides was calculated by inputting the presence of at least one Hexose (Hex) and two *N*-acetylhexoses (HexNAc), representing the conservative core.

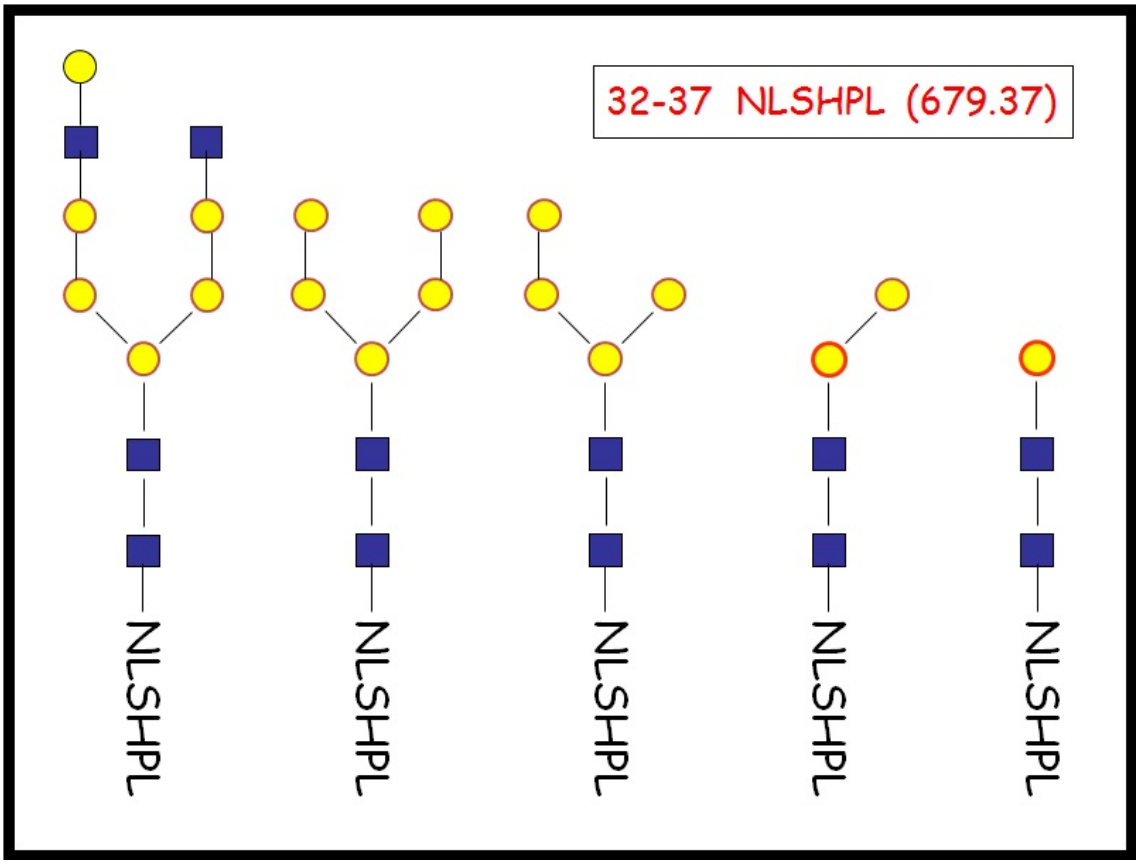


Figure 3.9 Oligosaccharide structures corresponding to NLSHPL.

Similarly, Peptide dansylation was performed on samples digested with α-chymotrypsin and analyzed by MALDI MS and MS/MS. The simple MALDI MS spectrum of the α-chymotrypsin-digested sample showed, in fact, the specific *m/z* spacing patterns (n203 and n162+n203) of glycoforms with in the 1-3 kDa mass range **(Fig. 3.10).**

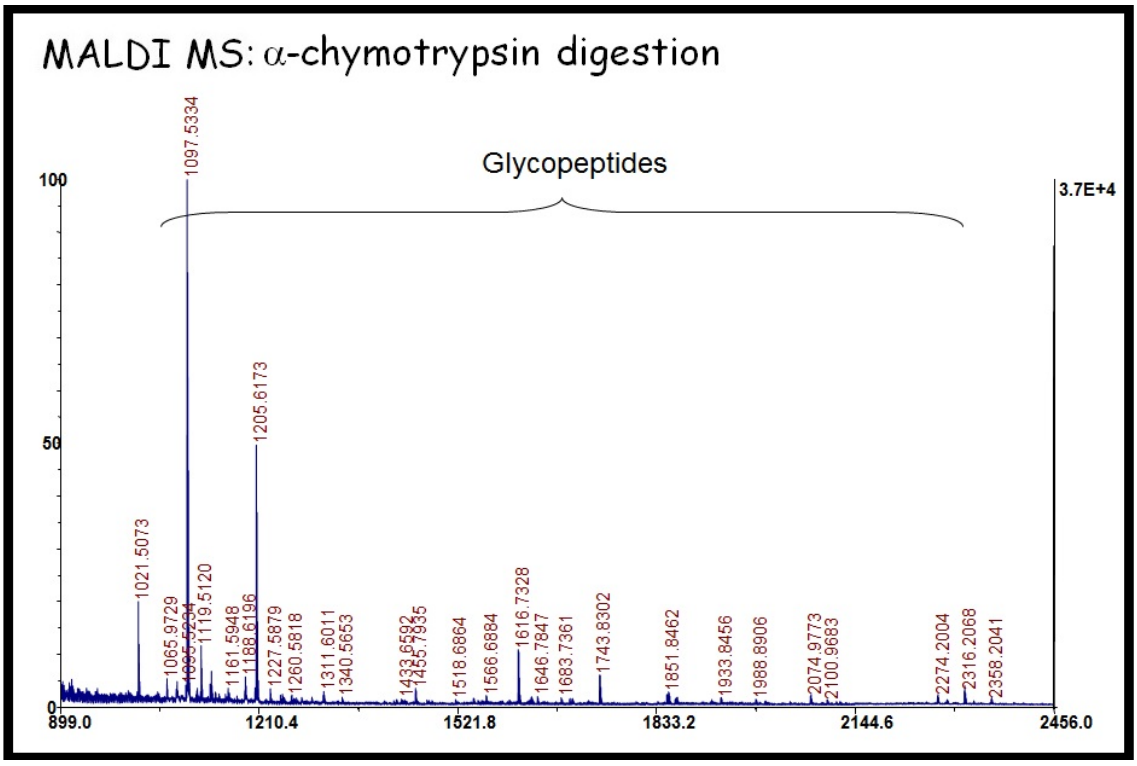


Figure 3.10 MALDI MS spectrum of α-chymotrypsin-digested sample (SkBr3 lysate).

We validated the second glycosylation site at position no. 44 in GPER and glycan composition including the corresponding amino acid sequence of glycolpeptide based on the structural information **(Table 3.4)** obtained by inputting the data derived from preliminary evaluation of the most likely constituents present in the examined glycopeptides. The high energy CID spectrum of *m/z* 1616.64 **(Fig. 3.11)** shows the MS/MS validation of predicted oligosaccharide structures corresponding to the glycopeptide ANGTGEL.

Table 3.4 Oligosaccharide structures predicted by Glycomod.

Glycan mass	Structure	Peptide mass	Position	Peptide sequence	Total Gly mass
552.217	$(HexNAc)_2$ $(Deoxyhexose)_1$	522.211	315-319	SNSCL	1097.418
1444.534	$(HexNAc)_2$ $(Deoxyhexose)_1$ + $(Man)_3$ $(GlcNAc)_2$	522.211	315-319	SNSCL	1989.735

1501.555	$(HexNAc)_3 + (Man)_3(GlcNAc)_2$	792.449	32-38	NLSHPLL	2316.994
933.344	$(Hex)_2 (HexNAc)_3$	660.308	43-49	ANGTGEL	1616.642
1590.592	$(HexNAc)_2 (Deoxyhexose)_2 + (Man)_3 (GlcNAc)_2$	660.308	43-49	ANGTGEL	2273.890
1939.729	$(HexNAc)_3 (Deoxyhexose)_3 + (Man)_3 (GlcNAc)_2$	660.308	43-49	ANGTGEL	2623.027
2391.819	$(Hex)_8(HexNAc)_1 + (Man)_3 (GlcNAc)_2$	660.308	43-49	ANGTGEL	3075.117
1524.534	$(Hex)_3 (Deoxyhexose)_1 + (Man)_3 (GlcNAc)_2$	811.353	312-319	AAFSNSCL	2358.877
2928.041	$(Hex)_7 (HexNAc)_3 (Deoxyhexose)_2 + (Man)_3 (GlcNAc)_2$	679.365	32-37	NLSHPL	3630.396

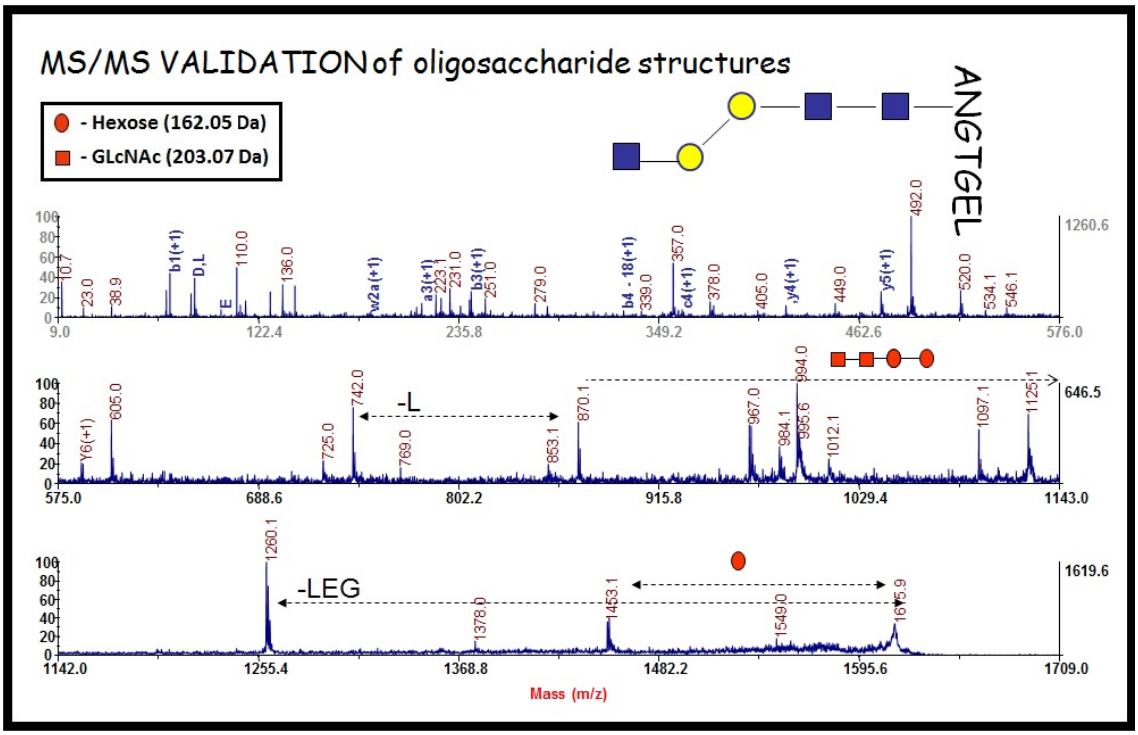


Figure 3.11 MS/MS validation of oligosaccharide structure.

3.4 GPER deglycosylation with endoglycosidases

Protein deglycosylation refers to the removal of the glycan moiety from a glycoprotein. This could be accomplished by either chemical or enzymatic methods. However, chemical methods such as β-elimination with mild alkali or mild hydrazinolysis can be harsh and may result in incomplete sugar removal and degradation of the protein; whereas, enzymatic methods are much gentler and can provide complete sugar removal with no protein degradation.

Use of the enzyme Peptide-*N*-Glycosidase F (PNGase F) is the most effective method for removing all *N*-linked oligosaccharides from glycoproteins.[237] PNGase F digestion deaminates the asparagine residue to aspartic acid, and leaves the oligosaccharide intact, keeping it suitable for further analysis. A tripeptide with the oligosaccharide-linked asparagine as the central residue is the minimal substrate for PNGase F. However, oligosaccharides containing a fucose α(1-3)-linked to the glycan core, commonly found in some plants and/or insect glycoproteins, are resistant to PNGase F. Peptide-*N*-Glycosidase A (PNGase A), isolated from almond meal, must be used in this situation.[238] Steric hindrance may slow or inhibit the action of PNGase F at certain glycosylation sites. Therefore, denaturation and reduction of the glycoprotein by heating with SDS and 2-mercaptoethanol greatly increases the rate of deglycosylation. Other commonly used endoglycosidases such as Endoglycosidase H and the Endoglycosidase F series are not suitable for general deglycosylation of *N*-linked sugars because of their limited specificities and because they leave one *N*-acetylglucosamine residue attached to the asparagine.[239][240]

In our study, we tried using both PNGase F and Endoglycosidase H for protein and peptide deglycosylation. Two sets of SkBr3 TPL were processed through HTP spin column and filtered using CA spin filter, individually. The membranes were then removed and the bound proteins were digested on-membrane overnight with pepsin (pH 2) at 37°C and α-chymotrypsin (pH 7.8) at 30°C, simultaneously. Further, the peptic and α-chymotryptic mixture were deglycosylated separately overnight with PNGase F at 37°C and analyzed by

MS and MS/MS. In later experiments, the deglycosylation was performed on-membrane (after CA membrane filtration), followed by proteolytic digestion with pepsin and α-chymotrypsin, and MS analysis. In some experiments, we used 0.1% SDS to assist the release of deglycosylated and proteolytically digested peptides from CA membrane. Sample clean-up using C18 ZipTip was implemented prior to MS analysis whenever SDS was used.

Similarly, deglycosylation experiments were performed on the proteins bound on CA membrane (after filtration) using Endoglycosidase H (Sigma Endo H kit) as per manufacturer instructions. The glycans were analyzed by MS and MS/MS. Then, the CA membrane was further subjected to proteolytic digestion and analyzed by MS and MS/MS after sample clean-up with C18 ZipTip. Unfortunately, the results from the deglycosylation experiments were not satisfactory, and thus, not reported in this thesis.

In this chapter, we reported the proteomic analysis of *N*-glycosylation in GPER by MALDI MS and MS/MS. We successfully validated 2 out of 3 *N*-glycosylation sites and the corresponding glycopeptides, NLSHPL at position 32-37 and ANGTGEL at position 43-49. Further, we predicted the possible oligosaccharide structures using Glycomod tool and manually validated the structures from obtained data by MS/MS.

CHAPTER 4

AFFINITY CAPTURE-RELEASE STRATEGY FOR GPER PURIFICATION: Design and synthesis of Polymer Nanoparticles with high affinity for GPER

Affinity chromatography is a method for separating biochemical mixture based on a highly specific interaction similar to the one between receptor and ligand. It is a variant of chromatography based on the ability of biomolecules (analytes) to bind certain ligands specifically and reversibly. These unique features of the analyte and the ligand interaction are then utilized for the separation of the analyte of interest from a complex mixture. From the first protein-protein interaction studies done in the late 1990´s[241][242], affinity separations have experienced a true renaissance in proteomics. A complete parade of affinity matrices and affinity-based experimental approaches has been developed that has found numerous applications ranging from subtraction of highly abundant proteins to study of drug target profiles to large scale mapping of posttranslational modifications.

4.1 Introduction to affinity purification

Proteins and other macromolecules of interest can be purified from crude extracts or other complex mixtures by a variety of methods. Selective precipitation is perhaps the simplest method for separating one type of macromolecule from another. Most purification methods, however, involve some form of chromatography whereby molecules in solution (mobile phase) are separated based on differences in chemical or physical interaction with a stationary material (solid phase). Gel filtration (also called size-exclusion chromatography or SEC) uses a porous resin material to separate molecules based on size (i.e., physical exclusion). In ion exchange chromatography, molecules are separated according to the strength of their overall ionic interaction with a solid phase material (i.e., nonspecific interactions).

By contrast, affinity chromatography (also called affinity purification) makes use of specific binding interactions between molecules. In the classical setup, a relevant ligand is attached to a solid, inert resin creating an affinity stationary phase (affinity matrix). When a sample containing desired analyte is passed over such affinity matrix, the analyte having specific binding affinity to the ligand become bound and retained by the matrix while the other molecules stay apart. After the other unwanted molecules are washed

away, the bound analyte is stripped from the affinity matrix, resulting in its purification from the original sample **(Fig. 4.1)**. This principle was discovered by P. Cuatrecasas and M. Wilchek, who applied it to the purification of *Staphylococcal nuclease* and avidine.[243][244] Since their discovery, numerous specialized affinity purification techniques appeared, but notably even today more than 90% of them apply the same general principles as reported in 1968.[245] Each specific affinity system requires its own set of conditions and presents its own peculiar challenges for a given research purpose.[246]

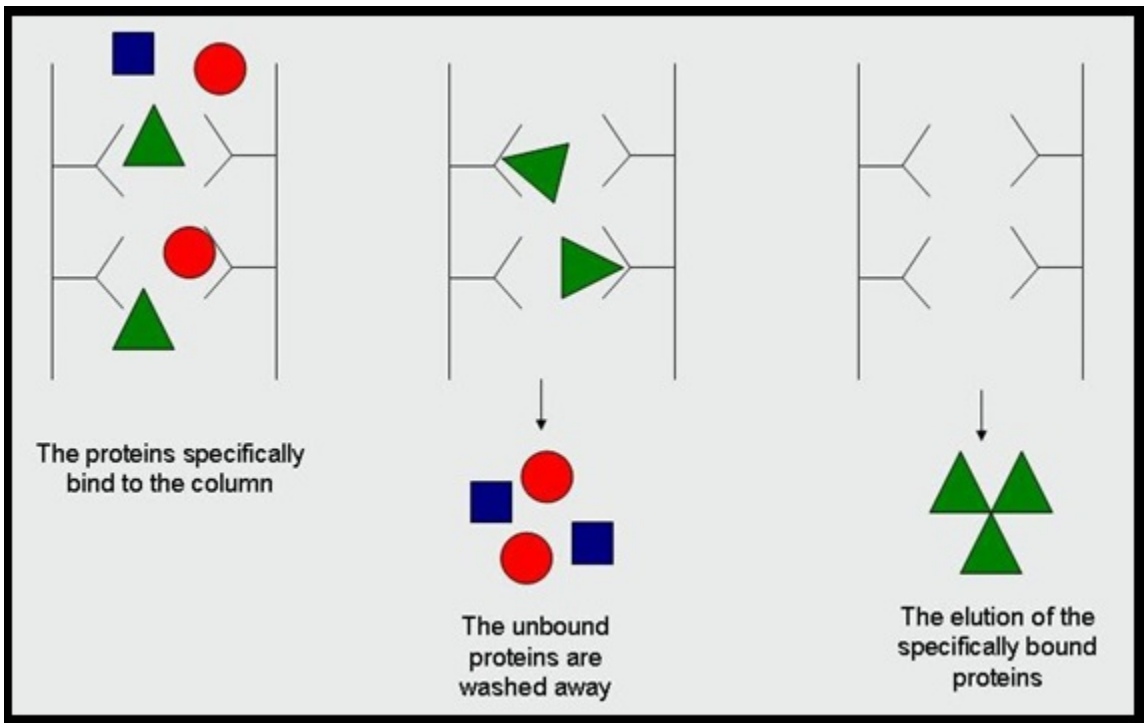


Figure 4.1 Protein purification using affinity chromatography.[247]

Basically, there are two main modes of affinity chromatography - a 'subtraction' mode and an 'enrichment' mode. If the aim of the purification is to specifically remove protein species that would hamper characterization of the sample, then affinity chromatography is employed in subtraction mode. Traditionally, the main purpose of affinity subtraction is the elimination of highly abundant proteins like albumin, immunoglobulin, etc., to achieve broader coverage of proteomes that suffer from wide

dynamic range such as human body fluids (blood plasma, cerebrospinal fluid, urine, saliva). Whereas, enrichment mode is applied for the isolation of selected protein species. At the protein level, the enrichment affinity chromatography permits the purification of a particular protein of interest or a group of low abundant proteins and/or proteins that share a specific structural feature. At peptide level, the enrichment affinity chromatography has attained an essential position in the purification of post-translationally modified species.

Moreover, the enrichment affinity chromatography of proteins can be used as a tool to obtain an information on specific protein affinities. In short, the basic principle of affinity chromatography allows to utilize the method for the isolation of protein partners of selected molecules.[248] During the affinity purification, the molecule of interest represents a 'bait' that is bound by its cellular protein counterparts, its 'preys'. These (protein) preys are then easily purified and consequently identified by MS[249][250], generating thus a map of the bait-protein interaction network. In this setup, affinity chromatography has facilitated the discernment of many different molecular relationships from protein-protein interactions to drug selectivity profiles.[251][252][253]

4.2 Engineered polymer nanoparticles for GPER purification

The support or matrix in affinity purification is any material to which a bio-specific ligand is covalently attached. Typically, the material to be used as an affinity matrix is insoluble in the system in which the target molecule is found. Usually, but not always, the insoluble matrix is a solid. Hundreds of substances have been described and utilized as affinity matrices, including agarose, cellulose, dextran, polyacrylamide, latex and controlled pore glass[254][255][256] Useful affinity supports are those with a high surface-area to volume ratio, chemical groups that are easily modified for covalent attachment of ligands, minimal nonspecific binding properties, good flow characteristics and mechanical and chemical stability.

Over the last decade, the use of magnetic nanoparticles[257][258][259][260] and polymer nanoparticles[261][262][263][264] for protein affinity purification have been widely reported. Engineered synthetic nanoparticles (NPs) with an intrinsic affinity and selectivity for target biomacromolecules are significant interest for use in diagnostics[265], therapeutics[266][267][268] and protein purification[269][270], and as a tool to investigate biochemical processes[271][272]. Recent studies show that synthetic NPs **(Fig. 4.2)** incorporating functional groups complementary to a surface domain of a target biomacromolecule can result in a high intrinsic affinity for target peptides[268], proteins[266][273], and polysaccharides[274][275]. These materials are attractive as an inexpensive and robust alternative to affinity reagents of biological origin, including antibodies.[276]

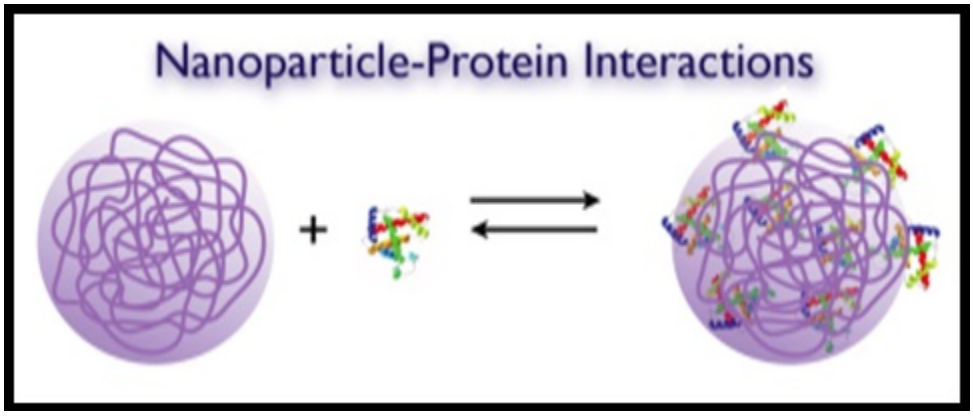


Figure 4.2 Synthetic polymer nanoparticles for protein purification.

It is well discussed in Chapter 1 that GPER may play a significant role in tamoxifen resistance in breast cancer cells. Although it is believed to be a key player in several other cancers, the background information at the molecular level is still limited. In order to study the complete proteome, post translational modifications, mutations and to better understand their molecular interactions with various other receptors and ligands, it is necessary to isolate pure GPER from crude cell lysate. The lack of promising separation and purification tools peaked our interest towards designing NPs that can capture this target biomolecule. In recent years, NPs with an intrinsic affinity have shown to be

successful in binding biomacromolecules like melittin[277][278], immunoglobulin G[279], histone[280], fibrinogen[280] and lysozyme[269] by controlling and optimizing the functional monomers composition.

In the presented work, we adopted a similar approach in an effort to capture GPER with high affinity and selectivity among a mixture of proteins that are expressed in cancer cells (SkBr3 & BG-1). Considering the fact that GPER is a membrane protein with many hydrophobic amino acid residues on the surface, a candidate NP was chosen from a library of nanoparticles that were prepared by combining different populations of functional groups on a poly-N-isopropylacrylamide (PNIPAm)-based polymer backbone. We started our experiment with GPER peptide selection and synthesis. Then, we evaluated the interaction between truncated-GPER (short peptide epitopes) and NPs by high performance liquid chromatography (HPLC). Finally, we moved on to evaluate the interaction between GPER (whole protein) and NPs by immunoblot analysis.

4.3 Solid phase peptide synthesis:

A list of amino acid sequence **(Table 4.1)** were prepared based on *in silico* data generated from UniProt, a database of protein sequence and functional information (accession number: Q99527).

Table 4.1 Truncated-GPER sequence information based on extracellular, cytoplasmic, and transmembrane domains.

