UNIVERSITÀ DEGLI STUDI DI PADOVA

University of Padua

Department of Pharmacological and Pharmaceutical Sciences
Doctoral School in Molecular Sciences
Pharmaceutical Sciences Curriculum
XXIX Cycle

PRODUCTION AND CHARACTERIZATION OF THERAPEUTIC PROTEINS/PEPTIDES:

Human Recombinant FSH-beta subunit expressed in Plant Cells and Chemical Synthesis of Human Osteocalcin and Neuritogenic Peptides.

This Work was financially supported by ABResearch

School Director: Prof. Antonino Polimeno

Supervisor: Prof. Vincenzo De Filippis

PhD Candidate: Simone Tescari

Abstract

Since 1980s proteins have emerged as a major new class of pharmaceuticals with over 350 marketed products that are mainly therapeutics with a small number of diagnostics and vaccines. In 2015 among the 10 best-selling drugs in the world, 7 are recombinant proteins or monoclonal antibodies. There is still considerable confusion about what therapeutic proteins are. Regulatory authorities (U S Food and Drug Administration and European Medicines Agency) provided different definitions describing them as *Biological medicinal products*, *Biopharmaceuticals*, *Biological* or *Biologic* depending on production system, biological source and pharmaceutical category. The European Medicines Agency (EMA) provided an actual definition describing *biological medicinal products* as "*a protein or nucleic acid–based pharmaceutical substance used for therapeutic or in vivo diagnostic purposes, which is produced by means other than direct extraction from a native (non-engineered) biological source*". This definition suggests that even small polypeptides and proteins obtained by chemical synthesis can be considered as *therapeutic proteins*.

Nowadays, most of approved therapeutic proteins are produced by DNA recombinant technologies. Recombinant proteins are mainly restricted to mammalian cells (CHO) as bioreactors which carry out post-translational modifications that significantly enhance the protein bioactivity. Recently, plants are emerging as an attractive alternative to conventional expression systems, due to its practical, economic and safety advantages, correct folding and similar glycosylation pattern like eukaryotes. In 2013 U S Food and Drug Administration approved the first plant-made pharmaceutical Elelyso® (Protalix and Pfizer) to treat Gaucher's Disease.

In this scenario, *Active Botanicals Research* (Brendola, VI, Italy) proposes suspension cell lines of *N. benthamiana* as alternative expression platform for recombinant protein and in particular for recombinant human follicle stimulating hormone beta subunit. Follicle Stimulating Hormone (FSH) is a gonadotropin that stimulates steroidogenesis and gametogenesis in the gonads in order to support and regulate the ovarian follicular maturation in women and sperm production in men. FSH is clinically used for controlled ovarian stimulation in women treated with assisted reproductive technologies, the treatment of anovulatory infertility in women and hypogonadotropic hypogonadism in men. FSH is most often administered in one of two forms: recombinant FSH expressed in CHO system (Gonal-F® (Merck Serono) and Puregon® (Merck Sharp and Dohme)) or highly purified human menopausal gonadotropin (Menopur® (Ferring) and Merional® (Pharmasure)). The first

chapter of this thesis focuses on extraction, purification and characterization of recombinant human follicle stimulating hormone beta subunit expressed in suspension cell lines of *N. benthamiana*. Extraction is based on KDEL-strategy, a retention signal which locks the protein in the endoplasmatic reticulum preventing the release into the cytoplasm. Purified recombinant human Follicle Stimulating Hormone beta subunit (rhFSHβ) was obtained with three consecutive chromatographic steps: IMAC chromatography, size-exclusion chromatography and ion-exchange chromatography.

Two purified isoforms of recombinant protein were chemically characterized by enzymatic deglycosylation with PNGase F and tryptic digestion, covering the total amino acid sequence while the assignment of disulphide bridges pattern (six internal disulphide bridges) confirmed the correct folding of protein.

It is known that FSH evokes the physiological response as a heterodimer. In all glycoproteins the common alpha subunit is non-covalently associated whit the beta subunit, which is structurally unique in its peptide sequence to each member of the family. However, some studies suggest a possible biological activity about the single beta subunit and different FSH beta subunit preparations are on market declaring its biological activity (FSHβ subunit ab191730 Abcam).

Purified rhFSHβ from ABResearch was tested on isolated Sertoli cells from pubertal porcine, unfortunately, preliminary data suggested no biological activity of the recombinant monomer. Waiting for recombinant human FSH alpha subunit from competent cell lines of *N. Benthamiana,* the biological activity of reconstituted heterodimer (rhFSHβ from ABResearch + Native Human Chorionic Gonadotropin) was evaluated and compared with commercial FSH Gonal-F®. Data showed equal biochemical responses suggesting the great potential of the new expression system for active recombinant therapeutic proteins.

The process has been improved to obtain 4.5/5 mg of purified rhFSHβ for kg of cells, exceeding abundantly average yields in plant expression system.

In the second chapter of this work, chemical synthesis of human osteocalcin (OC) 1-49 by solid-phase synthesis (SPPS) with Fmoc strategy is described. Human OC is a small protein of 49 amino acid residues mainly produced and secreted by osteoblasts, it represents one of the most abundant (10-20%) non-collagenous proteins in the bone tissue of vertebrates with a highly conserved primary structure. It has long been known that OC acquires high affinity for calcium ions through the post-translational modification of glutamate residues by carboxylation with a vitamin-K-dependent γ-glutamyl carboxylase (GGCX). Carboxylated residues, known as Gla residues, lead to a conformational change, stabilize the alpha-helical structure and confer a greater affinity for Ca^{2+} and hydroxyapatite. OC is involved in different

diagnostic fields due to its crucial roles in several physiological processes including remodeling of bone tissue, insulin production and regulation, testosterone secretion from testes and regulation of neurotransmitter levels in the brain. Recently osteocalcin has been proposed as potential therapeutic protein in regulation of androgen activity in a non-steroidal manner.

Nowadays purified human OC 1-49 is obtained by extraction from human bones or solid-phase synthesis with the Boc strategy. Due to limited supply of human osteocalcin from bone, the solid-phase synthesis represents the most accessible strategy.

In this study, SPPS Fmoc chemistry was proposed as alternative strategy due to the greater yields of synthesized peptides, minor side reactions during cleavage and high safety compared to Boc approach. Chemical synthesis was optimized getting final yield of 80/85% exceeding average yields of traditional chemical synthesis. The purified protein is subjected to the disulfide bond-forming reaction and subsequently it was chemically characterized by RP-HPLC, mass spectrometry analysis and enzymatic digestion. Conformational characterization and study of binding with Ca^{2+} were performed by spectroscopic measurements, hydrogen/deuterium exchange and ITC titration analysis comparing the new synthesized human osteocalcin with the commercial one (Bachem).

Peptides as therapeutics are the focus of the third chapter.

Peptides perform crucial roles in human physiology as growth factors, hormones, ion channel ligands, neurotransmitters and they are recognized for being highly selective and efficacious signaling molecules with attractive pharmacological profile. Peptides represent the new attractive perspective for therapeutics due to high safety, tolerability and efficacy properties. Lower production complexity and lower cost also increase interest in peptide drugs research and development compared with protein therapeutics.

Nowadays synthesized peptides are used in regenerative medicine due to their potential as guidance cues for neurite elongation and thus to activate intracellular pathways leading to cell differentiation.

Ten novel peptides derived from cell adhesion molecules and extracellular matrix proteins families (CHL1, Neurofascin, NrCAM, DCC, ROBO2 and 3, LINGO2, Contactin 1, 2 and 5) are designed to mimic guidance cues from the neural environment.

Peptides are synthesized with SPPS Fmoc strategy and characterized by RP-HPLC, mass spectrometry and circular dichroism analyses. According to previous experiments with L1-A and LINGO1-A, all synthetic peptides are tested on human neuroblastoma cell lines to evaluate their effect on neuronal differentiation and especially in neurite outgrowth and elongation.

Preliminary data suggest a prototype for the development of implants for long-term neuronal growth and differentiation.

CONTENTS

CHAPTER ONE

PRODUCTION AND CHARACTERIZATION OF HUMAN RECOMBINANT FSH-BETA SUBUNIT EXPRESSED IN PLANT CELLS

ABBREVIATIONS

N.benthamiana	Nicotiana benthamiana
hFSHβ	Human Follicle Stimulating Hormone beta subunit
rhFSHβ	Recombinant Human Follicle Stimulating Hormone beta subunit
FDA	Food and Drug Administration
EMA	European Medical Agency
WT	Wild Type
AA	Amino acids
MW	Molecular Weight
PTM	Post Translational Modification
GMP	Good Manufacturing Practice
SP	Signal Peptide
ER	Endoplasmic Reticulum
GA	Golgi Apparatus
IgG	Immunoglobulin G
CHO	Chinese Hamster Ovary
BHK	Baby Hamster Kidney
TFA	Trifluoroacetic Acid
TCA	Trichloroacetic Acid
Tris-HCl	2-amino-2(hydroxymethyl)-1,3-propanediol hydrochloride
TEMED	N,N,N',N'-tetramethylethylenediamine
DTT	Dithiothreitol
UV	Ultraviolet
PAGE	PolyAcrylamide Gel Electrophoresis
SDS	Sodium Dodecyl Sulphate
ES	Elettrospray
TOF	Time of Fly
IMAC	Immobilized Ion Metal Affinity Chromatography

SEC	Size Exclusion Chromatography
IEC	Ion-Exchange Chromatography
HPLC	High Pressure Liquid Chromatography
ECL	Enhanced Chemiluminescence
PVDF	Polyvinylidene Fluoride
A (Ala)	Alanine
C (Cys)	Cysteine
D (Asp)	Aspartic Acid
E (Glu)	Glutamic Acid
F (Phe)	Phenylalanine
G (Gly)	Glycine
H (His)	Histidine
I (Ile)	Isoleucine
K (Lys)	Lysine
L (Leu)	Leucine
M (Met)	Methionine
N (Asn)	Asparagine
P (Pro)	Proline
Hyp	Hydroxyproline
Q (Asn)	Glutamine
R (Arg)	Arginine
S (Ser)	Serine
T (Thr)	Threonine
V (Val)	Valine
W (Trp)	Tryptophan
Y (Tyr)	Tyrosine

1. INTRODUCTION

1.1 BIOPHARMACEUTICALS

The industry sector involved in "biopharmaceutical" development, manufacture and marketing is now over 30 years old, with over 350 marketed products. This includes over 125 recombinant proteins currently approved in the United States and European Union. The term is widely used but there is still considerable confusion about what biopharmaceutical is. Many authors (Rader 2008, Walsh 2002) have pointed out the lack of a definition by regulatory authorities. Food and Drug Administration (FDA) provides a very broad definition of "biological drugs", including recombinant proteins and human-derived products (blood and its derivatives): "*a virus, therapeutic serum, toxin, antitoxin, vaccine, blood, blood component or derivative, allergenic product protein (except any chemically synthesized polypeptide), or analogous product, or arsphenamine or derivative of arsphenamine (or anyothertrivalentorganicarsenic compound), applicable to the prevention, treatment, or cure of a disease or condition of human beings*". The European Medicines Agency (EMA) provides a more actual definition describing "biological medicinal products" as " *a protein or nucleic acid–based pharmaceutical substance used for therapeutic or in vivo diagnostic purposes, which is produced by means other than direct extraction from a native (non-engineered) biological source*" (Rader 2008). Notably, this definition suggests that even small polypeptides and proteins obtained by chemical synthesis can be considered as biological medicinal products. In 2001 Gary Walsh introduced the concept of "biotechnology medicine" as "*any pharmaceutical product used for therapeutic or in vivo diagnostic purposes, which is produced in full or in part by biotechnologically means*". Even though this definition isn't suitable, as small precursor molecules for the synthesis of antibiotics, or therapeutic proteins extracted from tissues or organs not genetically engineered (like the porcine pancreatic insulin) are not included.

The advent of recombinant DNA technologies has provided the tools for producing recombinant proteins that can be used as therapeutic agents. Different expression systems have been developed for the production of pharmaceutical products for the treatment of many different diseases as diabetes, anemia, hepatitis, rheumatic diseases and cancers

(Leader 2008). Since 1982 with the first genetically engineered drug approved by FDA (*Humulin,* human insulin) marketing, the number of biotech drugs on the market has increased (Mullard 2014). In 2015 among the 10 best-selling drugs in the world, 7 were considered biopharmaceuticals (**Tab. 1.1**). Biopharmaceuticals (large molecule drugs) are able to bind to specific cell receptors, associated with the disease process, performing highly specific and complex functions contrarily to the traditional chemical drugs (small molecule drugs). Their high specificity ensures minor collateral effects, allowing safer use of them. Furthermore their high activity allows the administration of smaller quantities of drug to the patient. Finally, the administration of the missing protein to treat disease caused by total or partial absence of the physiological protein activity (replacement therapy) represents the major advantage of biopharmaceuticals compared to the traditional drugs (ex. Insulin for diabetic patients).

DNA recombinant technologies allowed to overcome the limits of extracted and purified proteins from biological sources: limited availability of tissue, pathogens and immunogenic contaminants. Nowadays, the large amounts of recombinant proteins production, the high degrees of safety and purity have greatly encouraged the use of recombinant proteins in the medical field. Furthermore, the DNA recombinant technologies allowed to mutate the *wild type* protein sequence and express fusion proteins in order to improve the pharmacodynamic and/or pharmacokinetic characteristics of therapeutic proteins (Leader 2008).

Darbeoetin is a mutated erythropoietin with two additional N-glycosylation sites. The new iperglycosylated form has a greater plasma half-life improving the pharmacokinetics behavior. Insulin represents another example of protein engineering with available mutated forms in order to obtain different pharmacokinetic and pharmacodynamic responses. The Insulin Aspart (Pro28Asp of beta chain) showed a rapid adsorption, as Insulin Lispro (Pro28 and Lys29 are exchanged). The addition of two residues of arginine to the C-terminal of beta chain and the substitution of the residue of asparagine with glycine at the C-terminal of alpha chain slows the dissociation of hexamer reducing the rate of absorption.

	COMMERCIAL NAME	MOLECULE	PHARMACEUTICAL COMPANY	WORLDWIDE SALES IN 2014 ($ BILLION)	TECHNOLOGIES
1	Humira	adalimumab	AbbVie	12.543	Monoclonal Antibody
2	Sovaldi	sofosbuvir	Gilead Sciences	10.283	Synthesis Molecule
3	Remicade	infliximab	Johnson & Johnson/ Merck & Co.	9.240	Monoclonal Antibody
4	Rituxan	rituximab	Roche	8.678	Monoclonal Antibody
5	Enbrel	etanercept	Amgen/Pfizer	8.538	Recombinant Protein
6	Lantus	insulina glargina	Sanofi	7.279	Recombinant Protein
7	Avastin	bevacizumab	Roche	6.957	Monoclonal Antibody
8	Herceptin	trastuzumab	Roche	6.793	Monoclonal Antibody
9	Seretide	fluticasone + salmeterolo	GSK	6.431	Synthesis Molecule
10	Crestor	rosuvastatina	AstraZeneca/ Shionogi	5.869	Synthesis Molecule

Tab. 1.1. Top 10 pharmaceuticals best-selling.
Genetic Engineering & Biotechnology News

1.2 BIOLOGICAL EXPRESSION SYSTEMS

The manipulation of Escherichia coli provided the first expression system for therapeutic proteins, pioneering the production of recombinant proteins. Humulin (recombinant human insulin expressed in E. coli (Goeddel 1979)) is the first genetically engineered drug approved by the FDA for human use. As a prokaryote, E. coli cannot accurately express complex eukaryotic protein with essential post translational modifications (glycosylation, disulfide bond, carboxylation, hydroxylation, sulfation amidation, etc.). This problem has been overcome by the development of other expression systems, including the yeast, Baculovirus, the mammalian cells and the plant expression system.

Glycosylation pattern, typical for each species, can induce an immunological response which can cause hypersensitivity responses, anaphylaxis, anaphylactoid reactions, infusion reactions and decreased efficacy of the drug (Pineda 2016). Furthermore biopharmaceutical companies have to consider the economic advantages and disadvantages, production

timescale, scale up capacity, product quality, storage cost and the possibility of contaminant risks of each platform (**Tab. 1.2**).

SYSTEM	OVERALL COST	PRODUCTION TIMESCALE	SCALE-UP CAPACITY	PRODUCT QUALITY	GLYCOSYLATION	CONTAMINATION RISKS	STORAGE COST
Bacteria	Low	Short	High	Low	None	Endotoxins	Moderate
Yeast	Medium	Medium	High	Medium	Incorrect	Low risk	Moderate
Mammalian cell culture	High	Long	Very low	Very high	Correct	Viruses, prions and oncogenic DNA	Expensive
Transgenic animals	High	Very long	Low	Very high	Correct	Viruses, prions and oncogenic DNA	Expensive
Plant cell cultures	Medium	Medium	Medium	High	Minor differences	Low risk	Moderate
Transgenic plants	Very low	Long	Very high	High	Minor differences	Low risk	Inexpensive

Tab. 1.2. Comparison between major expression systems of recombinant proteins (Fahad 2015).

1.2.1 *Escherichia coli* system.

E. coli was the first expression system for recombinant proteins production. Its use as cell factory is well established and it has become the most popular expression platform. It is well documented for its easy transformation advantages (plasmid transformation: 5 min), rapid replicative cycle (doubling time: 20 min), high scalability, low cost and high protein yields (0.5-5g/L) (Pope and Kent 1996). *E. coli* is a suitable host for expressing stably folded, globular proteins from prokaryotes and eukaryotes. However, membrane proteins and proteins with molecular weights above 60 kDa are difficult to express. Furthermore solubility could represent a great obstacle since some proteins are insoluble and aggregate in the inclusion bodies. Notably, there are different factors that heavily affect expression levels and/or solubility: growing temperature, concentration of inducer (IPTG), host strain, protein size and structure, and toxicity (Rosano and Ceccarelli 2014). The inability of bacteria systems to produce active form of complex proteins with appropriate post-translational modifications represents the biggest obstacle of this expression platform.

Betaseron (Interferon beta-1b) is a non-glycosylated recombinant human IFNβ from *E. coli*. It is a prescription medicine used to reduce the number or relapses in people with relapsing forms of multiple sclerosis. Different cytokines, interleukins-2 (Proleukin), interferon

2α (Pegasys), granulocyte colony-stimulating factor (Neupogen), growth factor (Omnitrope) are produced in E. coli.

1.2.2 Yeast system.

Yeast protein expression is a common alternative to prokaryotic and higher eukaryotic expression. Yeast cells offer many of the advantages of producing proteins in microbes (easy transformation, growth speed, high scale up capacity and low cost media) as well as eukaryotic features including the secretory pathway leading to correct protein processing and post-translational modifications (Mattanovich 2012). Several yeast expression systems are used for recombinant proteins expression, including *Sacharomyces cerevisiae, Scizosacchromyces pombe, Pichia pastoris and Hansanuela polymorpha* (Sudbery 1996). *Saccharomyces cerevisiae* is able to secret heterologous proteins in the medium by adding a signal peptide (SP) improving and facilitating the purification step, *Pichia pastoris* allows the disulphide bond assignment increasing the final yield. One of the major advantages is that yeast cultures can be grown to very high densities (up to 130 g of cells/L of medium culture) in order to obtain high protein yields (0.1-2 g/L). The different post-translational modifications between yeast and human can reduce the biological activity and/or induce immunological response limiting the range of use. In yeast the *O-glycosylation* is formed only by mannose residues while *N-glycosylation* consists of a core of two N-acetylglucosamine linked to a high number of mannose residues.

Nowadays molecular biologists are overcoming the problem by engineering yeast to produce glycoproteins with human-like glycans, improve the performance of specific affinity tags, reduce proteolysis and insert non-native amino acids (e.g. selenomethionine) into proteins.

1.2.3 Insect.

Baculovirus expression system is widely used to produce large quantities of recombinant proteins in insect host cells. These systems are particularly well suited for proteins that are difficult to express in bacteria due to size, complexity, or post-translational processing requirements. The post translational processing and folding of recombinant proteins in insect cells are more similar to mammalian than bacterial cells (Miller 1988), so these proteins are more likely to have biological activities and immunological reactivity comparable to proteins

that naturally exist in mammalian cells. The baculovirus insect cell expression system is based on the use of recombinant baculoviruses (insect viruses) and their ability to produce high levels of biologically active proteins in cultured insect cells (30 % of total cellular proteins (Demain 2009)) or insect larvae. These viruses have very late genes (polyhedrin, P10), which are highly expressed due to very strong promoters, but are dispensable for infectious virus formation. Deletion of these genes and replacement with foreign genes convert these viruses into expression vectors for the production of recombinant proteins. The proteins produced can be easily purified from infected cells or their supernatants using tags and affinity chromatography. The system is considered safe as baculovirus infect only insects. The absence of a signal peptide that allows the secretion of recombinant proteins, the formation of aggregates and the misfolding of high molecular weight proteins represents the major disadvantages of the expression system (Loos 2012). Furthermore, potentially immunogenic glycoside residues (α1-3fucose (Shi 2007)) are added in the last step of *N-glycosylation*, while *O-glycosylation* mechanism is still little known.

1.2.4 Mammalian cells.

Expression systems based on mammalian cells are able to introduce proper protein folding, post translational modifications (glycosylation, disulfide bond, carboxylation, hydroxylation, sulfation amidation, etc) and possible assembly, which are essential for the biological activity. In the previous decade, protein therapeutics produced from mammalian cells revolutionary changed the landscape of human healthcare. Bioprocesses based on mammalian cell have been applied in the manufacture of viral vaccines, diagnostic and therapeutic proteins. Chinese hamster ovary (CHO) cells, mouse myeloma cells, including NS0 and Sp2/0 cells are the most commonly used expression system for recombinant proteins production (Griffin 2007). Two derivatives of the CHO cell line, CHO-K1 and CHO pro-3, gave rise to the two most commonly used cell lines in bio-processing today, DUKX-X11 and DG44. These two cell lines were engineered to be deficient in dihydrofolate reductase (DHFR) activity (Lee 2010, Wurm 2011). Vectors are autonomously replicating DNA molecules that can be used to carry foreign DNA fragments. For expressing heterologous genes in mammalian cells, usually vectors derived from mammalian viruses are used. These include viruses such as Simian Viruses 40 (SV40), polyomavirus, herpesvirus and papovirus. In order to design the vector, it is necessary to select an efficient promoter and also the selection marker. In last two decades, mammalian cell protein expression has become the dominant recombinant protein production

system for clinical applications, producing more than half of the biopharmaceutical products in the market and several hundreds of candidates in clinical development. The difficult to achieve transfection of gene of interest into these mammalian cells, low production rate, the contaminant risks and expensive time consuming for single clone selection represent major disadvantages of this expression system (Fahad 2015). Suitable expression system is selected based on the productivity, bioactivity, purpose, physicochemical characteristics of the interest protein, cost, convenience and safety of the system itself.

1.2.5 Plants.

In earlier days, the production of recombinant therapeutic proteins are mainly restricted to mammalian cells which are able to carry out post translational modifications that significantly enhance the protein bioactivity (Rasooly 2013).To date, most of the recombinant proteins approved by FDA are produced in mammalian cells (Lai 2013). Now plants are emerging as a potential competitor to conventional expression systems, due to its practical, economic and safety advantages. The benefits of using plants for therapeutic proteins production include: lower production costs of large scale production and convenient system when compared to transgenic animals, fermentation or bioreactors; facilities already exists for the planting, harvesting and processing of plant material; no contamination of therapeutic protein with any human pathogens when produced in the plants; protein produced with correct folding, more or less similar glycosylation pattern like eukaryotes (Twyman 2003, Horn 2004). Several plant-derived have reached and are reaching the late stages of commercial development. These products include antibodies, vaccines, human blood products, hormones and growth regulators (human growth hormone, human coagulation factor IX, NY-ESO-1 and cyanovirin (CV-N)) (Shanmugaraj 2014). Basically, there are three strategies for recombinant protein production in plant-based systems:(1) transient expression of foreign genes in plant tissues that are transformed by either agroinjection or by viral infection and (2) development of transgenic plants carrying stably integrated transgenes; (3) use of cell-culture-based systems that are equivalent to mammalian, microbial and insect cell systems (Rech 2008, Rech 2012). Nowadays, plants are considered more than expression platform; there are advanced studies about plant as oral delivery for chloroplast-derived proinsulin (Boyhan 2011) and vaccines (Takaiwa 2013). Plant cell wall prevents gastric degradation of therapeutic proteins, promoting the release and adsorption in the intestine. Lyophilized cells and whole seeds might represent

a new drug delivery system for biopharmaceuticals. Taliglucerase alfa, commercially known as Elelyso, is a biopharmaceutical drug developed by Protalix and Pfizer. The drug, a recombinant glucocerebrosidase used to treat Gaucher's Disease, was the first plant-made pharmaceutical to win approval by the U.S. Food and Drug Administration. The pharmaceutical company is developing a system to administer the drug orally by lyophilized carrot cells (Shaaltiel 2015).

1.2.5.1 Plants cells.

The therapeutic proteins production in plants can be achieved by developing stable transgenic lines or transient expression systems. Transiently transfected cells express the foreign gene but do not integrate it into their genome. Thus the new gene will not be replicated. These cells express the transiently transfected gene for a finite period of time after which the foreign gene is lost through cell division. Stable transfection begins with a transient transfection but in a small proportion of transfected cells, the foreign gene is integrated into the cells' genome. The hallmark of stably transfected cells is that the foreign gene becomes part of the genome and is therefore replicated. Descendants of these transfected cells, therefore, will also express the new gene, resulting in a stably transfected cell line (Shaaltiel 2015). Development of stable transgenic plants for the recombinant protein production provides the transgenic seed bank but with drawbacks as, it needs sterile environment and also its time consuming. Transient expression systems are very useful for research and are now being routinely used for the rapid production of valuable proteins. These systems allow high throughput production and straightforward manipulation, permitting the rapid validation of expression constructs and the production of large amounts of recombinant protein within a few weeks. As a direct consequence, the protein yields from transient expression in plants are normally higher than yields observed in other recombinant plant systems (the yield is 4-20 times higher than stable transfected cells (Xu 2012)). Transient expression is achieved either by epichromosomal expression of *Agrobacterium tumefaciens* on the infiltrated leaves (Agroinfiltration) or by viral expression vectors (Lombardi 2009). In viral expression vectors, the desired gene is cloned between the genetic materials of the plant virus, most preferably RNA virus and allowed to infect the plants to produce the target protein. Since the viral vectors do not get incorporated into the plant genome, the expression is transient that will yield more protein in less time (Scholthof 1996). Tobacco leaves are the dominant choice for the development of commercial platforms using transient expression (Ma 2003). Tobacco is being

used as a model plant for several in vitro studies. It has been proved to be an extremely versatile system for all the aspects of cell and tissue culture research. Although tobacco and Arabidopsis are the two facile plants that are easily to transform but it takes several months to produce to a stable transgenic line (Leckie 2011). This problem can be overcome by the transient expression of proteins. Seeds as bioreactors also provide a potential economical platform for the large-scale production and storage of recombinant proteins (Fisher 2000). In fact, the endosperm is considered an ideal environment for the protein storage because of low proteolytic activity (Xu 2012), small amounts of water and the presence of specific compartments dedicated to the preservation of proteins (Sabalza 2014). In that conditions the recombinant proteins can be stored without loss of activity up to three years in dried seeds (Twyman 2005). Recombinant protein can be produced in plants at the cost of 2-10% and 0.1% of the cost of microbial fermentation and cell culture systems respectively. Not all the proteins are expressed at high levels in plants, but yields of 0.1-1% Total Soluble Proteins (TSP) are sufficient to make plants economically viable. Above 85% of the total cost for the recombinant protein production is spent for downstream purification rather than production, regardless of the expression system (Shanmugaraj 2014).

1.2.5.2. Plant cell culture

Plant cell-suspension cultures also produce recombinant proteins with intermediate benefits between plant systems and mammalian cell lines. As plant cell media lack any mammalian components that are susceptible to the transmission of mammalian viruses or prions, plant cell systems naturally do not carry the risk of infection by, or transmission of, human or other animal pathogens. Moreover, plant cell cultures present a natural barrier to mammalian pathogen contamination, as attempts to propagate mammalian viruses in plant cells have been unsuccessful (EMA, 2013). These factors are not only important with respect to safety, but also significantly reduce the operational costs, compared to mammalian cell-based expression systems. The major resistance to mechanical stress of plant cells, due to cell wall rigidity, makes these eukaryotic cell the best choice to industrial scale up. The advantages of using plant cell-suspension cultures for therapeutic protein production includes (Shanmugaraj 2014): rapid and controlled growth, moderate costs (minor than mammalian cells but major than transgenic plants), rapidly production than transgenic plants because the development and testing schedule is much shorter, easily adoption of *Good Manufacturing Practice* and high reproducibility of process (**Tab. 1.3**). Suspension cells are particularly advantageous when

defined and sterile production conditions in addition to straightforward purification protocols are required, as these conditions are particularly applicable to the production of therapeutics. Furthermore, recombinant proteins expressed in plant-cell-suspension cultures can be secreted into the culture supernatant greatly reducing purification costs (Twyman 2003). Localization depends on protein targeting and the permeability of the plant cell wall to the protein in question. Plant cell-suspension cultures are able to overcome limitations related to ethical issues: the non-use of herbicides and pesticides and the least necessary amount of water have a lower environmental impact (Directive (EU) 2015/41224). Furthermore, plant cell cultures are not subject to the restrictive laws relating to the cultivation of *OGM* (forbidden in many countries including Italy).

PLANT		**PLANT CELL-SUSPENSION CULTURES**	
Advantages	**Disadvantages**	**Advantages**	**Disadvantages**
Costs	Contamination Risks	Safety	Costs
Scale-up	Production Time Scale	Production Time Scale	Yelds
Yields	GMP	GMP	
	Weather conditions	Controlled Growth	
	OGM	Reproducibility	

Tab. 1.3 Comparison of the major advantages and disadvantages between total-plant and plant cell-suspension cultures as expression systems.

1.2.5.3 Biopharmaceuticals expressed in plants.

The *Growth Factor* was the first recombinant human protein expressed in plant in 1986 (Obembe 2011), afterwards different antibodies (Hiatt 1989, During 1990) and human serum derivatives (1990) were produced in the plant system. Since then, the number of recombinant human proteins expressed in plant has increased: Epitelial Growth Factor (Bai 2007), interferon α, β e γ (Arlen 2007), erythropoietin (Weise 2007), interleukin 10 (Fujiwara 2010), etc. (**Tab. 1.4**). Human lactoferrin, lysozyme (*Ventria Bioscience*) and canine gastric lipase (*Meristem Therapeutics*) (Obembe 2011) are on market as nutraceutical products or reagents/standards for diagnostic kits. The first recombinant protein accepted by a regulatory

authority was a vaccine for poultry developed by *Dow Agro Sciences* containing the recombinant forms of the glycoproteins hemagglutinin and neuraminidase expressed in tobacco cell suspension cultures for Newcastle disease virus. In 2012 the first plant expressed protein for human application was approved by FDA to treat Gaucher's Disease: Taliglucerase alfa (*Elelyso*) is a biopharmaceutical drug developed by Protalix and Pfizer in carrot cells (Dandana 2015). The approval by the FDA of this first recombinant drug obtained by plant cells in suspension is an important precedent for obtaining the marketing authorization by future recombinant therapeutic proteins for human use.

PROTEIN	HOST PLANT SYSTEM	COMMENTS
Human biopharmaceuticals		
Growth hormone	Tobacco, sunflower	First human protein expressed in plants; initially expressed as fusion protein with *nos* gene in transgenic tobacco; later the first human protein expressed inchloroplasts, with expression levels ~7% of total leaf protein
Human serum albumin	Tobacco, potato	First full size native human protein expressed in plants; low expression levels in transgenics (0.1% of total soluble protein) but high levels (11% of total leaf protein) in trasformed chloroplasts
α-interferon	Rice, turnip	First human pharmaceutical protein produced in rice
Erythropoietin	Tobacco	First human protein produced in tobacco suspesion cells
Human-secreted alkaline phosphatase	Tobacco	Produced by secretion from roots and leaves
Aprotinin	Maize	production of a human pharmaceutical protein in maize
Collagen	Tobacco	First production on human structural-protein polymer; correct modification achieved by co-trasformation with modification enzyme
α1-antitrypsin	Rice	First use of rice suspension cells for molecular farming
Recombinant antibodies		
IgG1 (phosphonate ester)	Tobacco	First antibody expressed in plants; full lenght serum IgG produced by crossing plants that exoressed heavy and light chains
IgM (neuropeptide hapten)	Tobacco	First IgM expressed in plants and protein targeted to chloroplast for accumulation
SIA/G (*Streptococcus mutans* adhesin)	Tobacco	First secretory antibody expressed in plants; achieved by sequential crossing of four lines carrying individual components; at present the most advanced plant-derived pharmaceutical protein
scFv-bryodin 1 immunotoxin (CD 40)	Tobacco	First pharmaceutical scFv produced in plants; first antibody produced in cell-suspension culture
IgG (HSV)	Soybean	First pharmaceutical protein produced in soybean
LSC (HSV)	*Chlamidomonas reinhardtii*	First example of molecular farming in algae
Recombinant subunit vaccines		
Hepatitis B virus envelope protein	Tobacco	First vaccine candidate expressed in plants; third plant-derived vaccine to reach clinical trial stage
Rabies virus glycoprotein	Tomato	First example of an 'edible vaccine' expressed in edible plant tissue
Escherichia coli heat-labile enterotoxin	Tobacco, potato	First plant vaccine to reach clinical trials stage
Norwalk virus capsid protein	Potato	Second plant vaccine to reach clinical trials stage
Diabetes autogen	Tobacco, potato	First plant-derived vaccine for an autoimmune disease
Cholera toxin B subunit	Tobacco, potato	First vaccine candidate expressed in chloroplasts
Cholera toxin B and A2 subunits, rotavirus enterotox	Potato	First plant-derived multivalent recombinant antigen designed for protection against several enteric diseases
Porcine trasmissible gastroenteritis virus glycoprotein S	Tobacco, maize	First example of oral feeding induction protection in an animal

HSV, herpes simplex virus; IgG, immunoglobulin G; IgM, immunoglobulin M; LSC, long single chain; *nos*, nopaline synthase; scFv, single-chain FV fragment; SIgA, secretory immunoglobin A.

Tab 1.4 Major recombinant proteins expressed in plant system.

The important viral diseases that cause significant deaths or pandemics in human populations are influenza, measles, hepatitis B, hepatitis C, hepatitis E, human immunodeficiency virus -acquired immunodeficiency syndrome (HIV-AIDS), human papilloma virus (HPV) infection, and rabies, whereas considerable economic losses in animals are attributable to avian influenza, Norwalk virus, and foot and mouth disease. Cholera, tuberculosis, and diphtheria are among the bacterial diseases that cause considerable loss of life. Plants are used as expression systems to produce vaccines as prophylactic for these diseases too. Some biopharmaceuticals recently produced in plants against important viral, bacterial, and protozoan diseases and cancers are listed in **table 1.5** (Fahad 2014). Key events in the development of plant-derived pharmaceuticals were reported in the figure below (**Fig 1.1**).

Disease	Pathogen	Biopharmaceutics	Promoter/vector	Expression system
Protozoan infection				
Malaria	Plasmodium	$pyMSPI_{19}$	Deconstructed TMV vector	magnICON Tobacco
		CTB-MSP1 AMA-1	psba/rm Promoters	Transplastomic transformation lettuce/chloroplast
Bcterial diseases				
Cholera	Vibrio Cholarae	CTB	TMV vector	magnICON tobacco
		LTB	CaMV -35S Codon optimized carrot	Somatic embryogenesis
		LBT-ST	prm promoter	Transplastomic transformation tobacco
		Seed-specific LTB	Soybean glycinin promoter	Stable transformation somatic embryogenesis soybean
		CBT-MSP1 AMA-1	Psba/rrn promoters	Transplastomic transformation lettuce/chloroplast
Tubercolosis	Mycobacterium tuberclosis	Immune-dominant antigens	Patatin promoter	Stable transformation Tobacco
		CTB-ESTA6	psbA promoter	Transplastomic transformation tobacco/lettuce
		TB vaccine protein	CaMV -35S promoter	Agroinfiltertion tovacco
Diphtheria Pertussis and tetani (DPT)	Diphteria Pertussis and tetani (DPT)	sDPT polypeptide	CaMV -35S	Stable transformation tomato
Viral diseases				
SARS	Corona virus (human)	SARS-Cov	OCS_3MAS	Stable nuclear transformations cauliflower
Small pox	Variola virus (human)	Viral coat B5	CaMV -35S	Stable nuclear transferation, collard
		Candidate pB5	CaMV -35S	Magnifection, tobacco
Diarrhea	Rota virus gastroenteritis (Human)	RV VLPs	CaM -35S	Stable nuclear transformation, tobacco
Post weaning diarrea (PWD)	Procrine epidemic diarrhea virus (PEDV) (Pigs)	sLTB-sCOE	HMW-GS (Bx17)	Epitope presentation rice endosperms
		Functional recombinant FaeG	psbA promoter	Biolistic chloroplast transformation
Measles	Measles virus (human)	MV-H protein	CaMV -35S	Stable nuclear transformation, tobacco
Rabies	Rabies virus (human)	Rabies nucleoprotein	CaMV -35S	Nuclear transformation and agrofilteration,

				tomato
Influenza	H1N1	HA1-protein	Binary vector AscI-PacI	stable transformation tobacco
	H1N5	HA1-5 (VLPs)	Alfalfa plastocyanin promoter	Agroinfilteration Tobacco
	H7N7	HA7-7	CaMV -35S	Agroinfelteration Tobacco
	H1N5, H1N5	HA1-5/1	Launch vector	magnICON Tobacco
Hepatitis	HBV	S-HBsAg	CaMV -35S	Stable transformation lettuce
	HCV	Chimerie CMVs	Cucumber mosaic Virus	magnICON Tobacco
	HEV	pE2	Rice-psbA promoter E2	Biolistic chloroplasted transformation tobacco
AIDS	Human mmune deficiency virus HIV	C4V3 polypeptide	prm promoter	Biolistic chloroplasted transformation tobacco
		HIVmAbs	CaMV -35S	Agroinfilteration Tobacco
Cancer				
Cancer	Human papilloma virus (HPV)	HPV16-1.2 epitope	PVX	magnICON tobacco
	Non-Hodgkin's lymphoma	HPV16-L1mAbs	CaMV -35S	Stable transformation tobacco
		HPV11-L1-NLS proteins	CaMV -35S	Stable transformation arabidopsin/tobacco
Dengue	Dengue virus (DENV)	Dengue virus tetraepitope peptide (cE-DI/IIp)	Tobacco psbA promoter/pRL1001	lettuce plastid transformation

Tab 1.5 Biopharmaceutical compounds developed against infectious diseases and cancers using plants as biofactories.

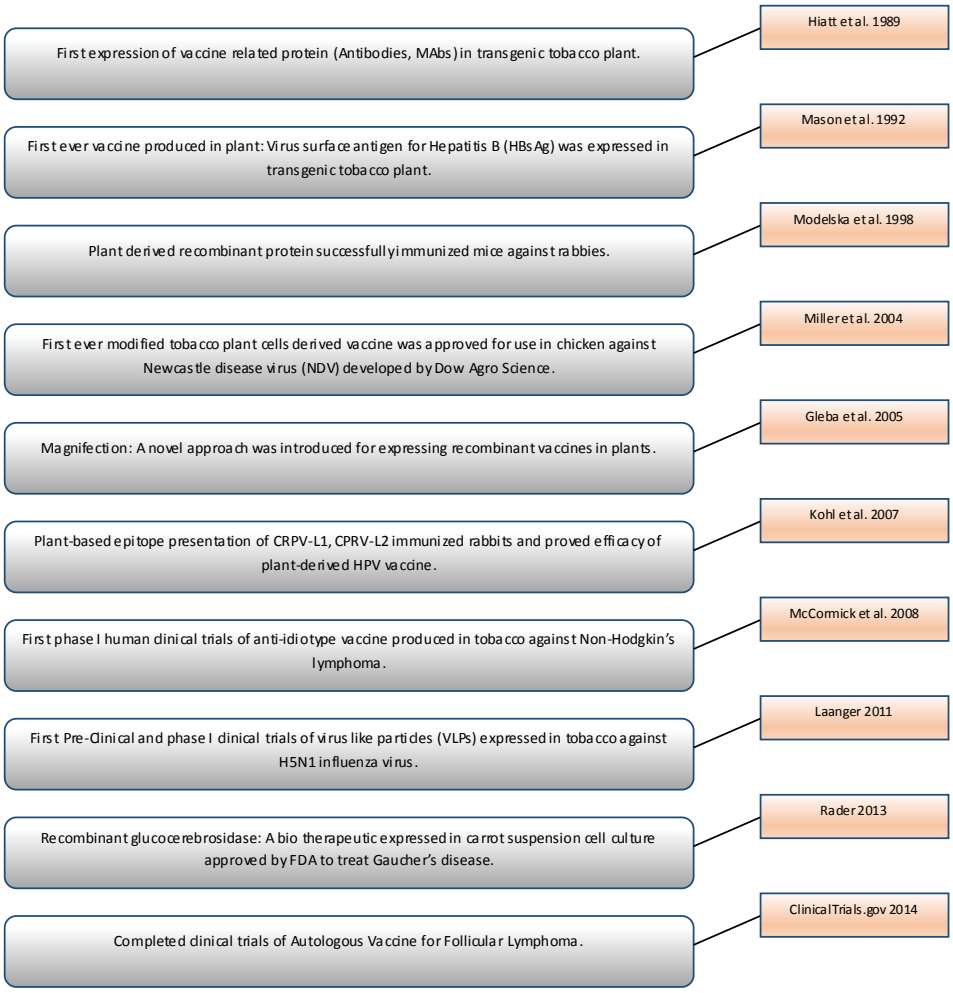


Fig 1.1 Key events in the development of plant-derived pharmaceuticals.

1.2.5.4 Post translational modifications in plants.

Post translational modifications (PTMs) are essential for the stability and biological action of recombinant therapeutic proteins produced in any host system. PTMs make the plant expression system more unique compared to prokaryotic expression systems (Shanmugaraj 2014). PTMs regard covalent changes of individual amino acid residues including: glycosylation, phosphorylation, methylation, disulfide bond formation, proteolytic processes of protein backbone and non-enzymatic modifications such as deamidation and the racemization (**Tab. 1.6 and 1.7**). Many plant proteins are synthesized as precursor proteins and have cleavable N- or C-terminal signal peptides for sorting or secretion to the relevant subcellular compartments for the subsequent maturation step and/or storage. The cleavages of these signal peptides and downstream signals are carried out by dedicated and specialized processing peptidases. In addition to the removal of sorting information, a subset of proteins undergo further N- or C-terminal trimming events by specific peptidase activity to fully activate and/or stabilize the proteins or release the protein from a membrane for full activity. In some cases, the newly formed termini then undergo specific PTMs, such as N-terminal acetylation, phosphorylation, or acylation (Friso 2015). In the endoplasmatic reticulum proteins undergo folding and/or assembling process. After that proteins usually reached Golgi apparatus to complete the maturation step (Matsuoka 1999). All mammalian PTMs occur in plants with some exceptions: hydroxylation of proline (Hyp) occurs much more frequently in plants and it constitutes a possible O-glycosylation site, gamma carboxymethylation is a post translational modification typical of mammals, the glycosilation process also occurs in plants although with alterations that will be explained later. The capability of plant system to couple properly the cysteine residues by disulfide bridges represents a great benefit improving the yields and lowering costs of production.

