# UNIVERSITÀ DEGLI STUDI DI MILANO

## DIPARTIMENTO DI SCIENZE BIOMEDICHE, CHIRURGICHE E ODONTOIATRICHE


**DOCTORAL SCHOOL IN:**

**PHYSIOPATHOLOGICAL, NEUROPSYCHOBIOLOGICAL SCIENCES AND LIFE-CYCLE CARE**

XXVIII cycle

# Production of Human Recombinant Differentiation Factors in Transgenic Tobacco Plants


TUTORS:

Prof. M. DEL FABBRO

Prof: E. PEDRAZZINI

VALENTINA CERESOLI

R09501

Ph.D. COORDINATOR:

Prof. ROBERTO L. WEINSTEIN



Academic Year 2014/2015

# Table of Content

# 1 Introduction

Research on bone regeneration began decades ago as a result of intensive studies on bone growth and healing. Bone has been recognised, among the many tissues in human body, as having the highest potential for regeneration [1], and it is the second most transplanted tissue following blood [2]. Due to both internal mediators and external mechanical demands, it possesses the intrinsic ability for regeneration and is constantly engaged in a cycle of resorption and renewal undergoing continual chemical exchange and structural remodelling throughout adult life as well as during repair process in response to injury. Despite these abilities, beyond a critical point clinical intervention measures are needed; there are different clinical conditions requiring a large quantity of bone regeneration, such as for skeletal reconstruction of large bone defects created by trauma, infection, or cases in which the regenerative process is compromised, including necrosis, atrophic non-unions and osteoporosis.

To describe the extent of this situation, it is estimated that annually more than 2.2 million patients receive bone defect repairs worldwide [3], with a cost greater than $2.5 billion just in the United States; this figure is expected to globally double by 2020 due to a variety of factors, including increased life expectancy [4,5]. The board of the Bone and Joint Decade in 2009 has assessed that half of the people aged over 65, affected by chronic conditions, suffers of joint diseases and that the number of osteoporotic fractures has doubled in the previous 15 years [4,6]. It should be also pointed out that the worldwide incidence of bone disorders and conditions is increasing in those societies where population ageing is combined with increased obesity and poor physical activity [6].

Shortcomings, limitations, and complications of current clinical treatments for bone repair and regeneration have been reported in different studies [7,8]. A variety of graft materials are currently used to enhance bone healing, and the relative success of these materials depends on many factors, not only on the specific properties of the graft itself. In addition to its physical properties, to be effective, a grafting material is required to even provide osteoconductive and/or osteoinductive activities. Osteoconduction, the ability of promoting bone growth by allowing bone formation on material's surface, may suffice in clinical condition of less severe defects, where sufficient quantities and margins of bone exist. Osteoinduction instead, is the capability of promoting *de novo* bone formation at soft or hard tissue sites, and offers needful advantages for biologic reconstruction of severe situations.

Among most commonly used materials there are allografts, cadaveric bone usually obtained from a bone bank, autologous grafts, bone harvested from the patient's own body, or synthetic ones, often made of hydroxyapatite or other naturally and biocompatible substances.

Allografts, mineralized or demineralized, are histocompatible, available in various forms including demineralized bone matrix and cortical grafts and whole-bone segments, depending on the host-site requirements, and provide an osteoconductive environment; however, their osteoinductive capacity is highly variable depending on the processing method and sourcing, and may be present in inadequate amounts for any bone-inductive effect [9].

To date, autografts serve as the gold standard for bone grafts because they are histocompatible, non-immunogenic, and they offer all of the properties required. Specifically, this material provides an osteoconductive environment (i.e., three-dimensional scaffolds and porous matrix) coupled with cells (i.e., osteoprogenitor cells) as well as growth and differentiation factors(i.e., growth factors) that can result in osteoinduction [10]. Nevertheless, autogenous grafting is sometimes an expensive procedure that has a number of shortcomings, including the need for secondary surgery to harvest the graft, donor site morbidity, irregular resorption of transplanted tissue, and limited availability of donor bone [11,12]. Furthermore autograft may be a useless or inadequate treatment option in cases where the defect site requires larger amounts of bone than is available.

Other commonly used bone repair techniques involve synthetic materials and fillers, and growth and differentiation factors, but, although all these clinical interventions have been shown to improve bone repair, none of them possess all of the necessary characteristics: high osteoinductive and angiogenic potentials, biological safety, low patient morbidity, ready access to surgeons, no size restrictions, long shelf life and reasonable cost; all these limitations have led to an extensive research for alternatives.

The discovery and subsequent production of the osteoinductive agents in bone, the Bone Morphogenetic Proteins (BMPs), have provided the possibility of reducing or avoiding the need for autograft, through a tissue-engineering product.

Clinically, BMPs have demonstrated the potential to replace the use of autogenous bone in many applications so that costs and complications related to harvesting autograft can be prevented [13]; in addition, the ability to control quality, activity and dose of the osteoinductive agent may provide greater assurance of bone induction and repair. Despite this, a widespread therapeutic use of BMPs has been hindered by difficulties in obtaining large amounts of pure, biologically active protein at a cost-effective price [13].

Thus, the aim of this study was to develop a plant based system for cost-effective production of active recombinant BMPs. This introduction provides a review of the relevant literature pertaining to BMPs (especially to BMP-2) synthesis, processing and recombinant production process.

## 1.1 Bone Morphogenetic Proteins.

As early as 1889 Senn observed that decalcified bone could induce healing of bone defects in dogs [14]. Later Levander provided the first evidence of ectopic bone formation after injecting bone crude extracts into muscle [15,16]. In 1965 a pioneering work by Urist established a landmark on the research in bone regeneration: he discovered that demineralized bone fragments are able to induce new bone formation by autoinduction when implanted either subcutaneously or intramuscularly in a study that involved more than 300 animals [17]. Urist discovered that the active compound responsible for bone formation was a mix of proteins and proposed the name "Bone Morphogenetic Protein" in the scientific literature in the Journal of Dental Research in 1971 [18].

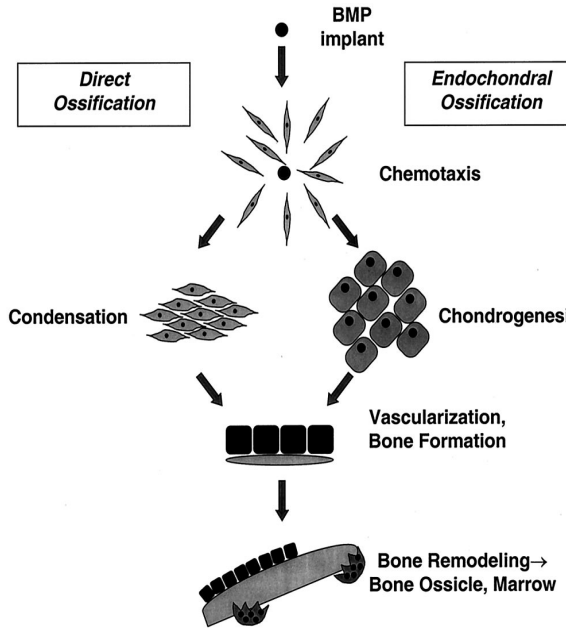



**Figure 1: Cellular events after implantation of BMP**. These proteins induce both endochondral and direct bone formation; the result in each case is woven bone that then remodels and becomes populated with hematopoietic bone marrow.
This figure is taken from Wozney, Spine 2002 [10].

In the years that followed, fundamental steps were made by Reddi and his colleagues, who proposed that BMPs are responsible for the initiation of a cascade of developmental events, in which progenitor cells in the bone marrow were induced to produce bone cells leading to bone regeneration [19–21].