UNIPROT accession : Q99527 (GPER1_HUMAN) / G-protein coupled estrogen receptor 1 / MW - 42,248 Da / Chain length - 375 AA / Theoretical pI - 8.63					
DOMAIN	**POSITION**	**AMINO ACID SEQUENCE**	**LENGTH**	**MW**	**pI**
N - Extracellular	1 – 62	MDVTSQARGV GLEMYPGTAQ PAAPNTTSPE LNLSHPLLGT ALANGTGELS EHQQYVIGLF LS	62	6453.2	4.56
Helical; **TM1**	63 – 84	CLYTIFLF PIGFVGNILI LVVN	22	2469.0	5.52
1st Cytoplasmic	85 – 96	ISFREK MTIPDL	12	1449.7	6.07
Helical; **TM2**	97 – 120	YFIN LAVADLILVA DSLIEVFNIH	24	2703.1	4.02
1st Extracellular	121 – 132	ERYYDIAVLC TF	12	1492.7	4.37
Helical; **TM3**	133 – 153	MSLFLQVN MYSSVFFLTW MSF	21	2579.0	5.27
2nd Cytoplasmic	154 – 175	DRYIALA RAMRCSLFRT KHHAR	22	2673.1	11.44
Helical; **TM4**	176 – 194	LSCGL IWMASVSATL VPFT	19	1996.4	5.52
2nd Extracellular	195 – 220	AVHLQH TDEACFCFAD VREVQWLEVT	26	3047.4	4.49
Helical; **TM5**	221 – 236	LGFIVPFAII GLCYSL	16	1726.1	5.52
3rd Cytoplasmic	237 – 259	IVRV LVRAHRHRGL RPRRQKALR	23	2844.4	12.85
Helical; **TM6**	260 – 280	M ILAVVLVFFV CWLPENVFIS	21	2440.0	4.00
3rd Extracellular	281 – 306	VHLLQRTQPG AAPCKQSFRH AHPLTG	26	2851.2	10.86
Helical; **TM7**	307 – 327	HIVN LAAFSNSCLN PLIYSFL	21	2336.7	6.73
C - Cytoplasmic	328 – 375	GET FRDKLRLYIE QKTNLPALNR FCHAALKAVI PDSTEQSDVR FSSAV	48	5437.1	8.16

R, K (+VE AA) H (10% +VE AA) D, E (-VE AA) A, F, I, L, M, P, V, W, Y (Hydrophobic AA)
N (Potential Glycosylation - 25, 32, 44) C (Disulfide bond) – Reported 130 ↔ 207

Five short peptide epitopes **(Table 4.2)** were selected for synthesis considering their domain (promoting NP accessibility), charge (electrostatic interaction), hydrophobicity, sequence length, and isoelectric point (pI).

Table 4.2. Selected short peptide epitopes for Solid Phase Peptide Synthesis.

P	Domain	Position	Amino Acid Sequence	L	MW	pI
1	3rd Cyto	237-259	IVRVLVRAHRHRGLRPRRQKALR	23	2844.4	12.85
2	2nd Cyto	154-175	DRYIALARAMRCSLFRTKHHAR	22	2673.1	11.44
3	3rd Extra	281-306	VHLLQRTQPGAAPCKQSFRHAHPLTG	26	2851.2	10.86
4	N-Extra	47-62	GELSEHQQYVIGLFLS	16	1820.0	4.51
5		1-24	MDVTSQARGVGLEMYPGTAQPAAP	24	2447.7	4.37

Note: Peptides (P1, P2, P3) – hydrophobic with slightly positive charge

Peptides (P4, P5) – hydrophobic with slightly negative charge

AAPPTec Apex 396 parallel synthesizer was used to synthesis peptide epitopes (truncated-GPER sequences). Fluorenylmethyloxycarbonyl chloride (FMOC-Cl) **(Fig. 4.3)** protected amino acids were used for solid phase peptide synthesis (SPPS).

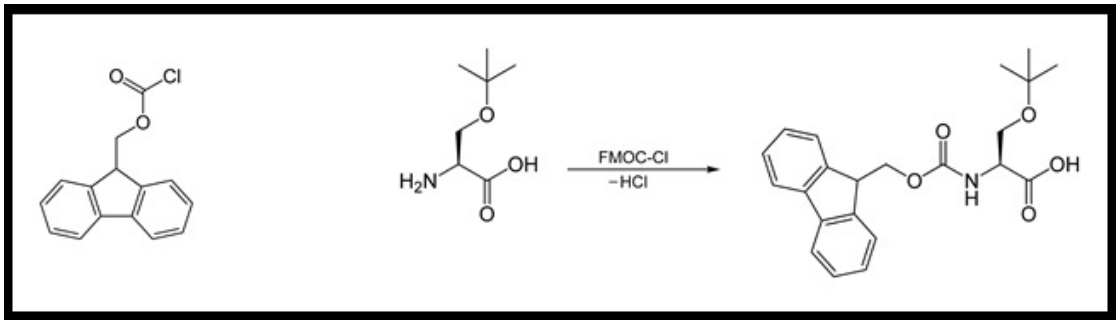


Figure 4.3 FMOC-Cl (left) and reaction showing FMOC introduction to amine (right).

Amino acid sequence were raised on a solid support, Nova-PEG rink amide resin (0.53 mmol/g loading, Novabiochem) (**Fig. 4.4**). 100 mg of resin was used for each peptide. Chemicals such as piperidine, dimethylformamide (DMF) and TFA were used for

deprotection, coupling and final cleavage, respectively. The amino acid sequences, [swell - (deprotect - wash deprotect - couple - wash couple)$_n$ - final wash] commands and number of programmed cycles (sequence length) were manually set in the AAPPTec software program.

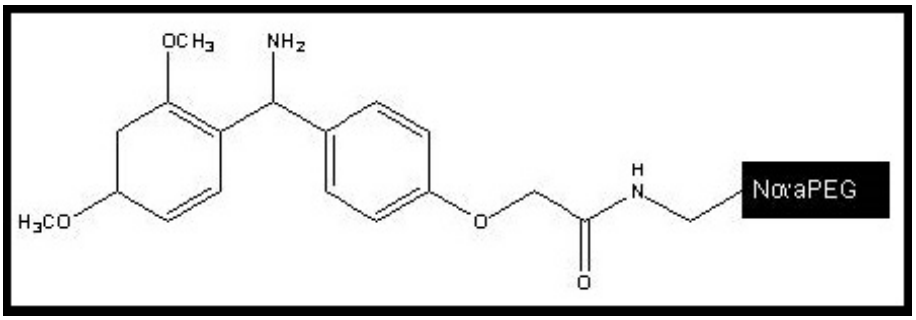


Figure 4.4 Nova-PEG Rink Amide Resin (Novabiochem).

Based on the preset parameters and (deprotecting - coupling) cycles, the automated peptide synthesizer synthesized the peptides utilizing the amino acid feed on continuous mode. Finally, based on the clevage information shown in **(Fig. 4.5)**, the resin was cleaved off from the synthesized peptides with TFA solution containing appropriate scavengers like water, ethanedithiol (EDT), thioanisole (TIS) **(Table 4.3)**.

Table 4.3 Cleavage cocktail used for synthesized peptides.

Peptide	*Sensitive amino acids* *Arg(Mtr),Cys(Trt), Met, Trp*	*Cleavage cocktail* *TFA + suitable scavengers*
P1	No Cys & Met	TFA 95% + H_2O 2.5% + TIS 2.5%
P2	Contain Cys & Met	TFA 94% + H_2O 2.5% + EDT 2.5% + TIS 1%
P3	Contain Cys	TFA 94% + H_2O 2.5% + EDT 2.5% + TIS 1%
P4	No Cys & Met	TFA 95% + H_2O 2.5% + TIS 2.5%
P5	Contain Met	TFA 94% + H_2O 2.5% + EDT 2.5% + TIS 1%

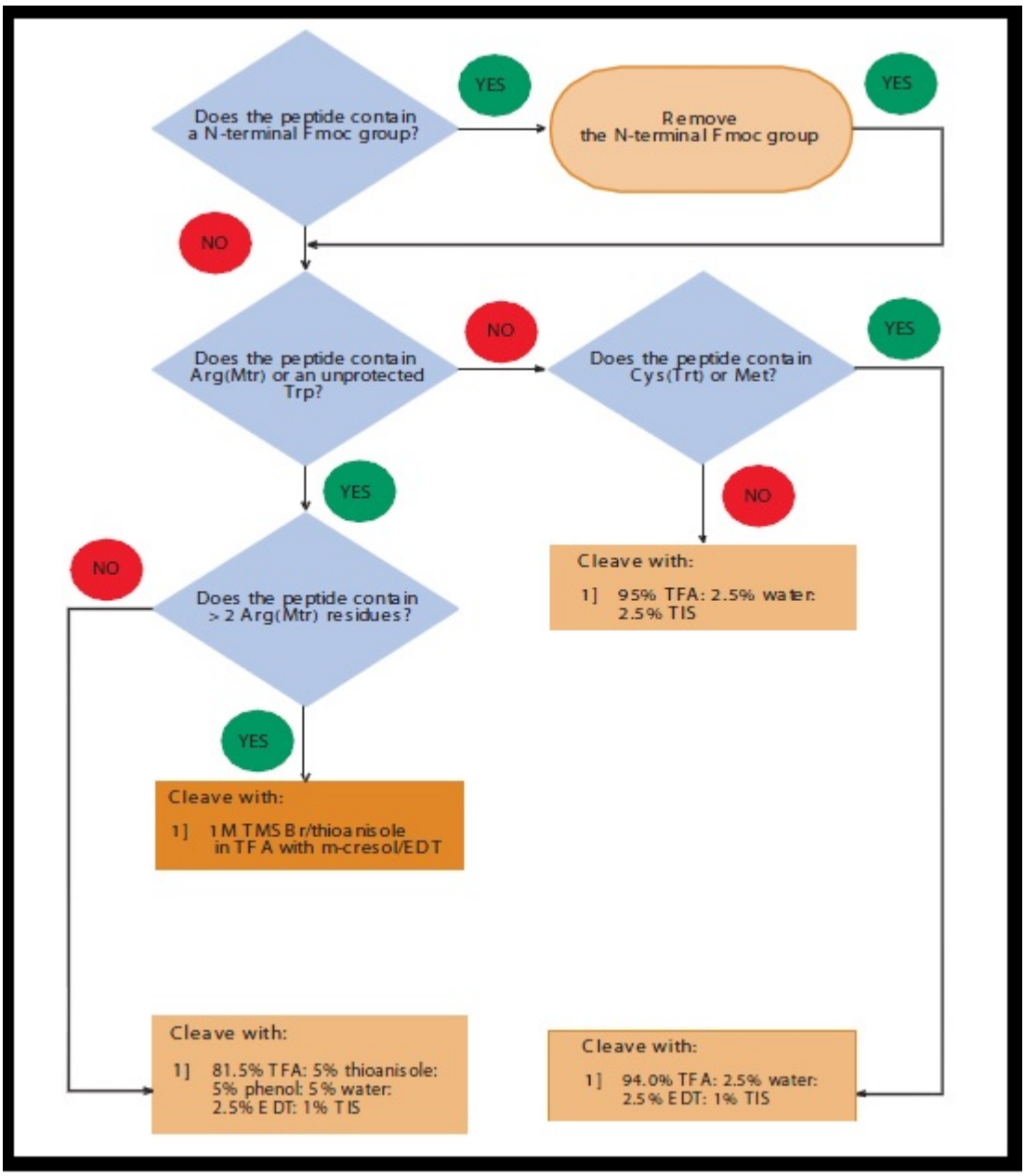


Figure 4.5 Flow-chart for selecting cleavage cocktail for FMOC SPPS.[281]

The resin was removed by filtration under reduced pressure and washed twice with 100% TFA. The filtrates were pooled together and added on an 8-10 fold volume of ice-cold diethy lether. The precipitated peptides were recovered after high-speed centrifugation, dissolved in suitable aqueous buffer and lyophilized. The % yield **(Table**

4.4) was calculated by comparing the dry mass of the product obtained to the theoretical yield calculated from the following equation

Theoretical yield (mg) = s.resin * m.resin * MW product

where, s.resin - resin substitution in mmol/g

m.reisn - resin dry mass in g

MW product - molecular weight of the peptide in mg/mmol

Table 4.4 % yield of the synthesized peptides (P1, P2, P3, P4 and P5).

Peptide	MW (Da)	Theoretical yield (mg)	Dry mass (mg)	Yield (%)
P1	2844.4	150.75	112	74.30
P2	2673.1	141.71	83	58.57
P3	2851.2	151.11	116	76.77
P4	1820.0	96.46	79	81.90
P5	2447.7	129.73	98	75.54

4.4 HPLC and MALDI MS analysis of synthesized peptides:

The analytical instruments such as high performance liquid chromatography (HPLC, Waters Corp.) and matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS, AB Sciex) were used to check the purity **(Fig. 4.6)** and verify the mass **(Fig. 4.7)** of synthesized peptides, respectively.

A small amount of lyophilized peptides (P1, P2, P3, P4 and P5) were dissoved separately in 500 uL of water/acetonitrile (50:50). All the sequence parameters and method were created and loaded on the HPLC software program. The sample injection volume was set at 1.5 mL/min. Then, 100 uL of each peptide sample was injected and analysed using gradient elution method from 0% to 60% (increasing gradient of acetonitrile) over 30 mins at 220 nm wavelength.

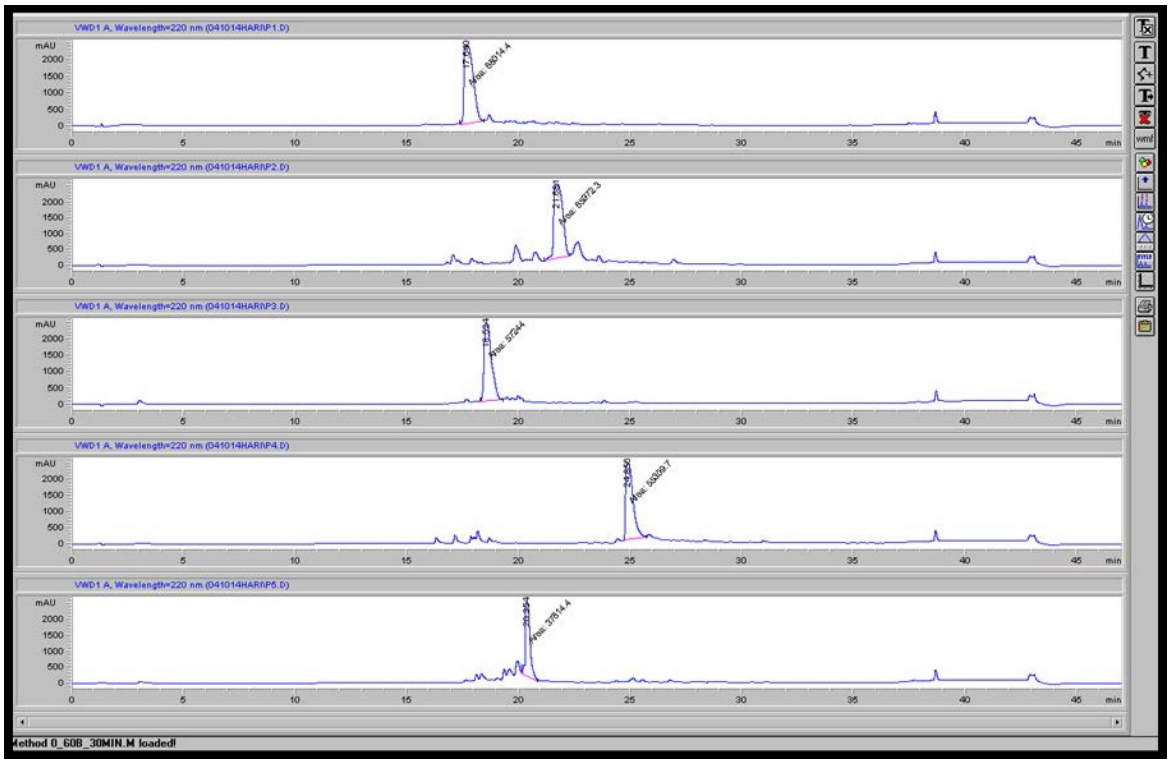


Figure 4.6 HPLC chromatograms of synthesized peptides (P1, P2, P3, P4 and P5).

MS experiments were carried out on AB SCIEX TOF/TOF™ 5800 System equipped with a 1 kHz variable rate laser. MS data were acquired at a laser repetition rate of 400 Hz with 4000 laser shots/spectrum (100 laser shots/sub-spectrum) with a mass accuracy of 50 ppm. CHCA was used as matrix. 10 mg of α-CHCA was dissoved in 1 mL of matrix solution (50:50 water/acetonitrile with 0.1% TFA) in a clean Eppendorf. The peptide samples prepared for HPLC was diluted with α-CHCA matrix solution in the ratio of 1:5, 1:10 or 1:20. 1 uL of final mix was spotted on to MALDI target, air-dried and analyzed.

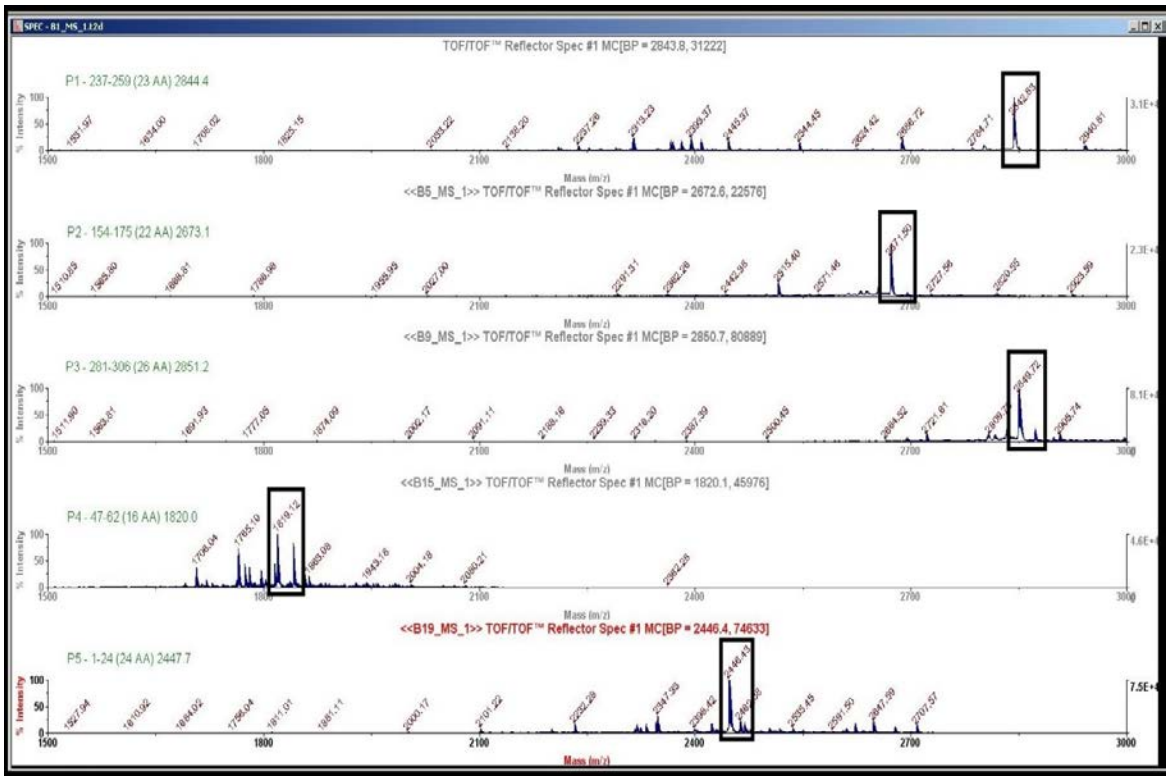


Figure 4.7 MALDI TOF MS spectra of synthesized peptides (P1, P2, P3, P4 and P5).

4.5 Preparation of synthetic polymer nanoparticles:

The nanoparticles were synthesized by a free radical precipitation polymerization method with no control over the functional monomer sequence **(Fig. 4.8)**. The selection of functional monomers were based on negative charge and hydrophobicity to complement the synthesized peptide epitopes. The nanoparticles were raised on a poly-N-isopropylacrylamide (PNIPAm)-based polymer backbone with 2% of a cross linker. Monomers such as acrylic acid (AAc, 5 mol%), N-tert butylacrylamide (TBAm, 40 mol%), N-isopropylacrylamide (NIPAm, 53 mol%), N,N'-methylene bis(acrylamide) (BIS, 2 mol%) were dissolved in 50 mL of water and filtered through Whatman filter paper before transferring the monomer solution to round bottom flask. TBAm was dissolved separately in 1 mL of ethanol before addition. The concentration of total monomer feed was calculated to be 65 mM. Nitrogen gas was purged into the monomer solution for 30 mins prior to initiating the polymerization reaction by addition of ammonium persulfate (APS,

30 mg in 500 uL of D.I. water). Following the addition of APS, the polymerization reaction was carried out at 60 °C for 3 hrs under nitogen atmosphere. The polymerization was stalled by the exposure of atmospheric oxygen to the reaction mixture. The polymerized solutions were purified by dialysis (using dialysis membrane with MWCO 12,000 – 14,000 Da) against an excess of pure water (changed at least twice a day) for ≥ 4 days. The nanoparticles with smaller diameter and/or controlled size can be prepared in similar method by adding few mg of SDS.

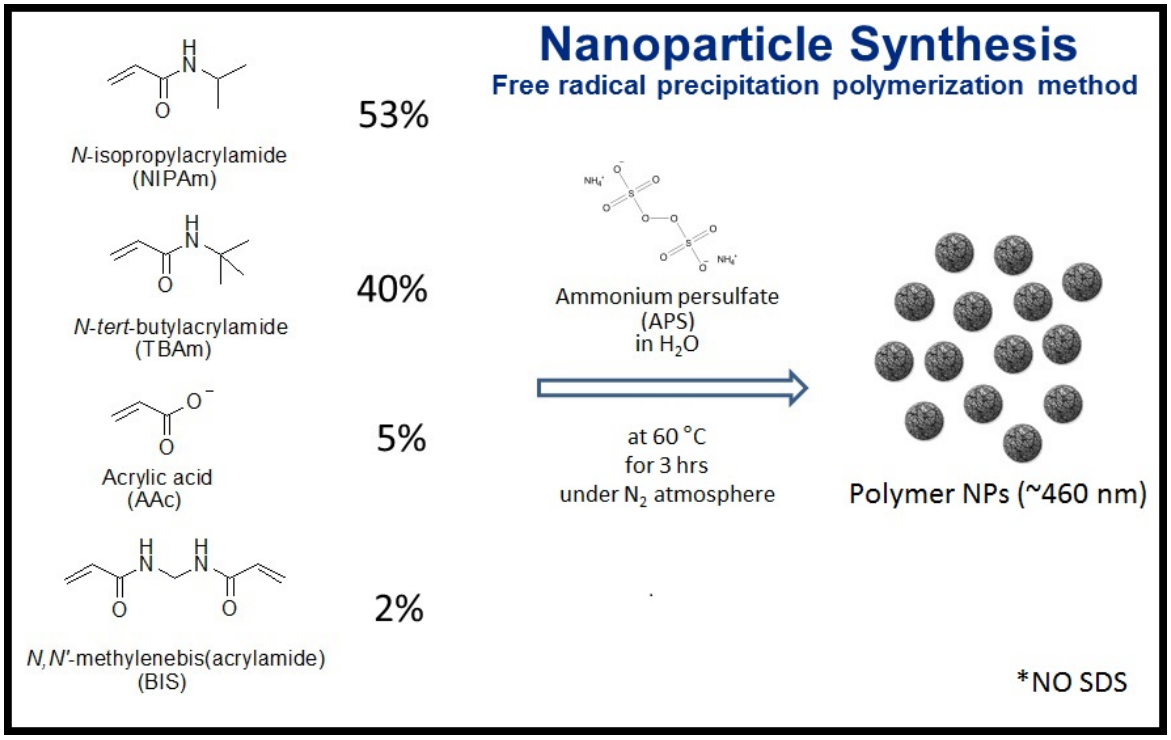


Figure 4.8 Preparation of PNIPAm-based synthetic polymer nanoparticles.

4.6 Nanoparticle characterization:

The hydrodynamic diameter (D_H) of nanoparticles was determined to be 462 nm with 0.007 PdI in d.H_2O at 25 °C by dynamic light scattering (DLS) instrument supported with Zetasizer software (Zetasizer Nano ZS, Malvern Instruments Ltd.). DLS results were obtained by following the manufacturer's protocol and all the results of DLS data fitting met the quality criteria set by Malvern. The yield (69.66%) and concentration (5.2 mg/mL)

of nanoparticles were determined by gravimetric analysis of lyophilized nanoparticles. 20 mg of lyophilized polymers were dissolved in 700 ul CD_3OD and used for NMR spectroscopy measurements. ^1H NMR and ^{13}C NMR were measured using Bruker DRX500 spectrometer with TCI (three channel inverse) cryoprobe. All measurements were run at 298 K and the peak of residuals CD_2HOD (δ 3.31 ppm for ^1H) and $^{13}CD_3OD$ (δ 49.15 ppm for ^{13}C) were used as a reference. The ^1H NMR and ^{13}C NMR characterization results were shown in **Fig. 4.9** and **Fig. 4.10**. The lightly cross-linked NPs had considerable chain flexibility as evidenced from sharp lines in their solution ^1H NMR **(Fig. 4.9)**.

