UNIVERSITY OF PISA

Department of Agriculture, Food and Environment

Master of Science

in

Plant and Microbe Biotechnologies

In vitro characterization of promiscuous biotinylation by biotin ligase

Supervisors:	*Candidate:*
Dr. Marek Mutwil	Camilla Ferrari
Dr. Rodolfo Bernardi	

Co-supervisor:

Prof. Annamaria Ranieri

Academic Year: 2013-2014

Contents

Abstract

Protein-protein interaction networks represent useful system to better understand the organization of the proteome and, moreover, provide a support for identifying protein functions. Thus, identification of these interactions would help to reach an improved comprehension of the entire cell. Most popular methods for detecting protein interactions are either *in vitro* assays or are based on heterologous systems, both of which do not represent true cellular environment. Thus, a lot of research is directed towards development of new methods that detect interactions under *in vivo* conditions, so as to preserve the spatio-temporal integrity of these interactions. One of such methods involve the biotin protein ligase (BirA) from *Escherichia coli*, which is responsible for catalyzing transfer of a small molecule, biotin, to a specific lysine residue (within a specific conserved sequence) of acceptor proteins. This reaction is highly specific, with only one protein in *E. coli* capable of being biotinylated *in vivo*. A mutation in BirA (R118G; BirA*), however, reduces this specificity such that the intermediate of reaction (biotinoyl-AMP) is released from the active site of the protein and can biotinylate any protein that it encounters. Based on this principle, if we fuse a protein with BirA*, the proteins that come in proximity of the fusion construct should be biotinylated and could further be identified by a streptavidin pull-down.

In order to develop this technique in plants, the enzymatic properties of BirA* need to be characterized. The current project was focused on purifying BirA* (codon optimized for expression in *Arabidopsis thaliana*) from bacteria and assessing its enzymatic properties under different physiological conditions. To this end, GST-6X Histidine fused BirA* was purified using inducible bacterial overexpression system and checked for promiscuous biotinylation of non-native target, *i.e.,* free glutathione S-transferase (GST) protein. Further, the substrate concentration, temperature, reaction time and pH conditions for the progress of reaction were optimized. The results generated by the current study verify that BirA* can catalyze promiscuous biotinylation, therefore it would be worthwhile to extend this system in plants for detection of protein-protein interactions. Moreover, if BirA* is selectively overexpressed in a specific cellular compartment, it has potential to biotinylate most of the proteins within this compartment. As different cellular compartments differ in pH, the present study also

helped to short-list the compartments whose proteome could be deciphered using BirA*.

To summarize, the present study verified that BirA* based proximity-biotinylation technique could be used to develop an *in vivo* assay for detecting protein interactions and to decipher organellar proteome. Also, we were able to optimize the reaction conditions at which the enzyme is most active.

Acknowledgements

I would like to express my gratitude and deepest appreciation to Dr. Marek Mutwil who gave me the great opportunity to work with his amazing group in a stimulating environment such as Max Planck Institute for Molecular Plant Physiology.

I gratefully acknowledge the precious advice of my supervisor Dr. Neha Vaid who helped me to carry out my project especially in writing this thesis, for her patience in answering my many doubts and for her support.

I also want to thank all of the colleagues and friends for the amazing time spent together and for the kind support they gave me in the discouraging moments.

A special thanks to Dr. Julia Smirnova for introducing me to ÄKTA Protein Purification Systems, a fundamental help to succeed in protein purification.

I would also like to express a sincere gratitude to Dr. Rodolfo Bernardi, who has always encouraged me to pursue this experience and motivate me in expanding my knowledge.

I also want to express my gratitude to my parents for their love and understanding.

1. Introduction

"Understanding how cells work requires understanding how proteins function". (Alberts et al., 2002)

Proteins are the most abundant macromolecule in the cell, constituting half of the dry-weight of an *E. coli* cell. Proteins are called the 'workhorses' of the cell, as they carry out the functions encoded by the genetic make-up of the living organism. They form structural and functional backbone of the entire cell and are involved in the majority of biological processes, such as DNA replication, transcription, protein translation, cell growth, proliferation, intracellular communication, apoptosis and many more. For many years, protein function analysis was based on the study of a single protein, however, it is now recognized that these processes are carried out not by proteins in isolation but instead complex protein networks, in which they interact with each other to carry out their biological activity. Consequentially, to understand the function of a protein within cell, it is of fundamental importance to analyze their function in context of their interacting partners.

1.1   Protein-protein interaction

Protein–protein interaction (PPI) is an event, due to electrostatic and/or biochemical forces, that leads to an intentional physical contact between two or more proteins.

The interactions established between two proteins can modulate the biological function of the molecules involved. PPIs are widely present virtually in every cellular pathway such as, transport machinery across membranes, packaging of chromatin, network of sub-membrane filaments and signal transduction, regulation of gene expression etc. (Royer, 1999).

Proteins can interact amongst themselves mainly by utilizing non covalent interactions such as, hydrogen bond, salt bridges, hydrophobic bonds, van der Waals forces, and rarely by covalent interactions such as disulphide bonds. These interactions are formed at specific binding domain of each protein that can be composed of few amino acids or can be represented by a large surface on their tertiary structure. The size of this domain

and the type of interaction determines the strength of the binding. According to their strength, the interactions could be classified as stable/strong and transient/weak interactions. For example Ras proteins, a family of GTP-ase involved in signal transduction within the cell, form tight complex with their partner when they are bound to GTP. This complex becomes weak when GTP is replaced with GDP. G-proteins show the similar behavior: the Gα subunit complexed with GDP forms a strong transient interaction with Gβγ subunits, but the trimer dissociates when Gα subunit binds GTP (Ozbabacan *et al.*, 2011).

1.2 Current methods

As PPIs are backbone of all cellular processes, several methods have been developed for their investigation. These include several biochemical methods (such as co-immunoprecipitation, yeast two-hybrid assay, bi-fluorescence complementation assay etc.) and biophysical approaches (such as fluorescence anisotropy, NMR spectroscopy, isothermal titration calorimetry etc.)

1.2.1 Biochemical methods

Amongst the most popular biochemical techniques is the yeast-two hybrid system (Y2H), which is based on the principle that eukaryotic transcriptional activators (GAL4, used in most cases) consist of two domains: a DNA binding domain that recognizes a specific DNA sequence and an activation domain, which summons and coordinates all of the elements required for transcription. An interaction between two proteins fused to these domains would bring the activation and binding domain in proximity, allowing transcriptional activation of GAL4, which is used as a marker (Chien *et al.*, 1991). Y2H system, developed to screen whole cDNA libraries to identify interacting partners for a protein of interest, has been successfully employed in identifying protein interactions from several organisms. For example, this technique has helped in identification of plant disease resistance as the interaction between the protein RPS2 of *Arabidopsis thaliana* and the product of the avirulence gene *avrRpt2* of *Pseudomonas syringae* (Betania *et al.*, 2004). It was also used to identify proteins involved in floral

development and their network of interactions, such as the products of most of floral homeotic genes belonging to the MADS box family and specific floral organ identity in *Antirrhinum* (Davies *et al.*, 1996). Another example of the application of Y2H is in the study of self-incompatibility mechanisms to further characterize the interactions between the *Brassica* S-receptor kinase (SRK) and its substrates (Mazzurco *et al.*, 2001).

Another classical method used to detect PPIs is pull-down assay, which is based on affinity purification of protein complexes from cell lysates. Affinity purification linked to mass-spectrometry has been successful in identifying proteins involved in plant immunity, as the interaction partners of the RPM1-interacting protein 4 (RIN4) complex from *A. thaliana*, target of AvrRpt2 of *Pseudomonas syringae* (Liu *et al.*, 2009). Apart from Y2H and protein pull-down assay, both of which are able to identify several interacting partners for a particular protein, several techniques such as bi-fluorescence complementation assay (BiFC), fluorescence resonance energy transfer (FRET) assay etc., are generally employed to test the interaction of one protein with another. In these techniques, proteins for which interaction is being tested, are fused with two different tags/fluorophores. For example, in case of BiFC, the two proteins are fused to two halves of fluorescence protein (such as YFP) so that there is YFP fluorescence only if the two proteins interact. All the above mentioned techniques have been successful in identification/confirmation of several protein interactions. However, these techniques have several application and efficiency based limitations. Techniques such as BiFC are successful in confirmation of putative interactions but they cannot be applied for detection of protein interactions in a high-throughput manner. On the other hand, affinity purification that helps in identification of protein complexes is based on an *in vitro* assay carried out on cell lysate where protein interactions lose their spatial integrity.

Another limitation of these techniques is the lower probability of detecting transient/weak interactions due to stringent washing steps required to eliminate false positive interactions. Contrary to affinity purification, Y2H is an *in vivo* technique but faces problems of detecting high percentage of false positive and false negative interactions due to heterologous post-translational modification machinery and overexpression of protein of interest, leading to detection of unspecific/false interactions. Moreover, Y2H system does not maintain spatial integrity of interactions as proteins are assessed for their interaction inside yeast nucleus, which for most of the

proteins is a non-native environment for interaction. Also, proteins with localization tags to other cellular compartments might not be transported to nucleus and thus they would remain undetected by the system. In addition to these problems, the currently available techniques have not been very successful in detecting interactions for membrane-bound proteins, proteins present in low-abundance and weak/transient interactions.

1.2.2 Biophysical methods

The study of the interactions between proteins is covered not only by biochemical methods: the developing of biophysical tools has drastically increased, so that biophysical techniques became an important branch of the study of proteins interactions.

Amongst them, FRET similarly to BiFC, uses two proteins fused to two fluorophores, one of which emits fluorescence when being excited by emission wavelength of the other fluorophore, when both are present in very close physical proximity, thus serving as an indicator of interaction between the proteins being attached to them.

It was used to show specific interaction between the two photoreceptors, phytochrome B (phyB) and cryptochrome 2 (cry2), in *Arabidopsis* living cells occurring in flowering time, hypocotyl elongation and the light-dependent translocation of the complex to the nuclear speckles (Más *et al.*, 2000). Similarly to BiFC, FRET cannot be applied for detection of protein interactions in a high-throughput manner. Differently from FRET, which needs the proteins to be tagged, the dual polarisation interferometry (DPI) does not require tags but one of the two proteins being investigated must be immobilized on the surface of a waveguide. Two light beams will pass through the waveguide producing an interference pattern. The presence of the interacting protein causes one of the light beams to form an altered polarization, thus changing the pattern, which is registered by measuring the refractive index and reflects a specific mass, density and thickness of the protein in real time (Freeman *et al.*, 2004).

Another recent biophysical technique is the microscale thermophoresis (MTS), similarly to DPI, it permits to detect conformation, charge and size changes of a complex. The method is based on the registration of motion of the molecules along microscopic temperature gradients. Differently from DPI, it does not need an immobilized protein,

so, the measurements are closer to native conditions. Even if this method is highly specific, it shows a low sensitivity; it can reveal a lot of interactions but wrong predictions as well.

Amongst the biophysical techniques to detect PPIs, one of the most popular techniques is nuclear magnetic resonance (NMR). Currently many NMR-based methods have been developed; the basic principle is represented by magnetically active nuclei orientated by a strong magnetic field that absorbs electromagnetic radiations at specific frequencies. The characteristics of these frequencies depend on the chemical environment, influenced mostly by chemical bonds and molecular conformations. An example of the use of this method is represented by the partially characterization of the protein interactions of phytochrome A (phyA) in *Avena sativa* (Song *et al.*, 2012). The advantage of this technique lays in its capability of capturing very weak interactions.

1.3   Upcoming techniques

New methods that can solve the limitations of existing techniques are currently being developed. One of these techniques is based on label transfer technology, in which bait protein is modified so as to covalently transfer a detectable probe on the proteins interacting with it. An example is given by the use of 3,4-dihydroxyphenylalanine-biotin-4′,5′-bis

(1,3,2-dithioarsolan-2-yl)fluorescein (DOPA-biotin-FlAsH) coupled with a tetra-cysteine sequence which represents the FlAsH receptor. Tagging a protein of interest with this sequence results, after the periodate oxidation, in the transfer of the biotin tag to the interacting protein. The proteins will not need a covalent modification (Liu *et al.*, 2007; Fancy, 2000).

Many varieties of cross-linking reagents had been developed in the past years; few examples are nickel-peptide chelate which proved the hydrophobic interactions between F-actin and skeletal myosin sub-fragment 1 (Bertrand *et al.*, 1997), oxidatively-activated manganese and iron porphyrins (Campbell *et al.*, 1998) and photogenerated oxidant for a photo-induced cross liking (Fancy and Kodadek, 1999).

Several modifications of these techniques are currently available and being developed further. Amongst these are several modifications of biotinylation based methods.