In 1981 Reddi created a reproducible bioassay for BMP mediated ectopic bone formation, that was based on the activity of the Alkaline Phosphatase (ALP) enzyme and on the content of calcium in the newly formed bone [21]. Moreover they observed that demineralization of the bone matrix in the solid state with dissociative extractants, that contained protease inhibitors in order to solubilize putative inductive molecules, resulted in the loss of the bone inductive property of the matrix, and that there was a complete recovery of biological activity after reconstitution of the residues with the extracts, thus demonstrating BMP to be part of the extracellular matrix of bone and showing unequivocally that the extracellular matrix of bone is a reservoir of morphogenetic soluble signals that per se are able to initiate the cascade of bone differentiation by induction [21–23].

The cellular events caused by BMP have been described by histological examination of BMPs implanted subcutaneously in a rat assay system [20,24]. These include chemotactic action, that mediate the infiltration of the implant site with cartilage and bone cell precursors, proliferative events, the differentiation of precursor cells into chondrocytes, the induction of vascularisation of the site, the maturation of the chondrocytes and the differentiation of cells into osteoblasts (Figure 1). This complex process results in a new piece of bone tissue, complete with osteoblasts, osteocytes, osteoclasts, and bone marrow elements [20,25]. A steep dose–response curve is observed, with low doses resulting in small amounts of cartilage and bone formation; however, at no concentration stable cartilage formation was observed, being cartilage always replaced by bone [26]. Larger doses, instead, result in more substantial bone formation and earlier osteoinduction [26]. Interestingly, when higher concentrations of BMPs are applied, bone forms at the same time as cartilage formation, suggesting that BMP can result in direct (intramembranous) ossification [10,26]. The great work of Urist and Reddi provided the reproducible evidence of the osteogenic activity of demineralized bone and dentine matrices, implying the presence of morphogenetic factors and, as per Urist speculation, of a bone morphogenetic protein complex within the bone matrix [18,21]. Despite these results, the characterization of this elusive complex has been hindered by the limited quantities of putative proteins within the extracellular matrix of bone, tightly bound to the organic and inorganic components of the bone matrix [27]. The purification process of BMP from bone includes: the removal of the mineral component of bone with acid, the extraction of the active component from the remaining organic matrix of bone using chaotropic agents, and then multiple column chromatography steps performed with the material in guanidine or urea due to the insolubility of the extract. This procedure results in a very low and still not pure yield. [28,29].

During the decades of 80s and 90s, the BMPs genes were cloned and the recombinant proteins demonstrated to be biologically potent [29,25]. Recombinant BMPs (rhBMPs) were used first time in late 90s for clinical applications such as spinal fusion, fracture healing and dental tissue engineering [30–33]. Human rhBMPs are currently manufactured using mammalian cell-line expression by recombinant technology. Briefly, a production cell line was generated by placing the DNA coding sequence of BMP in a vector system that contains appropriate promoter enhancer sequences for protein expression; this vector is then transfected into a mammalian cell host, such as Chinese Hamster Ovary (CHO) cells for rhBMP-2; after a characterization step to ensure fidelity of expression, the cell line is expanded in sequentially larger volumes of culture medium and

thousands of liters of the BMP-expressing cells in medium can be obtained. The protein is then purified from the conditioned medium. As mammalian cells are used to synthesize these BMPs, the obtained molecules are dimerized, processed and glycosylated like the human ones [34]. The advantages of this recombinant production system include the reproducibility, the ability to ensure purity and activity of the proteins and the guarantee of a freedom from impurity products, but the production process is expensive. [35]. In 2002 the Food and Drug Administration (FDA) gave approval of two recombinant products containing rhBMP-2 and rhBMP-7 (also known as Osteogenic Protein-1, OP-1) in absorbable collagen carriers for spinal fusion and long bone non-unions [36–39].

### 1.1.1 Properties and Classification

BMPs are members of the transforming growth factor beta (TGF-β) superfamily of secreted growth and differentiation factors [40,41]. The BMP subfamily comprises at least fifteen proteins: BMP-2, BMP-3 (also known as Osteogenin), BMP-3b (also known as Growth and Differentiation Factor 10, GDF-10), BMP-4, BMP-5, BMP-6, BMP-7 (OP-1), BMP-8 (OP-2), BMP-9, BMP-10, BMP-11 (also known as GDF-8 or Myostatin), BMP-12 (also known as GDF-7), BMP-13 (also known as GDF-6), BMP-14 (also known as GDF-5), and BMP-15 [42,1]. BMP-1 is a type I procollagen C-proteinase and doesn't belong to the TGF- β family [43]. Based on the amino acid (AA) sequence homologies of the mature domains, members of the BMP family are further divided into sub-families (Figure 2) [42]. Among these groups, BMP- 2 and BMP-4 have 92% AA sequence homology of molecules, and BMP5, BMP-6, BMP-7, and BMP-8 have 78% AA sequence homology. Between these two groups there is a 55% to 65% sequence homology of mature proteins [25,44].

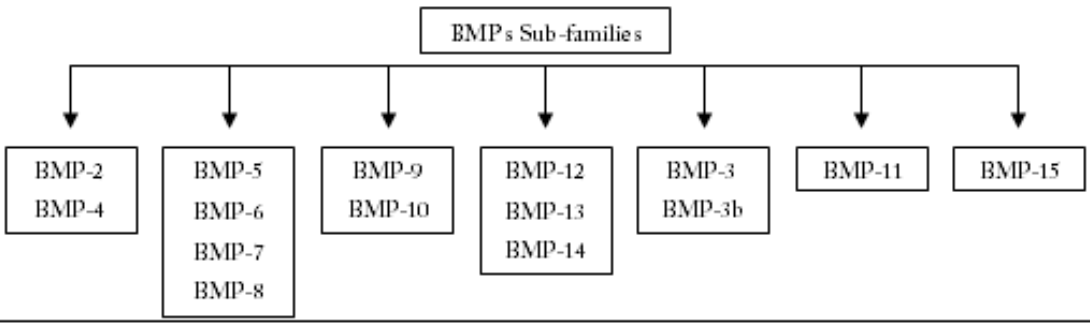



**Figure 2: Family tree of BMPs.** Sub-families within the BMP in relation to the similarity in the amino acid (AA) sequence homology of the 7-cysteine-TGF-β mature domain (Ozkaynak et al., 1992) [44].

Extensive research has shown that several BMPs, but not all, induce ectopic bone or cartilage formation; The physiologic roles of the BMPs include pivotal roles during embryonic development in the specification of the positional information and the development of organs and the requirement for skeletal and nonskeletal tissue formation.[45,40]. Analysis of gene knockouts, in which one or more BMP genes are inactivated, indicate some redundancy in the BMP family [46]. However, each BMP appears to have a unique role in the formation of one or more tissue types [47].

### 1.1.2 Receptor and Signaling

BMP dimers, like others members of the TGF-β superfamily, induce signaling binding with a heterotetrameric receptor complex consisting of two different types (Type 1(BMPR-1) and Type 2 (BMPR-2)) of serine-threonine kinase receptors, known as BMP receptors[42,48].

Even if Type 1 and Type 2 receptors can individually bind BMP, both type of receptors are indispensable for signal transduction [49]; ligands have to bind both receptors to form a high-affinity heteromeric ligand-receptor complex in order to initiate BMP-dependent intracellular signaling through the Smads, a family of intracellular signaling proteins [50].

The overall structures of Type 1 and Type 2 receptors are similar; they are composed of three parts: a short extracellular domain with 10-12 conserved cysteine residues, a single transmembrane domain and an intracellular domain with serine/threonine kinase motif [42]. Preceding the serine/threonine kinase domain, Type 1 receptors, but not Type 2 receptors, have a domain which contains a characteristic SGSGS motif (GS domain), which plays an important role in signal transduction [48]. In mammals, BMPs bind two different Type 1 receptors, BMP receptor Type 1A (BMPR-1A, also denoted Activin receptor-like kinase (ALK)-3) and Type 1B receptors (BMPR-1B, also called ALK-6), in presence of BMP Type 2 receptors [48,51]. The structures of the two BMP Type 1 receptors are similar, but they activate Smad proteins in a partly different pattern [52]. In the TGF-β receptor system, ligands bind to Type 2 receptors even in the absence of Type 1 receptors, but Type 1 receptors can bind ligands only in the presence of Type 2 receptors [53]; However, whereas BMP Type 1 receptors are able to bind ligands even in the absence of Type 2 receptors, the binding to Type 1 receptors is facilitated by the presence of the Type 2 [51,54]. So, signal transduction for BMPs requires both receptors: Type 1 receptors are activated by Type2 and signals are mediated through Type 1; thus, signal specificity is determined by Type 1 receptors [55]. Members of the BMP family show different binding affinities to different combinations of Type1

and Type 2 receptors, and these differences may be related to the overlapping but strictly different functional effects of BMPs [42,56].