Amino Acid Residue	*Observed Physiological PTM in Plants*	*PTMs Caused by Sample Preparation*
Ala (A)	Not known	
Arg (R)	Methylation, carbonylation	
Asn (N)	Deamination, N-linked glycosylation	Deamination
Asp (D)	Phosphorilation (in two-component system)	
Cys (C)	Glutathionylation (SSG), disulfide bonded (S-S), sulfenylation(-SOH), sulfonylation (-SO_3H), acylation, lipidation, acetylation, nitrosylation (SNO), methylation, palmitoylation, phosphorylation (rare)	Propionamide
Glu (E)	Carboxylation, methylation	Pyro-Glu
Gln (Q)	Deamination	Deamination, pyro-Glu
Gly (G)	N-Myristoylation (N-terminal Gly residue)	
His (H)	Phosphorylation (infrequent)	Oxidation
Ile (I)	Not known	
Leu (L)	Not known	
Lys (K)	N-ε-Acetylation, methylation, hydroxylation, ubiquitination, sumoylation, deamination, O-glycosylation, carbamylation, carbonylation, formylation	
Met (M)	(De)formylation, excision (NME), (reversible) oxydarion, sulfonation (-SO_2), sulfoxation (-SO)	Oxidation, 2-oxidation, formylation, carbamylation
Phe (F)	Not known	
Pro (P)	Carbonylation	Oxidation
Ser (S)	Phosphorylation, O- linked glycosylation, O- linked GlcNAc (O-GlcNAc)	Formylation
Thr (T)	Phosphorylation, O- linked glycosylation, O- linked GlcNAc (O-GlcNAc), carbonylation	Formylation
Trp (W)	Glycosylation (C-mannosylation)	Oxidation
Tyr (Y)	Phosphorilation, nitration	
Val (V)	Now known	
Free NH_2 of protein N termini	Preprotein processing, Met excision, formylation, pyro-Glu, N-myristoylation, N-acylation (i.e. palmitoylation), N-terminal α-amine acetylation, ubiquitination	Formylation (Met), pyro-Glu (Gln)

Tab. 1.6 Reactivity of amino acids and their PTMs observed in plants (Friso 2015)

Type of PTM (Reversible, Exept if Marked with an Asterisk)	*Spontaneous (S; Nonenzymatic) or Enzymatic (E)*	*Comment on Subcellular Location and Frequency*
Phosphorylaition (Ser, Thr, Tyr, His, Asp)	E	His and Asp phosphorilation have low frequency
S-Nitrosylation (Cys) and nitraion* (Tyr)	S (RNS), but reversal is enzymatic for Cys by thioredoxins	Throughout the cell
Acetylation (N-terminal α-amine, Lys ε-amine)	E	In mitochondria, very little N-terminal acetylation, but high Lys acetylation; Lys acetylation correlates to acetyl-CoA]
Deamination (Gln, Asn)	S, but reversal of isoAsp is enzymatic by isoAsp methyltransferase	Throughout the cell
Lipidation (S-acetylation, N-meristoylation*, prenylation*; Cys, Gly, Lys, Trp, N terminal)	E	Not (or rarely) within plastids, mitochondria, peroxisomes
N-linked glycosylation (Asp); O linked (Lys, Ser, Thr, Trp)	E	Only proteins though the secretory system; O linked in the cell wall
Ubiquitination (Lys, N terminal)	E	Not within plastids, mitochondria, peroxisomes
Sumoylation (Lys)	E	Not within plastids, mitochondria, peroxisomes
Carbonylation* (Pro, Lys, Arg, Thr)	S (ROS)	High levels in mitochondria and chloroplast
Methylation (Arg, Lys, N terminal)	E	Histones (nucleus) and chloroplasts; still underexlored
Glutathionylation (Cys)	E	High levels in chloroplasts
Oxiation (Met, Cys)	S (ROS) and E (by PCOs), but reversal is enzymatic by Met sulfoxide reductase, glutaredoxins, and thioredoxins, except id double oxidized	High levels in mitochondria and chloroplasts
Peptidase* (cleavage peptidyl bond)	E	Throughout the cell
S-Guanylation (Cys)	S (RNS)	Rare; 8-nitro-cGMP is signalling molecule in guard cells
Formilation (Met)	S, but deformylation is enzymatic by peptide deformylase	All chloroplasts and mitochondria-encoded proteins are synthesized with initiaing formylated Met

Tab. 1.7 Most significant and/or frequent PTMs observed in plants (Friso 2015)

1.2.5.5 Glycosylation in plants.

Glycosylation occurs when a carbohydrate is attached to a protein via a covalent bond. Most eukaryotic proteins that are secreted into the endoplasmatic reticulum become N-glycosylated via an Asn that is part of Asn-X-Ser/Thr and/or Asn-X-Cys motifs (X: every amino acid except for Pro) (Gomord 2010). The N-linked glycan subsequently undergoes different degrees of processing and modifications by enzymes that are spatially distributed over the endoplasmatic reticulum, Golgi, and post-Golgi apparatus. In particular, the initial glycosylation steps play an important role in protein folding and quality control, including recognition by the degradation pathway. Much less understood is O-glycosylation where glycans are linked to a Ser, Thr or Hyp residue (Friso 2015). The majority of approved therapeutic proteins are glycoproteins (Gomord 2010).The oligosaccharide chains linked to the polypeptide backbone seriously affect the physical and chemical properties of a protein, such as solubility, resistance to proteolytic degradation and denaturation. The results affect the half-life time and pharmacokinetic characteristics. The glycosidic pattern can also affect the immunogenicity, biological activity and possible ligand-receptor interaction. In plant cells, as in other eukaryotic systems, the N-glycosylation starts in the ER (Bosch, 2013) by moving the precursor oligosaccharide (Glc3Man9GlcNAc2) from a lipid careers (dolichol diphosphate, DPP) to an asparagine residue. Subsequently, the oligosaccharide chain loses 3 terminal residues of glucose before moving to *cis* region of Golgi apparatus. In higher eukaryotes processing in *cis* and *medial* Golgi compartments leads to the formation of the so-called complex N-glycans. Notably, the N-glycan processing steps are virtually identical in plants and mammals up to the formation of the vital intermediate GlcNAc2Man3GlcNAc2 (GnGn). In mammals, GnGn oligosaccharides provide the substrate for extensive elongation/modification processes to give rise to the final diversification of N-glycosylation. In plants, modifications of these oligosaccharides are more limited and the GnGn structures are normally decorated with β1,2-xylose and core α1,3-fucose residues (GnGnXF3) (**Fig 1.2**)). Although fucosylation is observed in mammals as well, the fucose residues are α1,3-linked in plants as opposed to α1,6-linkage in mammals. In some cases plant cells are able to further elongate the GnGnXF3 by attaching β 1,3-galactose and α1,4-fucose residues to form Lewis-a epitopes(Bosch 2013).In summary, although there are potentially immunogenic differences (Bardor 2003) in the final structure of N-glycans, mammals and plants share a remarkably high degree of homology during processing along the secretory pathway. Mammalian cells provide the advantage of producing

recombinant proteins with N-glycans very closely resembling those produced in the human body, however, differences exist. CHO cells for example lack the ability to sialylate proteins in the human-typical α 2,6-position and add sialic acid in α 2,3- linkage affecting biological activity of recombinant human proteins. In summary, mammalian cell lines, despite being the most used production platform, suffer from several drawbacks including the difficulty to produce single glycoforms, low batch-to-batch glycosylation reproducibility and attachment of non-human glycol-epitopes and absence of some human-type N-glycans. To overcome these drawbacks mammalian and plant expression hosts are being engineered to allow production with non- immunogenic epitopes and human-type glycosylations. A commonly approach to express recombinant proteins in plants is their retention in the endoplasmic reticulum by addition of C-terminal retention signal KDEL. Proteins containing the KDEL sequence at one end of its sequence do not undergo the processes of N-glycosylation and O-glycosylation which occur in the Golgi apparatus. The resulting glycoproteins will be equipped with a glycosidic profile identical to human protein (Bosch 2013).

O-glycosilation occurs inside Golgi apparatus in mammals and in plants and involves serine, theronine and hydroxylated proline residues. The residues of hydroxyproline are much more frequent in proteins expressed in plant, so the Hyp-glycosylated form is more abundant in plant system (Gomord 2010). There are two types of Hyp-glycosylation: the first one consists in chains of arabinose, the second one in a various mix of arabinose and galactose residues. In mammals the O-glycosylation is defined as "mucin-type": N-acetylgalactosamine, N-acetylglucosamine, galactose, sialic acid and often fucose, variously linked (**Fig. 1.3**). The O-glycosylation of plant is immunogenic in animal models (Gomord 2010), for this reason, techniques have been developed to lock it.

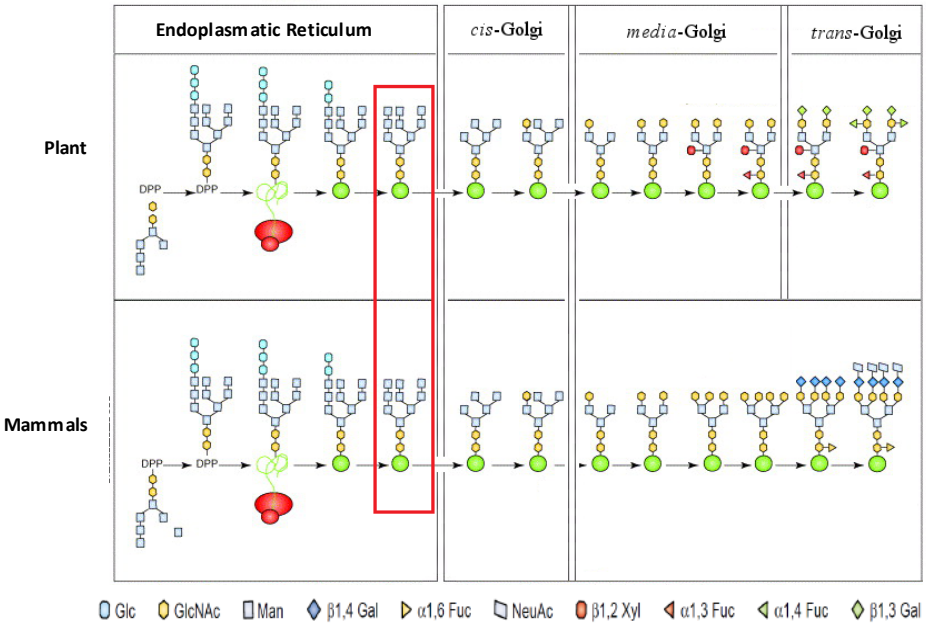


Fig 1.2 Protein N-Glycosylation pathway in plants and mammals. Sequential distribution of N-glycan processing enzymes across the Golgi apparatus, separated according to their action into early- and medial-acting. The red square underlines the common glycosylation form in plant and mammals. The KDEL technique lock the maturation step preventing the transport to Golgi apparatus.

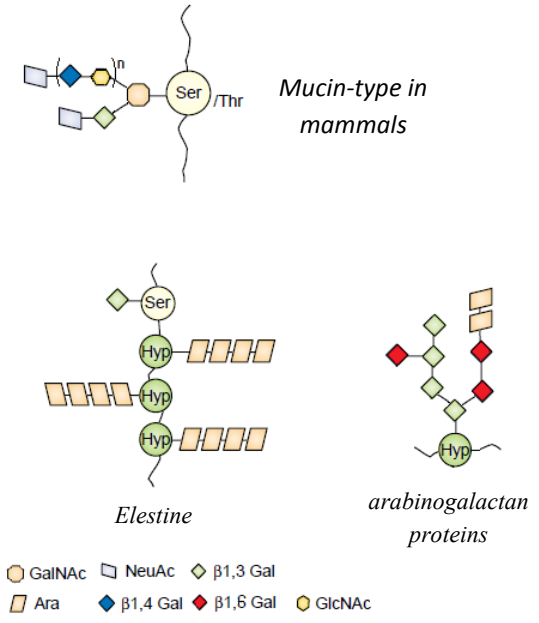


Fig. 1.3 Protein O-Glycosylation in plants and mammals. The arabinose residues are very frequent in plant O-glycosylation but completely absence in mammals. Their abundance is the major cause of immunogenic response.

1.3 FOLLICLE STIMULATING HORMONE

Follicle Stimulating Hormone (FSH) is a gonadotropin that stimulates steroidogenesis and gametogenesis in the gonads. Secreted by the anterior pituitary gland, FSH regulates the menstrual cycle and ovarian follicular maturation in women and supports sperm production in men. FSH acts by binding to the FSH receptor (FSHr) on the granulosa cell surface in ovaries and the Sertoli cell surface in testes (Simoni 1997). FSH belongs to the pituitary glycoprotein hormone family (GPH) which also comprises luteinizing hormone, chorionic gonadotropin and thyroid stimulating hormone. These hormones share a common α subunit non covalently associated to a specific β subunit to form a functional heterodimer. While the α-subunit primary structure is identical for all glycoprotein hormones within the same species, the oligosaccharide populations differ in a hormone-specific manner (Davis 2014).

Both subunits possess two potential N-glycosylation sites. While the FSHα subunit exhibits only micro-heterogeneity in its glycosylation, the FSHβ subunit exhibits both macro- and micro-heterogeneity reflecting the endocrine status and/or the stage of female reproductive cycle (Chappel 1998, Ulloa Aguirre 1998, Ulloa-Aguirre 2011, Bousfield 2015). FSH is a disulphide-rich heterodimer: FSHα subunit contains 92 amino acids and 5 disulphide bridges while the beta subunit consists of 111 amino acids with 6 disulphide bridges. Furthermore, FSH is a member of the glycoprotein hormone family, which is a subfamily of the cysteine knot growth factor superfamily (**Fig. 1.4**). The cystine-knot is a ring structure comprised of two disulfide bonds through which the third disulfide bond penetrates. Both subunits contain the cysteine knot motif (**Fig. 1.5 and 1.6**).

All glycoprotein hormones are important pharmaceutical drugs (PDR 2013). FSH is used clinically for controlled ovarian stimulation in women treated with assisted reproductive technologies and also for the treatment of anovulatory infertility in women and hypogonadotropic hypogonadism in men. The central role of FSH in human reproduction makes its receptor a unique pharmaceutical target in the field of fertility regulation. In conventional *in vitro* fertilization (IVF) treatment; gonadotrophins are administered in order to stimulate an ovarian cycle. Follicle stimulating hormone is universally recognized as the key driver of ovarian follicle growth and maturation, and is most often administered in one of two forms: recombinant FSH expressed in CHO system (rFSH; Gonal-F®(Merck Serono) or Puregon® (Merck Sharp and Dohme)) or highly purified human menopausal gonadotrophin

(hMG-HP; Menopur® (Ferring) or Merional® (Pharmasure)) which contains both FSH and luteinising hormone (LH) activity in a ratio of 1:1 (Threw 2010).

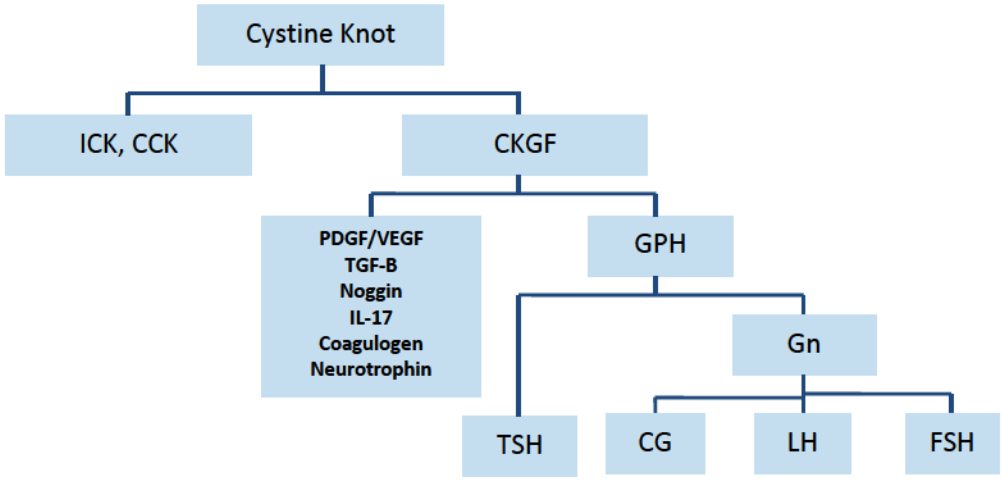


Fig. 1.4 Cystine knot classification. Abbreviations used in this figure are as follows. GPH: glycoprotein hormone Gn: gonadotropins; FSH: Follicle Stimulating Hormone; LH: luteinizing hormone; CG: chorionic gonadotropin; TSH: thyroid-stimulating hormone; PDGF: platelet-derived growth factor; TGF-β: transforming growth factor-β; IL-17: interleukin-17; CKGF: cystine-knot growth factor; ICK: inhibitor cystine knot; CCK: cyclic cystine knot.

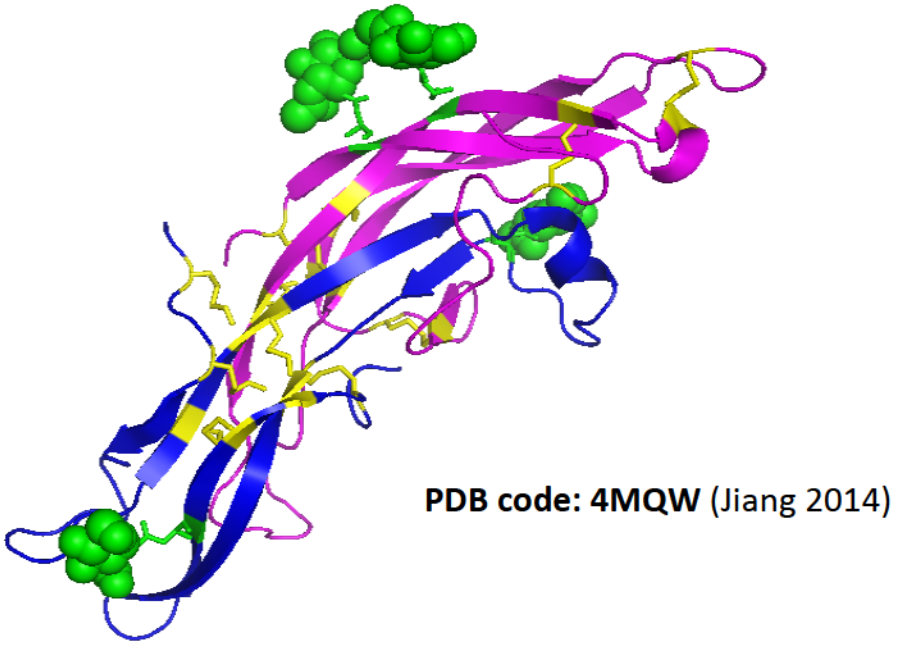


Fig.1.5 Crystallographic structure of Follicle Stimulating Hormone (PDB code: 4MQW). FSHβ subunit is colored in violet, FSHα subunit in blue, disulphide bridges in yellow. The N-glycosylation sites (N52-78 in alpha subunit and N7-24 in beta subunit) are colored in green as the N-acetylglucosamine residues (green spheres). The image was elaborated with the free software Pymol.

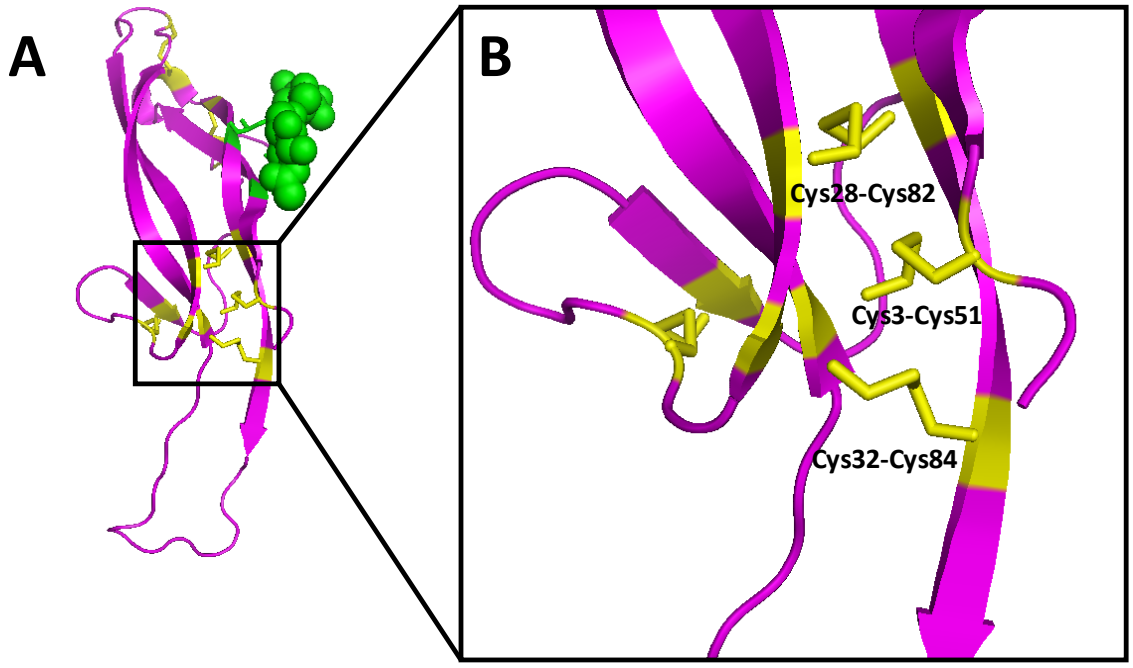


Fig. 1.6 Crystallographic structure of Follicle Stimulating Hormone (PDB code: 4MQW).
(**A**) Follicle stimulating hormone beta subunit (violet). The N-glycosylation sites N7-24) are colored in green as the N-acetylglucosamine residues (green spheres) while disulphide bridges are colored in yellow. (**B**) Cystine knot motif in FSHβ subunit (yellow): Cys3-51, Cys28-82, Cys32-54. The image was elaborated with the free software Pymol.

2. EXPERIMENTALS

2.1 REAGENTS

Competent cell lines of *Nicotiana benthamiana* for human follicle stimulating hormone beta subunit (rhFSHβ) were obtained in the department of Molecular Biology of Active Botanicals Research (ABResearch, Brendola, VI, Italy).

Sucrose Low Endotoxin (ARK2195B), L-Methionine (M5308), L-Ascorbic Acid (A5960), Citric Acid (791725), Sodium Phosphate Dibasic Anhydrous (76140), Sodium Chloride (S7653), Ethylenediaminetetraacetic Acid (EDS), Phenylmethanesulfonyl Fluoride (P7626), Tween20 (P1379), Amberlite XAD-4 (06444), Polyvinylpyrrolidone (P2307), Ammonium Sulfate (A4418), Imidazole (I5513), Silver Nitrate (209139), Ethanol (02854), Acetic Acid (A6283), Glutaric dialdehyde Solution 25% (G5882), Sodium Thiosulfate (S6672), Sodium Acetate (S8750), Formaldehyde Solution 37% (F8775), Sodium Carbonate (451614), Tris-HCl (857645), Glycine (G8898, 410225), Skim Milk Powder (70166), Dodecyl Sodium Sulfate (11667289004), 2-Mercaptoethanol (M3148), Bromophenol Blue (B8026), Acrylamide: Bis-Acrylamide 37:1 (A6050), N,N,N′,N′-Tetramethylethylenediamine (TEMED,T7024), Ammonium Persulfate (APS, A3678), Glycerol (G5516), Coomassie Brillant Blue R (B7920), Methanol (34860), 1-Butanol (34867), ColorBurst Electrophoresis marker (C1992), Formic Acid (14265), Trichloroacetic Acid (T9159), Trifluoroacetic acid (302031), Acetone (650501), QuantiPro BCA Assay Kit (QPBCA), Bradford Reagent (B6916), PNGase F (P7367), Iodoacetamide (I1149), Dithiothreitol (43815), Ammonium Bicarbonate (09830), 4-Chloro-7-Nitrobenzofurazan (25455), Sodium Citrate Monobasic (71497) and Subtilisin A (P5380) were purchased from Sigma Aldrich (St. Louis, Missouri, USA).

Methanol Ups ultra LC (H411) and Acetonitrile Ups ultra LC (H050) were purchased from Romil Pure Chemistry (Cambridge, UK).

Sequencing Grade Modified Trypsin (V5111) was purchased from Promega (Fitchburg, Wisconsin, USA).

BlueElf Prestained Protein marker 5-245 kDa (PS-105) was purchased from Jena Bioscience (Löbstedter, Germany).

4 ml Vivaspin *cutoff* 10 kDa (*VS0403*), 15 ml Vivaspin cutoff 3 kDa (VS2092) and 0.2 μm cellulose nitrate filters (11327-47-N) were purchased from Sartorius Stedim (Gottinga, Germany)

Rabbit polyclonal anti human FSH-beta antibody (Ab171431), Horseradish peroxidase-conjugated goat anti-rabbit IgG antibody (474-1506) and Standard rhFSHβ (H00002488-Q01) were purchased from Prodotti Gianni (Milan, Italy).

Rabbit polyclonal anti human FSH-beta antibody (130-10169) was purchased from Raybiotech (Atlanta, Georgia, USA)

Amersham ECL Western Blotting Detection Reagents (RPN2209) and Amersham Hybond blotting paper (RPN610M), Sepharose Ni-IDA resin (17-5318-01), Sephadex G-25 Medium resin (17-0033-01), Sephadex G-10 resin (17-0010-01) and SP Sepharose HP (17-1087-01) were purchased from GE Healthcare (Little Chalfont, UK).

Immobilon-Fl (Transfer Membranes) (IPFL00010) was purchased from Merk Millipore (Billerica, Massachusetts, USA).

RapiGest SF Surfactant was purchased from Waters (Milford, Massachusetts, USA).

Native Human Chorionic Gonadotropin (CGA) (CGA-8163H) was purchased from CreativeBioMart (New York, New York, USA).

The required material for biological activity evaluation: commercial recombinant follicle stimulating hormone (GONAL-F, Serono, Rome, Italy), testosterone enanthate (SIT, Pavia, Italy) and Isolated Porcine Sertoli cells (obtained by prepubertal (7-15 days) “large white” pigs), was generously offered by Carlo Foresta’s research group (Endocrinology and Male Fertility center, Padova, Italy).

2.2 METHODS

2.2.1 Cryo-Extraction in Liquid Nitrogen

An aliquot of competent cells was centrifuged with Avanti J-20 Beckman centrifuge (Brea, California, USA)

at 5000g for 15 minutes at 4°C in order to remove the culture medium. Cells were weighed and chopped in a mortar with liquid nitrogen to obtain a fine and homogeneous powder. The extraction buffer was added (1g: 3ml) maintaining manual grinding under constant addition of liquid nitrogen. Homogeneous powder was centrifuged at 5000g for 10 minutes at 4°C. Supernatant was saved and the pellet discarded.

Extraction Buffer: 50 mM Na_2HPO_4, 150 mM NaCl, 20 mM citric acid, 40mM ascorbic acid, 5 mM EDTA,1 mM PMSF, 0,05% Tween20, 1 % (w/v) Amberlite XAD-4, 1% (w/v) Polyvinylpyrrolidone, pH 6.5.

2.2.2 The Bradford Method for Protein Quantification

This protein assay is based on *Brilliant Blue G* staining. The protein sample was mixed with the reagent and absorbance was monitored at 595 nm after short room temperature incubation. Analyses were performed with *Bradford Reagent Kit* and PerkinElmer's VICTOR 3 multi-label plate reader (Waltham, Massachusetts, USA).

Bovine Serum Albumin stock solution (1mg/ml) was used to produce the standard curve by serially diluting, ranging from 0.2-1 mg/ml (**Tab. 2.1**). 5 μl of the protein standards were added to separate wells in the 96 well plates. 250 μl of the Bradford Reagent were added and mixed on a shaker for 30 seconds. Before absorbance reading, the samples were incubated for 30 minutes at 37°C. Absorbance was plotted versus the protein concentration of each standard (**Fig 2.1**).

Sample	BSA 1mg/ml (μl)	Deionized Water (μl)	BSA (mg/ml)	Bradford Reagent (μl)
Blank	0	5	0	250
1	1	4	0.2	250
2	2	3	0.4	250
3	3	2	0.6	250
4	4	1	0.8	250
5	5	0	1.0	250

Tab. 2.1 Standard Assay Set Up Table: standards were created by serially diluting of 1 mg/ml BSA protein standard.

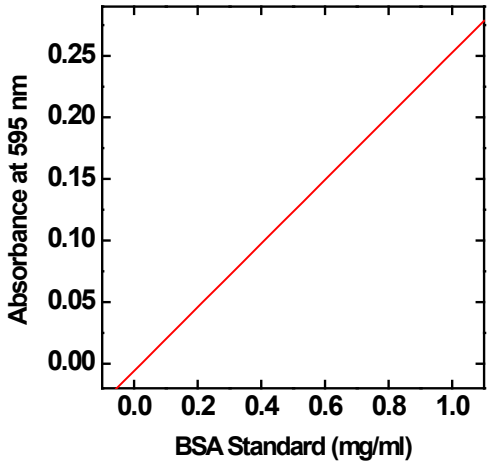


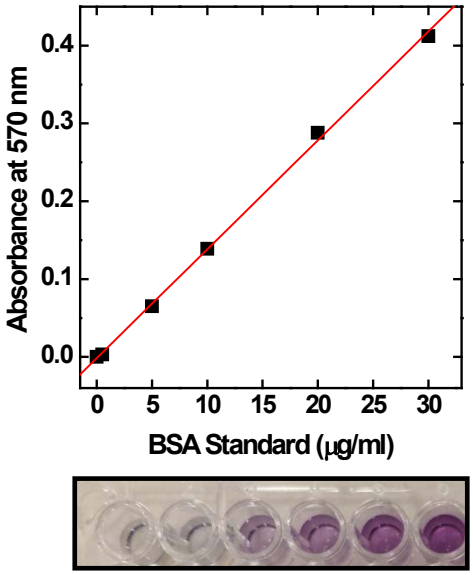
Fig 2.1 Standard Curve: standard curve produce by plotting the monitored absorbance at 595 nm versus the protein concentration of each standard.

2.2.3 Protein determination by the bicinchoninic acid assay (BCA)

This protein assay is based on the formation of a Cu^{2+}-protein complex under alkaline conditions followed by reduction of the Cu^{2+} to Cu^{1+}. The amount of reduction is proportional to the protein present. BCA forms a purple-blue complex with Cu^{1+} in alkaline environments with maximum absorbance at 562 nm (range of 540–590 nm can be substituted). *QuantiPro BCA Assay Kit* is used to measure very dilute protein concentrations (0.5 to 30 μg/ml) in very small sample volumes.

Bovine Serum Albumin stock solution (50 μg/ml) was used to produce the standard curve by serially diluting, ranging from 0.5-30 μg /ml (**Tab 2.2**). 150 μl of the protein standards were added to separate wells in the 96 well plates. 150 μl of the BCA Reagent were added and mixed on a shaker for 120 seconds. Before absorbance reading, the samples were incubated for 2 hours at 37°C. Absorbance was plotted versus the protein concentration of each standard (**Fig. 2.2**).

Sample	BSA 50 µg/ml (µl)	Deionized Water (µl)	BSA (µg/ml)	BCA Reagent (µl)
Blank	0	150	0	150
1	1.5	148.5	0.5	150
2	15	135	5	150
3	30	120	10	150
4	60	90	20	150
5	90	60	30	150



Tab. 2.2 Standard Assay Set up Table: standards were created by serially diluting of 50 µg/ml BSA protein standard. **Fig. 2.2 Standard Curve**: standard curve produce by plotting the registered absorbance at 570 nm versus the protein concentration of each standard.

2.2.4 Trichloroacetic Acid (TCA) Protein Precipitation

Milli-Q water was added to each sample in order to reach 200 µl. Subsequently, 50 µl of 100% (w/v) TCA solution freshly prepared were added. The samples were vortexed and placed at -20°C for 1 hour. After that, samples were centrifuged at 13000 g for 10 minutes at 4°C. The supernatant was accurately removed in order to save the precipitate. The excess of TCA was removed by repeated additions of cold acetone (-20°C). After every addition the samples were centrifuged at 13000g for 10 minutes at 4°C and the supernatant was discarded. Samples were dried and stored.

2.2.5 Sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis

Homemade polyacrylamide gels were used. Analyses were performed under denaturing conditions with or without reducing agents, using *Biorad MiniProtean II* apparatus (Hercules, California, USA). *Stacking gel* (4% acrylamide) solution, *Running gel* (10/12% acrylamide) solution and *Running Buffer* (25 mM Tris-HCl, 192 mM Glycine, 0.1% (w/v) SDS) were freshly prepared (**Tab. 2.3**). Two glass plates (10.1 cm x 7.1cm and 10.1 cm x 8.1 cm) with 1mm spacer were used for gel polymerization.

Running Gel	***Volume***	***Stacking Gel***	***Volume***
Acrylamide Solution	1.4 ml	Acrylamide Solution	200 μl
Tris-HCl pH 8.8 (1.875 M)	800 μl	Tris-HCl pH 6.8 (0.625	400 μl
Milli-Q Water	1.75 ml	Milli-Q Water	1.35 ml
SDS Solution 10% (w/v)	40 μl	SDS Solution 10% (w/v)	20 μl
TEMED	2.5 μl	TEMED	2 μl
APS Solution 20% (w/v)	12.5 μl	APS Solution 20% (w/v)	10 μl

Tab. 2.3 Composition of running and stacking gel. The reported quantities refer to one gel.

Samples were resuspended in Laemmli buffer (**Tab. 2.4**) and boiled for 10 minutes. Subsequently, they were centrifuged at 10000g for 30 seconds and loaded in wells. Electrophoretic runs were carried out by applying a constant current of 15 mA.

Laemmli Sample Buffer	***Amount***
Glycerol	500 μl
2-Mercaptoethanol	315 μl
Tris-HCl pH 6.8 (0.625 M)	50 μl
Milli-Q Water	140 μl
SDS Solution 10% (w/v)	250 μl
Bromophenol Blue	Some μg

Tab. 2.4 Laemmli Sample Buffer. Sample loading buffer was added to each sample before loading. The presence or absence of 2-Mercaptoethanol determines the reducing or no-reducing (respectively) conditions of analysis.

2.2.6 Coomassie Blue Staining method

After electrophoresis, gel was rinsed 3 times for 2 minutes with 20 ml of deionized water to remove SDS and buffer salts. Subsequently, gel was covered for 1 hour at room temperature under gentle shaking with staining solution freshly prepared by mixing alcoholic 0.5% *Coomassie Brillant Blue R-250* solution, with 20% (v/v) acetic acid solution in equal part. Gels were treated with destaining solutions: (1) 7% (v/v) acetic acid, 40% (v/v))methanol for 30 minutes, (2) 7% (v/v) acetic acid, 5% (v/v) methanol for 12 hours.

2.2.7 Silver Staining method

After electrophoresis, (**I**) gel was rinsed 3 times for 2-3 minutes with 20 ml deionized water to remove SDS and buffer salts. During the (**II**) fixing step, the gel was treated with water:ethanol:acetic acid 5:4:1 solution for 30 minutes. (**III**) Sensitization step was carried out by covering gel in 30% (v/v) ethanol, 0.72% (v/v) glutaric dialdehyde, 0.2% (w/v) sodium Thiosulfate and 6.8% (w/v) sodium acetate solution for 30 minutes. (**IV**) Gel was rinsed 4 times for 5 minutes with deionized water. (**V**) Gel was treated with 0.25% (w/v) silver nitrate, 0.04% (v/v) formaldehyde solution for 20 minutes. (**VI**) Rinse 2 times for 1 minute with deionized water. After that, the final development step was carried out by adding 2.5% (w/v) sodium carbonate, 0.03% (v/v) formaldehyde solution for some minutes until bands started to appear and the desired band intensity was reached. (**VII**) Stopper solution (0.5% (v/v) acetic acid) was added to prevent further reduction of silver ions.

2.2.8 Immunoblotting and Enhanced ChemiLuminescence

(**I**) The electrophoretic gel was washed with milliQ-water to remove the excess of SDS. (**II**) Protein was transferred completely onto Immobilon PVDF membrane with *Biorad Mini-Trans-Blot cell* (Hercules, CA, U.S.A.) in *Towbin Running Buffer* (25 mM Tris-HCl, 200 mM glycine, 0.05% (w/v) SDS pH=8) at 4°C, 400 mA for two hours. (**III**) The membrane was covered with phosphate buffer solution (PBS) and coated by adding 5% (w/v) skim milk powder. (**IV**) After 2 hours on orbital shaker it was abundantly washed with milli-Q water and (**VI**) incubated overnight with rabbit polyclonal anti human FSH-beta antibody (dilution 1:1000) in PBS at 4°C. (**VII**) Membrane was rinsed 3 times for 15 minutes with 0.1% (v/v) Tween20-PBS in order to remove the primary antibody in excess. (**VIII**) Immunoblotting analysis carried out by incubating the membrane with horseradish peroxidase-conjugated goat anti-rabbit IgG antibody (dilution 1:10000) in PBS at 4°C. (**IX**) After washing with milli-Q water, the ternary complex was detected by chemiluminescence with *ECL western blot detection reagent* with *Biorad VersaDocMP 4000* (Hercules, CA, U.S.A.).

PBS: 10 mM Na_2HPO_4, 150 mM NaCl, pH=7.5

2.2.9 Extraction and Purification or rhFSHβ in cell lines of *Nicotiana benthamiana*

2.2.9.1 Extraction

Suspension cell lines of *N. Bentamiana* were filtered with a cloth filter (*cut-off* 50 μm) in order to remove the medium culture and stored at -80°C. Frizzed cells were thawed at +4°C overnight. The washing buffer (20 mM Na_2HPO_4, 10 mM EDTA, pH 7.2) was added to cells (1g:2ml). The suspension was stirred for 1 hour and centrifuged at 18000 g for 30 minutes with Avanti J-20 Beckman centrifuge (Brea, California, USA.). The precipitate was resuspended in the extraction buffer (50 mM Na_2HPO_4, 150 mM NaCl, 20 mM citric acid, 40mM ascorbic acid, 5 mM EDTA, 1 mM PMSF, 0,05% Tween20, 1 % (w/v) XAD-4, 1% (w/v) polyvinylpyrrolidone, pH 6.5) (1g : 3ml) . After homogenization by mixer, the suspension was stirred for 1 hour and centrifuged at 18000g for 30 minutes. Ammonium sulfate was added up to 60% to the supernatant in order to isolate and concentrate proteins by mixing for 1 hour. The suspension was centrifuged at 18.000g for 30 minutes. The precipitate was resuspended in the IMAC buffer (20 mM Na_2HPO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0) and filtered with 0.2 μm cellulose nitrate membrane filters. All steps were performed at +4°C.

2.2.9.2 Purification

I: Immobilized-metal affinity chromatography (IMAC)

The first chromatographic step was performed by a homemade packed column with Sepharose Ni-IDA resin. The column was equilibrated with the IMAC buffer and it was eluted with 20 mM Na_2HPO_4, 300 mM NaCl, 500 mM imidazole, pH 8 at 1 ml/min with a multiple-steps gradient: 25 CV at 0%, 1 CV 0-5%, 5 CV 5%, 1 CV 5-10%, 5 CV 10%, 1 CV 10-40%, 5 CV 40%, 1 CV 40-100%, 5 CV 100%, 1 CV 100-0% (Column Volume (CV): 15 ml).

II: Size-exclusion chromatography (SEC)

The positive fractions were collected and eluted in a homemade packed column with Sephadex G-25 Medium resin. Equilibration and elution steps were performed with the 50 mM CH_3COONa, 150 mM NaCl, 60 μM Tween20, pH 5.5 buffer at 5 ml/min.

III: Ion-exchange chromatography (IEC)

The last step was carried out by ion exchange chromatography using a homemade packed column with Sulfopropyl resin. The column was equilibrated with a salt buffer 50 mM CH_3COONa, 150 mM NaCl, 60 μM Tween20, pH 5.5 and eluted with a multiple-steps gradient of 50 mM CH_3COONa, 1 M NaCl, 60 μM Tween20, pH 5.5 buffer at 3 ml/min: 2 CV at 0%, 5 CV 25%, 5 CV 65%, 1 CV 100% (Column Volume (CV): 50ml).

The chromatographic analysis were performed with the AKTA purifier UPC 100 system GE Healthcare (Fairfield, CT, U.S.A.) equipped with pumps P-900 and the UPC-900 UV detector, monitoring the absorbance at 280. After every step, the collected fractions were analyzed by SDS-PAGE and Immunoblotting by ECL.

2.2.10 Chemical Characterization

2.2.10.1 LC-MS Analysis

LC-MS analyses were performed with UPLC Agilent 1290 Infinity (Santa Clara, California, USA) combined with ESI-TOF XEVO G2-S Qtof (Milford, Massachusetts, USA). Data were acquired using the MassLinx Mass Spectrometry™ and processed with *BiopharmaLynks™* software (*Waters*, Milford, Massachusetts, USA).

2.2.10.2 Tryptic Digestion in Situ

The target protein band was cut and transferred to tube. (**I**)The sample was washed with milliQ-water and subsequently with acetonitrile in order to shrink the gel. (**II**)The residue was dried in *SAVANT Speed Vac Concentrator* (Louisville, Kentucky, USA) for 15 minutes and suspended in 0.1 M NH_4HCO_3, 10 mM DTT buffer at 56°C for 30 minutes. (**III**) After centrifugation the buffer was substituted with acetonitrile and *Speed Vac* treatment was repeated. (**IV**) 0.1 M NH_4HCO_3, 55 mM iodoacetamide buffer was added for 20 minutes and subsequently the sample was washed with 0.1 M NH_4HCO_3. (**V**) After drying process, the tryptic solution was added (12.5 ng/μl) at 37°C overnight. (**VI**) 0.25 M NH_4HCO_3 buffer and Acetonitrile were added in equal volume, the sample was centrifuged and the supernatant

saved. (**VII**) 5% HCOOH solution was used as final treatment of gel spot. (**VIII**) The supernatant was collected with the first one and dried in *Speed-Vac Concentrator*.

The residue was resuspended in 95% milliQ-water, 3% Acetonitrile and 2% formic acid. LC-MS analysis was performed with a C18 Vydac column 1x150 mm, 5 μm, 300 Å with a linear gradient of Acetonitrile (3-65% in 25 minutes) with flow rate of 40 μl/min. Data were acquired using the MassLinx Mass Spectrometry™ and processed with *BiopharmaLynks™* software.

2.2.10.3 Deglycosylation with PNGase F

Purified rhFSHβ fractions were collected and placed in 4 ml *Vivaspin* (*cutoff* 10kDa). Samples were concentrated with simultaneous buffer exchange at 10000 g, 4°C. Deglycosylation with PNGase F carried out in 50 mM NH_4HCO_3, 0.1% (v/v) Rapigest SF, pH 7.9 buffer solution with 1:100 molar ratio enzyme: substrate for 48 hours at 37°C. The reaction was stopped by trifluoroacetic acid (TFA) addition (0.1% v/v). The sample was centrifuged at 13.000g for 10 minutes and analyzed by SDS-PAGE.

2.2.11 Conformational Characterization

2.2.11.1 Disulphide-bond Assignment

100 μg of rhFSHβ were (**I**) incubated with PNGase F at 1:100 molar ratio (enzyme: substrate) for 48 hours at 37°C in 50 mM NH_4HCO_3, 1mM $CaCl_2$, pH 8.2 . The reaction was stopped by TFA addition (0.1% v/v). The sample was centrifuged at 13000 g for 10 minutes and analyzed by SDS-PAGE (no reduction conditions). The target protein band was cut, transferred to tube and de-stained by several wash cycles with milli-Q water and acetonitrile. Furthermore, sample was (**II**) digested with Sequencing Grade Modified Trypsin 1:50 molar ratio (enzyme:substrate) as described previously without addition of DTT and iodoacetamide to maintain the disulphide bridge pattern. The lyophilized residue was resuspended in 50 mM NH_4HCO_3, 1mM $CaCl_2$, pH 8.2 and (**III**) digested with subtilisin (enzyme:substrate, 1:50 molar ratio) at +37°C overnight for 6 hours. The reaction was stopped by addition of TCA solution.

Dried sample was resuspended in 95% milliQ-water, 3% Acetonitrile, 2% formic acid and 20 μl were analyzed by LC-MS (50 μl/min). LC-MS analysis was performed with a C18 Vydac column 1x150 mm, 5 μm, 300 Å with a linear gradient of Acetonitrile (3-65% in 12 min).

Analysis were scanned from m/z 400 to 2000 at 10s/scan, spectra acquired using MassLinx™ software and processed with with *BiopharmaLynks™* software.

2.2.12 *In vitro* Functional analyses

2.2.12.1 Samples and Buffer Solutions Preparation

Milli-Q water was sterilized with autoclave at 121°C for 21 minutes and filtered with 0.2 μm cellulose nitrate filter**.** Subsequently, 3 different solutions were prepared: (*S1*) 90 mg/ml *Sucrose Low Endotoxin* solution, (*S2*) 20 mg/ml L-methionine solution and (*S3*) 20 mM Na_2HPO_4, 0.0034% (v/v) Tween20, pH 7.0 buffer solution.

Collected rhFSHβ fractions were placed in 15 ml *Vivaspin* (*cutoff* 3kDa) and centrifuged at 7500 g for 1 hour at 4°C. *S3* was added to the sample up to 15 ml and it was centrifuged at 7500 g for 1 hour at 4°C. This step was repeated several times in order to concentrate the sample in S3 buffer solution. Subsequently, S2 and S1 were added with a final concentration of 3.2 mg/ml of sucrose and 0.1 mg/ml L-methionine. Sample were lyophilized and stored at -20°C.

Fluorescent labeled Gonal-F was obtained by selective N-terminal fluorescent labeling using 4-chloro-7-nitrobenzofurazan (NBD-Cl) (Lina 2012). A stock solution of 50 mM NBD-Cl was prepared in acetonitrile and stored at 4°C in the dark. The protein sample (27.5 μg/ml) was mixed with 0.5 mM NBD-Cl in 50 mM sodium citrate buffer containing 1 mM EDTA, pH 7.0 at room temperature and overnight.

Fluorescence derivatization was monitored with a Jasco Spectrofluorometer FP-6500 (Oklahoma City, Oklahoma, USA) equipped with *Peltier system*. Fluorescence spectra were measured at 25°C, samples (40 nM in 1.5 ml) were excited at 470 nm and the emission spectra were monitored in 500-750 nm range.

Excess of NBD-Cl was removed by size-exclusion chromatography with AKTA purifier UPC 100 system equipped with a homemade packed column with Sephadex G-10 resin (1x12 cm). Equilibration and elution steps were performed with 20 mM Na_2HPO_4, 0.0034% (v/v) Tween20, 3.2 mg/ml sucrose, 0.1 mg/ml L-methionine, pH 7.0 at 0.8 ml/min, monitoring the absorbance at 280 nm.

The heterodimer consisting of rhFSHβ and commercial CGA was associated (Thomas W. and David P. 1982). The reaction was performed at 37°C for 72 hours in 50 mM Na_2HPO_4 , pH 7.0 with 1:1 molar ratio (2 mg/ml). The reconstituted heterodimer was purified by size exclusion chromatography with Jasco HPLC Pu-1575 equipped with 1575 UV-Vis detector. Purified step was performed with *Phenomenex YARRA column 3 mm SEC 3000*, 7.8x150 mm (Torrance, California, USA) eluted in 20 mM Na_2HPO_4, 150 mM NaCl, pH 7.4 at 0.5 ml/min. Absorbance was monitored at 226 nm.

Test of functional activity has been performed at the Center of Endocrinology and Fertility of the University of Padova.

2.2.12.2 Binding Analyses to FSH receptor on *in vitro* porcine Sertoli cells.

Isolated porcine Sertoli cells (SCs) were obtained from prepubertal pig testes upon approval by Ethical Committe. The isolation of Sertoli cells process was developed in according to the Menegazzo's protocol (Menegazzo 2011). SCs ($20x10^6$ cells) were treated with fixed ammount of labbeled commercial rFSH (76 nM) and increasing amount of rhFSHβ/rhFSHd up to 7.6 μM. The interaction between recombinant protein from suspension cells of *N. Benthamiana* and FSH receptor was evaluated calculating the amount of displaced NBD-labelled. It was determined by monitoring the emission signal at 550 nm after 3 days of incubation, with Nikon ViCo equipped with microscope eclipse 80i (Shinagawa Intercity, Tokyo, Japan).

2.2.12.3 Evaluation of Biological Activity (Aromatase Activity)

SCs ($20x10^6$ cells) were incubated with 1μg/ml of rhFSHβ protein for three days. After that, the level of 17b-estradiol was evaluated. In parallel the same analysis was performed with addition of testosterone (0.2 mg/ml) 8 hours before the 17β-estradiol detection. 17β-estradiol was detected by ChemiLuminescence (ADVIA Centaur, Estradiol-6 III, Bayer Diagnostics, Germany).

3. RESULTS & DISCUSSION

3.1 EXPRESSION

Competent cell lines of *N. benthamiana* for recombinant human Follicle Stimulating Hormone beta subunit (rhFSHβ) were obtained in the department of Molecular Biology of Active Botanicals Research (ABResearch, Brendola, VI, Italy). rhFSHβ expression vector was designed in order to obtain the 6-His/KDEL-tagged protein developed from tumor-inducing (Ti) plasmid (Marillonnet 2005) (**Fig. 3.1**).