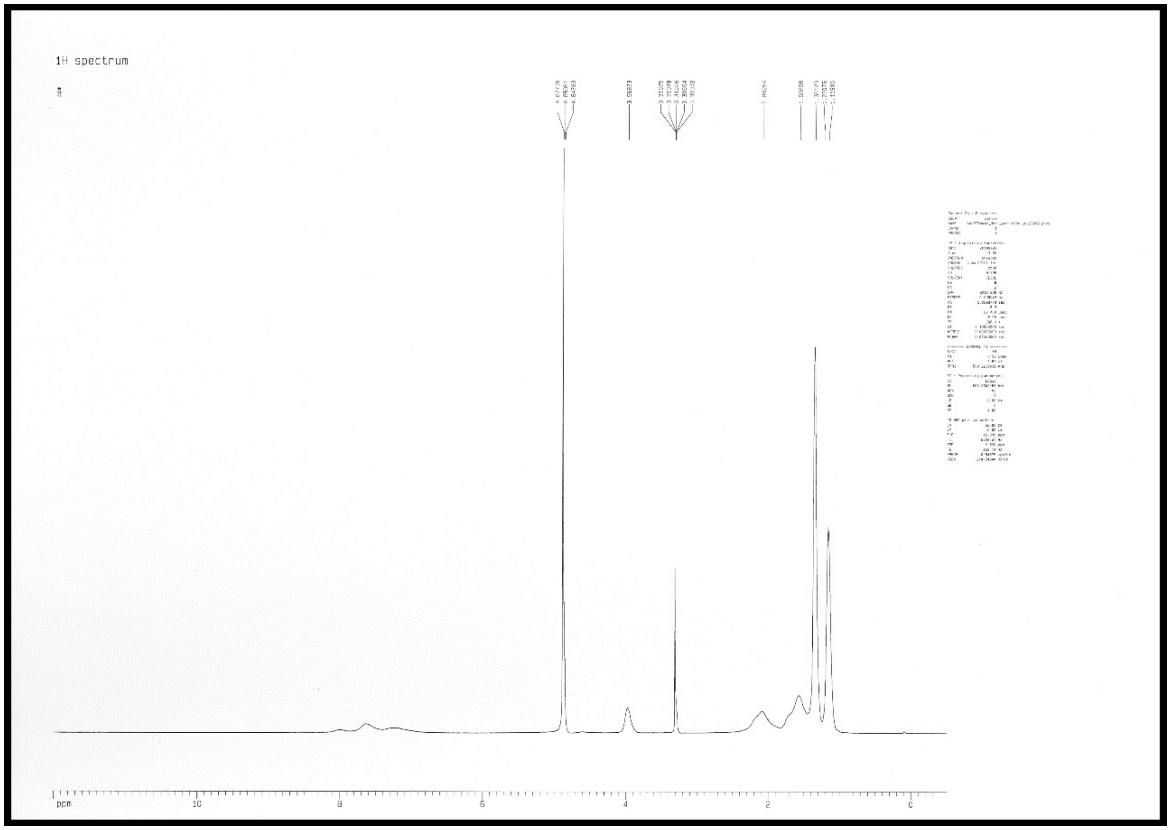


Figure 4.9 ^1H NMR spectrum of 462 nm NP in CD_3OD, 500 MHz, 298 K.

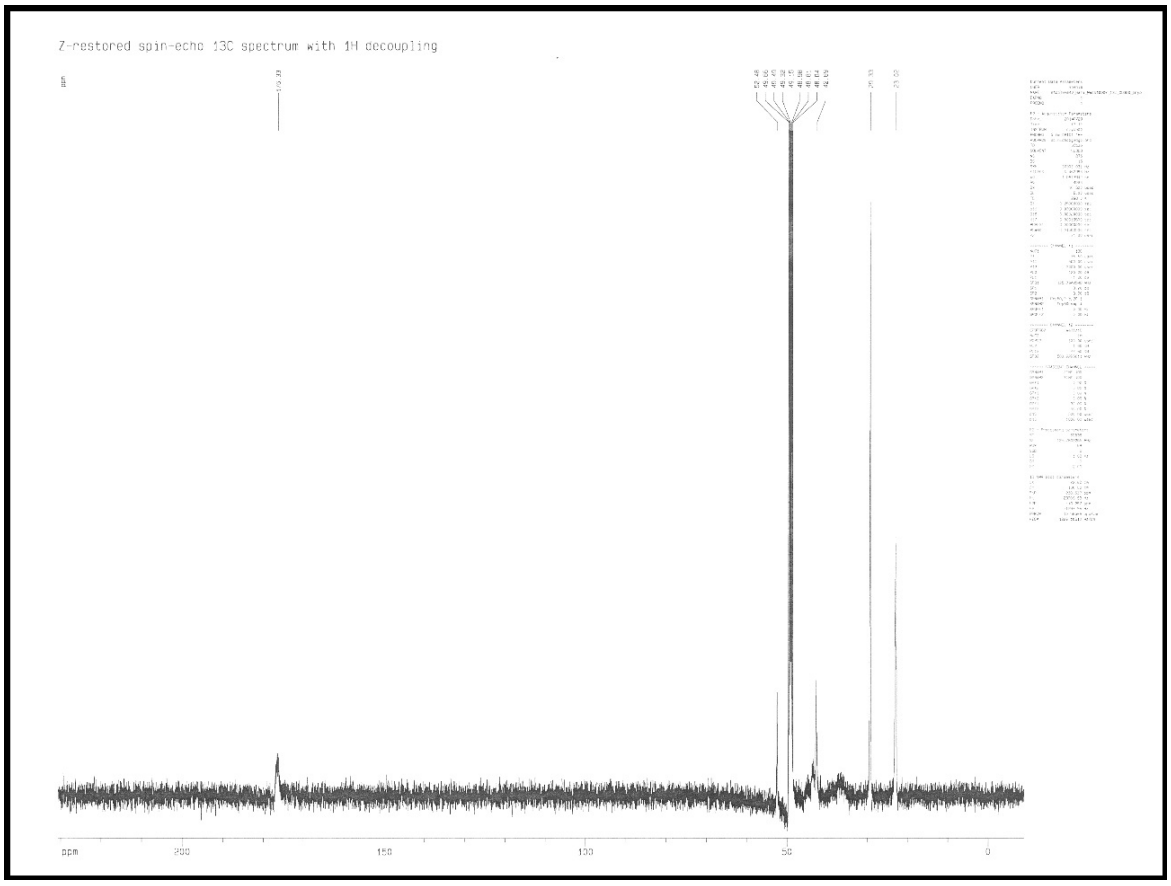


Figure 4.10 ^{13}C NMR spectrum of 462 nm NP in CD_3OD, 500 MHz, 298 K.

4.7 Interaction between peptides and nanoparticles by HPLC analysis:

Peptides (P1, P2 and P3) **(Fig. 4.11)** having rich hydrophobic and a decent number of positively charged residues were selected for peptide-NP interaction studies. To complement the peptides during interaction, the synthesized NPs contain hydrophobic and carboxylate group monomers incorporated on a PNIPAm backbone with 2% of a cross-linker giving considerable chain flexibility.

To determine the working concentration of nanoparticles, two different dilutions (i) 400 ug/mL (NP1, lowest conc.) and (ii) 2 mg/mL (NP2, highest conc.) were made from a stock solution of NPs having a concentration of 6.83 mg/mL. Similarly, 5 mM of peptide stock solutions were made from the lyophilized peptides (P1, P2 and P3) and diluted to

have three different concentrations, 10 uM, 20 uM and 50 uM to set the working concentration of peptides.

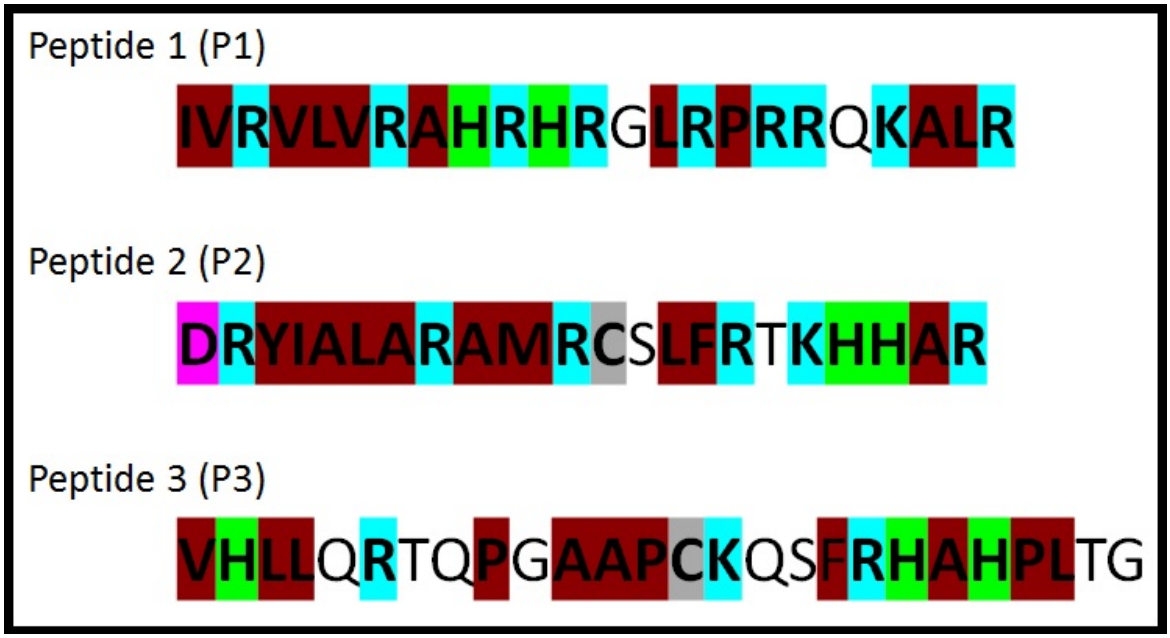


Figure 4.11 Selected peptides for Peptide-NP interaction studies.

For the peptide-NP binding experiment, 250 uL of NPs (NP1 & NP2, separately) were made to interact with 200 uL of peptides (P1, P2 & P3, separately) in the presence of 50 uL of 10X Phosphate Buffered Saline / Phosphate Buffer without salt (to supplement physiological pH ~7.3), as shown in **(Fig. 4.12)**. The NP-peptide mix was incubated at room temperature for 15 mins and centrifuged at 8000 g for 20 mins. 100 uL of supernatant from each mix was taken and injected into HPLC and analyzed for drop in intensity.

Note: If the peptide binds the NP, then most of the peptide will be bound and sedimented along with the NP in the eppendorf leaving the excess and/or unbound peptide in the supernatant. So, there will be a fall in peptide intensity on the HPLC chromatogram when compared against the reference peptides (standard).

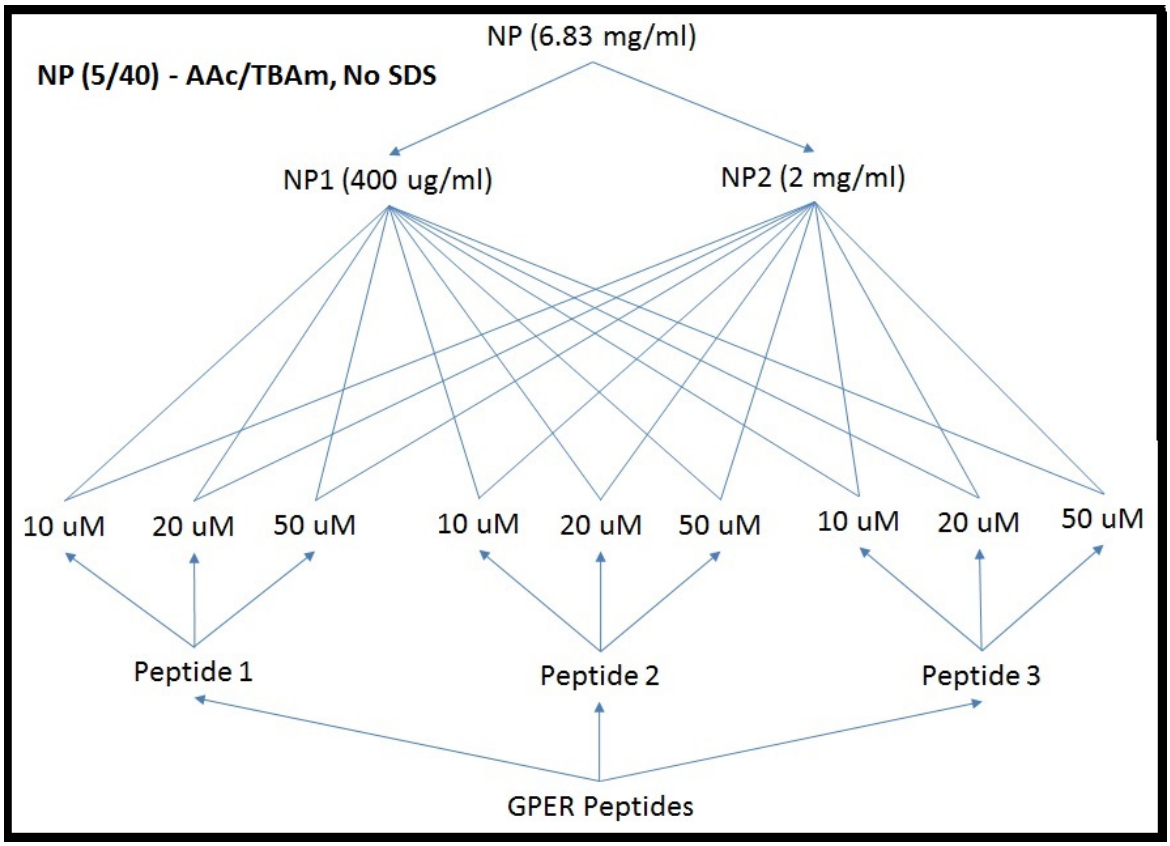


Figure 4.12 Peptide-NP interaction chart.

All three peptides showed notable binding affinity towards the NPs with P1 and P2 being the best **(Fig. 4.13)**. The above experiment was repeated several times by altering the concentrations of both peptides and NPs. Finally, 20 uM concentration of peptides were found to have comparitively strong binding affinity towards 2 mg/mL concentration of NPs as evidenced from their steep fall in peptide intensity on the HPLC chromatogram, and hence, considered as the working concentrations for similar experiments thereon.

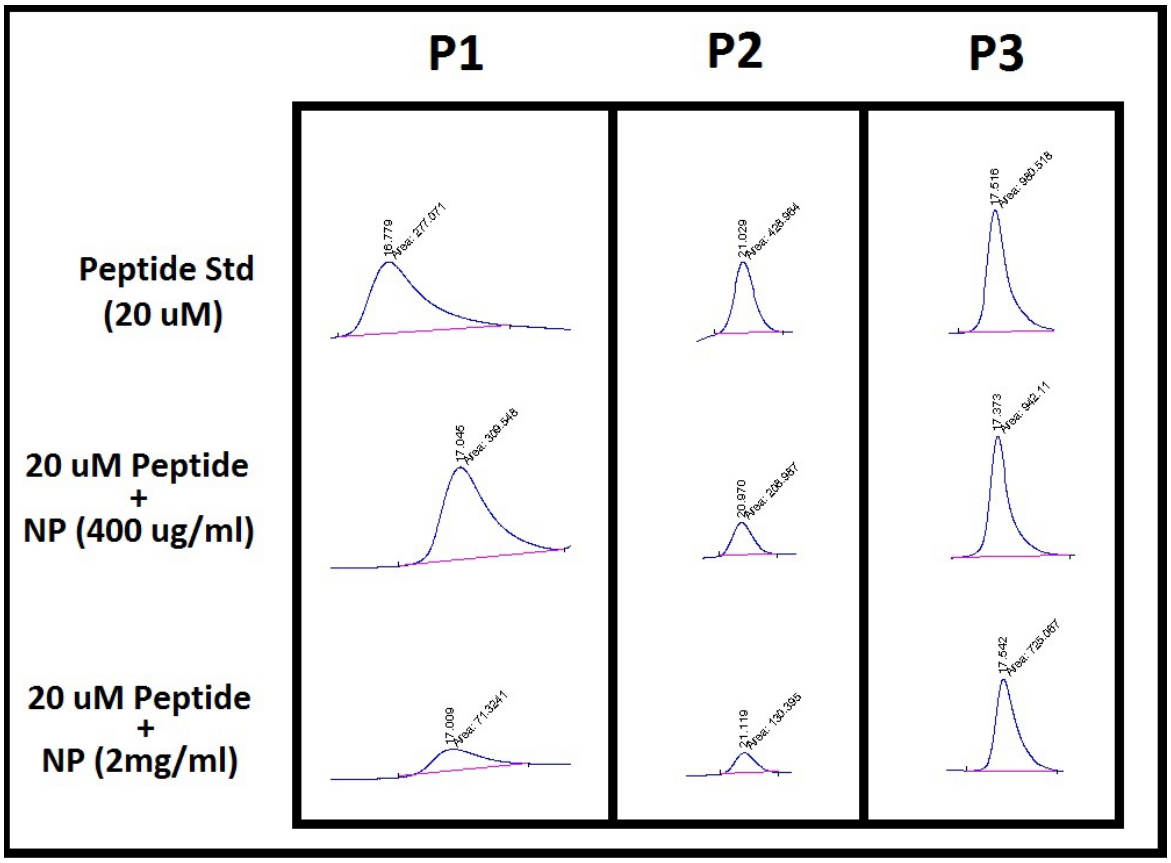


Figure 4.13 HPLC analysis of Peptide-NP binding affinity.

At this stage, these truncated-GPER peptides (P1 and P2) can be used as epitope for molecular imprinting. If imprinted, then the peptide imprinted polymer nanoparticles could be used as plastic antibodies for GPER. Plastic antibodies, also known as molecularly imprinted polymer nanoparticles (MIPs), are generic alternatives to antibodies that can recognize and capture the target biomacromolecules.[282][283] Due to time constraints, we moved on to protein-NP interaction studies without proceeding towards peptide imprinting.

4.8 Cell culture and cell lysis:

Four different cancer cell lines were used in the presented work. BG-1 ovarian cancer cells and Ishikawa endometrial cancer cells were maintained in DMEM without phenol red supplemented with 10% fetal bovine serum (FBS). Whereas, SkBr3 breast

cancer cells were maintained in RPMI 1640 without phenol red supplemented with 10% FBS. HEK293 cells were overexpressed with hGPER gene and maintained in DMEM high glucose supplemented with 10% FBS. All cells were grown in 10 cm petri dishes, and then lysed in 500 µL of 50 mM HEPES (pH 7.5), 150 mM NaCl, 1.5 mM MgCl2, 1 mM EGTA, 10% glycerol, 1% Triton X-100, 1% SDS, and a mixture of protease inhibitors containing 1 mM aprotinin, 20 mM PMSF and 200 mM sodium orthovanadate. The protein concentration **(Table 4.5)** was determined using Bradford or Bicinchoninic acid (BCA) protein assay dye reagent. The cell lysates were divided into small aliquots and stored at -20 °C for further use.

Table 4.5 Cell lysates and their total protein concentration.

Cell line	Type	Expression	Concentration
SKBR3	Breast cancer	ERα (-), GPER (+)	3.53 mg/ml
HEK293	hGPER over expressed	GPER (+) ↑	3.90 mg/ml
BG-1	Ovarian Cancer	ERα (+), GPER (+)	2.37 mg/ml
ISHIKAWA	Endometrial Cancer	ERα (+), GPER (+)	2.37 mg/ml

4.9 Interaction between protein and nanoparticles by WB analysis:

SDS-PAGE Electrophoresis: 12% sodium dodecyl sulfate-polyacrylamide precast gels form Bio-Rad were used for electrophoresis. Protein lysates treated with NPs were resolved simultaneously along with experimental controls and standard protein marker. After electrophoresis, the gel was stained with Coomassie blue stain solution from Bio-Rad followed by destaining with destain solution containing 40% CH_3OH, 10% CH_3COOH to 1 L d.H_2O.

Western blot analysis: To determine the working concentration of total protein content from cell lysates, equal volumes of the cell lysate from all four cell lines with varying dilutions were resolved on a 12% SDS-polyacrylamide precast gel and transferred to a nitrocellulose membrane. The membrane was then blocked with 5% non-fat milk

(blocking buffer) for 1 hr at room temperature under agitation, followed by probing the membrane overnight at 4 °C against two different primary antibodies [rabbit anti- GPER (*N*-terminal) polyclonal antibody and mouse anti-b-actin polyclonal antibody] under agitation. After overnight incubation, the unbound primary antibodies were washed off and the membrane was incubated with secondary antibodies [goat anti-rabbit antibody (GPER) and goat anti-mouse antibody (β-actin) conjugated with green (800 channel) and red (700 channel) fluorescent dyes, respectively] for 1 hr at room temperature under agitation. The processed blot was washed off from excess antibodies and visualized using Odyssey infrared imaging system **(Fig. 4.14**).

Notes: 1, the Odyssey imaging system detects infrared florescence from fluorochrome-conjugated secondary antibodies. 2, the use of two different antibodies in a same blot helps in distinguishing protein binding affinity (comparative analysis). 3, the molecular weight of β-actin (~42 kDa) is almost similar to that of GPER, hence, it is used as the second primary antibody.

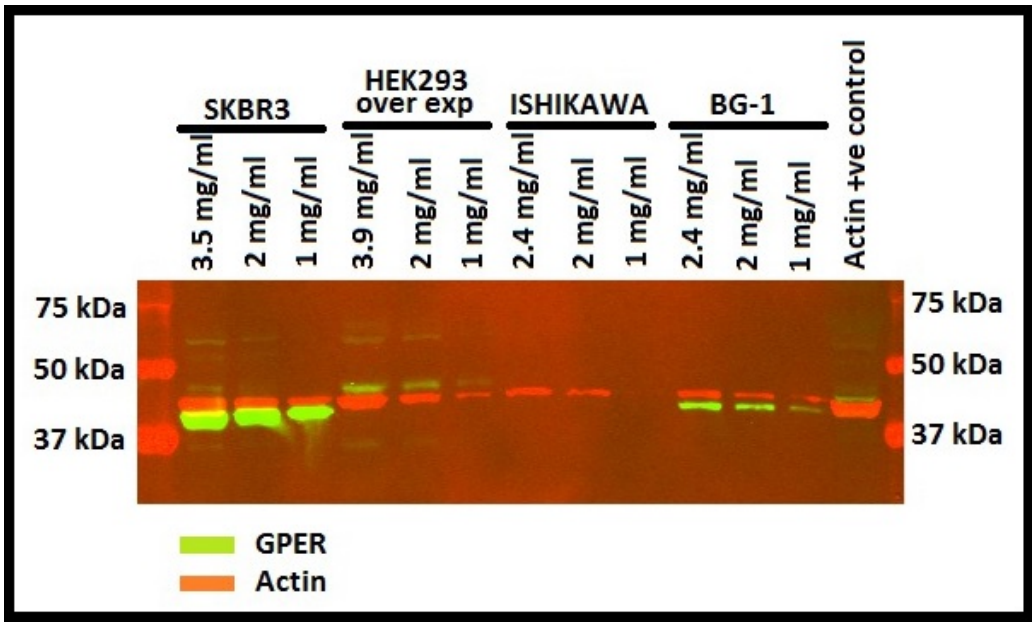


Figure 4.14 Immunoblot showing detectable protein concentrations.

To determine the working concentration of 5/40 (AAc/TBAm) nanoparticles, three different dilutions (1 mg/mL, 2 mg/mL and 3 mg/mL) were made in a similar way as mentioned in peptide-NP interaction studies. Likewise, various dilutions (0.25 mg/mL, 0.5 mg/mL and 1 mg/mL) of protein were prepared from SkBr3, hGPER (overexpressed) and BG-1 cell lysates. The protein-NP binding experiments were set in physiological pH ~7.3 supplementing PBS. The protein-NP mix was incubated at room temperature for 15 mins and centrifuged at 8000 g for 20 mins. 10 uL of supernatant from each mix was taken and analyzed by electrophoresis and Western blotting technique **(Fig. 4.15)**.

Note: If GPER binds the NP, then most of the bound GPER will be found sedimented along with the NP in the eppendorf leaving the unbound proteins in the supernatant. So, there will be a weak fluorescence intensity (green, 800 channel) on the imaged blot when compared against the standards.

The above experiment was repeated several times by altering both protein and NP concentrations. In conclusion, 0.5 mg/mL concentration of cell lysates were found to have notable GPER binding affinity towards 3 mg/mL concentration of 5/40 NPs as evidenced from their week fluorescence intensity on the western blot **(Fig. 4.15)**, and hence, considered as the working concentrations for protein-NP binding studies.

Few control experiments were also performed with 0/40 (AAc/TBAm) and 20/40 (AAc/TBAm) NPs **(Fig. 4.15)**. In the case of 20/40 NPs, 100 K MWCO centrifugal filters were used after 15 mins incubation of protein-NP mix. As a result of centrifugal filtration, most of the proteins remained on the filter membrane and failed to pass through. Whereas, the 0/40 NPs had no binding affinity towards any proteins as they completely lack carboxylate group monomers that are believed to have electrostatic contribution during macromolecular interactions. Among all three NPs (0/40, 5/40 and 20/40), 5/40 particle emerged as the winner in binding GPER from crude cell lysates (SkBr3 and BG-1) with competitive affinity as evidenced from different immunoblots **(Fig. 4.15)**.

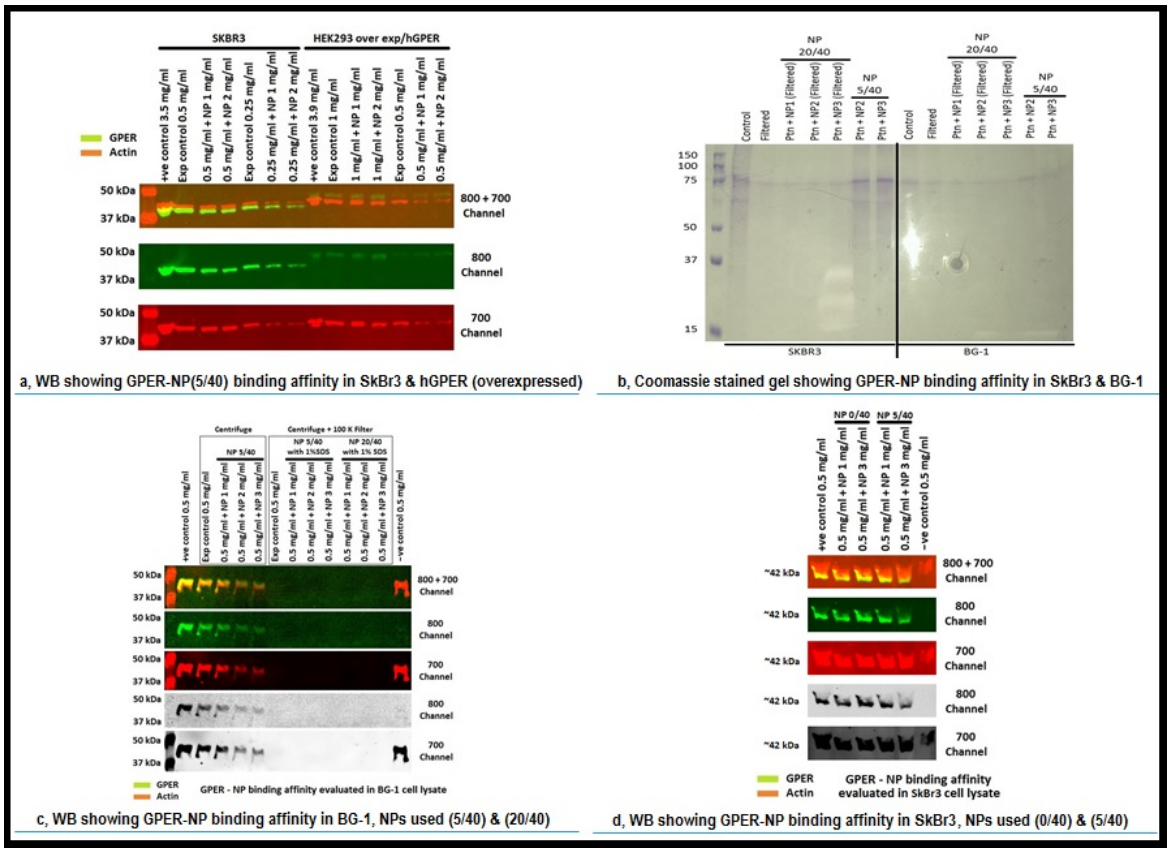
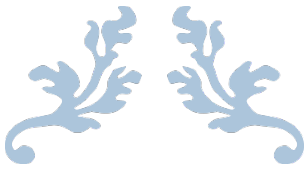


Figure 4.15 Western blot analysis of Protein-NP binding affinity.