On the inside of the cell membrane, BMPs activate the intracellular signalling through a receptor regulated class of Smads (tumor-suppressor Smad proteins), a family of eight intracellular signaling proteins phosphorylated by Type 1 receptor kinase [55]; in this way, receptors for different factors are able to transmit specific signals along the pathway (Figure 3).

There are three types of Smads: signal-transducing receptor-regulated Smads or R-Smads (Smad1, Smad5, and Smad8), common mediator Smad or Co-Smad (Smad 4 and Smad 3) and inhibitory Smads or I-Smads (Smad 6 and Smad 7). BMP receptors are able to target R-Smads; the phosphorylation of these proteins induces their association with a Co-Smad, and the formation of heterodimeric complexes induce their accumulation in the nucleus, where they recruit transcription factors, co-activators and co-repressors in order to regulate the transcription of target genes by binding to specific DNA sequence [50]. BMP targets the genes of tissue-determining transcription factors that promote differentiation of Mesenchymal Stem Cells (MSCs) toward the osseous cell phenotypes [57].

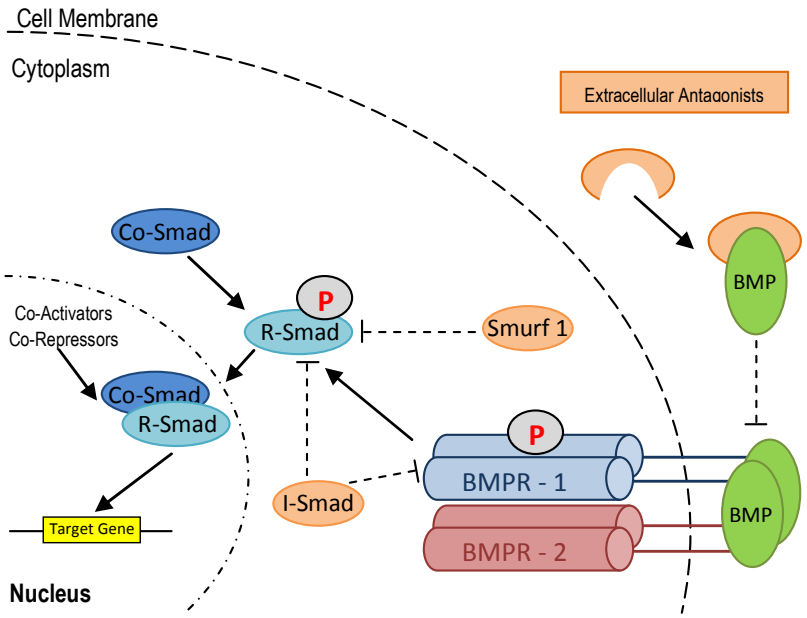



**Figure 3: BMP binding and intracellular signal transduction** [42,55]. A BMP dimer binds with the BMP receptors (BMPR). Following binding, BMPR-2 phosphorylates BMPR-1 and activates intracellular signaling through the Smads. R-Smads forms heterodimeric complexes with Co-Smad and enter the nucleus where alter the transcription of target genes. BMP antagonists can bind with BMP and inhibit the signal cascade. I-Smads can bind with BMPR-1 to prevent it from phosphorylating R-Smads. Smurf interacts with R-Smads and mediates the degradation of these proteins. In the nucleus, there are a number of co-activators and co-repressors that control the transcription of the target gene.

Between the transcription factors present in the nucleus two are specific targets for BMP signalling: Core Binding Factor Alpha 1 (Cfba1) and Osterix. Mice with a homozygous mutation in Cbfa1 died just after birth without breathing; examination of their skeletal systems showed a complete lack of ossification caused by the maturational arrest of osteoblasts, demonstrating that Cbfa1 plays an essential role in osteogenesis [58,59]. Cfba1 and Osterix expression is upregulated by the R-Smads/Co-Smad complex (the complex increase the level of mRNA expression) [60–62]. For example, after stimulation by BMP-2, Smad1 protein is phosphorylated mainly on serine residues and, associated with Smad4, translocated into the nucleus where it induces Osterix expression [57,63].

Smad6 and Smad7, instead, are I-Smad and were shown to bind BMPR-1B inhibiting the phosphorylation of Smad1, suggesting that this class of Smad members could interfere with BMP signalling [52,64].

### i. Regulation

BMP signaling can be tightly modulated at many levels (Figure 3). Extracellularly, soluble antagonists proteins, such as Noggin, Chordin, Gremlin, and Follistatin, can bind only some BMPs, inhibit their binding to cell surface receptors and prevent the signal cascade [50]. Noggin and Chordin were shown to be able to directly bind BMP-4 and BMP-2 and prevent the interaction with the receptors [65,66]; Follistatin was instead shown to bind BMP-7, acting as an inhibitor [67]. Moreover it has been found that BMPs can upregulate the expression of some of these inhibitors, suggesting a negative feedback loop that limits the activity of BMPs. The Tolloid Drosophila gene, on the contrary, encodes a metalloprotease similar to BMP-1 that was shown to inhibit Chordin activity by proteolytically cleavage, acting therefore as a coactivator of BMPs [68].

Intracellularly, the activity of BMPs can be modulated by the activation of I-Smads or by Smurf 1 (Smad ubiquitin regulatory factor, an ubiquitin ligase). As said before, I-Smads can bind with BMPR-1 to prevent it from phosphorylating R-Smads and the subsequent heteromerization with Co-Smad [52,69]; again, BMPs can upregulate expression of I-Smad proteins. Instead Smurf1 selectively interacts with receptor-regulated Smads (Smad1/5), specific to the BMP pathway, in order to trigger their ubiquitination and degradation, and hence their inactivation, thus inhibiting the transmission of BMP signals [70]. In addition, a number of regulators of BMP action exist within the nucleus [50]. All of these regulatory mechanisms together cause the bone-induction process to be controlled tightly and self-limiting [55].

### 1.1.3 Physiological Function

BMPs are a family of multifunctional proteins. Implantation of osteogenic BMPs, such as BMP-2, at an osseous or extraosseous site results in bone and cartilage formation [29]; in contrast, BMP-3 is known to be an inhibitor of bone and cartilage formation [71]. As mentioned before, these molecules are soluble, local-acting signaling proteins. *In vivo*, BMPs act primarily as differentiation factors, turning responsive mesenchymal cells into cartilage- and bone-forming cells [72]. This is supported by many *in vitro* studies showing that the BMPs turn on specific markers of the osteoblast or chondroblast phenotype in a number of cells; for example, rhBMP-2 induces the expression of Osteocalcin, a bone specific protein, in the W-20–17 bone marrow stromal cell line, ROB-C26 calvarial cells, and C3H10T1/2 embryonic cells [72,73]. In addition to inducing differentiation, BMPs can also change the phenotype of some cell types, in particular fat and muscle cells, into that of the osteoblastic lineage; the myoblastic cell line C2C12 stops expression the myoblast lineage markers and induces expression of the osteoblast lineage markers in response to rhBMP-2 [74].

BMPs activities can be categorized into two fundamental physiological functions: the skeletal and organogenic ones.