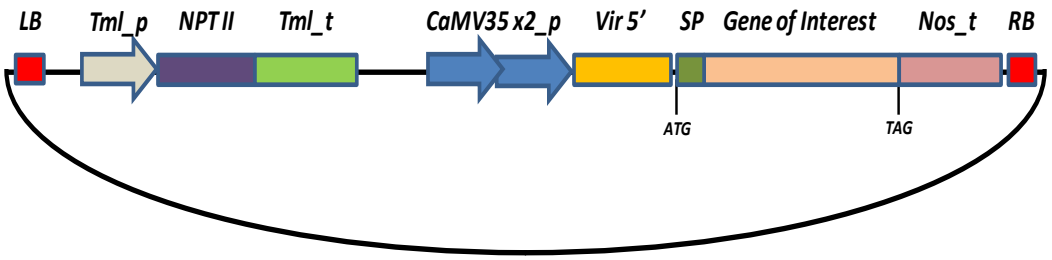


Fig.3.1 New Minimal ABR Binary Vector T-DNA Structure. The rhFSHβ expression vector designed from the Ti plasmid. **RB/LB :** Site-specific recombination, right and left border .**Tml p**:Tumor morphology rootym promoter. **NPT II**: Neomycin phosphotransferase II (kanamycin resistance)**.Tml t**: Tumor morphology rootym terminator. ***CaMV35 x2_p****:* Cauliflower mosaic virus promoter**. Vir 5':** 5'-untranslated region plant virus**. SP**: Plant signal peptide. **Nos t**: Nopaline synthase terminator.

3.2 EXTRACTION

3.2.1 Cryo-Extraction in Liquid Nitrogen

An aliquot of competent suspension cells of *N. benthamiana* was manually grinded with addition of liquid nitrogen and extraction buffer (50 mM Na_2HPO_4, 150 mM NaCl, 20 mM citric acid, 40mM ascorbic acid, 5 mM EDTA, 1 mM PMSF, 0.05% Tween20, 1 % (w/v) Amberlite XAD-4, 1% (w/v) Polyvinylpyrrolidone, pH 6.5). Because of cell wall composition (cellulose, polysaccharides, pectins, etc.), manual grinding of frozen cells in liquid nitrogen was the fastest and most efficient way to access plant cell proteins. The extraction buffer was added to frozen cells to obtain the first separation of protein component from lipids and glycans.

Amberlite XAD-4 is a polymeric adsorbent based on highly crosslinked, macroreticular polystyrene, aliphatic, or phenol-formaldehyde condensate polymers. Amberilte XAD-4 after extensive washing with water and methanol removes most of phenolic compounds and phenol (Li 2001). To improve the phenolic remotion from the extract, pretreated polyvinylpolypyrrolidone was added. PVPP free from metal ions and other contaminants was obtained by boiling in 10% HCl for 10 minutes and then washed extensively with glass-distilled water (Loomis 1979). In the context of protein purification, proteases are unwanted destructive contaminants that need to be inactivated and removed. Phenylmethylsulfonyl fluoride (PMSF) irreversibly inhibits serine proteases (James 1978) by selective covalent linkage to serine residue in the active site. Ethylenediaminetetraacetic acid (EDTA) was used to inhibit the metallopeptidases via chelation of metal ions required for catalytic activity. The buffer extraction was supplemented with citric and ascorbic acid to reduce oxidation damage, and Tween20 as detergent.

The extract was analyzed with SDS-PAGE (4-10% acrylamide) and Immunoblotting, comparing to *wild type* sample (**Fig. 3.2**). *Wild type* extract was obtained by transfected suspension cell lines of *N. benthamiana* with the same expression vector without the gene coding the protein of interest.

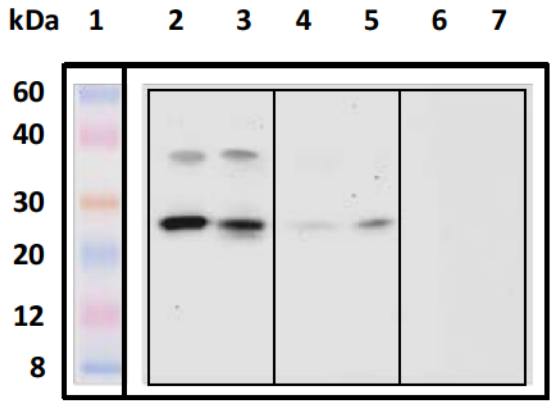


Fig. 3.2 Western Blot analysis of extract from competent suspension and wild type cells of N. benthamiana. (lane 1) 3 µl of Color Burst E. Marker, (lane 2-3) 15 and 30 µl of extract from competent cell lines for rhFSHβ. The membrane was incubated with rabbit polyclonal anti human FSHβ antibody (Ab171431) dilution 1:1000. (lane 4-5) 15 and 30 µl of extract from competent cell lines for rhFSHβ. The membrane was incubated with rabbit polyclonal anti human FSHβ antibody (130-10169) dilution 1:1000. (lane 6-7) 200 µl of *wild type* extract. The membrane was treated with (lane 6) rabbit polyclonal anti human FSHβ antibody (Ab171431) dilution 1:1000, (lane 7) rabbit polyclonal anti human FSHβ antibody (130-10169) dilution 1:1000.

Extract from competent suspension cell lines of *N. Benthamiana* were loaded in lanes 2-3-4-5. Lanes 2 and 3 (15 and 30 µl of sample respectively) were treated with rabbit polyclonal anti human FSHβ antibody (Abcam). Two species were identified at 25 and 35kDa. Lanes 4 and 5 (15 and 30 µl of sample respectively) were treated with rabbit polyclonal anti human FSHβ antibody (Raybiotech). Only species at 25 kDa was identified. Data suggested the greater specificity of polyclonal antibody purchased from Abcam.

The extract from *N. Benthamiana wild type* cells was loaded in lanes 6 and 7. Immobilon PVDF membrane incubated with anti human FSHβ antibody (Abcam) and anti human FSHb antibody (Raybiotech) respectively, suggested the absence of nonspecific interactions.

Dynamic populations of FSHβ subunit were widely studied (Bousfield 2015) demonstrating the importance of glycosylation isoforms to fully play biological activities. Two positive anti human FSHβ antibody species suggest the presence of different glycosylation isoforms of recombinant protein.

3.2.2 ABResearch Extraction Method

Cryo-extraction is suitable for small amount of cells, so a new extraction method was developed in order to increase the production yields. KDEL-labeled protein does not undergo the processes of N-glycosylation and O-glycosylation which occur in the Golgi apparatus. KDEL retention signal, in fact, lock the protein within endoplasmic reticulum preventing the release into the cellular cytoplasm. New extraction method, developed from KDEL strategy, allowed the complete removal of cytoplasmic environment saving the integrity of internal organelles. The slow thawing at 4°C and causes the cellular lysis because of the incomplete maturation of cell wall in the callus. Under 500 g of cells, protocol is not recommended: thawing of minor amount of cells occurs in less time allowing the protease to act. The washing buffer (20 mM Na_2HPO_4, 10 mM EDTA, pH 7.2) was added to cells (1g : 2ml) in order to remove the large amounts of cytoplasmic macromolecules. The suspension was stirred for 1 hour and centrifuged at 18000g for 30 minutes. Supernatant was discarded and extraction buffer (composition above) was added to precipitate (1g : 3ml). Suspension was homogenized and stirred for 1 hour at 4°C in order to obtain the first separation of proteins from the lipid and glucose components. Supernatant from "washing" step and homogenized suspension were analyzed by SDS-PAGE (4-10% acrylamide) and Immunoblotting (**Fig. 3.3**).

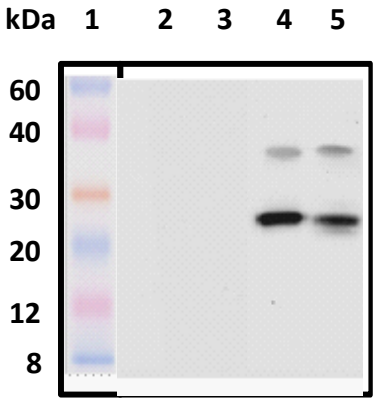


Fig. 3.3 Western Blot analysis of extract from competent suspension and wild type cells of N. Benthamiana and "Washing buffer". (lane 1) 3 μl of Color Burst E. Marker, (lane 2-3) 15 and 30 μl of supernatant post washing step. (lane 4-5) 15 and 30 μl precipitate resuspended in extraction buffer. The membrane was incubated with rabbit polyclonal anti human FSHβ antibody (Ab171431) dilution 1:1000.

The absence of signals in lanes 2 and 3 suggested the storage of recombinant protein in endoplasmatic reticulum as expected from "KDEL strategy". In the extract, two molecular species were identified (25 and 35 kDa) corresponding to proteins obtained with cryo-extraction.

3.2.3 Recombinant human FSH beta subunit quantification

Cryo-extraction versus ABResearch extraction method

Bradford assay was used for protein determination in cryo- and ABRsearch extract. This protein assay is found to be the most reliable since it is less affected by the composition of both buffers. Different conditions were explored in order to optimize protein extraction (**Tab. 3.1**); however, there were no significant differences in amount of total extracted proteins (0.77 mg/ml). To avoid proteolytic processes and aggregation, minor timing protocol was chosen (1 hour).

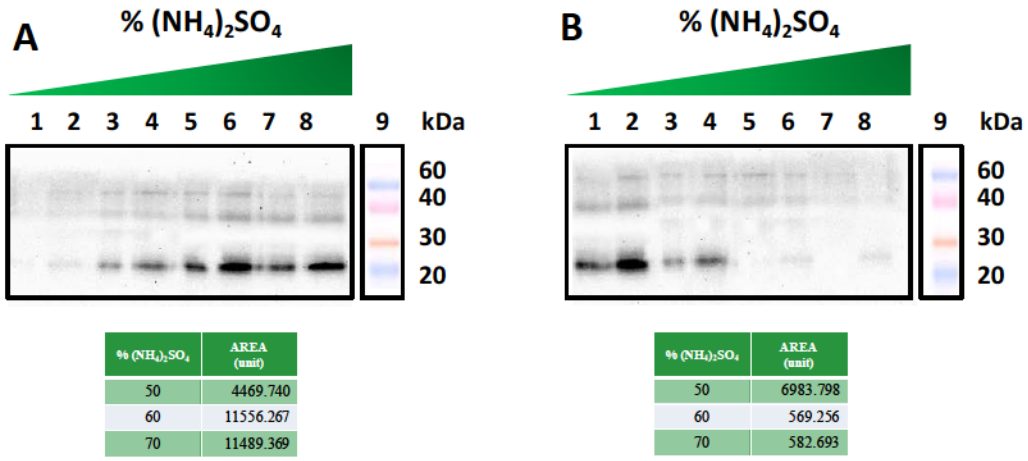
Sample	Abs 1 at 595 nm	Abs 2 at 595 nm	Abs average 595 nm	[mg/ml]	mg of protein/ g cells
Cryo- Extract	0.100	0.125	0.113	0.497	1.491
Washing 0 h	0.000	0.006	0.003	0.091	0.273
Washing 1 h	-0.006	0.008	0.001	0.083	0.249
Washing 2 h	-0.004	0.012	0.004	0.094	0.282
Washing 3 h	0.010	0.012	0.011	0.120	0.360
Extract 0 h	0.015	0.011	0.013	0.128	0.384
Extract 1 h	0.042	0.035	0.049	0.259	0.777
Extract 2 h	0.035	0.061	0.049	0.258	0.774
Extract 3 h	0.046	0.049	0.048	0.256	0.768

Tab. 3.1 Protein determination of cryo-extract, washing buffer and extract.

Washing step decreased the total amount of extracted protein: 1.5 mg/ml to 0.77 mg/ml, compared with liquid nitrogen method. However there wasn't the loss of interest protein because of its retention in the endoplasmatic reticulum, obtaining the suitable enriched sample. Direct quantification of rhFSHβ was not allowed as kit ELISA is not yet available.

3.2.4 Ammonium Sulphate precipitation

His-tagged rhFSHβ was designed in order to facilitate the purification step with Immobilized metal affinity chromatography (IMAC). Ammonium sulphate precipitation ($(NH_4)_2SO_4$) allowed to concentrate the sample with subsequently buffer exchange. Homogenized suspension was stirred for 1 hour and centrifuged at 18000g for 30 minutes at 4°C. 60% ammonium sulphate was added to supernatant and mixed for 1 hour at 4°C. Different saturation conditions were explored (40 to 70%) in order to obtain the more enriched target protein sample with lower contaminants contribution. Supernatant and salt-precipitate were analyzed with Immunoblotting. Target protein was estimate with densitometric analysis using free software *ImageJ* (**Fig. 3.4**).



% $(NH_4)_2SO_4$	AREA (unif)
50	4469.740
60	11556.267
70	11489.369

% $(NH_4)_2SO_4$	AREA (unif)
50	6983.798
60	569.256
70	582.693

Fig. 3.4 Supernatant and salt-precipitate immunoblotting.
(**A**) Immunoblotting and densitometric analysis of protein component in the precipitate, at increasing concentration of ammonium sulfate. (**B**) Immunoblotting and densitometric analysis of the protein component in the supernatant at increasing concentration of ammonium sulfate. (lane 1-2) 15 and 30 µl of 40% ammonium sulfate, (lane 3-4) 15 and 30 µl of 50% ammonium sulfate, (lane 5-6) 15 and 30 µl of 60% ammonium sulfate, (lane 7-8) 15 and 30 µl of 70% ammonium sulfate. Target protein was estimate with densitometric analysis

96% of recombinant protein was recovered in precipitate with ammonium sulfate concentration up to 60%. In order to decrease contaminants contribution caused by high salt concentration, 60% was considered the optimal condition.

3.3 PURIFICATION

1: Immobilized-metal affinity chromatography (IMAC)

Precipitate was resuspended in *IMAC buffer* (20 mM Na_2HPO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0) and filtered with 0.2 μm cellulose nitrate membrane filters. Sodium phosphate buffer was used for equilibration and elution buffer including sodium chloride to reduce ion-exchange effects. Low concentration of imidazole was added to minimize non-specific binding. The chromatographic step was performed by a homemade packed column with Sepharose Ni-IDA resin equilibrated with the *IMAC buffer.* Column was eluted with 20 mM Na_2HPO_4, 300 mM NaCl, 500 mM imidazole, pH 8 by a multiple-steps gradient: 25 CV at 0%, 1 CV 0-5%, 5 CV 5%, 1 CV 5-10%, 5 CV 10%, 1 CV 10-40%, 5 CV 40%, 1 CV 40-100%, 5 CV 100%, 1 CV 100-0% (Column Volume (CV): 15 ml) (**Fig. 3.5/A**).

Eluted sample was collected and each fraction (8.5 ml) was analyzed with immunoblotting in order to identify the rhFSHβ subunit. Western blot analysis (data not shown) confirmed the optimal chromatographic conditions in facts the absence of positive fractions in the *flow through* suggested the suitable binding conditions of His-tagged protein to the resin. Furthermore, nonspecific interactions were discarded with lower concentrations of imidazole (peaks at 5% and 10%) while the positive fraction was eluted at 40% of elution buffer (**Fig. 3.5/A**).

2: Size-exclusion chromatography (SEC)

Positive fractions were collected and eluted in a homemade packed column with Sephadex G-25 Medium resin which is well established gel filtration medium for desalting and buffer exchange in industrial applications. Equilibration and elution steps were performed with 50 mM CH_3COONa, 150 mM NaCl, 60 μM Tween20, pH 5.5 buffer. With the exclusion limit up to 5 kDa, SEC allowed to discard the large amount of imidazole which would compromise the final chromatography step.

The interested sample was eluted with the *dead volume* whilst imidazole was subsequently eluted due to own molecular weights (**Fig. 3.5/B**).

3: Ion-exchange chromatography (IEC)

Ion-exchange chromatography was carried out with the homemade packed column with Sulfopropyl resin. The isoelectric point (pI) of rhFSHβ subunit produced by ABResearch was estimated with online tool *Expasy Protparam*. Isoelectric protein point (pI 6.3) was calculated on the amino acidic composition without considering the contribution of glycoside portion. To enable the interaction between rhFSHβ subunit and the negatively charge resin, the purification step was performed at ph 5.5: the lower pH than pI allows to positively charge the wanted protein. The column was equilibrated with a salt buffer 50 mM CH3COONa, 150 mM NaCl, 60 μM Tween20, pH 5.5 and eluted with a multiple-steps gradient of 50 mM CH3COONa, 1 M NaCl, 60 μM Tween20, pH 5.5 buffer at 3 ml/min: 2 CV at 0%, 5 CV 25%, 5 CV 65%, 1 CV 100% (Column Volume (CV): 50ml).

Eluted sample was collected and each fraction (10 ml) was analyzed with immunoblotting. The positive fraction corresponding to the unique peak at 65% of elution buffer (**Fig. 3.5/C**)was treated with TCA precipitation and analyzed with SDS-PAGE and Silver Staining detection (**Fig. 3.5/D/E**).

SDS-PAGE analysis showed the presence of two molecular species at 25 and 35 kDa, both positive to rabbit polyclonal anti human FSHβ antibody. These species were chemically and conformational characterized.

Quantification of final protein sample was estimated with bicinchoninic acid assay due to its reliability with glycosylated species. Bradford assay, in fact, is known to underestimate glycosylated proteins.

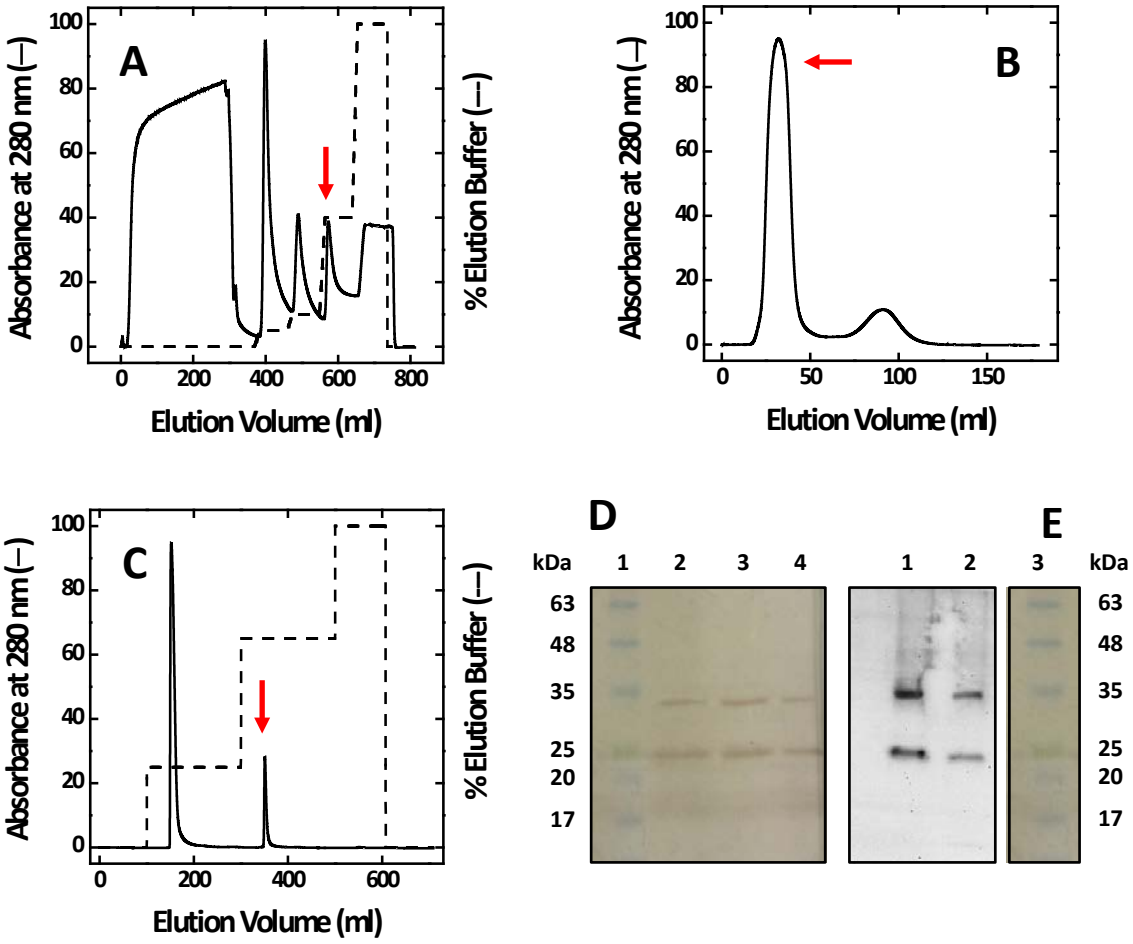


Fig. 3.5 Purification and Characterization of rhFSHβ. (**A**) Immobilized-metal affinity chromatography, (**B**) Size exclusion chromatography, (**C**) Ion-exchange chromatography. After each purification step the eluted sample was analyzed by SDS-PAGE and Immunoblotting. The red arrows indicate the positive fractions. (**D**) SDS-PAGE analysis (4-12% acrylamide) of purified rhFSHβ (lane 1) BlueELf Prestained Protein Marker (Jena Bioscence), (lane 2) 480 ng of rhFSHβ sample, (lane 3) 480 ng of rhFSHβ sample, (lane 4) 240 of rhFSHβ sample. (**E**)Immunoblotting analysis of purified rhFSHβ : (lane 1) 50 ng of rhFSHβ sample, (lane 2) 25 ng of rhFSHβ sample, (lane 3) BlueELf Prestained Protein Marker (Jena Bioscence).

3.4. CHEMICAL CHARACTERIZATION

3.4.1 Mass Spectrometry analysis

After TCA precipitation of positive fraction from ion-exchange chromatography, precipitate was resuspended in 3% acetonitrile,2% formic acid solution and analyzed with Xevo G2-S Q-tof. LC-MS analysis was performed with a C4 Vydac column 1x150mm, 5µm, 300A with a linear gradient of Acetonitrile (2-65% in 20 minutes). Data were acquired using the MassLinx Mass Spectrometry™ software and processed with *MaxEnt* tool for protein mass deconvolution (**Fig. 3.6**). The theoretical molecular weight was calculated with *Protparam* software based on amino acid sequence with all cysteine involved into disulphide bridge (13781.5 Da). Experimental molecular weights (24081.5 and 34998.3 Da) are in agreement with SDS-Page analysis.Dissimilarity between theoretical and experimental values depends on post translational glycosylation of synthesized protein.

Two species in the purified sample suggested different glycosylation isoforms of recombinant human protein. This idea was confirmed by the deglycosylation analysis with PNGase F as described above.

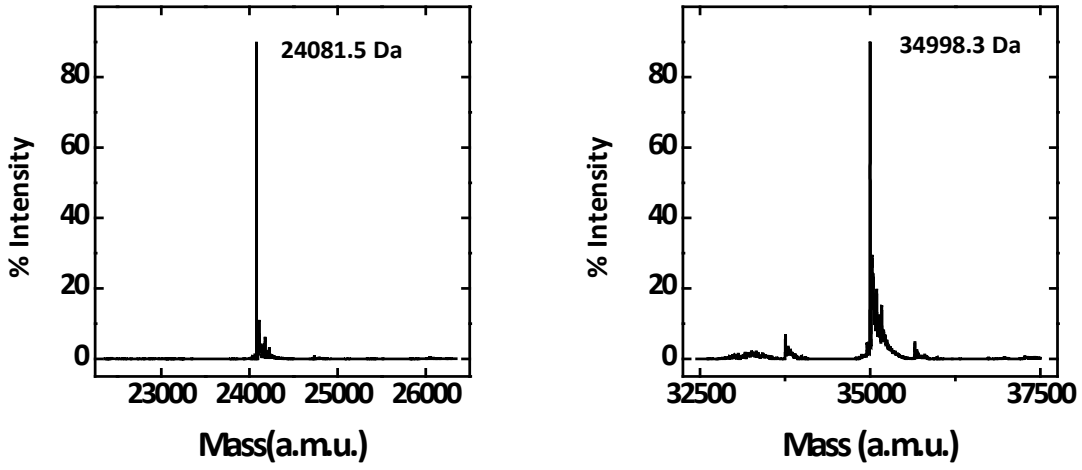


Fig. 3.6 Deconvoluted mass spectra of the two purified species.
Experimental molecular weight (24081.5 ± 0.2 and 34998.3 ± 0.3 Da) of both species is in agreement with data from SDS-Page analysis. Data were acquired using the MassLinx Mass Spectrometry™ software and processed with *MaxEnt* tool for protein mass deconvolution

3.4.1 Deglycosylation

Purified sample was enriched using 4 ml Vivaspin *cutoff* 10 kDa. Subsequently, sample was diluted with buffer solution (50 mM NH_4HCO_3, 0.1% (v/v) Rapigest SF, pH 7.9) and centrifuged. This step was repeated until complete replacement of the buffer. Rapigest SF prevents aggregation improving the enzymatic activity.

It is known that PNGase F is not able to cleave N-linked glycans from glycoproteins when the innermost Nag residue is linked to an α1-3 fucose residue. This modification is most commonly found in glycoproteins expressed in plant but it is not allowed in rhFSHβ due to KDEL retention signal which prevents addition of α1-3 fucose keeping the recombinant protein in endoplasmic reticulum.

Deglycosylation was carried out at 37°C with 1:100 molar ratio enzyme: substrate (0.5 mg/ml). Samples were taken at different times and analyzed by SDS-PAGE .After 48 hours the two glycoproteins (25 and 35 kDa) were quantitatively converted into one glycosylated species with apparent molecular weight of 14 kDa (**Fig. 3.7**) demonstrating that both species are isoforms of the same protein with different glycosylation pattern.

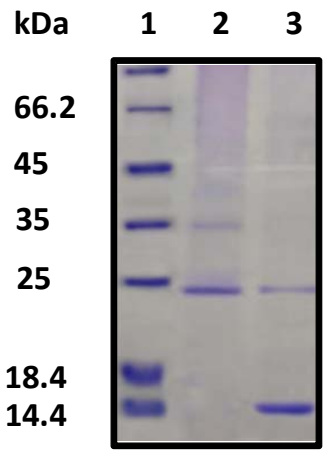


Fig. 3.7 SDS-PAGE analysis of glycosylated and deglycosylated rhFSHβ.
SDS-PAGE analysis (4-12% acrylamide): (lane 1) BlueELf Prestained Protein Marker (Jena Bioscence), (lane 2) 200 ng of rhFSHβ, (lane 3) 1.6 μg of deglycosylated rhFSHβ.

3.4.2 Tryptic digestion

The 14 kDa species was identified with tryptic digestion in situ. Gel was intensively washed in order to discard SDS contaminants which would invalidate enzymatic digestion.

Trypsin is a serine protease that specifically cleaves at the carboxylic group of lysine and arginine residue, this stringent specificity is essential for protein identification. Sequencing grade modified Trypsin has been manufactured to provide maximum specificity and decrease autolysis product that could interfere with database analysis of the mass fragment detected by mass spectrometry. To eliminate the contribution of additional peptides fragments from autolysis, the enzymatic digestion was carried out to a non-spotted gel portion (blank).

Tryptic digest was performed following an established protocol developed in the laboratory of professor Vincenzo De Filippis with minor modification. The reducing condition cause misfolding of the protein making them more susceptible to enzymatic cleavage while iodoacetamide prevents the pairing or disulphide brides by carbamidomethylation of cysteine.

After enzymatic treatment the residue was resuspended in 95% milliQ-water, 3% Acetonitrile and 2% formic acid. LC-MS analysis was performed with a C18 Vydac column 1x150mm, 5µm, 300 Å with a linear gradient of Acetonitrile (3-65% in 25 minutes) at flow rate of 40 µl/min. Mass spectrometry analysis was performed with XEVO G2-S Qtof (**Fig. 3.8**).

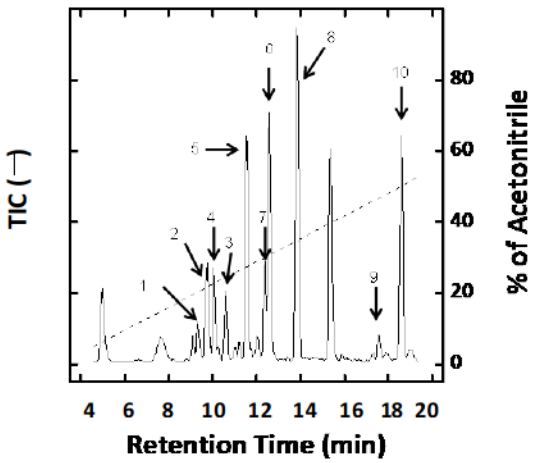


Fig 3.8 LC-MS analysis of digested sample.
Chromatographic analysis of digested protein band at 14 kDa. The residue was resuspended in 95% milliQ-water, 3% Acetonitrile and 2% formic acid. LC-MS analysis was performed with a C18 Vydac column 1x150mm, 5µm, 300A with a linear gradient of Acetonitrile (3-65% in 25 minutes) at flow rate of 40 µl/min. The material eluted was analyzed by mass spectrometry. The peaks were numbered according to the matched peptides reported in the following table.

```
         10         20         30         40         50         60
HHHHHHNSCE LTNITIAIEK EECRFCISIN TTWCAGYCYT RDLVYKDPAR PKIQKTCTFK
         70         80         90         100        110        120
ELVYETVRVP GCAHHADSLY TYPVATQCHC GKCDSDSTDC TVRGLGPSYC SFGEMKEKDE L
```

Peak	Sequence	Region	Modifications	Theoretical M.W. (u.m.a.)	Experimental M.W. (u.m.a.)	ΔPM (u.m.a.)
1	EECR	21-24	CAM (Cys 23)	592.2275	592.2172	0.0103
2	EKDEL	117-121	M.C. (K 118)	632.3017	632.3001	0.0016
3	ELVYETVR	61-68		1007.5287	1007.54	0.0113
4	IQKTCTFK	53-60	CAM (Cys57), M.C. (K 55)	1024.5376	1024.5588	0.0212
5	CDSDSTDCTVR	93-103		1200.4387	1200.4399	0.0012
6	DLVYKDPARPK	42-52	1 M.C. (K46, R50)	1300.7139	1300.7200	0.0061
7	GLGPSYCSFG EMK	104-116	CAM (Cys 110)	1431.6163	1431.6215	0.0182
8	FCISIN(*)TTWC AGYCYTR	25-41	1 CAM (Cys 26, Cys 34, Cys 38), Asn 30 → Asp 30	2058.8799	2058.8788	0.0011
9	HHHHHHNSCE LTN(*)ITIAIEK	1-20	Asn 13 → Asp 13	2371.1240	2371.1356	0.0043
10	VPGCAHHADSLY TYPVATQCHCGK	69-92	3 CAM (Cys 72, Cys 88, Cys 390	2728.1946	2728.1845	0.0101

Tab. 3.2 Detected peptides post tryptic digestion and LC-MS analysis. Complete coverage of the sequence. (*) conversion of N13 and N30 (glycosylation sites) into aspartic acid residues by PNGase F deglycosylation. CAM: carbamidomethylation of cysteine. M.C.: Missed Cleavage.

Total coverage of amino acids sequence confirmed the identity of rhFSHβ (**Tab. 3.2/ Fig. 3.9**). Data were acquired with MassLynx Mass Spectrometry™ software and processed with BiopharmaLynks™ software. Experimental mass values of peptides are in agreement with the theoretical values calculated with BiopharmaLynks™ software.

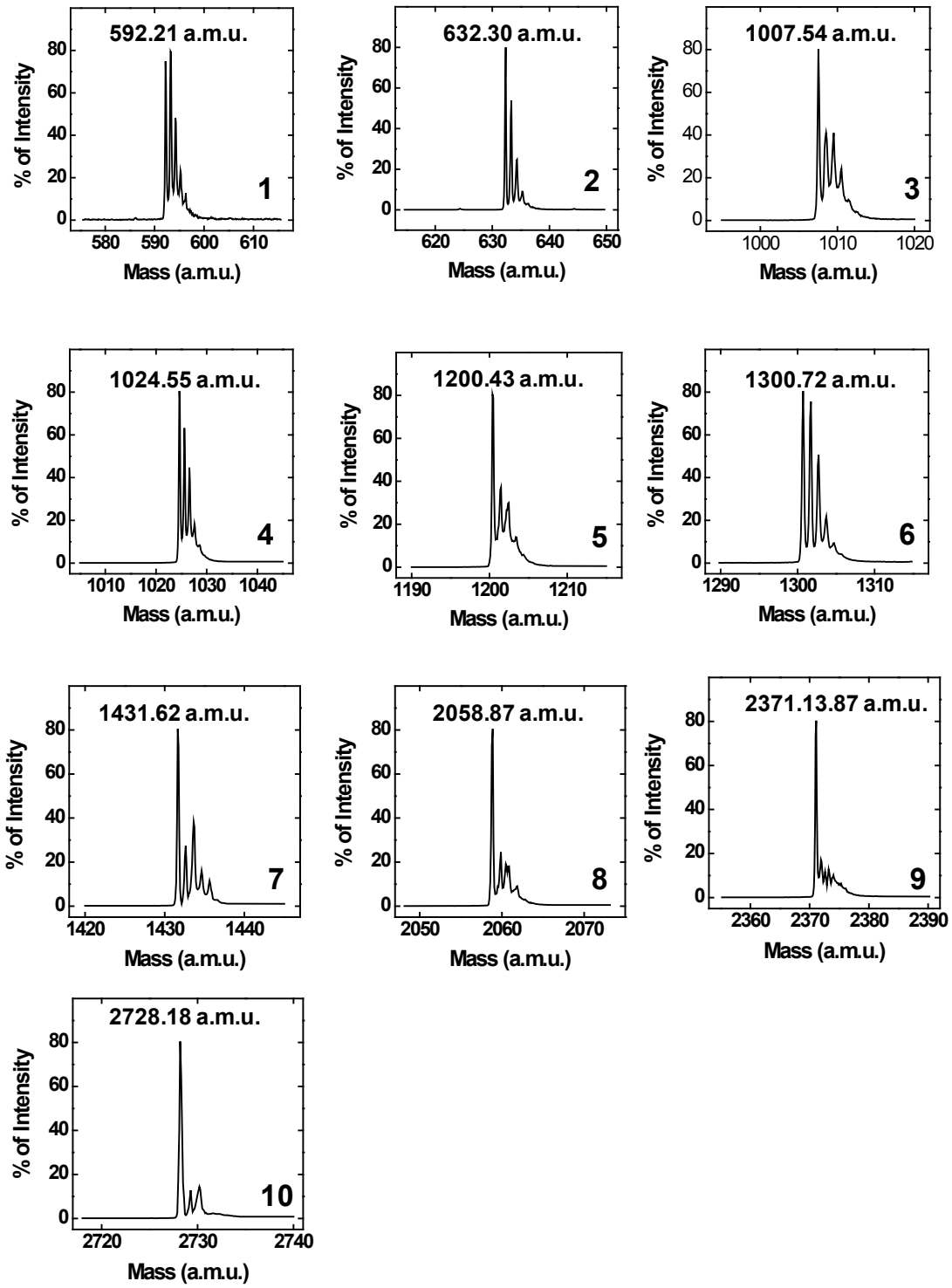


Fig.3.9 Deconvoluted mass spectra of generated peptides by tryptic digestion.

3.5 CONFORMATIONAL CHARACTERIZATION

Disulphide-bond assignment

The chemical assessment of the complete disulphide bridge pattern in rhFSHβ subunit was performed by integrating classical biochemical strategies with mass spectrometric procedures. Six internal disulphide bonds stabilize the structure (Amoresano 2001) (**Fig. 3.10**): Cys9-Cys57, Cys23-Cys72, Cys26-Cys110, Cys34-Cys88, Cys38-Cys90 and Cys93-Cys110 (ABResearch numeration). A subset of the SS bridge pattern comprising Cys9-Cys57, Cys34- Cys88 and Cys38-Cys90 constitutes a cysteine knot motif common in the growth factor superfamily. Cysteine residues are particularly abundant (10 %) and not all interspersed by cleavage site for high specifically protease (e.g. Cys88 and Cys 90).

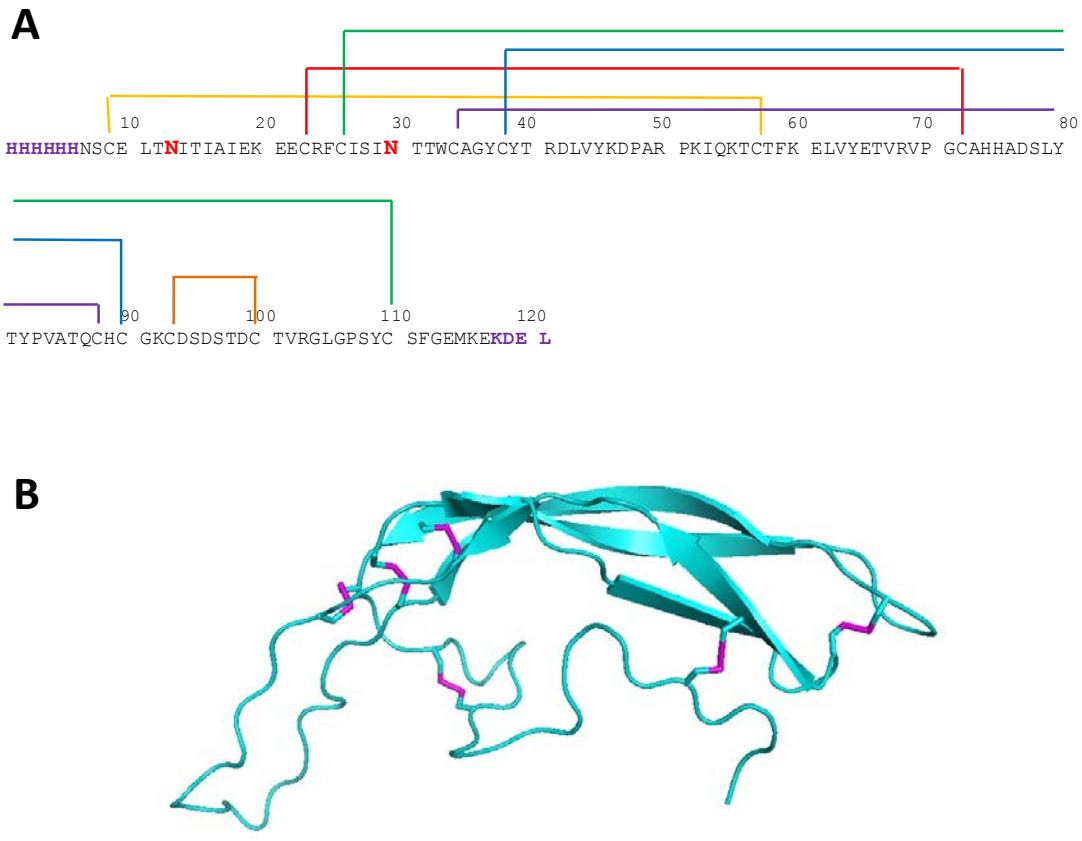


Fig. 3.10 (A) Primary structure and disulphide bonds pairing of recombinant follicle stimulating hormone beta subunit. The extra N- and C-terminal sequences of recombinant protein are colored in violet. The two N-glycosylation sites are colored in red**. (B) Crystal structure of recombinant follicle stimulating hormone beta subunit.** Disulphide bridge pattern is colored in violet (PDB code: **1FL7**).

Multiple digestion procedure allowed to overcome the well known FSHβ subunit resistance to proteases. The protocol was developed referring to Amoresano method (Amoresano 2001) with minor modifications.

Glycosylation represent a physical obstacle for proteases preventing the access to cleavage sites, so the glycosylated sample was analyzed with SDS-Page with non reduction condition in order to preserve disulphide bridge pattern. Protein band at 14 kDa was treated with sequencing grade modified trypsin in non reduction conditions. The generated peptides were further digested with the non specific subtilisin *Carlsberg* enzyme in order to obtain more detectable and suitable peptides.

Lyophilized sample was resuspended in 95% milliQ-water, 3% Acetonitrile and 2% formic acid and analyzed by LC-MS (**Fig. 3.11**). Data were acquired with MassLynx Mass SpectrometryTM software and processed with BiopharmaLynks™ software. Experimental mass values of peptides are in agreement with the theoretical values calculated with BiopharmaLynks™ software (**Fig. 3.12**).

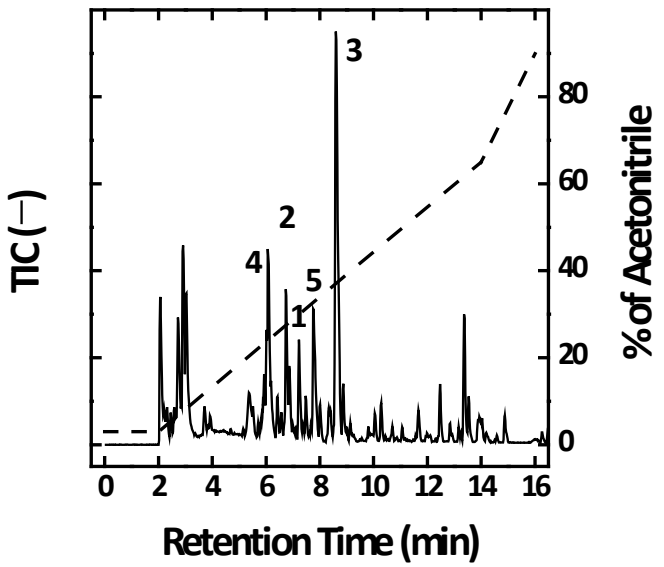
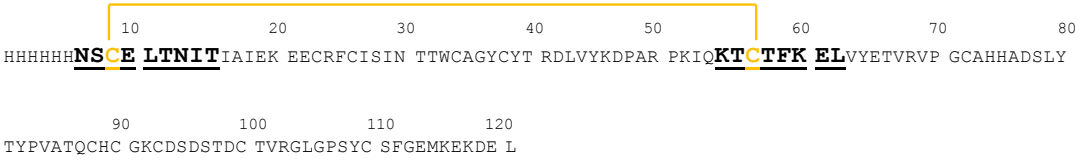


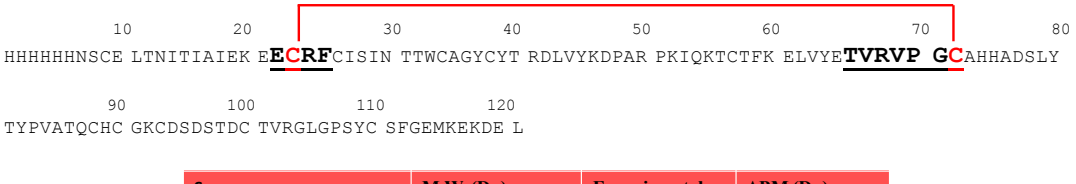
Fig. 3.11 LC-MS analysis of digested rhFSHβ.
After multiple enzymatic treatment with PNGase F, sequencing grade modified Trypsin and non specific subtilisin *Carlsberg,* the residue was resuspended in 95% milliQ-water, 3% Acetonitrile and 2% formic acid and analyzed with Xevo G2-S Q-tof LC-MS analysis was performed with a C18 Vydac column 1x150mm, 5µm, 300A with a linear gradient of Acetonitrile (3-65% in 12 minutes). The peaks were numbered according to the matched peptides reported below.

Cys9-Cys57 (1)



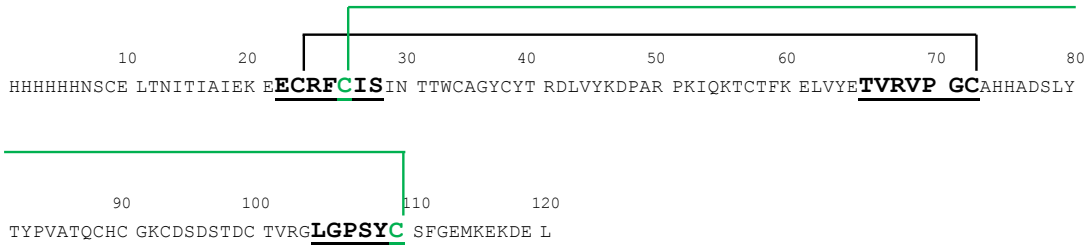
Sequence	M.W. (Da)	Experimental M.W. (Da)	ΔPM (Da)
(7-15)+(55-62)	1962.19	1962.28	0.09

Cys23-Cys72 (2)



Sequence	M.W. (Da)	Experimental M.W. (Da)	ΔPM (Da)
(22-25)+(66-72)	1282.49	1282.52	0.03

Cys26-Cys110 (3)



Sequence	M.W. (Da)	Experimental M.W. (Da)	ΔPM (Da)
(22-28)+(66-72)+(104-110)	2222.58	2222.48	0.1

Cys34-Cys88 (4)

```
          10        20        30        40        50        60        70        80
HHHHHHNSCE LTNITIAIEK EECRFCISIN TTWCAGYCYT RDLVYKDPAR PKIQKTCTFK ELVYETVRVP GCAHHADSLY

          90        100       110       120
TYPVATQCHC GKCDSDSTDC TVRGLGPSYC SFGEMKEKDE L
```

Sequence	M.W. (Da)	Experimental M.W. (Da)	ΔPM (Da)
(33-35)+(86-88)	726.79	726.78	0.01

Cys93-Cys100 (5)

```
          10        20        30        40        50        60        70        80
HHHHHHNSCE LTNITIAIEK EECRFCISIN TTWCAGYCYT RDLVYKDPAR PKIQKTCTFK ELVYETVRVP GCAHHADSLY

          90        100       110       120
TYPVATQCHC GKCDSDSTDC TVRGLGPSYC SFGEMKEKDE L
```

Sequence	M.W. (Da)	Experimental M.W. (Da)	ΔPM (Da)
(93-103)	1199.18	1199.17	0.01

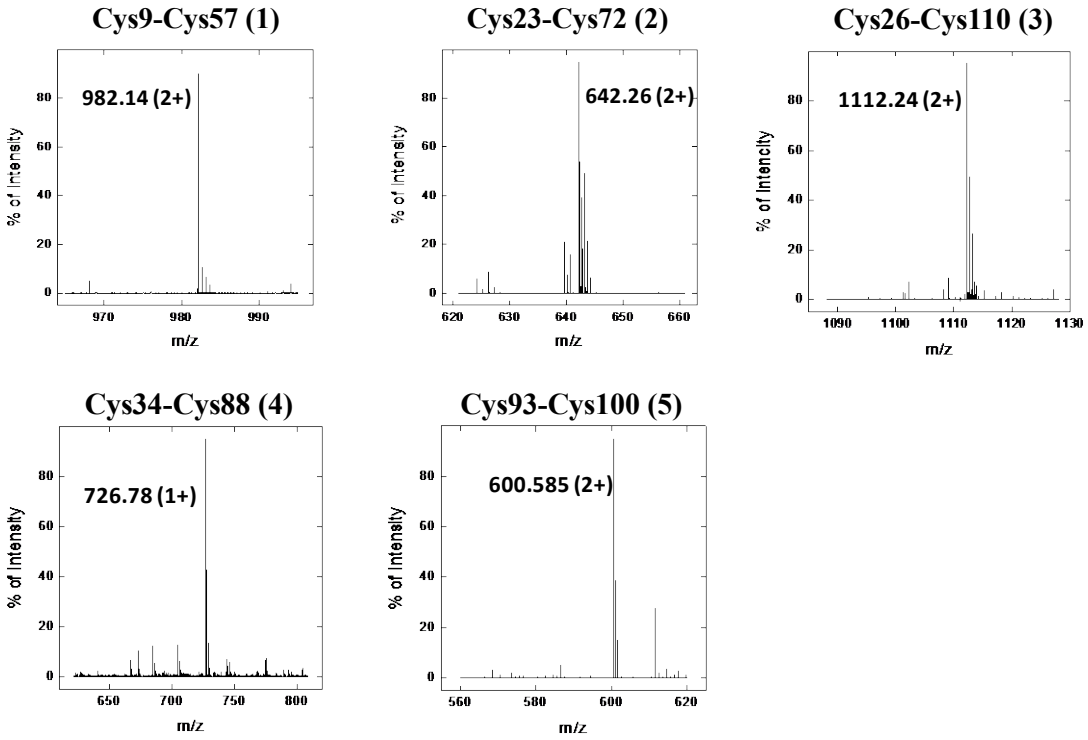


Fig3.12 Disulphide bridge pattern assignment. Underlined peptides were identified with mass spectrometry analysis in order to assign the correspondent disulphide bridge. Experimental values are in agreement with theoretical ones considering the experimental error of 5 ppm. The m/z value of each peptide is reported with the correspondent charge number of ion (z).