1.4   Biotinylation reaction

Biotinylation is a biochemical process that allows the attachment of a small molecule, biotin, to a protein, nucleic acid or to another molecule (Samols *et al.*, 1988) (Fig. 1) and it is catalyzed by biotin protein ligase (BPL) family proteins. BPL family is composed of two groups: group I ligases, without an N-terminal DNA-binding domain and group II ligases (such as BirA) with an N-terminal DNA-binding domain (Mukhopadhyay *et al.*, 2001). BirA represents the BPL of *E. coli* and carries out biotinylation reaction in two steps: in the first step biotin interacts with ATP to form the activated biotin called biotinoyl-5`-AMP (bio-AMP), while releasing a pyrophosphate. This bio-AMP remains hidden in BirA's active site until it reacts with a specific lysine residue (Lane *et al.*, 1964) of the biotin acceptor tag (BAT) sequence, which is a 15-amino acid consensus sequence (GLNDIFEAQKIEWHE) (Beckett *et al.*, 1999), thus transferring biotin to the ε-amino group of the lysine residue in the target protein.
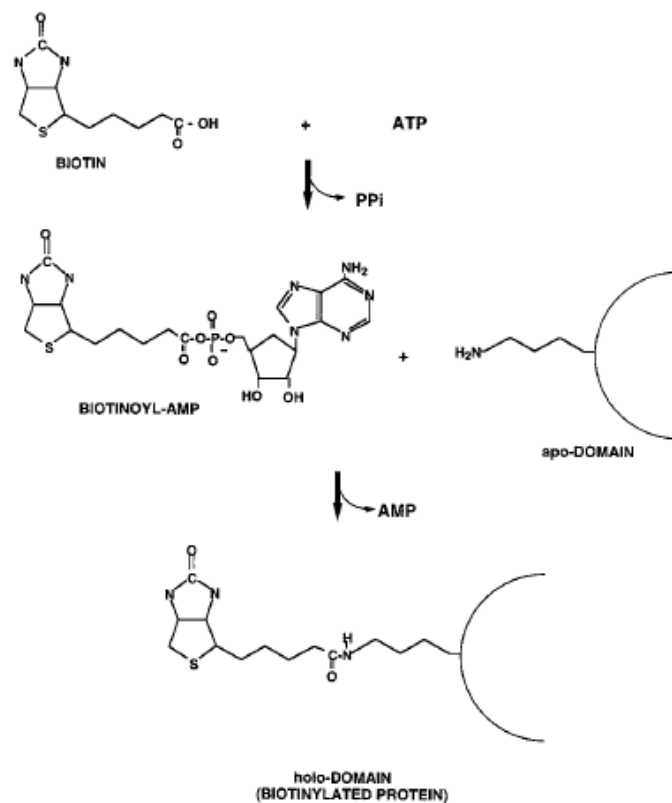
Fig. 1. Biotinylation reaction (Chakravartty and Cronan, 2012)

BirA is a monomeric 35 kDa protein composed by 321-amino acids arranged in three different domains (Wilson *et al.*, 1992): N-terminal domain, central catalytic core, C-terminal domain (Fig. 2). The first domain is responsible for the DNA-binding capacity and contains an N-terminal winged helix-turn-helix (wHTH) structure. This kind of domain is common for a large number of bacteria and represents the conservation of the transcription repression function in ligases (Beckett, 2007). The central domain binds ATP and biotin to catalyze bio-AMP synthesis. It has a sequence that is well conserved between all of BPL members. The C-domain is involved in DNA binding, dimerization and catalytic function (Streaker and Beckett, 1999; Chakravartty and Cornan 2012). Enzymatic biotinylation is an extremely specific reaction and only one protein in *E. coli* (biotin carboxyl carrier protein subunit of acetyl-CoA carboxylase) and four proteins in *A. thaliana* (homomeric acetyl-CoA carboxylase, heteromeric acetyl-CoA carboxylase, geranoyl-CoA carboxylase, and methylcrotonyl-CoA carboxylase) are biotinylated in nature. All of these proteins are key enzymes in metabolic pathways and biotin represents an important cofactor. Furthermore, biotin also seems to be involved in gene regulation.
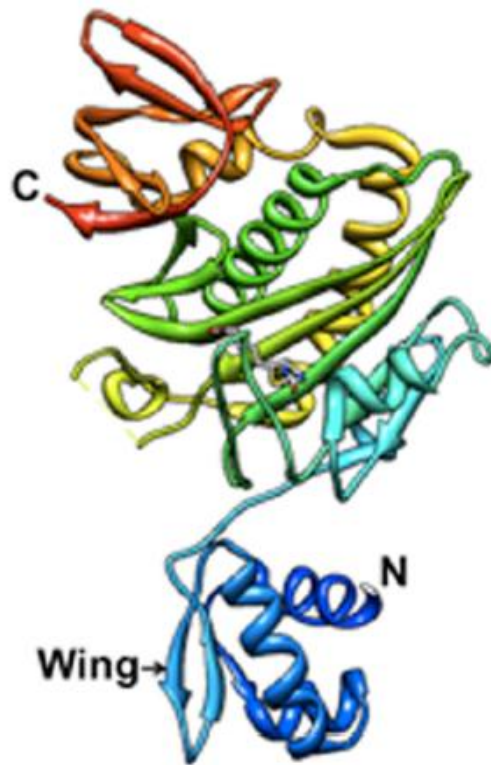

Fig. 2. BirA structure (Chakravartty and Cronan, 2013)

Biotinylation is a highly regulated process. Kinetic studies and thermodynamic measurements revealed that the affinity of the central domain is higher for biotin compared to ATP. It is suggested that biotin binds the domain as first, leading to reorganization thus facilitating ATP binding (Wilson *et al.*, 1992). Biotinylation reaction is also regulated at transcriptional level. In presence of an excess of biotin, BirA dimerizes and binds the biotin biosynthetic operon (bioABFCD) leading to the inhibition (repression) of biotin synthesis. If there is, instead, a lack of biotin or an excess of acceptor proteins, the dimeric form will not be formed, the operon will be active and it will transcribe biotin (derepression) to carry out biotinylation reaction (Fig. 3).
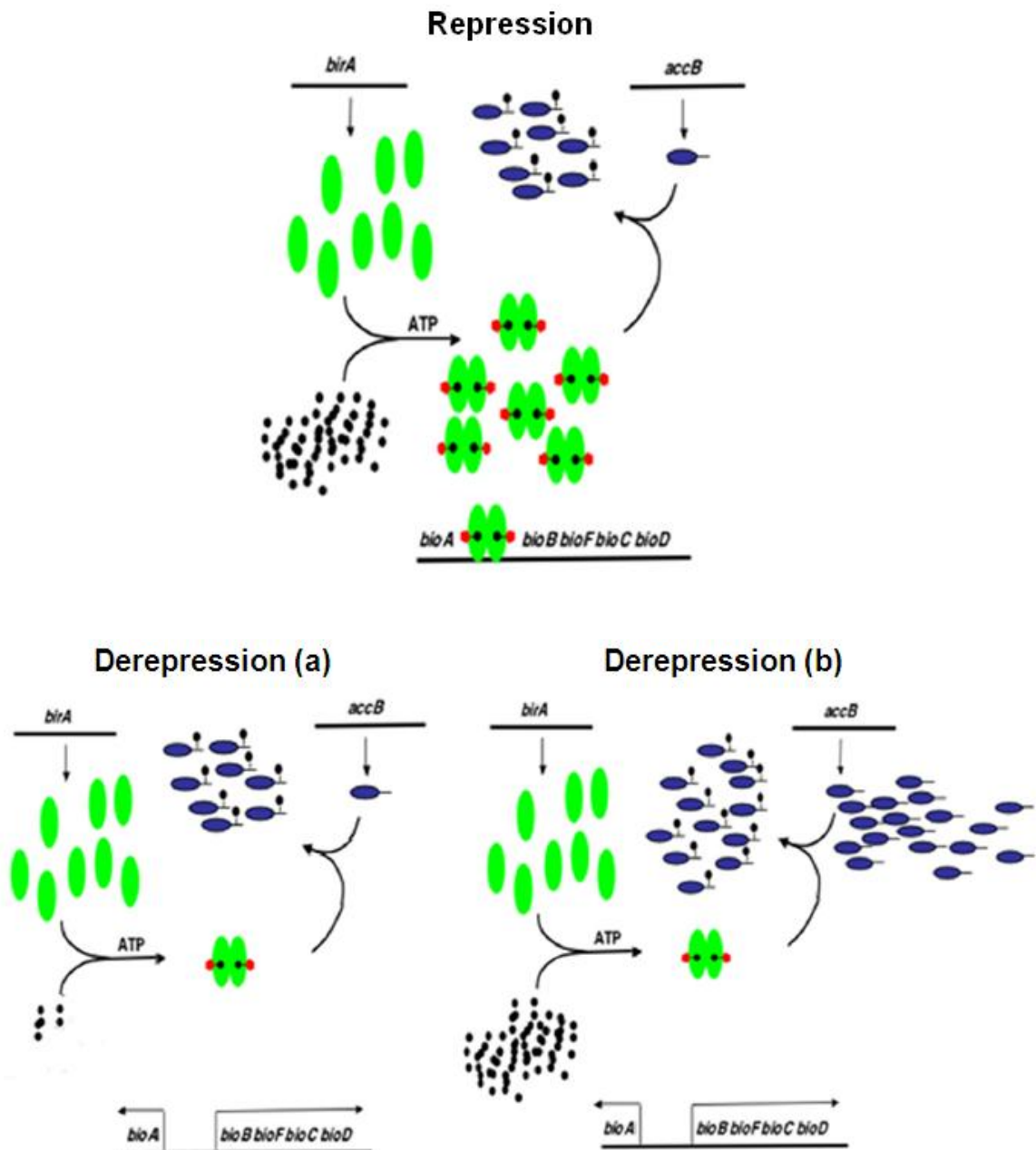
Fig. 3. Biotin transcription regulation. BirA is represented by the green ovals, the black dots are biotin, the reds are bio-AMP and the blue tailed ovals are the natural acceptors, AccB. Repression occurs when biotin is abundant and this leads to BirA dimerization. In the dimeric form BirA binds biotin operon and biotin synthesis is stopped (repression). This will not occur when there is a lack of biotin or an excess of acceptor protein, so that there is biotin transcription (derepression) (Chakravartty and Cronan, 2013).

1.5 Current techniques involving proximity dependent biotinylation

The current techniques utilizing biotin labeling to detect protein interactions are based on the specific biotinylation of BAT sequence by BirA. BirA is able to recognize this minimal recognition sequence with a high specificity and biotinylate it. In several variations of this method, BAT sequence and BirA are fused to two proteins under investigation. If they interact, then BirA would come in proximity of the BAT sequence and thus biotinylate it (Fig. 4). Some of these methods are described below.
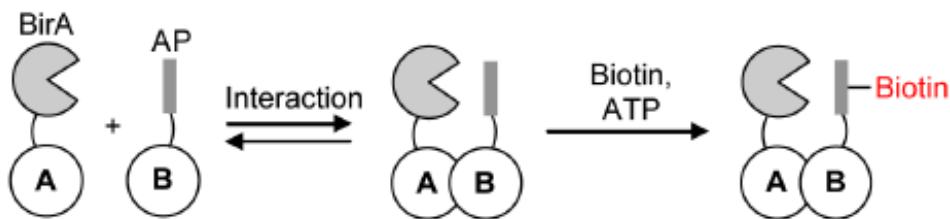

Fig. 4. General model of BirA use to detect protein-protein interaction. BirA could be fused to a protein A and the acceptor peptide (AP) to a protein B that could interact with BirA. The interaction will be detected by the biotinylation of the acceptor peptide (Fernández-Suárez *et al.*, 2008).

1.5.1 INTACT

Isolation of Nuclei TAgged in specific Cell Types (INTACT) has been developed to isolate and purify cell-type-specific nuclei from the total pool of nuclei in a tissue. This method uses the BirA expression coupled with the expression of a nuclear targeting fusion protein (NTF). The fusion protein has to be composed by three different parts: a protein domain which allows the nuclear envelope association, a portion that allows the visualization (green fluorescent protein, GFP), and, finally, the biotin ligase recognition peptide (BAT) (Fig. 5). The expression of NTF and BirA in the same cell type leads the biotin-labeling of the nuclei only in those cells, which can subsequently be purified using streptavidin mediated affinity chromatography.

The same strategy has been successfully used in *Arabidopsis thaliana* (Deal and Henikoff, 2011), in *Caenorhabditis elegans* and in *Drosophila melanogaster* (Steiner *et al.*, 2012).
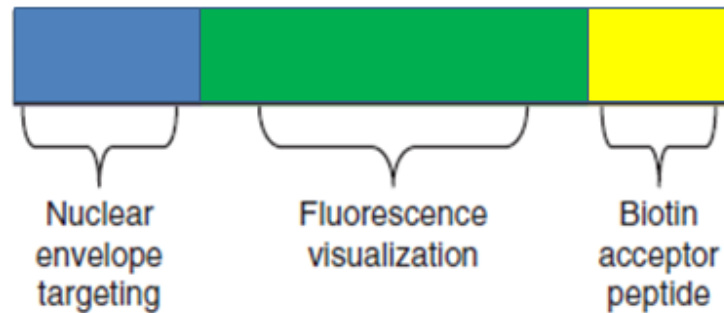

Fig. 5. The nuclear targeting fusion protein (NTF). The blue part is the nuclear domain; the green one represents GFP and the yellow, the peptide that will be biotinylated by BirA. (Modified from Deal and Henikoff, 2010)

1.5.2 PUB-NChIP

Proximity Utilizing Biotinylation with Native ChIP (PUB-NChIP) has been developed to detect the interaction between chromatin and a protein of interest. It is based on chromatin immunoprecipitation (ChIP), widely used to study *in vivo* association between a DNA sequence and a regulatory protein. In this strategy, *E. coli* BirA is fused to a protein of interest (transcription factor or another nuclear protein) and is co-expressed with a BAT-tagged histone. Under these conditions the histone will be incorporate into chromatin and BirA will biotinylate it every time the protein of interest interacts or will be close to it (Shoaib *et al.*, 2013). The labeling gives the possibility of isolating DNA associated with the histone subpopulation in proximity to the protein of interest. It is called native ChIP because exploits the strong interaction established between core histones and DNA, escaping the risk of cross-linking during the purification (O'Neill and Turner, 2003; Ooi *et al.*, 2010). With this modified approach it is also possible analyze the protein which interacts with the purified DNA and not only the DNA portion of the complex. Despite this seems that its resolution is lower compared to ChIP (Shoaib *et al.*, 2013). The methods described above utilize specific

biotinylation reaction under *in vivo* conditions and can overcome limitation of identifying false positives and can detect transient and weak interactions. However, their application in identifying PPIs is still limited to identifying one-to-one interactions. Another modification in the BirA dependent biotinylation system called promiscuous biotinylation, however, seems to overcome this limitation.

1.6   BirA dependent promiscuous biotinylation

BirA mediated biotinylation reaction is a highly specific reaction which occurs between BirA and BAT sequence of the acceptor protein. The specificity is mainly due to the structure of the central domain of BirA, which after formation of bio-AMP in its active site, restructures itself, thus preventing leakage of bio-AMP into vicinity and also prevents access of non-specific proteins to the active site. The analysis of X-ray crystallographic structure of BirA showed that a loop like structure within active site controls the sequestration ability of the bound bio-AMP. This loop appears disordered in native state of protein and more ordered when biotin occupies the active site (Weaver *et al.*, 2001). It has been shown that in two mutants, G115S and R118G, with a defective binding of biotin and bio-AMP, but not ATP, this loop is altered. The same altered bound capacity has been shown in BirA Δ1-34, protein lacking the N-terminal domain. In these mutants the biotin bound is weak but nevertheless, they are able to catalyze the formation of bio-AMP. Since the dissociation constants of these mutants for bio-AMP is greater compared to the wild-type, this molecule is released immediately (Kwon and Beckett, 2000; Xu and Beckett, 1996).