#### i. Skeletal Function

When implanted *in vivo* BMPs initiate a complex series of cellular events culminating in cartilage, bone, and other connective tissue formation [41,75]. It has been demonstrated that BMPs can regulate growth and differentiation of several cell types [72]; BMPs 2, 4, 5, 6, 7, and 8 are growth and differentiation factors that can individually induce de novo bone formation both *in vitro* and at heterotopic sites *in vivo* [34]. BMP-2 was the first of them that showed its ability to induce the formation of new cartilage and bone tissues, demonstrating that members of this class of molecules were sufficient for osteoinduction [26]. BMPs induce bone formation by promoting the differentiation of MSCs in the bone marrow into osteoblasts and enabling the proliferation of osteogenic cells [10]. The injection of BMP-2 locally, over the surface of calvaria of mice, induces periosteal bone formation on the surface of calvaria without a prior cartilage phase [76]. The over-expression of a dominant-negative truncated BMPR-1B in osteoblast precursor cells inhibits osteoblast-specific gene expression and mineralized bone matrix formation [77]. Transgenic mice in which expression of a dominant negative truncated BMPR-1B transgene is targeted to the osteoblast lineage using the osteoblast-specific Type I collagen promoter, demonstrated a decreased

postnatal bone formation [78]. All these results together demonstrate that BMPs and their signalling receptor play a necessary role in physiological postnatal bone formation [70].

Osseous wound healing requires instead the induction of mesenchymal stromal cells to differentiate along the osteoblastic lineage for new bone formation. BMPs expression by osteoblastic cells has been demonstrated *in vivo* and *in vitro*, however, during fracture healing the immunolocalization of BMP-2 and BMP-4 revealed their limited expression in a small number of primitive cells of the fracture callus [79] and their secretion by macrophages [80]; this data suggest a "domino effect", where cells produce BMPs in order to activate other osteoprogentior cells, which produce more BMPs [81].

#### ii. Organogenic Functions

BMPs are also involved in a number of non-osteogenic developmental processes, regulating growth, differentiation, chemotaxis and apoptosis of cells. These proteins play pivotal roles during embryonic development by specification of the positional information in the embryo and the requirement for skeletal and nonskeletal tissue formation, in the development of several organs and tissues including the nervous system, tooth buds, eye, lung, kidney, prostate, gonads, and hair follicles [40]. Moreover it was well demonstrated that during embryonic development, BMP expression by non-osteoblastic cells results in osteoinduction and bone morphogenesis, and that mesenchymal cells in developing tissues are influenced by BMP prior to any innate ability to produce BMP themselves [82].

Analysis of gene knockouts demonstrated that the elimination of BMP-7 gene results in developmental abnormalities of the kidney and eye, with skeletal deficits [83]. Because BMP-2 is required much earlier in embryonic development mice deficient for BMP-2 are nonviable: homozygous BMP-2 mutant embryos die between embryonic day 7.5 and 10.5 and have defects in cardiac development, manifested by the abnormal development of the heart in the exocoelomic cavity [24,84].

### 1.1.4 Clinical Applications of BMPs

As said before, the repair and replacement of bone is a major clinical problem. The need for functional treatments of fracture non-unions, spinal injuries, and bone loss associated with trauma and cancer has become increasingly common and remains a significant challenge in surgery [6].

The osteoinductive capacity of BMPs has been demonstrated in preclinical models and evaluated in clinical trials [85], and nowadays two rhBMP-based commercial products are on market. OP-1™ (rhBMP-7, Stryker Biotech, Hopkinton, MA), a rhBMP-7 bioimplant, was the first BMP to be approved by the Food and Drug Agency (FDA) in 2001 for the treatment of long bone non-unions [36]. today, OP-1® is also approved for use in spinal fusion surgeries [39]. Infuse® (rhBMP-2, Medtronic, Minneapolis, MN), a rhBMP-2 bioimplant, is approved by regulatory agencies for spinal fusion, tibial fracture repair and for maxillary sinus and alveolar ridge grafting [37,38].

**Table 1: FDA-approved clinical applications of recombinant BMP-2 and BMP-7.**

| Recombinant BMP isoforms | rhBMP-2 | rhBMP-7 |
|---|---|---|
| FDA approval | -Spinal fusion (anterior lumbar interbody fusion)<br>-Open tibial fractures<br>-Sinus lift<br>-Alveolar ridge augmentation | *-Spinal fusion (posterolateral lumbar fusion)<br>*-Long bone non-union |
| | | *Under a humanitarian device exemption (HDE). |

rhBMPs were used in clinical setting at a pharmacologic dose; their pleiotropic effects may be due to a threshold paradigm, in which a low threshold dose of fentomolar concentration is optimal for chemotaxis, a nanomolar range supports mesenchymal cell proliferation, and a low micromolar concentration promotes bone differentiation [20]. Since the half-life of rhBMP-2 is about 6-7 min (tested on non-human primates) due to enzymatic degradation and rapid clearance rate, to increase its effectiveness of healing non-union fractures, rhBMPs are often combined with biocompatible carriers such as absorbable collagen sponges [86,87]. Loading rhBMP into an absorbable collagen sponge (ACS) allows for a gradual proteins release over time, which stimulates new bone formation in the implant site more effectively [33]. In addition to collagen-based scaffolds, calcium phosphate scaffolds have shown promise for use in spinal fusion surgery [88]; since calcium phosphates can provide structural support and osteoinductivity, composite scaffolds composed of BMP-2 loaded collagen and calcium phosphate have been studied and shown a great potential[89,90]. rhBMP-2 delivery and subsequent spinal fusion has also been mediated by biocompatible polymers like polylactide-co-glycolide (PLGA) [91] and polyether-ether-ketone (PEEK) [92]. Specifically, CD-Horizon®, a PEEK-based material made by Medtronic that is supplemented with rhBMP-2 and a

compression resistant matrix composed of calcium phosphate and collagen has been shown to enhance spinal fusion in a Phase III clinical trial [93].

Current clinical applications of rhBMP-based products include spinal fusion, long bone non-unions, and oral and craniomaxillofacial surgeries [94,95] (see the resume in Table 1).

### i. Spinal Surgery

Recombinant human BMP-based therapies have greatly enhanced the outcomes of spinal fusion surgery [96]. This procedure is typically conducted in order to reduce pain associated with abnormal vertebrae motion or to treat spinal deformities; the gold standard graft for spinal fusion surgery has been harvested tissue from the iliac crest which, unfortunately, can often lead to significant pain and morbidity at the donor site [97]. For these reasons synthetic grafts coupled with BMPs have been shown to be clinically viable alternatives.

rhBMP-2 has been studied extensively in preclinical spine fusion models in several species, including non-human primates, demonstrating rhBMP-2 to be equivalent or, in some cases, superior to autogenous bone [31]. These results led to the first human prospective, randomised, multi-centre clinical trial that involved 279 patients with degenerative disc disease (DDD) [98]. in that study rhBMP-2/ACS was added to cortical allografts and it was shown to induce better anterior lumbar interbody fusion (ALIF) results when compared to iliac crest autografts leading to clearance by the FDA for this application in 2002 [38,99]. Spine surgeons then began to develop surgical techniques for conducting interbody fusion procedures through posterior lumbar interbody fusion (PLIF) and, later, through transforaminal lumbar interbody fusion (TLIF) procedures. A clinical trial that involved 67 patients with single-level DDD treated with cylindrical threaded cages filled with either rhBMP-2/ACS or iliac crest autograft using a PLIF procedure demonstrated a better fusion rate for rhBMP-2/ACS [100]. A similar result was obtained in a study on TLIF procedures on 40 patients [101]. More recent studies on the use of rhBMP-2/ACS in single- and multi-level posterolateral spine fusions have shown less complications, decreased need for additional treatment or revision surgery, cost less to treat and an equivalent fusion rate than patients who received an ileac crest bone graft [102,103]. Similar to rhBMP-2, rhBMP-7 showed the capacity to mediate enhanced spinal fusion. To date, research evaluating rhBMP-7 as an

osteoinductive protein for spinal applications shown itself to be safe and effective for the treatment of posterolateral lumbar fusion and cervical non-unions [39,104].

BMP dose escalation studies in animal models have indicated that 3 to 3.5 mg of BMP is sufficient in virtually all cases of spinal fusion to induce new bone and to bridge the osseous defect, and additional doses do not lead to any benefit in terms of fusion rate or the time taken for the fusion to occur [105].