In the low mass region two signals were recorded at m/z 642.26 and 600.58, and were assigned to the peptides pair (22-25) + (66-72) and (93-103), the first linked by the disulphide bridge Cys23-Cys72 and the second containing the disulphide bridge Cys93-Cys100. In the medium mass region two signals led to assignment the two disulphide bridges involving Cys9-Cys57 and Cys34-Cys88. Values at m/z 982.14 and 726.78 were interpreted as arising respectively from the peptide (7-15) linked to the fragments (55-62), and peptide (33-25) linked to fragments (86-88).

In the high mass region the signal at m/z 1112.24 was interpreted as arising from the peptide (22-28) linked to the fragments (66-72) and (104-110) respectively, through the SS bridge Cys23-Cys72 and Cys26-Cys110.

The previous determination of Cys23-Cys72 bond allowed to assign the Cys26-Cys110. The last S-S bridge was indirectly assigned to Cys38 and Cys90, since no peptide fragments were found to contain this bond.

After glycosylation step, an aliquot of sample was treated with trichloroacetic acid (TCA) in order to precipitate protein material. The residue was resuspended and analyzed by mass spectrometry to estimate the correct molecular weight (**Fig. 3.13**).

The experimental molecular weight was in agreement with the theoretical one calculated with *Protparam* software based on primary sequence with all cysteine involved into disulphide bridge. The treatment of the same sample with iodoacetamide (1:40 molar ratio) which reacts with thiol groups of reduced cysteine, further confirmed the absence of "free" cysteine with mass spectrometry analysis: no alteration of the molecular weight between rhFSHβ and rhFSHβ incubated with iodoacetamide was detected.

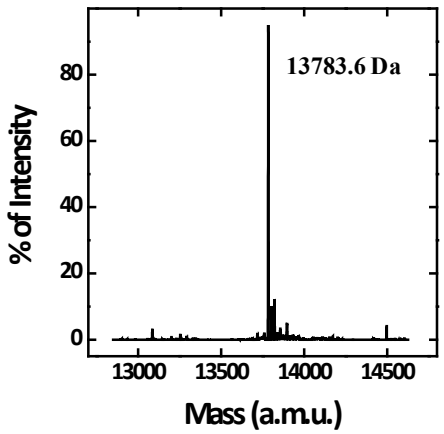


Fig. 3.13Deconvoluted mass spectrum of deglycosylated rhFSHβ. The experimental mass (13783.6 ± 0.1 Da) is in agreement with theoretical value estimated with Protparam software based on amino acid sequence considering all paired cysteine residues (13783.7 Da).

Assignment of five disulphide bridges and the absence of free cysteine residues suggested the pairing between Cys38 and Cys90 and the complete assignment of SS bridges demonstrating the correct conformation of rhFSHβ subunit expressed in *N. benthamiana*.

3.6 BIOLOGICAL EVALUATION

Evaluation of Biological Activity

Follicle stimulating hormone evokes the physiological response as a result of alpha and beta subunit association (Chappel 1998). In all glycoproteins (FSH, LH, GC, TSH) the common alpha subunit is non-covalently associated whit the beta subunit which is structurally unique in its peptide sequence to each member of the family. However, different studies suggest a possible biological activity about the single beta subunit (Fares 1992; Keene 1989). Furthermore, Abcam (Cambridge, UK) provides the single FSH beta subunit (ab191730 Abcam) declaring an equal biological activity to heterodimer.

ABResearch planned to produce both subunits in order to test the single one and the reconstituted dimer.

Waiting for the optimization of alpha monomer production, the rhFSHβ was tested. Test of functional activity has been performed at the Center of Endocrinology and Fertility of the University of Padova. Recombinant human FSHβ was tested on isolated porcine Sertoli cells which express the follicle stimulating hormone receptor (FSHR), monitoring the aromatase activity in according to Menegazzo's protocol (Menegazzo 2011). FSH, in fact, improves aromatization of testosterone into 17β-estradiol and enzyme gene expression (Stocco 2009). Evaluation of biological activity was performed detecting the 17β-estradiol production by the rhFSHβ from ABResearch and the commercial one (Gonal-F). SCs ($20x10^6$ cells) were incubated with 0.8-1-1.2 μg/ml of rhFSHβ protein for three days. After that, the level of 17β-estradiol was evaluated. In parallel the same analysis was performed with addition of testosterone (0.2 mg/ml) 8 hours before the 17β-estradiol detection (**Fig. 3.14/A**).

Administration of 0.8 μg/ml of Gonal-F increased 3 times the amount of estradiol than basal activity. Improving FSH administration, the level of estradiol decreased as

expected by pharmacological overdose. Otherwise, addition of exogenous testosterone increased 4 times the amount of detected species.

Estradiol level after administration of 800 ng/ml of rhFSHβ from ABResearch is comparable with the basal activity suggesting that beta subunit didn't trigger the aromatase activity. The higher level of estradiol in exogenous testosterone treated cells was caused by increased availability of substrate for aromatase. Data show that single subunit is not able to evoke the biological response (**Fig 3.14/B**).

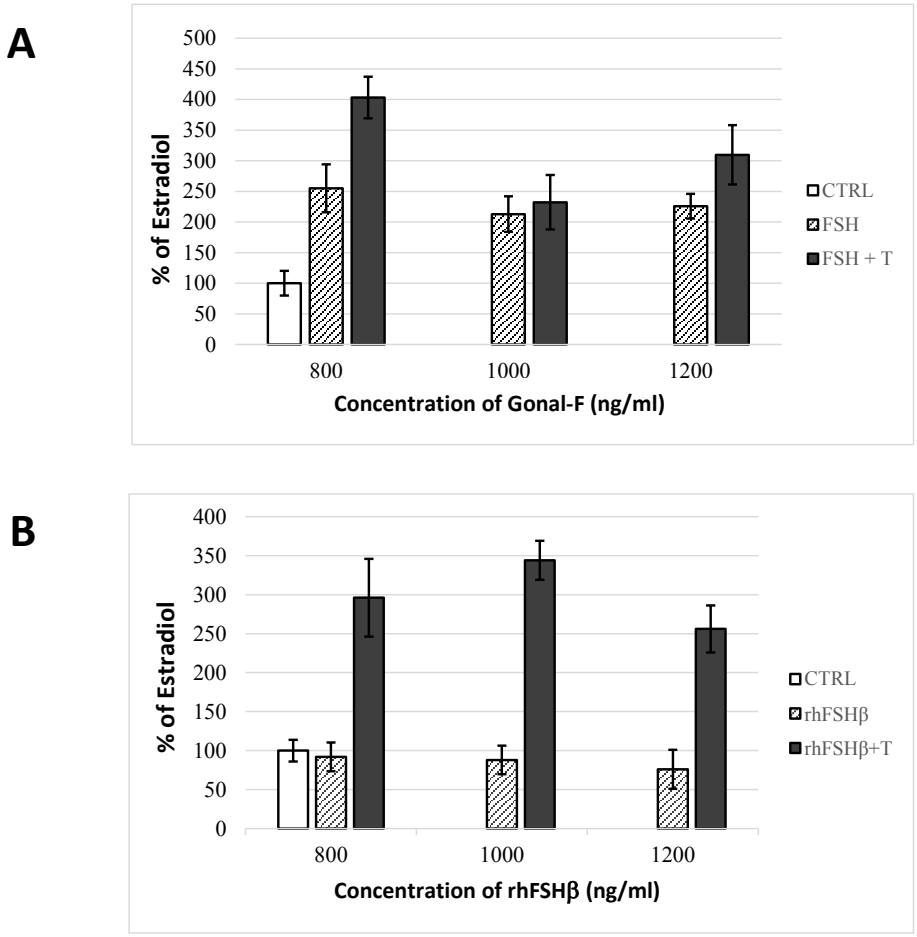


Fig. 3.14 Aromatase activity on Sertoli cells of (A) recombinant human FSH (Gonal-F, Merck Serono), (B) rhFSHβ. The level of estradiol released from no-treated cells (CTRL) corresponds to 100%. FSH: recombinant human FSH (Gonal-F, Merck Serono). rhFSHb: recombinant human FSH beta subunit. +T:testosterone.

In order to evaluate the real potential of this new expression system the heterodimer was associated. Waiting for the recombinant human FSHα subunit from *N. benthamiana*, the heterodimer between rhFSHβ and commercial FSH alpha subunit was tested.

SCs (20x10^6 cells) were incubated with 250-500 ng/ml of heterodimer for three days. After that, the level of 17β-estradiol was evaluated. In parallel the same analysis was performed with addition of testosterone (0.2 mg/ml) 8 hours before the 17β-estradiol detection (**Fig. 3.15**). Notably, the associated heterodimer is able to evoke the biochemical response comparable to the commercial one. This preliminary data are crucial as it show the capability of suspension cell lines of *N. benthamiana* to express active recombinant human proteins.

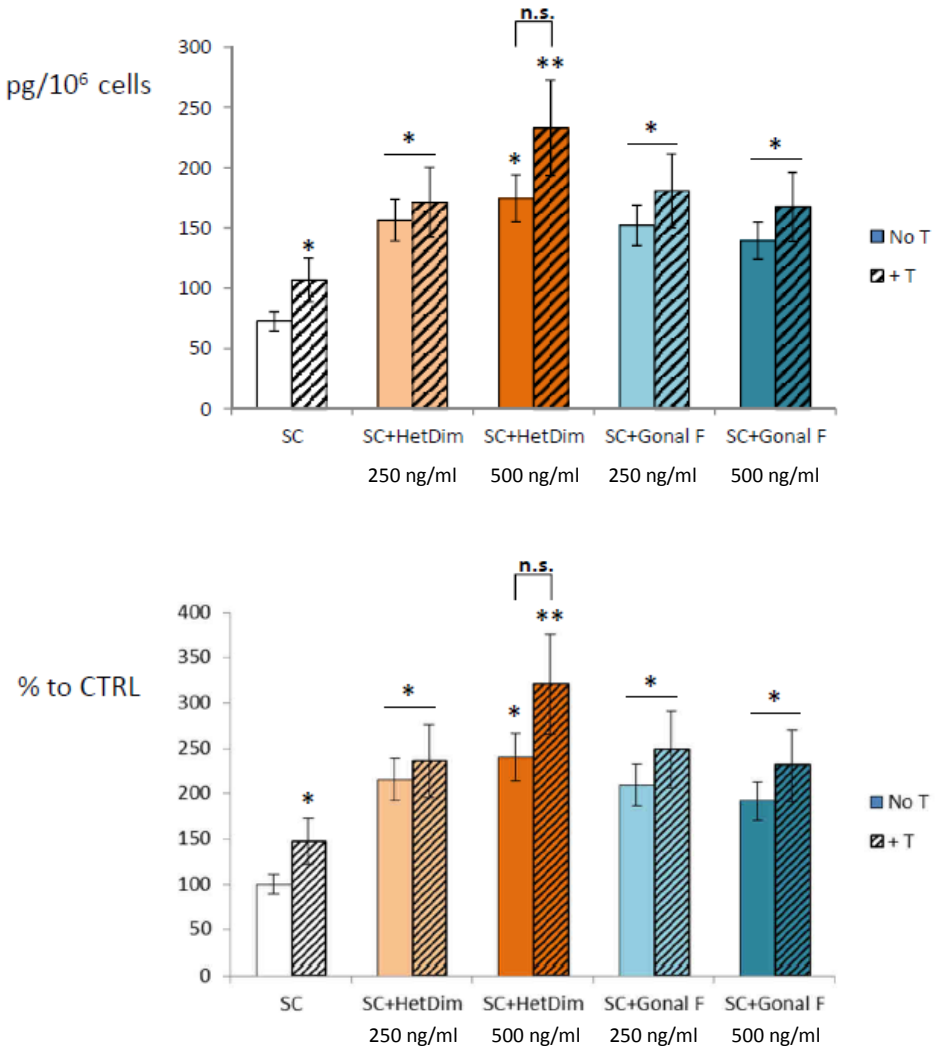


Fig. 3.15 Aromatase activity on Sertoli cells of associated heterodimer (HetDim) and Gonal F.
(**A**) Total amount of released estradiol from no-treated cells (SC), HetDim treated cells and Gonlal F treated cells. (**B**) Percent of estradiol production compared to control cells (100%).
No T: No testosterone administration
T: testosterone administration.

CONCLUSION

Competent cell line of *N. benthamiana* for recombinant human follicle stimulating hormone beta subunit (rhFSHβ) was engineered by ABResearch (Brendola, VI, Italy). Extraction step was developed by combing the KDEL-strategy with ammonium sulphate salting out, saving up to 95% of recombinant protein. The purification step was designed and improved by three consecutive chromatographic processes: IMAC chromatography, size-exclusion chromatography and ion-exchange chromatography. The process has been improved to obtain 4.5/5 mg of purified rhFSHβ for kg of cells, exceeding abundantly average yields in plant expression system.

Two isolated species were obtained and subsequentely chemically and conformationally characterized in order to meet the selective and rigorous demands of regulatory authorities. Enzymatic deglycosylation with PNGase F and tryptic digestion allowed to cover the total amino acid sequence suggesting two different isoforms of the same recombinant protein. The recombinant protein correct folding was confirmed by the assignment of its six disulphide bonds.

Recombinant protein was tested on isolated Sertoli cells from pubertal porcine in order to evaluate the biological activity of the single monomer. Data confirmed that the single beta subunit isn't able to evoke the physiological response as reported in literature. Subsequently, recombinant beta subunit expressed in plant and commercial native human chorionic gonadotropin (alpha subunit) were associated and the heterodimer was tested on Sertoli cells. Notably, preliminary data showed that the biological activity of associated heterodimer is comparable to the commercial follicle stimulating hormone (Gonal-F®).

Experimental evidences support the idea of suspension plant cells as alternative expression system for active therapeutic proteins. In particular, rhFSHβ in association with the alpha human subunit is able to evoke the same biological response of commercial FSH expressed in CHO.

This revolutionary expression platform will be further confirmed by the production of both subunits of follicle stimulating hormone in suspension plant cells and biological evaluation. Production of FSH in plant cells would completely change the world market by proposing an attractive alternative to recombinant FSH expressed in CHO system (Gonal-F®, Puregon®) and highly purified human menopausal gonadotropin (Menopur®, Merional®).

REFERENCES

Altmann F, Staudacher E, Wilson IBH, März L. Insect cells as hosts for the expression of recombinant glycoproteins. Glycoconj J. 1999 Feb 1;16(2):109–23.

Altmann F. The role of protein glycosylation in allergy. Int Arch Allergy Immunol. 2007;142(2):99–115.

Amoresano A, Orrù S, Siciliano RA, De Luca E, Napoleoni R, Sirna A, et al. Assignment of the complete disulphide bridge pattern in the human recombinant follitropin beta-chain. Biol Chem. 2001 Jun;382(6):961–8.

Arlen PA, Falconer R, Cherukumilli S, Cole A, Cole AM, Oishi KK, et al. Field production and functional evaluation of chloroplast-derived interferon-α2b. Plant Biotechnol J. 2007 Jul;5(4):511–25.

Auld DS. (1995) *Removal and replacement of metal ions in metallopeptidases Methods.* Enzymol 248, 228-242. Cleland WW. (1964). *Dithiothreitol, a new protective reagent for SH groups*. *Biochemistry* 3, 480–482.

Bai J-Y, Zeng L, Hu Y-L, Li Y-F, Lin Z-P, Shang S-C, et al. Expression and characteristic of synthetic human epidermal growth factor (hEGF) in transgenic tobacco plants. Biotechnol Lett. 2007 Dec;29(12):2007–12.

Bardor M, Faveeuw C, Fitchette A-C, Gilbert D, Galas L, Trottein F, et al. Immunoreactivity in mammals of two typical plant glyco-epitopes, core α(1,3)-fucose and core xylose. Glycobiology. 2003 Jun 1;13(6):427–34.

Benham CJ, Saleet Jafri M. Disulfide bonding patterns and protein topologies. Protein Science. 1993 Jan 1;2(1):41–54.

Bosch D, Castilho A, Loos A, Schots A, Steinkellner H. N-glycosylation of plant-produced recombinant proteins. Curr Pharm Des. 2013;19(31):5503–12.

Bousfield GR, Butnev VY, White WK, Hall AS, Harvey DJ. Comparison of Follicle-Stimulating Hormone Glycosylation Microheterogenity by Quantitative Negative Mode Nano-Electrospray Mass Spectrometry of Peptide-N Glycanase-Released Oligosaccharides. J Glycomics Lipidomics . 2015 [cited 2016 Nov 13];5(1).

Boyhan D, Daniell H. Low-cost production of proinsulin in tobacco and lettuce chloroplasts for injectable or oral delivery of functional insulin and C-peptide. Plant Biotechnol J. 2011 Jun;9(5):585–98.

Cann JR. Multiple electrophoretic zones arising from protein-buffer interaction. Biochemistry. 1966 Mar;5(3):1108–12.

Chappel S, Buckler D, Kelton C, Tayar NE. Follicle stimulating hormone and its receptor: future perspectives. Hum Reprod. 1998 Jun;13 Suppl 3:18-35-51.

Charlton AJ, Baxter NJ, Khan ML, Moir AJG, Haslam E, Davies AP, et al. Polyphenol/peptide binding and precipitation. J Agric Food Chem. 2002 Mar 13;50(6):1593–601.

da Cunha NB, Vianna GR, da Almeida Lima T, Rech E. Molecular farming of human cytokines and blood products from plants: Challenges in biosynthesis and detection of plant-produced recombinant proteins. Biotechnology Journal. 2014 Jan 1;9(1):39–50.

Dandana A, Ben Khelifa S, Chahed H, Miled A, Ferchichi S. Gaucher Disease: Clinical, Biological and Therapeutic Aspects. Pathobiology. 2015 Nov 21;83(1):13–23.

Davis JS, Kumar TR, May JV, Bousfield GR. Naturally Occurring Follicle-Stimulating Hormone Glycosylation Variants. Journal of Glycomics & Lipidomics [Internet]. 2014 Feb 8 [cited 2016 Nov 13]; Available from: http://www.omicsonline.org/open-access/naturally-occurring-folliclestimulating-hormone-glycosylation-variants-2153-0637.1000e117.php?aid=24866

Demain AL, Vaishnav P. Production of recombinant proteins by microbes and higher organisms. Biotechnol Adv. 2009 Jun;27(3):297–306.

Ds A. Removal and replacement of metal ions in metallopeptidases. Methods Enzymol. 1994 1995;248:228–42.

Düring K, Hippe S, Kreuzaler F, Schell J. Synthesis and self-assembly of a functional monoclonal antibody in transgenic Nicotiana tabacum. Plant Mol Biol. 1990 Aug 1;15(2):281–93.

Fahad S, Khan FA, Pandupuspitasari NS, Ahmed MM, Liao YC, Waheed MT, et al. Recent developments in therapeutic protein expression technologies in plants. Biotechnol Lett. 2015 Feb 1;37(2):265–79.

Fares FA, Suganuma N, Nishimori K, LaPolt PS, Hsueh AJ, Boime I. Design of a long-acting follitropin agonist by fusing the C-terminal sequence of the chorionic gonadotropin beta subunit to the follitropin beta subunit. Proc Natl Acad Sci U S A. 1992 May 15;89(10):4304–8.

Faye L, Boulaflous A, Benchabane M, Gomord V, Michaud D. Protein modifications in the plant secretory pathway: current status and practical implications in molecular pharming. Vaccine. 2005 Mar 7;23(15):1770–8.

Fischer B, Sumner I, Goodenough P. Isolation, renaturation, and formation of disulfide bonds of eukaryotic proteins expressed in Escherichia coli as inclusion bodies. Biotechnol Bioeng. 1993 Jan 5;41(1):3–13.

Fischer R, Drossard J, Commandeur U, Schillberg S, Emans N. Towards molecular farming in the future: moving from diagnostic protein and antibody production in microbes to plants. Biotechnol Appl Biochem. 1999 Oct;30 (Pt 2):101–8.

Fischer R, Emans N. Molecular farming of pharmaceutical proteins. Transgenic Res. 2000 Aug 1;9(4–5):279–99.

Fontana A, Fassina G, Vita C, Dalzoppo D, Zamai M, Zambonin M. Correlation between sites of limited proteolysis and segmental mobility in thermolysin. Biochemistry. 1986 Apr 1;25(8):1847–51.

Friso G, van Wijk KJ. Posttranslational Protein Modifications in Plant Metabolism. Plant Physiol. 2015 Nov;169(3):1469–87.

Fujiwara Y, Aiki Y, Yang L, Takaiwa F, Kosaka A, Tsuji NM, et al. Extraction and purification of human interleukin-10 from transgenic rice seeds. Protein Expr Purif. 2010 Jul;72(1):125–30.

Gellissen G, Janowicz ZA, Weydemann U, Melber K, Strasser AW, Hollenberg CP. High-level expression of foreign genes in Hansenula polymorpha. Biotechnol Adv. 1992;10(2):179–89.

Giddings G. Transgenic plants as protein factories. Curr Opin Biotechnol. 2001 Oct;12(5):450–4.

Goeddel DV, Kleid DG, Bolivar F, Heyneker HL, Yansura DG, Crea R, et al. Expression in Escherichia coli of chemically synthesized genes for human insulin. Proc Natl Acad Sci USA. 1979 Jan;76(1):106–10.

Gomord V, Faye L. Posttranslational modification of therapeutic proteins in plants. Curr Opin Plant Biol. 2004 Apr;7(2):171–81.

Gomord V, Fitchette A-C, Menu-Bouaouiche L, Saint-Jore-Dupas C, Plasson C, Michaud D, et al. Plant-specific glycosylation patterns in the context of therapeutic protein production. Plant Biotechnology Journal. 2010 Jun 1;8(5):564–87.

Griffin TJ, Seth G, Xie H, Bandhakavi S, Hu W-S. Advancing mammalian cell culture engineering using genome-scale technologies. Trends Biotechnol. 2007 Sep;25(9):401–8.

Hefferon K. Plant-derived pharmaceuticals for the developing world. Biotechnol J. 2013 Oct;8(10):1193–202.

Hiatt A, Cafferkey R, Bowdish K. Production of antibodies in transgenic plants. Nature. 1989 Nov 2;342(6245):76–8.

Hirano M, Igarashi A, Suzuki M. Dynamic changes of serum LH and FSH during pregnancy and puerperium. Tohoku J Exp Med. 1976 Mar;118(3):275–82.

Horn ME, Woodard SL, Howard JA. Plant molecular farming: systems and products. Plant Cell Rep. 2004 May;22(10):711–20.

Huang T-K, McDonald KA. Molecular Farming Using Bioreactor-Based Plant Cell Suspension Cultures for Recombinant Protein Production. In: Wang A, Ma S, editors. Molecular Farming in Plants: Recent Advances and Future Prospects [Internet]. Springer Netherlands; 2012 [cited 2016 Nov 13]. p. 37–67. Available from: http://link.springer.com/chapter/10.1007/978-94-007-2217-0_3

James GT. Inactivation of the protease inhibitor phenylmethylsulfonyl fluoride in buffers. Anal Biochem. 1978 Jun 1;86(2):574–9.

Jiang X, Dias JA, He X. Structural biology of glycoprotein hormones and their receptors. Molecular and cellular endocrinology, Molecular and cellular endocrinology. 2014 Jan 25;382(1):424–51.

Jiang X, Fischer D, Chen X, McKenna SD, Liu H, Sriraman V, et al. Evidence for Follicle-stimulating Hormone Receptor as a Functional Trimer. J Biol Chem. 2014 May 16;289(20):14273–82.

Jiang X, Liu H, Chen X, Chen P-H, Fischer D, Sriraman V, et al. Structure of follicle-stimulating hormone in complex with the entire ectodomain of its receptor. Proc Natl Acad Sci U S A. 2012 Jul 31;109(31):12491–6.

Johnson IS. Human insulin from recombinant DNA technology. Science. 1983 Feb 11;219(4585):632–7.

Kauzmann, W., 1959. *Relative Probabilities Of Isomers In Cystine- Containing Randomly Coiled Polypeptides*, in *Advances in Protein Chemistry, pp. 93-108*

Keene JL, Matzuk MM, Otani T, Fauser BC, Galway AB, Hsueh AJ, et al. Expression of biologically active human follitropin in Chinese hamster ovary cells. J Biol Chem. 1989 Mar 25;264(9):4769–75.

Khan KH. Gene Expression in Mammalian Cells and its Applications. Adv Pharm Bull. 2013 Dec;3(2):257–63.

Kwon K-C, Daniell H. Oral Delivery of Protein Drugs Bioencapsulated in Plant Cells. Mol Ther. 2016 Aug;24(8):1342–50.

Kwon K-C, Verma D, Singh ND, Herzog R, Daniell H. Oral delivery of human biopharmaceuticals, autoantigens and vaccine antigens bioencapsulated in plant cells. Adv Drug Deliv Rev. 2013 Jun 15;65(6):782–99.

Laemmli UK. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature. 1970 Aug 15;227(5259):680–5.

Lai T, Yang Y, Ng SK. Advances in Mammalian Cell Line Development Technologies for Recombinant Protein Production. Pharmaceuticals. 2013 Apr 26;6(5):579–603.

Lakshmi PS, Verma D, Yang X, Lloyd B, Daniell H. Low cost tuberculosis vaccine antigens in capsules: expression in chloroplasts, bio-encapsulation, stability and functional evaluation in vitro. PLoS ONE. 2013;8(1):e54708.

Leader B, Baca QJ, Golan DE. Protein therapeutics: a summary and pharmacological classification. Nat Rev Drug Discov. 2008 Jan;7(1):21–39.

Leão R de BF, Esteves SC. Gonadotropin therapy in assisted reproduction: an evolutionary perspective from biologics to biotech. Clinics (Sao Paulo). 2014 Apr;69(4):279–93.

Leckie BM, Neal Stewart C. Agroinfiltration as a technique for rapid assays for evaluating candidate insect resistance transgenes in plants. Plant Cell Rep. 2011 Mar;30(3):325–34.

Lee JS, Park HJ, Kim YH, Lee GM. Protein reference mapping of dihydrofolate reductase-deficient CHO DG44 cell lines using 2-dimensional electrophoresis. Proteomics. 2010 Jun;10(12):2292–302.

Li A, Zhang Q, Chen J, Fei Z, Long C, Li W. Adsorption of Phenolic Compounds on Amberlite XAD-4 and Its Acetylated Derivative MX-4. ResearchGate. 2001 Oct 1;49(3):225–33.

Lombardi R, Circelli P, Villani ME, Buriani G, Nardi L, Coppola V, et al. High-level HIV-1 Nef transient expression in Nicotiana benthamiana using the P19 gene silencing suppressor protein of Artichoke Mottled Crinckle Virus. BMC Biotechnol. 2009 Nov 20;9:96.

Loomis WD, Lile JD, Sandstrom RP, Burbott AJ. Adsorbent polystyrene as an aid in plant enzyme isolation. Phytochemistry. 1979 Jan 1;18(6):1049–54.

Loos A, Steinkellner H. IgG-Fc glycoengineering in non-mammalian expression hosts. Archives of Biochemistry and Biophysics. 2012 Oct 15;526(2):167–73.

Lucht JM. Public Acceptance of Plant Biotechnology and GM Crops. Viruses. 2015 Jul 30;7(8):4254–81.

Ma JK-C, Christou P, Chikwamba R, Haydon H, Paul M, Ferrer MP, et al. Realising the value of plant molecular pharming to benefit the poor in developing countries and emerging economies. Plant Biotechnol J. 2013 Dec;11(9):1029–33.

Ma JK-C, Drake PMW, Christou P. The production of recombinant pharmaceutical proteins in plants. Nat Rev Genet. 2003 Oct;4(10):794–805.

Marillonnet S, Thoeringer C, Kandzia R, Klimyuk V, Gleba Y. Systemic Agrobacterium tumefaciens–mediated transfection of viral replicons for efficient transient expression in plants. Nat Biotech. 2005 Jun;23(6):718–23.

Matsuoka K, Neuhaus J-M. Cis-elements of protein transport to the plant vacuoles. J Exp Bot. 1999 Feb 1;50(331):165–74.

Mattanovich D, Branduardi P, Dato L, Gasser B, Sauer M, Porro D. Recombinant Protein Production in Yeasts. In: Lorence A, editor. Recombinant Gene Expression [Internet]. Humana Press; 2012 [cited 2016 Nov 13]. p. 329–58. (Methods in Molecular Biology). Available from: http://dx.doi.org/10.1007/978-1-61779-433-9_17

Medrano G, Reidy M, Liu J, Ayala J, Dolan M, Cramer C. Rapid System for Evaluating Bioproduction Capacity of Complex Pharmaceutical Proteins in Plants. In: Faye L, Gomord V, editors. Recombinant Proteins From Plants [Internet]. Humana Press; 2009 [cited 2016 Nov 13]. p. 51–67. (Methods in Molecular Biology™). Available from: http://dx.doi.org/10.1007/978-1-59745-407-0_4

Menegazzo M, Zuccarello D, Luca G, Ferlin A, Calvitti M, Mancuso F, et al. Improvements in human sperm quality by long-term in vitro co-culture with isolated porcine Sertoli cells. Hum Reprod. 2011 Oct;26(10):2598–605.

Merlin M, Gecchele E, Capaldi S, Pezzotti M, Avesani L. Comparative evaluation of recombinant protein production in different biofactories: the green perspective. Biomed Res Int. 2014;2014:136419.

Miller LK. Baculoviruses as gene expression vectors. Annu Rev Microbiol. 1988;42:177–99.

Mor TS. Molecular pharming's foot in the FDA's door: Protalix's trailblazing story. Biotechnol Lett. 2015 Nov;37(11):2147–50.

Mullard A. 2014 FDA drug approvals. Nat Rev Drug Discov. 2015 Feb;14(2):77–81.

Obembe OO, Popoola JO, Leelavathi S, Reddy SV. Advances in plant molecular farming. Biotechnol Adv. 2011 Apr;29(2):210–22.

Pazzagli M. (1996) La chemiluminescenza - sviluppo tecnologico ed attuali applicazioni in laboratorio. Edizioni Sorbona.

Pannekoek H, Noordermeer I, van de Putte P. Expression of the cloned uvrB gene of Escherichia coli: mode of transcription and orientation. J Bacteriol. 1979 Jul;139(1):54–63.

Papadimitriou K, Kountourakis P, Kottorou AE, Antonacopoulou AG, Rolfo C, Peeters M, et al. Follicle-Stimulating Hormone Receptor (FSHR): A Promising Tool in Oncology? Mol Diagn Ther. 2016 Dec 1;20(6):523–30.

Pineda C, Hernández GC, Jacobs IA, Alvarez DF, Carini C. Assessing the Immunogenicity of Biopharmaceuticals. BioDrugs. 2016 Jun 1;30(3):195–206.

Pope B, Kent HM. High Efficiency 5 Min Transformation of Escherichia Coli. Nucl Acids Res. 1996 Feb 1;24(3):536–7.

Rader RA. (Re)defining biopharmaceutical. Nat Biotech. 2008 Jul;26(7):743–51.

Radu A, Pichon C, Camparo P, Antoine M, Allory Y, Couvelard A, et al. Expression of Follicle-Stimulating Hormone Receptor in Tumor Blood Vessels. New England Journal of Medicine. 2010 Oct 21;363(17):1621–30.

Ramessar K, Capell T, Christou P. Molecular pharming in cereal crops. Phytochem Rev. 2008 Oct 1;7(3):579–92.

Rasooly R, Hernlem B, Friedman M. Low Levels of Aflatoxin B1, Ricin, and Milk Enhance Recombinant Protein Production in Mammalian Cells. PLOS ONE. 2013 ago;8(8):e71682.

Rech EL, Vianna GR, Aragão FJL. Rech, E.L. , Vianna, G.R. & Aragao, F.J. High-efficiency transformation by biolistics of soybean, common bean and cotton transgenic plants. Nat. Protoc. 3, 410-418. ResearchGate. 2008 Feb 1;3(3):410–8.

Rech EL. Seeds, recombinant DNA and biodiversity. Seed Science Research. 2012 Feb;22(S1):S36–44.

Rosano GL, Ceccarelli EA. Recombinant protein expression in Escherichia coli: advances and challenges. Front Microbiol [Internet]. 2014 Apr 17 [cited 2016 Nov 13];5. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4029002/

Roda, A., Girotti, S., Grigiolo, B. and Ghini, S. (1983) Bioluminescenza: Teoria ed Applicazioni Analitiche. *Giornale italiano di Chimica Clinica*. **8** (4).

Sabalza M, Christou P, Capell T. Recombinant plant-derived pharmaceutical proteins: current technical and economic bottlenecks. Biotechnol Lett. 2014 Dec;36(12):2367–79.

Scholthof HB, Scholthof KB, Jackson AO. Plant virus gene vectors for transient expression of foreign proteins in plants. Annu Rev Phytopathol. 1996;34:299–323.

Shaaltiel Y, Gingis–Velitski S, Tzaban S, Fiks N, Tekoah Y, Aviezer D. Plant-based oral delivery of β-glucocerebrosidase as an enzyme replacement therapy for Gaucher's disease. Plant Biotechnol J. 2015 Oct 1;13(8):1033–40.

Shanmugaraj BM, Ramalingam S. Plants as an Alternate System for the Large Scale Production of Recombinant Therapeutic Proteins. Biochemistry & Analytical Biochemistry [Internet]. 2014 Nov 10 [cited 2016 Nov 13]; Available from: http://www.omicsonline.org/open-access/plants-as-an-alternate-system-for-the-large-scale-production-2161-1009.1000i101.php?aid=33404

Shi X, Jarvis DL. Protein N-Glycosylation in the Baculovirus-Insect Cell System. Curr Drug Targets. 2007 Oct;8(10):1116–25.

Showalter AM. Arabinogalactan-proteins: structure, expression and function. CMLS, Cell Mol Life Sci. 2001 Sep 1;58(10):1399–417.

Simoni M, Gromoll J, Nieschlag E. The follicle-stimulating hormone receptor: biochemistry, molecular biology, physiology, and pathophysiology. Endocr Rev. 1997 Dec;18(6):739–73.

Spadaro ACC, Assis-Pandochi AI, Lucisano-Valim YM, Rothschild Z. Salt fractionation of plasma proteins: A procedure to teach principles of protein chemistry. Biochem Mol Biol Educ. 2003 Jul 1;31(4):249–52.

Stocco C. Aromatase expression in the ovary: hormonal and molecular regulation. Steroids. 2008 May;73(5):473–87.

Stoger E, Fischer R, Moloney M, Ma JK-C. Plant molecular pharming for the treatment of chronic and infectious diseases. Annu Rev Plant Biol. 2014;65:743–68.

Takaiwa F. Update on the use of transgenic rice seeds in oral immunotherapy. Immunotherapy. 2013 Mar;5(3):301–12.

Tekoah Y, Shulman A, Kizhner T, Ruderfer I, Fux L, Nataf Y, et al. Large-scale production of pharmaceutical proteins in plant cell culture—the protalix experience. Plant Biotechnol J. 2015 Oct 1;13(8):1199–208.

Thorpe GH, Kricka LJ. Enhanced chemiluminescent reactions catalyzed by horseradish peroxidase. Meth Enzymol. 1986;133:331–53.

Trew GH, Brown AP, Gillard S, Blackmore S, Clewlow C, O'Donohoe P, et al. In vitro fertilisation with recombinant follicle stimulating hormone requires less IU usage compared with highly purified human menopausal gonadotrophin: results from a European retrospective observational chart review. Reproductive Biology and Endocrinology. 2010;8:137.

Twyman RM, Schillberg S, Fischer R. Transgenic plants in the biopharmaceutical market. Expert Opin Emerg Drugs. 2005 Feb;10(1):185–218.

Twyman RM, Stoger E, Schillberg S, Christou P, Fischer R. Molecular farming in plants: host systems and expression technology. Trends Biotechnol. 2003 Dec;21(12):570–8.

Ulloa-Aguirre A, Timossi C. Structure-function relationship of follicle-stimulating hormone and its receptor. Hum Reprod Update. 1998 May 1;4(3):260–83.

Vermij, P., 2006. USDA approves the first plant-based vaccine. *Nature biotechnology*, **24**(3), p.234.

Walsh G, Jefferis R. Post-translational modifications in the context of therapeutic proteins. Nat Biotechnol. 2006 Oct;24(10):1241–52.

Walsh G. Biopharmaceuticals and biotechnology medicines: an issue of nomenclature. European Journal of Pharmaceutical Sciences. 2002 Mar;15(2):135–8.

Weber K, Osborn M. The Reliability of Molecular Weight Determinations by Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis. J Biol Chem. 1969 Aug 25;244(16):4406–12.

Weise A, Altmann F, Rodriguez-Franco M, Sjoberg ER, Bäumer W, Launhardt H, et al. High-level expression of secreted complex glycosylated recombinant human erythropoietin in the Physcomitrella Delta-fuc-t Delta-xyl-t mutant. Plant Biotechnol J. 2007 May;5(3):389–401.

Wurm FM, Hacker D. First CHO genome. Nat Biotech. 2011 Aug;29(8):718–20.

Xu J, Dolan MC, Medrano G, Cramer CL, Weathers PJ. Green factory: plants as bioproduction platforms for recombinant proteins. Biotechnol Adv. 2012 Oct;30(5):1171–84.

Xu J, Ge X, Dolan MC. Towards high-yield production of pharmaceutical proteins with plant cell suspension cultures. Biotechnol Adv. 2011 Jun;29(3):278–99.

CHAPTER TWO

PRODUCTION AND CHARACTERIZATION OF CHEMICALLY SYNTHESIZED HUMAN OSTEOCALCIN (1-49)

ABBREVIATIONS

AA	Amino acid
Abs/A	Absorbance
ACN	Acetonitrile
CD	Circular Dichroism
Da	Dalton
DCM	Dichloromethane
DIEA	Diisopropylethylamine
DMF	Dimethylformamide
DTT	Dithotreithol
EDT	1,2-Ethanedithiol
ESI	Electrospray Ionization
Fmoc	9-Flurenylmethyloxycabonyl
Gla	Gamma carboxyglutamic acid
Gnd-HCl	Guanidinium Chloride
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5b]pyridinium 3-oxid hexafluorophosphate
HBTU	3-[Bis(dimethylamino)methyliumyl]-3H-benzotriazol-1-oxide hexafluorophosphate
HCTU	[(6-Chloro-1H-benzotriazol-1-yl)oxy](dimethylamino)-N,N-dimethylmethaniminium hexafluorophosphate
HOBt	Hydroxybenzotriazole
ITC	Isothermal Titration Calorimetry
K_d	Dissociation constant
K_a	Association constant
MS	Mass Spectroscopy
NMP	N-Methyl-pyrrolidone
OC	Osteocalcin
OtBu	O-tert- Butyl
PAGE	Poluacrylamide gel electrophoresis

Pbf	2,2,4,6,7-pentemetyldihydrobenzofuran-5-sulfonyl
PEG	Polyethylene glycol
MW	Molecular Weight
Q-TOF	Quadrupole- Time Of Flight
RP-HPLC	Reversed-phase high-performance liquid chromatography
SDS	Sodium dodecyl sulfate
SPPS	Solide Phase Peptide Synthesis
tBoc	tert-Butyloxycarbonyl
tBu	tert-Butyl
TEMED	Tetramethylethylendiamine
TFA	Trifluoride Acetic Acid
TIPS	triisopropylsilyl chloride
TOF	Time of Flight
Tris-HCl	tris-(hydroxymethyl)-aminomethane chloride
Trt	Trytil group
UV	Ultra Violet
A (Ala)	Alanine
C (Cys)	Cysteine
D (Asp)	Aspartic Acid
E (Glu)	Glutamic Acid
F (Phe)	Phenylalanine
G (Gly)	Glycine
H (His)	Histidine
I (Ile)	Isoleucine
K (Lys)	Lysine
L (Leu)	Leucine
M (Met)	Methionine
N (Asn)	Asparagine
P (Pro)	Proline
Q (Asn)	Glutamine

R (Arg)	Arginine
S (Ser)	Serine
T (Thr)	Threonine
V (Val)	Valine
W (Trp)	Tryptophan
Y (Tyr)	Tyrosine

1. INTRODUCTION

1.1 THERAPEUTIC PEPTIDES

More than 7000 natural peptides have been identified. Peptides perform crucial roles in human physiology as growth factors, hormones, ion channel ligands, neurotransmitters, anti-infectives, etc. (Padh 2014, Buchwald 2014, Giordano 2014), recognized for being highly selective and efficacious signaling molecules with attractive pharmacological profile. Peptides represent a new alternative perspective for biopharmaceuticals due to high safety, tolerability and efficacy properties. Furthermore, lower production complexity and lower cost increased interest in peptides in pharmaceutical research and development compared with protein based therapeutics (Fosgerau and Hoffmann 2015). The peptide-based medicine *Lupron*^{TM} from Abbott Laboratories (Chicago, Illinois, USA) for the treatment of prostate cancer achieved global sales of more than US$ 2.3 billion in 2011 (Kaspar 2013). *Lantus*^{TM} from Sanofi (Paris, France) reached sales of US$ 7.9 billion in 2013. The most recent example of a novel peptide drug class is the group of glucagon-like peptide-1 (GLP-1) agonists for the treatment of type 2 diabetes mellitus, which reached total sales of over US$2.6 billion in 2013, with *Victoza*^{TM}, the most prominent member of the class, reaching blockbuster status (Transparency Market Research 2012). Nowadays, there are more than 60 FDA-approved therapeutic peptides on market, 140 in clinical trials and more than 500 in preclinical development including antimicrobial, anticancer activity toward solid and hematologic tumors and immune stimulation (Kaspar 2013). The global peptide drug market has been predicted to increase from US$14.1 billion in 2011 to an estimated US$25.4 billion in 2018 (Transparency Market Research 2012).

Intrinsic weaknesses including short circulating plasma half-life, poor chemical and physical stability prevent direct use of peptides as therapeutics (**Fig. 1.1**). Some of these disadvantages have been successfully resolved with chemical design strategies.

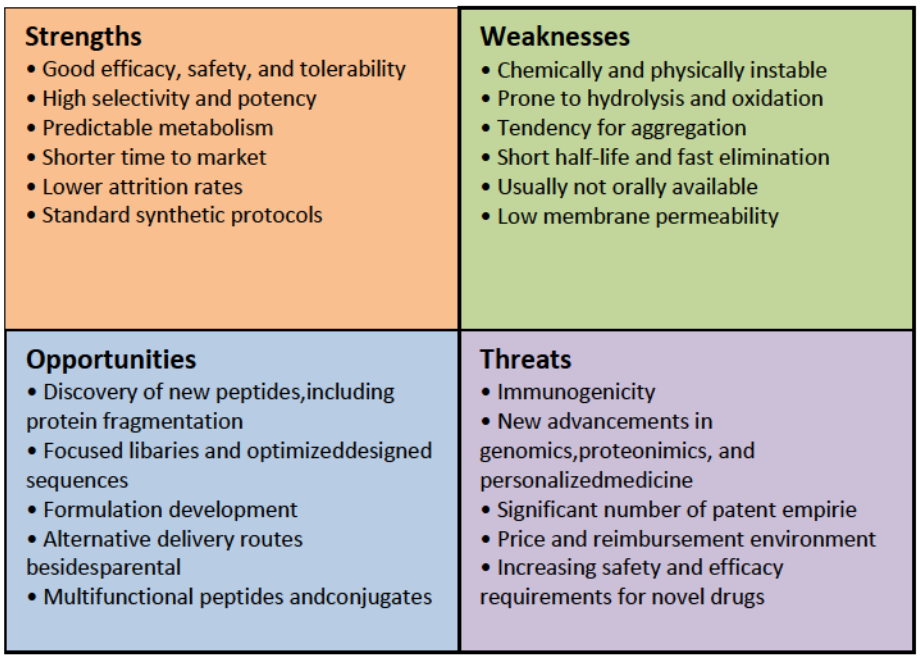


Fig. 1.1 Analysis of the strengths, weaknesses, opportunities, and threats of naturally occurring peptides in their use as therapeutics.

1.2 RATIONAL DESIGN

Rational design allows to improve the physicochemical properties of natural peptides, matching desired pharmacological and pharmacokinetic properties. Analyzing secondary and tertiary crystal structure of the peptide and via input from various analyses, such as alanine substitutions and small focused libraries, the structure-activity relation (SAR) is built in sequential steps that allow to identify essential amino acids and sites for possible substitution or modifications as enzymatically cleavable amino acids and chemically labile sites prone to isomerization, glycosylation, or oxidation (Manning 2010) (**Fig. 2.1**).

Another relevant aspect in rational peptide drug design is to avoid aggregate formation and improve water solubility especially when liquid drug formulations are the desired final product (Hamley 2007). N-methylation, introduction of stabilizing alpha-helixes, salt bridge formations are some of common strategies to avoid aggregation.