Moreover, it has been described that one of this mutant form of BirA (BirA*, R118G), which shows a dissociation constant for biotin 100-fold greater than the wild-type and for bio-AMP 400-fold greater, is getting itself biotinylated *in vivo*. This level of self-biotinylation increases with the enzyme concentration. This mutant acts exactly as a promiscuous biotinylation reagent, since it was found that it biotinylated not only itself, but also several different proteins that normally do not carry this modification (Choi-Rhee *et al.*, 2004). Two hypothesis have been proposed to explain the promiscuous biotinylation of these mutants: it could be due to the chemical acylation by bio-AMP released from the mutant's active site or, alternatively, to the capacity of the lysine side

chain of a promiscuous acceptor protein to access to the active site of the enzyme due to disordering caused by the mutation. The first hypothesis is the preferred as it seems highly unlikely that BirA* active site could rearrange itself to allow a generic lysine residue side chain the access, without losing the ability of sequestering bio-AMP (Choi-Rhee *et al.*, 2004). Thus if BirA* is introduced in a cellular system, it would retain its catalytic activity to synthesize bio-AMP from biotin and ATP, but it would release bio-AMP in its vicinity soon after its formation. The released bio-AMP, due to its reactive nature, should then be able to biotinylate any protein that comes in its close vicinity. To develop this characteristic as a technique to identify PPIs, if BirA* is fused to a protein of interest, the proteins present in its proximity will be biotinylated (Choi-Rhee *et al.*, 2004). This includes the proteins that interact directly or indirectly (via complex formation) with the protein of interest. Thereafter, utilizing the extraordinary affinity between biotin and avidin/streptavidin ($Kd = 10^{-14}$ M; Green, 1963), simple affinity purification would allow isolation of the tagged proteins in a rapid and easy manner. The formation of this bond is extremely rapid and is a permanent modification that persists through the life of the protein and once formed is not affected by variables such as pH, temperature, organic solvents and other denaturing agents.

This technique, if employed to the living system, should be able to overcome all the limitations encountered by the current techniques as it is an *in vivo* technique, that can maintain spatio-temporal integrity of the interactions, also can be utilized to identify whole complexes interacting with a protein including transient and weak interactions. Also, this technique is suitable to be extended to low expressing or membrane bound proteins. One could, however, contemplate about few possible problems associated with the technique. First, leakage of bio-AMP in a large cellular area rather than restricted vicinity could result in biotinylation of non-specific proteins leading to false positives. However, bio-AMP as such acyl-adenylate is an unstable molecule and, as a mixed anhydride, is highly labile to hydrolytic degradation. Thus even if the active molecule is diffused, it would be rendered ineffective soon after its leakage. Secondly, biotinylation being a very specific modification in nature could render the non-native biotinylated protein inactive, thus affecting its interactions. However, since biotin is a very small molecule, even smaller than a globular protein, it should not significantly interfere with the normal activity of the protein. Moreover, the small size permits many biotin molecules to bind to the same protein, thus allowing a better detection. Third problem could be associated with the control of the reaction. Biotin is always present in a cell

and can thus increase the bio-AMP cloud in an uncontrollable manner. However, Choi-Rhee *et al.* (2004) have shown also that biotin and ATP are necessary for the activity of this enzyme. Also, experiments by Roux *et al.* (2012) revealed that biotin level primarily regulates the biotinylation reaction; and its concentration in a standard complete media and thus in cellular environment is not sufficient to cause the reaction. Therefore, to build up the bio-AMP cloud, external supplementation of biotin is necessary. Based on this principle, proximity based biotinylation technique called BioID has been developed in yeast.

1.6.1 BioID

BioID, uses the modified form of *E. coli* biotin ligase (BirA*) to tag proteins. As mentioned above, the fusion between this enzyme and a protein of interest would allow identification of the other proteins that come in contact with it (Fig. 6). This methods has been used in mammalian cells by Roux *et al.* (2012), applying BioID to an intermediate filament protein in the nuclear lamina, lamin-A (LaA), and, by Morriswood *et al.* (2013) in *Trypanosoma brucei*, a protozoan parasite, to study its cytoskeletal structure exploiting the TbMORN1 protein. This technique allows to detect potential candidate interactors in their natural environment, included the membrane protein. Since the biotinylation occurred before solubilization, it is enough to detect even weak interactions.

It has been shown that this biotin supplementation has an effect only on BirA* and not on the wild-type form of the enzyme; in fact it was detected a massive stimulation of promiscuous biotinylation after the biotin addition (Fig. 7) (Roux *et al.*, 2012).
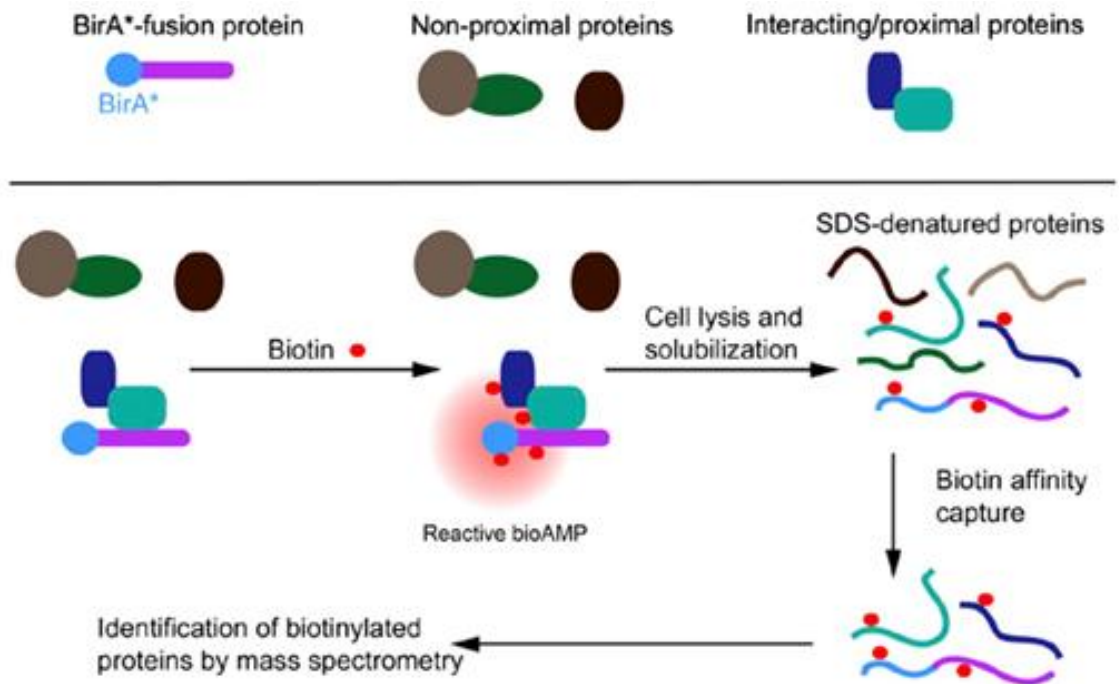
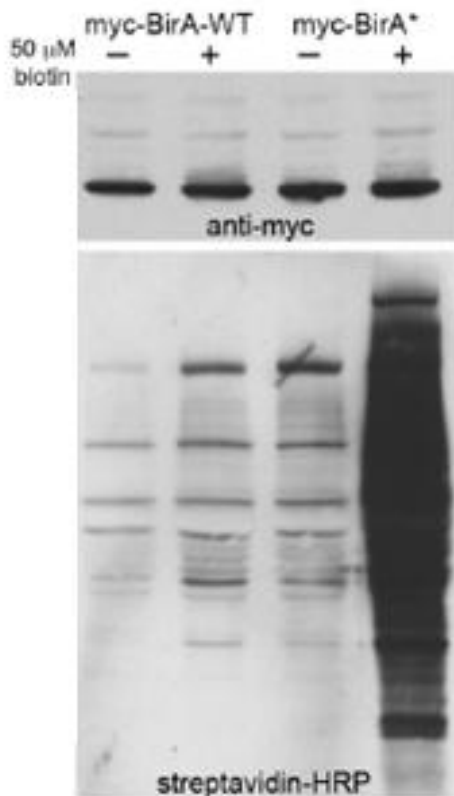Fig. 6. Model of BioID (Roux *et al.*, 2012)

Fig. 7. Comparison between HeLa cells transfected with myc-BirA-WT or myc-BirA* with and without biotin supplementation (Roux *et al.*, 2012).

1.7   Aim of the project

The above mentioned technique till date has only been applied in yeast or animal system. The project which I was a part of, aimed to transfer and further develop this method to detect PPIs in plants, the originality and the innovation of this, lies in its utilization. In addition to utilizing of this technique to identify PPIs, the project aims to extend the same principle to prepare an inventory of the proteins present in a particular cellular compartment. To achieve this, compartment localization tag would be fused to N-terminus of BirA* so that it is targeted to the compartment, where upon supply of excess of biotin, the whole compartment would essentially be filled with bio-AMP cloud, thus biotinylating all the proteins present in this compartment, without affecting proteins of any other compartment. This technique has advantage over the existing technique of sub-cellular fractionation that is generally employed to identify organelle proteome inventories, that it would be less time and skill consuming and would also prevent the problem of cross-contamination of proteins from other compartments, a problem most generally faced in sub-cellular fractionation technique.

Prior to its transfer to the plant system, it was essential to ascertain if the enzymatic characteristics of *E. coli* BirA* protein, allowed it to be utilized in plants. I was assigned to work on the *in vitro* characterization of protein activity under different physiological conditions. To achieve this, the gene was codon optimized for the expression in plant. The expression of prokaryotic proteins in other organisms usually leads to lower translation efficiency, so, some rare codons need to be replaced with others more frequent in plant cells. After that, the codon optimized gene was expressed in *E. coli* cells and purified. The purified protein was used in different assays to test the performance of protein activity under different conditions. The results indicate that the protein could be utilized for expression in plant system, and probably would be amenable to technology transfer to other crop species in future, thus opening new avenues in the field of proteomics.

2. Materials and Methods

2.1 Cloning of *BirA* and GST in protein expression vectors

Wild-type *BirA* (EG10123) from *E. coli* with R118→G mutation was codon optimized (from here forth referred to as BirA*) to allow efficient translation of the mRNAs in *Arabidopsis thaliana*, and designed to express the enzymes with C-terminal HA-epitope and FLAG tags to allow easier detection of the protein via immunoblotting. The synthesized gene was PCR amplified with high fidelity polymerase (Phusion Taq, Thermo Scientific, Hudson, NH, USA) using *BirA* specific primers (forward primer: CACCAAGCTTAGCCTCGAGAGCC, reverse primer: GCTAGCGCTGGTACCGCTGAATTC) and, cloned in pENTR™/SD/D-TOPO® vector (Invitrogen, Life Technologies, Grand Island, NY). The construct was subsequently sub-cloned in Gateway® Nova pET-60-DEST™ vector (Merck Millipore, Merck KGaA, Darmstadt, Germany) using LR clonase reaction (Invitrogen), as per manufacturer's instructions. In order to purify GST protein as control, Gateway® Nova pET-60-DEST™ vector was modified by removing *ccdB* gene, lethal to most *E. coli* strains, using BamHI and PstI restriction enzymes (FastDigest, Thermo Scientific) followed by agarose gel purification using Zymoclean™ Gel DNA Recovery Kit (Zymo Research, Orange, CA) and nick filling (to generate blunt ends from the sticky ends) using Klenow Enzyme (Roche, Mannheim, Germany) as per manufacturer's instructions. The vector was then re-circularized by ligating the blunt ends using T4 DNA Ligase (Roche) as per manufacturer's instructions and transforming in *E. coli* DH5α competent cells (Invitrogen). The clones from both constructs were verified by sequencing analysis and transformed to Rosetta BL21(DE)pLysS chemically competent cells (Invitrogen). Briefly, 50 μL of Rosetta BL21(DE)pLysS competent cells were thawed on ice for 10 minutes and then incubated with 100 ng of purified pET-60-DEST™-BirA* or modified Gateway® Nova pET-60-DEST plasmid for another 10 minutes on ice, following which the cells were heat shocked by incubating in 42°C water bath for 30 seconds to allow the DNA enter into the cells and immediately incubated on ice for 2 minutes. The heat-shocked cells were further incubated for 1 h at 37°C at 200 rpm after addition of 1 mL LB medium. The transformation mixture (100 μL) was spread on LB plates with ampicillin (50 μg/mL, selection for the vector) and

chloramphenicol (50 μg/mL, selection for the competent cells) and incubated overnight at 37°C for selection of positive transformant, which were further verified by colony PCR. The PCR reaction mix composed of 0.25 mM dNTPs mix, 0.5 μM BirA* specific or Gateway® Nova pET-60-DEST, specific forward primer and reverse primers, 1.5 U DreamTaq DNA Polymerase (Thermo Scientific), 1X DreamTaq Buffer (containing MgCl$_2$) and single bacterial colonies (picked by toothpick) was cycled for initial denaturation at 95°C for 10 min, 30 cycles of 95°C for 30 s, 58°C for 30 s, 72°C for 1 min and final extension at 72°C for 5 min. To check the PCR products, 5μL of reaction was loaded and run on 1% agarose (Sigma-Aldrich, Saint Louis, MO) gel, containing 5% GelRed™ (Biotium Inc., Hayward, CA) added for visualization of DNA bands and run at 100V till bands of DNA ladder (Thermo scientific) resolved and was visualized in UV-transilluminator 75S/02701 (Bio-Rad, Hercules, CA).

2.2 Expression analysis

A single PCR positive colony was used to inoculate 10 mL LB medium supplemented with ampicillin (50 μg/mL) and chloramphenicol (50 μg/mL), incubated overnight at 37°C on a shaker at 200 rpm and used as a primary inoculum to inoculated secondary culture (1%) supplemented with same antibiotics. The secondary culture was grown at 37°C at 200 rpm till the OD$_{600}$ 0.6-0.8, followed by the induction of the culture with 1 mM isopropyl ß–D-1-thiogalactopyranoside (IPTG) (Roche) and the incubation at 28°C for 6 hours at 200 rpm. IPTG is a synthetic analog of lactose and it is used to induce expression of cloned genes under control of the lac operon (Sambrook and Russell, 2001).

At the end of incubation period, the induced culture was harvested by centrifugation (Eppendorf Centrifuge 5810R, Swing-bucket rotor A-4-81) at 8000 rpm at 4°C for 20 min, and stored at -20°C until further used.