### ii. Fractures and Bone Non-Unions

Long bone fractures make up a large portion of clinically reported fractures[5]. While many long bone fractures can be repaired without surgery, a significant portion of fractures are considered critical-size defects, meaning they commonly form non-unions without a surgical intervention with a bone graft to assist the repair and regeneration. Autografts or allografts are often used, but recently, after many preclinical studies that clearly demonstrated the capability of rhBMP-2/ACS to induce new bone formation in a number of different orthotropic locations in animal models, BMP-loaded implants have become more frequently utilized in clinical practice [106]. The pivotal study on this field was a large clinical study performed in 2002 by a group of surgeons collectively named the BMP-2 Evaluation in Surgery for Tibial Trauma Study Group (BESTT) on 450 patients, showing that rhBMP-2 loaded collagen sponges greatly benefited patients undergoing severe, open tibial fracture repair surgeries [106]. Overall, in this study 74% of rhBMP-2/ACS patients healed without secondary intervention compared to 54% of control patients [106]. Further analyses showed that rhBMP-2/ACS patients had fewer hardware failures and significantly faster fracture healing than control patients, as confirmed even by a subsequent analysis performed on a subgroup patients with the most severe fractures. [106,107]. These results lead to FDA approval in 2004 for the treatment of acute, open tibial fractures [37]. Moreover, a follow-up economical analysis on the BESTT study in the United Kingdom, Germany, and France found that the medical cost-savings of the BMP-2 loaded sponges outweighed their product cost [108]. With the demonstrated ability of rhBMP-2/ACS to induce bone formation there has been interest in applying rhBMP-2/ACS to other orthopaedic trauma applications. In a randomised, controlled, prospective clinical trial, that compared rhBMP-2/ACS combined with cancellous allograft to iliac crest autograft in tibial fractures with critical bone loss, the fractures healed in 13 of 15 patients in the rhBMP-2 group and 10 of 15 in the autograft group [109].

Clinical trials using rhBMP-7 for fracture repair started nearly 20 years ago and, like rhBMP-2, also rhBMP-7 has shown clinical benefits in promoting long bone fracture healing [110]. In 2001, rhBMP-7 was approved by the FDA under the Humanitarian Device Exemption (HDE) for the use in long bone non-unions [36]. The early studies showed that collagen sponges loaded with rhBMP-7 have the same effectiveness in healing tibial fracture non-unions as autografts and that they were able to induce bone healing in critical-size fibular defects similar to demineralized bone matrix [110]. These results have been supported by more recent studies that used rhBMP-7 to guide repair of a wide-range of non-unions (tibial, femoral, humeral, ulnar, and radial) with an high fusion rate [111].

### iii. Dental and Maxillofacial Surgery

Bone grafts are utilized to repair mandibular and maxillofacial defects caused by traumatic injuries, congenital defects, or surgeries for tumor removal [112]; moreover they could be also required in order to create a solid base for dental implants [113]. As already mentioned, autologous bone grafts harvested from various body districts, such as the tibia, ilium or chin, are considered the gold standard for almost all bone surgical procedures, even in oral and maxillofacial reconstructions. However, after the rhBMP-2 bone graft FDA approval for oral and maxillofacial surgeries requiring a large amounts of bone, such as for sinus augmentations, for alveolar ridge augmentations, or for defects associated with extraction sockets, rhBMP products are recommended by surgeons as an alternative to autografts, used alone or in combination to other bone substitute materials [114].

Similar to spine and trauma application, prior to FDA approval, preclinical and clinical research was performed to examine feasibility, safety and efficacy of using rhBMP-2/ACS for the treatment of common oral maxillofacial defects in a number of animal species (included monkeys) and humans [30,115]. First important steps were made in 2003, when a feasibility study on sinus floor augmentation was performed in order to test a single rhBMP-2 concentration (0.43 mg/cc) on ACS, demonstrating that it was successful at inducing bone formation in a non-human primate segmental defect model [116].

After this encouraging result, Boyne performed a large randomised, controlled, rhBMP-2 dosing phase II study for maxillary sinus floor augmentation [117]. Patients were treated with either two different concentrations of rhBMP-2/ACS or with bone allograft as control; histological bone core

biopsies taken at the time of dental implant placement demonstrated normal bone growth and CT scan bone density readings were comparable between all treatment groups following 6 months of functional loading [117]. After identifying 1.5 mg/cc of rhBMP-2 as the most effective concentration, a large randomised, multi-centre, pivotal study was performed examining the safety and efficacy of rhBMP-2/ACS in sinus floor augmentations [118]. A total of 160 patients were treated with either 1.5 mg/cc rhBMP-2/ACS or autogenous bone graft; CT scans prior to and following implant placement and bone core biopsies for histological analysis were analysed, demonstrating that both groups experienced significant formation of new trabecular bone that was biologically and structurally similar to the host site and an implant success rates comparable between groups, concluding that rhBMP-2 and autograft groups performed similarly [118].

Similar results were found in a large clinical study performed on patients requiring extraction socket augmentation; the results demonstrated that sites treated with rhBMP-2/ACS had about two times the amount of bone compared to the empty control group; in addition core bone biopsies showed no differences between the rhBMP-2-induced bone and native one [119].

These research studies demonstrated that rhBMP-2/ACS at 1.5 mg/cc concentration is effective at inducing viable de novo bone formation capable of implant osseointegration and functional loading of dental prostheses [120].

Despite this positive data, compared to spinal fusion and non-union fractures, there are limited cases of clinical studies involving oral and maxillofacial surgeries using rhBMPs, probably due to an unfavourable cost/effective ratio [13].

## 1.2 Biosynthesis, Processing and Secretion in Eukaryotes

Proteins constitutes more than half of the total dry mass of a cell, and their synthesis is fundamental to cell maintenance, growth and development [121]. The synthesis starts when a molecule of mRNA (messenger RNA) is copied from the genomic DNA that encodes for a specific protein and binds a ribosome in the cytoplasm [122]; ribosomes, interacting with tRNA (transfer RNA), mediate the translation of proteins from the nucleotide sequences to the amino acid one [123].

Most of the proteins intended for secretion (like BMPs), or membrane-bound, or that will reside in an organelle, use a pathway named "co-translational translocation". It means that these proteins present an N-terminal signal peptide that is recognized by a Signal Recognition Particle (SRP)

while the protein is still being synthesized on the ribosome; this recognition pauses the synthesis while the ribosome-protein complex is transferred to a specific receptor on the endoplasmic reticulum (ER), where the nascent protein is inserted through a translocon complex named Sec61 [124]. A schematic outline of the organelles that compose the exocytotic pathway in eukaryotic cells is presented in (Figure 4) [125]. It comprise the ER, which is in continuity with the nuclear envelope, the Golgi apparatus, the endosomal/lysosomal system, the secretory granules, the plasma membrane and various type of vesicular intermediates that mediate the dynamic connection between these compartments [126,127].