Short circulating plasma half-life represents one of the most disadvantages for peptide drug administration. A first approach is to limit enzymatic degradation with identification of cleavable sites followed by substitution or modification of amino acids. A second approach consists to protect peptide drug with enhancement of secondary structure including insertion of lactam bridges, cyclization and stapling or clipping of peptide sequences (Timmerman 2009, Houston 1996, Sim 2012). Nowadays, there are several strategies about binding to the circulating protein albumin as a vehicle to increase plasma half-life in order to decrease the number of administrations. These strategies include peptide acylation (GLP-1 agonist *Victoza*) (Knudsen 2010), insertion of albumin- binding peptide elements or conjugation to albumin-binding antibody fragments (*AlbudAb*^{TM}) (Bao 2013). (PEG)-ylation has been used to limit globular filtration and thereby increasing plasma half-life by limiting the elimination of peptides [19].

An alternative and attractive strategy was developed by Intarcia (Henry 2014) using an implantable device for peptide drugs delivery from a dry reservoir for up to 1 year through an osmotic pump mechanism. This approach allowed to overcome the disadvantages about low plasma half-life and multiple administrations.

More than 75% of peptide drugs are administered by injection, however alternative and more attractive administrations including oral, intranasal, transdermal and gold nanoparticles delivery have been developed (Transparency Market Research 2012). Midasol Therapeutics is currently carrying out clinical development of a transbuccal delivery system that utilizes insulin-passivated gold glyconanoparticles (Sheridan 2012). Another example is the *TopAct*^{TM}

technology platform from ActoGeniX which might enable oral delivery of peptides directly expressed in the gastrointestinal tract (Steidler 2009).

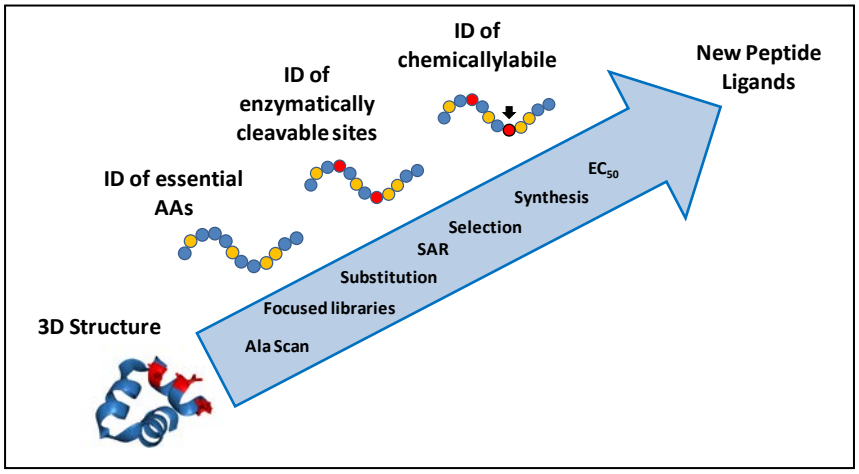


Fig. 2.1 The traditional structure-based design strategies that are used in peptide drug discovery. This includes substitution of amino acids (AA), and the building of structure–activity relations (SAR) via elements such as an alanine (Ala) scan and estimation of EC50 (half maximal effective concentration).

1.3 PEPTIDE SYNTHESIS

1.3.1 Development of solid-phase peptide-synthesis methodology

The global peptide drug market has been predicted to increase from US$14.1 billion in 2011 to an estimated US$25.4 billion in 2018 (Transparency Market Research 2012). In order to meet the large market growth; rapid, efficient and reliable methodology for the chemical synthesis is the utmost interest. The stepwise assembly of peptides consists in a vector elongation via a coupling reaction between amino acids, followed by removal of reversible protecting group. Fischer and Fourneau were the pioneers of peptide synthesis introducing, as well, the term “peptide” (Fischer and Fourneau 1901). Bergmann and Zervas created the first reversible N-alpha protecting group for peptide synthesis, the carbobenzoxy (Cbz) group (Bergmann and Zervas, 1932). DuVigneaud successfully applied early “classical” strategies to construct a peptide with oxytocin-like activity (Vigneaud 1953).

Nowadays weaknesses of classical solution-phase strategy were overcome with the advent of solid-phase techniques. In solid-phase peptide synthesis (SPPS) the peptide elongation occurs by addition of protected amino acids to an insoluble polymeric support (resin). Excess soluble reagents are removed by simple filtration and washing without losses. Once chain has been completed, the crude peptide is released from the support. In 1963 Merrifield introduced a polystyrene-based solid support for peptide synthesis (Merrifield 1963). Peptides are synthesized from the C-terminus to N-terminus, in the opposite direction in cells, using N-alpha protected amino acids. In 1967 SPPS was optimized with the tert-butyloxycarbonyl (Boc) protecting group chemistry for N-alpha protection (Merrifield 1967). Final removal of the peptide from resin occurred with anhydrous hydrogen fluoride (HF) via hydrolytic cleavage (Sakakibara 1967). Boc group was removed with moderate acid as trifluoroacetic acid (TFA), while side-chain protecting group (benzyl or benzyl-based groups) were simultaneously removed with the final peptide removal by HF addition. In 1966 the first automated synthesizer based on Boc SPPs was built by Merrifield Stewart and Jernberg (Merrifield 1966). Boc SPPS was extensively improved in 1980s (Merrifield 1986), it has been utilized for proteins such as interlukin-3 and active enzyme including ribonuclease A.

In 1970 Carpino introduced a new attractive alternative form of SPPS based on 9-fluorenylmethoxycarbonyl (Fmoc) protecting group chemistry for N-alpha protection (Carpino and Han 1970). Contrarily to Boc-strategy, Fmoc-SPPS requires moderate base (20-50% piperidine) for removal and moderate acid condition (TFA) for side-chain deprotection (tbutyl

or tbutyl-based groups) and peptide removal to the resin (hydroxymethylphenoxy-based linkers). Repetitive cycles of base and acid condition allows an "orthogonal" scheme in Fmoc strategy. The milder conditions of Fmoc chemistry allowed greater compatibility with synthesis of peptides that are labile to acid conditions as modification of the indole ring of Trp (Barany 1987). Furthermore, Fmoc SPPS doesn't require special equipment due to its safety conditions. Improvements in these areas allowed for the synthesis of proteins such as bovine pancreatic trypsin inhibitor analogs, ubiquitin, yeast actin-binding protein 539–588, human beta-chorionic gonadotropin 1–74, mini-collagens, HIV-1 Tat protein, HIV-1 nucleocapsid protein NCp7, and active HIV-1 protease (Atherton and Sheppard 1987, Fields 2001, Fields and Noble 1990). Since 1991 Fmoc strategy has become the mostly used form of SPPS: 50% of participating laboratories. By 1994, 98% of participating laboratories were using Fmoc chemistry due to the greater yields of synthesized peptides, minor side reactions during cleavage and high safety process (Angeletti 1997). Currently SPPS development includes addition of non-natural amino acids, post-translationally modified amino acids and labeled amino acids. Modern peptide synthesis is typically limited to peptides consisting of no more than 70 residues, longer lengths and protein can be accessed by using native chemical ligation to couple two or more peptides (described below). Chemical synthesis and semi-synthesis of protein overcome many weaknesses of current protein production. Rapid to effect, easily automated and facilitate purification (80% of recombinant protein cost is about purification step) represent the major advantages of SPPS technique (Casi 2003, Borgia 2000).

1.3.2 The solid support

Peptide synthesis starts with the choice of the solid support, linker (between the solid support and the peptide), appropriately protected amino acid, coupling methodology and protocol for cleaving the peptide from the resin (Fields 1997). Suitable solid support includes physical stability, efficient solvation and no collateral reactions with solvent. Currently there are three primary solid supports: Gel-type (polystyrene, polyacrylamide, polyethylene glycol, polyethylene glycol-based supports), Surface-type and composites supports (Albericio 2000). Solid phase peptide syntheses carry out in the C-terminal to N-terminal direction. For synthesis of C-terminal modified peptides one can take advantage of many linkers that are available (Guillier 2000).

1.3.3Coupling Reagents

Coupling refers to peptide bond formation between two amino acids. Peptide bond consists in condensation with the amino group and the activated carboxyl group by an electron withdrawing group.

The classical examples of in situ coupling reagents are N,N'-dicyclohexylcarbodiimide (DCC) and the related N,N'-diisopropylcarbodiimide (Rieh and Singh 1979). The generality of carbodiimide-mediated couplings is extended significantly by the use of either 1-hydroxybenzotriazole (HOBt) or 1-hydroxy-7-azabenzotriazole (HOAt) as an additive, either of which accelerates carbodiimide-mediated couplings, suppresses racemization, and inhibits dehydration of the carboxamide side chains of Asn and Gln to the corresponding nitriles. Protocols involving benzotriazol-1-yl-oxy-tris (dimethylamino) phosphonium hexafluorophosphate (BOP), benzotriazol-1-yl-oxy-tris(pyrrolidino)phosphonium hexafluorophosphate (PyBOP),7 azabenzotriazol-1-yl-oxytris(pyrrolidino)phosphonium hexafluorophosphate (PyAOP), O-benzotriazol-1-yl-N,N,N',N'-tetramethyluronium hexafluorophosphate (HBTU), O-(7-azabenzotriazol-1-yl)-N,N,N'N' tetramethyluronium hexafluorophosphate (HATU), O-(6-Chlorobenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (HCTU), and O-benzotriazol-1-yl-N,N,N',N'- tetramethyluronium tetrafluoroborate (TBTU) result in coupling kinetics even more rapid than that obtained with carbodiimides.

Other coupling agents that result in low levels of epimerization, and thus are particularly useful for head-to-tail peptide cyclizations and fragment condensations, include O-(3,4-dihydro-4-oxo-1,2,3-benzotriazine-3-yl)-N,N,N',N'-tetramethyluronium tetrafluoroborate (TDBTU) and 3 (diethylphosphoryloxy)-1,2,3- benzotriazin-4(3H)-one (DEPBT). All of the coupling reagents and additives discussed here are commercially available (Stawikowski and Fields 2002).

1.3.4 Synthesis of modified residues and structures

The 20 amino acids that are encoded directly by the codons of the universal genetic code are called natural peptide amino acids. Often they undergo post-translationally modifications including phosphorylation, glycosylation, carboxylation and disulphide bridges. Incorporation of phosphorylated Ser and Thr by SPPS occurs with some care in order to avoid phosphate decomposition by strong acid and lost with base in a beta-elimination process. Boc-Ser(PO_3phenyl$_2$) and Boc-Thr(PO_3phenyl$_2$) are useful derivatives, where HF cleaves the peptide/linker and removes the phenyl groups. Fmoc strategy allows the same addition with Fmoc-Ser(PO_3Bzl,H) and Fmoc-Thr(PO_3Bzl,H) (Perich and Reynolds 1999, Wakamiya 1994). Alternative strategy consists in Fmoc chemistry method to include unprotected Ser and Thr side chains followed by post-assembly phosphorylation (Otvös 1989). Phosphorylated Tyr is less susceptible to chemical decomposition so Fmoc-Tyr(PO_3methyl$_2$) is directly added in SPPS. Fmoc-Tyr(PO_3tBu_2) (Perich and Reynolds 1991), Fmoc-Tyr(PO_3H_2) (Ottinger 1993), and Boc-Tyr(PO_3H_2) (Zardeneta 1990).

Site-specific incorporation of carbohydrates during chemical synthesis of peptides has developed rapidly.

The mild conditions of Fmoc strategy are suitable for glycosylated peptides syntheses compared to strong Boc conditions. Fmoc-Ser, -Thr, -5-hydroxylysine (-Hyl), -4-hydroxyproline (-Hyp), and -Asn have all been incorporated successfully with glycosylated side chains (Cudic and Burstein 2008).

There are different methods to achieve disulfide-bond formation on the SPPS: $K_3Fe(CN)_6$, dithiobis(2 nitrobenzoic acid), diiodoethane oxidation of free sulfhydryls, direct deprotection/oxidation of Cys(acetamidomethyl) residues using thallium trifluoroacetate or I_2, direct conversion of Cys(9 fluorenylmethyl) residues using piperidine and by nucleophilic attack by a free sulfhydryl on either Cys(3-nitro-2-pyridinesulfenyl) or Cys(Scarboxymethylsulfenyl). Direct conversion of Cys(acetamidomethyl) residues by thallium trifluoroacetate is considered the most suitable strategy (Tam 1991).

In order to increase biological activity and specificity, intra chain lactams are formed to conformationally define synthetic peptides. The links occur between side chain of Lys or Orn and Asp or Glu selectively deprotected . Selective deprotection is best achieved by using orthogonal sidechain protection, such as allyloxycarbonyl or 1-(4,4-dimethyl-2,6-dioxocyclohex-1- ylidene)ethyl protection for Lys and allyl or N-[1-(4,4,-dimethyl-2,6-

dioxocyclohexylidene)-3-methyl butyl]aminobenzyl protection for Asp/Glu in combination with an Fmoc/tBu strategy. Lactams can also involved in cyclization reactions (Kates 1994).

The orthogonal strategy of Fmoc/tBu/allyl protecting group is the chosen method for head (N-terminal) to tail (C-terminal) cyclizations. An amide bond is used for side chain attachment of a C-terminal Asp/Glu and the alpha carboxyl group is protected as an allyl ester (Stawikowski and Cudic 2006).

1.3.5 Protein Synthesis

SPPS technique is often used in small protein production. Modern peptide synthesis is tipically limited to peptides consisting of no more than 70 residues, however three different approaches were developed for constructing longer peptides and proteins. First is “stepwise synthesis” in which the entire protein is assembled one amino acid at a time. Second is “fragment assembly” where single peptides are synthesized, purified and subsequently covalently linked to desired protein. The last one is “direct assembly” where single peptide strands are constructed stepwise, purified and non covalently driven to built protein-like structures (Albericio 1997).

If each coupling step were to have 99% yield, a 26 amino acid peptide would be synthesized in 77% final yield (assuming 100% yield in each deprotection step). The advantage of “convergent” protein synthesis by peptide fragment association effects of stepwise synthetic errors and minimized.

An alternative and attractive approach has been developed by which unprotected peptide strands may be linked. “Chemical ligation” consists in an amide bond generated between a peptided bearing a C-terminal thioacid converted to a 5-thio-2-nitrobenzoic acid ester and then reacted with a peptide bearing an N terminal Cys residue (Dawson 1994). The initial thioester ligation product undergoes spontaneous rearrangement, leading to an amide bond and regeneration of the free sulfhydryl on Cys. The method was later refined in order to use unreactive thioester in the ligation reaction (Ayers 1999, Dawson 1997). “Safety-catch” linkers are used with Fmoc strategy to produce the wanted peptide thioester (Shin 1999). Safety-catch linkers anchor the crude peptide to the resin and are stable throughout the synthesis. These linkers then allow the release of a C-terminally modified peptide from the solid support under mild conditions following an additional activation step (Stawikowski and Fields 2002).

1.3.6 Side-reactions

TFA treatment during Fmoc SPPS strategy allows simultaneously deprotection of side-chains groups and peptide/resin removal. Cleavage and liberation of these reactive species can modify susceptible residues such as Met, Tyr and Trp. Modifications can be minimized during TFA cleavage by utilizing effective scavengers. Three efficient cleavage “cocktails” quenching reactive species and preserving amino acid integrity, are (1) TFA-phenolthioanisole-1,2-ethanedithiol-H2O (82.5:5:5:2.5:5) (King 1990), (2) TFA-thioanisole-1,2- ethanedithiol-anisole (90:5:3:2) (Albericio 1990), and (3) TFA-phenol-H2O-triisopropylsilane (88:5:5:2) (Solé and Barany 1992). The use of Boc sidechain protection of Trp also significantly reduces alkylation by Pmc or 2,2,4,6,7- pentamethyldihydro-benzofuran-5-sulfonyl (Pbf) groups.

1.3 HUMAN OSTEOCALCIN (1-49)

Osteocalcin (OC) is a small protein/petide of 46-50 amino acids residues mainly produced and secreted by osteoblasts. It represents one of the most abundant (10-20%) noncollagenous protein in the bone tissue of vertebrates with a highly conserved primary structures (Laizè 2005) (**Fig. 1.3/A**). It has long been known that OC acquires high affinity for calcium ions through the posttranslational modification of glutamate residues (Glu) by carboxylation with a vitamin-K-dependent γ-glutamyl carboxylase (GGCX) (Poser 1979). Carboxylated residues known as Gla residues lead to a conformational change and stabilize the alpha-helical structure and confer a greater affinity for Ca^{2+} and hydroxyapatite (Cristiani 2014). Human OC contains three Gla residues at position 17, 21 and 24 (Hauschka 1989). Circular dichroism analyses indicate that human carboxylated OC in solution is largely unstructured in the absence of calcium. After addition of physiological Ca^{2+} concentrations (1.0 – 1.3 mM) OC undergoes a transition to a folded state formed by three α-helical segment, two β-turns and an αβ sheet stabilized by a single disulphide bridge (Cristiani 2014) (**Fig. 1.3/B**). Ca^{2+} levels are generally associated with high activity and increased bone formation (Garnero 2008). However, within the bone extracellular matrix the glutamic residue at position 17 is γ-carboxylated only in about 9% of human OC (Poser 1980) suggesting a possible decarboxylation process in vivo. Decarboxylation process was confirmed in bone resorption lacunae through acidification of extracellular matrix operated by osteoclast under the control of insulin signaling in osteoblasts (Ferron 2010). Nowadays, different forms of OC (0, 1, 2 or 3 carboxylated glutamate residues) and fragments including N- and C-terminal OC midfragments were identified. The complexity of OCN's status reflects new biological actions in different tissue. OCN is thought to play a vital role in skeletal development and is considered a marker bone formation and remodeling (Kruse 1986, Szulc 1996). However, in vitro experiments and animal models have demonstrated that osteocalcin is a negative regulator of bone mineralization (Ducy 1996) preventing the precipitation of calcium salts from saturated solutions. The chemical modifications, previously described, shift the target tissue of OC from the bone to extraosseous organs where it was demonstrated to regulate male fertility, steroidogenesis, neurotransmitter production and energy metabolism through a new pancreas-bone-testis axis (Ferron 2010). Circulating undercarboxylated OC, which is regulated by insulin, acts in a feed-forward loop to increase beta cell proliferation and consequently insulin production and secretion, while skeletal muscle and adipose tissue respond to OC by

increasing their sensitivity (Lee 2007). OC also acts in the brain to increase neurotransmitter production and in the testes to stimulate testosterone (De Toni 2014).

In the last years, interaction between OC and GPRC6a receptor expressed by adipose, skeletal muscle and the Leyding cells of testes has been extensively studied in order to better understand functions and mechanism of OC and overcome conflicting results obtained from researchers worldwide. New discoveries related to functions will assist its potential clinical application. Nowadays, OC 14-28 fragment in urine was used as bone metabolism marker (Srivastava 2002). OC is considered an important marker of bone activity, and has been used to diagnose disease and evaluate the effects of pharmaceuticals in bone metabolism. Recent studies demonstrated that OC and Sex Hormone Binding Globulin compete on a Specific Binding Site of GPRC6a in testes adding bases for the possible regulation of androgen activity in a non-steroidal manner (De Toni 2016).

Nowadays purified human osteocalcin 1-49 is obtained by extraction from human bones or solid-phase synthesis with the Boc strategy. Human osteocalcin is prepared from extracts of acid-demineralized bone and purified using the ion metal affinity chromatography (ex. *Haematologic Technologies Inc*, Essex Junction, New Hampshire, USA). Due to limited supply of human osteocalcin from bone, the solid-phase synthesis represents the most accessible strategy (ex. *Bachem*, Bubendorf, Switzerland) (Kurihara 1994). Recombinant human OC was produced in *E. Coli* lacking the Gla residues; γ-carboxylation is a specific posttranslational modification of mammalian cells.

A

YLYQWLGAPVPYPDPLEPRREVCELNPDCDELADHIGFQEAYRRFYGPV	Human
YLDHWLGAPAPYPDPLEPKREVCELNPDCDELADHIGFQEAYRRFYGPV	Bovine
YLDHGLGAPAPYPDPLEPRREVCELNPDCDELADHIGFQEAYRRFYGIA	Ovine
YLNNGLGAPAPYPDPLEPHREVCELNPNCDELADHIGFQDAYKRIYGTTV	Rat
YLGASVPSPDPLEPTREQCELNPACDELSDQYGLKTAYKRIYGITI	Mouse

Human vs	***Similarity***	***Identity***
Bovine	93.9%	91.8%
Ovine	87.8%	87.8%
Rat	82%	76%
Mouse	66%	58%

B

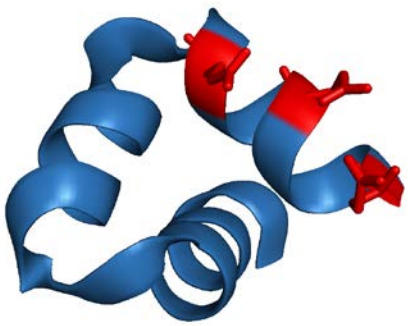


Fig 1.3 Sequence alignment of Osteocalcin from different species. (**A**) Alignment shows the highly conserved primary structures of OC in different mammalians (similarity and identity data are reported in the table). (**B**) Crystal structure of human osteocalcin with Gla residues in red.

2. EXPERIMENTALS

2.1 REAGENTS

H-Val-HMPB-ChemMatrix® (CFL-820-2001) was purchased from PCAS BioMatrix Inc (Quebec, Canada).

Fmoc-L-Tyr(tBU)-OH (FAA1230), Fmoc-L-Asp(tBu)-OH (FAA1020), Fmoc-L-Glu(tBu)-OH (FAA1045), Fmoc-L-Gla(OtBu)2-OH (FAA1368), Fmoc-L-Trp(Boc)-OH (FAA1225), Fmoc-L-Cys(Trt)-OH (FAA1040), Fmoc-L-His(Trt)-OH (FAA1090), Fmoc-L-Gln(Trt)-OH (FAA1043), Fmoc-L-Asn(Trt)-OH (FAA1015), Fmoc-L-Arg(Pbf)-OH (FAA1040), N-methylpirrolidone (SOL-009), piperidine (SOL-010), dymethylformamide (SOL-004) and trifluoroacetic acid (SOL-011) were purchased from Iris Biotech GMBH (Marktredwitz, Germany).

Diisopropylethylamine (387649), Acetonitrile (34998), 1,2-Ethanedithiol (02390), triisopropylsilyl chloride (241725), ammonium acetate (A7330), Ammonium Bicarbonate (09830), Formic Acid (14265), phosphorus pentoxide (214701) and tertbutyl-methylether (306975),), Tris-HCl (857645), Sodium Chloride (S7653), Calcium Chloride (C1016), Deuterium oxide (151882), were purchased from Sigma Aldrich (St. Louis, Missouri, USA).

Sequencing Grade Modified Trypsin (V5111) was purchased from Promega (Fitchburg, Wisconsin, USA).

Osteocalcin (1-49) (human) acetate salt (H-4912) was purchased from Bachem (Bubendorf, Switzerland).

2.2 METHODS

2.2.1 Chemical Synthesis

The synthesis of human Osteocalcin (1-49) was performed by the solid-phase method using the 9-fluorenylmethyloxycarbonyl (Fmoc) strategy on a PS3 automated synthesizer from Protein Technologies International (Tucson, Arizona). The peptide chain was assembled stepwise on a HMPB-ChemMatrix resin derivatized with Fmoc-Val (0.58 mmol/g). Tert-butyl side-chain protecting group was used for Tyr, Asp, Glu and Gla (γ-carboxyglutamic amino acid); tert-butyloxycarbonyl for Trp; triphenylmethyl for Cys, His, Gln and Asn; and 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl group was used for Arg. Removal of N^{α}-Fmoc-protecting groups was achieved by 20% piperidine in N-methylpirrolidone (NMP) treatment for 20 minutes at room temperature. Standard coupling reactions were performed with an equal molar ratio of 2-(1H-benzotriazol-1-yl)-1,1,3,3,tetramethyluronium hexafluorophosphate (HBTU) and 1H-hydroxy-benzotriazole (HOBt) as activating agents, with a fourfold molar excess of the N^{α}-Fmoc-protected amino acids in the presence of diisopropylethylamine (0.35 M) in dymethylformamide (DMF). For double couplings at peptide bonds, the stronger activator 2-(7-aza-1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate was used (HATU). After peptide assembly, the side chain-protected peptidyl resin was treated for 110 min at 10°C with a mixture of TFA/H_2O/ethandithiol/triisopropylsilane (92:3:3:2 v/v/v/v) (**Fig. 2.1**). The resin was removed by filtration, and the acidic solution, containing the unprotected peptide, was precipitated with ice-cold tertbutyl-methylether and then dried. The crude product with Cys residues in the reduced state was analyzed by RP-HPLC with Jasco HPLC Pu-1575 equipped with 1575 UV-Vis detector on a Vydac C18 analytical column (4.6 x 150 mm, 5μm particle size, 300 Å) (Hesperia, California). The column was equilibrated with 0.1% (v/v) aqueous TFA and eluted with a linear 0.078% (w/w) TFA-acetonitrile gradient (20-60% in 30 minutes) at a flow rate of 0.8 ml/min. Absorbance was monitored at 226 nm. The chemical identity of the purified product was established by ESI-TOF mass spectrometry (MS) on a Mariner Workstation (Stafford, Texas).

The purified peptide was subjected to the disulfide bond-forming reaction by air oxidation at room temperature for 72 h in 0.1 M ammonium acetate buffer (pH 7.8) at $2*10^{-4}$ M (Takashi 1994).

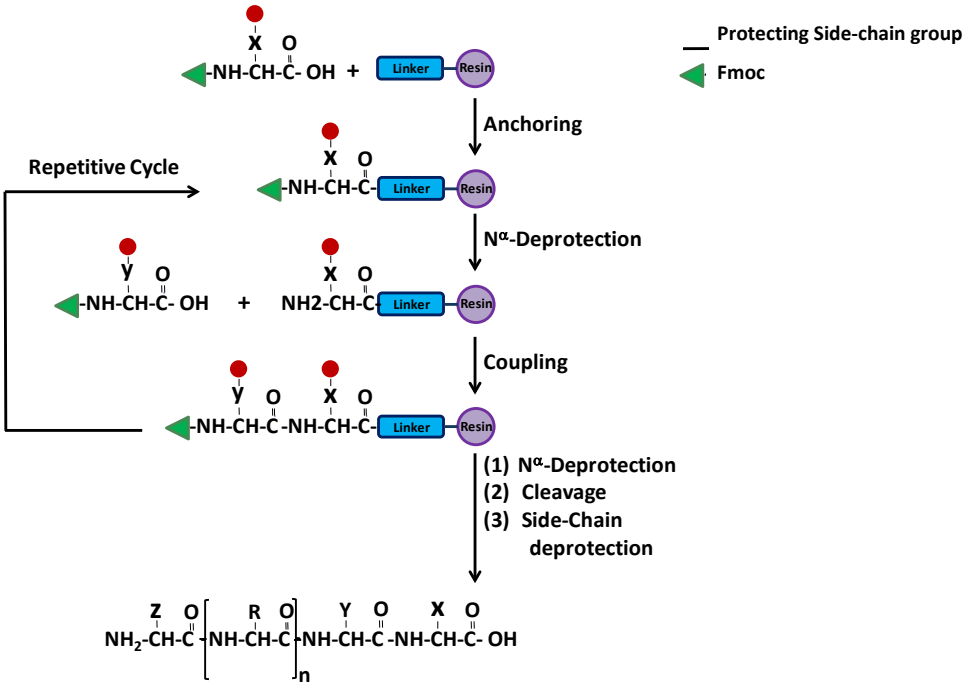


Fig. 2.1 Schematic approach of Fmoc SPPS strategy.

2.2.2 Characterization of the synthetic human osteocalcin 1-49 by tryptic digestion.

The synthetic human osteocalcin 1-49 was purified by reverse-phase HPLC.

SDS-PAGE (4-15%) and Tryptic digestion analysis were performed in *non-reducing* conditions in order to maintain the disulphide bridge in the protein structure. After enzymatic digestion, the residue was resuspended in H2O/TFA 0.1 % and 20 ml 20 µl were injected in a Vydac C18 analytical column (4.6 x 150 mm, 5µm particle size, 300 Å) and eluted with a linear acetonitrile-0.078% TFA gradient. Eluted material was analyzed with ESI-TOF XEVO G2-S Qtof (Milford, Massachusetts, USA). Data were acquired using the MassLinx Mass Spectrometry™ and processed with *BiopharmaLynks™* software (*Waters*, Milford, Massachusetts, USA).

2.2.3 Spectroscopic measurements

UV-Vis spectra were recorded on a V-630 spectrophotometer (Jasco, Tokyo, Japan) at 25 ± 0.1°C, using a 1 cm pathlength cuvette. The concentration of synthesized human Ostaocalcin were determined by measuring the absorbance at 280 nm, using a molar extinction coefficient of 13075 $M^{-1}cm^{-1}$ for the native protein and 12950 $M^{-1}cm^{-1}$ for the reduced form.

Circular dichroism (CD) spectra were recorded on a Jasco J-810 spectropolarimeter equipped with a thermostatted cell holder and a Peltier PTC-423S temperature control system (Jasco, Tokyo, Japan). Spectra were recorded on 0.1 cm pathlength cuvette, 0.1 mg/ml at 25 °C. Each spectrum resulted from the average of four accumulations after base line subtraction. Ellipticity data were expressed as the mean residue ellipticity, $[\theta] = \theta_{obs}*MRW/(10*l*c)$, where θ_{obs} is the measured ellipticity in degrees, MRW is the mean residue weight, l is the cuvette pathlength, and c is the protein concentration in g/ml.

Fluorescence spectra were recorded on a Jasco FP-6500 spectrofluorimeter equipped with a thermostatted cell holder and a Peltier ETC-273T temperature control system (Jasco, Tokyo, Japan) using a 1 cm pathlength cuvette. The samples were excited at 280 nm and the signals were recovery in 290-500 nm range, with an excitation/emission slit of 5/10 nm.

All the measurements were conducted in TBS buffer (20 mM Tris-HClpH 7.4)/ (20 mM Tris-HCl, 0.15 NaCl pH 7.4) at 25 ± 0.1°C.

In order to evaluate the conformational modifications induced by the binding with Ca^{2+} ions, the protein solution was added with $CaCl_2$ solution.

2.2.4 ITC

ITC titrations were performed at 25 ± 0.1°C in 50 mM HEPES pH 7.5, using a MicroCal VP7ITC (Malvern, United Kingdom). The sample solution (63 μM, 1400 μl) was treated by sequentially injections of a $CaCl_2$ stock solution (5.5 mM). After thermal equilibration (60 s), 1 μl of $CaCl_2$ solution was added to the sample, followed by 21 serial injections (5 μl) at intervals of 240 s with continuous stirring (300 r.p.m.).

Each injection generated a heat-burst curve (μcal/s) versus time (min). The area under each peak was determined by integration using Origin 5.0 software (Microcal, Inc.) to give the measure of the heat associated with the injection. The resulting associated temperatures were plotted against the molar ratio. The resulting experimental binding isotherm was corrected for the effect of titrating ligand into binding buffer. The binding affinity and stoichiometry parameters of the binding process were obtained by fitting the integrated heats of binding using the MicroCal ITC Data Analysis software.

2.2.5 H/D exchange

H/D exchange in Osteocalcin was initiated by diluting 4 μl of the stock solution (1 mg/ml, 20 mM ammonium acetate pH 7.8) to 50 μl of the D_2O in absence or in presence of $CaCl_2$ (1.5 mM) at 25 ± 0.1°C. The final concentration of protein was 74 μg/ml. The exchangeable hydrogens in proteins include polar side-chain hydrogens bound to heteroatoms (N, O, and S), the N- and C-terminal hydrogens, and the backbone amide hydrogens.

5 μl of the deuterated protein solution was analyzed by Waters XevoG-2S Q-Tof (Milford, Massachusetts, United States) mass spectrometer at different times in the range 0-15 min. The sample was eluted in D_2O/ACN/HCOOH (50:50:1 v/v/v) at 50 μl/min flow rate.

The extent of H/D exchange was calculated using the equation $D = H^*(m - m_{0\%})/(m_{100\%} - m_{0\%})$ where m is the mass of protein at *t* time; $m_{0\%}$ is the measured mass of a "zero-deuteration" control; $m_{100\%}$ is the measured mass of a "full-deuteration" control.

3. RESULTS & DISCUSSION

3.1 CHEMICAL SYNTHESIS AND CHARACTERIZATION

3.1.1 Chemical synthesis strategy

Since 1994 Fmoc chemistry has become the primary choice in SPPS due to the greater yields of synthesized peptides, minor side reactions during cleavage and high safety process (Angeletti 1997). Classic polystyrene resin was replaced by ChemMatrix resin due to its swelling capacity in different organic solvent. Resin solvation prevents shrinking episodes caused by hydrophobic interaction between adjacent growing peptide chains which compromise peptide elongation. Furthermore the 4-(4-Hydroxymethyl-3-methoxyphenoxy) butyric acid (HMPB) linker is used as a hyperacid sensitive linker in solid-phase synthesis.

Deprotection solution (20% piperidine in NMP) and activation solution (0.35 M diisopropylethylamine in DMF) were freshly prepared to overcome collateral reaction caused by ammine degradation. Anhydrous solvent are strictly required in order to improve production yields, the presence of water molecules, in fact, would hydrolyze the amidic bond.

Solid-phase synthesis of human Osteocalcin was carried out on a PS3 automated synthesizer from Protein Technologies International using the 0.1 mmol scale protocol. Each amino acid was weighed (0.4 mmol) and placed in a vial. HOBt/HBTU mixture (1:1) or HATU were added to the correspondent vials and dried with phosphorous pentoxide. HATU is strictly used for difficult couplings in order to increase the nucleophilic properties of carboxylic carbon in incoming amino acid. Restricted use is recommended to minimize racemization reaction.

The peptide chain was assembled stepwise on a HMPB-ChemMatrix resin derivatized with Fmoc-$C_{terminal}$ amino acid. Solid phase peptide synthesis carries out in the C-terminal to N-terminal direction.

Synthesis was planned considering the aggregation potential (AP) of each amino acid. AP refers to hydrophobic proprieties of each amino acid relating to its position in the sequence. It was estimated with the *Peptide Companion* software (CoshiSoft /PeptiSearch, San Diego, CA, USA): easy, intermediate or difficult coupling are respectively marked in green, yellow and red (**Fig. 3.1**).Moreover, synthesis was designed considering steric hindrance and the effects of

side-chain protecting groups in each coupling reaction in order to identify the critical control point of peptide elongation.

Human OC (1-49) was assembled stepwise, including four check breaks to identify and optimize critical control points in peptide elongation (**Fig. 3.1**). Analytical checks were carried out by chemical characterization of step by step synthesized peptides. Intermediate peptides (Gly37-Val49, Leu25–Val49, Pro18–Val49 and Pro11–Val49) were analyzed by reverse-phase HPLC and mass spectrometry (**Fig. 3.2/1-4**). Upon optimization, peptide elongation was carried out to next check point.

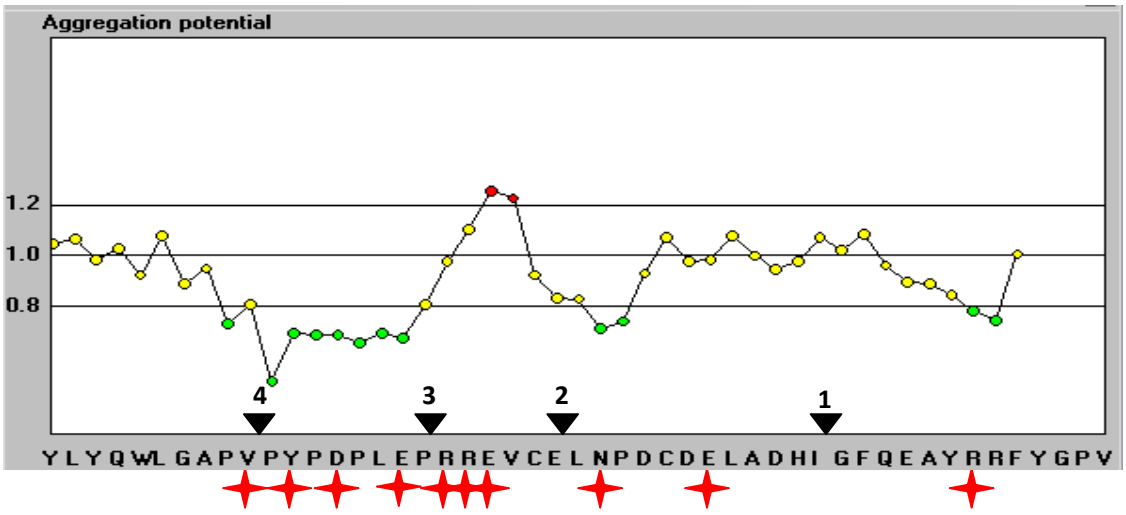


Fig. 3.1 Aggregation potential and SPPS strategy. Peptide synthesis was planned considering the aggregation potential of each amino acid in the primary sequence. Easy, intermediate or difficult coupling are respectively marked in green, yellow and red. (▼) Check point in peptide elongation. (✦) Incoming activated amino acids with HATU.

After peptide assembly, the side chain-protected peptidyl resin was treated for 110 min at 10°C with a mixture of TFA/H_2O/ethandithiol/triisopropylsilane (92:3:3:2 v/v/v/v). TFA treatment allows simultaneously deprotection of side-chains groups and peptide/resin removal. Cleavage of these reactive species can modify susceptible residues such as Met, Tyr and Trp. Modifications were minimized during TFA treatment by suitable scavengers.

During TFA cleavage, decarboxylation of Gla residues was drastically minimized by working at 10°C. Decarboxylation of γ-carboxyglutamic amino acid, in fact, occurs in acidic condition at high temperature (De Toni 2014).

The resin was removed by filtration, and the acidic solution, containing the unprotected peptide, was precipitated with ice-cold tertbutyl-methylether and then dried.

Synthesized human OC 1-49 was resuspended in H_2O/TFA 0.1% and analyzed by reverse-phase HPLC and mass spectrometry (**Fig. 3.2/5**).

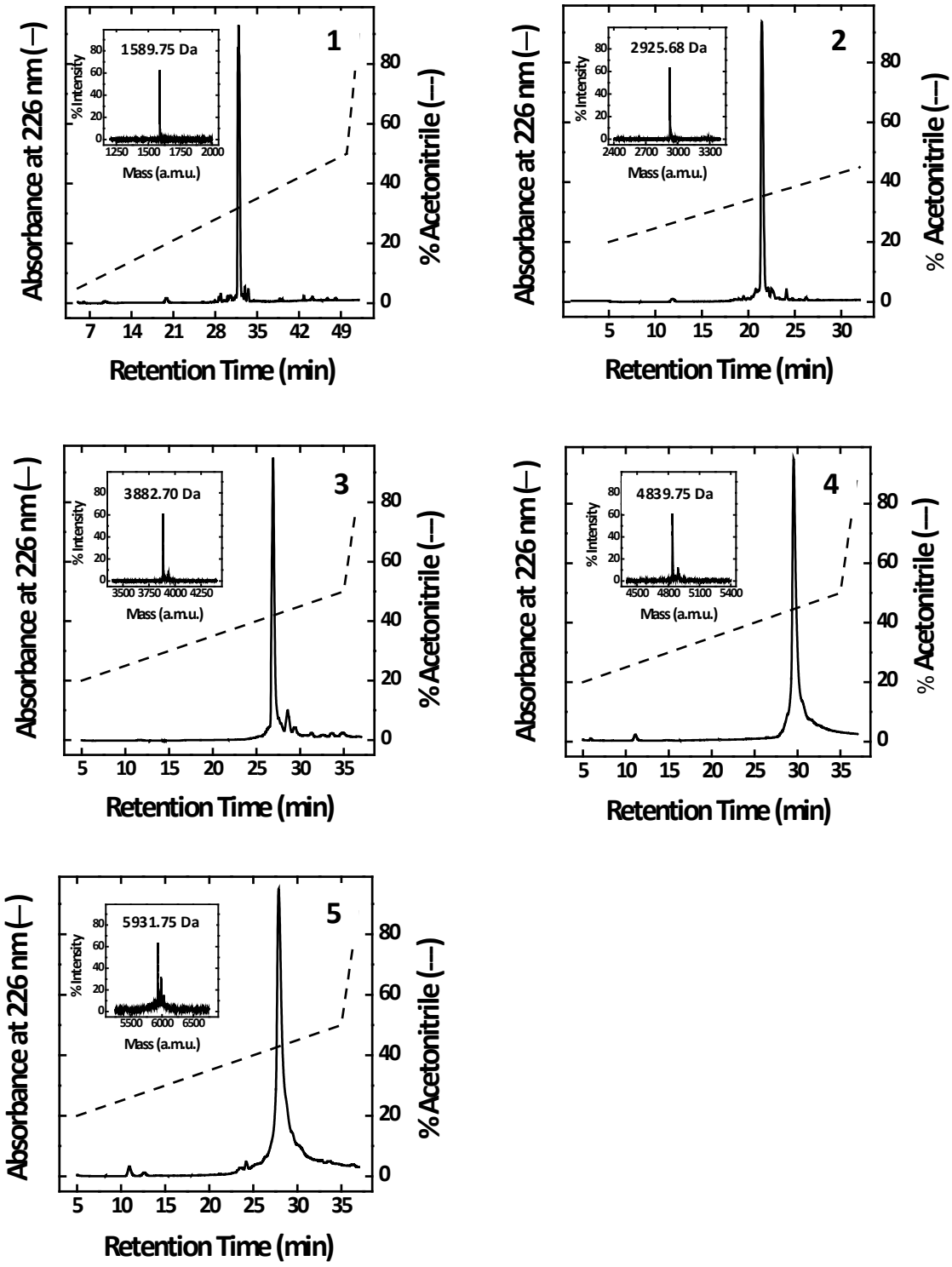


Fig. 3.2 RP-HPLC analysys of chemically synthesized human Osteocalcin 1-49.
Human OC (1-49) was assembled stepwise. Intermediate fragments: (**1**) Gly37-Val49, (**2**) Leu25–Val49, (**3**) Pro18–Val49, (**4**) Pro11–Val49 and total protein (**5**) Tyr1–Val49 were analyzed by reverse-phase HPLC. Peptides and total protein were eluted with an acetonitrile-0.078% TFA gradient (---): (**1**) 5-50% in 45 minutes, (**2**,**3**,**4** and **5**) 20-50% in 30 minutes. (**Inset**) Deconvoluted ESI-TOF mass spectrum of RP-HPLC purified peptides. The estimated molecular masses are in agreement with the theoretical values.

Crude human Osteocalcin was purified and quantified with UV-Vis spectra analysis. Coefficient molar extinction was estimated with online tool *ProtParam* (12950 $M^{-1}cm^{-1}$).

Reduced human OC was air-oxidized to generate the intramolecular disulphide formation in 0.1 ammonium acetate buffer (pH 7.8) at $2*10^{-4}$ M at room temperature for 72 h (Kakashi 1994). Mixture reaction was analyzed by RP-HPLC (**Fig. 3.3**).

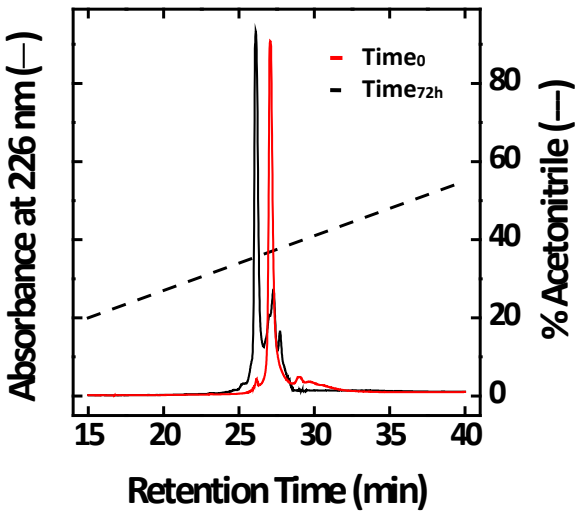


Fig. 3.3 RP-HPLC analysys of oxidative reaction of synthesized human Osteocalcin 1-49.
Samples were taken at different time from oxidative reaction and analyzed with RP-HPLC. 20 µl were injected in a Vydac C18 analytical column (4.6 x 150 mm, 5µm particle size, 300 Å) and eluted with a linear acetonitrile-0.078% TFA gradient (---) (20-60% in 35 minutes). The absorbance was monitored at 226 nm. Flow rate 0.8 ml/min.

As shown in figure 3.3, reduced OC 1-49 ($Time_0$) peak decreased in 72 hours, whilst a new peak, eluted at minor retention time, appeared. The peak eluted was purified and characterized by mass spectrometry (**Fig. 3.4**).

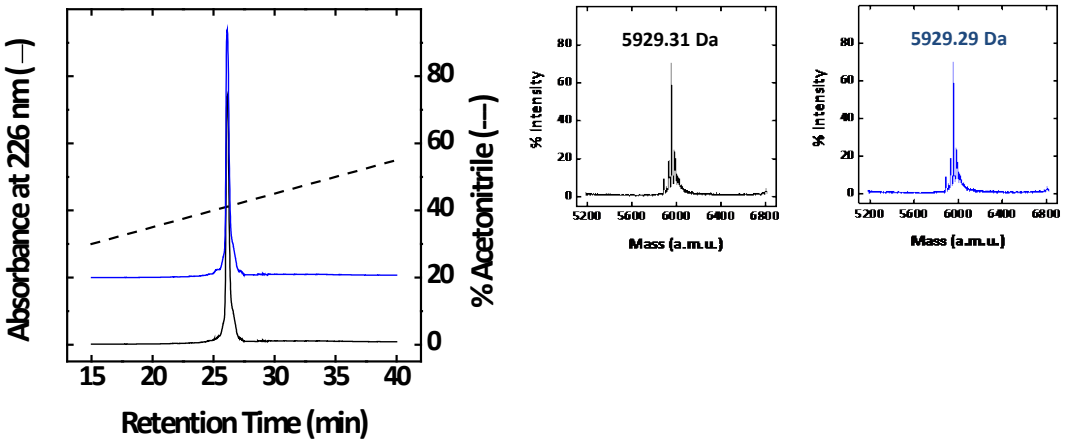


Fig. 3.4 RP-HPLC and MS analysis of purified commercial OC 1-49 and synthesized OC after oxidative reaction. Commercial human Osteocalcin 1-49 (▬) and synthesized OC after oxidative reaction (▬) were analyzed by RP-HPLC. Purified species were eluted with an acetonitrile-0.078% TFA gradient (---) (20-60% in 35 minutes) in a Vydac C18 analytical column (4.6 x 150 mm, 5µm particle size, 300 Å). Species eluted were analyzed by mass spectrometry.