As the NPs were designed with complementary functionality against the peptides/ protein, the NP-peptide/protein binding were through multipoint interactions. The qualitative results obtaibed from the immunoblot analysis revealed some interesting hints on GPER's competitive affinity towards NPs when probed against multiple antibodies. Peptides (P1, P2 and P3) showed notable binding affinty towards 5/40 (AAc/TBAm) NPs with P1 & P2 being the best. These could be used as peptide epitopes for imprinting to develop plastic antibodies. Further, the 5/40 (AAc/TBAm) NPs captured GPER with competetive affinty over other proteins in crude cell lysates, and thus, could be immobilized on agarose and used for GPER affinity purification. Moreover, these polymer-based NPs were reported to have the capability to 'catch and release' a target protein in a reversible and temperature-sensitive manner.[269] Below lowest critical solution temperature (LCST) these particles gradually swell and release the captured

protein. We anticipate to use this strategy as a sample purification step before MS-based proteomic analysis.



CHAPTER 5

SUMMARY & OUTLOOK

All living organisms produce hormones which act as chemical messengers in transmitting signals between cells. Hormones, including oestrogens, bind to specific receptors leading to the activation of multiple transduction pathways. Numerous physiological processes in mammals are influenced by estrogen and the estrogen receptors. Estrogen is implicated in the development or progression of a number of diseases, such as human cancers, endometriosis, fibroids, and cardiovascular diseases. The biological actions of estrogen are mainly mediated by classical oestrogen receptors, ERα & ERβ that belongs to the nuclear receptor superfamily. In recent years, a class of membrane-associated oestrogen receptors are found to mimic the functions of classical ERs, including genomic (transcriptional) as well as non-genomic (rapid) signaling. These non-genomic signaling events include pathways that are usually thought of as arising from transmembrane growth factor receptors and G protein-coupled receptors (GPCRs). Almost half of all medications achieve their effect through GPCRs. A member of the GPCR family, G Protein-coupled Estrogen Receptor 1 (GPER) formerly known as GPR30, is found to mediate rapid biological responses to estrogen in diverse normal and cancer cells, as well as transformed cell types. GPER acts independently of ERα & ERβ and triggers ER-dependent EGFR action. It is believed to play a significant role in the development of tamoxifen resistance in breast cancer. GPER also involves in several other cancers, but the background information at molecular level is still limited. Hence, it is of vital interest to understand the nature and behaviour of GPER signaling mechanisms.

GPER is a member of the rhodopsin-like family of G protein-coupled receptors and is a multi-pass membrane protein that localises predominantly to the endoplasmic reticulum. Estrogen binding to GPER, stimulates cAMP production, intracellular calcium mobilisation and indirectly activates MAPK and PI3K-Akt pathways that can induce additional non-genomic effects, or genomic effects regulating gene transcription. According to UNIPROT (a comprehensive protein database), GPER comprises of 375 amino acids with a molecular mass of about 42 kDa.

Initial experiments concentrated on the isolation and identification of GPER from total protein lysate obtained from different cancer cells, such as SkBr3 (breast cancer), MCF-7 (breast cancer), Ishikawa (endometrial cancer), and BG-1 (ovarian cancer). A suitable protocol was established for protein extraction, purification, and MALDI MS and MS/MS analyses. GPERs are difficult to extract from gels post SDS-PAGE electrophoresis, as they are transmembrane receptors. Also, there is no proven protocol published till date to isolate and enrich GPCRs from complex protein mixture. Therefore, we focussed on developing a method to achieve the same. For GPER isolation, the total protein extract is subjected to spin column containing hydroxyapatite and the bound proteins are eluted with buffers at varying pH. The eluate is concentrated and subjected to SDS-PAGE, followed by WB analysis to check the efficiency and consistency of the proposed HTP enrichment method prior to MS analysis. For GPER identification, the concentrated eluate is filtered through Cellulose Acetate membrane spin filter. Then, the CA membrane is removed and the proteins are digested on-membrane using different proteases, such as trypsin, pepsin, and α-chymotrypsin. The protease digested samples were analysed using MALDI MS and MS/MS. MASCOT search (search engine which uses mass spectrometry data to identify proteins from primary sequence database) revealed the presence of GPCR family proteins. Protein Pilot, a software package is also used to identify and quantify the peptides found. We were able to identify GPER by peptide mass fingerprinting with top score. However, this approach is not only limited to GPER, but can also be applied to study several other GPCRs. The developed proteomics workflow is very simple, unique and cost-effective.

Secondly, we focussed on studying all the possible post-translational modifications (PTMs) in GPER. As per UNIPROT, GPER has 3 potential glycosylation sites and 1 disulfide bond. In order to execute glycosylation studies, the peptides mixture are derivatized by dansylation and directly analysed by MALDI MS and MS/MS. From this experiment we manually validated 2 out of 3 glycosylation sites, one from pepsin digestion and the other from α-chymostrypsin digestion, by matching the MALDI MS/MS

spectra of natural peptides generated by protease digestion against dansylated peptides. All the possible glycan compositions were predicted using Glycomod tool and validated manually from the obtained MS/MS data. In addition, we used PNGase F and Endoglycosidase H for peptide deglycosylation to perform a comparative analysis.

In early 2014, we established a new project on affinity capture-release strategy for GPER purification in collaboration with Shea Lab (UCI). The lack of promising separation and purification tools peaked our interest towards synthetic polymer nanoparticles that can capture this target biomolecule. We successfully designed and synthesized polymer nanoparticles with high affinity for GPER.

Recently, NPs with an intrinsic affinity have shown to be successful in binding biomacromolecules like melittin, immunoglobulin G, histone, fibrinogen and lysozyme by controlling and optimizing the functional monomers composition. We adopted a similar approach in an effort to capture GPER with high affinity and selectivity among a mixture of proteins that are expressed in cancer cells (SkBr3 & BG-1). Considering the fact that GPER is a membrane protein with many hydrophobic amino acid residues on the surface, a candidate NP was chosen from a library of nanoparticles that were prepared by combining different populations of functional groups on a poly-N-isopropylacrylamide (PNIPAm)-based polymer backbone. The chosen NPs contain hydrophobic and carboxylate group monomers incorporated on a PNIPAm backbone with 2% of a cross-linker. The NPs were synthesized by a free radical precipitation polymerization method with no control over the functional monomer sequence. The lightly cross-linked NPs had considerable chain flexibility as evidenced from sharp lines in their solution ^{1}H NMR. The NP binding affinity was evaluated against both truncated-GPER (short peptide epitopes) and GPER (whole protein). As the NPs were designed with complementary functionality against the peptides/protein, the NPs-peptide/protein binding will be through multipoint interactions. The initial qualitative results were obtained by HPLC and immunoblotting analyses. The immunoblots revealed some interesting hints on GPER's competitive affinity towards NPs when probed against multiple antibodies. We anticipate to use this

strategy as a sample purification step before mass spectrometry-based proteomic analysis.

Complicated protein receptors like GPER could not be characterized to the fullest within the timescale of a normal PhD, three years. So, still there is much more to explore and understand about the receptor. The proposed methods for GPER isolation and purification, and the initial results reported in this thesis will definitely help to answer some of the puzzles related to GPER. As we spent much of the time on method development to isolate GPER from complex biological sample, very minimal results were reported on protein characterization. Though, GPER was identified by PMF, further validation by MS/MS peptide-sequencing is necessary. The mass spectrometry experiments were performed only on TPL obtained from SkBr3, so it would be an added advantage if the results are recreated from different TPL samples using the same method. We were able to validate 2 out of 3 glycosylation sites and the corresponding glycopeptides in GPER from our experimental data, it is necessary to validate the third glycosylation site in GPER. In future, research should be directed on studying all other PTMs and mutations in GPER among different cancer cells, and also their molecular interactions with other receptors and ligands by tandem mass spectrometry. The NP-peptide/protein binding studies were tested successfully, further studies should focus on releasing and eluting the purified GPER from NPs with suitable condition. Efforts should be made on to develop plastic antibodies for GPER by peptide imprinting technology.

BIBLIOGRAPHY

Bibliography

[1] D. L. Rubbelke, "Tissues of the Human Body," in *Understanding Human Anatomy & Physiology*, New York, McGraw-Hill Companies, 2000.

[2] D. Purves, G. J. Augustine, D. Fitzpatrick, L. C. Katz, A.-S. LaMantia, J. O. McNamara and M. Williams, "Chapter 8. Intracellular Signal Transduction: Receptor Types," in *Neuroscience, 2nd edition*, Sunderland, Sinauer Associates, Inc., 2001.

[3] H. Lodish, A. Berk, L. Zipursky, P. Matsudaira, D. Baltimore and J. Darnell, "20.1 Overview of Extracellular Signaling," in *Molecular Cell Biology, 4th edition*, New York, W H Freeman and Company, 2000.

[4] L. A. A. de Jong, D. R. A. Uges, J. P. Franke and R. Bischoff, "Receptor–ligand binding assays: Technologies and Applications," *Journal of Chromatography B,* vol. 829, no. 1-2, pp. 1-25, 2005.

[5] J. C. Matthews, "Receptor-ligand interactions that generate proportional physiological effects," in *Fundamentals of Receptor, Enzyme, and Transport Kinetics*, Boca Raton, CRC Press, Inc., 1993, pp. 33-40.

[6] C. M. O'Connor and J. U. Adams, "Essentials of Cell Biology," NPG Education, Cambridge, MA, 2010.

[7] N. L. Novére and J.-P. Changeux, "The Ligand Gated Ion Channel database: an example of a sequence database in neuroscience," *Philos Trans R Soc Lond B Biol Sci.,* vol. 356, no. 1412, pp. 1121-1130, 2001.

[8] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, "Signaling through Enzyme-Linked Cell-Surface Receptors.," in *Molecular Biology of the Cell. 4th edition.*, New York, Garland Science, 2002.

[9] Boundless, "Boundless Human Anatomy & Physiology," Boundless Learning, Inc., Boston, MA, 2013.

[10] T. P. Iismaa and J. Shine, "G protein-coupled receptors," *Current Opinion in Cell Biology,* vol. 4, no. 2, pp. 195-202, 1992.

[11] P. G. Strange, "G-protein Coupled Receptors: Conformations and States," *Biochemical Pharmacology,* vol. 58, no. 7, pp. 1081-1088, 1999.

[12] R. A. Bowen, "Mechanism of Action: Hormones with Intracellular Receptors," Colorado State University, Fort Collins, CO, 1998.

[13] P. H. Raven, G. B. Johnson, K. A. Mason, J. B. Losos and S. R. Singer, Biology, Ninth edition, New York: The McGraw-Hill Companies, Inc., 2011.

[14] M.-J. Tsai and B. W. O'Malley, "Molecular Mechanisms of Action of Steroid/Thyroid Receptor Superfamily Members," *Annual Review of Biochemistry,* vol. 63, pp. 451-486, 1994.

[15] H. P. Rang, M. M. Dale, J. M. Ritter, R. J. Flower and G. Henderson, Rang & Dale's Pharmacology, 7th Edition, London, UK: Elsevier Churchill Livingstone, 2012.

[16] D. E. Golan, A. H. Tashjian, E. J. Armstrong and A. W. Armstrong, Principles of Pharmacology: The Pathophysiologic Basis of Drug Therapy, Philadelphia: Wolters Kluwer, 2012.

[17] A. E. Brady and L. E. Limbird, "G protein-coupled receptor interacting proteins: Emerging roles in localization and signal transduction," *Cellular Signalling,* vol. 14, no. 4, pp. 297-309, 2002.

[18] T. K. Bjarnadóttir, D. E. Gloriam, S. H. Hellstrand, H. Kristiansson, . R. Fredriksson and H. B. Schiöth, "Comprehensive repertoire and phylogenetic analysis of the G protein-coupled receptors in human and mouse," *Genomics,* vol. 88, no. 3, pp. 263-273, 2006.

[19] R. Fredriksson and H. B. Schiöth, "The repertoire of G-protein-coupled receptors in fully sequenced genomes," *Molecular Pharmacology,* vol. 67, no. 5, pp. 1414-1425, 2005.

[20] L. Chen, L. Jin and N. Zhou, "An update of novel screening methods for GPCR in drug discovery," *Expert Opinion on Drug Discovery,* vol. 7, no. 9, pp. 791-806, 2012.

[21] R. J. Lefkowitz and B. K. Kobilka, "The Nobel Prize in Chemistry 2012," Royal Swedish Academy of Sciences, Stockholm, 2012.

[22] A. Ulloa-Aguirre, D. Stanislaus, J. A. Janovick and P. M. Conn, "Structure-activity relationships of G protein-coupled receptors," *Archives of Medical Research,* vol. 30, no. 6, pp. 420-435, 1999.

[23] J. Bockaert and J. P. Pin, "Molecular tinkering of G protein-coupled receptors: an evolutionary success," *The EMBO Journal,* vol. 18, no. 7, pp. 1723-1729, 1999.

[24] C. R. McCudden, M. D. Hains, R. J. Kimple, D. P. Siderovski and F. S. Willard, "G-protein signaling: back to the future," *Cellular and Molecular Life Sciences,* vol. 62, no. 5, pp. 551-577, 2005.

[25] G. G. J. Hazell, C. C. Hindmarch, G. R. Pope, J. A. Roper, S. L. Lightman, D. Murphy, A.-M. O'Carroll and S. J. Lolait, "G protein-coupled receptors in the hypothalamic paraventricular and supraoptic nuclei - serpentine gateways to neuroendocrine homeostasis," *Frontiers in Neuroendocrinol,* vol. 33, no. 1, pp. 45-66, 2012.

[26] M. I. Simon, M. P. Strathmann and N. Gautam, "Diversity of G proteins in signal transduction," *Science,* vol. 252, no. 5007, pp. 802-808, 1991.

[27] R. K. Sunahara and R. Taussig, "Isoforms of mammalian adenylyl cyclase: Multiplicities of signaling," *Molecular Interventions,* vol. 2, no. 3, pp. 168-184, 2002.

[28] J. D. Hildebrandt, R. D. Sekura, J. Codina, R. Iyengar, C. R. Manclark and L. Birnbaumer, "Stimulation and inhibition of adenylyl cyclases mediated by distinct regulatory proteins," *Nature,* vol. 302, no. 5910, pp. 706-709, 1983.

[29] S. G. Rhee and Y. S. Bae, "Regulation of phosphoinositide-specific phospholipase C isozymes," *The Journal of Biological Chemistry,* vol. 272, pp. 15045-15048, 1997.

[30] F. Horn, E. Bettler, L. Oliveira, F. Campagne, F. E. Cohen and G. Vriend, "GPCRDB information system for G protein-coupled receptors," *Nucleic Acids Research,* vol. 31, no. 1, pp. 294-297, 2003.

[31] L. F. Kolakowski, "GCRDb: a G-protein-coupled receptor database," *Receptors Channels,* vol. 2, no. 1, pp. 1-7, 1994.

[32] M. N. Davies, A. Secker, M. Halling-Brown, D. S. Moss, A. A. Freitas, J. Timmis, E. Clark and D. R. Flower, "GPCRTree: online hierarchical classification of GPCR function," *BMC Research Notes,* vol. 1, no. 67, pp. 1-5, 2008.

[33] H. B. Schiöth and R. Fredriksson, "The GRAFS classification system of G-protein coupled receptors in comparative perspective," *General and Comparative Endocrinology,* vol. 142, no. 1-2, pp. 94-101, 2005.

[34] C. Branden and J. Tooze, Introduction to Protein Structure, 2nd Edition, New York: Garland Science, 1999.

[35] S. Watson and S. Arkinstall, The G-protein linked receptor factsbook, London: Academic Press, 1994.

[36] T. D. Werrya, P. M. Sexton and A. Christopoulos, "'Ins and outs' of seven-transmembrane receptor signalling to ERK," *Trends in Endocrinology & Metabolism,* vol. 16, no. 1, pp. 26-33, 2005.

[37] A. M. Spiegel, "Inborn errors of signal transduction: mutations in G proteins and G protein-coupled receptors as a cause of disease," *Journal of inherited metabolomic disease,* vol. 20, no. 2, pp. 113-121, 1997.

[38] B. G. Katzung, A. J. Trevor and S. B. Masters, Basic and Clinical Pharmacology, 12th Edition, San Francisco: McGraw-Hill Medical, 2012.

[39] J. R. Hepler, "RGS protein and G protein interactions: a little help from their friends," *Molecular Pharmacology,* vol. 64, no. 3, pp. 547-549, 2003.

[40] M. L. Burcha, N. Osman, R. Getachew, S. Al-aryahi, P. Poronnik, W. Zheng, M. A. Hill and P. J. Little, "G protein coupled receptor transactivation: extending the paradigm to include serine/threonine kinase receptors," *The International Journal of Biochemistry & Cell Biology,* vol. 44, no. 5, pp. 722-727, 2012.

[41] S. Offermanns, "G-proteins as transducers in transmembrane signalling," *Progress in Biophysics and Molecular Biology,* vol. 83, no. 2, pp. 101-130, 2003.

[42] L. M. Luttrell, "Reviews in molecular biology and biotechnology: transmembrane signaling by G protein-coupled receptors," *Molecular Biotechnology,* vol. 39, no. 3, pp. 239-264, 2008.

[43] L. M. Luttrell, "Transmembrane signaling by G protein-coupled receptors," *Methods in Molecular Biology,* vol. 332, pp. 3-49, 2006.

[44] R. Irannejad and M. von Zastrow, "GPCR signaling along the endocytic pathway," *Current Opinion in Cell Biology,* vol. 27, pp. 109-116, 2014.

[45] M. T. Drake, S. K. Shenoy and R. J. Lefkowitz, "Trafficking of G Protein–Coupled Receptors," *Circulation Research,* vol. 99, pp. 570-582, 2006.

[46] N. Kawakami, K. Miyoshi, S. Horio and H. Fukui, "β2-Adrenergic Receptor-Mediated Histamine H1 Receptor Down-Regulation: Another Possible Advantage of β2 Agonists in Asthmatic Therapy," *Journal of Pharmacological Sciences,* vol. 94, no. 4, pp. 449-458, 2004.

[47] R. Lappano and M. Maggiolini, "GPCRs and cancer," *Acta Pharmacologica Sinica,* vol. 33, pp. 351-362, 2012.

[48] J.-F. Chena, P. K. Sonsalla, . F. Pedata, A. Melani, M. R. Domenici, P. Popoli, J. Geiger, L. V. Lopes and A. de Mendonça, "Adenosine A2A receptors and brain injury: Broad spectrum of neuroprotection, multifaceted actions and “fine tuning” modulation," *Progress in Neurobiology,* vol. 83, no. 5, pp. 310-331, 2007.

[49] W. Duan, L. Gui, Z. Zhou, Y. Liu, H. Tian, J.-F. Chen and J. Zheng, "Adenosine A2A receptor deficiency exacerbates white matter lesions and cognitive deficits induced by chronic cerebral hypoperfusion in mice," *Journal of the Neurological Sciences,* vol. 285, no. 1-2, pp. 39-45, 2009.

[50] P. A. Insel, C.-M. Tang, I. Hahntow and M. C. Michel, "Impact of GPCRs in clinical medicine: monogenic diseases, genetic variants and drug targets," *Biochimica et Biophysica Acta,* vol. 1768, no. 4, pp. 994-1005, 2007.

[51] G. Innamorati, M. T. Valenti, F. Giovinazzo, L. D. Carbonare, M. Parenti and C. Bassi, "Molecular approaches to target GPCRs in cancer therapy," *Pharmaceuticals,* vol. 4, pp. 567-589, 2011.

[52] R. T. Dorsam and J. S. Gutkind, "G-protein-coupled receptors and cancer," *Nature Reviews Cancer,* vol. 7, pp. 79-94, 2007.

[53] R. Lappano and M. Maggiolini, "G protein-coupled receptors: novel targets for drug discovery in cancer," *Nature Reviews Drug Discovery,* vol. 10, pp. 47-60, 2011.

[54] Z. Kan, B. S. Jaiswal, J. Stinson, V. Janakiraman, D. Bhatt, H. M. Stern, P. Yue, P. M. Haverty, R. Bourgon, J. Zheng, M. Moorhead, S. Chaudhuri, L. P. Tomsho, B. A.

Peters, K. Pujara, S. Cordes, D. P. Davis, V. E. H. Carlton, W. Yuan, L. Li, W. Wang, C. Eigenbrot, J. S. Kaminker, D. A. Eberhard, P. Waring, S. C. Schuster, Z. Modrusan, Z. Zhang, D. Stokoe, F. J. de Sauvage, M. Faham and S. Seshagiri, "Diverse somatic mutation patterns and pathway alterations in human cancers," *Nature,* vol. 12, pp. 869-873, 2010.

[55] E. A. Ariazi, J. L. Ariazi, F. Cordera and V. C. Jordan , "Estrogen receptors as therapeutic targets in breast cancer," *Current Topics in Medicinal Chemistry,* vol. 6, no. 3, pp. 181-202, 2006.

[56] Y. Daaka, "G proteins in cancer: the prostate cancer paradigm," *Science STKE,* vol. 2004, no. 216, p. re2, 2004.

[57] D. P. McDonnell and J. D. Norris, "Connections and Regulation of the Human Estrogen Receptor," *Science,* vol. 296, no. 5573, pp. 1642-1644, 2002.

[58] B. J. Feldman and D. Feldman, "The development of androgen-independent prostate cancer," *Nature Reviews Cancer,* vol. 1, pp. 34-45, 2001.

[59] E. F. Foley, A. A. Jazaeri, M. A. Shupnik, O. Jazaeri and L. W. Rice, "Selective loss of estrogen receptor beta in malignant human colon," *Cancer Research,* vol. 60, no. 2, pp. 245-248, 2000.

[60] G. P. Risbridger, S. J. Ellem and S. J. McPherson, "Estrogen action on the prostate gland: a critical mix of endocrine and paracrine signaling," *Journal of Molecular Endocrinology,* vol. 39, no. 3, pp. 183-188, 2007.

[61] R. Dai, R. A. Phillips, E. Karpuzoglu, D. Khan and S. A. Ahmed, "Estrogen regulates transcription factors STAT-1 and NF-kappaB to promote inducible nitric oxide synthase and inflammatory responses," *Journal of Immunology,* vol. 183, no. 11, pp. 6998-7005, 2009.

[62] J. A. Guzzo, "Selective estrogen receptor modulators--a new age of estrogens in cardiovascular disease?," *Clinical cardiology,* vol. 23, no. 1, pp. 15-17, 2000.

[63] E. B. Friedrich, Y. P. Clever, S. Wassmann, C. Hess and G. Nickenig, "17Beta-estradiol inhibits monocyte adhesion via down-regulation of Rac1 GTPase," *Journal of Molecular & Cellular Cardiology,* vol. 40, no. 1, pp. 87-95, 2006.

[64] P. D. Hurn and I. M. Macrae, "Estrogen as a neuroprotectant in stroke," *Journal of Cerebral Blood Flow & Metabolism,* vol. 20, no. 4, pp. 631-652, 2000.

[65] J. D. Termine and M. Wong, "Post-menopausal women and osteoporosis: available choices for maintenance of skeletal health," *Maturitas,* vol. 30, no. 3, pp. 241-245, 1998.

[66] S. C. Hewitt, J. C. Harrell and K. S. Korach, "Lessons in estrogen biology from knockout and transgenic animals," *Annual Review of Physiology,* vol. 67, pp. 285-308, 2005.

[67] J. Matthews and J. A. Gustafsson, "Estrogen signaling: a subtle balance between ER alpha and ER beta," *Molecular Interventions,* vol. 3, no. 5, pp. 281-292, 2003.

[68] J. E. Visvader and G. J. Lindeman, "Transcriptional regulators in mammary gland development and cancer," *The International Journal of Biochemistry & Cell Biology,* vol. 35, no. 7, pp. 1034-1051, 2003.

[69] J. F. Couse and K. S. Korach, "Estrogen receptor null mice: what have we learned and where will they lead us?," *Endocrine Review,* vol. 20, no. 3, pp. 358-417, 1999.

[70] J. E. Rossouw, G. L. Anderson, R. L. Prentice, A. Z. LaCroix, C. Kooperberg, M. L. Stefanick, R. D. Jackson, S. A. A. Beresford, B. V. Howard, K. C. Johnson, J. M. Kotchen and J. Ockene, "Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women: Principal Results From the Women's Health Initiative Randomized Controlled Trial," *The Journal of the Americal Medical Association,* vol. 288, no. 3, pp. 321-333, 2002.

[71] B. L. Herrmann, B. Saller, O. E. Janssen, P. Gocke, A. Bockisch, H. Sperling, K. Mann and M. Broecker, "Impact of estrogen replacement therapy in a male with congenital aromatase deficiency caused by a novel mutation in the CYP19 gene," *The Journal of Clinical Endocrinology & Metabolism,* vol. 87, no. 12, pp. 5476-5484, 2002.