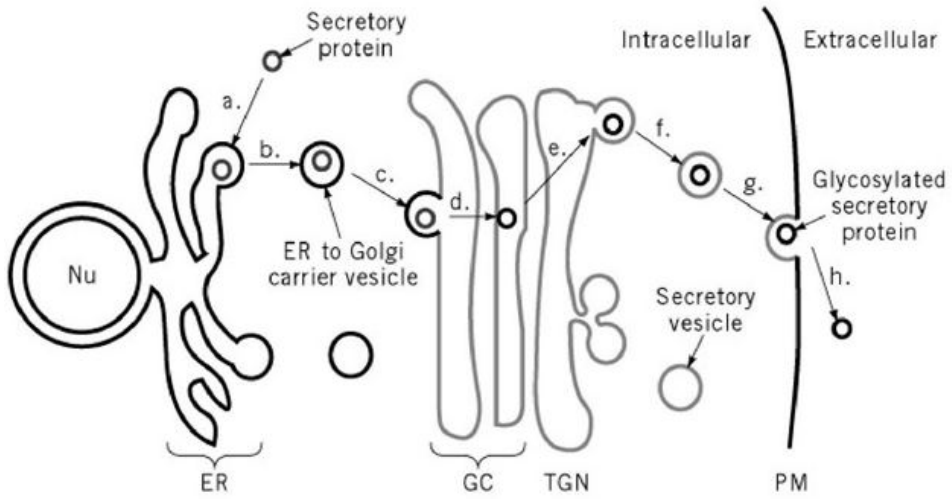



**Figure 4: The eukaryotic secretory pathway.** Secretory proteins enter the membrane trafficking system by (a) translocation from the cytoplasm into the endoplasmic reticulum (ER) (b) In the ER, proteins are packaged into ER to Golgi complex (GC) carrier vesicles that (c) deliver proteins to the GC. As proteins progress through the GC they become (d) glycosylated. Once the (e) reach the /trans-Golgi network (TGN), they (f) are sorted into secretory vesicles in which proteins are (g) transported to the plasma membrane (PM). At the cell surface secretory proteins are (h) released into the extracellular environment.

The ER is the entry site of the pathway and, as said before, targeting to this compartment is mediated by various type of targeting sequence existing in mRNA of proteins [124]. In the ER proteins undergo various co- and post-translational modification, including proteolitic cleavage, glycosylation, folding, and formation of disulfide bonds [128]. If proteins are properly folded they can exit the ER [129]; the passage of proteins from the ER to their final localization involves sequential transit through a series of compartments, collectively referred as the Golgi complex, mediated by carrier vesicles that bud from each compartment and then fuse with the next [127]. The Golgi apparatus is organized into three functionally distinct (but physically associated) regions and is a

complex and dynamic organelle that plays a pivotal role in controlling trafficking and sorting of proteins to various branched routes in the exocytotic pathway [130].

The Golgi apparatus is organized into three sub-compartments regions; the cis-Golgi network (CGN), the medial Golgi (or Golgi stack) and the trans-Golgi network (TGN). The CGN and the TGN are the entry and exit faces of the stack, respectively, and are primarily sorting and distribution centres. The CGN, located on the cis side and including the cistenae and the vacuolar elements linked to the cis side of Golgi apparatus, serves mainly as a filter to remove misfolded or uncorrected proteins [130,131]. The stack cisternae functions as a series of different processing stations that are involved protein modifications, such as the maturation of the saccharide side chain, as well as in forwarding the anterograde traffic to the TGN [127]. The TGN, composed by an extensive tubular/vesicular network on the trans side of the Golgi apparatus, is where proteins with different final destinations diverge to their final destinations [132]; these include the transport to secretory granules, to the endosomal/lysosomal system and the plasma membrane [133]. This mechanism has been confirmed by examining epithelial cells in vitro: it has been found that proteins from ER destined for secretion travel together until they reach the trans Golgi network, where proteins are separated and dispatched in vesicles to the appropriate plasma membrane domains [134]. In general, there are two pathways of protein secretion in eukaryotes: the constitutive and the regulated pathways [135]. In the constitutive pathway, proteins are transported in vesicles to the cell surface at a constant rate without storage or regulation; the vesicles then fuse with the plasma membrane to release their contents by exocytosis [135]. In the regulated pathway instead, newly synthesized proteins destined for secretion are stored at high concentration in secretory vesicles until the cell receives an appropriate stimulus. Unlike the transport vesicles of constitutive secretory cells, secretory vesicles are prevented from fusing with the plasma membrane until the level of a cytoplasmic messenger is altered [135].

For BMP synthesis, the DNA encoding BMP is first transcribed into mRNA in the nucleus. Then, the mRNA is translated by ribosomes into amino acids (AAs) to form a precursor BMP polypeptide. To mature into dimeric BMP, the newly synthesized proBMP has to go through a series of post-translational modifications before finally being secreted outside the cell.

### 1.2.1 Biosynthesis of BMPs

Like other molecules of the TGF-β family, the BMPs are synthesized as pre-pro-peptides of approximately 400 to 525 amino acid (AA) in length[25,40,42]. The pre-pro-peptide consists of an amino terminal hydrophobic secretory signal peptide, a pro-domain, and a carboxyl terminal mature protein. The mature domain of all BMPs contains in their primary structure seven highly preserved cysteine residues. Six of these cysteines form a rigid "cysteine-knot", a structural motif that is highly conserved for all members of the TGF-β superfamily [136,137,41]. A cysteine knot is a structural signature that describes disulfide connectives in which two pairs of disulfide bridges form a ring through which a third disulfide bridge is threaded. The 7th cysteine, instead, is involved in linking two BMP monomers to form a dimer. Three of the BMP family members, BMP-2, BMP-4, and BMP-7 do not contain cysteine residues in their propeptides; the presence of cysteine residues required for dimerization of the propeptide have been implicated in the formation of a latent (inactive) complex between TGF-ββ and its precursor [138]. This indicates that probably these three BMPs are secreted in active forms.

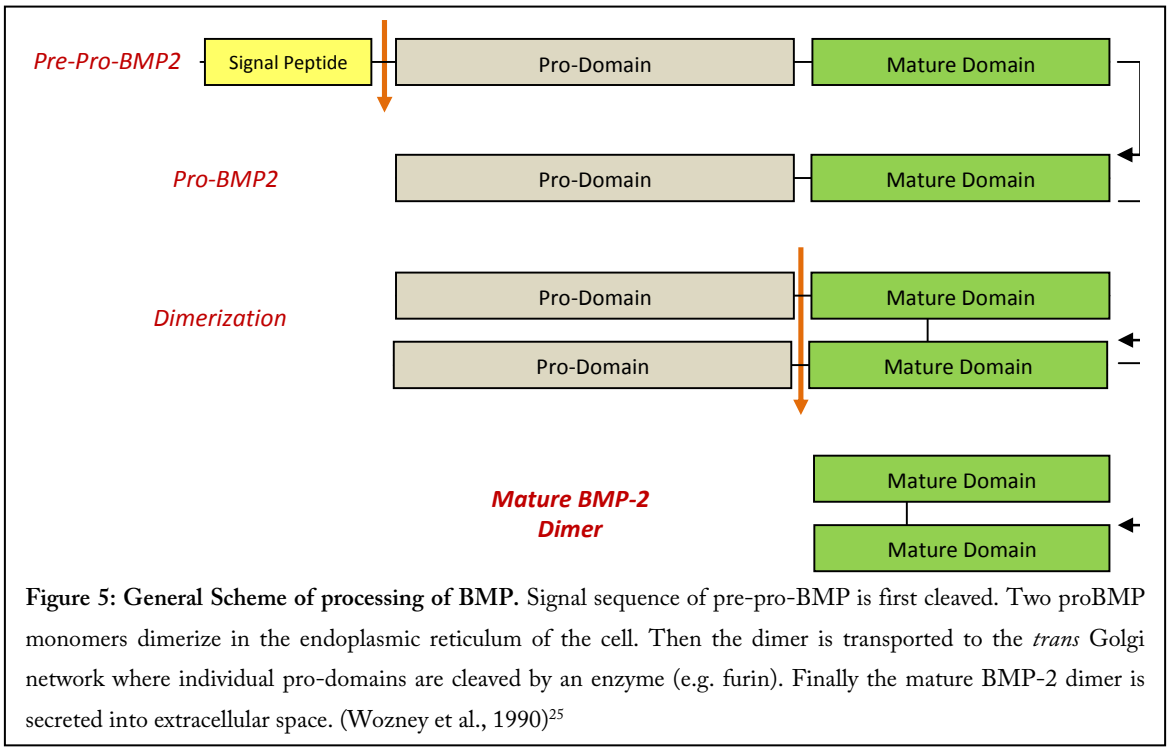



**Figure 5: General Scheme of processing of BMP.** Signal sequence of pre-pro-BMP is first cleaved. Two proBMP monomers dimerize in the endoplasmic reticulum of the cell. Then the dimer is transported to the *trans* Golgi network where individual pro-domains are cleaved by an enzyme (e.g. furin). Finally the mature BMP-2 dimer is secreted into extracellular space. (Wozney et al., 1990)[25]

### 1.2.2 Post-Translational Modifications of BMPs

There are different types of post-translational modifications. These can be grouped into: modification of AAs (such as by addition of phosphate, acetate, or various lipids and

carbohydrates) or modifications resulting in structural changes (like the formation of disulfide bonds, or proteolytic cleavage) [121]. Pre-pro-BMPs have to go through several post-translational modifications before they can be secreted as active proteins.