Retention time and the molecular weight compared to commercial OC, confirmed the formation of disulphide bond in synthesized OC 1-49 ($M.W._{OX}$: 5929.31 ± 0.2 Da, $M.W._{RED}$: 5931.75 ± 0.2 Da).

Oxidation ration was determined up to 88% by the UV-Vis spectra analysis (ε_{ox}: 15075 $M^{-1}cm^{1}$) quantification.

3.1.2 Tryptic digestion

The synthetic human Osteocalcin 1-49 was further characterized by tryptic digestion. Tryptic digestion generated three major fragments, 1-19, 20-43 and 45-49 of the peptide, which were separated on the RP-HPLC and identified by mass spectrometry (**Fig. 3.5**). Minor peptide was identified as decarboxylation product (Gla to Glu) or autolysis fragment from trypsis. Data confirmed formation of disulphide bond in synthesized OC 1-49.

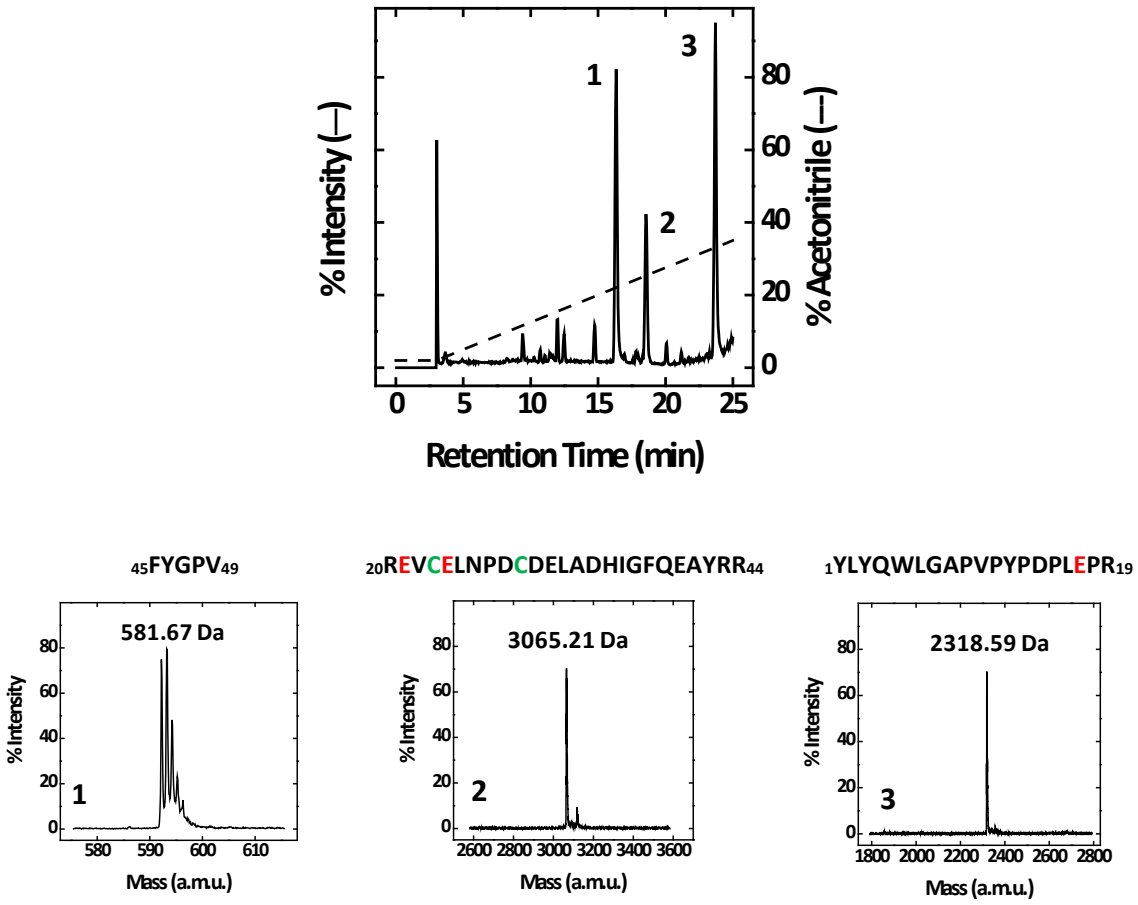


Fig. 3.5. Elution profile of tryptic digest of synthesized human Osteocalcin. HPLC analysis was performed with Jasco HPLC Pu-1575 equipped with 1575 UV-Vis detector on a Vydac C18 analytical column (4.6 x 150 mm, 5μm particle size, 300 Å). Absorbance was monitored at 226 nm with flow rate of 0.8 ml/min. Fractions species were eluted with an acetonitrile-0.078% TFA gradient (---) (5-40% in 25 minutes). Major peaks correspond to 1-19 (3), 20-43 (2) and 45-49 (1) fragments. Gla residues are colored in red while cysteine residues involved in disulphide bridge in green.

3.2 CONFORMATIONAL CHARACTERIZATION

Synthesized human Osteocalcin (1-49) was conformationally characterized by spectroscopic measurements and Hydrogen/deuterium exchange by mass spectrometry.

3.2.1 Far-UV Circular Dichroism

The conformational properties of synthesized human Osteocalcin were investigated by CD spectroscopy in the far UV region which is a sensitive probe of protein secondary structure. Spectra were acquired at increasing Ca^{2+} concentrations (0-25.6 mM) with and without 0.15 M NaCl. In the absence of calcium both spectra shared a similar shape with a minimum at 204-208 nm and a shallow band at 220 nm: in both conditions the polypeptide chain was rather flexible with crude helical structure.

Ca^{2+} induces alpha-helical structure in Osteocalcin and Gla residues are required for this transformation. At saturating [Ca^{2+}], both the shape and intensity of spectra were changed more closely to high helical secondary structure, with minima at 208 and 222 nm.

However, under physiological conditions (0.15 M NaCl) the amount of alpha-helix decreases by 8-10% suggesting the competition among metal ions Ca^{2+} and $Na^{1+}$and the major contribution of more flexible and disordered Gla-OC in physiological scenario (**Fig. 3.6**).

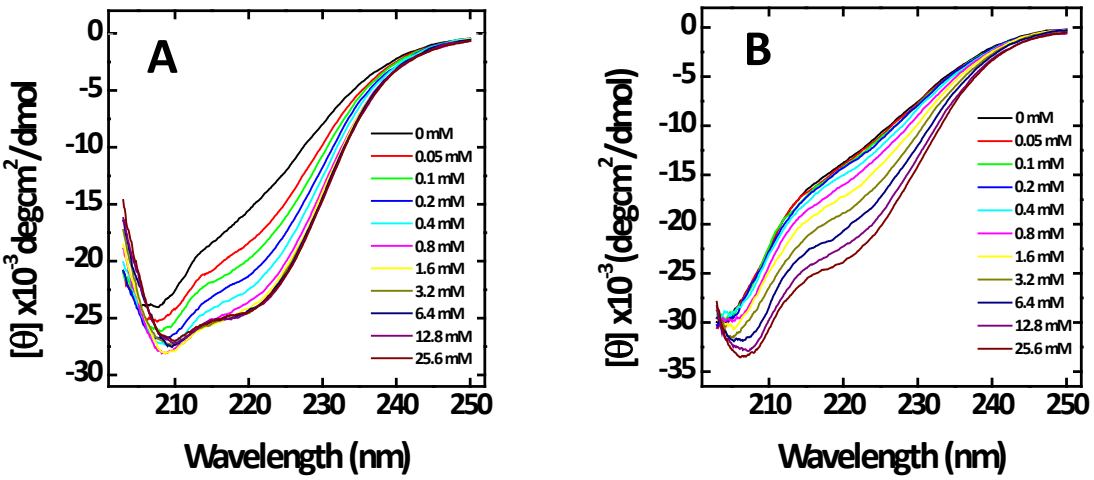


Fig. 3.6 Far-UV CD spetra of synthesized human Osteocalcin (1-49). Spectra were recorded on 0.1 cm pathlength cuvette, 0.1 mg/ml at 25 °C in (**A**) 20 mM Tris-HCl, pH 7.4 and (**B**) 20 mM Tris-HCl, 0.15 M NaCl.

3.2.2 Fluorescence

Intrinsic Trp fluorescence in Osteocalcin is decreased by Ca^{2+}. The unique tryptophan residue was excited at 295 nm and fluoresce emission was monitored in the 290-500 nm range. Spectra were acquired at increasing Ca^{2+} concentrations (0-25.6 mM) with or without 0.15 M NaCl (**Fig. 3.7**). At saturating [Ca^{2+}], fluorescence intensity decreased of 23%. As in all interactions of Osteocalcin with metal cations, decreased ionic strength increases the affinity of the protein for Ca^{2+} (Hauschka 1982). The midpoint for completion of fluorescence decrease is 0.28 mM (20 mM Tris-HCl, pH 7.4), while at high ionic strength 0.63 mM (20 mM Tris-HCl, 0.15 M NaCl, pH 7.4). Interpretation of these fluorescence changes in terms of polarity of Trp environment is not warranted by the available data, yet the conformational transition clearly affects Trp residue (**Fig 3.7**).

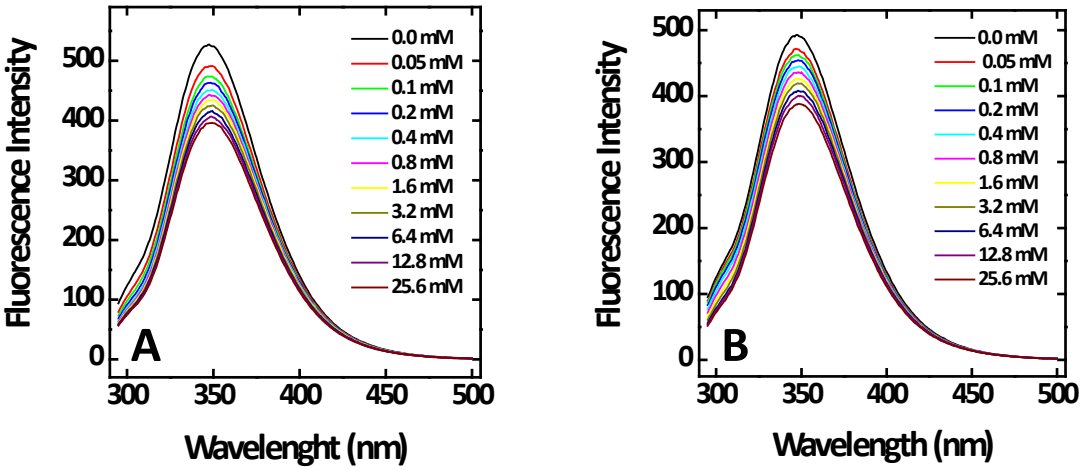


Fig. 3.7 Fluorescence spetra of synthesized human Osteocalcin (1-49). Spectra were recorded on 1 cm pathlength cuvette, 1μM at 25 °C in (**A**) 20 mM Tris-HCl, pH 7.4 and (**B**) 20 mM Tris-HCl, 0.15 M NaCl.

3.2.3 Hydrogen/Deuterium Exchange

Conformational studies were further performed by Hydrogen/deuterium exchange. Availability at hydrogen exchange depends on different factors including pH, temperature, flexibility and exposed sites of protein (Katta 1993). D_2O was added to synthesized osteocalcin solutions 20 mM ammonium acetate, pH 7.8 in the presence and in the absence of 1.5 mM $CaCl_2$. Samples were taken from both reaction environments at same time and analyzed by ESI-QTOF mass spectrometry.

The extent of H/D exchange was calculated using the equation $D = H^*(m - m_{0\%})/(m_{100\%} - m_{0\%})$ where m is the mass of protein at t time; $m_{0\%}$ is the measured mass of a "zero-deuteration" control; $m_{100\%}$ is the measured mass of a "full-deuteration" control.

Higher mass increment in reaction environment with lower ionic strength confirmed data obtained with circular dichroism and fluorescence (**Fig. 3.8**). Ca^{2+} induces alpha-helical structure in Osteocalcin, decreasing flexibility and consequently the number of exchangeable hydrogen atoms.

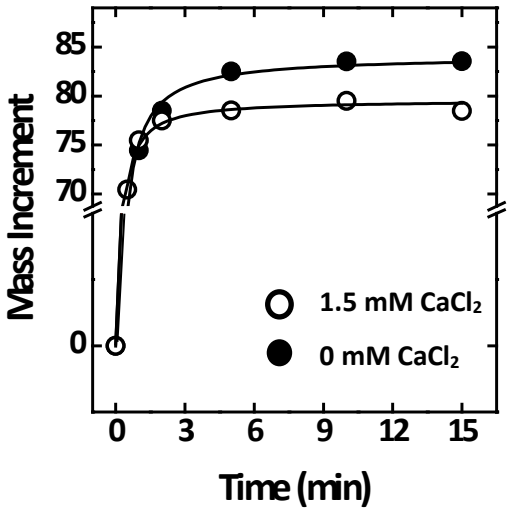


Fig. 3.8 Hidrogen/Deuterium exchange reaction in synthesized human Osteocalcin with and without $CaCl_2$.

3.3 Binding studies of Human Osteocalcin (1-49) and Ca^{2+}

Binding analyses of synthesized human Osteocalcin (1-49) with Ca^{2+} were performed by spectroscopic measurements and Isothermal Titration Calorimetry. Analyses were performed comparing synthesized and commercial Osteocalcin.

3.3.1 Far-UV Circular Dichroism

The effect of Ca2+ binding to human Osteocalcin (1-49) was investigated by recording the CD spectra at increasing [Ca^{2+}] and plotting the CD signal at 222 nm versus [Ca^{2+}] (**Fig. 3.9**). Interpolation of the data with *one site binding equation* ($Y = (Y_{MAX}*X) / (K_d+X)$)) (Lakowicz 1999) *y*ielded a dissociation constant (K_d) of 122 ± 15 μM and 1.5 ± 0.3 mM for the Ca^{2+}-OC complex and the Ca^{2+}-OC complex with 0.15 M NaCl respectively. These data suggest that in presence on 0.15 M NaCl the affinity of Ca^{2+} decreases by metal cations competition.

Higher K_d in physiological condition decreases the amount of Ca^{2+} saturated Gla-OC in serum.

Estimated K_d are in agreement with values obtained from commercial OC binding analysis: 146 ± 12 μM for the Ca^{2+}-OC complex and 1.1 ± 0.3 mM for the Ca^{2+}-OC complex with 0.15 M NaCl (**Fig. 3.9**).

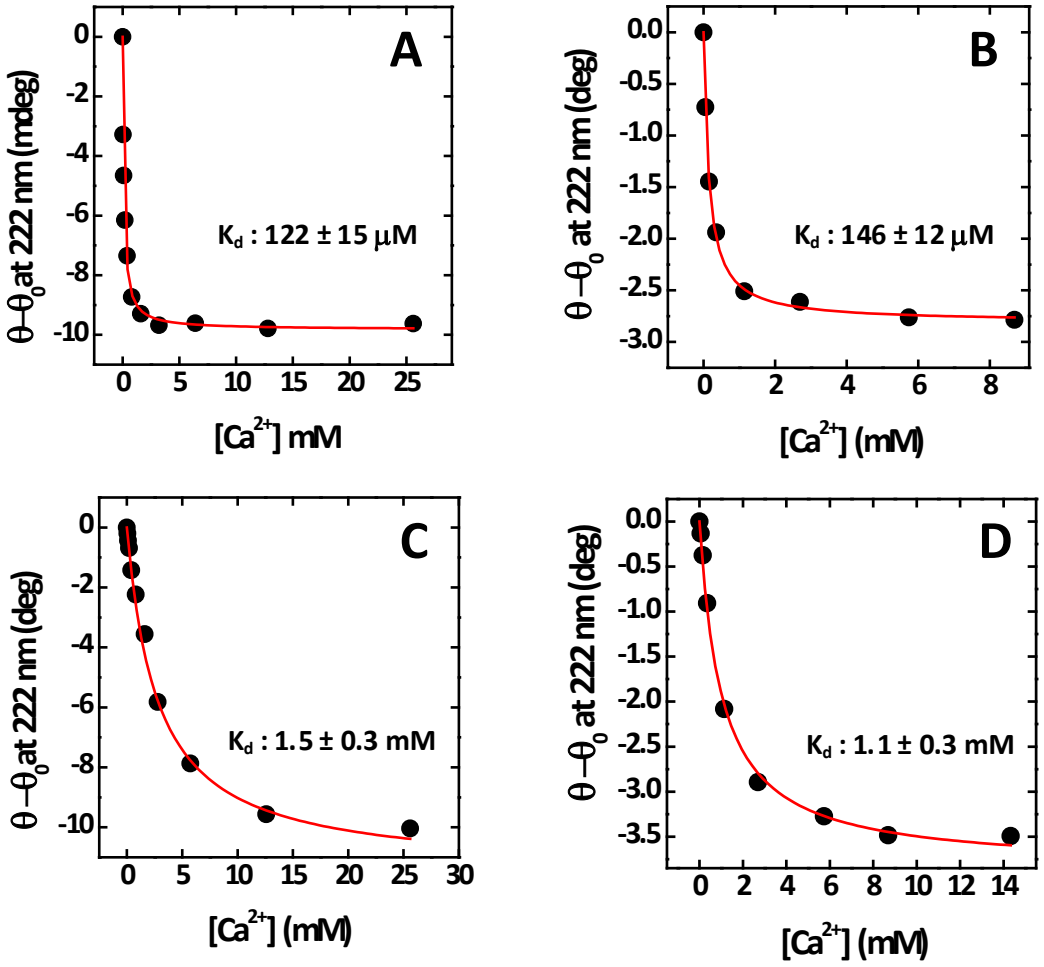


Fig. 3.9 Binding of human Osteocalcin (1-49) with Ca2+ by Far-UV CD. The effect of Ca2+ binding to human Osteocalcin (1-49) was investigated by recording the CD spectra at increasing [Ca^{2+}]. (A) Synthesized human OC in 20 mM Tris-HCl, pH 7.4, (B) Commercial human OC in 20 mM Tris-HCl, pH 7.4, (C) Synthesized human OC in 20 mM Tris-HCl, 0.15 M NaCl and (D) Commercial human OC in 20 mM Tris-HCl, 0.15 M NaCl. Dissociation constant is calculated by plotting the CD signal at 222 nm versus [Ca^{2+}] with *one site binding equation.*

3.3.2 Fluorescence

Intrinsic Trp fluorescence in Osteocalcin is decreased by Ca^{2+}. Fluorescence spectra were acquired at increasing [Ca^{2+}] and fluorescence emission monitored in the 290-500 nm range with en emission maximum at 334 nm. K_d was determined plotting the fractional occupancy (B) versus [Ca^{2+}](**Fig. 3.10**).

Contrarily to data obtained with circular dichroism, fluorescence data suggests the *two binding sites model* for Ca^{2+}-OC binding reaction. As shown in figure x, *one binding site equation* (Copeland 2000) doesn't plot experimental data which reflect typical behavior of *two sites model.*

One Binding Site

$$Y = \frac{Y_{max} * X}{K_D + X}$$

Two Binding Sites

$$Y = \frac{Y_{max} * X}{K_{D1} + X} + \frac{Y_{max} * X}{K_{D2} + X} \qquad K_{D1} \neq K_{D2}$$

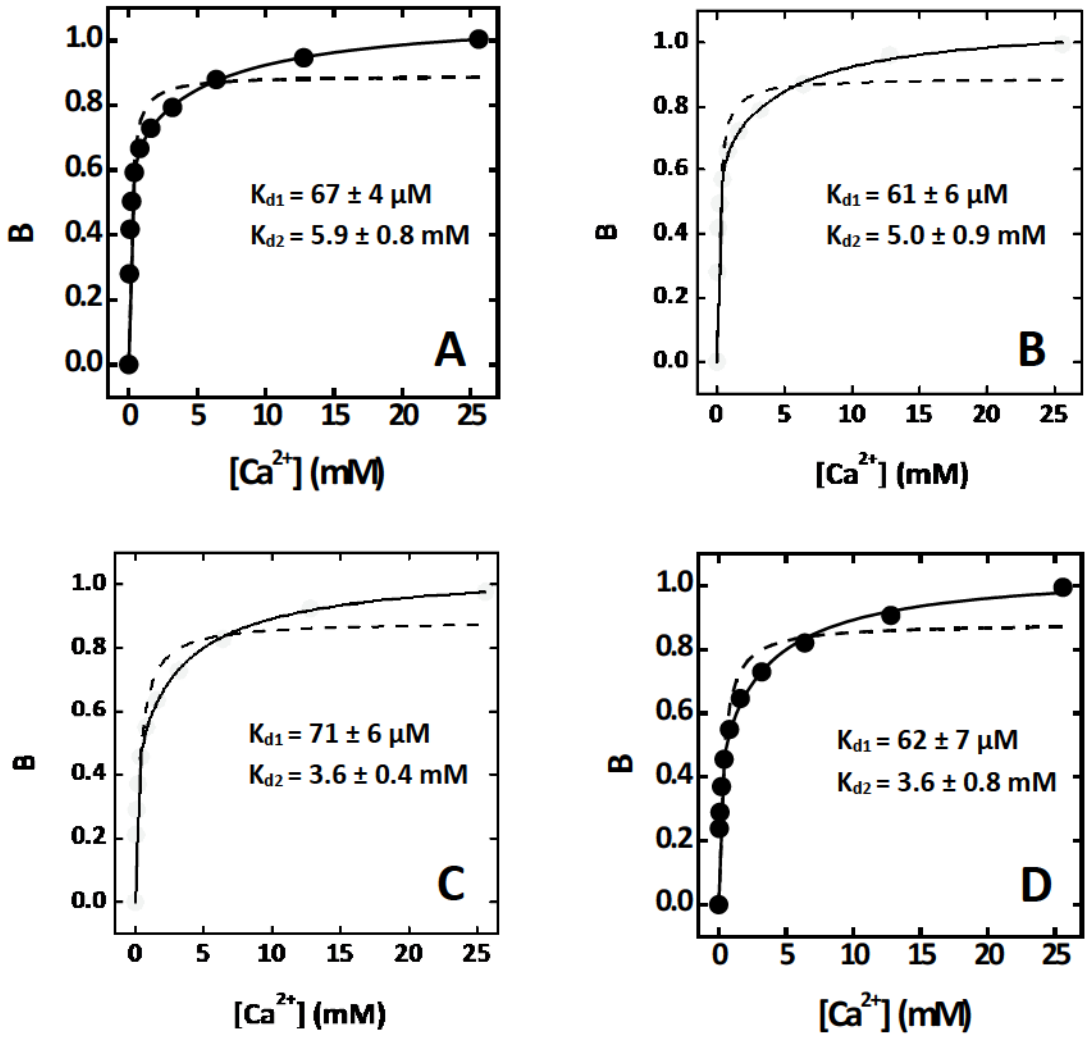


Fig. 3.10 Binding of human Osteocalcin (1-49) with Ca2+ by Fluorescence. The effect of Ca2+ binding to human Osteocalcin (1-49) was investigated by recording by fuorescence at increasing [Ca^{2+}]. (A) Synthesized human OC in 20 mM Tris-HCl, pH 7.4, (B) Commercial human OC in 20 mM Tris-HCl, pH 7.4, (C) Synthesized human OC in 20 mM Tris-HCl, 0.15 M NaCl and (D) Commercial human OC in 20 mM Tris-HCl, 0.15 M NaCl. Dissociation constants are calculated by plotting the fractional occupancy (B) versus [Ca^{2+}] with *two biding sites equation* (**—**). Uncorrected one site binding plot is reported (---).

Dissociation constants for synthesized human OC in 20 mM Tris-HCl, pH 7.4 (K_{d1}: 67 ± 4 μM, K_{d2}: 5.9 ± 0.8 mM) and in 20 mM Tris-HCl, 0.15 M NaCl, pH 7.4 (K_{d1}: 71 ± 6 μM, K_{d2}: 3.6 ± 0.4 mM) are in agreement with dissociation constants evaluated in commercial human OC in 20 mM Tris-HCl, pH 7.4 (K_{d1}: 61 ± 6 μM, K_{d2}: 5.0 ± 0.9 mM) and in 20 mM Tris-HCl, 0.15 M NaCl, pH 7.4 (K_{d1}: 62 ± 7 μM, K_{d2}: 3.6 ± 0.8 mM). Data suggest the presence of two binding sites for the metal cation with different affinity: K_{d1} values are constant in each condition whilst the affinity for the second binding site is higher in presence of metal cation competitor.

Controversial and ambiguous data from fluorescence analyses in Ca2+-OC binding are known (Hauschka 1982) and still investigated. Interpretation of these fluorescence changes in terms of polarity of Trp environment is not warranted by the available data, yet the conformational transition clearly affects Trp residue.

3.3.3 Isothermal Titration Calorimetry (ITC)

Ca^{2+}-OC binding studies were performed by Isothermal Titration Calorimetry (ITC). Titrations were performed at 25 ± 0.1°C in 50 mM HEPES pH 7.5. The sample solution (63 μM, 1400 μl) was treated by sequentially injections of a $CaCl_2$ stock solution (5.5 mM). After thermal equilibration (60 s), 1 μl of $CaCl_2$ solution was added to the sample, followed by 21 serial injections (5 μl) at intervals of 240 s with continuous stirring (300 r.p.m.).

Acquired data were plotted with the *one site binding equation*, and the dissociation constant was calculated (Kd: 99 ± 12 μM) (**Fig. 3.11**). Titration of synthesized human osteocalcin confirms data obtained with circular dichroism analysis, suggesting the presence of only one binding site for the metal cations.

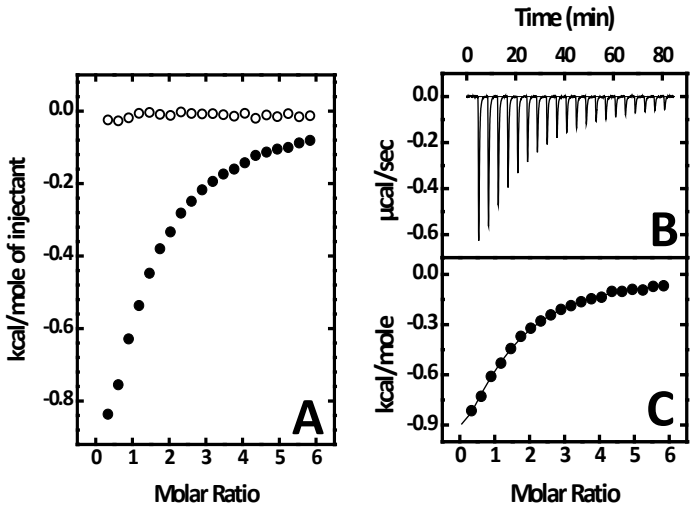


Fig. 3.11 Isothermal Titration Calorimetry of synthesized human OC with Ca^{2+}. (**A**, **C**) Enthalpy variation of human OC by sequential addition of $CaCl_2$ (●Sample, ○: blank). (**B**) Thermogram of OC by Ca^{2+} additions.

CONCLUSION

Chemical synthesis of human osteocalcin (OC) 1-49 was performed by the solid-phase method using the 9-fluorenylmethyloxycarbonyl (Fmoc) strategy on a PS3 automated synthesizer. The peptide chain was assembled stepwise on a HMPB-ChemMatrix resin including four check breaks to identify and optimize critical control points in peptide elongation. Synthesized peptide was subjected to the disulfide bond-forming reaction between Cys26 and Cys 29 and chemically characterized by RP-HPLC, mass spectrometry analysis and enzymatic digestion.

Fmoc chemistry represents a new attractive alternative strategy to synthesized human osteocalcin thanks to the higher final yield (about 80/85%), minor side reactions during cleavage and high safety process.

Conformational characterization by circular dichroism in the Far UV region shows that the Ca^{2+} induces alpha-helical structure in Osteocalcin. Under physiological conditions (0.15 M NaCl) the amount of alpha-helix decreases by 8-10% suggesting the competition among metal ions Ca^{2+} and Na^{1+}.

Furthermore, structural induction of human OC by Ca^{2+} was suggested by hydrogen/deuterium exchange analysis: Ca^{2+} decreases the flexibility and consequently the number of exchangeable hydrogen atoms as shown by the molecular mass improvement in absence or presence of calcium ions.

Binding studies of human OC and Ca^{2+} were performed by spectroscopic measurements and Isothermal Titration Calorimetry comparing the synthesized protein with the commercial sample.

Circular dichroism data suggested the one binding site model with estimated K_d of 122 ± 15 μM and 1.5 ± 0.3 mM for the Ca^{2+}-OC complex in absence and presence of 0.15 M NaCl (data in agreement with commercial OC analysis). The one binding site model was confirmed by ITC analysis with the estimated K_d: 99 ± 12 μM.

Controversial and ambiguous data from fluorescence analyses in the Ca^{2+}-OC binding are known (Hauschka 1982) and still investigated. Interpretation of fluorescence changes in terms of polarity of Trp environment is not warranted by the available data, yet the conformational transition clearly affects Trp residue.

REFERENCES

Albericio F, Lloyd-Williams P, Giralt E. Convergent solid-phase peptide synthesis. Meth Enzymol. 1997;289:313–36.

Albericio F, Kneib-Cordonier N, Biancalana S, Gera L, Masada RI, Hudson D, et al. Preparation and application of the 5-(4-(9-fluorenylmethyloxycarbonyl)aminomethyl-3,5-dimethoxyphenoxy)-valeric acid (PAL) handle for the solid-phase synthesis of C-terminal peptide amides under mild conditions. J Org Chem. 1990 Jun 1;55(12):3730–43.

Albericio F. Solid-Phase Synthesis: A Practical Guide. CRC Press; 2000. 852 p.

Angeletti RH, Bonewald LF, Fields GB. Six-year study of peptide synthesis. Meth Enzymol. 1997;289:697–717.

Aoki A, Muneyuki T, Yoshida M, Munakata H, Ishikawa S, Sugawara H, et al. Circulating osteocalcin is increased in early-stage diabetes. Diabetes Res Clin Pract. 2011 May;92(2):181–6.

Atherton E, Logan CJ, Sheppard RC. Peptide synthesis. Part 2. Procedures for solid-phase synthesis using Nα-fluorenylmethoxycarbonylamino-acids on polyamide supports. Synthesis of substance P and of acyl carrier protein 65–74 decapeptide. J Chem Soc, Perkin Trans 1. 1981 Jan 1;(0):538–46.

Ayers B, Blaschke UK, Camarero JA, Cotton GJ, Holford M, Muir TW. Introduction of unnatural amino acids into proteins using expressed protein ligation. Biopolymers. 1999;51(5):343–54.

Bao W, Holt LJ, Prince RD, Jones GX, Aravindhan K, Szapacs M, et al. Novel fusion of GLP-1 with a domain antibody to serum albumin prolongs protection against myocardial ischemia/reperfusion injury in the rat. Cardiovascular Diabetology. 2013;12:148.

Barany G, Kneib-Cordonier N, Mullen DG. Solid-phase peptide synthesis: a silver anniversary report*. International Journal of Peptide and Protein Research. 1987 Dec 1;30(6):705–39.

Bergmann M, Zervas L. Über ein allgemeines Verfahren der Peptid-Synthese. Ber dtsch Chem Ges A/B. 1932 Jul 6;65(7):1192–201.

Borgia JA, Fields GB. Chemical synthesis of proteins. Trends Biotechnol. 2000 Jun;18(6):243–51.

Buchwald H, Dorman RB, Rasmus NF, Michalek VN, Landvik NM, Ikramuddin S. Effects on GLP-1, PYY, and leptin by direct stimulation of terminal ileum and cecum in humans: implications for ileal transposition. Surg Obes Relat Dis. 2014 Oct;10(5):780–6.

Carpino LA, Han GY. 9-Fluorenylmethoxycarbonyl function, a new base-sensitive amino-protecting group. J Am Chem Soc. 1970 Sep 1;92(19):5748–9.

Casi G, Hilvert D. Convergent protein synthesis. Curr Opin Struct Biol. 2003 Oct;13(5):589–94.

Cristiani A, Maset F, De Toni L, Guidolin D, Sabbadin D, Strapazzon G, et al. Carboxylation-dependent conformational changes of human osteocalcin. Front Biosci (Landmark Ed). 2014 Jun 1;19:1105–16.

Cudic M, Burstein GD. Preparation of glycosylated amino acids suitable for Fmoc solid-phase assembly. Methods Mol Biol. 2008;494:187–208.

Dawson PE, Muir TW, Clark-Lewis I, Kent SB. Synthesis of proteins by native chemical ligation. Science. 1994 Nov 4;266(5186):776–9.

Dawson PE, Churchill MJ, Ghadiri MR, Kent SBH. Modulation of Reactivity in Native Chemical Ligation through the Use of Thiol Additives. J Am Chem Soc. 1997 May 1;119(19):4325–9.

De Toni L, De Filippis V, Tescari S, Ferigo M, Ferlin A, Scattolini V, et al. Uncarboxylated osteocalcin stimulates 25-hydroxy vitamin D production in Leydig cell line through a GPRC6a-dependent pathway. Endocrinology. 2014 Nov;155(11):4266–74.

De Toni L, Guidolin D, De Filippis V, Tescari S, Strapazzon G, Santa Rocca M, et al. Osteocalcin and Sex Hormone Binding Globulin Compete on a Specific Binding Site of GPRC6A. Endocrinology. 2016 Nov;157(11):4473–86.

Ducy P, Desbois C, Boyce B, Pinero G, Story B, Dunstan C, et al. Increased bone formation in osteocalcin-deficient mice. Nature. 1996 Aug 1;382(6590):448–52.

Ferron M, Wei J, Yoshizawa T, Del Fattore A, DePinho RA, Teti A, et al. Insulin signaling in osteoblasts integrates bone remodeling and energy metabolism. Cell. 2010 Jul 23;142(2):296–308.

Fields GB, Noble RL. Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. Int J Pept Protein Res. 1990 Mar;35(3):161–214.

Fields GB. Introduction to peptide synthesis. Curr Protoc Protein Sci. 2002 Feb;Chapter 18:Unit 18.1.

Fischer E, Fourneau E. Ueber einige Derivate des Glykocolls. Ber Dtsch Chem Ges. 1901 May 1;34(2):2868–77.

Fosgerau K, Hoffmann T. Peptide therapeutics: current status and future directions. Drug Discov Today. 2015 Jan;20(1):122–8.

Garnero P. Biomarkers for osteoporosis management: utility in diagnosis, fracture risk prediction and therapy monitoring. Mol Diagn Ther. 2008;12(3):157–70.

Giordano C, Marchiò M, Timofeeva E, Biagini G. Neuroactive peptides as putative mediators of antiepileptic ketogenic diets. Front Neurol. 2014;5:63.

Grant GA. Synthetic Peptides: A User's Guide. Oxford University Press; 2002. 401 p.

Guillier F, Orain D, Bradley M. Linkers and cleavage strategies in solid-phase organic synthesis and combinatorial chemistry. Chem Rev. 2000 Jun 14;100(6):2091–158.

Hamley IW. Peptide fibrillization. Angew Chem Int Ed Engl. 2007;46(43):8128–47.

Hauschka PV, Lian JB, Cole DE, Gundberg CM. Osteocalcin and matrix Gla protein: vitamin K-dependent proteins in bone. Physiol Rev. 1989 Jul;69(3):990–1047.

Henry RR, Rosenstock J, Logan D, Alessi T, Luskey K, Baron MA. Continuous subcutaneous delivery of exenatide via ITCA 650 leads to sustained glycemic control and weight loss for 48 weeks in metformin-treated subjects with type 2 diabetes. Journal of Diabetes and its Complications. 2014 May;28(3):393–8.

Houston ME, Campbell AP, Lix B, Kay CM, Sykes BD, Hodges RS. Lactam Bridge Stabilization of α-Helices: The Role of Hydrophobicity in Controlling Dimeric versus Monomeric α-Helices. Biochemistry. 1996 Jan 1;35(31):10041–50.

Käkönen SM, Hellman J, Pettersson K, Lövgren T, Karp M. Purification and characterization of recombinant osteocalcin fusion protein expressed in Escherichia coli. Protein Expr Purif. 1996 Sep;8(2):137–44.

Kaspar AA, Reichert JM. Future directions for peptide therapeutics development. Drug Discov Today. 2013 Sep;18(17–18):807–17.

Kates SA, Solé NA, Albericio F, Barany G. Solid-Phase Synthesis of Cyclic Peptides. In: Basava C, Anantharamaiah GM, editors. Peptides [Internet]. Birkhäuser Boston; 1994 p. 39–58. Available from: http://link.springer.com/chapter/10.1007/978-1-4615-8176-5_4

King DS, Fields CG, Fields GB. A cleavage method which minimizes side reactions following Fmoc solid phase peptide synthesis. Int J Pept Protein Res. 1990 Sep;36(3):255–66.

Knudsen LB. Liraglutide: the therapeutic promise from animal models. Int J Clin Pract Suppl. 2010 Oct;(167):4–11.

Kruse K, Kracht U. Evaluation of serum osteocalcin as an index of altered bone metabolism. Eur J Pediatr. 1986 Apr;145(1–2):27–33.

Kurihara T, Taniyama E, Hane M, Saito T, Hirose S, Ohashi S. Solid-phase synthesis of human osteocalcin by using a gamma-carboxyglutamic acid derivative. Int J Pept Protein Res. 1994 Apr;43(4):367–73.

Laizé V, Martel P, Viegas CSB, Price PA, Cancela ML. Evolution of matrix and bone gamma-carboxyglutamic acid proteins in vertebrates. J Biol Chem. 2005 Jul 22;280(29):26659–68.

Lee NK, Sowa H, Hinoi E, Ferron M, Ahn JD, Confavreux C, et al. Endocrine regulation of energy metabolism by the skeleton. Cell. 2007 Aug 10;130(3):456–69.

Li J, Zhang H, Yang C, Li Y, Dai Z. An overview of osteocalcin progress. J Bone Miner Metab. 2016 Jul;34(4):367–79.

Manning MC, Chou DK, Murphy BM, Payne RW, Katayama DS. Stability of protein pharmaceuticals: an update. Pharm Res. 2010 Apr;27(4):544–75.

Merrifield B. Solid phase synthesis. Science. 1986 Apr 18;232(4748):341–7.

Merrifield RB, Stewart JM, Jernberg N. Instrument for automated synthesis of peptides. Anal Chem. 1966 Dec;38(13):1905–14.

Merrifield RB. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. J Am Chem Soc. 1963 Jul 1;85(14):2149–54.

Merrifield RB. New approaches to the chemical synthesis of peptides. Recent Prog Horm Res. 1967;23:451–82.

Ottinger EA, Shekels LL, Bernlohr DA, Barany G. Synthesis of phosphotyrosine-containing peptides and their use as substrates for protein tyrosine phosphatases. Biochemistry. 1993 Apr 1;32(16):4354–61.

Otvos L, Elekes I, Lee VM. Solid-phase synthesis of phosphopeptides. Int J Pept Protein Res. 1989 Aug;34(2):129–33.

Oury F, Ferron M, Huizhen W, Confavreux C, Xu L, Lacombe J, et al. Osteocalcin regulates murine and human fertility through a pancreas-bone-testis axis. J Clin Invest. 2013 Jun;123(6):2421–33.

Padhi A, Sengupta M, Sengupta S, Roehm KH, Sonawane A. Antimicrobial peptides and proteins in mycobacterial therapy: current status and future prospects. Tuberculosis (Edinb). 2014 Jul;94(4):363–73.

Perich JW, Reynolds EC. Fmoc/solid-phase synthesis of Tyr(P)-containing peptides through t-butyl phosphate protection. International Journal of Peptide and Protein Research. 1991 Jun 1;37(6):572–5.

Poser JW, Esch FS, Ling NC, Price PA. Isolation and sequence of the vitamin K-dependent protein from human bone. Undercarboxylation of the first glutamic acid residue. J Biol Chem. 1980 Sep 25;255(18):8685–91.

Poser JW, Price PA. A method for decarboxylation of gamma-carboxyglutamic acid in proteins. Properties of the decarboxylated gamma-carboxyglutamic acid protein from calf bone. J Biol Chem. 1979 Jan 25;254(2):431–6.

Rieh D, Singh J. In The Peptides: Analysis, Synthesis, Biology; Gross, E., Meienhofer, J., Eds. 1979;(Journal Article).

Sakakibara S, Shimonishi Y, Kishida Y, Okada M, Sugihara H. Use of anhydrous hydrogen fluoride in peptide synthesis. I. Behavior of various protective groups in anhydrous hydrogen fluoride. Bull Chem Soc Jpn. 1967 Sep;40(9):2164–7.

Sheridan C. Proof of concept for next-generation nanoparticle drugs in humans. Nat Biotechnol. 2012 Jun 7;30(6):471–3.

Shin Y, Winans KA, Backes BJ, Kent SBH, Ellman JA, Bertozzi CR. Fmoc-Based Synthesis of Peptide-αThioesters: Application to the Total Chemical Synthesis of a Glycoprotein by Native Chemical Ligation. J Am Chem Soc. 1999 Dec 1;121(50):11684–9.

Sim S, Kim Y, Kim T, Lim S, Lee M. Directional assembly of α-helical peptides induced by cyclization. J Am Chem Soc. 2012 Dec 19;134(50):20270–2.

Sole NA, Barany G. Optimization of solid-phase synthesis of [Ala8]-dynorphin A. J Org Chem. 1992 Sep 1;57(20):5399–403.

Srivastava AK, Mohan S, Singer FR, Baylink DJ. A urine midmolecule osteocalcin assay shows higher discriminatory power than a serum midmolecule osteocalcin assay during short-term alendronate treatment of osteoporotic patients. Bone. 2002 Jul;31(1):62–9.

Stawikowski M, Cudic P. A novel strategy for the solid-phase synthesis of cyclic lipodepsipeptides. Tetrahedron Lett. 2006 Nov 27;47(48):8587–90.

Steidler L, Rottiers P, Coulie B. Actobiotics as a novel method for cytokine delivery. Ann N Y Acad Sci. 2009 Dec;1182:135–45.

Szulc P, Chapuy MC, Meunier PJ, Delmas PD. Serum undercarboxylated osteocalcin is a marker of the risk of hip fracture: a three year follow-up study. Bone. 1996 May;18(5):487–8.

Tam JP, Wu CR, Liu W, Zhang JW. Disulfide bond formation in peptides by dimethyl sulfoxide. Scope and applications. J Am Chem Soc. 1991 Aug 1;113(17):6657–62.

Timmerman P, Puijk WC, Boshuizen RS, Dijken P, Slootstra JW, Beurskens FJ, et al. Functional reconstruction of structurally complex epitopes using CLIPS^{TM} technology. The Open Vaccine Journal. 2009;2:56–67.

Vigneaud V du, Ressler C, Swan CJM, Roberts CW, Katsoyannis PG, Gordon S. THE SYNTHESIS OF AN OCTAPEPTIDE AMIDE WITH THE HORMONAL ACTIVITY OF OXYTOCIN. J Am Chem Soc. 1953 Oct 1;75(19):4879–80.

Wakamiya T, Saruta K, Yasuoka J, Kusumoto S. An Efficient Procedure for Solid-Phase Synthesis of Phosphopeptides by the Fmoc Strategy. Chem Lett. 1994 Jun 1;23(6):1099–102.

Zardeneta G, Chen D, Weintraub ST, Klebe RJ. Synthesis of phosphotyrosyl-containing phosphopeptides by solid-phase peptide synthesis. Analytical Biochemistry. 1990 Nov 1;190(2):340–7.

Zoch ML, Clemens TL, Riddle RC. New insights into the biology of osteocalcin. Bone. 2016 Jan;82:42–9.

CHAPTER THREE

PRODUCTION AND CHARACTERIZATION OF NEURITOGENIC PEPTIDES

ABBREVIATIONS

a.m.u	atomic mass unit
AA	aminoacid
Abs/A	Absorbance
ACN	Acetonitrile
Calcein-AM	Calcein acetoxymethyl ester
CAMs	Cell adhesion molecules
CD	Circular dichroism
CNS	Central nervous system
CNTs	Carbon nanotubes
Da	Dalton
DAT	Dopamine transporter
DCM	Dichloromethane
DIEA	Diethylamine
DMEM/F-12	Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DTT	Dithotreithol
EDT	Ethandithiol
ESI	Electrospray Ionization
FBS	Foetal bovine serum
F-moc	9-Flurenylmethyloxycabonyl
Fn	Fibronectin
GFP	Green fluorescent protein
HBSS	Hank's Balanced Salt Solution
Gnd-HCl	Guanidinium Chloride
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5b]pyridinium 3-oxid hexafluorophosphate
HBTU	3-[Bis(dimethylamino)methyliumyl]-3H-benzotriazol-1-oxide hexafluorophosphate

HCTU	[(6-Chloro-1H-benzotriazol-1-yl)oxy](dimethylamino)-N,N-dimethylmethaniminium hexafluorophosphate
HOBt	Hydroxybenzotriazole
Ig	Immunoglobulin
LINGO	LRR and Ig-like domain containing Nogo receptor interacting protein
LRR	Leucine rich repeat
MS	Mass Spectroscopy
MW	Molecular Weight
NCAM	Neural cell adhesion molecule
NGF	Nerve growth factor
NMP	N-Methyl-pyrrolidone
OtBu	O-tert- Butyl
Pbf:	2,2,4,6,7-pentemetyldihydrobenzofuran-5-sulfonyl
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
PD	Parkinson's disease
PhOMe	Methoxyphenyl group
PLLA	Poly(L-lactic acid)
Q-TOF	Quadrupole- Time Of Flight
RA	all-trans-retinoic acid
RARE	Retinoic acid response element
RARs	Retinoic acid receptors
RP-HPLC	Reversed-phase high-performance liquid chromatography
RXRs	Retinoid X receptors
SEM	Scanning electron microscopy
SPPS	Solide Phase Peptide Synthesis
tBoc	tert-Butyloxycarbonyl
tBu	tert-Butyl
TFA	Trifluoride Acetic Acid
TIPS	triisopropylsilyl chloride

TOF	Time of Flight
Tris-HCl	tris-(hydroxymethyl)-aminomethane chloride
Trt	Trytil group
UV	Ultra Violet
A (Ala)	Alanine
C (Cys)	Cysteine
D (Asp)	Aspartic Acid
E (Glu)	Glutamic Acid
F (Phe)	Phenylalanine
G (Gly)	Glycine
H (His)	Histidine
I (Ile)	Isoleucine
K (Lys)	Lysine
L (Leu)	Leucine
M (Met)	Methionine
N (Asn)	Asparagine
P (Pro)	Proline
Hyp	Hydroxyproline
Q (Asn)	Glutamine
R (Arg)	Arginine
S (Ser)	Serine
T (Thr)	Threonine
V (Val)	Valine
W (Trp)	Tryptophan
Y (Tyr)	Tyrosine

1. INTRODUCTION

1.1 REGENERATIVE MEDICINE

Regenerative medicine is a broad field that includes engineering and research on self-healing evolved as interdisciplinary technologies including life science, material and engineering knowledge. The purpose consists in restoring or maintaining physiological functions by delivering biological factors, cells and biomaterial scaffolds to support cell maturation, differentiation and proliferation as to native tissue.