2.3   Protein purification

2.3.1 Protein purification using histidine-tag

For protein fused with 6x Histidine tag, nickel immobilized metal affinity chromatography (IMAC) was performed due to highly selective affinity of hexa-histidine for divalent metal ions. The purification was either performed manually with Ni-NTA agarose beads (Quiagen, Hilden, Germany) and HisPur Cobalt Spin Columns (Thermo Scientific) or using automated ÄKTA protein purification systems. For manual purification, the bacterial pellet was re-suspended in 10 mL of sonication buffer and incubated with 1 mg/mL lysozyme for 30 min at 28°C. The re-suspended culture was sonicated (Sonopuls homogenisator; Bandelin, Berlin, Germany), while being incubated on ice, using 3 cycles of 1 minute sonication at 70 amplitude and 1 minute of ice incubation to avoid heat induced protein degradation. The sonicated solution was then centrifuged (14000 g, 4°C, 2 min) and the supernatant was collected and used for protein purification. Prior to use, the Ni-NTA agarose beads were washed thrice with 3 bed volumes of native buffer (see appendix I). The washed beads were then incubated with the bacterial crude protein extract for 1 h at 4°C on a rotary shaker. After incubation, the unbound proteins were removed from column by two washes of 10 bed volumes of sonication buffer (see appendix I) and other two washes of 5 bed volumes of sonication buffer containing 20 mM imidazole. For protein elution, a step gradient of imidazole (40 mM, 100 mM, 150 mM, 300 mM, 500 mM, 1 M) was used and different fractions were analyzed using sodium dodecyl sulphate – polyacrylamide gel electrophoresis (SDS-PAGE) followed by Coomassie/Silver staining and western blot analysis.

In case of purification with HisPur Cobalt Spin Columns, the resin was washed once with 2 bed volumes of equilibration/wash buffer (see appendix I) and then incubated with the bacterial crude protein extract for 1 h at 4°C on a rotary shaker. Afterwards, the unbound proteins were washed out through three washes with 2 bed volumes of equilibration/wash buffer and then the protein were eluted with 1 bed volume of elution buffer (see appendix I).

For automated protein purification, the bacterial pellet was re-suspended in 30 mL of IMAC-buffer A (see appendix I) containing 1X protease inhibitors cocktail (Halt™

Protease Inhibitor Cocktail EDTA-Free, Thermo Scientific), 2 mM benzamidine and 2 mM aminocaproic acid, sonicated and centrifuged to collect crude bacterial protein extract which was subsequently filtered through 0.45 μm cellulose acetate syringe filter (Whatman, Florham Park, NJ).

The GE ÄKTA Explorer (Amersham Pharmacia Biotech, Piscataway, NJ) chromatography system, UNICORN 5.01 software with HisTrap-HP affinity column (GE Healthcare, Munich, Germany) were used for this purification as per manufacturer's instructions. Briefly, 1 mL HisTrap-HP affinity prepacked column (Ni^{2+} immobilized on sepharose matrix) was washed with 2 column volumes of buffer A at flow rate of 1 mL/min. After washes, the protein was eluted by a linear imidazole gradient from 20 mM to 500 mM. The protein peak, which appeared at 370 mM imidazole concentration was collected and applied to the Superdex G-25 desalting column (HiTrap-Desalting; GE Healthcare) in order to change the buffer and remove contaminating salts. After desalting, the protein was eluted in 20 mM sodium phosphate buffer pH 7.

2.3.2 Protein purification using GST tag

For purification using GST tag, the bacterial pellet was sonicated and processed as mentioned above. The crude bacterial protein extract was subjected to two different types of purification, one, with magnetic GST beads (MagneGST™ Protein Purification System; Promega, Madison, WI) and another by Glutathione Sepharose 4B (GE Healthcare).

In the first case, the crude bacterial protein extract was incubated with magnetic GST beads for 1 h at 4°C under constant shaking. The beads were washed thrice using binding/wash buffer provided by manufacturer and eluted in three fractions of 100 μL elution buffer (see appendix I).

For purification with glutathione sepharose 4B, the resin was washed twice with 10 bed volumes of PBS (see appendix I) and incubated with crude bacterial protein extract for 1 h at 4°C on a rotary shaker. Before the elution, the unbound proteins were removed by three washes with 10 bed volumes of PBS. The elution was performed by three elutions with 1 bed volume of elution buffer (see appendix I).

The fractions were checked on 12% SDS-PAGE.

2.4   SDS-PAGE analysis

SDS-Polyacrylamide (PAGE) were cast as described by Harlow and Lane (1988). The composition of resolving and stacking gels is described in appendix I. Prior to loading in gels, each sample was mixed with 1X Laemmli buffer (see appendix I), immediately boiled for 10 min and centrifuged at 13000 rpm for 15 min. The gels were run at 80V till the loading front entered the resolving gel and at 120V till loading front reached the bottom of the resolving gel. The gels were then carefully removed from the glass plates and used for either Coomassie/Silver staining or western blot analysis.

Gels stained through Coomassie brilliant blue staining were incubated for 1 h to overnight with Coomassie solution at gentle agitation and then subjected to de-staining solution for three hours. Fresh de-staining solution was added every hour.

Silver staining procedure was carried out using silver staining plus kit (Bio-Rad) as per manufacturer's instructions.

2.5   Western blot analysis

Western blot analysis was performed to confirm presence of proteins of interest in bacterial protein extracts and purifications.

The proteins were resolved on 12% SDS-PAGE and then transferred on PVDF membrane (Immobilon-P, Merck Millipore) using Trans-Blot® SD Semi-Dry Transfer Cell (Bio-Rad). Briefly, the PVDF membrane (after 5 min incubation in 100% methanol) and blotting sheets were incubated in transfer buffer (see appendix I) for 2 min and assembled in the semi-dry transfer apparatus as per manufacturer's instructions. The proteins were transferred on to PVDF membrane by applying 18 V for 1 h. The membrane was blocked by incubating in 3% blocking solution (3% Blotting Grade Blocking, Bio-Rad) in TBST buffer (see appendix I) for 1 h to overnight under constant shaking at 75 rpm at room temperature. The membranes were washed once with TBS (see appendix I) for 5 min and thrice with TBST for 5 min each, followed by antibody incubation 1:12,500 dilution of streptavidin Horseradish Peroxidase (HRP) Conjugate (Invitrogen), 1:20,000 dilution of anti-GST (Pierce, Thermo Scientific) as primary antibody followed by 1:20,000 dilution of HRP conjugated anti-rabbit antibody

(Sigma-Aldrich) in TBST and then they were subjected to several washing steps in TBS and TBST buffers. Detection of HRP conjugated antibody (and hence protein of interest) was carried out using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) for 5 min in dark followed by exposure to Amersham Hyperfilm ECL (GE Healthcare) in dark room and processed with an automated film developer (Hyperprocessor, Amersham Pharmacia Biotech, Piscataway, NJ) for variable times according to the signal strength.

2.6   Protein quantification

The protein concentration measurements were performed by Bradford assay (Bradford, 1976).

10 µL of protein standards/protein sample was mixed with 200 µL of 1:4 dilution (1 part Bradford reagent, 4 parts Milli-Q water) Bradford Reagent (Bio-Rad) in a 96 well microplate and incubated at room temperature for 30 minutes. The absorbance was measured at 595 nm using Synergy™ HT multiwell plate reader (BioTek Instruments, Winooski, VT) and Gen5 2.04 software. The protein concentration was determined by comparison with a standard curve prepared with BSA samples of known concentration ranging from 0.05 to 0.5 mg/mL for BirA*-GST and from 0.1 to 1 mg/mL for GST.

2.7   Protein activity assays

The activity of BirA*-GST protein was first optimized for different substrate concentrations (ATP and biotin concentration) and then tested under different physiological conditions, such as: time of incubation, temperature, and pH optima.

2.7.1   Optimizing substrate concentrations

First of all, the test to establish the concentration of biotin and ATP which enhance the protein activity was performed. We chose to test three different concentrations of ATP (5 mM, 10 mM, 15 mM) and three other for biotin (100 µM, 200 µM, 300 µM). To

enhance the solubility of biotin dimethylformamide (DFM) was used. All of the possible combinations were tested. The reaction was stopped by adding 1X Laemmli buffer to the reaction mix and boiling the mixture for 10 min. The samples were run in duplicate on 12% polyacrylamide gel and analyzed by silver staining and western blot analysis.

2.7.2 Reaction time

This experiment was carried out to estimate the minimum time required by BirA*-GST mediated GST biotinylation to reach its saturation. The reaction mixture was composed of 10 mM ATP (Thermo Scientific), 200 μM biotin (Sigma-Aldrich), 5.5 mM $MgCl_2$, 100 mM KCl, 0.4 μg BirA*-GST, 0.8 μg GST. Each reaction tested for proximity biotinylation had a volume of 12 μL. The reaction was incubated at 37°C for 0 h, 2 h, 4 h or 20 h, then it was stopped as mentioned above and analyzed by silver staining and western blot analysis.

2.7.3 Deducing temperature and pH optima for the reaction

To optimize temperature for the reaction, the reaction mixture composed of 10 mM ATP, 200 μM biotin, 5.5 mM $MgCl_2$, 100 mM KCl, 0.4 μg BirA*-GST, 0.8 μg GST was incubated at different temperatures for one hour. We tried first a wider range of temperatures (0, 4, 10, 16, 20, 28, 37, 42, 50°C) which were then narrowed to 28, 30, 35, 37, 40, 47, 50, 52°C and further to 35, 36, 37, 38, 39, 40, 41 and 42°C.

To assess the effect of different pH on the BirA* biotinylation, the reaction mixture (as explained above) was supplemented with either HCl or NaOH to achieve a pH range of 1 to 14 and incubated for 1 h at 37°C. The reaction mixtures were then inactivated by addition of 1X Laemmli buffer and analyzed by silver staining and western blot analysis.

3. Results

3.1   Protein expression

For the purpose of protein purification, they are generally overexpressed in bacterial system which can be a challenging task because several factors influence the process. The conditions for the expression have to be optimized for each protein, in order to obtain high yield of recombinant protein that are present in soluble form rather than inclusion bodies. The inclusion bodies are insoluble aggregates composed of densely packed denatured proteins (Kane and Hartley, 1988) and they are formed due to an inappropriate aggregation of partially folded or malfolded polypeptides (Mitraki and King, 1989), hence they represent an inactive form of the protein. As different proteins behave differently when subjected to induced overexpression, multiple experiments were set up to test different conditions of temperature and time of induction with IPTG. The conditions tested were: 3 h at 37°C; overnight at 20°C; 6 h at 28°C. After checking and comparing the uninduced and induced fractions, (separated in soluble and insoluble fraction, as mentioned in material and methods), 6 h induction at 28°C was chosen as the best induction conditions (Fig. 8) as under these conditions, there was an increase of the protein expression in the soluble fraction.
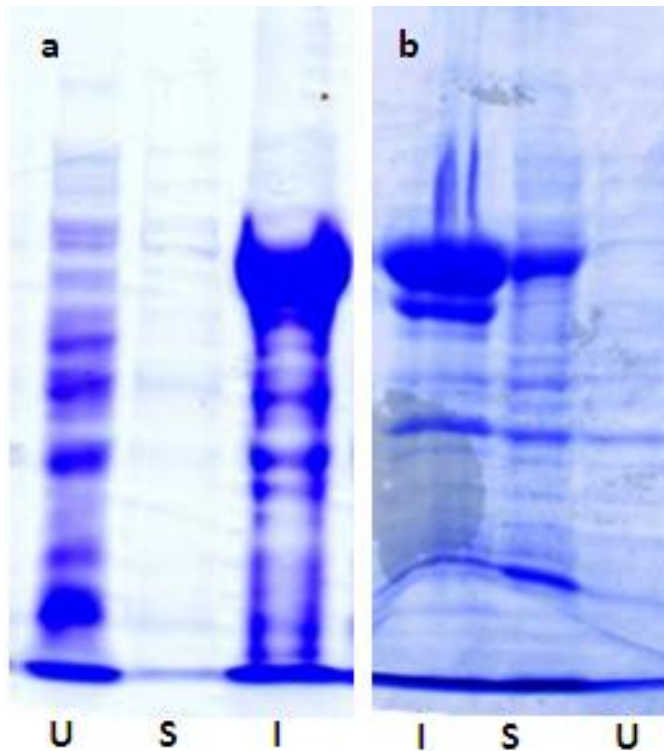
Fig. 8. Expression pattern of pET-60-DEST-CoBirA* under different conditions. (a) 3 hours incubation at 37°C. (b) 6 hours incubation at 28°C.

Respectively **U** is the uninduced fraction, **S** is the soluble part of the induced culture, **I** is the pellet from the induced culture.

3.2   Proteins purification

3.2.1   Batch purifications

Gateway® Nova pET-60-DEST™ vector used for BirA*protein expression carries two different tags fused in frame with the gene namely: N-terminal GST tag and a C-terminal 6x-His-tag. Both tags were used for the purifications.