**Table 2: Primary sequence of a full-length human BMP-2 precursor shows signal peptide, pro-domain and mature BMP-2 region.**

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | MVAGTRCLLA | LLLPQVLLGG | AAGLVPELGR | RKFAAASSGR | PSSQPSDEVL | SEFELRLLSM |
| 61 | FGLKQRPTPS | RDAVVPPYML | DLYRRHSGQP | GSPAPDHRLE | RAASRANTVR | SFHHEESLEE |
| 121 | LPETSGKTTR | RFFFNLSSIP | TEEFITSAEL | QVFREQMQDA | LGNNSSFHHR | INIYEIIKPA |
| 181 | TANSKFPVTR | LLDTRLVNQN | ASRWESFDVT | PAVMRWTAQG | HANHGFVVEV | AHLEEKQGVS |
| 241 | KRHVRISRSL | HQDEHSWSQI | RPLLVTFGHD | GKGHPLHKRE | KRQAKHKQRK | RLKSSCKRHP |
| 301 | LYVDFSDVGW | NDWIVAPPGY | HAFYCHGECP | FPLADHLNST | NHAIVQTLVN | SVNSKIPKAC |
| 361 | CVPTELSAIS | MLYLDENEKV | VLKNYQDMVV | EGCGCR | | |

Two cleavage sites RISR (S2 site)and REKR (S1 site) are both located on the pro-domain (underlined and bolded). Seven cysteines involved in cysteine knot and dimerization are located within the mature domain (bolded, underlined and red). There are three potential N-linked glycosylation sites in pro-domain and one in the mature region (highlighted in yellow).

### i. Dimerization

BMP proteins are active as dimer molecules [25]. They can form either homodimers (made up of monomers of a single BMP subfamily member) or heterodimers (made up of monomers of two different BMP subfamily members) connected by a disulfide bond in a head-to-tail arrangement in the ER [139]. Transient co-expression of BMP-2 with either BMP-5, BMP-6 or BMP-7, or BMP-4 transiently co-expressed with BMP-7, resulted in greater BMP activity than expression of any single BMP [140]; particularly, under certain conditions, heterodimers of BMP-2 and BMP-7 are more osteoinductive than the corresponding homodimers [140]. Currently, both rhBMP therapeutic formulations (rhBMP-2 or Infuse® and rhBMP-7 or OP-1®) are composed of BMP homodimers.

### ii. Glycosylation

Glycosylation, a factor affecting solubility of proteins, is another major post-translational modification with significant effect on protein folding, conformation, stability and activity, that begins in the ER and continues in Cis-Golgi [121]. Carbohydrates in the form of Asparagine-linked (N-linked) or Serine/Threonine-linked (O-linked) oligosaccharides are major structural components of many secreted proteins. In N-linked glycosylation, the oligosaccharide chain is attached by an oligotransferase to Asparagine (Asn) in the tri-peptide sequence Asn-X-Ser, Asn-

X-Thr, or Asn-X-Cys, where X can be any AA, except Proline [121]. BMPs are glycosylated proteins [141]. Unlike TGF-β, all the mature BMP proteins contain between one and three potential N-linked glycosylation sites [142]; the propeptide portion also contains potential N-linked glycosylation sites, similar to TGF-β [25]. A glycosylation site in the centre of mature protein domain is shared by BMPs 2, 4, 5, 6, 7, and 8, but is absent in BMP-3 that has a single glycosylation site at the C-terminus of the mature domain [44].

BMP-2 has one glycosylation site in the mature domain (see Table 2), instead BMP-7 has three potential sites. Recombinant BMPs produced in mammalian cells are post-translationally modified through N-linked glycosylation, ensuring correct folding e activity of the protein [35]; on the contrary, lack of glycosylation lead to a decrease in solubility of BMPs that might be useful to enhance the retention of rhBMPs in organic matrices in order to allow for lower BMP doses to induce bone formation, as first animal study suggests [143].

### iii. Proteolytic Cleavage

BMPs, like many other proteins and like all TGF-β family members , are first synthesized as inactive precursor proteins, which are then turned into the active form via proteolytic cleavage of a multibasic amino acid motif to yield a C-terminal active peptide [25,144]. This processing event has been proposed to regulate the secretion and/or the diffusion of BMPs, thereby controlling the range over which these molecules can signal during embryonic development [145]. Organisms often use pro-proteins when the subsequent protein is potentially harmful, but needs to be available on short notice and/or in large quantities [121]. The cleavage of pro-proteins begins in the trans Golgi network and it continues in the secretory vesicles and sometimes in the extracellular fluid after secretion has occurred [121]. Precursors of BMP are cleaved by proteases that belong to the Proprotein Convertase (PC, or Subtilisin-like Proprotein Convertases (SPCs)) family [144]. Pro-protein convertases are calcium-dependent serine proteases that process a wide variety of pro-protein substrates in eukaryotic cells. They have been shown to cleave at either specific single or pairs of basic residues of the general formula (Arg/Lys)–Xn–(Arg/Lys), where n = 0, 2, 4, or 6 [146]. In mammals, seven members of this family have been characterized and different studies have revealed that in some cases, more than one PC is able to cleave the same substrate [144]; for instance, radiolabeled proBMP-4 incubated with four different PCs in vitro was cleaved by Furin and PC6, but not by PACE4 and PC7 [144]. This suggests that there may be some overlap in the activity of

PCs [147]. The number of cleavage sites varies with different BMPs. For example, both BMP-2 and BMP-4 have two cleavage consensus sites named S1 and S2 [26,35], while BMP-7 possesses only one consensus site [148]. For the BMPs with more than one cleavage site, the order of cleavage may be fundamental; studies have demonstrated that proBMP-4 is cleaved in a sequential manner [144,149]. Cleavage at the S1 site adjacent to the mature domain of BMP-4 allows for the subsequent cleavage at an upstream S2 site within the pro-domain. Study on a mutant form of proBMP-4 with modified S2 site showed that it was cleaved at S1 site to generate 18 kDa mature protein, but the cleaved pro-domains were 35 kDa (larger than the normal 32 kDa pro-peptides because they retains the S1-S2 peptide), indicating that the S2 site cleavage was required to generate 32 kDa pro-peptides. A mutated proBMP-4 with disrupted S1 site was completely resistant to cleavage by Furin, despite containing the native S2 site [149]. From these results, it emerges that cleavage at the S1 site is required for subsequent S2 site cleavage.

Despite proBMP-2 and proBMP-4 have identical cleavage sites, an optimal cleavage site REKR (Arg-Glu-Lys-Arg, named S1 site) and a minimal cleavage site RISR (Arg-Ile-Ser-Arg, named S2 site) (see Table 2), only recent published works demonstrate that PC6, during decidualization, processes proBMP-2 into mature and biologically active BMP-2 [150]. Nevertheless, it is very likely that Furin as well is one of the enzymes that can process proBMP-2. Due to the similarity of cleavage site specificity between the PCs and their overlapping expression in different tissues, it is often difficult to assign cleavage of a given precursor to particular convertases and, as told before, it is also probable that more than one PC is involved in proteolytic cleavage of a BMP.