[72] B. B. Sherwin, "Estrogen and Cognitive Functioning in Women," *Endocrine Reviews,* vol. 24, no. 2, pp. 133-151, 2003.

[73] L. J. Currie, M. B. Harrison, J. M. Trugman, J. P. Bennett and G. F. Wooten, "Postmenopausal estrogen use affects risk for Parkinson disease," *JAMA Neurology,* vol. 61, no. 6, pp. 886-888, 2004.

[74] K. M. Dhandapani and D. W. Brann, "Protective effects of estrogen and selective estrogen receptor modulators in the brain," *Biology of Reproduction,* vol. 67, no. 5, pp. 1379-1385, 2002.

[75] L. Baker, K. K. Meldrum, M. Wang, R. Sankula, R. Vanam, A. Raiesdana, B. Tsai, K. Hile, J. W. Brown and D. R. Meldrum, "The role of estrogen in cardiovascular disease," *Journal of Surgical Research,* vol. 115, no. 2, pp. 325-344, 2003.

[76] M. E. Mendelsohn, "Mechanisms of estrogen action in the cardiovascular system," *The Journal of Steroid Biochemistry and Molecular Biology,* vol. 74, no. 5, pp. 337-343, 2000.

[77] E. J. Filardo and P. Thomas, "Minireview: G Protein-Coupled Estrogen Receptor-1, GPER-1: Its Mechanism of Action and Role in Female Reproductive Cancer, Renal and Vascular Physiology," *Endocrinology,* vol. 153, no. 7, pp. 2953-2962, 2012.

[78] M. Razandi, G. Alton, A. Pedram, S. Ghonshani, P. Webb and E. R. Levin, "Identification of a structural determinant necessary for the localization and function of estrogen receptor alpha at the plasma membrane," *Molecular and Cellular Biology,* vol. 23, no. 5, pp. 1633-1646, 2003.

[79] E. R. Prossnitz, J. B. Arterburnd and L. A. Sklar, "GPR30: A G protein-coupled receptor for estrogen," *Molecular and Cellular Endocrinology,* Vols. 265-266, pp. 138-142, 2007.

[80] M. H. Wyckoff, K. L. Chambliss, C. Mineo, I. S. Yuhanna, M. E. Mendelsohn, S. M. Mumby and P. W. Shaul, "Plasma membrane estrogen receptors are coupled to endothelial nitric-oxide synthase through Galpha(i)," *The Journal of Biological Chemistry,* vol. 276, no. 29, pp. 27071-27076, 2001.

[81] M. J. Kellya and E. R. Levin, "Rapid actions of plasma membrane estrogen receptors," *Trends in Endocrinology & Metabolism,* vol. 12, no. 4, pp. 152-156, 2001.

[82] E. J. Filardo and P. Thomas, "GPR30: a seven-transmembrane-spanning estrogen receptor that triggers EGF release," *Trends in Endocrinology & Metabolism,* vol. 16, no. 8, pp. 362-367, 2005.

[83] C. Owmana, P. Blay, C. Nilsson and S. J. Lolait, "Cloning of human cDNA encoding a novel heptahelix receptor expressed in Burkitt's lymphoma and widely

distributed in brain and peripheral tissues," *Biochemical and Biophysical Research Communications,* vol. 228, no. 2, pp. 285-292, 1996.

[84] Y. Takada, C. Kato, S. Kondo, R. Korenaga and J. Ando, "Cloning of cDNAs Encoding G Protein-Coupled Receptor Expressed in Human Endothelial Cells Exposed to Fluid Shear Stress," *Biochemical and Biophysical Research Communications,* vol. 240, no. 3, pp. 737-741, 1997.

[85] C. Carmeci, D. A. Thompson, H. Z. Ring, U. Francke and R. J. Weigel, "Identification of a Gene (GPR30) with Homology to the G-Protein-Coupled Receptor Superfamily Associated with Estrogen Receptor Expression in Breast Cancer," *Genomics,* vol. 45, no. 3, pp. 607-617, 1997.

[86] B. F. O'Dowd, T. Nguyen, A. Marchese, R. Cheng, K. R. Lynch, H. H. Q. Heng, L. F. Kolakowski Jr and S. R. George, "Discovery of three novel G-protein-coupled receptor genes," *Genomics,* vol. 47, no. 2, pp. 310-313, 1998.

[87] Y. Feng and P. Gregor, "Cloning of a novel member of the G protein-coupled receptor family related to peptide receptors," *Biochemical and Biophysical Research Communications,* vol. 231, no. 3, pp. 651-654, 1997.

[88] E. J. Filardo, J. A. Quinn, K. I. Bland and A. R. Frackelton, "Estrogen-induced activation of Erk-1 and Erk-2 requires the G protein-coupled receptor homolog, GPR30, and occurs via trans-activation of the epidermal growth factor receptor through release of HB-EGF," *Molecular Endocrinology,* vol. 14, no. 10, pp. 1649-1660, 2000.

[89] D. Wang, L. Hu, G. Zhang, L. Zhang and C. Chen, "G protein-coupled receptor 30 in tumor development," *Endocrine,* vol. 38, no. 1, pp. 29-37, 2010.

[90] S. P. H. Alexander, A. Mathie and J. A. Peters, "Guide to Receptors and Channels (GRAC), 5th edition," *British Journal of Pharmacology,* vol. 164, no. Suppl. 1, pp. S1-S324, 2011.

[91] C. M. Revankar, D. F. Cimino, L. A. Sklar, J. B. Arterburn and E. R. Prossnitz, "A transmembrane intracellular estrogen receptor mediates rapid cell signaling," *Science,* vol. 307, no. 5715, pp. 1625-1630, 2005.

[92] C. M. Revankar, H. D. Mitchell, A. S. Field , R. Burai, C. Corona , C. Ramesh, L. A. Sklar , J. B. Arterburn and E. R. Prossnitz, "Synthetic Estrogen Derivatives

Demonstrate the Functionality of Intracellular GPR30," *ACS Chemical Biology,* vol. 2, no. 8, pp. 536-544, 2007.

[93] F. Gobeil, A. Fortier, T. Zhu, M. Bossolasco, M. Leduc, M. Grandbois, N. Heveker, G. Bkaily, S. Chemtob and D. Barbaz, "G-protein-coupled receptors signalling at the cell nucleus: an emerging paradigm," *Canadian Journal of Physiology and Pharmacology,* vol. 84, no. 3-4, pp. 287-297, 2006.

[94] P. Thomas, Y. Pang, E. J. Filardo and J. Dong, "Identity of an estrogen membrane receptor coupled to a G protein in human breast cancer cells," *Endocrinology,* vol. 146, no. 2, pp. 624-632, 2005.

[95] T. Funakoshi, A. Yanai, K. Shinoda, M. M. Kawano and Y. Mizukami, "G protein-coupled receptor 30 is an estrogen receptor in the plasma membrane," *Biochemical and Biophysical Research Communications,* vol. 346, no. 3, pp. 904-910, 2006.

[96] M. Maggiolini, A. Vivacqua, G. Fasanella, A. G. Recchia, D. Sisci, V. Pezzi, D. Montanaro, A. M. Musti, D. Picard and S. Andò, "The G protein-coupled receptor GPR30 mediates c-fos up-regulation by 17beta-estradiol and phytoestrogens in breast cancer cells," *The Journal of Biological Chemistry,* vol. 279, no. 26, pp. 27008-27016, 2004.

[97] L. Albanito, A. Madeo, R. Lappano, A. Vivacqua, V. Rago, A. Carpino, T. I. Oprea, E. R. Prossnitz, A. M. Musti, S. Andò and M. Maggiolini, "G protein-coupled receptor 30 (GPR30) mediates gene expression changes and growth response to 17beta-estradiol and selective GPR30 ligand G-1 in ovarian cancer cells," *Cancer Research,* vol. 67, no. 4, pp. 1859-1866., 2007.

[98] E. J. Filardo, J. A. Quinn, A. R. Frackelton and K. I. Bland, "Estrogen action via the G protein-coupled receptor, GPR30: stimulation of adenylyl cyclase and cAMP-mediated attenuation of the epidermal growth factor receptor-to-MAPK signaling axis," *Molecular Endocrinology,* vol. 16, no. 1, pp. 70-84, 2002.

[99] E. R. Prossnitz, J. B. Arterburn, H. O. Smith, T. I. Oprea, L. A. Sklar and H. J. Hathaway, "Estrogen signaling through the transmembrane G protein-coupled receptor GPR30," *Annual Review of Physiology,* vol. 70, pp. 165-190, 2008.

[100] E. R. Prossnitza, T. I. Oprea, L. A. Sklar and J. B. Arterburn, "The ins and outs of GPR30: a transmembrane estrogen receptor," *The Journal of Steroid Biochemistry and Molecular Biology,* vol. 109, no. 3-5, pp. 350-353, 2008.

[101] C. G. Bologa, C. M. Revankar, S. M. Young, B. S. Edwards, J. B. Arterburn, A. S. Kiselyov, M. A. Parker, S. E. Tkachenko, N. P. Savchuck, L. A. Sklar, T. I. Oprea and E. R. Prossnitz, "Virtual and biomolecular screening converge on a selective agonist for GPR30," *Nature Chemical Biology,* vol. 2, no. 4, pp. 207-212, 2006.

[102] M. K. Dennis, R. Burai, C. Ramesh, W. K. Petrie, S. N. Alcon, T. K. Nayak, C. G. Bologa, A. Leitao, E. Brailoiu, E. Deliu, N. J. Dun, L. A. Sklar, H. J. Hathaway, J. B. Arterburn, T. I. Oprea and E. R. Prossnitz, "In vivo effects of a GPR30 antagonist," *Nature Chemical Biology,* vol. 5, no. 6, pp. 421-427, 2009.

[103] M. K. Dennis, A. S. Fielda, R. Burai, C. Ramesh, W. K. Petrie, C. G. Bologa, T. I. Oprea, Y. Yamaguchi, S.-I. Hayashi, L. A. Sklar, H. J. Hathaway, J. B. Arterburn and E. R. Prossnitz, "Identification of a GPER/GPR30 antagonist with improved estrogen receptor counterselectivity," *The Journal of Steroid Biochemistry and Molecular Biology,* vol. 127, no. 3-5, pp. 358-366, 2011.

[104] A. Vivacqua, D. Bonofiglio , A. G. Recchia, A. M. Musti, D. Picard and S. Andò, "The G protein-coupled receptor GPR30 mediates the proliferative effects induced by 17beta-estradiol and hydroxytamoxifen in endometrial cancer cells," *Molecular Endocrinology,* vol. 20, no. 3, pp. 631-646, 2006.

[105] A. Vivacqua, D. Bonofiglio, L. Albanito, A. Madeo, V. Rago, A. Carpino, A. M. Musti, D. Picard, S. Andò and M. Maggiolini, "17beta-estradiol, genistein, and 4-hydroxytamoxifen induce the proliferation of thyroid cancer cells through the g protein-coupled receptor GPR30," *Molecular Pharmacology,* vol. 70, no. 4, pp. 1414-1423, 2006.

[106] D. P. Pandey, R. Lappano, L. Albanito, A. Madeo, M. Maggiolini and D. Picard, "Estrogenic GPR30 signalling induces proliferation and migration of breast cancer cells through CTGF," *The EMBO Journal,* vol. 28, no. 5, pp. 523-532, 2009.

[107] P. Thomas and J. Dong, "Binding and activation of the seven-transmembrane estrogen receptor GPR30 by environmental estrogens: a potential novel mechanism of endocrine disruption," *The Journal of Steroid Biochemistry and Molecular Biology,* vol. 102, no. 1-5, pp. 175-179, 2006.

[108] V. Gigoux and D. Fourmy, "Acting on hormone receptors with minimal side effect on cell proliferation: a timely challenge illustrated with GLP-1R and GPER," *Front Endocrinol,* vol. 4, no. 50, pp. 1-16, 2013.

[109] J. A. Quinn, C. T. Graeber, A. R. Frackelton Jr, M. Kim, J. E. Schwarzbauer and E. J. Filardo, "Coordinate regulation of estrogen-mediated fibronectin matrix assembly and epidermal growth factor receptor transactivation by the G protein-coupled receptor, GPR30," *Molecular Endocrinology,* vol. 23, no. 7, pp. 1052-1064, 2009.

[110] I. Vivanco and C. L. Sawyers, "The phosphatidylinositol 3-Kinase–AKT pathway in human cancer," *Nature Reviews Cancer,* vol. 2, no. 7, pp. 489-501, 2002.

[111] M. Maggiolini and D. Picard, "The unfolding stories of GPR30, a new membrane-bound estrogen receptor," *Journal of Endocrinology,* vol. 204, no. 2, pp. 105-114, 2010.

[112] E. P. Samartzis, A. Noske, A. Meisel, Z. Varga, D. Fink and P. Imesch, "The G Protein-Coupled Estrogen Receptor (GPER) Is Expressed in Two Different Subcellular Localizations Reflecting Distinct Tumor Properties in Breast Cancer," *PLoS ONE 9,* vol. 9, no. 1, pp. 1-8, 2014.

[113] E. J. Filardo, C. T. Graeber, J. A. Quinn, M. B. Resnick, D. Giri, R. A. DeLelli, M. M. Steinhoff and E. Sabo, "Distribution of GPR30, a seven membrane-spanning estrogen receptor, in primary breast cancer and its association with clinicopathologic determinants of tumor progression," *Clinical Cancer Research,* vol. 12, no. 21, pp. 6359-6366, 2006.

[114] L. Albanito, D. Sisci, S. Aquila, E. Brunelli, A. Vivacqua, A. Madeo, R. Lappano, D. P. Pandey, D. Picard, L. Mauro, S. Andò and M. Maggiolini, "Epidermal growth factor induces G protein-coupled receptor 30 expression in estrogen receptor-negative breast cancer cells," *Endocrinology,* vol. 149, no. 8, pp. 3799-3808, 2008.

[115] Y.-Y. He, B. Cai, Y.-X. Yang, X.-L. Liu and X.-P. Wan, "Estrogenic G protein-coupled receptor 30 signaling is involved in regulation of endometrial carcinoma by promoting proliferation, invasion potential, and interleukin-6 secretion via the MEK/ERK mitogen-activated protein kinase pathway," *Cancer Science,* vol. 100, no. 6, pp. 1051-1061, 2009.

[116] K. Leblanc, É. Sexton, S. Parent, G. Bélanger, M.-C. Déry, V. Boucher and E. Asselin, "Effects of 4-hydroxytamoxifen, raloxifene and ICI 182 780 on survival of uterine

cancer cell lines in the presence and absence of exogenous estrogens," *International Journal of Oncology,* vol. 30, no. 2, pp. 477-487, 2007.

[117] E. Henic, V. Noskova, G. Høyer-Hansen , S. Hansson and B. Casslén, "Estradiol attenuates EGF-induced rapid uPAR mobilization and cell migration via the G-protein-coupled receptor 30 in ovarian cancer cells," *International Journal of Gynecological Cancer,* vol. 19, no. 2, pp. 214-222, 2009.

[118] L. Albanito, R. Lappano, A. Madeo , A. Chimento, E. R. Prossnitz, A. R. Cappello, V. Dolce, S. Abonante, V. Pezzi and M. Maggiolini, "G-protein-coupled receptor 30 and estrogen receptor-alpha are involved in the proliferative effects induced by atrazine in ovarian cancer cells," *Environental Health Perspectives,* vol. 116, no. 12, pp. 1648-1655, 2008.

[119] W.-H. Kuo, L.-Y. Chang, D. L.-Y. Liu, H.-L. Hwa, J.-J. Lin, P.-H. Lee, C.-N. Chen, H.-C. Lien, R.-H. Yuan, C.-T. Shun, K.-J. Chang and F.-J. Hsieh, "The interactions between GPR30 and the major biomarkers in infiltrating ductal carcinoma of the breast in an Asian population," *Taiwanese Journal of Obstetrics and Gynecology,* vol. 46, no. 2, pp. 135-145, 2007.

[120] Y. Mizukami, "In vivo functions of GPR30/GPER-1, a membrane receptor for estrogen: from discovery to functions in vivo," *Endocrine journal,* vol. 57, no. 2, pp. 101-107, 2010.

[121] E. R. Prossnitz, L. A. Sklar, T. I. Oprea and J. B. Arterburn, "GPR30: a novel therapeutic target in estrogen-related disease," *Trends in Pharmacological Science,* vol. 29, no. 3, pp. 116-123, 2008.

[122] Y. Shang, "Molecular mechanisms of oestrogen and SERMs in endometrial carcinogenesis," *Nature Reviews Cancer,* vol. 6, no. 5, pp. 360-368, 2006.

[123] H. O. Smith, K. K. Leslie, M. Singh, C. R. Qualls, C. M. Revankar, N. E. Joste and E. R. Prossnitz, "GPR30: a novel indicator of poor survival for endometrial carcinoma," *GPR30: a novel indicator of poor survival for endometrial carcinoma,* vol. 196, no. 4, pp. 386.e1-386.e11, 2007.

[124] H. O. Smith, H. Arias-Pulido, D. Y. Kuoe, T. Howard, C. R. Qualls, S.-J. Lee, C. F. Verschraegen, H. J. Hathaway, N. E. Joste and E. R. Prossnitz, "GPR30 predicts poor

survival for ovarian cancer," *Gynecologic Oncology,* vol. 114, no. 3, pp. 465-471, 2009.

[125] A. Madeo and M. Maggiolini, "Nuclear Alternate Estrogen Receptor GPR30 Mediates 17β-Estradiol–Induced Gene Expression and Migration in Breast Cancer–Associated Fibroblasts," *Cancer Research,* vol. 70, no. 14, pp. 6036-6046, 2010.

[126] B. Olde and L. M. F. Leeb-Lundberg, "GPR30/GPER1: searching for a role in estrogen physiology," *Trends in Endocrinology and Metabolism,* vol. 20, no. 8, pp. 409-416, 2009.

[127] G. Sharma and E. R. Prossnitz, "Mechanisms of estradiol-induced insulin secretion by the G protein-coupled estrogen receptor GPR30/GPER in pancreatic beta-cells," *2011,* vol. 152, no. 8, pp. 3030-3039, 2011.

[128] M. R. Meyer, E. R. Prossnitz and M. Barton, "The G protein-coupled estrogen receptor GPER/GPR30 as a regulator of cardiovascular function," *Vascular Pharmacology,* vol. 55, no. 1-3, pp. 17-25, 2011.

[129] M. V. Hofmeister, H. H. Damkier, B. M. Christensen , B. Olde, L. M. F. Leeb-Lundberg , R. A. Fenton, H. A. Praetorius and J. Praetorius, "17β-Estradiol induces nongenomic effects in renal intercalated cells through G protein-coupled estrogen receptor 1," *American Journal of Physiology - Renal Physiology,* vol. 302, no. 3, pp. F358-F368, 2012.

[130] S. H. Lindsey, L. M. Yamaleyeva, K. B. Brosnihan, P. E. Gallagher and M. C. Chappell, "Estrogen receptor GPR30 reduces oxidative stress and proteinuria in the salt-sensitive female mRen2.Lewis rat," *Hypertension,* vol. 58, no. 4, pp. 665-671, 2011.

[131] M. Barton, "The membrane estrogen receptor GPER - clues and questions," *Steroids,* vol. 77, pp. 935-942, 2012.

[132] C. Otto, B. Rohde-Schulz, G. Schwarz, I. Fuchs, M. Klewer, D. Brittain, G. Langer, B. Bader, K. Prelle, R. Nubbemeyer and K. H. Fritzemeier, "G protein-coupled receptor 30 localizes to the endoplasmic reticulum and is not activated by estradiol," *Endocrinology,* vol. 149, no. 10, pp. 4846-4856, 2008.

[133] C. Otto, I. Fuchs , G. Kauselmann, H. Kern, B. Zevnik , P. Andreasen, G. Schwarz, H. Altmann, M. Klewer, M. Schoor M, R. Vonk and K. H. Fritzemeier, "GPR30 does not

mediate estrogenic responses in reproductive organs in mice," *Biology of Reproduction,* vol. 80, no. 1, pp. 34-41, 2009.

[134] E. R. Levin, "G Protein-Coupled Receptor 30: Estrogen Receptor or Collaborator?," *Endocrinology,* vol. 150, no. 4, pp. 1563-1565, 2009.

[135] D. M. Rosenbaum, S. G. F. Rasmussen and B. K. Kobilka, "The structure and function of G-protein-coupled receptors," *Nature,* vol. 459, no. 7245, pp. 356-363, 2009.

[136] M. R. Wilkins, J. C. Sanchez, A. A. Gooley, R. D. Appel, I. Humphery-Smith, D. F. Hochstrasser and K. L. Williams, "Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it," *Biotechnology & Genetic Engineering Reviews,* vol. 13, pp. 19-50, 1996.

[137] R. Cammack, T. Atwood, P. Campbell, H. Parish, A. Smith and F. Vella, Oxford Dictionary of Biochemistry and Molecular Biology, 2nd edition, New York: Oxford University Press, 2006.

[138] A. Shevchenko , O. N. Jensen , A. V. Podtelejnikov, F. Sagliocco, M. Wilm, O. Vorm, P. Mortensen, A. Shevchenko, H. Boucherie and M. Mann, "Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels," *Proc Natl Acad Sci U S A,* vol. 93, no. 25, pp. 14440-14445, 1996.

[139] M. Uhlén M and S. Hober, "Generation and validation of affinity reagents on a proteome-wide level," *Journal of Molecular Recognition,* vol. 22, no. 2, pp. 57-64, 2009.

[140] P. V. Edman, "Method for determination of the amino acid sequence in peptides," *Acta Chemica Scandinavica,* vol. 4, pp. 283-293, 1950.

[141] P. Edman and G. Begg, "A protein sequenator," *European Journal of Biochemistry,* vol. 1, pp. 80-91, 1967.

[142] R. Laursen, "Solid-phase Edman degradation. An automatic peptide sequence," *European Journal of Biochemistry,* vol. 20, pp. 89-102, 1971.

[143] H. Niall, "Automated Edman degradation: the protein sequenator," *Methods in Enzymology,* vol. 27, pp. 942-1010, 1973.

[144] E. De Hoffmann and V. Stroobant, Mass spectrometry: principles and applications, 3rd Edition, Wiley-Interscience, 2007.

[145] F. W. Aston, A positive ray spectrograph, London: Taylor & Francis, 1919.

[146] M. Mann , R. C. Hendrickson and A. Pandey, "Analysis of proteins and proteomes by mass spectrometry," *Annu Rev Biochem,* vol. 70, pp. 437-473, 2001.

[147] W. J. Henzel , T. M. Billeci , J. T. Stults, S. C. Wong, C. Grimley and C. Watanabe, "Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases," *Proc Natl Acad Sci U S A,* vol. 90, no. 11, pp. 5011-5015, 1993.

[148] S. D. Patterson and R. Aebersold, "Mass spectrometric approaches for the identification of gel-separated proteins," *Electrophoresis,* vol. 16, no. 10, pp. 1791-1814, 1995.

[149] J. 3. Yates, S. Speicher , P. R. Griffin and T. Hunkapiller, "Peptide mass maps: a highly informative approach to protein identification," *Anal Biochem,* vol. 214, no. 2, pp. 397-408, 1993.

[150] M. Yamashita and J. B. Fenn, "Electrospray ion source. Another variation on the free-jet theme," *J Phys Chem,* vol. 88, no. 20, pp. 4451-4459, 1984.

[151] M. Karas and F. Hillenkamp , "Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons," *Anal Chem,* vol. 60, no. 20, pp. 2299-2301, 1988.

[152] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong and C. M. Whitehouse, "Electrospray ionization for mass spectrometry of large biomolecules," *Science,* vol. 246, pp. 64-71, 1989.

[153] F. Hillenkamp , M. Karas, R. C. Beavis and B. T. Chait, "Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers," *Anal Chem,* vol. 63, pp. 1193A-1203A, 1991.

[154] I. V. Chernushevich, A. V. Loboda and B. A. Thomson, "An introduction to quadrupole-time-of-flight mass spectrometry," *J Mass Spectrom,* vol. 36, pp. 849-865, 2001.

[155] P. Roepstorff, "MALDI-TOF mass spectrometry in protein chemistry," *EXS,* vol. 88, pp. 81-97, 2000.

[156] R. A. Yost and R. K. Boyd, "Tandem Mass-Spectrometry - Quadrupole and Hybrid Instruments," *Methods in Enzymology,* vol. 193, pp. 154-200, 1990.

[157] S. M. Peterman, C. P. Dufresne and S. Horning, "The use of a hybrid linear trap/FT-ICR mass spectrometer for on-line high resolution/high mass accuracy bottom-up sequencing," *J Biomol Tech,* vol. 16, pp. 112-124, 2005.

[158] R. A. Zubarev and A. Makarov, "Orbitrap Mass Spectrometry," *Anal Chem,* vol. 85, no. 11, pp. 5288-5296, 2013.

[159] A. V. Loboda, A. N. Krutchinsky, M. Bromirski, W. Ens and K. G. Standing, "A tandem quadrupole/time-of-flight mass spectrometer with a matrix-assisted laser desorption/ionization source: design and performance," *Rapid Commun Mass Spectrom,* vol. 14, pp. 1047-1057, 2000.

[160] N. V. Gogichaevaa, T. Williamsa and M. A. Alterman, "MALDI TOF/TOF Tandem Mass Spectrometry as a New Tool for Amino Acid Analysis," *Journal of the American Society for Mass Spectrometry,* vol. 18, no. 2, pp. 279-284, 2007.

[161] J. K. Lewis, J. Wei and G. Siuzdak, "Matrix-assisted Laser Desorption/Ionization Mass Spectrometry," *Encyclopedia of Analytical Chemistry R.A. Meyers (Ed.),* pp. 5880-5894, 2000.

[162] L. F. Marvin, M. A. Roberts and L. B. Fay, "Matrix-assisted laser desorption/ionization time-of-flight mass," *Clinica Chimica Acta,* vol. 337, pp. 11-21, 2003.