The soluble part obtained after sonication was used to perform the purification. As a first attempt to purify GST-6X Histidine fused BirA* protein, batch purification with Ni-NTA agarose beads was tried. This resin selectively binds proteins with hexa-histidine tag. NTA has four chelating sites for nickel ions and histidine is the amino acid that exhibits the strongest interaction with immobilized metal ion matrices, as nickel. Imidazole was used as an eluting agent because it is able to compete with the polyhistidine tag for binding to the resin. Several concentration of imidazole (from 20 mM to 1 M) and variable salt (NaCl) concentrations were used for elution. However, high protein purity could not be achieved by this method. According to the literature, Ni^{+2} chelate resins seem to achieve high protein yields with suboptimal purity. Cobalt

25

instead, achieves high protein yield and purity with minimal optimization; it exhibits a more specific interaction resulting in less nonspecific interaction. Furthermore, cobalt resins display less metal leaching compared to Ni^{+2} resins. Therefore, we decided to switch to batch purification using cobalt spin columns. In this case as well, a protocol optimization with both salt and imidazole concentration was tried by increasing imidazole and NaCl concentrations in the washes to avoid the nonspecific interactions between proteins and the resins. These attempts did improve the purity level of protein, however, several impurities could still be detected (Fig. 9). The same results were obtained after a second purification, exploiting the GST protein fused to BirA*. We tried a second run of purification on the fractions coming from the two histidine purifications. Two different GST purifications were tested: one, using iron paramagnetic particles with reduced glutathione attached through its sulfhydryl group to the surface and encapsulated in cross-linked agarose and, another one, where the glutathione ligand was coupled via a 10-carbon linker to cross-linked 4% agarose. Glutathione is immobilized through its central carbon to preserve the essential structure necessary for efficient GST binding. In this case, the eluting agent used was glutathione as it competitively displaces the fusion protein bound to the resin, allowing it to be eluted. In this case as well, either the purity of the protein was not high or in the cases it was, the concentration of the protein was very low.
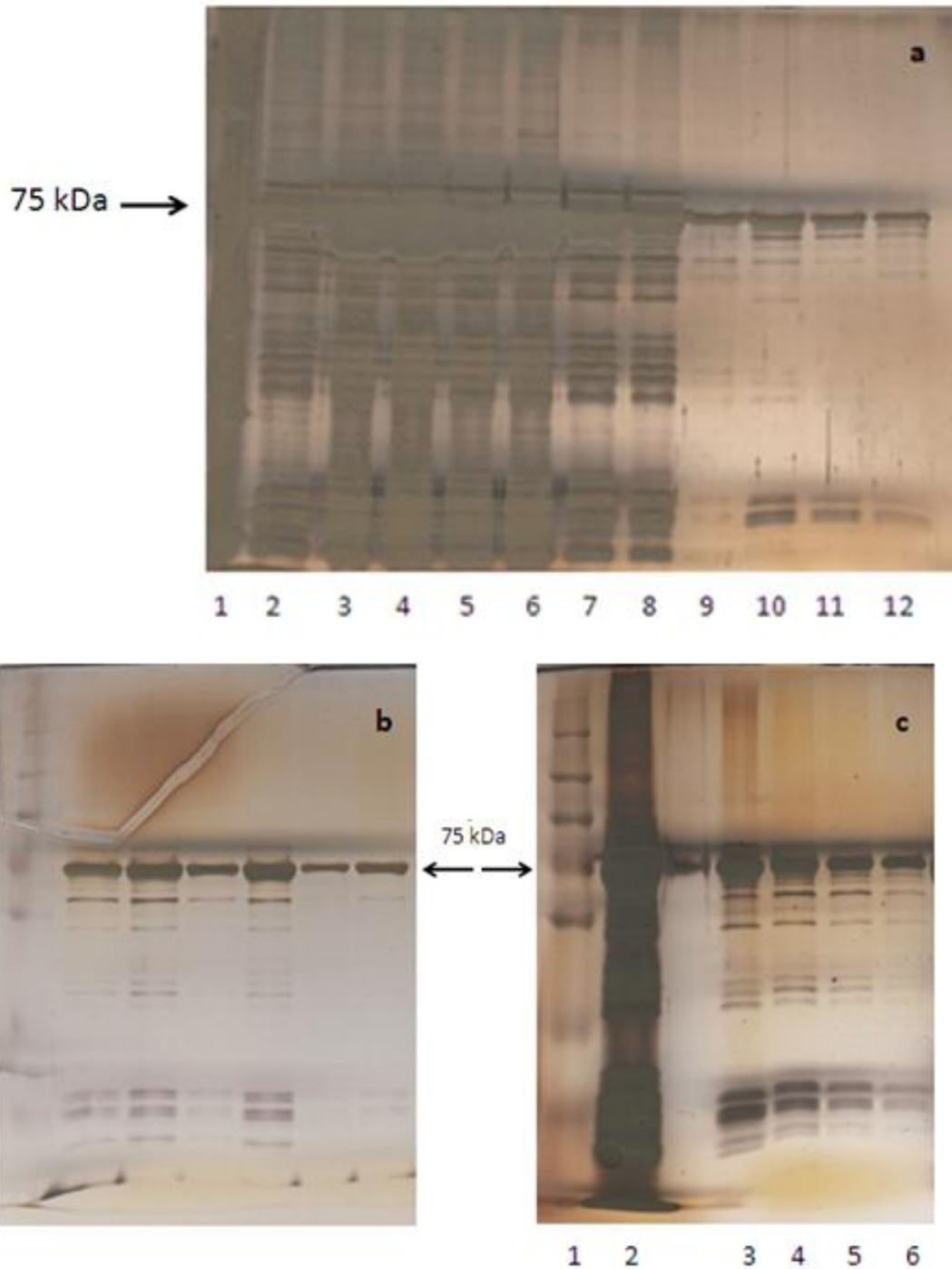
Fig. 9. Silver stained gels of GST-6X Histidine fused BirA*. (a) lane 1 marker, 2-9 elutions with a scale gradient of imidazole (100-150-300-500 mM) from Ni-NTA purification. Lane 10-12 elutions from GST purification using paramagnetic beads. The second purification was conducted on the sample in lane 2. (b) elutions with 150 mM imidazole from cobalt purification. (c) lane 1 marker, lane 2 elution from cobalt purification on which the second purification was based, lane 3-6 elutions from GST purification using sepharose beads.

3.2.2 Affinity chromatography

After attempts with batch purifications, a nickel immobilized metal affinity chromatography (IMAC) using the UNICORN™ controlled ÄKTAexplorer™ system was tried. The protein was eluted applying a linear gradient of imidazole and the peak appeared in correspondence to an imidazole concentration between 260 mM and 500 mM (Fig. 10). This peak included six fractions. These fractions were checked on 12% SDS-PAGE with silver staining (Fig. 11) and six of these were further chosen to be applied to the desalting column to prevent protein precipitation due to high ionic condition as well as to prevent interference of imidazole in subsequent assays. The fractions that appeared purer, albeit with a lower protein concentration (A12, B5), were mixed together. The same was done for the fractions more concentrated but with more impurities (B1, B2, B3, B4). After the desalting runs, two fractions of each pool were obtained. This run allowed to get rid of most of impurities, even though a loss of sample material was registered (Fig. 12).
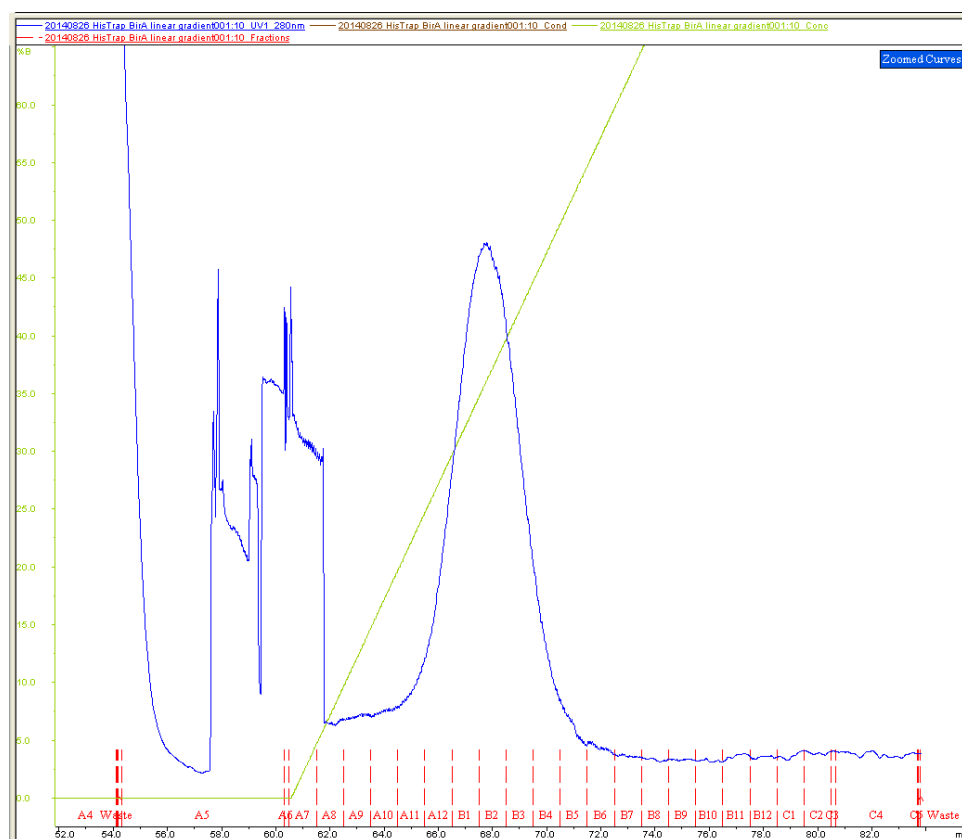

Fig. 10. GST-6X Histidine fused BirA* IMAC chromatogram illustrates steadily increasing imidazole gradient (green) from 20 mM (100% buffer A) to 1 M (100% buffer B). The peak in the chromatogram obtained from absorbance values at 280 nm (blue) is caused by the elution of the protein. The previous not well defined peaks are because of impurities in the column.
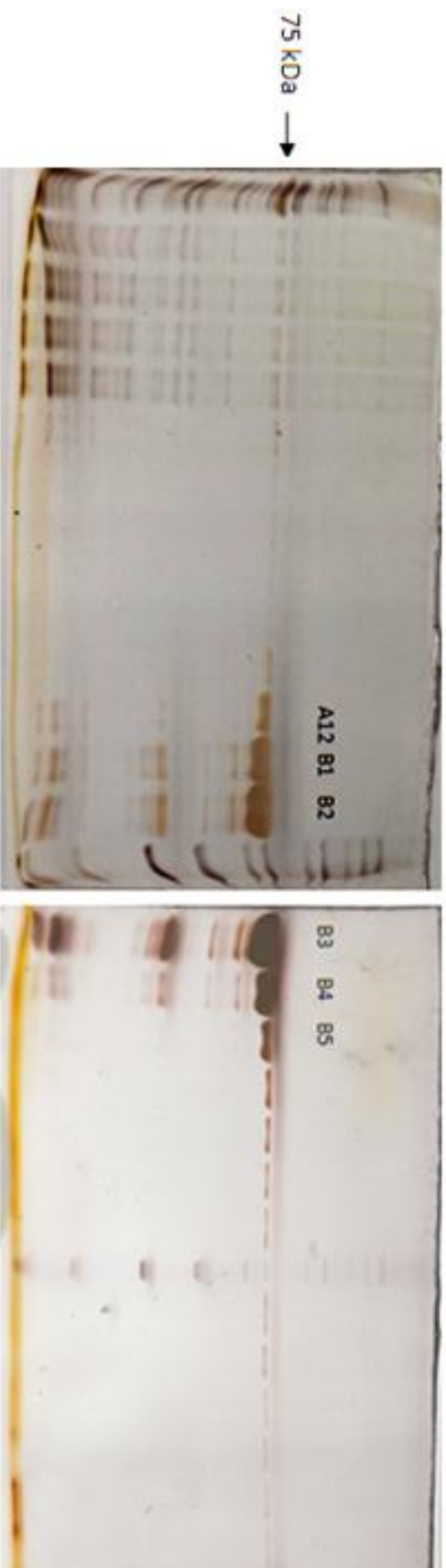
Fig. 11. HisTrap purification of BirA*-GST using an imidazole linear gradient

Fig. 12. Desalted BirA*-GST fractions. Respectively **1** comes from the mixture between B1, B2, B3, B4 fractions and **2**, from A12, B5 mixed together.

3.3 GST cloning and purification

As a control for the BirA* (GST fused) activity assays, GST protein purification was also attempted. As Gateway® Nova pET-60-DEST™ contains toxic *ccdB* cassette that prevents it from being cloned as an empty vector, a modified vector in which the *ccdB* gene was removed by enzymatic digestion was created. In order to generate blunt ends from sticky ends (caused by restriction digestion) nick filling by klenow enzyme was used. The sample was transformed initially in *E. coli* DH5α and later, after clone verification, in competent cells suitable for protein expression (RosettaDE3pLysS strain) (data not shown).

To purify GST protein, glutathione immobilized-paramagnetic particles were used (see materials and methods). The results as analyzed on 12% SDS-PAGE using Coomassie and silver staining revealed GST fractions with a high level of purity and good concentration (Fig. 13).
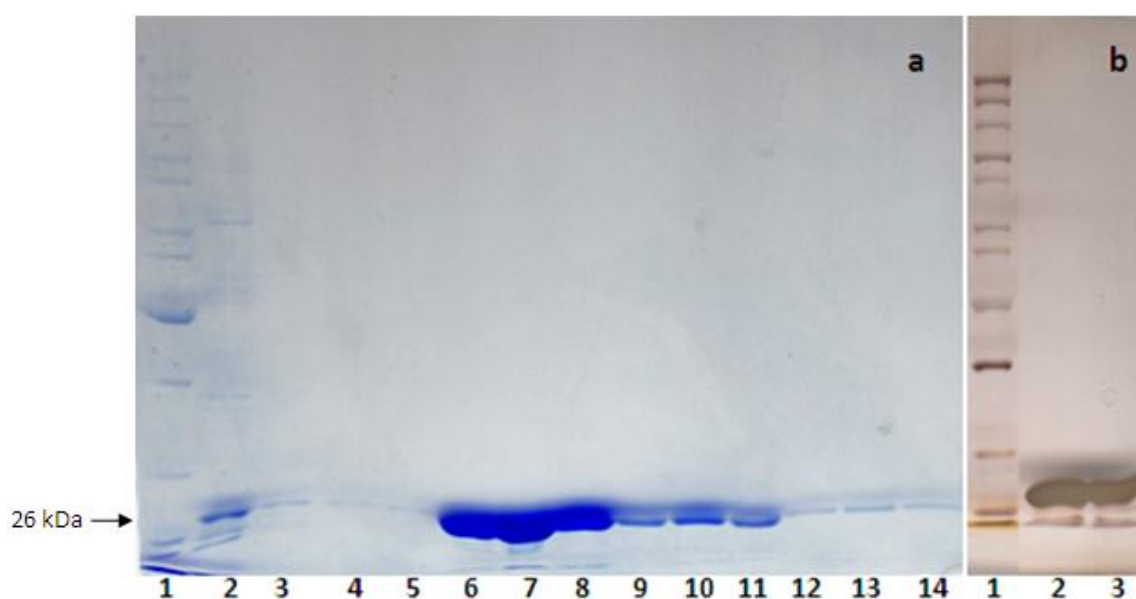
Fig. 13. GST purification on the mixed washes from the previous histidine purification. Coomassie (a) and silver staining (b). (a) lane 1 marker, 2-4 washes, 6-8 first elutions, 9-11 second elutions, 12-14 third elutions. (b) lane 1 marker, 2-3 first elutions.

3.4   Activity assays

The aim of the project was to study BirA*'s enzymatic activity *in vitro* in order to assess the possible use of this protein as a biotinylation agent. In order to test the promiscuous biotinylation we set up the reaction, using GST-6X Histidine fused BirA* and GST protein with the supplementation of ATP and biotin. As per the literature, under these conditions we should observe self-biotinylation of BirA* because of the presence of lysine residue within the N-terminus/HTH domain (Choi-Rhee *et al.*, 2004) of BirA* and biotinylation of GST by the promiscuous BirA* due to interaction between GSTs (between GST provided as substrate and GST fused to BirA*).

In order to extend the usage of this protein as a detection method in plants, we performed different assays. The purpose was to identify the ranges in which BirA* showed higher promiscuous biotinylation so as to understand if this technique could be developed for several plant species (grown in different temperature conditions) and the cellular compartments (different pH). Since the most influential variables are temperature and pH, we performed two assays to test these two factors and other two

31

assays to identify the best concentration of ATP and biotin to supplement the reaction, and the time in which the reaction gives the best results.