Multiple approaches have been developed, depending on variety of injured tissue type and injury entity: (I) scaffolds functionalized with factors are implanted to recruit progenitor cells at the defective site to stimulate their differentiation, (II) Stem cells are loaded onto the scaffold *in vitro*, and after their differentiation the cell-scaffold composite is implanted into the defective site.

1.1.1 MULTIPLE CUES TO IMPROVE REGENERATION

The cells reorganize via multiple interactions with the extracellular environment that provides topographical and mechanical stimuli, concentration gradients of growth factors or extracellular matrix (ECM) molecules and electrical signals. Therefore, scaffolds for regenerative medicine should recapitulate/mimic as more such features of the native tissue environment as possible in order to properly provide cells with suitable information for differentiation and tissue development (Dvir 2011).

1.1.1.1 Biochemical cues

Biochemical cues are supplied by interactions between cells, soluble bioactive agents, and the ECM. Such biochemical cues are able to interact with cell receptors activating signaling cascade and determining cell differentiation. They include: (I) insoluble ECM macromolecules (collagens, elastin, etc.), glycoproteins (fibrinonectin, vitronectin, etc.), as well as polysaccharides such as heparin sulfate and hyaluronic acid, (II) soluble molecules such as chemokines, cytokines and growth factos and (III) cell-cell receptors such as cadherins, cell adhesion molecules (CAMs) and ephrins.

1.1.1.2 Physical cues

Physical cues are mostly provided by the ECM, which acts as a cellular scaffold. In vivo the ECM present a variety of defined, three dimensional physical cues due to its structure and molecular composition, referred to as topographies. Cell response to topographies is mediated by a phenomenon called contact guidance, which is known to affect cell adhesion, morphology, migration, and differentiation (Nikkhah 2012). Furthermore, ECM is able to modulate different cellular functions by its mechanical stiffness. Matrix sensing requires the ability of cells to pull against the matrix and cellular mechano-transducers to generate signals based on the force that the cell must generate to deform the matrix. Mechano-sensitive pathways subsequently convert these biophysical cues into biochemical signals that commit the cells to a specific lineage (Engler 2006). For examples neurons are able to reacts to topographical stimuli in order to direct neurite extension and axonal growth (Pettikiriarachchi 2010).

1.1.1.3 Electrical cues

Electrical biosignals are present in many developing systems and in normal cell cycle (division, migration, differentiation) (Yao 2009). In particular, electric signals activate membrane receptor and downstream intracellular signaling elements in neuronal cells leading to asymmetrical activation/redistribution of the cytoskeleton and the consequent cell polarization. Furthermore, many studies agree with the hypothesis of the involvement of electrical activity in growing axons mechanism and the formation of initial connection in the early nervous system (Yao 2009). Currently it was discovered that an endogenous electric field arises at the onset of neural tissue repair (Huang 2012).

1.2 THE NERVOUS SYSTEM: INJURY AND REGENERATION

1.2.1 Alternative and innovative approach for treating nerve injuries

Central nervous system (CNS) injury, like traumatic brain injury (TBI) or stroke, can result in several symptoms including cognitive, motor and psychotic dysfunction. Injury in peripheral nerves could lead to loss of neuronal communication along sensory and motor nerves between the CNS and the peripheral organs causing reduction in motor and sensory functions (Pettikiriarachchi 2010, Arslantunali 2014).The development of scaffolds for neural regeneration could help in the repair of damaged CNS tissue which has a limited regenerative potential as well as in peripheral nerve injuries when the nerve defect is too extended to be naturally regenerated (Harrison 2007, Saracino 2013).

1.2.1.1 Scaffold

Carbon nanotubes (CNTs) are attractive candidates for neurological applications thanks to their tunable electrical conductivity, excellent mechanical properties, morphological similarity to neurites and chemically modifiable surface (Dai 1996, Wong 1997, Mattson 2000, Hu 2004). CNTs are rolled layers of sp2 carbon resembling in form and dimensions neural processes and proteins of the extracellular matrix (e.g. collagen); this ability to mimic neural environment topography is very important to boost neuronal cell adhesion, growth and differentiation (Chao 2010, Fabbro 2013, Chua 2014, Resende 2014, Boccafoschi 2015, Bokara 2013) .CNTs are among the toughest and strongest nanomaterials; moreover, they are flexible and their tunable conductivity , stable in a biological environment (Chao 2010), can help to stimulate cultured neurons and to control cell migration, proliferation and differentiation (Mazzatenta 2007, Liopo 2006). CNT substrates can enhance neuronal network activity under culturing conditions, suggesting they might mediate charge redistribution along membrane surfaces, increasing neuronal excitability (Lovat 2005, Cellot 2011). Furthermore, tunable CNT electro-conductive scaffolds promote the functional maturation of neurons (Fabbro 2013) and can stimulate neurite outgrowth (Malarkey 2009).

Nevertheless, CNTs suffer from biocompatibility issues depending on type and doses of CNTs employed and on cell population tested (Cui 2010). A convenient way to improve biocompatibility involves the functionalization of CNTs (Prato 2008) by covalent or non-covalent chemistry to modify surface charge or to attach biologically active molecules able to

act as guidance cues to elongating neurites (Matsumoto 2007) These characteristics make CNTs ideal candidates for designing new biomaterials, which can reestablish the connections between neurons after injury. In this context, CNTs can find applicability as implants where long-term extracellular molecular cues for neurite outgrowth are necessary (Fabbro 2013).

1.2.1.2 Synthetic peptides for mimicking protein regulatory motifs

Concerning the extracellular environment, biologically active molecules can be potentially used in regenerative medicine to act as guidance cues for neurite elongation and thus to activate intracellular pathways leading to cell differentiation. To this aim, peptides corresponding to motifs likely to mediate regulatory signals present significant advantages compared to using entire recombinant proteins: (I) low immunogenic activity, (II) increased stability, (III) low production costs and (IV) simplified preparation and immobilization onto substrates (Chen 2008).

1.3 PROTEINS INVOLVED IN NEURITE OUTGROWTH

1.3.1 L1 family

The L1-CAM family proteins include: L1, DTT, Neurofascin and CHL1 (Close Homolog of L1). Their extracellular domains are composed by six Ig-like domains (Ig1-Ig6) followed by a variable number of FnIII and a single-pas transmembrane region. The short and highly conserved cytoplasmic tail, missing any catalytic domain, is functional to the downstram-signalling pathway (Liu 2011).

In the central and peripheral nervous system, L1 family members expression in mainly restricted to neurons and in particular to the axons. These proteins can interact both homophilically and heterophically with cis or trans bindings at long (as secreted factors) or short distance

1.3.1.1 L1

L1 cell adhesion molecule (CAM) is extremely important for a correct development of the nervous system in humans controlling neurite outgrowth, adhesion, fasciculation, migration, myelination and axon guidance. L1 is a transmembrane protein with an ectodomain consisting of six immunoglobulin-like (Ig) domains and five fibronectin type-III (FnIII) repeats (Haspel 2003). L1 is exposed at the surface of both neurons and glial cells and binds to a diverse set of molecules sending intracellular signals through its cytoplasmic region (Kenwirck 2000). In particular, neurite outgrowth is triggered by homophilic binding in trans occurring between two L1 molecules during axonal development. Even though it has been demonstrated that Ig1-Ig4 region is the minimal segment necessary to mediate homophilic binding (Gouveia 2008), Ig2 in particular possesses a homophilic interaction motif (Zhao 1998).

1.3.1.2 Neurofascin

Neurofascin has different splicing variants with various ectodomain architectures and expression patterns during CNS development, suggesting its relevant influence in the establishment and maintenance of the axonal pathways. In particular neurofascin enhances neurite outgrowth and controls postsynaptic structure isoforms, it is involved in the stabilization of synapses and clustering on Na^+ channels (Basak 2007, Kriebel 2012). Furthermore, neurofascin is involved the organization of $GABA_A$ receptors (Burkarth 2007) and regulates stabilization of paranodes (Kriebel 2012).

1.3.1.3 CHL1 (Close Homolog of L1)

The close homologue of L1 (CHL1) is detectable in subpopulations of neurons, astrocytes, oligodendrocyte precursors and Schwann cells both in CNS and PNS (Hillenbrand 1996). CHL1 shares high homology to L1 and in particular several residues are conserved in the hinge between Ig domains and FnIII repeats.

CHL1 function is controversial because it can mediate opposite effects depending on both the cellular subpopulation and the isoform: while the membrane-bound isoform may be involved in repulsive interactions, the soluble one can enhance adhesion and growth. CHL1 expression is upregulated when injury occurs, preventing L1 from enhancing neurite repair. The observation that CHL1 and L1 show overlapping and distinct patterns of synthesis in neurons and glia, suggests differential effects of L1-like molecules on neurite outgrowth

(Hillenbrand 1996). CHL1 is expressed even outside the CNS or PNS: it has been found in heart, intestine and gonads (Wei 1998).

In contrast to L1, CHL1 was suggested to neither act as a homophilic adhesion molecule (both in presence nor absence of Ca^{2+}) not mediate heterophilic adhesion with L1 (Hillenbrand 1999). However, more recent evidence for homophilic CHL1 interaction in trans was related to block the promoting function of CHL1 itself (Jakovcevski 2007) and to enhance differentiation of cerebellar granule cells in early development (Katic 2014).

1.3.1.4 DCC- Deleted in Colorectal Cancer

Deleted in Colorectal Cancer (DCC) is another Type I transmembrane glycoprotein, with an extracellular domain consisting of four Ig-like and six FnIII domains (Hedrick 1994) It was originally identified as a tumor suppressor gene located on chromosome 18q: its deletion in several tumor forms is implicated in differentiation and proliferation. DCC is constitutively expressed in all adult tissues, but the highest levels are found in neural tissue.

DCC is involved in complex regulatory mechanism of axon wiring and, according to the several combinations obtained, it may act as both attractive and repulsive cue for the growth cone: for this reason DCC is considered as a molecular switcher

1.3.2 ROBO family

Midline crossing between cerebral hemispheres is finely regulated by the Robo/Slit and DCC/ Netrin systems, preventing aberrant axonal overpassing.

In humans, three Slit proteins (Slit1-3) and four Robos (Robo1-3 and Robo4, involved in vascular angiogenesis) have been identified. Slits are large glycoproteins secreted by the midline cells, constituted by several LRR (Leucine-Rich Repeats), EGF like domains and a Laminin G domain. Instead, Robo1-3 are expressed at the surface of the growing axons. Robos are Type I transmembrane proteins, and their ectodomains share several features with other neuronal CAMs, in that consisting of five Ig like domains in the distal portion (Ig1-5) and three FnIII modules.

The signal mediated by this interaction is generally repulsive, expelling axons from the midline and preventing their recrossing (Ypsilanti 2010). As far as cell adhesion is concerned, Robo/Slit system can mediate either inhibition or enhancement depending on the type of cells and coreceptor involved. Finally, Robo/Slit play an important role in defining the topographical maps that help axonal pathfinding and in silencing repulsion effects.

1.3.3 Contactin family

The Contactin family includes six proteins (Contactin-1/F3/F11, Contactin-2/TAG-1/Axonin-1, Contactin-3/Big-1, Contactin-4/Big-2, Contactin-5/FAR-2/ NB-2 and Contactin-6/NB-3) mainly expressed in the nervous system. They are characterized by an extracellular domain consisting of six Ig-like and four FnIII-like domains, followed by a glycosyl phosphatidyl inositol (GPI)-anchoring domain by which it binds to the extracellular part of the cell membrane (Shimoda and Watanabe 2009).

Contactins are protagonists in neural cell adhesion and migration, neurite outgrowth and fasciculation, axon guidance and myelination although the precise molecular mechanisms involved in their effects on neurite outgrowth have not been identified yet. Since they lack any citoplasmatic domain (they are GPI-anchored proteins), they can exert their function only through homo- and heterophilic extracellular domain interactions that occur in cis and/or in trans with a number of transmembrane proteins (including those from L1CAM family) and ECM components. Therefore their ability to modulate neurite outgrowth depends on their adhesive properties and it may also involve other molecules and receptors that can transduce signals to cytoplasm. In particular contactins have been shown to interact with amyloid precursor protein (APP), Notch, and other Ig-CAMs (Mercati 2013).

Contactins are reported to regulate apical dendrite orientation in the visual cortex (Shimoda and Watanabe 2009).

1.3.4 LINGO family

LINGOs are transmembrane proteins localized at the PM. All LINGO members have a consistent ectodomain, made of 12 LRR (flanked at N- and C-termini with capping domains), followed by a single Ig module, then a single transmembrane domain and a short cytoplasmic tail containing an EGFR Tyrosine phosphorylation site (Mi 2004, Inoue 2007), that is conserved in all LINGOs except for LINGO4 (Llorens 2008).

In humans, LINGO1 is selectively expressed in brain, while the other LINGOs have been found ubiquitously (Mi 2004). LINGO1 is widely described in the previous chapter while other members of this family are still poorly understood.

1.3.4.1 LINGO1 protein

Leucine rich repeat (LRR) and Ig-like domain containing Nogo receptor interacting protein (LINGO1) is a transmembrane protein selectively expressed in CNS and playing an important role in neurite outgrowth. LINGO1 ectodomain is formed by twelve LRR motifs and one Ig-like domain (Mi 2004). LINGO1 forms at the neuron surface a complex with Nogo66-Receptor NgR1 and the neurotrophin receptor p75. NgR1/LINGO1/p75 complex interacts with Mag/NogoA/OMgp complex on oligodendrocyte surface and inhibits neurite elongation (Petrinovic 2010). NgR1/LINGO1/p75 complex also interacts with neuronal NogoA, thus blocking neurite elongation and generating repulsion between neurites (Petrinovic 2010). Furthermore, LINGO1 can homophilically interact in trans, suggesting that LINGO1 molecules from adjacent neuronal and oligodendrocytic membranes could interact inhibiting oligodendrocyte final differentiation (Jepson 2012). Recent discoveries showed that the LINGO1 Ig domain alone mediates hemophilic binding and heterotypic interactions (Stein 2012). LINGO1 expression is developmentally regulated, and it is upregulated after injury and in a number of CNS diseases (Inoue 2007, Mi 2013, Fernandez-Enright 2014). In vitro and in vivo studies demonstrated that when inhibiting LINGO1 function, its inhibitory effect is reversed and neurite elongation is promoted (Ji 2006). LINGO1 seems to be one of the molecules responsible for CNS neurons inability to regenerate and thus it represents an ideal candidate for therapeutic targets in neural dysfunctions because of its restricted tissue distribution.

1.4 STATE OF THE ART

In 2015 we presented the design and preparation of a nanocomposite scaffold made of multi-walled CNTs (MWCNTs) in a biocompatible poly-L-lactic acid (PLLA) matrix (CNT-PLLA scaffold) able to support neuronal cell growth and differentiation thanks to their ability to conduct electrical stimuli, to interface with cells and to mimic the neural environment. In parallel, we developed biomimetic peptides, derived from L1 and LINGO1 proteins, to further modulate neuronal differentiation.

Such peptides positively modulate neuronal differentiation, which is synergistically improved by the combination of the nanocomposite scaffold and the peptides, thus suggesting a prototype for the development of implants for long-term neuronal growth and differentiation (Scapin 2015, Scapin 2016).

In order to develop and characterize a set of novel, putatively biomimetic peptides, different neuritogenesis and axon guidance proteins were considered.

L1 family proteins, Contactin family proteins, LINGO family proteins, roundabout (ROBO) receptors and Netrin receptor Deleted in Colorectal Cancer (DCC) are known to be widely involved in differentiation, fasciculation, neurite outgrowth and axon guidance. These multidomain proteins organize into dynamic complexes, capable of sensing the signal from the extracellular space and modulating cellular activities by homo-and/or heterophilic interactions. The biological interactions are fully mediated by the complete extracellular portion; notably all reported protein families share Immunoglobulin (Ig)-like domains and fibronectin type-III (FnIII)-like domains in the extracellular portion. Emerging evidence suggest that a prominent role in homo or heterophilic binding is played by the Ig-like domains and in particular by the Ig2 repeat, from which the biomimetic L1A peptide was obtained (Scapin 2015), suggesting its capability to mimic the homophilic L1 binding motif (Zhao 1988). Furthermore, proteins from the LINGO family share a single Ig-like domain, which shows both sequence homology and structural similarity to Ig2. Considerably, LINGO1-A biomimetic peptide was dervide from the unique Ig-like domain of LINGO1.

2. EXPERIMENTALS

2.1 REAGENTS

Fmoc-Gly-Wang Resins (WAAA41313) and Fmoc-Lys(Boc)-Wang Resin (WAA41317), Fmoc-L-Tyr(tBU)-OH (FAA1230), Fmoc-L-Asp(tBu)-OH (FAA1020), Fmoc-L-Glu(tBu)-OH (FAA1045), Fmoc-L-His(Trt)-OH (FAA1090), Fmoc-L-Gln(Trt)-OH (FAA1043), Fmoc-L-Asn(Trt)-OH (FAA1015), Fmoc-L-Arg(Pbf)-OH (FAA1040), Fmoc-L-Ser(tBu)-OH (FSaa1592), Fmoc-L-Thr(tBu)-OH (FAA1210), Fmoc-L-Lys(Boc)-OH (FSP1125), Fmoc-L-Phe-OH (FAA1175), Fmoc-Gly-OH (FAA1050), Fmoc-L-Ile-OH (FAA1110), Fmoc-L-Leu-OH (FAA1120), Fmoc-L-Met-OH (FAA1150), Fmoc-L-Pro-OH (FAA1185), Fmoc-L-Val-OH (FAA1245), HATU (RL1190),HOBt, HBTU (RL1030), N-methylpirrolidone (SOL-009), piperidine (SOL-010), dymethylformamide (SOL-004) and trifluoroacetic acid (SOL-011) were purchased from Iris Biotech GMBH (Marktredwitz, Germany).

Diisopropylethylamine (387649), acetonitrile (34998), 1,2-Ethanedithiol (02390), triisopropylsilyl chloride (241725),phosphorus pentoxide (214701) and tertbutyl-methylether (306975) Sodium Chloride (S7653), gelatine (G9136), gentamicin (G1272), retinoic acid (R3255), poly-L-lysine solution (P8920), calcein acetoxymethyl ester (56496), Hank's balanced salt solution (H4641) and Sodium Phosphate Dibasic Anhydrous (76140) were purchased from Sigma Aldrich (St. Louis, Missouri, USA).

Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F-12) GlutaMAX^{TM} supplement (31331-028) 10% heat-inactivated FBS were purchased from ThermoFisher Scientific (Waltham, MA, USA).

2.2 METHODS

2.2.1 Bionformatic analysis

UniprotKB server was used to check out general information about function and structure, as well as to retrieve reference sequences and links to PDB structures. LALIGN server was used for pairwise global sequence alignment of the Ig and Ig2 domains.

Phyre2 server was used to obtain structural models while UCSF Chimera was used for structure viewing and structure superposition.

2.2.2 Synthesis and Purification

The synthesis of biomimetic peptides was performed by the solid-phase method using the 9-fluorenylmethyloxycarbonyl (Fmoc) strategy on a PS3 automated synthesizer from Protein Technologies International (Tucson, Arizona). Peptides were assembled stepwise on a Wang resin derivatized with the correspondent C-terminal amino acid residue. Tert-butyl side-chain protecting group was used for Tyr, Asp, Glu, Ser and Thr; tert-butyloxycarbonyl for Lys; trityl for His, Gln and Asn; and 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl group was used for Arg. Removal of N^{α}-Fmoc-protecting groups was achieved by 20% piperidine in N-methylpirrolidone (NMP) treatment for 20 minutes at room temperature. Standard coupling reactions were performed with an equal molar ratio of 2-(1H-benzotriazol-1-yl)-1,1,3,3,tetramethyluronium hexafluorophosphate (HBTU) and 1H-hydroxy-benzotriazole (HOBt) as activating agents, with a fourfold molar excess of the N^{α}-Fmoc-protected amino acids in the presence of diisopropylethylamine (0.35 M) in dymethylformamide (DMF). For double couplings at peptide bonds, the stronger activator 2-(7-aza-1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate was used (HATU). After peptide assembly, the side chain-protected peptidyl resin was treated for 90 min at room temperature with a mixture of TFA/H_2O/ethandithiol/triisopropylsilane. The resin was removed by filtration, and the acidic solution, containing the unprotected peptide, was precipitated with ice-cold tertbutyl-methylether and then dried. Peptides were analyzed by RP-HPLC with Jasco HPLC Pu-1575 equipped with 1575 UV-Vis detector on a Vydac C18 analytical column (4.6 x 150 mm, 5μm particle size, 300 Å) (Hesperia, California). The column was equilibrated with 0.1% (v/v) aqueous TFA and eluted with a linear 0.078% (w/w) TFA-acetonitrile gradient at a flow rate of

0.8 ml/min. Absorbance was monitored at 226 nm. Eluted material was analyzed by ESI-TOF mass spectrometry (MS) on a Mariner Workstation (Stafford, Texas).

Peptides containing aromatic chromophores are quantificated by UV-Vis specrotometry absorption analysis, using a Jasco UV-Vis spectrometer mod. 630 supplied with a peltier system PAC-743 (Jasco Corporation, Tokyo, Japan).

Peptide molar extiction factors were determined on the presence of aromatic aminoacids (F, Y, W) by online software EXPASY (Protparam tool): molar absorbibity could be calculated only for CHL1-A, DCC-A (ε=1390 M^{-1} cm^{-1}), CNTN1-A, CNTN2-A and CNTN5-A (ε=400 M^{-1} cm^{-1}).

In absence of chromophores, quantification was evaluated by analytic scaling (E/50 Gibertini, Novate Milanese, Italy).

2.2.3 Conformational characterization

Circular dichroism (CD) spectra were recorded on a Jasco J-810 spectropolarimeter equipped with a thermostatted cell holder and a Peltier PTC-423S temperature control system (Jasco, Tokyo, Japan). Spectra were recorded on 0.1 cm pathlength cuvette, 0.1 mg/ml in PBS buffer (10 mM Na_2HPO_4, 150 mM NaCl pH 7.4) at 25 °C. Each spectrum resulted from the average of four accumulations after base line subtraction. Ellipticity data were expressed as the mean residue ellipticity, $[\theta] = \theta_{obs}*MRW/(10*l*c)$, where θ_{obs} is the measured ellipticity in degrees, MRW is the mean residue weight, l is the cuvette pathlength, and c is the protein concentration in g/ml.

2.2.4 SH-SY5Y cell line growth and differentiation

Exponential growing human neuroblastoma SH-SY5Y cells were cultured with growth medium, composed of Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F-12) GlutaMAX™ supplement supplemented with 10% heat-inactivated FBS and 25µg/ml gentamicin, in humidified atmosphere of 5% of CO_2 in air at 37 °C. Cultures were maintained by subculturing 900000 cells into 25 cm^2 flasks every 2 days (once they reached 90% confluence).

Cell differentiation was induced by treating cells with all-trans-retinoic acid at 10 µM concentration and decreasing the FBS in the culture medium to 2% (differentiation medium) 24 h after seeding. In undifferentiated control samples, Dimethyl sulfoxide (DMSO) was added as equivalent amount (in which RA is dissolved). In experiments with peptides added to culture

medium, cells were seeded in a 24- well plate (15000 cells/well) coated with gelatine (porcine skin 0.005% in H2O milliQ)/poly-L-lysine (Invitrogen, 1μg/ml) solution. 24h after cell seeding (day 0), the growth medium was replaced by the differentiation medium (day 1); then, 24h after RA induction (day 2) peptides were added (either individually or in combination) at 1 μM concentration to the culture medium to asses if they can influence cell morphology in promoting/inhibiting neurite elongation. Peptides are not present in control samples. Neurite length was measured 24 h after the peptides addition (day 3).

2.2.5 Evaluation of neuronal differentiation

Neurite outgrowth was measured after staining the cells with calcein acetoxymethyl ester (Calcein-AM, Biotium), a non-fluorescent cell permeable compound that when hydrolyzed by intracellular esterases in live cells converts it to the strongly green fluorescent calcein (excitation 490 nm, emission 539 nm). This staining is particularly useful to clearly detect, along their whole length, the neural processes. Briefly, cells were incubated with calcein-AM (2 μM in HBSS, Hank's Balanced Salt Solution) for 30 minutes in dark at 37°C and 5% CO2 and visualized under a fluorescent microscope (Leica DM4000b) using green fluorescent protein (GFP) filter. Each experiment was performed in duplicate (two wells *per* conditions). Five images/well were recorded (ten images for each condition), taken with a 20X objective, 1.6X magnification. The first field was set to correspond to the centre of the well. Next fields were then selected following each of the four directions (N, S, W, E) from the first field. Neurite length was measured using ImageJ software by tracing the trajectory of the neurite from the tip to the junction between the neurite and cell body. If a neurite exhibited branching, the measure from the end of the longest branch to the soma was recorded, then each branch was measured from the tip of the neurite to the neurite branch point. The neuritogenic properties were analyzed in terms of total neurite length/n°.of cells (aggregate length of all cellular processes divided by cell number), n°.of neurites/n°.of cells, the percent of neurites longer than 100 μm,. Only neurites longer than 50 μm were considered (Hu et al., 2004; Munnamalai et al., 2014).

2.2.6 Statistical analysis

Statistical analysis was performed using paired Student's *t* test, and results were considered significant when $p < 0.05$. Values are expressed as mean ± standard error of the mean (M ± SEM).

3. RESULTS & DISCUSSION

3.1 DESIGN AND SYNTHESIS

3.1.1 Identification of Biomimetic peptides and Bioinformatic analysis

L1-CAM and LINGO1 posses multiple Ig-like domains (L1-CAM) or a single Ig (LINGO1) sharing a putative conserved binding motif at the Ig2/Ig domain from which biomimetic peptides L1A and LINGO1A were obtained (Scapin 2015). Considering all members of each family protein and interacting partners as Contactin family proteins, DCC/Netrin receptor and ROBO proteins, the amino acid sequence of the only Ig-like domain (or Ig2 for multi-Ig proteins) of each protein was used as a LALIGN probe for pairwise comparison of identity and similarity levels (**Tab 3.1**)

	L1	CHL1	Neurofascin	NrCAM	CNTN1	CNTN2	CNTN5	DCC	ROBO2	ROBO3	LINGO1	LINGO2
L1	100 100	59.1 80.7	50 77.9	48.8 79.1	38.8 63.8	38.2 63.2	40.8 64.8	26.5 57.4	30 62.9	27.8 59.7	27.8 58.2	24.3 63.5
CHL1	59.1 80.7	100 100	38.1 73.3	46.4 75.3	39.0 67.1	40.3 64.9	32.3 63.1	27.7 50.8	31.1 67.2	29.2 60.0	25.7 52.7	27.8 60.8
Neurofascin	50 77.9	38.1 73.3	100 100	56.3 81.6	40.0 65.3	34.5 57.1	34.3 68.7	31.0 59.2	34.3 65.7	31.8 60.6	28.9 51.8	28.8 56.2
NrCAM	48.8 79.1	46.4 75.3	56.3 81.6	100 100	30.5 59.8	50.0 60.0	33.3 60.9	33.3 59.3	31.2 59.4	31.8 54.5	30.8 58.5	25.3 51.9
CNTN1	38.8 63.8	39.0 67.1	40.0 65.3	30.5 59.8	100 100	61.7 76.6	51.8 79.5	22.6 50.9	32.8 59.7	28.2 54.9	23.0 55.2	25.0 52.9
CNTN2	38.2 63.2	40.3 64.9	34.5 57.1	50.0 60.0	61.7 76.6	100 100	52.5 81.2	31.9 54.2	28.0 57.3	29.7 56.2	28.8 50.7	31.3 50.7
CNTN5	40.8 64.8	32.3 63.1	34.3 68.7	33.3 60.9	51.8 79.5	52.5 81.2	100 100	28.0 57.3	39.2 58.9	34.8 62.3	25.6 50.0	32.4 59.2
DCC	26.5 57.4	27.7 50.8	31.0 59.2	33.3 59.3	22.6 50.9	31.9 54.2	26,7 51.2	100 100	25.9 61.4	25.9 59.1	31.9 65.3	36.2 65.2
ROBO2	30.0 62.9	31.1 67.2	34.3 65.7	31.2 59.4	32.8 59.7	28.0 57.3	39.2 58.9	25.9 61.4	100 100	25.9 59.1	29.3 53.7	35.4 62.2
ROBO3	27.8 59.7	29.2 60.0	31.8 60.6	31.8 54.5	28.2 54.9	29.7 56.2	34.8 62.3	25.9 59.1	25.9 59.1	100 100	28.8 56.2	34.8 60.3
LINGO1	27.8 58.2	25.7 52.7	28.9 51.8	30.8 58.5	23.0 55.2	28.8 50.7	25.6 50.0	31.9 65.3	29.3 53.7	28.8 56.2	100 100	56.7 86.7
LINGO2	24.3 63.5	27.8 60.8	28.8 56.2	25.3 51.9	25.0 52.9	31.3 50.7	32.4 59.2	36.2 65.2	35.4 62.2	34.8 60.3	56.7 86.7	100 100

Tab 3.1 Global Alignment among Ig domains. Number on top cell: % of identity in amino acidic sequence. Number on bottom cell: % of homologous amino acid that occur among sequences. All data are obtained from LALIGN online server.

The overall conservation is clearly evident: identity ranges 23/60%, while similarity ranges 50/80%. Structures of considered Ig domains were analyzed by superposition to identify conserved peptide motifs. Comparing structures of the different Ig repeats of the ectodomain, Ig2 domains resulted to show the highest structural closeness with conserved peptide motifs (**Fig. 3.1 A/B**).

Fig 3.1 (A) Superposition among Ig2 domains from Neurofascin, Contactin 2, ROBO1 and LINGO1 (obtained with UCSF Chimera). Peptide regions are highlighted in orange.

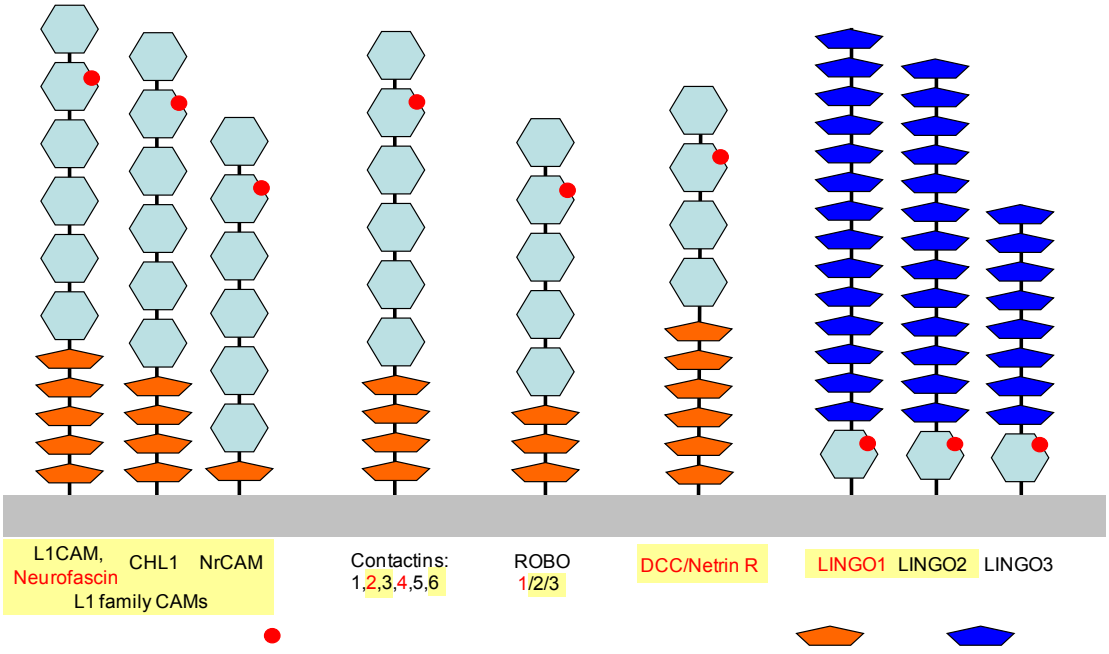


Figure 3.1 (B) Extracellular domain architecture of CAM and LINGO proteins recognized by a regular expression for the active binding motif.
Graphical legend for the Immunoglobulin-like (Ig-like), Fibronectin III-like (FnIII-like) and Leucine Rich Repeat (LRR) domains and for the active binding motifs (full red circles) are presented in the figure. When active binding motifs were reproduced and tested as synthetic peptides, the protein names are highlighted in yellow. Names of proteins are in red when a complete or partial solved structure including at least the Ig-like domain harbouring the active binding motif is available in Protein Data Bank (PDB). Hereafter are reported the UniprotKB accession numbers for all proteins in figure, followed between parentheses by corresponding PDB accessions (when available) : L1CAM, P32004; Neurofascin, O94856 (3P3Y,3P40); CHL1, O00533; NrCAM, Q92823; Cnt1, Q12860; Cnt2, Q02246 (1CS6); Cnt3, Q9P232; Cnt4, Q8IWV2 (3JXA); Cnt5, O94779; Cnt6, Q9UQ52; ROBO1, Q9Y6N7 (2V9R); ROBO2, Q9HCK4; ROBO3, Q96MS0; DCC/Netrin R, P43146 (3LAF); LINGO1, Q96FE5 (2ID5); LINGO2, Q7L985; LINGO3, P0C6S8.

3.1.2 Biomimetic peptide Synthesis, Purification and Chemical Characterization

Solid-phase synthesis of biomimetic peptides (**Tab. 3.2**) were carried out on a PS3 automated synthesizer from Protein Technologies International using the 0.1 mmol scale protocol based on manufacturer's instruction. Each amino acid was weighed (0.4 mmol) and placed in a vial. HOBt/HBTU mixture (1:1) or HATU (difficult coupling) were added amino acids and dried with phosphorous pentoxide.

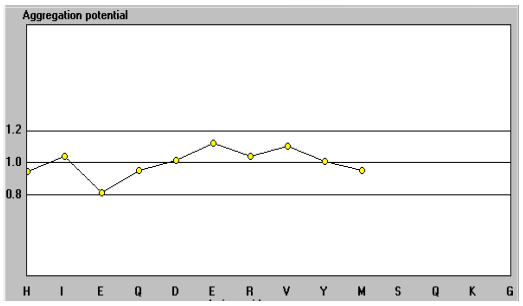
Peptide chains were assembled stepwise on a classical Wang resin derivatized with Fmoc-Cterminal amino acid. Solid phase peptide synthesis carries out in the C-terminal to N-terminal direction. Deprotection solution (20% piperidine in NMP) and activation solution (0.35 M diisopropylethylamine in DMF) were freshly prepared to overcome collateral reaction caused by ammine degradation. Anhydrous solvent are strictly required in order to improve production yields. The presence of water molecules would hydrolyze the amidic bond.

Peptide Companion software (CoshiSoft /PeptiSearch, San Diego, CA, USA) was used to predict aggregation potential of peptides (**Fig. 3.2**).

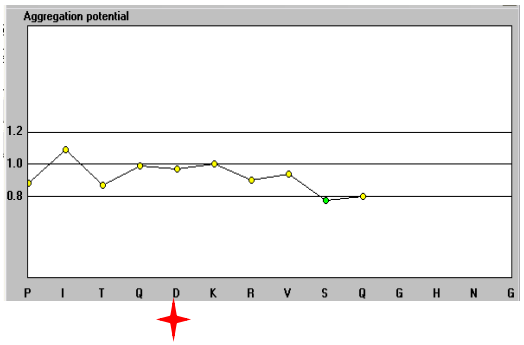
Peptide	Sequence
CHL1-A	HIEQDERVYMSQKG
Neurofascin-A	PITQDKRVSQGHNG
NrCAM-A	RLPQSERVSQGLNG
DCC-A	PIPGDSRVVVLPSG
LINGO2-A	TTKSNGRATVLGDG
ROBO2-A	IDDKEEIRISIRGGK
ROBO3-A	LKEEEGRITIRGGK
CNTN1-A	FITMDKRRFVSQTNG
CNTN2-A	FIPTDGRHFVSQTTG
CNTN5-A	FVAEDSRRFISQETG

Tab. 3.2 Synthesized Biomimetic peptides.

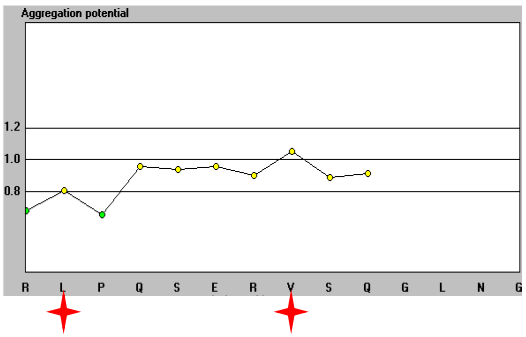
CHL1A



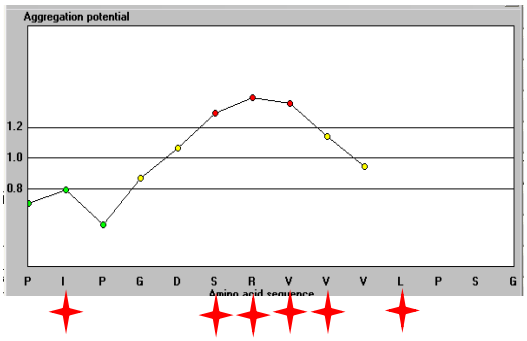
Neurofascin-A



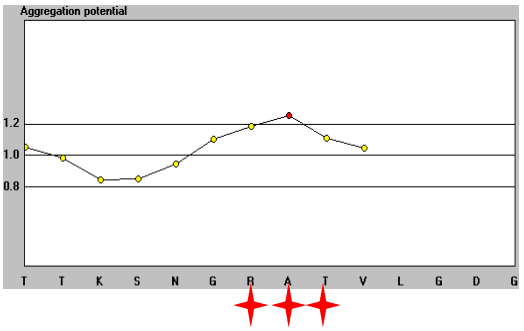
NrCAM-A



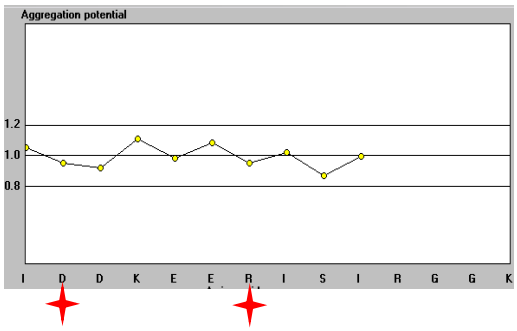
DCC-A



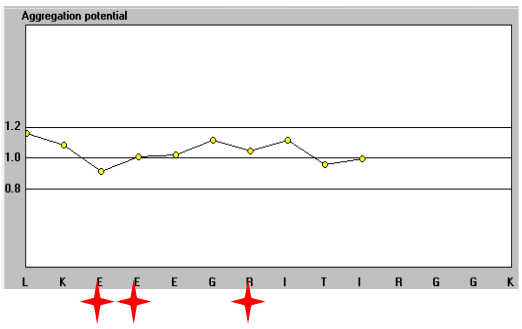
LINGO2-A



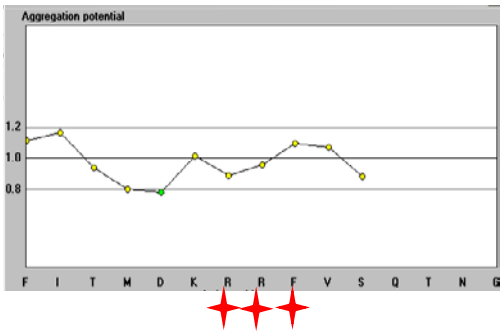
ROBO2-A



ROBO3-A



CNTN1-A



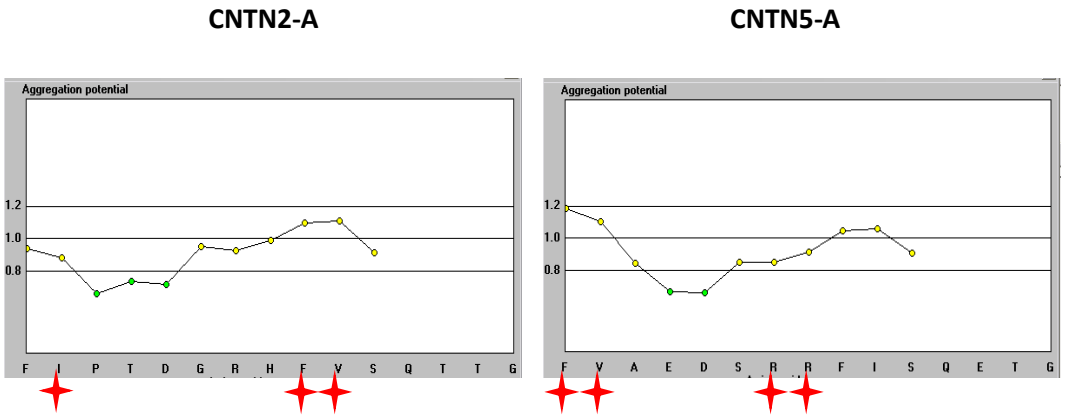
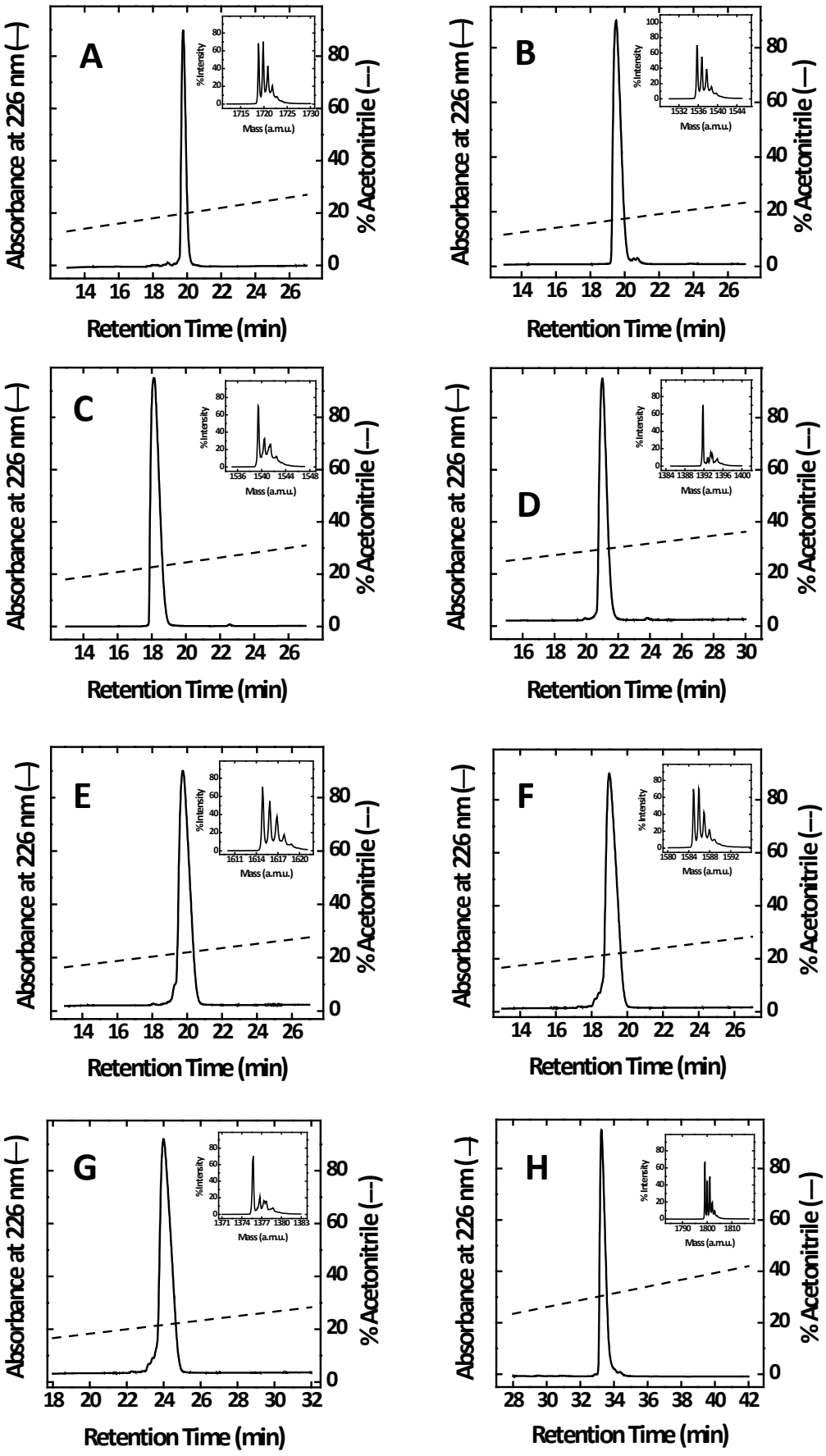


Fig 3.2 Aggregation potential and SPPS strategy. Peptides synthesis was planned considering the aggregation potential of each amino acid in the primary sequence. Easy, intermediate or difficult coupling are respectively marked in green, yellow and red. ( ) Incoming amino acids activated with HATU.

Two different solutions were prepared: TFA 92% (v/v), EDT 3% (v/v), H_2O (v/v) and TIPS 2% (v/v) for CHL1-A, Neurofascin-A, NrCAM-A, LINGO2-A, CNTN1-A, CNTN2-A, CNTN5-A; then a solution composed of TFA 95% (v/v), 3%H_2O (v/v) e TIPS 2% (v/v) was used for DCC-A, ROBO2-A and ROBO3-A. TFA treatment allows simultaneously deprotection of side-chains groups and peptide/resin removal. Cleavage and liberation of these reactive species can modify susceptible residues such as Met, Tyr and Trp. Modifications were minimized during TFA cleavage by utilizing suitable scavengers.