[163] C. Wang, E. R. Prossnitz and S. K. Roy, "Expression of G protein-coupled receptor 30 in the hamster ovary: differential regulation by gonadotropins and steroid hormones," *Endocrinology,* vol. 148, no. 10, pp. 4853-4864, 2007.

[164] R. Falk, M. Ramström, S. Ståhl and S. Hober, "Approaches for systematic proteome exploration," *Biomolecular Engineering,* vol. 24, no. 2, pp. 155-168, 2007.

[165] R. J. Pietras and C. M. Szego, "Partial purification and characterization of oestrogen receptors in subfractions of hepatocyte plasma membranes," *Biochem J,* vol. 191, pp. 743-760, 1980.

[166] K. S. Bloom and J. N. Anderson, "Fractionation and characterization of chromosomal proteins by the hydroxyapatite dissociation method.," *The journal of Biological Chemistry,* vol. 253, no. 12, pp. 4446-4450, 1978.

[167] M. I. Kay, R. A. Young and A. S. Posner, "Crystal structure of hydroxyapatite," *Nature,* vol. 204, pp. 1050-1052, 1964.

[168] "Osteopontin and related phosphorylated sialo proteins: effects on mineralization," *Ann NY Acad Sci,* vol. 760, pp. 249-256, 1995.

[169] L. J. Cummings, M. A. Snyder and K. Brisack, "Protein chromatography on hydroxyapatite columns," *Methods Enzymol,* vol. 463, pp. 387-404, 2009.

[170] S. R. Schmidt, F. Schweikart and M. E. Andersson, "Current methods for phosphoprotein isolation and enrichment," *J Chromatogr,* vol. 849, no. 1-2, pp. 154-162, 2007.

[171] G. Mamone, G. Picariello, P. Ferranti and F. Addeo, "Hydroxyapatite affinity chromatography for the highly selective enrichment of mono-and multi-phosphorylated peptides in phosphoproteome analysis," *Proteomics,* vol. 10, no. 3, pp. 380-393, 2010.

[172] J. Krenkova, N. A. Lacher and F. Svec, "Control of selectivity via nanochemistry: monolithic capillary column containing hydroxyapatite nanoparicles for separation of proteins and enrichment of phosphopeptides," *Anal Chem,* vol. 82, no. 19, pp. 8335-8341, 2010.

[173] B. R. Fonslow, S. M. Niessen, M. Singh, C. C. Wong, T. Xu, P. C. Carvalho, J. Choi, S. K. Park and J. 3. Yates, "Single-step inline hydroxyapatite enrichment facilitates identification and quantification of phosphopeptides from mass-limited proteomes with MudPIT," *Journal of Proteome Research,* vol. 11, pp. 2697-2709, 2012.

[174] S. Noguchi and Y. Satow, "Purification of human β2-adrenergic receptor expressed in methylotrophic yeast Pichia pastoris," *J Biochem,* vol. 140, no. 6, pp. 799-804, 2006.

[175] S. Fischer, K. Thummler, B. Volkert, K. Hettrich, I. Schmidt and K. Fischer, "Properties and applications of cellulose acetate," *Macromol Symp,* vol. 262, pp. 89-96, 2008.

[176] Z. Chen, M. Deng, Y. Chen, G. He, M. Wu and J. Wang, "Preparation and performance of cellulose acetate/polyethyleneimine blend microfiltration membranes and their applications," *Journal of Membrane Science,* vol. 235, no. 1-2, pp. 73-86, 2004.

[177] I. H. G. S. Consortium, "Finishing the euchromatic sequence of the human genome," *Nature ,* vol. 431, pp. 931-945, 2007.

[178] O. N. Jensen, "Modification-specific proteomics: Characterization of post-translational modifications by mass spectrometry," *Curr Opin Chem Biol,* vol. 8, pp. 33-41, 2004.

[179] T. A. Ayoubi and W. J. Van De Ven, "Regulation of gene expression by alternative promoters," *FASEB J,* vol. 10, pp. 453-460, 1996.

[180] L. Beevers, D. Boulter and B. Parthier, "Post-Translational Modifications," in *Nucleic Acids and Proteins in Plants I: Structure, Biochemistry and Physiology of Proteins*, Berlin, Springer Berlin Heidelberg, 1982, pp. 136-168.

[181] Y.-C. Wang, S. E. Peterson and J. F. Loring, "Protein post-translational modifications and regulation of pluripotency in human stem cells," *Cell Research,* vol. 24, pp. 143-160, 2014.

[182] C. T. Walsh, Posttranslational modification of proteins : Expanding nature's inventory, Englewood, Colorado: Roberts and Co., 2006.

[183] T. Hunter, "Tyrosine phosphorylation: thirty years and counting," *Curr Opin Cell Biol,* vol. 21, pp. 140-146, 2009.

[184] A. Ciechanover, "Proteolysis: from the lysosome to ubiquitin and the proteasome," *Nat Rev Mol Cell Biol,* vol. 6, pp. 79-87, 2005.

[185] A. Varki, Essentials of Glycobiology, 2nd Edition, New York: Cold Spring Harbor Laboratories Press, 2009.

[186] G. Burnett and E. P. Kennedy, "The enzymatic phosphorylation of proteins," *J Biol Chem,* vol. 211, no. 2, p. 969–980, 1954.

[187] C. M. Pickart and M. J. Eddins, "Ubiquitin: structures, functions, mechanisms," *Biochim Biophys Acta,* vol. 1695, no. 1-3, pp. 55-72, 2004.

[188] T. Brook, "Protein Acetylation: Much More than Histone Acetylation," Cayman Chemical, [Online]. Available: https://www.caymanchem.com/app/template/Article.vm/article/2152.

[189] G. Stefanidakis and J. E. Gwyn, "Alkylation," in *Chemical Processing Handbook*, Florida, CRC Press, 1993, pp. 80-138.

[190] C. Walsh, "Protein Methylation," in *Posttranslational modification of proteins: expanding nature's inventory*, Colorado, Roberts and Company, 2006.

[191] F. R. Bettelheim, "Tyrosine-O-sulfate in a peptide from fibrinogen," *J Am Chem Soc,* vol. 76, no. 10, pp. 2838-2839, 1954.

[192] P. Anand and S. J. S, "Enzymatic mechanisms regulating protein S-nitrosylation: implications in health and disease," *J Mol Med (Berl),* vol. 90, no. 3, pp. 233-244, 2012.

[193] M. J. Nadolski and M. E. Linder, "Protein lipidation," *FEBS J,* vol. 274, no. 20, pp. 5202-5210, 2007.

[194] D. N. Crowell and D. H. Huizinga, "Protein isoprenylation: the fat of the matter," *Trends in Plant Science,* vol. 14, no. 3, pp. 163-170, 2009.

[195] D. L. Cox and M. M. Nelson, Lehninger principles of biochemistry, 4TH Edition, New York: W.H. Freeman, 2005.

[196] O. Alzate, "Mass Spectrometry for Post-Translational Modifications," in *Neuroproteomics*, Boca Raton, Florida, CRC Press, 2010.

[197] B. Edde, J. Rossier, J. P. Le Caer, E. Desbruyeres, F. Gros and P. Denoulet, "Posttranslational glutamylation of alpha-tubulin," *Science,* vol. 247, no. 4938, pp. 83-85, 1990.

[198] C. Mencarelli, D. Caroti, M. H. Bré, N. Levilliers and R. Dallai, "Tubulin glycylation and glutamylation deficiencies in unconventional insect axonemes," *Cell Motil Cytoskeleton,* vol. 61, no. 4, pp. 226-236, 2005.

[199] J. Beld, E. C. Sonnenschein, C. R. Vickery, J. P. Noel and M. D. Burkart, "The phosphopantetheinyl transferases: catalysis of a post-translational modification crucial for life," *Nat Prod Rep,* vol. 31, no. 1, pp. 61-108, 2014.

[200] T. A. Brown and A. Shrift, "Identification of Selenocysteine in the Proteins of Selenate-grown Vigna radiata," *Plant Physiol,* vol. 66, no. 4, pp. 758-761, 1980.

[201] W. J. Driscoll, S. A. Mueller, B. A. Eipper and G. P. Mueller, "Differential Regulation of Peptide-Amidation by Dexamethasone and Disulfiram," *Molecular Pharmacology,* vol. 55, pp. 1067-1076, 1999.

[202] K. Fujiwara, K. Okamura-Ikeda and Y. Motokawa, "Lipoylation of Acyltransferase Components of α-Ketoacid Dehydrogenase Complexes," *The Journal of Biological Chemistry,* vol. 271, pp. 12932-12936, 1996.

[203] L. D. Rogers and C. M. Overall, "Proteolytic post-translational modification of proteins: proteomic tools and methodology," *Mol Cell Proteomics,* vol. 12, no. 12, pp. 3532-3542, 2013.

[204] Y. L. Deribe, T. Pawson and I. Dikic, "Post-translational modifications in signal integration," *Nat Struct Mol Biol,* vol. 17, pp. 666-672, 2010.

[205] J. Seo and K. J. Lee, "Post-translational Modifications and Their Biological Functions: Proteomic Analysis and Systematic Approaches," *Journal of Biochemistry and Molecular Biology,* vol. 37, no. 1, pp. 35-44, 2004.

[206] F. Meng, A. J. Forbes, L. M. Miller and N. L. Kelleher, "Detection and localization of protein modifications by high resolution tandem mass spectrometry," *Mass Spectrom Rev,* vol. 24, no. 2, pp. 126-134, 2005.

[207] H. Li, C. Liu, L. Burge and W. Southerland, "Identification of two post-translational modifications via tandem mass spectrometry," *International Journal of Computational Biology and Drug Design,* vol. 5, no. 3-4, pp. 314-324, 2012.

[208] D. Shental-Bechor and Y. Levy, "G-1-activated membrane estrogen receptors mediate increased contractility of the human myometrium," *Curr Opin Struct Biol,* vol. 19, pp. 524-533, 2009.

[209] R. S. Haltiwanger, "Regulation of signal transduction pathways in development by glycosylation," *Curr Opin Struct Biol,* vol. 12, pp. 593-598, 2002.

[210] C. I. Balog, O. A. Mayboroda, M. Wuhrer, C. H. Hokke, A. M. Deelder and P. J. Hensbergen, "Mass spectrometric identification of aberrantly glycosylated human

apolipoprotein C-III peptides in urine from Schistosoma mansoni-infected individuals," *Mol Cell Proteomics,* vol. 9, pp. 667-681, 2010.

[211] A. I. Fogel, Y. Li, J. Giza, Q. Wang, T. T. Lam, Y. Modis and T. Biederer, "N-glycosylation at the SynCAM (synaptic cell adhesion molecule) immunoglobulin interface modulates synaptic adhesion," *J Biol Chem,* vol. 285, pp. 34864-34874, 2010.

[212] C. W. Reid, K. M. Fulton and S. M. Twine, "Never take candy from a stranger: the role of the bacterial glycome in host-pathogen interactions," *Future Microbiol,* vol. 5, pp. 267-288, 2010.

[213] P. M. Rudd, T. Elliott, P. Cresswell, I. A. Wilson and R. A. Dwek, "Glycosylation and the immune system," *Science,* vol. 291, pp. 2370-2376, 2001.

[214] J. D. Marth and P. K. Grewal, "Mammalian glycosylation in immunity," *Nat Rev Immunol,* vol. 8, pp. 874-887, 2008.

[215] J. J. Bustamante, L. Gonzalez, C. A. Carroll, S. T. Weintraub, R. M. Aguilar, J. Muñoz, A. O. Martinez and L. S. Haro, "O-Glycosylated 24 kDa human growth hormone has a mucin-like biantennary disialylated tetrasaccharide attached at Thr-60," *Proteomics ,* vol. 9, pp. 3474-3488, 2009.

[216] A. Varki, R. Kannagi and B. P. Toole, "Glycosylation Changes in Cancer," in *Essentials of Glycobiology, 2nd Edition*, New York, Cold Spring Harbor Laboratory Press, 2009.

[217] Q. Yan, Q. Yao, L. L. Wei, Y. Huang, J. Myers, L. Zhang, W. Xin, J. Shim, Y. Man, B. Petryniak, S. Gerson, J. B. Lowe and L. Zhou, "O-fucose modulates notch-controlled blood lineage commitment," *The American Journal of Pathology,* vol. 176, no. 6, pp. 2921-2934, 2010.

[218] A. Seth, Q. J. Machingo, A. Fritz and B. D. Shur, "Core fucosylation is required for midline patterning during zebrafish development," *Dev Dyn,* vol. 239, pp. 3380-3390, 2010.

[219] R. Apweiler, H. Hermjakob and N. Sharon, "On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database," *Biochim Biophys Acta,* vol. 1473, p. 4–8, 1999.

[220] K. W. Moremen, M. Tiemeyer and A. V. Nairn, "Vertebrate protein glycosylation: diversity, synthesis and function," *Nat Rev Mol Cell Biol,* vol. 13, pp. 448-462, 2012.

[221] S. Mishra, S. R. Ande and N. W. Salter, "O-GlcNAc modification: why so intimately associated with phosphorylation?," *Cell Communication and Signaling,* vol. 9, no. 1, pp. 1-4, 2011.

[222] "Glycosylation analysis of protein," Creative Proteomics, 2014. [Online]. Available: http://www.creative-proteomics.com/Services/Glycosylation-analysis-of-protein.htm.

[223] A. Dell and H. R. Morris, "Glycoprotein structure determination by mass spectrometry," *Science,* vol. 291, pp. 2351-2356, 2001.

[224] W. Morelle, C. Flahaut, J. C. Michalski, A. Louvet, P. Mathurin and A. Klein, "Mass spectrometric approach for screening modifications of total serum N-glycome in human diseases: application to cirrhosis," *Glycobiology,* vol. 16, no. 4, pp. 281-293, 2006.

[225] M. Ethier, J. A. Saba, M. Spearman, O. Krokhin, M. Butler, W. Ens, K. G. Standing and H. Perreault, "Application of the StrOligo algorithm for the automated structure assignment of complex N-linked glycans from glycoproteins using tandem mass spectrometry," *Rapid Commun Mass Spectrom,* vol. 17, no. 24, pp. 2713-2720, 2003.

[226] A. E. Manzi, K. Norgard-Sumnicht, S. Argade, J. D. Marth, H. van Halbeek and A. Varki, "Exploring the glycan repertoire of genetically modified mice by isolation and profiling of the major glycan classes and nano-NMR analysis of glycan mixtures," *Glycobiology,* vol. 10, no. 7, pp. 669-689, 2000.

[227] P. M. Rudd, C. Colominas, L. Royle, N. Murphy, E. Hart, A. H. Merry, H. F. Hebestreit and R. A. Dwek, "A high-performance liquid chromatography based strategy for rapid, sensitive sequencing of N-linked oligosaccharide modifications to proteins in sodium dodecyl sulphate polyacrylamide electrophoresis gel bands," *Proteomics,* vol. 1, no. 2, pp. 285-294, 2001.

[228] N. Hashii, N. Kawasaki, S. Itoh, M. Hyuga, T. Kawanishi and T. Hayakawa, "Glycomic/glycoproteomic analysis by liquid chromatography/mass spectrometry:

analysis of glycan structural alteration in cells," *Proteomics,* vol. 5, no. 18, pp. 4665-4672, 2005.

[229] O. Blixt, S. Head, T. Mondala, C. Scanlan, M. E. Huflejt, R. Alvarez, M. C. Bryan, F. Fazio, D. Calarese, J. Stevens, N. Razi, D. J. Stevens, J. J. Skehel, I. van Die, D. R. Burton, I. A. Wilson, R. Cummings, N. Bovin, C. H. Wong and J. C. Paulson, "Printed covalent glycan array for ligand profiling of diverse glycan binding proteins," *Proc Natl Acad Sci U S A,* vol. 101, no. 49, pp. 17033-17038, 2004.

[230] X. Song, B. Xia, Y. Lasanajak, D. F. Smith and R. D. Cummings, "Quantifiable fluorescent glycan microarrays," *Glycoconj J,* vol. 25, pp. 15-25, 2008.

[231] M. Wuhrer, M. I. Catalina, A. M. Deelder and C. H. Hokke, "Glycoproteomics based on tandem mass spectrometry of glycopeptides," *J Chromatogr B Analyt Technol Biomed Life Sci,* vol. 849, no. 1-2, pp. 115-128, 2007.

[232] K. Deguchi, H. Ito, Y. Takegawa, N. Shinji, H. Nakagawa and S. Nishimura, "Complementary structural information of positive- and negative-ion MSn spectra of glycopeptides with neutral and sialylated N-glycans," *Rapid Commun Mass Spectrom,* vol. 20, no. 5, pp. 741-746, 2006.

[233] H. Ito, Y. Takegawa, K. Deguchi, S. Nagai, H. Nakagawa, Y. Shinohara and S. Nishimura, "Direct structural assignment of neutral and sialylated N-glycans of glycopeptides using collision-induced dissociation MSn spectral matching," *Rapid Commun Mass Spectrom,* vol. 20, no. 23, pp. 3557-3565, 2006.

[234] M. Wuhrer, C. A. Koeleman and A. M. Deelder, "Hexose rearrangements upon fragmentation of N-glycopeptides and reductively aminated N-glycans," *Anal Chem,* vol. 81, no. 11, pp. 4422-32, 2009.

[235] A. Napoli, D. Aiello, L. Di Donna, P. Moschidis and G. Sindona, "Vegetable Proteomics: The Detection of Ole e 1 Isoallergens by Peptide Matching of MALDI MS/MS Spectra of Underivatized and Dansylated Glycopeptides," *J. Proteome Res,* vol. 7, no. 7, pp. 2723-2732, 2008.

[236] K. D. W. Roth, Z. H. Huang, N. Sadagopan and J. T. Watson, "Charge derivatization of peptides for analysis by mass spectrometry," *Mass Spectrometry Reviews,* vol. 17, no. 4, pp. 255-274, 1998.

[237] A. L. Tarentino and T. H. Plummer, "Enzymatic deglycosylation of asparagine-linked glycans: purification, properties, and specificitly of oligosaccharide-cleaving enzymes from Flavobacterium meningosepticum," *Methods in Enzymol,* vol. 230, pp. 44-57, 1994.

[238] E. M. Taga, A. Waheed and R. L. Van Etten, "Structural and Chemical characterization of a homogeneous peptide N-glycosidase from almond," *Biochemistry,* vol. 23, no. 5, pp. 815-822, 1984.

[239] A. Kobata, "Use of endo- and exoglycosidases for structural studies of glycoconjugates," *Anal. Biochem,* vol. 100, pp. 1-14, 1979.

[240] R. B. Trimble and A. L. Tarentino, "Identification of Distinct Endoglycosidase (Endo) Activities in Flavobacterium meningosepticum: Endo F1, Endo F2 and Endo F3," *J Biochem,* vol. 266, pp. 1646-1651, 1991.

[241] G. Neubauer, A. Gottschalk, P. Fabrizio, B. Séraphin, R. Lührmann and M. Mann, "Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry," *Proc Natl Acad Sci USA,* vol. 94, no. 2, pp. 385-390, 1997.

[242] G. Neubauer, A. King, J. Rappsilber, C. Calvio, M. Watson, P. Ajuh, J. Sleeman, A. Lamond and M. Mann, "Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex," *Nat Genet,* vol. 20, no. 1, pp. 46-50, 1998.

[243] P. Cuatrecasas, M. Wilchek and C. B. Anfinsen, "Selective enzyme purification by affinity chromatography," *Proc Natl Acad Sci USA,* vol. 61, no. 2, pp. 636-643, 1968.

[244] P. Cuatrecasas and M. Wilchek, "Single-step purification of avidine from egg white by affinity chromatography on biocytin-Sepharose columns," *Biochem Biophys Res Commun,* vol. 33, no. 2, pp. 235-239, 1968.

[245] M. Wilchek, "My life with affinity," *Protein Sci,* vol. 13, no. 11, pp. 3066-3070, 2004.

[246] P. Cuatrecasas and C. B. Anfinsen, "Affinity Chromatography," *Annual Review of Biochemistry,* vol. 40, pp. 259-278 , 1971.

[247] J. Tőzsér, T. Emri and É. Csősz, "Protein purification (chromatographic techniques) and analysis (SDS-PAGE, 2DE, mass spectrometry)," in *Protein Biotechnology*, Budapest, 2011.

[248] V. I. Muronetz, M. Sholukh and T. Korpela, "Use of protein-protein interactions in affinity chromatography," *J Biochem Biophys Methods,* vol. 49, no. 1-3, pp. 29-47, 2001.

[249] A. Bauer and B. Kuster, "Affinity purification-mass spectrometry: Powerful tools for the characterization of protein complexes," *Eur J Biochem,* vol. 270, pp. 570-578, 2003.

[250] W. H. Dunham, M. Mullin and A. C. Gingras, "Affinity-purification coupled to mass spectrometry: Basic principles and strategies," *Proteomics,* vol. 12, pp. 1576-1590, 2012.

[251] K. D. Belanger, "Using affinity chromatography to investigate novel protein-protein interactions in an undergraduate cell and molecular biology lab course," *CBE Life Sci Educ,* vol. 8, no. 3, pp. 214-225, 2009 Fall.

[252] T. W. Hutchens, T. T. Yip and J. Porath, "Protein interaction with immobilized ligands: Quantitative analyses of equilibrium partition data and comparison with analytical chromatographic approaches using immobilized metal affinity adsorbents," *Analytical Biochemistry,* vol. 170, no. 1, pp. 168-182, 1988.

[253] A. Hoffmann and R. G. Roeder, "Purification of his-tagged proteins in non-denaturing conditions suggests a convenient method for protein interaction studies," *Nucleic Acids Res,* vol. 19, no. 22, pp. 6337-6338, 1991.

[254] P. Cuatrecasas, "Protein Purification by Affinity Chromatography: derivatizations of agarose and polyacrylamide beads," *The Journal of Biological Chemistry,* vol. 245, pp. 3059-3065, 1970.

[255] J. Porath, "Immobilized metal ion affinity chromatography," *Protein Expression and Purification,* vol. 3, no. 4, pp. 263-281, 1992.

[256] J. Porath and B. Olin, "Immobilized metal affinity adsorption and immobilized metal affinity chromatography of biomaterials. Serum protein affinities for gel-immobilized iron and nickel ions," *Biochemistry,* vol. 22, no. 7, pp. 1621-1630, 1983.

[257] I. Safarik and M. Safarikova, "Magnetic techniques for the isolation and purification of proteins and peptides," *BioMagnetic Research and Technology,* vol. 2, no. 7, pp. 1-17, 2004.

[258] F. Franzreb, M. Siemann-Herzberg, T. J. Hobley and O. R. T. Thomas, "Protein purification using magnetic adsorbent particles," *Applied Microbiology and Biotechnology,* vol. 70, no. 5, pp. 505-516, 2006.

[259] J. F. Peter and A. M. Otto, "Magnetic particles as powerful purification tool for high sensitive mass spectrometric screening procedures," *Proteomics,* vol. 10, no. 4, pp. 628-633, 2010.

[260] I. Fischera, C. C. Hsub, M. Gärtnera, C. Müller, T. W. Overtonb, O. R. T. Thomas and M. Franzreb, "Continuous protein purification using functionalized magnetic nanoparticles in aqueous micellar two-phase systems," *Journal of Chromatography A,* vol. 1305, pp. 7-16, 2013.

[261] J. H. Kim and J. Y. Yoon, "Protein adsorption on polymer particles," in *Encyclopedia of Surface and Colloid Science*, 2002, pp. 4373-4381.

[262] Y. Hoshino, W. W. Haberaecker, T. Kodama, Z. Zeng, Y. Okahata and K. J. Shea, "Affinity Purification of Multifunctional Polymer Nanoparticles," *J Am Chem Soc,* vol. 132, no. 39, pp. 13648-13650, 2010.

[263] C. Boyer, X. Huang, M. R. Whittaker, V. Bulmus and T. P. Davis, "An overview of protein–polymer particles," *Soft Matter,* vol. 7, pp. 1599-1614, 2011.

[264] Y. Hoshino, H. Lee and Y. Miura, "Interaction between synthetic particles and biomacromolecules: fundamental study of nonspecific interaction and design of nanoparticles that recognize target molecules," *Polymer Journal,* vol. 46, pp. 537-545, 2014.

[265] K. Saha, S. S. Agasti, C. Kim, X. Li and V. M. Rotello, "Gold nanoparticles in chemical and biological sensing," *Chem Rev,* vol. 112, no. 5, pp. 2739-2779, 2012.

[266] J. Khandare, M. Calderón, N. M. Dagia and R. Haag, "Multifunctional dendritic polymers in nanomedicine: opportunities and challenges," *Chem Soc Rev,* vol. 41, no. 7, pp. 2824-2848, 2012.

[267] M. H. Smith and L. A. Lyon, "Multifunctional nanogels for siRNA delivery," *Acc Chem Res,* vol. 45, no. 7, pp. 985-993, 2012.

[268] Y. Hoshino, H. Koide, K. Furuya, W. W. Haberaecker, S. H. Lee, T. Kodama, H. Kanazawa, N. Oku and K. J. Shea, "The rational design of a synthetic polymer nanoparticle that neutralizes a toxic peptide in vivo," *Proc Natl Acad Sci U S A,* vol. 109, no. 1, pp. 33-38, 2012.

[269] K. Yoshimatsu, B. K. Lesel, Y. Yonamine, J. M. Beierle, Y. Hoshino and K. J. Shea, "Temperature-Responsive “Catch and Release” of Proteins by using Multifunctional Polymer-Based Nanoparticles," *Angew Chem Int Ed Engl,* vol. 51, no. 10, pp. 2405-2408, 2012.

[270] Y. Yonamine, K. Yoshimatsu, S. H. Lee, Y. Hoshino, Y. Okahata and K. J. Shea, "Polymer nanoparticle-protein interface. Evaluation of the contribution of positively charged functional groups to protein affinity," *ACS Appl Mater Interfaces,* vol. 5, no. 2, pp. 374-379, 2013.

[271] M. P. Monopoli, C. Aberg, A. Salvati and K. A. Dawson, "Biomolecular coronas provide the biological identity of nanosized materials," *Nat Nanotechnol ,* vol. 7, no. 12, pp. 779-786, 2012.