3.4.1   Optimization of substrate composition

According to the literature, intracellular level of biotin is insufficient for BirA to carry out biotinylation reaction, thus, supplementation of media with biotin and ATP is essential (Choi-Rhee *et al.*, 2004). Moreover, it has been shown in mammalian cells that a concentration of biotin-AMP similar or higher than 100 µM results in nonspecific biotinylation. On the contrary, 1 mM biotin gives lower signal compared to the one gave by 10 µM biotin; the explanation could be attributed to the difficulty to wash off biotin resulting in competition during streptavidin mediated pull-down assay (Howarth and Ting, 2008).

Therefore, optimization of biotin concentration was essential. Also, optimum ATP concentration was tested performing biotinylation experiments at three different concentrations which were chosen according to the ATP concentration in plant compartments.

The activity was checked by western blot using streptavidin antibodies after 1 and 24 hours incubation of the reactions. In the first experiment (1 h incubation, Fig. 14 a, b, c), GST biotinylation is lightly visible in most of the fractions except for the fraction where ATP has a medium concentration and biotin has the highest concentration (lane 7). Comparing this result with the silver stained gel the explanation become clear, the amount of BirA* is higher in this fraction compared with the others. A deeper analysis shows that in the lanes where ATP is at the lowest concentration (lane 2, 3, 4) we can observe the self-biotinylation of the BirA* but not of GST; however in the other lanes, where ATP is higher the GST biotinylation occurs. Except for the fraction mentioned before with a higher amount of BirA*, in all of the others the amount of BirA* loading was relatively similar.

In the other experiment (24 h incubation, Fig. 14d) we can see that after 24 hours the signal for BirA* self-biotinylation is present in every sample as well as the GST biotinylation. The difference resulted in the intensity of the signal reveals a higher biotinylation, both for the enzyme itself and the GST, with the middle conditions of ATP and biotin. The difference between the others is rather marked.

These results and those from the previous experiment led us to conclude that medium concentration for both the substrates would be the best to carry out all of the experiments.

Another thing came out from the second blot; unexpectedly the higher concentration of ATP seems to partially inhibit the reaction. In fact, if we compare the samples with the same amount of biotin but lower amount of ATP (Fig. 14d, lane 4) with those with the highest amount (Fig. 14d, lane 5), we can easily notice that in the latter lane the signal is lower.
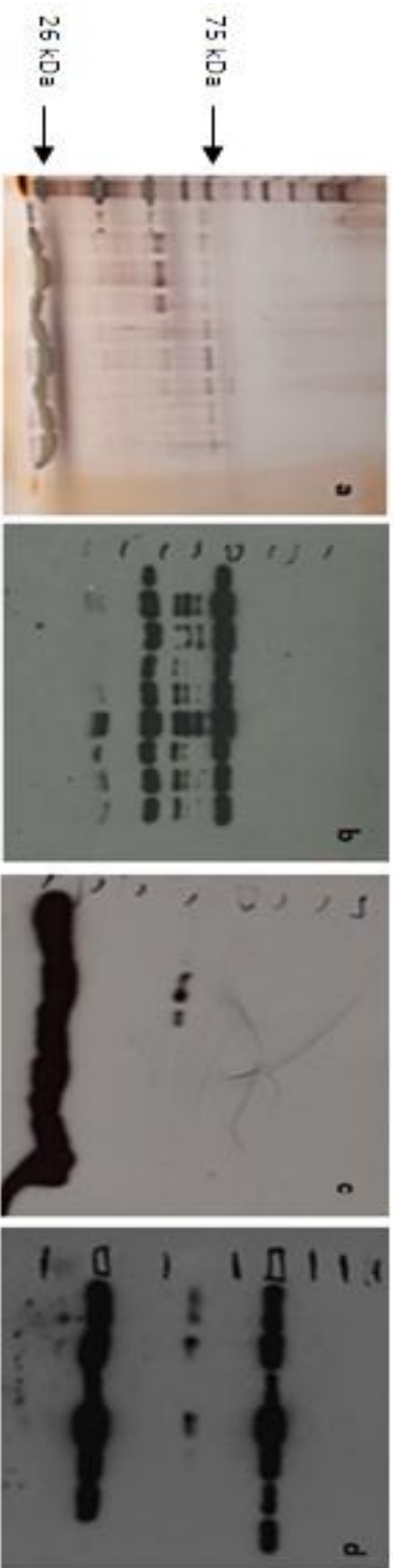
Fig 14. Substrate's optimisation assays. (a) silver stained gel of 1 h incubation experiments, 1 lane is marker, lanes 2, 3 and 4 have lower ATP concentration but respectively low, medium and high biotin concentration; lanes 5, 6 and 7 have medium concentration of ATP and respectively low, medium and high of biotin and lane 8, 9 and 10 have higher ATP concentration and respectively low, medium and high biotin concentration. (b) strp-HRP blot and (c) anti-GST blot. The two blots show the same loading order than the silver stained gel. (d) strp-HRP blot of 24 h incubation experiments. Lane 1 marker, lanes 2, 3 and 4 have lower biotin concentration and respectively low, medium and high of ATP, lanes 5 and 6 have medium biotin concentration and respectively medium and high concentration of ATP, lane 6 has higher concentration of ATP and biotin and lane 7 is the negative control with higher ATP and biotin concentration but without GST.

3.4.2 Reaction time

Biotinylation, as all of the other enzymatic processes, is a time-dependent reaction; so, it became necessary monitoring the reaction progress over the time. The enzyme needs time to start the reaction and, after a while, it is possible detect the substrate saturation. In order to establish the time period in which the enzyme gives the best results in terms of GST-biotinylation, biotinylation reactions were incubated for different incubation times. For the reaction, three controls were set-up, two negative controls (BirA* without GST and second negative control containing GST without BirA*) and a proximity control. As the biotinylation is a proximity-dependent reaction, this control was set up to test if higher reaction volume would reduce GST biotinylation as compared to a reaction set-up in lower volume (same concentration of BirA* and GST). This experiment would help us deduce if GST biotinylation was due to crowding of reaction with GST and bio-AMP cloud or due to the fact that GST interacted with BirA*-GST and hence got biotinylated.

The control reactions were blocked after 4 hours.

As shown in figure 10, at time-point 0 (start of the reaction), no GST biotinylation was visible. The signal was already strong and saturated after 2 hours incubation at 37°C as no difference in level of biotinylation could be seen after 4 hours incubation at 37°C (Fig. 15).

Afterwards, we decided to use 1 hour of incubation for all of the other assays.
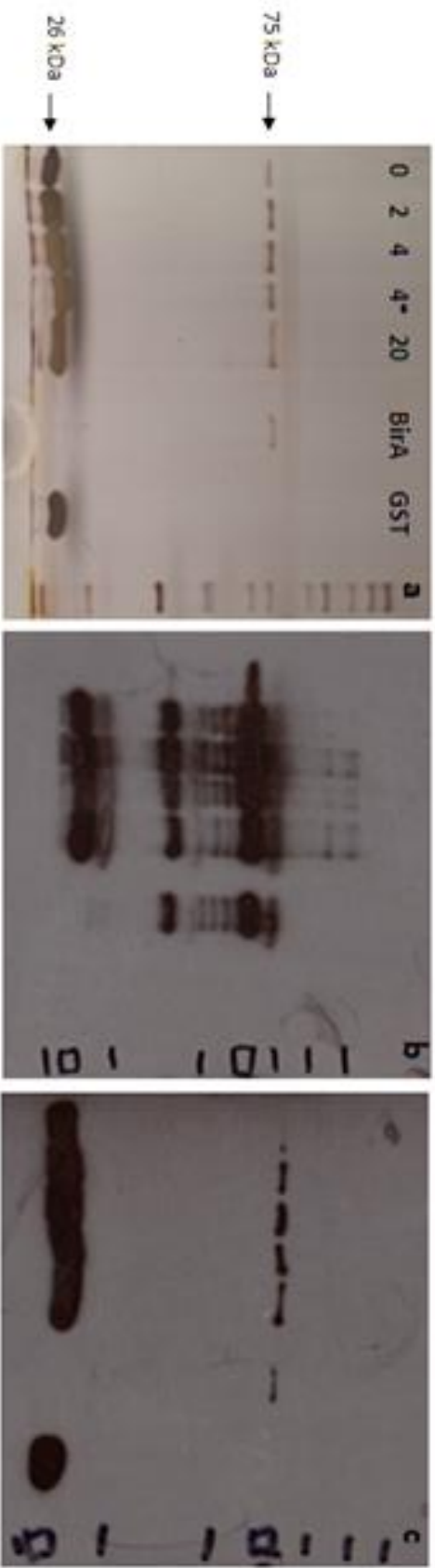
Fig. 15. Time assay monitoring the activity of BirA* enzyme through the time. Six experiments were conducted. First lane is 0 hour experiment, then 2, 4 and 20. 4* represent the proximity control, BirA is the control without GST and GST is the control without BirA* . (a) silver stained gel, (b) is strp-HRP blot and (c) anti-GST blot. The loading order is the same for all of them.

3.4.3 Optimal temperature

The enzyme activity is mostly influenced by the temperature; in fact, for any chemical reaction, we can see an increase in the velocity of the reaction with the temperature. According to an empirical rule, the velocity increases two or three times every 10°C (Connors, 1990). Monitoring the enzyme activity at different temperature values generally gives a bell-shaped curve; the reaction rate increases with temperature to a maximum level, then, declines with further increases of temperature. This behavior is frequently described as temperature optimum. Majority of enzymes rapidly denature at temperatures above 40°C due to the instability of their 3-D structure, consequentially, leading to a decrease of product formation (Bisswanger, 2014).

The temperature dependency of BirA* activity was analyzed, initially, in large temperature range, between 0 and 50°C (Fig. 16A). The results showed higher activity between 28 and 42°C. After that, we tried the same assay with a narrower range, focusing only on the activity between 28 and 52°C (Fig. 16B). The activity was detected again between 28 and 42°C but was higher around 35-37°C. Finally, difference in the biotinylation activity was observed per degree rise in temperature. It showed that BirA* activity is almost the same between 35 and 42°C (Fig. 16C).

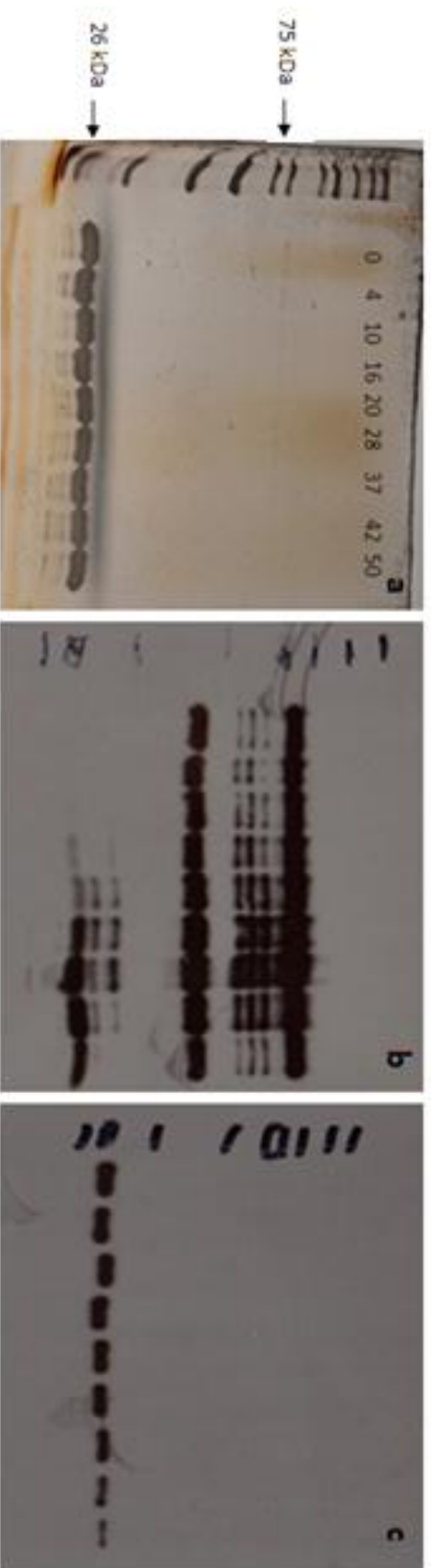Fig. 16A. Effect of temperature on biotinylation activity of BirA*. (a) silver stained gel representing 0-50°C temperature range, (b), (c) blots, strp-HRP and anti-GST respectively, corresponding to the fractions of gel (a).
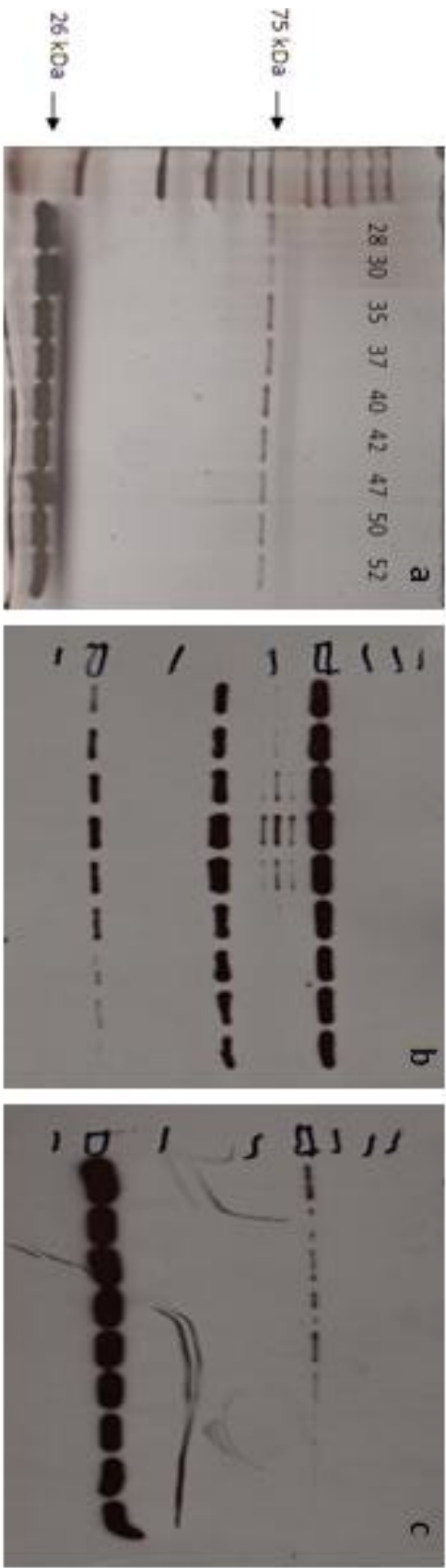
Fig. 16B. Effect of temperature on biotinylation activity of BirA* : (a) silver stained gel for 28-30°C temperature range, (b) strp-HRP and (c) anti-GST blot corresponding to these fractions.
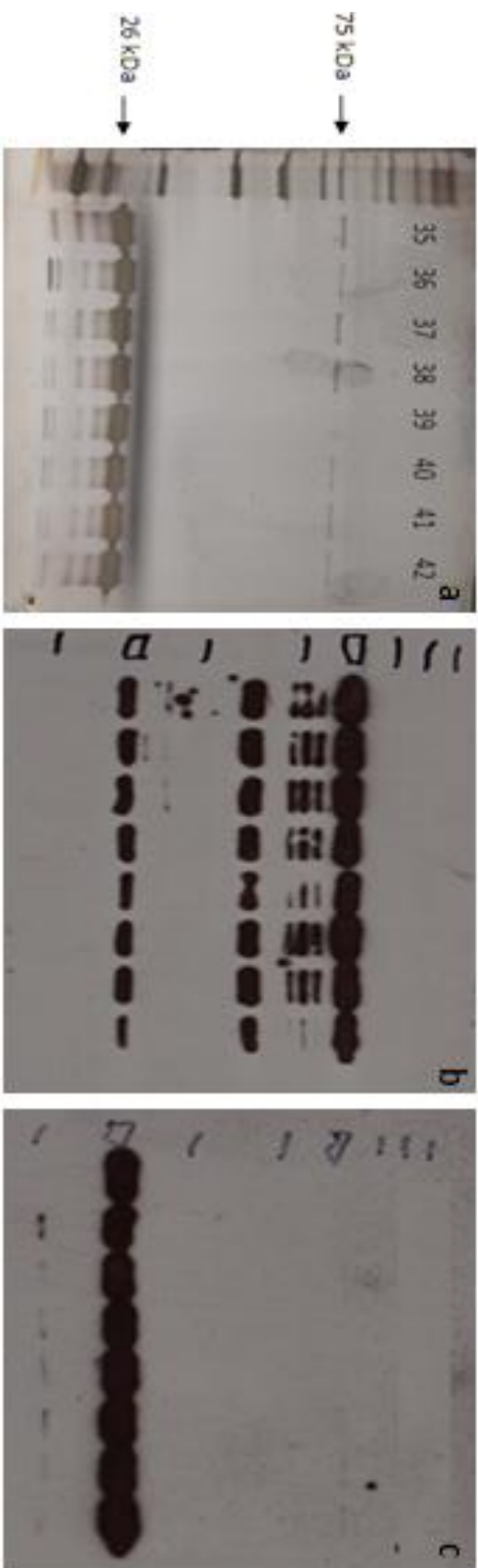
Fig. 16C. Effect of temperature on biotinylation activity of BirA*. (a), (b), (c) silver stained gel, strp-HRP blot and anti-GST blot from the 35-42°C range experiments.

