# MSc in Biochemistry for Health

**Dissertation Project – 2nd Cycle**

Student´s Name:
Student email address: No.
Supervisor(s): Dr Benedita Pinheiro
Supervisor(s) email address: b.pinheiro@fct.unl.pt
Lab/Institution: Macromolecular Crystallography Laboratory & GlycoLab, UCIBIO@REQUIMTE, FCT/UNL
**TITLE:** *Assignment of new roles for malectin-like domains to understand their divergent evolution*

## BACKGROUND

Malectin was discovered as a lectin from the endoplasmic reticulum (ER) of animal cells, being highly conserved in the animal kingdom. In mammals, malectin is highly specific for di-glucosylated *N*-glycan ligands ($Glc_2$-*N*-glycans) present on *N*-glycosylated proteins in transit in the ER (Schallus et al.,2008) and is known to have a potential role in their genesis, processing and secretion, acting as backup protein quality control chaperone in conditions of ER stress. Recent reports have shown that malectin is upregulated in specific brain tumors, such as glioblastoma, and that can be used as cancer biomarker. Due to its important function in mammals, we have resolved the human malectin structure in complex with its endogenous glycan ligands, which explained at molecular level its highly restricted binding specificity to di-glucosylated glycan ligands. Since its identification, various malectin-like domain proteins have been discovered in every kingdom of life and have been assigned to a novel carbohydrate binding module (CBM) family that is believed to have a common ancestor (CAZY database). In plants, this ancestor has evolved into various different recognition receptors present at the plasma membrane. These are believed to recognize different glycan ligands as they are putative extracellular domains of receptor-like kinases, proteins that regulate many processes during the reproductive and vegetative development (Lidner et al., 2012). Several of these malectin-like domains are also identified in bacteria but their function is still unknown.

## OBJECTIVES

The central objective of this project is to study bacterial malectin-like domains by resolving their 3D-structure and identifying their glycan ligands. The information derived will be important to assign new roles for these protein domains in bacteria, and understand the divergent evolution of these domains present in different kingdoms of life.

It is intended that the student will receive training in different but complementary state-of-the-art methodologies in areas of Glycobiology, Structural and Molecular Biology.

## PROJECT DESCRIPTION

**Task 1: Bioinformatics analysis**

First the sequences of the malectin-like domains of bacteria and eukaryote will be analyzed and compared using several bioinformatic tools. The sequences will be fetched using the Carbohydrate-Active enZYme (CAZY), Uniprot and GenBank databases. The sequences will be analyzed using bioinformatic tools such InterPro, Pfam, SwissProt, CLustalOmega and others. From this we will create our own database.

**Task 2: Protein expression and purification**

From the information generated in Task 1, such as differences in their protein sequence, a first selection of the protein domains will be made and these will be recombinantly expressed for their structure elucidation and biochemical and functional characterization. The selected proteins will be cloned in the pET system with histidine tags. Expression tests will be performed varying the *E. coli* strain, temperature, type of expression induction (auto- and IPTG) and time of induction. These clones will be then expressed in small scale cultures in order to have enough protein to be tested in the microarrays. The proteins will be purified using immobilized metal affinity chromatography (IMAC).

**Task 3: Assessment of the glycan ligand specificities.**

High-throughput screening of the ligand specificity of these expressed bacterial malectin-domains will be performed using state-of-the-art carbohydrate-microarray technology. Binding to the glycan ligands immobilized into the microarray slides will be detected using a fluorescence immuno- and biotin-streptavidin detection systems.

**Task 4: Structural studies and biophysical characterization**

From the all the protein domains tested, 2 will be expressed in a larger scale for biochemical studies and preliminary crystallization and co-crystallization trials will be performed with the aim to resolving the 3D-structure by X-Ray crystallography. For the crystallization trials a crystallization robot will be used and several crystallization solutions tested. The mode of binding these malectin-like domains will be investigated using isothermal titration calorimetry (ITC), surface plasmon resonance (SPR) and non-denaturing gels.

## TIMELINE (use fill tool for the cells)

| | Month 1 | Month 2 | Month 3 | Month 4 | Month 5 | Month 6 | Month 7 | Month 8 | Month 9 | Month 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Task 1 | | | | | | | | | | |
| Task 2 | | | | | | | | | | |
| Task 3 | | | | | | | | | | |
| Task 4 | | | | | | | | | | |
| Thesis | | | | | | | | | | |