[272] S. T. Kim, K. Saha, C. Kim and V. M. Rotello, "The role of surface functionality in determining nanoparticle cytotoxicity," *Acc Chem Res,* vol. 46, no. 3, pp. 681-691, 2013.

[273] N. O. Fischer, C. M. McIntosh, J. M. Simard and V. M. Rotello, "Inhibition of chymotrypsin through surface binding using nanoparticle-based receptors," *Proc Natl Acad Sci USA,* vol. 99, no. 8, pp. 5018-5023, 2002.

[274] Z. Zeng, J. Patel, S. H. Lee, M. McCallum, A. Tyagi, M. Yan and K. J. Shea, "Synthetic Polymer Nanoparticle-Polysaccharide Interactions: A Systematic Study," *J Am Chem Soc,* vol. 134, no. 5, pp. 2681-2690, 2012.

[275] W. Sun, H. Bandmann and T. Schrader, "A Fluorescent Polymeric Heparin Sensor," *Chemistry - A European Journal,* vol. 13, no. 27, pp. 7701-7707, 2007.

[276] K. Yoshimatsu, T. Yamazaki, Y. Hoshino, P. E. Rose, L. F. Epstein, L. P. Miranda, P. Tagari, J. M. Beierle, Y. Yonamine and K. J. Shea, "Epitope Discovery for a Synthetic

Polymer Nanoparticle: A New Strategy for Developing a Peptide Tag," *J Am Chem Soc,* vol. 136, no. 4, pp. 1194-1197, 2014.

[277] Y. Hoshino, T. Urakami, T. Kodama, H. Koide, N. Oku, Y. Okahata and K. J. Shea, "Design of synthetic polymer nanoparticles that capture and neutralize a toxic peptide," *Small,* vol. 5, no. 13, pp. 1562-1568, 2009.

[278] Y. Hoshino, H. Koide, T. Urakami, H. Kanazawa, T. Kodama, N. Oku and K. J. Shea, "Recognition, neutralization and clearance of target peptides in the blood stream of living mice by molecular imprinted polymer nanoparticles: a plastic antibody," *J Am Chem Soc,* vol. 132, no. 19, pp. 6644-6645, 2010.

[279] S. H. Lee, Y. Hoshino, A. Randall, Z. Zeng, P. Baldi, R. Doong and K. J. Shea, "Engineered synthetic polymer nanoparticles as IgG affinity ligands," *Journal of the American Chemical Society,* vol. 134, no. 38, pp. 15765-15772, 2012.

[280] Y. Yonamine, Y. Hoshino and K. J. Shea, "ELISA-mimic screen for synthetic polymer nanoparticles with high affinity to target proteins," *Biomacromolecules,* vol. 13, no. 9, pp. 2952-2957, 2012.

[281] Novabiochem, "Webbrochure: Fmoc resin cleavae and deprotection," Novabiochem Publications, 2007.

[282] Y. Hoshino, T. Kodama, Y. Okahata and K. J. Shea, "Peptide Imprinted Polymer Nanoparticles: A Plastic Antibody," *J Am Chem Soc,* vol. 130, no. 46, pp. 15242-15243, 2008.

[283] K. Haupt, "Biomaterials: Plastic antibodies," *Nature Materials,* vol. 9, pp. 612-614, 2010.

APPENDIX

ALLEGATO C – TANGAVEL HARIPRASAD

Il Candidato Thangavel Hariprasad ha usufruito di una borsa di studio "Marie Curie" fp 7, nell'ambito del progetto *CHEBANA.*

Il Collegio dei Docenti ha valutato l'attività di ricerca del candidato che si è sviluppata nel campo della identificazione e determinazione strutturale di proteine recettori di membrana del tipo GPER, ed ha preso in esame i risultati conseguiti, riportati in n° 1 lavoro a stampa su riviste internazionali con referee a ottimo IF medio, n° 3 comunicazioni orali in congressi internazionali, n° 1 comunicazione orale in congressi nazionali, 3 comunicazioni poster in convegni internazionali e 1 comunicazione poster in convegni nazionali.

Il Collegio ha inoltre valutato:

- La partecipazione a corsi formativi e workshop a Ginevra, Amburgo, Lund e Madrid;
- L'attività di ricerca relativa all'oggetto della tesi svolta presso l'Università della California, IRVINE, dal 8/2/2014 al 2/8/2014;
- L'attività formativa del candidato/a che si è realizzata attraverso la partecipazione a a n° 6 Convegni internazionali e n° 3 Convegni nazionali:

- l'attività formativa del candidato/a che si è realizzata a seguito della assidua frequenza all'attività didattica proposta dalla Scuola di Dottorato.

Con riferimento a quanto sopra richiamato, il Collegio dei Docenti del corso di Dottorato di Ricerca in *Scienza e Tecnica – Curriculum OMPI (Organic Materials of Pharmaceutical Interest)*, giudica l'attività del candidato/a ampiamente positiva e lo presenta con piena soddisfazione al giudizio della Commissione.

Il Coordinatore OMPI

(Prof. B. Gabriele)

ACADEMIC PUBLICATIONS, CONFERENCE & TRAINING PROGRAM PARTICIPATIONS

PUBLICATION

- Furia E, Aiello D, Di Donna L, Mazzotti F, Tagarelli A, Thangavel H, Napoli A, Sindona G. (2014). **Mass spectrometry & potentiometry studies of Pb(II)-, Cd(II)- & Zn(II)-cystine complexes"**, *Dalton Transactions*, 43(3):1055-62.
- D. Aiello, A. Napoli, S. Materazzi, R. Risoluti, F. Casadonte, H. Thangavel, C. Siciliano, G. Sindona, (2014), **Characterization of PRVB isoforms in Oncorhynchus mykiss: a mass spectrometry-based proteomics approach**, submitted to the Journal of Proteome Research.

CONFERENCE PARTICIPATIONS

- Participated in "MASSA 2012", congress organized by Division of Mass Spectrometry, Italian Chemical Society. (Palermo, Italy, 1 - 5 July 2012)
- Participated in "2nd CHEBANA summer school" jointly organized by Roche and CHEBANA network. (Mannheim, Germany, 17 - 20 September 2012)
- Participated in "2nd CHEBANA winter school" jointly organized by TuTech Innovation and CHEBANA network. (Munich, Germany, 18 - 28 February 2013)
- Participated in "VIII ItPA Annual Congress" jointly organized by Italian Proteomics Association and University of Padova. (Padova, Italy, 18 - 21 June 2013)
- Participated in "3rd CHEBANA summer school" jointly organized by Université de Technologie de Compiègne and CHEBANA network. (Biarritz, France, 1 - 7 September 2013)
- Participated in "4-MS-J-Day 'I giovani e la spettrometria di massa 2013" jointly organized by Università degli Studi della Basilicata and Italian Chemical Society - Division of Mass Spectrometry. (Potenza, Italy, 14 November 2013)
- Participated in "4th CHEBANA summer school" jointly organized by University of Regensburg and CHEBANA network. (Regensburg, Germany, 1 - 3 September 2014)
- Participated in "13th Human Proteome Organization world congress" jointly organized by HUPO, EuPA and SEProt. (Madrid, Spain, 5-8 October 2014)

POSTER COMMUNICATIONS

- Presented a poster titled **Mass spectrometry-based isolation and identification of G protein-coupled estrogen receptor (GPER)** in *2nd CHEBANA summer school* (Mannheim, Germany, Sep 2012)
- Presented a poster titled **a proteomic approach for the identification and characterization of G protein-coupled estrogen receptor 1 (GPER1)** in *VIII ItPA Annual Congress* (Padova, Italy, Jun 2013)

- Presented a poster titled **Custom-made polymer nanoparticles that capture GPER with high affinity in crude lysate** in *4th CHEBANA summer school* (Regensburg, Germany, Sep 2014)
- Presented a poster titled **Proteomic analysis of G Protein-coupled Estrogen Receptor 1 (GPER) by MALDI-TOF/TOF tandem mass spectrometry** in *13th Human Proteome Organization world congress* (Madrid, Spain, Oct 2014)

ORAL COMMUNICATIONS

- Thangavel H, et al. **Mass Spectrometry-based isolation & identification of G Protein-coupled estrogen receptor (GPER)**. *2nd CHEBANA Summer School* (Mannheim, Germany, Sep 2012)
- Thangavel H, et al. **Proteomic approaches to study glycosylation of G Protein-coupled estrogen receptor 1 (GPER1)**. *3rd CHEBANA Summer School* (Biarritz, France, Sep 2013)
- Thangavel H, et al. **Proteomic analysis of N-glycosylation in G Protein-coupled estrogen receptor 1 (GPER)**. *4-MS-J-Day 'I giovani e la spettrometria di massa 2013* (Potenza, Italy, Nov 2013)
- Thangavel H, et al. **Proteomic analysis of estrogen-reponsive G protein-coupled receptor (GPCR) by MALDI-TOF MS and MS/MS**. *4th CHEBANA summer school* (Regensburg, Germany, Sep 2014)

WORKSHOP AND TRAINING PARTICIPATIONS

- Attended training course entitled "**Bioinformatics for Mass spectrometry**" in Geneva, Switzerland from 15-19 October 2012, jointly organized by *SIB Swiss Institute for Bioinformatics and European Proteomics Association (EuPA)*.
- Attended training course on complementary skills "**Grant acquisition – How to produce a winning proposal!**" in Germany on 7 December 2012, given by *TuTech Innovation, Hamburg, Germany*.
- Attended training course entitled "**Mass spectrometry for Proteomics**" in Lund, Sweden from 10 - 14 December 2012, jointly organized by *Lund University and European Proteomics Association (EuPA)*.
- Attended training course on complementary skills "**An introduction to Innovation Management for Early Stage Researchers**" in Germany from 20-22 February 2013, given by *TuTech Innovation, Hamburg, Germany*.
- Attended workshop on "**Proteomics informatics: Mascot, Maxquant, Progenesis and Skyline**" in Madrid, Spain from 3-4 October 2014, jointly organized by *EuPA and HUPO education committees*.
- Attended one day workshop on "**Clinical proteomics**" in Madrid, Spain on 5 October 2014, organized by *HUPO*.

Mass Table *

Seq.	Mass	Seq.	Mass	Seq.	Mass	Seq.	Mass
A	71.0371	DT	216.0746	HH	274.1178	NR	270.1440
AA	142.0742	DV	214.0954	HK	265.1539	NS	201.0750
AC	174.0463	DW	301.1063	HL	250.1430	NT	215.0906
AD	186.0641	DY	278.0903	HM	268.0994	NV	213.1113
AE	200.0797	E	129.0426	HN	251.1018	NW	300.1223
AF	218.1055	EE	258.0852	HP	234.1117	NY	277.1063
AG	128.0586	EF	276.1110	HQ	265.1175	P	97.0528
AH	208.0960	EG	186.0641	HR	293.1600	PP	194.1055
AK	199.1321	EH	266.1015	HS	224.0909	PQ	225.1113
AL	184.1212	EK	257.1375	HT	238.1066	PR	253.1539
AM	202.0776	EL	242.1267	HV	236.1273	PS	184.0848
AN	185.0801	EM	260.0831	HW	323.1382	PT	198.1004
AP	168.0899	EN	243.0855	HY	300.1223	PV	196.1212
AQ	199.0957	EP	226.0953	K	128.0950	PW	283.1321
AR	227.1382	EQ	257.1012	KK	256.1899	PY	260.1161
AS	158.0691	ER	285.1437	KL	241.1790	Q	128.0586
AT	172.0848	ES	216.0746	KM	259.1354	QQ	256.1172
AV	170.1055	ET	230.0903	KN	242.1379	QR	284.1597
AW	257.1164	EV	228.1110	KP	225.1477	QS	215.0906
AY	234.1004	EW	315.1219	KQ	256.1535	QT	229.1063
C	103.0092	EY	292.1059	KR	284.1960	QV	227.1270
CC	206.0184	F	147.0684	KS	215.1270	QW	314.1379
CD	218.0361	FF	294.1368	KT	229.1426	QY	291.1219
CE	232.0518	FG	204.0899	KV	227.1634	R	156.1011
CF	250.0776	FH	284.1273	KW	314.1743	RR	312.2022
CG	160.0307	FK	275.1634	KY	291.1583	RS	243.1331
CH	240.0681	FL	260.1525	L	113.0841	RT	257.1488
CK	231.1042	FM	278.1089	LL	226.1681	RV	255.1695
CL	216.0933	FN	261.1113	LM	244.1246	RW	342.1804
CM	234.0497	FP	244.1212	LN	227.1270	RY	319.1644
CN	217.0521	FQ	275.1270	LP	210.1368	S	87.0320
CP	200.0620	FR	303.1695	LQ	241.1426	SS	174.0641
CQ	231.0678	FS	234.1004	LR	269.1852	ST	188.0797
CR	259.1103	FT	248.1161	LS	200.1161	SV	186.1004
CS	190.0412	FV	246.1368	LT	214.1317	SW	273.1113
CT	204.0569	FW	333.1477	LV	212.1525	SY	250.0954
CV	202.0776	FY	310.1317	LW	299.1634	T	101.0477
CW	289.0885	G	57.0215	LY	276.1474	TT	202.0954
CY	266.0725	GG	114.0429	M	131.0405	TV	200.1161
D	115.0269	GH	194.0804	MM	262.0810	TW	287.1270
DD	230.0539	GK	185.1164	MN	245.0834	TY	264.1110
DE	244.0695	GL	170.1055	MP	228.0933	V	99.0684
DF	262.0953	GM	188.0620	MQ	259.0991	VV	198.1368
DG	172.0484	GN	171.0644	MR	287.1416	VW	285.1477
DH	252.0859	GP	154.0742	MS	218.0725	VY	262.1317
DK	243.1219	GQ	185.0801	MT	232.0882	W	186.0793
DL	228.1110	GR	213.1226	MV	230.1089	WW	372.1586
DM	246.0674	GS	144.0535	MW	317.1198	WY	349.1426
DN	229.0699	GT	158.0692	MY	294.1038	Y	163.0633
DP	212.0797	GV	156.0899	N	114.0429	YY	326.1267
DQ	243.0855	GW	243.1008	NN	228.0859		
DR	271.1281	GY	220.0848	NP	211.0957		
DS	202.0590	H	137.0589	NQ	242.1015		

* To calculate total mass of any two amino acid residues, first put the 2 single-letter codes in aphabetical order, then look them up in the table.

Reverse Mass Table*

Mass	Seq.	Mass	Seq.	Mass	Seq.	Mass	Seq.
16.0313	$[CH_4]$	199.0957	AQ, AAG	231.0855	GSS	256.1172	QQ, AAN, AQG
17.0265	$[NH_3]$	199.1321	AK	231.1042	CK	256.1535	QK, AGK
18.0106	$[H_2O]$	200.0620	CP	232.0518	CE	256.1899	KK
57.0215	G	200.0797	AE	232.0882	MT	257.0834	PC(+57), CGP
71.0371	A	200.1161	LS, TV	234.0497	CM	257.1012	EQ, AAD, ABG
87.0320	S	201.0750	NS, GGS	234.0674	SM(+16)	257.1164	AW
97.0528	P	202.0590	DS	234.1004	AY, FS	257.1376	EK, ASV, GLS GTV
99.0684	V	202.0776	AM, CV	234.1117	HP		
101.0477	T	202.0954	TT	236.1273	HV	257.1488	RT
103.0092	C	204.0569	CT, GM(+16)	238.0355	AS(+80), GT(+80)	258.0852	EE
113.0841	L	204.0899	GF	238.1066	HT	258.0964	NGS, GGGS
114.0429	N, GG	206.0184	CC	239.1270	AAP	259.0804	DGS
115.0269	D	208.0960	AH	240.0681	CH	259.0991	QM, VC(+57), AGM, CGV
128.0586	Q, AG	210.1368	LP	241.1063	GPS		
128.0950	K	211.0957	NP, GGP	241.1426	QL, AAV, AGL	259.1103	RC
129.0426	E	212.0797	DP	241.1790	LK	259.1168	AST, GTT
131.0405	M	212.1525	LV	242.1015	NQ, ANG, QGG, AGGG	259.1355	KM
137.0589	H	213.1113	NV, AAA, GGV			260.0831	EM
142.0742	AA	213.1226	RG	242.1266	EL	260.1161	PY
144.0535	GS	214.0954	DV	242.1379	NK, GGK	260.1195	LM(+16)
147.0354	M(+16)	214.1317	LT	243.0297	Y(+80)	260.1525	LF
147.0684	F	215.0906	NT, QS, AGS, GGT	243.0855	NE, DQ, ADG, EGG	261.0783	ACS, CGT, NM(+16), TC(+57), GGM(+16)
154.0742	GP						
156.0899	GV	215.1270	KS	243.1008	GW		
156.1011	R	216.0746	DT, ES	243.1219	DK, AAT, GSV	261.0961	SSS
158.0691	AS, GT	216.0933	CL	243.1331	RS	261.1113	NF, GGF
160.0307	C(+57), CG	217.0521	NC, GC(+57), CGG	244.0695	DE	262.0623	DM(+16)
163.0633	Y	218.0361	DC	244.0882	PM(+16)	262.0810	MM
166.9984	S(+80)	218.0725	AM(+16), MS	244.1212	FP	262.0954	DF
168.0899	AP	218.1055	AF	244.1246	LM	262.1317	YV
170.1055	AV, GL	220.0848	GY	245.0834	NM, AAC, GGM	263.0398	CC(+57), CCG
171.0644	NG, GGG	224.0198	GS(+80)	245.1012	ASS, GST	264.0511	PS(+80)
172.0484	DG	224.0909	HS	246.0674	DM	264.1110	TY
172.0848	AT	225.1113	QP, AGP	246.1038	VM(+16)	265.1175	QH, AGH
174.0463	AC	225.1477	KP	246.1368	FV	265.1426	APP
174.0641	SS	226.0954	EP	247.0627	SC(+57), CGS	265.1539	HK
181.0140	T(+80)	226.1681	LL	248.0831	TM(+16)	266.0668	VS(+80)
184.0848	PS	227.1270	NL, QV, AGV, GGL	248.1161	FT	266.0725	CY
184.1212	AL			250.0446	CM(+16)	266.1015	EH
185.0800	AN,QG,AGG	227.1382	AR	250.0776	CF	267.1583	APV, GLP
185.1164	GK	227.1634	KV	250.0954	SY	268.0460	ST(+80), TS(+80)
186.0640	EG	228.0858	GGGG	250.1430	HL	268.0994	HM
186.0641	AD	228.0859	NN, NGG	251.1018	NH, GGH	268.1171	GGGP
186.0793	W	228.0933	MP	251.1270	GPP	268.1172	NGP
186.1004	SV	228.1110	DL, EV	252.0511	AT(+80)	269.1012	DGP
188.0620	GM	229.0699	ND, DGG	252.0859	DH	269.1376	APT
188.0797	ST	229.1063	QT, AAS, AGT	253.1426	GPV	269.1739	AVV, GLV
190.0412	CS	229.1426	KT	253.1539	RP	269.1852	RL
194.0804	GH	230.0539	DD	254.0304	SS(+80)	270.0076	CS(+80)
194.1055	PP	230.0903	ET	255.1219	APS, GPT	270.1328	AAQ, NGV, AAAG, GGGV
196.1212	PV	230.1089	MV	255.1583	AAL, GVV		
198.1004	PT	231.0678	AC(+57), CQ, ACG	255.1695	RV	270.1440	RN, RGG
198.1368	VV			256.1171	AAGG	270.1692	AAK

* Three common PTMs are represented by their nominal mass in the sequence:
+57 - Carbamidomethyl (57.02146), +80 - Phosphorylation (79.96633), +16 - Oxidation on M (15.99491)

Common Amino Acid Residues

Name	3-letter Symbol	1-letter Symbol	Monoisotopic Mass	Average Mass	Residue Composition	Residue Structure
Alanine	Ala	A	71.03711	71.08	C_3H_5NO	CH_3 -NH-CH-CO-
Arginine	Arg	R	156.10111	156.2	$C_6H_{12}N_4O$	$-CH_2-(CH_2)_2-NH-C(=NH)-NH_2$ -NH-CH_2-CO-
Asparagine	Asn	N	114.04293	114.1	$C_4H_6N_2O_2$	CH_3-CONH_2 -NH-CH-CO-
Aspartic Acid	Asp	D	115.02694	115.1	$C_4H_5NO_3$	CH_3-COOH -NH-CH-CO-
Cysteine	Cys	C	103.00919	103.1	C_3H_5NOS	CH_2-SH -NH-CH-CO-
Glutamic Acid	Glu	E	129.04259	129.1	$C_5H_7NO_3$	CH_2-CH_2-COOH -NH-CH-CO-
Glutamine	Gln	Q	128.05858	128.1	$C_5H_8N_2O_2$	$CH_2-CH_2-CONH_2$ -NH-CH-CO-
Glycine	Gly	G	57.02146	57.05	C_2H_3NO	-NH-CH_2-CO-
Histidine	His	H	137.05891	137.1	$C_6H_7N_3O$	CH_2-(imidazole)- -NH-CH_2-CO-
Isoleucine	Ile	I	113.08406	113.2	$C_6H_{11}NO$	$CH(CH_3)-CH_2-CH_3$ -NH-CH-CO-
Leucine	Leu	L	113.08406	113.2	$C_6H_{11}NO$	$CH_2-CH(CH_3)_2$ -NH-CH-CO-
Lysine	Lys	K	128.09496	128.2	$C_6H_{12}N_2O$	$CH_2-(CH_2)_3-NH_2$ -NH-CH-CO-
Methionine	Met	M	131.04049	131.2	C_5H_9NOS	$CH_2-CH_2-S-CH_3$ -NH-CH-CO-
Phenyalanine	Phe	F	147.06841	147.2	C_9H_9NO	CH_2-(benzene)- -NH-CH-CO-
Proline	Pro	P	97.05276	97.12	C_5H_7NO	-N-(cyclopentane)-CH-CO-
Serine	Ser	S	87.03203	87.08	$C_3H_5NO_2$	CH_2-OH -NH-CH-CO-
Threonine	Thr	T	101.04768	101.1	$C_4H_7NO_2$	$CH(OH)-CH_3$ -NH-CH-CO-
Tryptophan	Trp	W	186.07931	186.2	$C_{11}H_{10}N_2O$	$-CH_2$-(indole)- -NH-CH_2-CO-
Tyrosine	Tyr	Y	163.06333	163.2	$C_9H_9NO_2$	$-CH_2$-(benzene)-OH -NH-CH_2-CO-
Valine	Val	V	99.06841	99.13	C_5H_9NO	$CH(CH_3)_2$ -NH-CH-CO-

Common PTMs

Monoisotopic Mass Δ	Position	Modification
-29.99281	[M]@C-term	Homoserine
-18.01057	Y, T, S, [N, Q]@C-term, [C]@N-term	Dehydration
-17.02655	[Q]@N-term	Pyroglutamic Acid from Gln
-2.01565	C	Disulphide Bond Formation
-0.98402	[X]@C-term	Amidation
0.98402	R	Citrullination
0.98402	R, N, Q	Deamidation
14.01565	T, S, E, D, L, I, [X]@N-term, R, Q, N, K, H, C, [X]@C-term	Methylation
15.99491	W, H, C, M	Oxidation
27.99491	T, K, S, [X]@N-term	Formylation (CHO)
42.01056	C, S, K, [X]@N-term	Acetylation
42.02180	K	Homoarginine
43.98983	W, K, D, E, [M]@N-term	Carboxylation
57.02146	C	Carbamidomethyl
58.00548	C	Carboxymethyl
79.96633	R, C, D, Y, H, T, S	Phosphorylation
79.95682	Y	Sulphation
119.00410	C	Cysteinylation
132.04226		Pentoses (Ara, Rib, Xyl)
146.05791		Deoxyhexoses (Fuc, Rha)
161.06881	K, N, T, W	Hexosamine (GalN, GlcN)
162.05282	Y, R, C, T, W, N, K, [X]@N-term	Hexoses (Fru, Gal, Glc, Man)
188.03296	K	Lipoic Acid (Amide Bond to Lys)
203.07937	S, T, N	N-acetylhexosamine (GalNAc, GlcNAc)
204.18780	C	Farnesylation
210.19836	K, C, [G]@N-term	Myristoylation
226.07760	K, [X]@N-term	Biotinylation (Amide Bond to Lys)
238.22966	C, K, S, T, [X]@N-term	Palmitoylation
272.25040	C	Geranylgeranlylation
365.13220		Hex-HexNAc
541.06111	R, C, N, S, E	ADP-Ribosylation (from NAD)

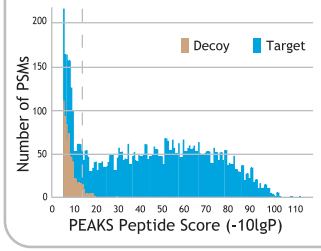
Fragment Ions

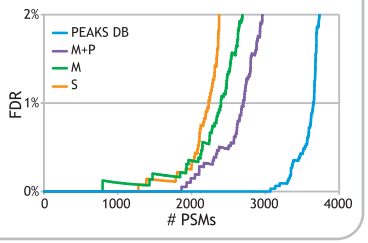
Ion Type	Mass	Nominal Mass
a	[M]- 26.9871	[M]- 27
b	[M]+ 1.0078	[M]+ 1
c	[M]+ 18.0344	[M]+ 18
x	[M]+ 44.9977	[M]+ 45
y	[M]+ 19.0184	[M]+ 19
z	[M]+ 1.9918	[M]+ 2
z°	[M]+ 2.9997	[M]+ 3

PEAKS®
www.bioinfor.com

False Discovery Rate

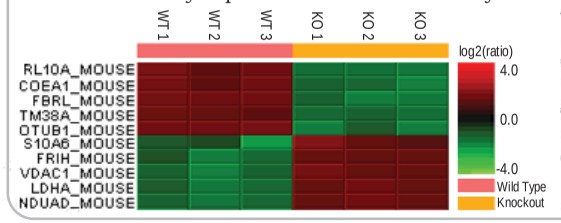
A target protein DB is mixed with a decoy DB, and searched by peptide ID software. FDR = #decoy hits/#target hits. FDR is used to validate ID results (left fig.) and compare different software (right fig.)