3.4.4 Optimal pH

The pH, as well as the temperature, influences the activity of the enzymes. Similar to temperature, enzyme activity in variable pH generally follows a bell-shape curve. The pH at which maximum activity is obtained is called the pH optimum and it corresponds to the highest value for $V_{max.}$ This effect is also due to possible alteration in the 3D structure of protein as are ionic interactions or hydrogen bonds could have profound influence in the tertiary and quaternary protein structure. Also, protonation/deprotonation can modify the functional groups of the enzymes (Bisswanger, 2014). In order to test the optimal pH range for BirA* activity, biotinylation reaction was carried out at different pH values, from 1 to 14 (Fig. 17). The result indicated that there was no/very weak GST biotinylation at extreme pH values (pH 1-5, 13-14), and the best activity was registered from pH 7 to 10.
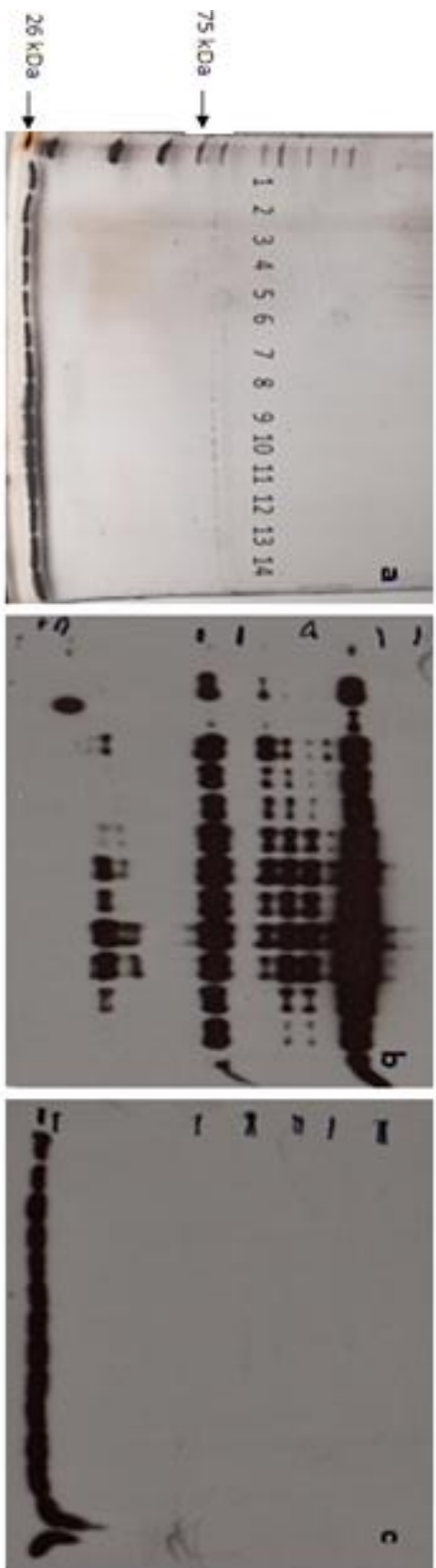
Fig. 17. Biotinylation reaction at different pH values, from 1 to 14. (a) silver stained gel, (b) strp-HRP blot and (c) anti-GST blot.

4. Discussion

The advent of the whole genome sequencing allowed characterization of increasing amounts of genomic information, which is eventually translated to proteomic studies. Proteomics is a branch of molecular biology that includes study of entire set of proteins expressed by a genome, cell, tissue or organism. It includes studies of protein modifications, abundance, expression, transport and interaction of proteins, all which have profound effect on activity of a certain protein and on cell expressing these proteins. We are able to identify proteins which are responsible for cellular regulation but knowing their role in the regulatory pathway is not immediately known. In order to understand the behavior of a protein in the whole cell we need to analyze the protein-protein interaction network (Pawson and Nash, 2003), which itself is largely restricted by the lack of information available about protein interactions. PPIs and association studies represent useful elements that could be used to build a large interconnectivity network within the cell. Once this network is built, it can further be used to predict the function of unknown proteins.

My study was the preliminary part of a bigger project aimed at developing a new method for identifying PPIs. As an extension of the project, the new method could further be used to create an inventory of specific cellular compartments.

The primary goal of this thesis research was to purify BirA* and check its activity under different physiological conditions. BirA is biotin protein ligase of *E. coli*, whose primary function is to biotinylate proteins of secondary metabolism, especially fatty-acid metabolism, as a part of their post-translational modification. Biotinylation reaction proceeds with formation of a highly reactive intermediate, bio-AMP, whose synthesis from biotin and ATP is catalyzed by BirA. Bio-AMP is a highly reactive compound that has capacity to biotinylate any acceptor lysine group it encounters, leading to non-specific biotinylation reactions. In nature, this phenomenon is prevented by sequestration of this molecule in the active site of BirA, till acceptor peptide with a specific sequence (as that of enzymes involved in fatty acid metabolism) approaches the active site of BirA (Li and Cronan, 1992). This property of BirA has been used to develop one-on-one protein interactions where, one protein was expressed in frame with BirA while the other protein is fused with the consensus sequence of the biotin acceptor peptide. The advantage of this technique was due to the specificity of the reaction,

chance of identifying false positive interactions was minimized. The disadvantage was that this technique could only be used for confirmation of one to one interactions. However, in certain mutant forms of BirA, the ability to sequester bio-AMP in active site is lost; but the ability to catalyze reaction still persists. One of these mutants has an arginine to glycine conversion at amino acid residue 118 (Kwon *et al.*, 2000), which shows active catalysis of bio-AMP reaction but also has high dissociation constant for the mixed anhydride (400-fold greater than wild-type BirA) (Know and Beckett, 2000). This mutant opened up new avenues for development of BirA mediated PPI techniques, based on the principle that if BirA* (BirA R118→G) is fused to a protein of interest, the fusion protein and all other proteins which are involved in complex formation with this protein would come in contact with highly reactive bio-AMP cloud and it would be biotinylated on their lysine residues (Roux *et al.*, 2002). This technique has been put into practice in two studies where, in mammalian cells was used to identify candidate proteins associated with the intermediated filament protein lamin-A (LaA) of the nuclear lamina, from the nuclear envelope (Roux *et al.,* 2012) whereas, in yeast cells it was used to characterized several component of the cytoskeletal structure, the bilobe, of *Trypanosoma brucei* (Morriswood *et al.*, 2013). The project in which I was working was aimed at developing this technique for plants. Also, as an extension to this, the project aimed at utilizing this property of BirA* in deciphering organelle proteome. This is based on a simple hypothesis that, if we overexpress BirA* thus forming bio-AMP cloud in a confined area, such as chloroplast or mitochondrial stroma, we could potentially biotinylate all the proteins present in the organelle. Although techniques such as sub-cellular fractionation is available and is generally used to meet this aim, it suffers a major problem of cross-contamination during the process of fractionation, mainly due to lysis of organelles etc. Use of BirA, however, would potentially solve this problem as only proteins proximal to the bio-AMP cloud would be biotinylated and could be pull-down using streptavidin from whole cell lysate (therefore time and labor consuming step of fractionation could be avoided).

In order to realize the above mentioned projects in plants, certain pre-experiments were required to ascertain if BirA* activity was preserved in the physiological conditions inside the plant cell. These included testing different physiological conditions such as temperature, pH and substrate concentration.

Choi-Rhee *et al*. (2004) reported in their studies that a supplementation of ATP and biotin is necessary to make the biotinylation reaction takes place. This is an especially

interesting aspect of the BirA* mediated *in vivo* biotinylation technique as the reaction could be controlled with the help of substrate supplementation. We tested different concentrations of these two compounds (ATP and biotin) to identify the concentrations at which the reaction works best. The concentrations chosen for biotin were lower compared to these chosen for ATP. That difference was due to the different affinity that these two molecules have for the enzyme. In fact, the affinity for biotin is higher, even in a low intracellular biotin concentration 10-100 nM, a tight bond between biotin and the enzyme is formed (Xu and Beckett, 1994). On the other hand, the K_M for ATP is high, so the enzyme requires a higher concentration, in the millimolar range (Beckett, 2007). The experiments showed that the self-biotinylation of the enzyme always occurs, independently from the ATP and biotin concentration, but the GST biotinylation, during short time incubation, is visible only with concentrations higher than 200 μM and 10 mM for biotin and ATP respectively. These results compared with the others obtained from the long time incubation led us to think that these concentrations were the best to perform all of the assays. We also observed a slight inhibition of the biotinylation reaction by high ATP concentration. This could be due to an allosteric regulation, similar to what happens for phosphofructokinase (PFK-1), whose activity is inhibited by high ATP concentration (Walker *et al.*, 2012).

Temperature plays an huge role in plant growth and in every biological process linked to them as flowering, germination, photosynthesis. The temperatures preferred for most of the physiological process are in the range of 0-40°C (Went, 1953). Different plant species have different optimum temperatures of growth and their thermotolerance also varies to a great extent as exemplified from following data: cotton can tolerate 45°C (Rehman *et al.*, 2004), brassica and tomato have thermotolerance till 29°C and 30°C, respectively (Morrison and Stewart, 2002; Camejo *et al*., 2005). Wheat instead, which is amongst the most popular crop plants, prefers lower temperature nearing 26°C (Stone and Nicolás, 1994) (Fig. 18). As the aim of the project is to develop a technique that could be extended to crop and other economically important crops at a later stage, it was very important to test if BirA* could be functionally active at these variable temperature ranges. Our results indicate that BirA* got biotinylated at every temperature we tested, but the GST biotinylation was slightly detected at 16 and 20°C; the signal became stronger from 28°C till 42°C. From the results of experiment, we could conclude that since BirA is functionally active from 28-42°C, thus this technique could easily be applied to a wide range of crop and other commercially important plants.

This method could be used in model plant, *Arabidopsis thaliana* because, even though the optimal growing temperature is between 23 and 25°C, older plants can tolerate higher temperatures, at least up to 30°C (ABRC, Handling Arabidopsis Plants and Seeds,

http://www.biosci.ohiostate.edu/~plantbio/Facilities/abrc/abrc_handling_mw2004.pdf).

Proteome of *Nicotiana tabacum* (optimal growth temperature is between 20 and 30°C, FAO) a model plant for solanaceae family could also be deciphered and protein interactions revealed from this approach.

Crop plants	Threshold temperature (°C)	Growth stage
Wheat	26	Post-anthesis
Corn	38	Grain filling
Cotton	45	Reproductive
Pearl millet	35	Seedling
Tomato	30	Emergence
Brassica	29	Flowering
Cool season pulses	25	Flowering
Groundnut	34	Pollen production
Cowpea	41	Flowering
Rice	34	Grain yield

Fig. 18. Table of the main crop plants and their threshold temperature (Modified from Wahida *et al.*, 2007)

All of the organelles within the cell are well interconnected with each other but, at the same time, most of them can act as a singular compartment. They have different functions and structures and most of the times they show different pH values. The pH range of different plants is shown in figure 19. The conservation of these different pH is essential to allow all of the biological processes of the cell, in fact pH affects metabolism, membrane trafficking, ion channel activity and proteins structures and stability (Casey *et al.*, 2010). In order to identify the compartments whose proteome could be deciphered using BirA*, it was essential to know the pH range within which BirA* is functionally active. From the results we obtained exposing the reaction mixture

to different pH values, we can see that, even though BirA* is able to biotinylated itself in every condition, the GST biotinylation occurs just in a small range, from pH 7 to 11. The highest activity was registered at pH 8, because, even though the streptavidin signal is rather the same in the reaction carried out at pH 10 and the one at pH 8, we can see from the silver stained gel that the amount of protein in the first sample is higher than the second one. So, if we try to normalize the signal with the enzyme amount, we can conclude that pH 8 seems to be the best pH condition. Shen *at al.* (2013) have developed genetically encoded pH sensors and they have used them to measure the organelle pH in the *Arabidopsis* endomembrane system. With this knowledge we can assume that this method could be applied to detect nuclear, cytosolic, mitochondrial, peroxisomial and plastidial proteins.