The resin was removed by filtration, and the acidic solution, containing the unprotected peptide, was precipitated with ice-cold tertbutyl-methylether and then dried.

Synthesized peptides were resuspended in H_2O/TFA 0.1% and analyzed by reverse-phase HPLC and mass spectrometry (**Fig. 3.3, Tab. 3.3**)


A
B
C
D
E
F
G
H
Absorbance at 226 nm (—)
% Acetonitrile (---)
Retention Time (min)
%Intensity
Mass (a.m.u.)

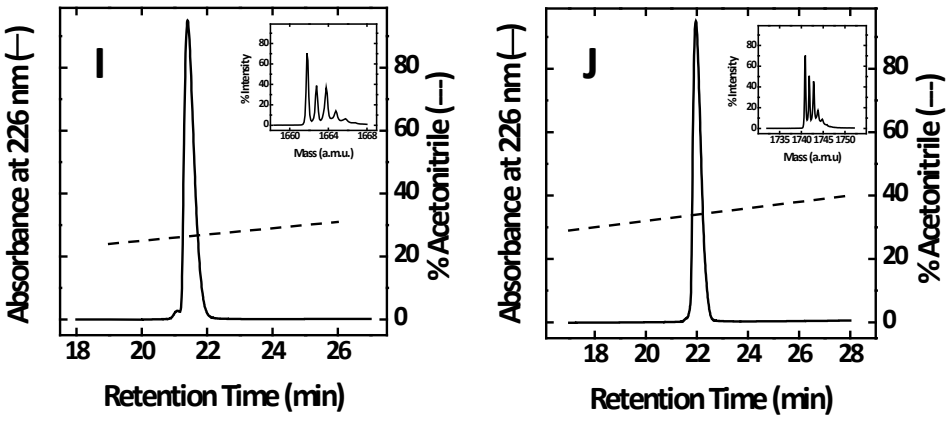


Fig. 3.3 Chemical characterization of synthetized peptides. RP-HPLC analysis of purified (**A**) CHL1-A, (**B**) Neurofascin-A, (**C**) NrCAM-A, (**D**) DCC-A (**E**) Lingo2-A, (**F**) ROBO2-A, (**G**) ROBO3-A, (**H**) Contactin1-A, (**I**) Contactin2-A and (**J**) Contactin5-A peptides on a Vydac C18 analytical column (4.6 x 150 mm, 5μm particle size, 300 Å) (Hesperia, California). The column was equilibrated with 0.1% (v/v) aqueous TFA and eluted with a linear 0.078% (w/w) TFA-acetonitrile gradient at a flow rate of 0.8 ml/min (tab 3.2.2). Absorbance was monitored at 226 nm. Eluted material was analyzed by mass spectrometry on a Mariner Workstation (Stafford, Texas).(**Inset**) Deconvoluted ESI-TOF mass spectra of RP-HPLC purified peptides. The estimated molecular masses are in agreement with the theoretical values (tab 3.2.2).

Peptide	Sequence	Theoretical MW [Da]	Experimental MW [Da]
CHL1-A	HIEQDERVYMSQKG	1718.80	1719.78 ± 0.02
Neurofascin-A	PITQDKRVSQGHNG	1535.78	1535.77± 0.01
NrCAM-A	RLPQSERVSQGLNG	1539.81	1539.84 ± 0.03
DCC-A	PIPGDSRVVVLPSG	1391.78	1391.82 ± 0.04
LINGO2-A	TTKSNGRATVLGDG	1375.71	1375.74 ± 0.03
ROBO2-A	IDDKEEIRISIRGGK	1614.87	1614.90 ± 0.03
ROBO3-A	LKEEEGRITIRGGK	1584.89	1585.91 ± 0.02
CNTN1-A	FITMDKRRFVSQTNG	1798.91	1799.89 ± 0.02
CNTN2-A	FIPTDGRHFVSQTTG	1661.82	1661.80 ± 0.02
CNTN5-A	FVAEDSRRFISQETG	1740.84	1740.91 ± 0.05

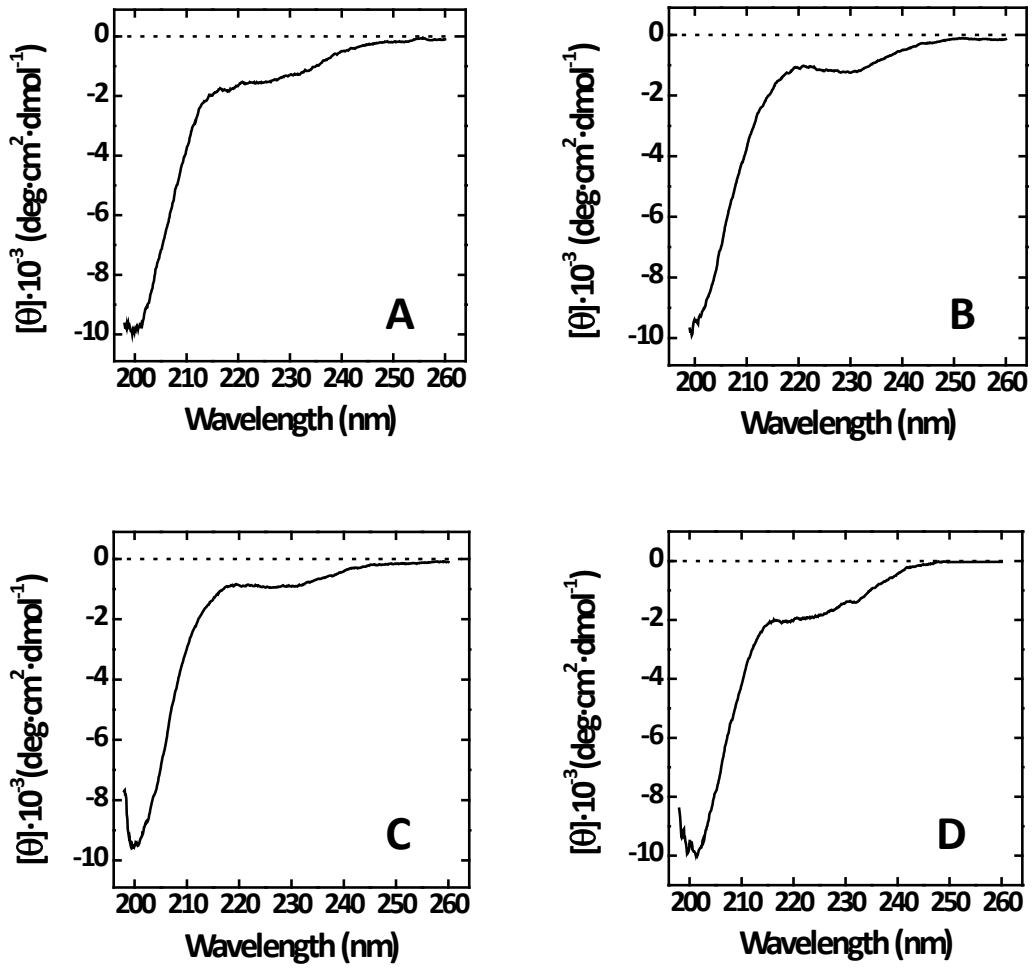
Tab. 3.3 Biomimetic peptides and relative theoretical/experimental molecular weight.

3.2 CONFORMATION CHARACTERIZATION

3.3 Far-UV Circular Dichroism

Conformation of purified peptides was investigated by circular dichroism (CD) on a Jasco J-810 spectropolarimeter, equipped with a Peltier temperature control system. Spectra were recorded using a 0.1 cm quartz cuvette, in the 260-198 nm range at 25°C. A 0.1 mg/ml sample concentration allowed to obtain a suitable dichroic signal. Since the instrument cannot overcome the limit of 600V, the analyses were performed at a wavelength above 198 nm.

As expected from peptide localization in flexible loop regions, conformation of the synthetic peptides in solution is essentially disordered as documented by the far-UV CD spectra (**Fig. 3.4**), resembling those typical of an unfolded polypeptide chain: positive band at 210 nm (typical of a n $\rightarrow\pi$ transition) and a negative one at 198 nm (due to $\pi \rightarrow \pi^*$ transition). Data are also in accordance to the predicted peptide structure obtained with Chimera software.



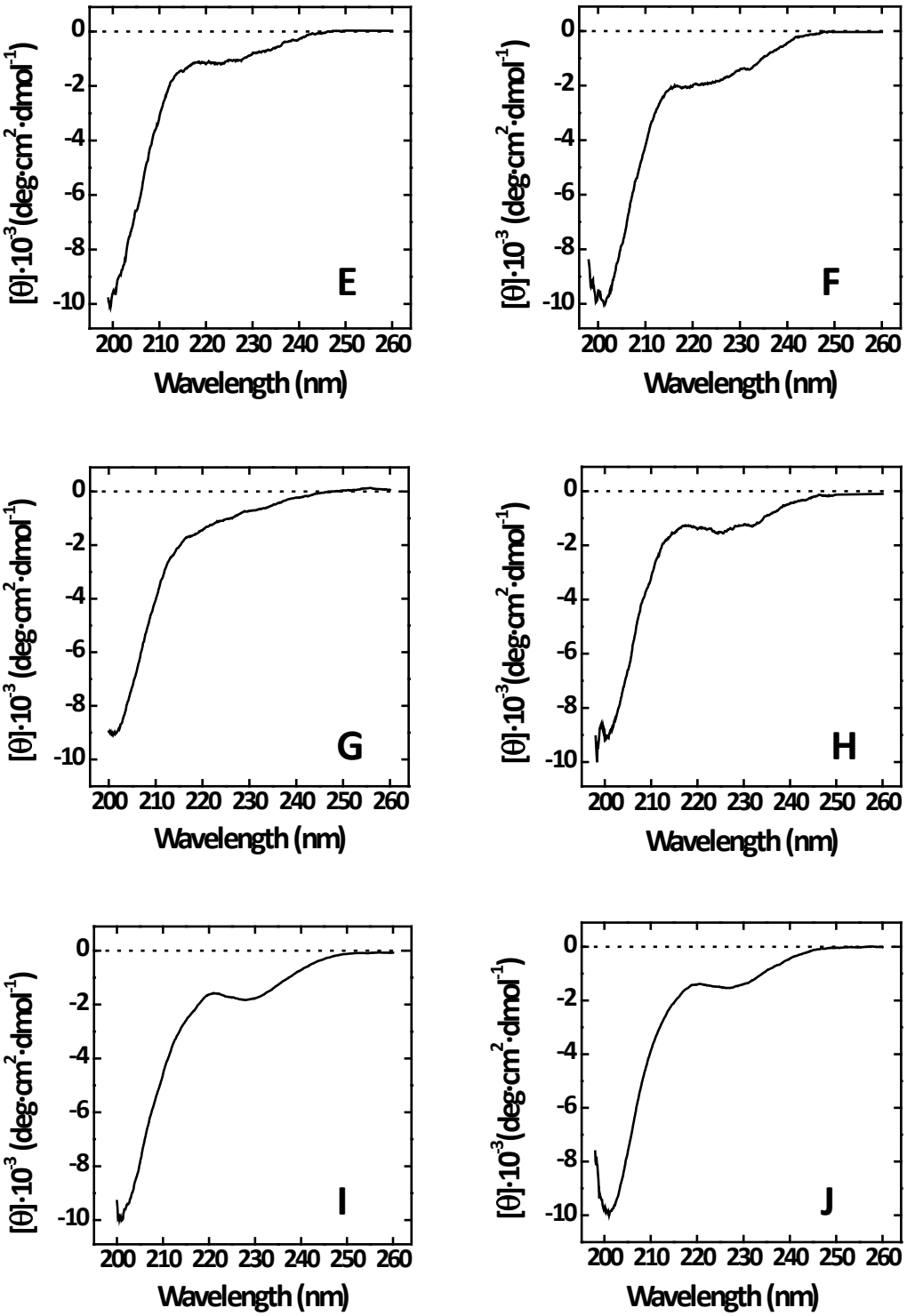


Fig. 3.4 Far-UV CD analyses of synthesized peptides. (**A**) CHL1-A, (**B**) Neurofascin-A, (**C**) NrCAM-A, (**D**) DCC-A (**E**) Lingo2-A, (**F**) ROBO2-A, (**G**) ROBO3-A, (**H**) Contactin1-A, (**I**) Contactin2-A and (**J**) Contactin5-A peptide. All samples are resuspended in phosphate buffer solution (0.1 mg/ml). Spectra are recorded using a 0.1 cm cuvette, in a range from 260 to 198 nm at 25°C.

3.3 ASSESSMENT OF PEPTIDE EFFECTS ON NEURONAL DIFFERENTIATION

3.3.1 Peptides in solution

Putatively biomimetic peptides were separately added to the cell culture medium in order to check their neuritogenic properties. According to Scapin's (Scapint 2015) protocol, analyses with synthesized peptides were performed as well with the SH-SY5Y cell line, using LINGO1A and L1A peptides as positive controls. As preliminary assessment of eventual biomimetic capacity, all new peptides were employed at 1 μM concentration, which was the one found to be optimal with both biomimetic L1-A and LINGO1-A (Scapin 2015).

L1-A scrambled peptide (L1-A_scr) was used as a second negative control (Scapin 2015), in addition to the untreated sample. Therefore, the effect of novel peptides was compared to two independent positive and two independent negative controls. In samples treated with scrambled peptide, values for total neurite length, number of neurites per neuron and percentage of neurites longer than 100 μm are comparable, or even lower than untreated control.

Neurites longer than 50 μm were considered as significant for morphological evaluation, according to the estabilished literature (Hu 2004, Munnamalai 2014). The evaluated parameters are: number of neurite/cells (**Fig. 3.5**), total neurite length/cell (**Fig. 3.6**) and % of neurites > 100 μm (**Fig. 3.7**).

In the following graphs, morphological evaluation for three independent experiments (in technical duplicate) are reported.

As far as number of neurites/cell is concerned (**Fig. 3.5**), all biomimetic peptides have a significant positive effect when compared to the negative control. Their activity is comparable or in some cases higher than the positive controls. Neurite outgrowth is mostly improved in samples treated with RA.

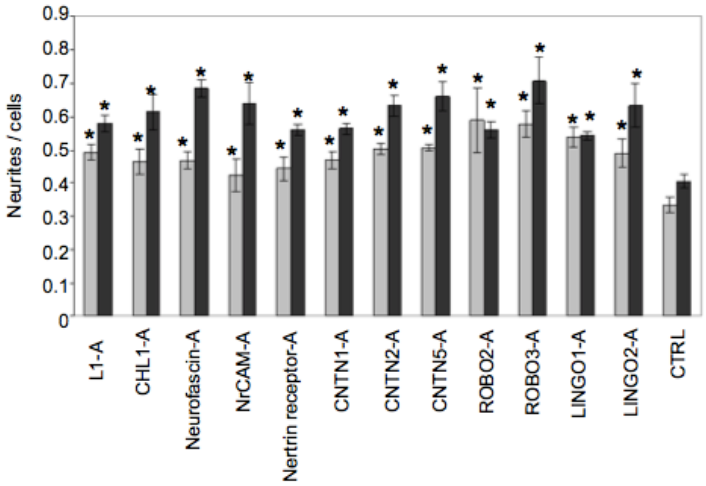


Fig. 3.5 Effect of peptides added to the culture medium on SH-SY5Y cell differentiation. Number of neurites/cell from SH-SY5Y cells treated with 1 μM peptide concentration either in the absence or presence of RA. Grey bars, cells treated with control medium (DMEM-F12; 2% FBS; 25 ug/ml gentamycin; 10uM DMSO); black bars: cells treated with differentiation medium (DMEM-F12; 2% FBS; 25 ug/ml gentamycin; 10 μM RA). Data represent the mean ± SEM of three independent experiments performed in duplicate. * shows significance at $p<0.05$ between samples treated with peptides and the control without peptides.

Concerning % of neuritis longer than 100 μm (**Fig 3.6**), all peptides show significantly higher data when compared to the control. Even in this case effects are more pronounced in samples treated with RA, except for peptides belonging to the ROBO and LINGO families. In these cases, % of long neurites is basically higher both for RA-treated samples and RA-untreated ones.

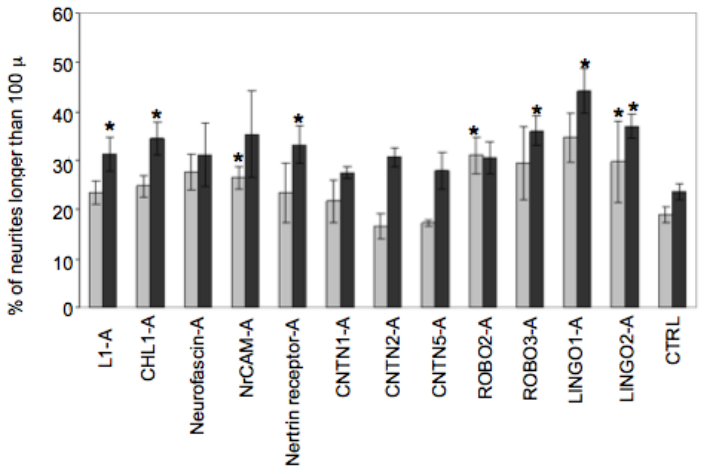


Fig. 3.6 Effect of peptides added to the culture medium on SH-SY5Y cell differentiation. Neurites longer than 100 μm from SH-SY5Y cells treated with 1 μM peptide concentration either in the absence or presence of RA. Grey bars, cells treated with control medium (DMEM-F12; 2% FBS; 25 ug/ml gentamycin ; 10uM DMSO); black bars: cells treated with differentiation medium (DMEM-F12; 2% FBS; 25 ug/ml gentamycin; 10 μM RA). Data represent the mean ± SEM of three independent experiments

performed in duplicate. * shows significance at p<0.05 between samples treated with peptides and the control without peptides.

Concerning total neurite length/cell (**Fig. 3.7**), all peptides show significantly higher data when compared to the control. Even in this case effects are more pronounced in samples treated with RA. The most evident effect is shown by ROBO3-A and Neurofascin-A. However, higher number of replicates will be needed to reach a more robust comparative picture.

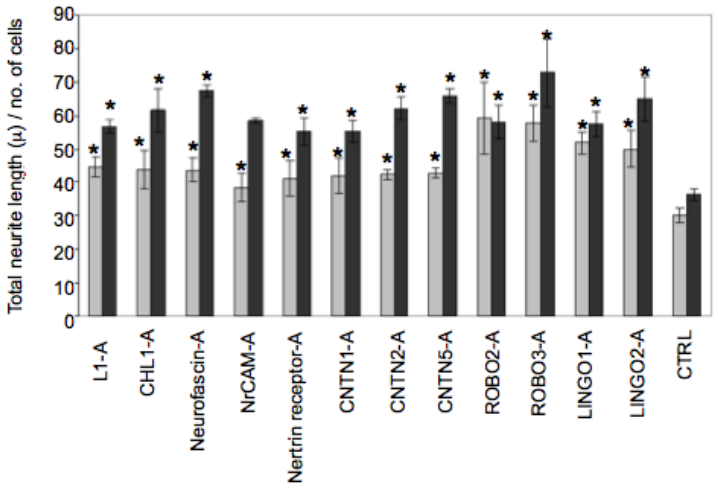


Fig. 3.7 Effect of peptides added to the culture medium on SH-SY5Y cell differentiation. Total neurite length from SH-SY5Y cells treated with 1 μM peptide concentration either in the absence or presence of RA. Grey bars, cells treated with control medium (DMEM-F12; 2% FBS; 25 ug/ml gentamycin ; 10uM DMSO); black bars: cells treated with differentiation medium (DMEM-F12; 2% FBS; 25 ug/ml gentamycin; 10 μM RA). Data represent the mean ± SEM of three independent experiments performed in duplicate. * shows significance at p<0.05 between samples treated with peptides and the control without peptides.

3.3.2 Preliminary analysis of peptide combinations

From the prevoius experiments, ROBO3-A and Neurofascin-A were selected for further experiments with L1-A and LINGO1-A because of the high biomimetic activity and the different biologic functions of the corresponding proteins (mediating either attractive or repulsive interactions). In order to check biomimetic peptide combinations, two preliminary independent experiments were performed in technical duplicate following the protocol used for peptide activity assessment. Peptide combinations were employed at 1 μM concentration.

As negative controls, untreated samples and cells treated with L1-A_scr were employed, to discriminate aspecific effects; while L1-A, LINGO1-A, ROBO3-A, Neurofascin-A and the combination L1-A+LINGO1-A were used as both positive controls and for comparison. The following combinations were tested: L1-A+Neurofascin-A, L1-A+ROBO3-A, LINGO1-A+ROBO3-A, LINGO1-A+Neurofascin-A and ROBO3-A+Neurofascin-A.

As shown in following figures, the effects mediated by the scrambled version of L1-A are comparable to those observed in untreated samples, confirming the sequence-dependence of the effect.

Concerning number of neurites *per* cell, all peptide combinations show both in the presence or absence of RA values clearly higher than the negative controls (**Fig. 3.8**). However, all combinations show results comparable to the samples treated with single peptides, and sometimes (e.g. with ROBO3-A+Neurofascin-A, L1-A+ROBO3-A) the neuritogenic effect is even lower than individual peptides.

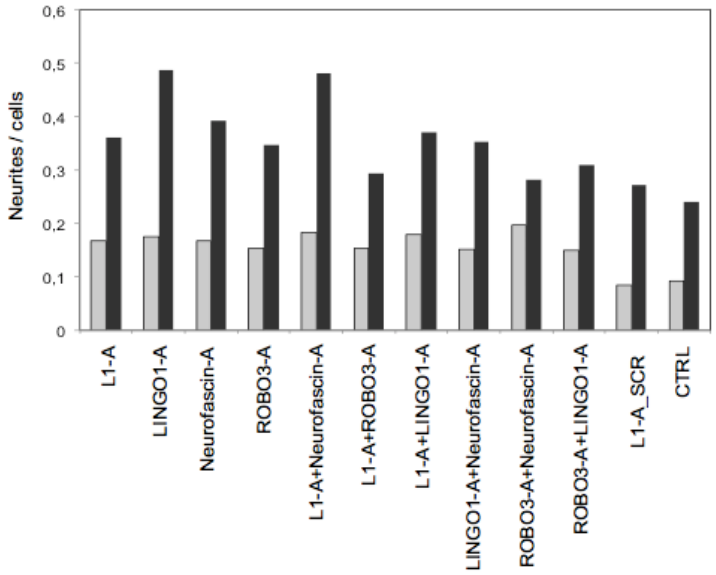


Fig. 3.8 Effect of peptide combinations added to the culture medium on SH-SY5Y cell differentiation. Number of neurites/cell from SH-SY5Y cells treated with 1 μM peptide concentration compared to peptide combinations either in the absence or presence of RA. Grey bars, cells treated with control medium (DMEM-F12; 2% FBS; 25 ug/ml gentamycin ; 10μM DMSO); black bars: cells treated with differentiation medium (DMEM-F12; 2% FBS; 25 ug/ml gentamycin; 10 μM RA). Data represent two independent experiments performed in duplicate.

As far as long neurites are concerned, the same trend can be observed (**Fig. 3.9**). The extrusion of long neurites is promoted especially for the positive control LINGO1-A alone or in combination with Neurofascin-A. As shown above, the effect is particularly evident in differentiative conditions. Moreover, in samples treated with peptide combinations but not with RA the effect is in some cases lower than the positive controls.

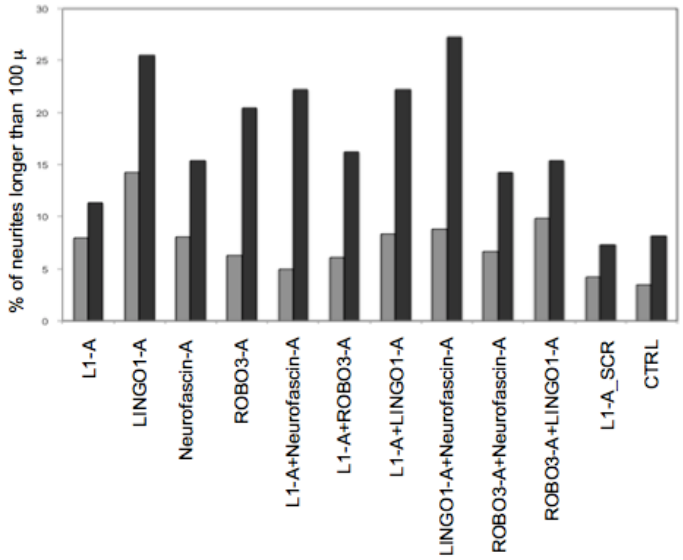


Fig. 3.9 Effect of peptide combinations added to the culture medium on SH-SY5Y cell differentiation. % of neurites longer than 100 µm from SH-SY5Y cells treated with 1 µM peptide concentration compared to peptide combinations either in the absence or presence of RA. Grey bars, cells treated with control medium (DMEM-F12; 2% FBS; 25 ug/ml gentamycin; 10µM DMSO); black bars: cells treated with differentiation medium (DMEM-F12; 2% FBS; 25 ug/ml gentamycin; 10 µM RA). Data represent two independent experiments performed in duplicate.

As far as total neurite length per cell is concerned (**Fig. 3.10**), peptides in combination give an overall effect comparable to the individual peptides: either in absence or presence of RA there is a positive effect, but it is not addictive. Concerning cells treated with RA the effect is much more pronounced.

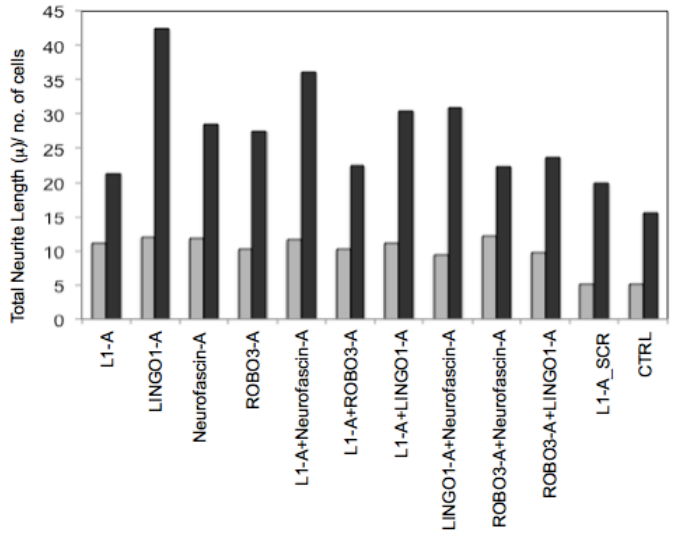
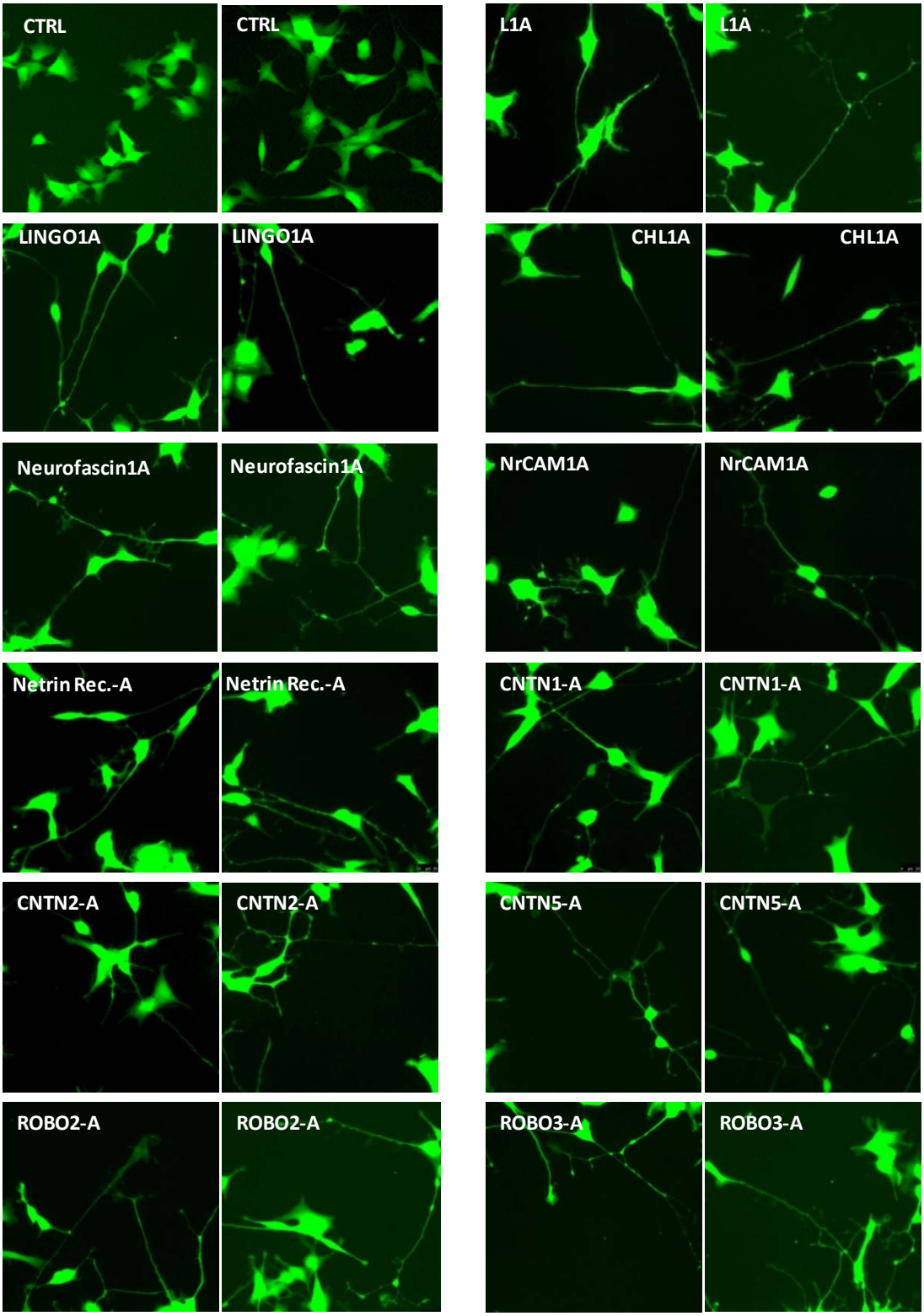
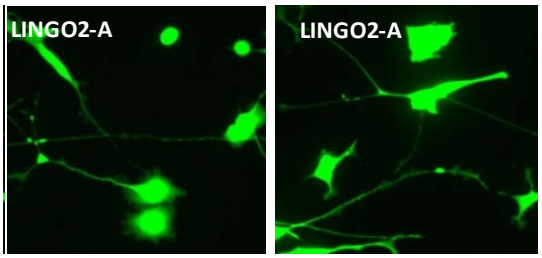


Fig. 3.10 Effect of peptide combinations added to the culture medium on SH-SY5Y cell differentiation. Total neurite length/cell from SH-SY5Y cells treated with 1 µM peptide concentration compared to peptide combinations either in the absence or presence of RA. Grey bars, cells treated with control medium (DMEM-F12; 2% FBS; 25 ug/ml gentamycin; 10µM DMSO); black bars: cells treated with differentiation medium (DMEM-F12; 2% FBS; 25 ug/ml gentamycin; 10 µM RA). Data represent two independent experiments performed in duplicate.

3.4 SUPPLEMENTARY MATERIALS





Biomimetic-treated SH-SY5Y cells. Cell are stained with calcein acetoxymethyl ester: a non-fluorescent cell permeable compound that when hydrolyzed by intracellular esterases in live cells converts it to the strongly green fluorescent calcein (excitation 490 nm, emission 539 nm).
Left Image: Biomimetic peptide -RA
Righ Image: Biomimetic peptide +RA

CONCLUSIONS

The preliminary step of this project consisted in biomimetic peptide design based on both literature and database search, performed to find structurally conserved proteins. In the end, ten Ig proteins were chosen (CHL1, Neurofascin, NrCAM, DCC, ROBO2 and 3, LINGO2, Contactin 1, 2 and 5) and then the corresponding peptides localized in Ig2 or the single Ig were identified. All such peptides show sequence conservation, and localization into the protein structure shows that they are inserted into a superimposable pocket structure.

The following step consisted in obtaining the peptides, so they were synthetized in SPPS exploiting F-moc chemistry and then purified through RP-HPLC. Species were confirmed by Mass Spectroscopy analysis. After chemical characterization, peptides are analysed from their conformational point of view with Circular Dichroism in the far-UV, assesting they all assume random coil conformation, notwithstanding they are inserted into a β-hairpin structure. Once peptides were characterized, their effects on cells were tested, in order to prove their biomimetic activity.

In agreement to the previous experiments performed for L1-A and LINGO1-A, all putatively biomimetic peptides were tested on SH-SY5Y cell line evaluating morphological changes (number of neurites per cell, total neurite length per cell and neurites longer than 100 βm). The neuritogenic effect is evidenced for the entire set, in fact they have a similar or in some cases higher effect than the positive controls L1-A and LINGO1-A. Moreover, the effect was tested also for peptides used in combination: in all such cases they are effective, but the entire effect is shaded and comparable to the individual peptides.

Since peptide combination does not give additional or synergistic effect on neuronal differentiation, a more complex mechanism of action might be involved. In facts all these homologuos peptides have comparable effects on neuronal differentiation, independently on they are derived from attractive or repulsive moieties. In every case, the effect is sequence specific: in facts it is significantly higher than the negative controls.

For a better characterization of the molecular mechanism underlying the neuritogenic properties, synthesis and test of consensus or chimeric peptides is proposed. Moreover, further experiments can be molecular characterization *via* SPR techniques, to prove peptide binding to the L1 ectodomain.

REFERENCES

Arslantunali D, Dursun T, Yucel D, Hasirci N, Hasirci V. Peripheral nerve conduits: technology update. Med Devices (Auckl). 2014;7:405–24.

Basak S, Raju K, Babiarz J, Kane-Goldsmith N, Koticha D, Grumet M. Differential expression and functions of neuronal and glial neurofascin isoforms and splice variants during PNS development. Dev Biol. 2007 Nov 15;311(2):408–22.

Boccafoschi F, Rasponi M, Ramella M, Ferreira AM, Vesentini S, Cannas M. Short-term effects of microstructured surfaces: role in cell differentiation toward a contractile phenotype. J Appl Biomater Funct Mater 2015 Jul 4

Bokara KK, Kim JY, Lee YI, Yun K, Webster TJ, Lee JE. Biocompatability of carbon nanotubes with stem cells to treat CNS injuries. Anat Cell Biol. 2013 Jun;46(2):85–92.

Burkarth N, Kriebel M, Kranz EU, Volkmer H. Neurofascin regulates the formation of gephyrin clusters and their subsequent translocation to the axon hillock of hippocampal neurons. Mol Cell Neurosci. 2007 Sep;36(1):59–70.

Cellot G, Toma FM, Varley ZK, Laishram J, Villari A, Quintana M, et al. Carbon nanotube scaffolds tune synaptic strength in cultured neural circuits: novel frontiers in nanomaterial-tissue interactions. J Neurosci. 2011 Sep 7;31(36):12945–53.

Chao T-I, Xiang S, Chen C-S, Chin W-C, Nelson AJ, Wang C, et al. Carbon nanotubes promote neuron differentiation from human embryonic stem cells. Biochem Biophys Res Commun. 2009 Jul 10;384(4):426–30.

Chen H, Yuan L, Song W, Wu Z, Li D. Biocompatible polymer materials: Role of protein–surface interactions. Progress in Polymer Science. 2008;11(33):1059–87.

Chua JS, Chng C-P, Moe AAK, Tann JY, Goh ELK, Chiam K-H, et al. Extending neurites sense the depth of the underlying topography during neuronal differentiation and contact guidance. Biomaterials. 2014 Sep;35(27):7750–61.

Cui H-F, Vashist SK, Al-Rubeaan K, Luong JHT, Sheu F-S. Interfacing carbon nanotubes with living mammalian cells and cytotoxicity issues. Chem Res Toxicol. 2010 Jul 19;23(7):1131–47.

Dai HJ WE Lieber CM. Probing electrical transport in nanomaterials: Conductivity of individual carbon nanotubes [Internet]. [cited 2016 Nov 13]. Available from: http://search.proquest.com/openview/fc89728e7218c62e8198412232c26cd2/1?pq-origsite=gscholar&cbl=1256

Dvir T, Timko BP, Kohane DS, Langer R. Nanotechnological strategies for engineering complex tissues. Nat Nanotechnol. 2011 Jan;6(1):13–22.

Engler AJ, Sen S, Sweeney HL, Discher DE. Matrix elasticity directs stem cell lineage specification. Cell. 2006 Aug 25;126(4):677–89.

Fabbro A, Prato M, Ballerini L. Carbon nanotubes in neuroregeneration and repair. Adv Drug Deliv Rev. 2013 Dec;65(15):2034–44.

Fabbro A, Sucapane A, Toma FM, Calura E, Rizzetto L, Carrieri C, et al. Adhesion to Carbon Nanotube Conductive Scaffolds Forces Action-Potential Appearance in Immature Rat Spinal Neurons. PLOS ONE. 2013 ago;8(8):e73621.

Fernandez-Enright F, Andrews JL, Newell KA, Pantelis C, Huang XF. Novel implications of Lingo-1 and its signaling partners in schizophrenia. Transl Psychiatry. 2014 Jan 21;4(1):e348.

Gouveia RM, Gomes CM, Sousa M, Alves PM, Costa J. Kinetic analysis of L1 homophilic interaction: role of the first four immunoglobulin domains and implications on binding mechanism. J Biol Chem. 2008 Oct 17;283(42):28038–47.

Harrison BS, Atala A. Carbon nanotube applications for tissue engineering. Biomaterials. 2007 Jan;28(2):344–53.

Haspel J, Grumet M. The L1CAM extracellular region: a multi-domain protein with modular and cooperative binding modes. Front Biosci. 2003 Sep 1;8:s1210-1225.

Hedrick L, Cho KR, Fearon ER, Wu TC, Kinzler KW, Vogelstein B. The DCC gene product in cellular differentiation and colorectal tumorigenesis. Genes Dev. 1994 May 15;8(10):1174–83.

Hillenbrand R, Molthagen M, Montag D, Schachner M. The close homologue of the neural adhesion molecule L1 (CHL1): patterns of expression and promotion of neurite outgrowth by heterophilic interactions. Eur J Neurosci. 1999 Mar;11(3):813–26.

Hu H, Ni Y, Montana V, Haddon RC, Parpura V. Chemically Functionalized Carbon Nanotubes as Substrates for Neuronal Growth. Nano Lett. 2004 Mar;4(3):507–11.

Huang Y-J, Wu H-C, Tai N-H, Wang T-W. Carbon nanotube rope with electrical stimulation promotes the differentiation and maturity of neural stem cells. Small. 2012 Sep 24;8(18):2869–77.

Inoue H, Lin L, Lee X, Shao Z, Mendes S, Snodgrass-Belt P, et al. Inhibition of the leucine-rich repeat protein LINGO-1 enhances survival, structure, and function of dopaminergic neurons in Parkinson's disease models. PNAS. 2007 Sep 4;104(36):14430–5.

Jakovcevski I, Wu J, Karl N, Leshchyns'ka I, Sytnyk V, Chen J, et al. Glial scar expression of CHL1, the close homolog of the adhesion molecule L1, limits recovery after spinal cord injury. J Neurosci. 2007 Jul 4;27(27):7222–33.

Jepson S, Vought B, Gross CH, Gan L, Austen D, Frantz JD, et al. LINGO-1, a transmembrane signaling protein, inhibits oligodendrocyte differentiation and myelination through intercellular self-interactions. J Biol Chem. 2012 Jun 22;287(26):22184–95.

Ji B, Li M, Wu W-T, Yick L-W, Lee X, Shao Z, et al. LINGO-1 antagonist promotes functional recovery and axonal sprouting after spinal cord injury. Mol Cell Neurosci. 2006 Nov;33(3):311–20.

Katic J, Loers G, Kleene R, Karl N, Schmidt C, Buck F, et al. Interaction of the cell adhesion molecule CHL1 with vitronectin, integrins, and the plasminogen activator inhibitor-2 promotes CHL1-induced neurite outgrowth and neuronal migration. J Neurosci. 2014 Oct 29;34(44):14606–23.

Kenwrick S, Watkins A, Angelis ED. Neural cell recognition molecule L1: relating biological complexity to human disease mutations. Hum Mol Genet. 2000 Apr 1;9(6):879–86.

Kriebel M, Wuchter J, Trinks S, Volkmer H. Neurofascin: a switch between neuronal plasticity and stability. Int J Biochem Cell Biol. 2012 May;44(5):694–7.

Liopo AV, Stewart MP, Hudson J, Tour JM, Pappas TC. Biocompatibility of native and functionalized single-walled carbon nanotubes for neuronal interface. J Nanosci Nanotechnol 2006;6(5):1365-1374.

Liu H, Focia PJ, He X. Homophilic adhesion mechanism of neurofascin, a member of the L1 family of neural cell adhesion molecules. J Biol Chem. 2011 Jan 7;286(1):797–805.

Llorens F, Gil V, Iraola S, Carim-Todd L, Martí E, Estivill X, et al. Developmental analysis of Lingo-1/Lern1 protein expression in the mouse brain: interaction of its intracellular domain with Myt1l. Dev Neurobiol. 2008 Mar;68(4):521–41.

Lovat V, Pantarotto D, Lagostena L, Cacciari B, Grandolfo M, Righi M, et al. Carbon nanotube substrates boost neuronal electrical signaling. Nano Lett. 2005 Jun;5(6):1107–10.

Malarkey EB, Fisher KA, Bekyarova E, Liu W, Haddon RC, Parpura V. Conductive single-walled carbon nano tube substrates modulate neuronal growth. Nano Lett 2009;9(1):264-268.

Matsumoto K, Sato C, Naka Y, Kitazawa A, Whitby RLD, Shimizu N. Neurite outgrowths of neurons with neurotrophin-coated carbon nanotubes. J Biosci Bioeng. 2007 Mar;103(3):216–20.

Mattson MP, Haddon RC, Rao AM. Molecular functionalization of carbon nanotubes and use as substrates

Mazzatenta A, Giugliano M, Campidelli S, Gambazzi L, Businaro L, Markram H, et al. Interfacing Neurons with Carbon Nanotubes: Electrical Signal Transfer and Synaptic Stimulation in Cultured Brain Circuits. J Neurosci. 2007 Jun 27;27(26):6931–6.

Mercati O, Danckaert A, André-Leroux G, Bellinzoni M, Gouder L, Watanabe K, et al. Contactin 4, -5 and -6 differentially regulate neuritogenesis while they display identical PTPRG binding sites. Biol Open. 2013 Mar 15;2(3):324–34.

Mi S, Lee X, Shao Z, Thill G, Ji B, Relton J, et al. LINGO-1 is a component of the Nogo-66 receptor/p75 signaling complex. Nat Neurosci. 2004 Mar;7(3):221–8.

Mi S, Pepinsky RB, Cadavid D. Blocking LINGO-1 as a therapy to promote CNS repair: from concept to the clinic. CNS Drugs. 2013 Jul;27(7):493–503.

Nikkhah M, Edalat F, Manoucheri S, Khademhosseini A. Engineering microscale topographies to control the cell-substrate interface. Biomaterials. 2012 Jul;33(21):5230–46.

Petrinovic MM, Duncan CS, Bourikas D, Weinman O, Montani L, Schroeter A, et al. Neuronal Nogo-A regulates neurite fasciculation, branching and extension in the developing nervous system. Development. 2010 Aug 1;137(15):2539–50.

Pettikiriarachchi JTS, Parish CL, Shoichet MS, Forsythe JS, Nisbet DR. Biomaterials for Brain Tissue Engineering. Aust J Chem. 2010 Aug 31;63(8):1143–54.

Prato M, Kostarelos K, Bianco A. Functionalized carbon nanotubes in drug design and discovery. Acc Chem Res. 2008 Jan;41(1):60–8.

Resende RR, Fonseca EA, Tonelli FMP, Sousa BR, Santos AK, Gomes KN, et al. Scale/Topography of Substrates Surface Resembling Extracellular Matrix for Tissue Engineering. Journal of Biomedical Nanotechnology. 2014 Jul 1;10(7):1157–93.

Roman JA, Niedzielko TL, Haddon RC, Parpura V, Floyd CL. Single-Walled Carbon Nanotubes Chemically Functionalized with Polyethylene Glycol Promote Tissue Repair in a Rat Model of Spinal Cord Injury. J Neurotrauma. 2011 Nov;28(11):2349–62.

Saracino GAA, Cigognini D, Silva D, Caprini A, Gelain F. Nanomaterials design and tests for neural tissue engineering. Chem Soc Rev. 2013 Jan 7;42(1):225–62.

Scapin G, Bertalot T, Vicentini N, Gatti T, Tescari S, De Filippis V, et al. Neuronal commitment of human circulating multipotent cells by carbon nanotube-polymer scaffolds and biomimetic peptides. Nanomedicine (Lond). 2016 Aug;11(15):1929–46.

Scapin G, Salice P, Tescari S, Menna E, De Filippis V, Filippini F. Enhanced neuronal cell differentiation combining biomimetic peptides and a carbon nanotube-polymer scaffold. Nanomedicine. 2015 Apr;11(3):621–32.

Shimoda Y, Watanabe K. Contactins: emerging key roles in the development and function of the nervous system. Cell Adh Migr. 2009 Mar;3(1):64–70.

Stein T, Walmsley AR. The leucine-rich repeats of LINGO-1 are not required for self-interaction or interaction with the amyloid precursor protein. Neurosci Lett. 2012 Feb 10;509(1):9–12.

Wei MH, Karavanova I, Ivanov SV, Popescu NC, Keck CL, Pack S, et al. In silico-initiated cloning and molecular characterization of a novel human member of the L1 gene family of neural cell adhesion molecules. Hum Genet. 1998 Sep;103(3):355–64.

Wong EW, Sheehan PE, Lieber CM. Nanobeam Mechanics: Elasticity, Strength, and Toughness of Nanorods and Nanotubes. ResearchGate. 1997 Sep 26;277(5334):1971–1975.

Yao L, McCaig CD, Zhao M. Electrical signals polarize neuronal organelles, direct neuron migration, and orient cell division. Hippocampus. 2009 Sep;19(9):855–68.

Ypsilanti AR, Zagar Y, Chédotal A. Moving away from the midline: new developments for Slit and Robo. Development. 2010 Jun;137(12):1939–52.

Zhao X, Yip PM, Siu CH. Identification of a homophilic binding site in immunoglobulin-like domain 2 of the cell adhesion molecule L1. J Neurochem. 1998 Sep;71(3):960–71.