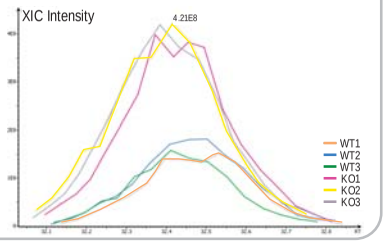




PEAKS® Q Label Free Quantification

Beyond accurate and sensitive protein identification, optional PEAKS Q adds a quantification solution. For label free quantification, peptide ratios are calculated based on peak intensities, and protein ratios are then calculated using the peak intensities of the top-3 unique peptide features, as seen in the heatmap (left fig.). Additional utilities include: peptide features quantified without identification can provide valuable new information; replicates are used for significance assessment, as demonstrated in figure XIC (right fig.); and intensities of peptide features are easily exported for customized analyses.





Manual MS/MS Interpretation

Manual Interpretation of Peptide MS/MS spectra

Interpretation check list

---

1 Accurate mass measurement of native parent

2.1 Depending on parent charge, deconvolute spectrum to all 1+ ions. Remove all spikes

2.2 Print out spectrum as is, and then with all regions scaled up to the same size

2.3 Label possible a-b pairs (a=b-28)

2.4 Label -18 and -17 losses of ^{0} water and * ammonia respectively

3.1 Examine low mass end for R/K y_1 ions (175 or 147) if tryptic

3.2 Look for PHW immonium ions (Table 1)

3.3 If PHW present, look for intense internal cleavage series

3.4 Inspect high mass end for b and y ions and residue mass loss (Table 2)
Last b_{ion} is MH^+ - 18 - residue mass; first y_{ion} is MH^+ - residue mass
Ignore ions between MH^+ and MH^+ -60 (loss of H_2O, NH_3, HCOOH, CH_3COOH)

4.1 Look for ion series by sequentially subtracting residue masses (Table 2)

4.2 Look for corresponding b/y ion pairs (y + b = parent mass +1)

4.3 For each residue look for diagnostic losses from it and respective immonium ions

4.4 Be careful about dipeptide masses that match residue masses (Table 3)

5.1 Assign tentative sequence. Calculate MH+ from sequence

5.2 Compare all theoretical masses to data. Are all ions accounted for? If not repeat process

---

Manual MS/MS Interpretation

Table 2. Residue masses

Residue	Mass	Possible equivalent mass	Behaviour in MS/MS
Gly	57		Gives weak signal
Ala	71		-
Ser	87		Loses –18 (water)
Pro	97		Strong signal after C-term cleavage giving internal ions. Look for 70
Val	99	AcGly (N-term only)	
Thr	101		Loses –18 (water)
Cys	103		Unusual, always modified. Loses –34 (H2S)
Pyro-Glu	111		N-terminal only
Leu/Ile	113	AcAla (N-term only)	-
Asn	114	Gly-Gly	Loses 17 (ammonia)
Asp	115		Cleaves N-terminal to give strong signal
Lys	128.09	Gly-Ala	-
Gln	128.06	Gly-Ala	Strong loss of –17
Glu	129	AcSer (N-term only)	-
Met	131		Loss of CH3SH -48
His	137		Like Pro but weaker internal. Look for 110
Phe	147		Can be Met(Ox)
Arg	156	Gly-Val/AcAsn (N-term)	-
CmCys	161		-
Tyr	163		No loss of 18
Trp	186	Gly-Glu or Ala-Asp or Ser-Val	Like Pro and His but much weaker internals

Table 3. Dipeptide masses

	Gly	Ala	Ser	Pro	Val	Thr	Cys	L/I	Asn	Asp	K/Q	Glu	Met	His	Phe	Arg	CmC	Tyr	Trp
Gly	114																		
Ala	128	142																	
Ser	144	158	174																
Pro	154	168	184	194															
Val	156	170	186	196	198														
Thr	158	172	188	198	200	202													
Cys	160	174	190	200	202	204	206												
L/I	170	184	200	210	212	214	216	226											
Asn	171	185	201	211	213	215	217	227	228										
Asp	172	186	202	212	214	216	218	228	229	230									
K/Q	185	199	215	225	227	229	231	241	242	243	256								
Glu	186	200	216	226	228	230	232	242	243	244	257	258							
Met	188	202	218	228	230	232	234	244	245	246	259	260	262						
His	194	208	224	234	236	238	240	250	251	252	265	266	268	274					
Phe	204	218	234	244	246	248	250	260	261	262	275	276	278	284	294				
Arg	213	227	243	253	255	257	259	269	270	271	284	285	287	293	303	312			
CmC	218	232	248	258	260	262	264	274	275	276	289	290	292	298	308	317	322		
Tyr	220	234	250	260	262	264	266	276	277	278	291	292	294	300	310	319	326	326	
Trp	243	257	273	283	285	287	289	299	300	301	314	315	317	323	333	342	349	349	272

Commonly observed ESI contaminant, polymer & cluster peaks

Ion mass (Da)	*Ion Mode*	*Notes*	*Ref*
+18n	+ / −	Water clusters	a
+32n	+	Methanol clusters	a
+41n	+	Acetonitrile clusters	a
+44n	+	Polyethylene glycol related $(-CH_2CH_2O-)_n$	a
+53n	+	Ammonium chloride adducts (NH_4Cl)	a
+58n	+	Polypropylene glycol related $(CH_2CH_2CH_2O)_n$	a
+58n	+	NaCl adducts (distinguishable from PPG by Cl isotope pattern)	a
+63n	+	Ammonium formate adducts ($HCOONH_4$)	a
+68n	+ / −	Sodium formate adducts (HCOO-Na)	a
+72n	+	Replacement of –OH by $–OSi(CH_3)_3$	a
+74n	+ / −	Polysiloxanes ($Si(CH_3)_2O$)	a
+74n	+	KCl adducts (distinguishable by Cl isotope pattern)	a
+77n	+	Ammonium acetate salts (CH_3COONH_4)	a
+78n	+	DMSO adducts	a
+82n	+ / −	Sodium acetate adducts (CH_3COONa)	a
+114n	−	TFA (trifluoroacetic acid) adducts (CF_3COOH)	a
+122n	−	Sodium perchlorate adducts	a
+136n	+ / −	Sodium TFA (trifluoroacetic acid) adducts (CF_3COONa)	a
+162n	+	Polysaccharides	a
+288n	−	SDS (sodium dodecylsulfate) adducts	a
231, 522, 550	+	From rubber tip of disposable syringe plunger	a
297, 371, 445, 519 (+74n)	+	Polysiloxanes, from air, silicone rubber	c
355, 429, 503, 593, 667, 741, 815 (+74n)	+	Polysiloxanes, from air, silicone rubber	a
391, 413	+	Dioctyl pthlalate & sodium salt	c
538, 556, 574	+	Iron carboxylate complexes: 538 = $Fe_3O(AcO)_6$, 556 = $Fe_3O(AcO)_6(H_2O)$, 574 = $Fe_3O(AcO)_6(H_2O)_2$	b

[a] Tong, Bell, Tabei, Siegel, *JASMS* **1999**, *10*, 1174-1187.
[b] Ijames, Dutky, Fales, *JASMS* **1995**, *7*, 1226-1231.
[c] unpublished

Please send suggestions and corrections to allis@stanford.edu

Mutation Mass Shifts (2/2: Monoisotopic Mass)

Koichi Tanaka Laboratory of Advanced Science and Technology

Full Name	3-letter Code	Empirical Formula (Residual)	Monoiso. Mass (Res.)	Gly	Ala	Ser	Pro	Val	Thr	Cys	Ile	Leu	Asn	Asp	Gln	Lys	Glu	Met	His	Phe	Arg	Tyr	Trp
RNA Codon List			Stop: UAA UAG UGA	GGU GGC GGA GGG	GCU GCC GCA GCG	UCU UCC UCA UCG AGU AGC	CCU CCC CCA CCG	GUU GUC GUA GUG	ACU ACC ACA ACG	UGU UGC	AUU AUC AUA	UUA UUG CUU CUC CUA CUG	AAU AAC	GAU GAC	CAA CAG	AAA AAG	GAA GAG	AUG	CAU CAC	UUU UUC	CGU CGC CGA CGG AGA AGG	UAU UAC	UGG
				H–	CH_3–	HO–CH_2–	N, CO (ring)	$(H_3C)_2CH$–	$CH_3CH(OH)$–	HS–CH_2–	$CH_3CH_2CH(CH_3)$–	$(H_3C)_2CHCH_2$–	H_2NCOCH_2–	(−) $HOOCCH_2$–	$H_2NCOCH_2CH_2$–	(+) $H_2NCH_2CH_2CH_2CH_2$–	(−) $HOOCCH_2CH_2$–	CH_3–S–CH_2CH_2–	(+) N, NH (ring)–CH_2–	(ring)–CH_2–	(+) H_2N, HN=CNHCH$_2$CH$_2$CH$_2$–	HO–(ring)–CH_2–	(ring, NH)–CH_2–
				G	A	S	P	V	T	C	I	L	N	D	Q	K	E	M	H	F	R	Y	W
				57.02146	71.03711	87.03203	97.05276	99.06841	101.04768	103.00918	113.08406	113.08406	114.04293	115.02694	128.05858	128.09496	129.04259	131.04048	137.05891	147.06841	156.10111	163.06333	186.07931
Glycine	Gly	C_2H_3NO	57.02146	0.0000	14.0157	30.0106	40.0313	42.0470	44.0262	45.9877	56.0626	56.0626	57.0215	58.0055	71.0371	71.0735	72.0211	74.0190	80.0374	90.0470	99.0796	106.0419	129.0578
Alanine	Ala	C_3H_5NO	71.03711	-14.0157	0.0000	15.9949	26.0157	28.0313	30.0106	31.9721	42.0470	42.0470	43.0058	43.9898	57.0215	57.0578	58.0055	60.0034	66.0218	76.0313	85.0640	92.0262	115.0422
Serine	Ser	$C_3H_5NO_2$	87.03203	-30.0106	-15.9949	0.0000	10.0207	12.0364	14.0157	15.9772	26.0520	26.0520	27.0109	27.9949	41.0265	41.0629	42.0106	44.0085	50.0269	60.0364	69.0691	76.0313	99.0473
Proline	Pro	C_5H_7NO	97.05276	-40.0313	-26.0157	-10.0207	0.0000	2.0157	3.9949	5.9564	16.0313	16.0313	16.9902	17.9742	31.0058	31.0422	31.9898	33.9877	40.0061	50.0157	59.0483	66.0106	89.0265
Valine	Val	C_5H_9NO	99.06841	-42.0470	-28.0313	-12.0364	-2.0157	0.0000	1.9793	3.9408	14.0157	14.0157	14.9745	15.9585	28.9902	29.0265	29.9742	31.9721	37.9905	48.0000	57.0327	63.9949	87.0109
Threonine	Thr	$C_4H_7NO_2$	101.04768	-44.0262	-30.0106	-14.0157	-3.9949	-1.9793	0.0000	1.9615	12.0364	12.0364	12.9952	13.9793	27.0109	27.0473	27.9949	29.9928	36.0112	46.0207	55.0534	62.0157	85.0316
Cysteine	Cys	C_3H_5NOS	103.00918	-45.9877	-31.9721	-15.9772	-5.9564	-3.9408	-1.9615	0.0000	10.0749	10.0749	11.0337	12.0178	25.0494	25.0858	26.0334	28.0313	34.0497	44.0592	53.0919	60.0541	83.0701
Isoleucine	Ile	$C_6H_{11}NO$	113.08406	-56.0626	-42.0470	-26.0520	-16.0313	-14.0157	-12.0364	-10.0749	0.0000	0.0000	0.9589	1.9429	14.9745	15.0109	15.9585	17.9564	23.9748	33.9843	43.0170	49.9793	72.9952
Leucine	Leu	$C_6H_{11}NO$	113.08406	-56.0626	-42.0470	-26.0520	-16.0313	-14.0157	-12.0364	-10.0749	0.0000	0.0000	0.9589	1.9429	14.9745	15.0109	15.9585	17.9564	23.9748	33.9843	43.0170	49.9793	72.9952
Asparagine	Asn	$C_4H_6N_2O_2$	114.04293	-57.0215	-43.0058	-27.0109	-16.9902	-14.9745	-12.9952	-11.0337	-0.9589	-0.9589	0.0000	0.9840	14.0157	14.0520	14.9997	16.9976	23.0160	33.0255	42.0582	49.0204	72.0364
Aspartic acid	Asp	$C_4H_5NO_3$	115.02694	-58.0055	-43.9898	-27.9949	-17.9742	-15.9585	-13.9793	-12.0178	-1.9429	-1.9429	-0.9840	0.0000	13.0316	13.0680	14.0157	16.0135	22.0320	32.0415	41.0742	48.0364	71.0524
Glutamine	Gln	$C_5H_8N_2O_2$	128.05858	-71.0371	-57.0215	-41.0265	-31.0058	-28.9902	-27.0109	-25.0494	-14.9745	-14.9745	-14.0157	-13.0316	0.0000	0.0364	0.9840	2.9819	9.0003	19.0098	28.0425	35.0048	58.0207
Lysine	Lys	$C_6H_{12}N_2O$	128.09496	-71.0735	-57.0578	-41.0629	-31.0422	-29.0265	-27.0473	-25.0858	-15.0109	-15.0109	-14.0520	-13.0680	-0.0364	0.0000	0.9476	2.9455	8.9639	18.9735	28.0061	34.9684	57.9843
Glutamic acid	Glu	$C_5H_7NO_3$	129.04259	-72.0211	-58.0055	-42.0106	-31.9898	-29.9742	-27.9949	-26.0334	-15.9585	-15.9585	-14.9997	-14.0157	-0.9840	-0.9476	0.0000	1.9979	8.0163	18.0258	27.0585	34.0207	57.0367
Methionine	Met	C_5H_9NOS	131.04048	-74.0190	-60.0034	-44.0085	-33.9877	-31.9721	-29.9928	-28.0313	-17.9564	-17.9564	-16.9976	-16.0135	-2.9819	-2.9455	-1.9979	0.0000	6.0184	16.0279	25.0606	32.0228	55.0388
Histidine	His	$C_6H_7N_3O$	137.05891	-80.0374	-66.0218	-50.0269	-40.0061	-37.9905	-36.0112	-34.0497	-23.9748	-23.9748	-23.0160	-22.0320	-9.0003	-8.9639	-8.0163	-6.0184	0.0000	10.0095	19.0422	26.0044	49.0204
Phenylalanine	Phe	C_9H_9NO	147.06841	-90.0470	-76.0313	-60.0364	-50.0157	-48.0000	-46.0207	-44.0592	-33.9843	-33.9843	-33.0255	-32.0415	-19.0098	-18.9735	-18.0258	-16.0279	-10.0095	0.0000	9.0327	15.9949	39.0109
Arginine	Arg	$C_6H_{12}N_4O$	156.10111	-99.0796	-85.0640	-69.0691	-59.0483	-57.0327	-55.0534	-53.0919	-43.0170	-43.0170	-42.0582	-41.0742	-28.0425	-28.0061	-27.0585	-25.0606	-19.0422	-9.0327	0.0000	6.9622	29.9782
Tyrosine	Tyr	$C_9H_9NO_2$	163.06333	-106.0419	-92.0262	-76.0313	-66.0106	-63.9949	-62.0157	-60.0541	-49.9793	-49.9793	-49.0204	-48.0364	-35.0048	-34.9684	-34.0207	-32.0228	-26.0044	-15.9949	-6.9622	0.0000	23.0160
Tryptophan	Trp	$C_{11}H_{10}N_2O$	186.07931	-129.0578	-115.0422	-99.0473	-89.0265	-87.0109	-85.0316	-83.0701	-72.9952	-72.9952	-72.0364	-71.0524	-58.0207	-57.9843	-57.0367	-55.0388	-49.0204	-39.0109	-29.9782	-23.0160	0.0000

⊥ : N–C(H)–C(=O), H

	Monoisotopic Mass
H	1.00782504
C	12.0000000
N	14.0030740
O	15.9949146
S	31.9720710

0<|Δm|<10	10<|Δm|<20	20<|Δm|<30	30<|Δm|<40	40<|Δm|<50	50<|Δm|<70	70<|Δm|<100	100<|Δm|

Residues DOWN the left indicate the EXPECTED residues.
Residues ACROSS the top indicate the MUTANT residues.

Ref: <http://en.wikipedia.org/wiki/Genetic_code>, <http://prowl.rockefeller.edu/aainfo/mutation.html>

ms3d FIRST Program

SHIMADZU

Rev. 3.0 17/October/2010

GlycoMod

GlycoMod / GlycanMass tools

Mass values in Dalton, used in GlycoMod and GlycanMass

		underivatised		permethylated		peracetylated	
		monoisotopic	average	monoisotopic	average	monoisotopic	average
Hexose (Hex)	e.g. galactose, mannose, glucose	162.0528	162.1424	204.0998	204.2230	288.0845	288.2542
HexNAc	e.g. N-acetylglucosamine, N-acetylgalactosamine	203.0794	203.1950	245.1263	245.2756	287.1005	287.2695
Deoxyhexose	e.g. fucose, rhamnose	146.0579	146.1430	174.0892	174.1968	230.0790	230.2176
Pentose	e.g. xylose	132.0423	132.1161	160.0736	160.1699	216.0634	216.1907
NeuAc	= N-acetyl neuraminic acid (sialic acid)	291.0954	291.2579	361.1737	361.3923	417.1271	417.3698
NeuGc	= N-glycolyl neuraminic acid	307.0903	307.2573	391.1842	391.4186	475.1326	475.4064
KDN	= Ketodeoxynonulosonic acid	250.0689	250.2053	320.1472	320.3397	376.1006	376.3171
HexA	e.g. hexuronic acid	176.03209	176.1259	218.0790	218.2066	260.0532	260.2005
Acetate		42.0106	42.0373	-	-	-	-
H^+		1.00727	1.00739	-	-	-	-
H		1.00783	1.00794	-	-	-	-
H_2O		18.01056	18.01524	-	-	-	-
K^+		38.963707	39.0983	-	-	-	-
Na^+		22.989768	22.998977	-	-	-	-
Phosphate		79.9663	79.9799	93.9820	94.0068	37.9558	37.9426
Sulphate		79.9568	80.0642	65.9412	66.0373	37.9463	38.0269
Trifluoroacetic acid		112.9850391	113.0160096	–	–	–	–

SDS-PAGE & Western Blotting Protocol

SAMPLE PREPARATION

1, Different cancer cell lines were used in the study.

Cancer cells	Cancer type	Growth media
BG-1	Ovarian	DMEM without phenol red supplemented with 10% FBS
Ishikawa	Endometrial	DMEM without phenol red supplemented with 10% FBS
SkBr3	Breast	RPMI 1640 without phenol redsupplemented with 10% FBS

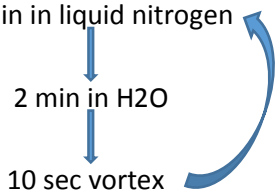
DMEM - Dulbecco's Modified Eagle Medium, FBS - Fetal Bovine Serum

2, all cells were grown in 10 cm petri dishes and harvested.
3, during cell lysis, the culture medium was aspirated out.
4, the cells were washed with ice cold 1x PBS for 2 times and the excess 1x PBS was removed.
5, the grown cells were lysed in 200 µL of the following lysis buffer containing a mixture of protease inhibitors.

Lysis Buffer	Protease inhibitors
50 mM HEPES, pH 7.5 150 mM NaCl 1.5 mM MgCl2 1 mM EGTA 10% glycerol 1% Triton X-100 1% SDS	1 mM aprotinin 20 mM PMSF 200 mM sodium orthovanadate

*As soon as lysis occurs, proteolysis, dephosphotylation and denaturation begin. These events can be slowed down tremendously if samples are kept on ice or at 4 ^{o}C at all times and appropriate inhibitors are added fresh to the lysis buffer.
6, the cells were scraped thoroughly after addition of 200 µL lysis buffer with protease inhibitors.
7, the contents on the 10 cm petridishes were transferred to 1.5 ml tubes.
8, followed by sonication or 3 cycles of freeze thawing to rupture the cell wall and release the cellular contents.

Freeze thaw (3 cylcles) –



9, the tubes were spun down at maximum speed (14,000 rpm) for 5 to 8 minutes.
10, 200 µL of supernatant were taken and transferred to a new 1.5 mL tubes.
11, all supernatant fractions from same batch (same cells grown in different petridishes) were pooled together.
12, total protein concentration was determined using Bradford assay or BCA assay. BSA was used as protein standard.
*The samples can be stored at this point at -20 ^{o}C or -80 ^{o}C for later use.

ELECTROPHORESIS

13, the protein samples were denatured by adding 2x loading (Laemmli) buffer in the ratio 1:1 and boiling the mixture at 95 ^{o}C – 100 ^{o}C for 5 minutes.

2x Laemmli sample buffer (Bio-rad)

65.8 mM Tris-HCl, pH 6.8
26.3% (w/v) glycerol
2.1% SDS
0.01% bromophenol blue

950 µL + 50 µL 2-mercaptoethanol

14, while loading, the protein concentration and the sample volume should be same across all the lanes.
15, 12% Mini-PROTEAN TGX precast polyacrylamide gels with 15 sample loading wells from Bio-rad were used for electrophoretic protein separation.
16, 15 µL (7.5 µL protein sample + 7.5 µL loading buffer) were loaded onto each well on the gel along with positive control and protein standard (Bio-rad Precision Plus Protein Kaleidoscope standards).
17, the proteins were resolved in the presence of 1x running buffer and by applying a voltage of 160V - 200V to the system for about 30 - 45 minutes.

5x Running buffer (for 1L)

15.1g Tris base, pH 8.3
94g glycine
5g SDS
to 1L water

18, the power was turned off when the migration front reaches the bottom of the gel.
*Proteins will slowly elute from the gel at this point, so do not store the gel; proceed immediately to transfer.

PROTEIN TRANSFER

19, 1x Transfer buffer was prepared freshly and chilled at 4 ^{o}C.

20x Transfer buffer (for 1L)

31.32g Tris-HCl
150.14g glycine
To 1L water

20, the sample loading wells were cut off and top left-hand corner of the gel was nicked for orientation.
21, the dimensions of the gel was measured and the positions of the protein standards' bands were noted.
22, the gel was transferred to a box containing 1x transfer buffer and agitated for 10 min at RT ^{o} to remove salts and SDS.
23, a piece of nitrocellulose membrane was cut to the size of the gel and immersed in 1x transfer buffer for 10 mins prior to use.
24, 4 - 6 pieces of 3mm filter paper was cut to the dimensions of the gel (or slightly bigger).
25, a gel holder cassette was opened in a casserole dish filled 1/4th with 1x transfer buffer, black side down and hinges to the left and below the black side.
26, a fiber pad was soaked in 1x transfer buffer and placed in the center of the black side.
27, 2 pieces of filter paper was soaked in 1x transfer buffer and placed on top of the fiber pad.
28, the gel was removed with a glass plate and placed on top of the filter paper.
29, the nitrocellulose membrane was placed on top of the gel followed by 2 pieces of filter paper soaked with 1xt transfer buffer.
30, bubbles were rolled out with a glass tube.
31, a second fiber pad was soaked in 1x transfer buffer and placed on top of stack.
32, the gel holder cassette was closed and locked.
*2 mL of 1x transfer buffer was added to the stack after each steps (26-31) to ensure wet transfer.

33, the cassette was placed in a transfer tank (orient the white and black sides of the cassette with the red and black panels of the electrode respectively) and filled with ~800 mL of 1x transfer buffer.
34, the entire tank setup was placed in a Styrofoam box or plastic tub containing ice.
35, the protein transfer was performed at 100 V for about 2 hr.

VISUALIZATION IN MEMBRANE

36, after electro transfer, the membrane was rinsed in 1x TBST.

1x TBST (for 1L)

0.606g Tris base, pH 7.6
8.77g NaCl
0.5 mL Tween-20
to 1L water

37, to check for success of transfer, the membrane was stained with Ponceau Red diluted 1:10 in 1x TBST for 5 min on an agitator. (Stock: 2% Ponceau S in 30% trichoroacetic acid and 30% sulfosalicylic acid).
38, the membrane was washed in water until the protein bands are well-defined and scanned to save the image.
39, the membrane was destained completely by repeated washing in water.

IMMUNODETECTION

40, the membrane was blocked in 5% non-fat milk (Odyssey blocking buffer) for 1 hr at RT ^{o} with gentle shaking. (5% non-fat milk – 5g in 100 mL 1x TBST)
41, after incubation, rinsed in TBST for 5 sec.
42, Primary and secondary antibody dilutions were prepared in 1x TBST

Primary antibody dilutions

A, Rabbit anti-GPR30 polyclonal IgG - (1:1000 to 1:500)
B, Mouse anti-β-actin polyclonal IgG - (1:15000)

Secondary antibody dilutions

A, Goat ant-rabbit secondary IgG 800CW - (1:15000)
B, Goat ant-mouse secondary IgG 680RD - (1:15000)

*In case of two colour western blot, combine and dilute both the primary antibodies together in 1x TBST and incubate simultaneously with membrane. The primary antibodies must be from different host species. Similarly, combine and dilute the 2 different IRDye conjugated secondary antibodies in 1x TBST and incubate simultaneouly with membrane.
43, the blot was first incubated in diluted primary antibody solution overnight at 4 ^{o}C with gentle shaking.
44, after overnight incubation, the primary antibody solution was poured off and the membrane was rinsed 3 times with 1x TBST for 5 min, while agitating to remove residual primary antibody.
45, the blot was then incubated in diluted secondary antibody solution for 1 hr at RT ^{o}, while agitating.
46, after incubation, the secondary antibody solution was poured off and the blot was rinsed 3 times with 1x TBST for 5 min, while agitationg to remove residual secondary antibody.
47, followed by rinsing with 1x TBS to remove residual Tween-20.
*Protect membrane from light during washes after secondary antibody incubation. The blot can be stored in 1x TBS for 48 hr in dark at 4 ^{o}C.
48, the processed blot can be imaged in Odyssey infrared imaging system.