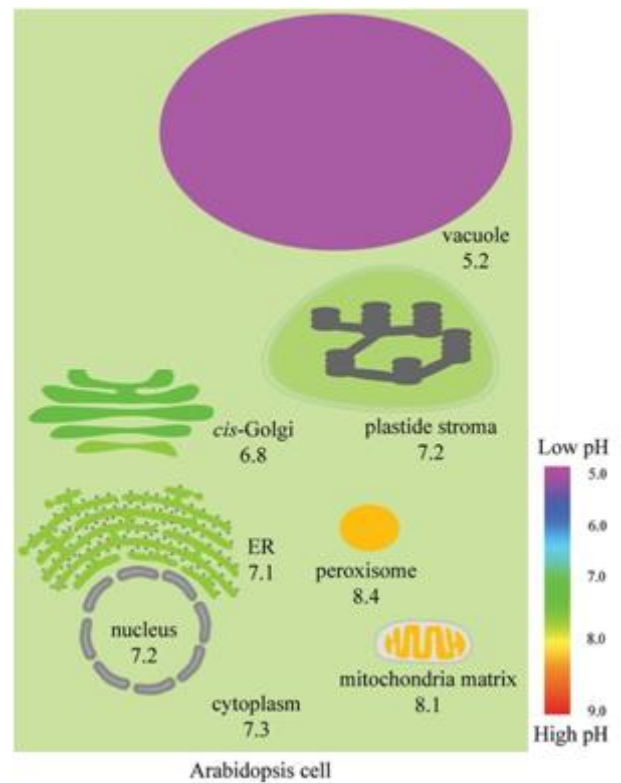| pH 5.2	in the vacuole
pH 5.5	in the lysosome
pH 7.1	in the endoplasmic reticulum (ER)
pH 7.2	in the nucleus
pH 7.2	in the plastid stroma
pH 7.3	in the cytosol
pH 8.1	in the mitochondrial matrix
pH 8.4	in the peroxisome

Fig. 19. Different pH values shown by different cell compartments (Modified from Shen *et al.*, 2013).

To summarize, the work presented in this thesis is a preliminary study to assess if BirA* could be used in an easy and rapid method to isolate and characterize the proteins of a specific cellular compartment and to identify PPIs. The purpose of this work was to understand in advance the behavior of this enzyme and its capacity to operate in different physiological conditions. The results show that BirA* could be used to identify PPIs in different plant systems due to its ability to function in wide temperature ranges. However, due to its inability to maintain its activity over a range of pH, it might not be applied to decipher organelle proteomes such as vacuole or of apoplastic region. Despite of that, BirA* would be overexpressed in chloroplast and mitochondria to decipher the organellar proteome. As an extension of this work, in future, BirA* would be expressed in constitutive and inducible manner in *A. thaliana*, while being fused to microtubule associated proteins to identify PPIs in cytoskeletal system, as a proof of concept.

Appendix I

LB medium	10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl
Running buffer 10X	250 mM Tris base, 2.5 M Glycine, 1% SDS
Resolving gel (12%, 5 mL)	1.6 mL H_2O, 2 mL 30% Acrylamide/Bis Solution, 19:1 (Bio-Rad), 1.3 mL 1.5 M Tris-Cl pH 8.8, 0.05 mL 10% SDS, 0.05 mL 10% ammonium persulfate, 0.002 mL TEMED
Stacking gel (5%, 2 mL)	1.4 mL H_2O, 0.33 mL 30% Acrylamide/Bis Solution 19:1 (Bio-Rad), 0.25 mL 1 M Tris-Cl pH 6.8, 0.02 mL 10% SDS, 0.02 mL 10% ammonium persulfate, 0.002 mL TEMED
Coomassie solution	45% methanol, 10% glacial acetic acid, 0.05% Coomassie Brilliant Blue R-250
Destain solution	45% methanol, 10% glacial acetic acid
Laemmli buffer 5X	250 mM Tris-Cl pH 6.8, 10% SDS, 0.5% Bromophenol blue, 50% glycerol
Tris-Buffered Saline (TBS)	25 mM Tris-Cl pH 7.5, 150 mM NaCl
TBST buffer	TBS, 0.05% Tween 20
Transfer buffer	39 mM Glycine, 48 mM Tris, 0.03% SDS, 20% Methanol
Native buffer	50 mM Tris-Cl pH 7, 100 mM NaCl, 10% Glycerol
Sonication buffer	50 mM native buffer pH 7 with 5 mM DTT, 10 mM benzamidine, Halt™ Protease Inhibitor Cocktail EDTA-Free
Equilibration/Wash Buffer	50 mM NaH_2PO_4, 250 mM NaCl, 10mM imidazole, pH 7.4
Elution Buffer	50 mM NaH_2PO_4, 250 mM NaCl, 150mM imidazole; pH 7.4

Sodium phosphate buffer	57.7 mL 1 M Na_2HPO_4 and 42.3 mL 1 M NaH_2PO_4, pH 7
IMAC buffer A	20 mM NaH_2PO_4, 500 mM NaCl, 20 mM imidazole, 2 mM DTE, pH 7.4
IMAC buffer B	20 mM NaH_2PO_4, 500 mM NaCl, 1 M imidazole, 2 mM DTE, pH 7.4
Elution buffer MagneGST™	50 mM Tris-Cl, pH 8, 50 mM glutathione
PBS	140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3
Elution buffer Glutathione Sepharose 4B	50 mM Tris-HCl pH 8.0, 10 mM reduced glutathione

Bibliography

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P.(2002). Analyzing Protein Structure and Function. *Molecular Biology of the Cell*, 4th ed. New York: Garland Science. p 519.

Beckett D. (2007). Biotin sensing. Universal influence of biotin status on transcription. *Annu Rev Genet*. 41: 443–464.

Beckett D, Kovaleva E, Schatz PJ. (1999). A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci*. 8(4): 921-929.

Bisswanger H. (2014). Enzyme assays. *Perspectives in Science*. 1(1–6): 41–55.

Bradford MM. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*. 72(1-2): 248-254.

Camejo D, Rodrıguez P, Morales MA, Dell'amico JM, Torrecillas A, Alarcon JJ. (2005). High temperature effects on photosynthetic activity of two tomato cultivars with different heat susceptibility. *J Plant Physiol*. 162(3): 281–289.

Casey JR, Grinstein S, Orlowski J. (2010). Sensors and regulators of intracellular pH. *Nat Rev Mol Cell Biol*. 11(1): 50–61.

Chakravartty V, Cronan JE. (2012). Altered regulation of *Escherichia coli* biotin biosynthesis in BirA superrepressor mutant strains. *J Bacteriol*. 194(5): 1113–1126.

Chakravartty V, Cronan JE. (2013). The wing of a winged helix-turn-helix transcription factor organizes the active site of BirA, a bifunctional repressor/ligase. *J Biol Chem*. 288(50): 36029-36039.

Chien CT, Bartel PL, Sternglanz R, Fields S. (1991). The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc Natl Acad Sci U S A*. 88(21): 9578-9582.

Choi-Rhee E, Schulman H, Cronan JE. (2004). Promiscuous protein biotinylation by *Escherichia coli* biotin protein ligase. *Protein Sci*. 13(11): 3043–3050.

Connors KA. (1990). Chemical Kinetics: the study of reaction rates in solution. *VCH Publishers*. p 14.

Deal RB, Henikoff S. (2011). The INTACT method for cell type-specific gene expression and chromatin profiling in *Arabidopsis thaliana*. *Nat Protoc*. 6(1): 56-68.

Fernández-Suárez M, Chen TS, Ting AY. (2008). Protein-protein interaction detection *in vitro* and in cells by proximity biotinylation. *J Am Chem Soc*. 130(29): 9251-9253.

Green NM. (1963). Avidin. 1. The use of [14C]biotin for kinetic studies and for assay. *Biochem J*. 89(3): 585–591.

Harlow E, Lane D. Antibodies. A laboratory manual. (Cold Spring Harbor Laboratory, 1988).

Howarth M, Ting AY. (2008). Imaging proteins in live mammalian cells with biotin ligase and monovalent streptavidin. *Nat Protoc*. 3(3): 534-545.

Kane JF, Hartley DL. (1988). Formation of recombinant protein inclusion bodies in *Escherichia coli*. *Trends Biotechnol.* 6(5): 95–101.

Kwon K, Beckett D. (2000). Function of a conserved sequence motif in biotin holoenzyme synthetases. *Protein Sci*. 9(8): 1530-1539.

Kwon K, Streaker ED, Ruparelia S, Beckett D. (2000). Multiple disordered loops function in corepressor-induced dimerization of the biotin repressor. *J Mol Biol*. 304(5): 821-833.

Lane MD, Rominger KL, Young DL, Lynen F. (1964). The enzymatic synthesis of holotranscarboxylase from apotranscarboxylase and (+)-biotin. *J Biol Chem*. 239(9): 2865–2871.

Li SJ, Cronan JE Jr. (1992). The gene encoding the biotin carboxylase subunit of Escherichia coli acetyl-CoA carboxylase. *J Biol Chem*. 267(2): 855-863.

Mitraki A, King J. (1989). Protein Folding Intermediates and Inclusion Body Formation. *Nature Biotechnology*. 7(7): 690–697.

Morrison MJ, Stewart DW. (2002). Heat stress during flowering in summer brassica. *Crop Sci*. 42(3): 797-803.

Morriswood B, Havlicek K, Demmel L, Yavuz S, Sealey-Cardona M, Vidilaseris K, Anrather D, Kostan J, Djinovic-Carugo K, Roux KJ, Warren G. (2013). Novel bilobe components in *Trypanosoma brucei* identified using proximity-dependent biotinylation. *Eukaryot Cell.* 12(2): 356-367.

Mukhopadhyay B, Purwantini E, Kreder CL, Wolfe RS. (2001). Oxaloacetate synthesis in the methanarchaeon Methanosarcina barkeri. Pyruvate carboxylase genes and a putative *Escherichia coli*-type bifunctional biotin protein ligase gene (bpl/birA) exhibit a unique organization. *J.Bacteriol*. 183(12): 3804–3810.

O'Neill LP, Turner BM. (2003). Immunoprecipitation of native chromatin: NChIP. *Methods*. 31(1): 76-82.

Ooi SL, Henikoff JG, Henikoff S. (2010). A native chromatin purification system for epigenomic profiling in *Caenorhabditis elegans*. *Nucleic Acids Res.* 38(4): e26.

Pawson T, Nash P. (2003). Assembly of Cell Regulatory Systems Through Protein Interaction Domains. *Science.* 300(5618): 445-452

Rehman H, Malik SA, Saleem M. (2004). Heat tolerance of upland cotton during the fruiting stage evaluated during cellular membrane thermostability. *Field Crops Res.* 85(2-3): 149–158.

Roux KJ, Kim DI, Raida M, Burke B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol.* 196(6): 801-810.

Royer C. (1999). Protein-protein interactions. *Outline of the Thermodynamic and Structural Principles Governing the Ways that Proteins Interact with Other Proteins.* Previously Published in the Biophysics Textbook Online (BTOL).

Sambrook J, Russell DW. (2001). *Molecular Cloning: A Laboratory Manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, p 1124-1125.

Samols D, Thornton CG, Murtif VL, Kumar GK, Haase FC, Wood HG. (1988). Evolutionary conservation among biotin enzymes. *J Biol Chem.* 263(14): 6461-6464.

Shen J, Zeng Y, Zhuang X, Sun L, Yao X, Pimpl P, Jiang L. (2013). Organelle pH in the Arabidopsis endomembrane system. *Mol Plant.* 6(5): 1419-1437.

Shoaib M, Kulyyassov A, Robin C, Winczura K, Tarlykov P, Despas E, Kannouche P, Ramanculov E, Lipinski M, Ogryzko V. (2013). PUB-NChIP--"*in vivo* biotinylation" approach to study chromatin in proximity to a protein of interest. *Genome Res.* 23(2): 331-340.

Steiner FA, Talbert PB, Kasinathan S, Deal RB, Henikoff S. (2012). Cell-type-specific nuclei purification from whole animals for genome-wide expression and chromatin profiling. *Genome Res.* 22(4): 766-777.

Stone PJ, Nicolas ME. (1994). Wheat cultivars vary widely in their responses of grain yield and quality to short periods of post-anthesis heat stress. *Aust J Plant Physiol.* 21(6): 887–900.

Streaker ED, Beckett D. (1999). Ligand-linked structural changes in the Escherichia coli biotin repressor. The significance of surface loops for binding and allostery. *J Mol Biol.* 292(3): 619–632.

Wahida A, Gelania S, Ashrafa M, Fooladb MR. (2007). Heat tolerance in plants: An overview. *Environ Exp Bot.* 61(3): 199–223.

Walker LR, Simcock DC, Pedley KC, H.V. Simpson HV, Brown S. (2012). The kinetics and regulation of phosphofructokinase from *Teladorsagia circumcincta. Exp Parasitol.* 13(4): 348-353.

Weaver LH, Kwon K, Beckett D, Matthews BW. (2001). Corepressor-induced organization and assembly of the biotin repressor: a model for allosteric activation of a transcriptional regulator. *Proc Natl Acad Sci U S A*. 98(11): 6045-6050.

Went FW. (1953). The Effect of Temperature on Plant Growth. *Annu Rev Plant Physiol*. 4: 347-362.

Wilson KP, Shewchuk LM, Brennan RG, Otsuka AJ, Matthews BW. (1992). *Escherichia coli* biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains. *Proc Natl Acad Sc. U.S.A*. 89(19): 9257–9261.

Xu Y, Beckett D. (1994). Kinetics of biotinyl-5'-adenylate synthesis catalyzed by the Escherichia coli repressor of biotin biosynthesis and the stability of the enzyme-product complex. *Biochemistry*. 33(23): 7354–7360.

Xu Y, Beckett D. (1996). Evidence for interdomain interaction in the Escherichia coli repressor of biotin biosynthesis from studies of an N-terminal domain deletion mutant. *Biochemistry*. 35(6): 1783-1792.