Like other highly related proteins, mature TGF-β is well known to be secreted by most cell types in noncovalent association with a disulfide-linked homodimer of its pro domain, masking its activity and prolonging its in vivo half-life [151]. It has been reported that after cleavage, even BMPs pro-domain may remain bound to the mature domain non-covalently, thus forming a large complex [35,152]; whether or not the pro-domain remains noncovalently associated with processed BMP-2 has not been established, but is known that an in-frame deletion of the pro-peptide of BMP-2 yielded a polypeptide that was not secreted from the cell, suggesting that the pro-peptide may therefore be involved in processing and secretion of mature BMP-2 protein [35].

### 1.2.3 Secretion of BMPs

BMPs are secretory proteins; however, little is known about BMP secretion. Considering their synthesis as pro-proteins and the presence of multiple basic AA motif Arg-X-Lys/Arg-Arg among various BMPs (e.g. BMP-2/4/7), BMPs may be secreted through the constitutive secretory pathway, like TGF-β1, another member of the TGF-β superfamily [153]. However, currently, there is no direct evidence that supports that BMPs are also secreted from the constitutive pathway.

## 1.3 Recombinant Protein Production for Pharmaceuticals

Once a rarely used subset of medical treatments, protein therapeutics have increased greatly in number and frequency of use since the introduction of the first recombinant protein therapeutic, human insulin, 30 years ago [154]. Only a small number of proteins are expressed in their native cell type under physiological conditions in amounts that permit convenient purification of the relatively large quantities required for research and clinical use; accordingly, various expression systems have been developed to produce recombinant products. In the late 1970s and early 1980s recombinant proteins began to be produced and tested, and during last decades several of them reach the market under a strict regulation [155]. In USA certain of them, like insulin and human growth hormone, were regulated by the Centre for Drug Evaluation and Research under the Federal Food, Drug, and Cosmetic (FDC) Act as drugs, and others, like cytokines, proteins that are involved in the immune response, and blood factors, such as factor VIII for the treatment of haemophilia were regulated by the Centre for Biologics Evaluation and Research under the Public Health Service (PHS) Act [155].
Compared to many small molecule drug products, proteins are usually larger and complex molecules, which, however, have several advantages, mainly because proteins often serve a highly specific and complex set of functions that cannot be mimicked by simple chemical compounds [154]. Particularly, the specificity of proteins activity is an advantage because there is less potential for protein therapeutics to interfere with normal biological processes and consequently cause adverse effects, and because the body naturally produces these proteins, so there is less chance to elicit immune responses. Protein therapeutics can also provide effective replacement treatment for diseases in which a gene is mutated or deleted, without needing for gene therapy, which is not currently available for most of genetic disorders. A study published in 2003 showed that the average clinical development and approval time was more than 1 year faster for 33 protein therapeutics approved between 1980 and 2002 than for 294 small-molecule drugs approved during the same

time period, demonstrating that clinical development and FDA approval time of protein therapeutics is faster than that of chemical drugs [156]. This makes proteins very attractive from a commercial perspective for companies that could be able to obtain far-reaching patent protection for protein therapeutics.

Close monitoring operated by the FDA is necessary because even well-characterized, highly purified recombinant proteins may exhibit minor degrees of structural variability between different batches resulting from variations depending on condition of the source material, processes used to extract and purify the product, and other factors. It is fundamental to remember that the amino acid sequence is the most rudimentary characteristic of a protein; thus technologies based on peptide mapping, protein sequencing, and mass spectroscopy are not sufficient to lead to an FDA approval. Other structural complexities that must be evaluated include: folding of the protein's amino acid chain into organized structures, post-translational modification of the protein and, eventually, association of multiple protein molecules into aggregates. It is the combination of amino acid sequence and structural properties that give a protein its specific activity, and the functional characterization using *in vitro* and *in vivo* tests is also mandatory during the process of development. Potential immunogenicity is another important issue that needs to be strictly investigated; an immune response to a therapeutic protein can range from development of detectable but not clinically significant antibodies, to an immune response with impact on safety or effectiveness [155].

Despite this strict control, to date, more than 100 recombinant proteins with human therapeutic use produced from single cell expression systems have entered the market, and more than 350 are under development, demonstrating usefulness and benefits of this technology [154,157,158].

### 1.3.1 Common Problems Encountered During Production of Recombinant Proteins

Even nowadays, when hundreds of proteins are produced at commercial scale, their production still constitutes a challenge in many cases [157]; moreover, many applications would benefit with higher production efficiencies and consequent lower costs of the final product.

Typical problems encountered during recombinant protein production concern loss of expression, inefficient post-translational processing and inadequate final localization of the proteins.

### i. Loss of Expression

A necessary condition for adequate recombinant protein production is the efficient expression of the gene of interest; however, expression can be lost due to structural changes in the recombinant gene or disappearance of the gene from host cells. The gene of interest can be located in plasmids, or integrated to the host's chromosome, or delivered by a virus, and the strategies in order to avoid a loss of expression are different [157].

- *Plasmid-Based Systems*

Plasmids are extrachromosomal self-replicating cytoplasmic DNA elements that have been used as molecular vehicles for recombinant genes since the dawn of genetic engineering because of their easy genetic manipulation. Plasmid-based expression is the most popular choice when using prokaryotes as hosts, and plasmid copy number is an inherent property of each expression system and depends on the kind of plasmid, the host, and the culture conditions[159,160]. In particular, plasmid copy-number is regulated by copy-number control genes and it can range from a few up to 200 [161]; the metabolic load imposed by plasmids presence increases with increasing in the size of the insert, temperature, expression level, recombinant protein yield, and toxicity of the expressed protein toward the host [162]. As copy number increases, the metabolic load increases, consequently, growth rate decreases and faster-growing plasmid-free cells eventually overtake the culture [162]. Plasmid loss is the main cause of reduced recombinant protein productivity in plasmid-based systems; an unequal plasmid distribution upon cell division, called plasmid segregational instability, will eventually lead to plasmid-free cells, depending from a random distribution between daughter cells [161]. The higher the number of plasmid copies, the lower the probability of a plasmid-free daughter cell.

Another factor that induce loss of expression is plasmid multimerization, an event that occurs when plasmid copies having the same sequence recombine forming a single dimeric circle with two origins of replication. This results in fewer independent units to be segregated between daughter cells, and consequently plasmid loss can increase [163]. It is possible to prevent plasmid instability by adding a competitive advantage for plasmid-bearing cells over plasmid-free cells; the strategy most commonly used is to introduce into the plasmid a gene that provides resistance to particular antibiotics [164]. Unfortunately, a large use of antibiotics has many disadvantages, because they are expensive and their presence is undesirable in food and therapeutic products as well as in the exhausted culture broth that is discharged to wastewater treatment facilities of large scale

fermentation operations. Accordingly, other forms of selective pressure have been explored, such as deletion of an essential gene from the bacterial chromosome and its inclusion in the plasmid, or the introduction of a growth repressor in the bacterial genome and its antidote in the plasmid [165].

Another important issue in plasmid-based systems is plasmid copy number. Although high plasmid copy numbers are generally used for improving recombinant protein yield, this might not always be true. For instance, high copy numbers may drive high protein production rates, which can result in protein aggregation and deficient posttranslational modification [165]. Low recombinant protein yields can also occur in cells with a high plasmid copy number, possibly because of a reduction in translation efficiency. Accordingly, different production strategies should be chosen for different plasmid copy numbers in order to obtain a productive process.

- *Chromosomal Integration*

Chromosome integration is the strategy of choice for the commercial expression of recombinant proteins by mammalian cells. In this case, the long and intricate procedure invested in host development is easily compensated with a stable host [166]. This strategy is a powerful alternative for overcoming problems of expression stability in plasmid-based systems, even though there are several disadvantages. Indeed, adequate integration of a foreign gene in host's chromosome is intricate and time-consuming. Still, a major problem encountered with chromosomal integration is the possibility that the gene of interest will become integrated into an inactive region of chromatin, moreover, chromosome integration typically results in lower production rates than with plasmid-based systems due to a low copy number of the recombinant gene [167]. Nevertheless, methods for obtaining multiple gene integration into the chromosome that yield similar expression levels to those achieved by plasmid systems have been described and the recombinant cells obtained are able to grow in the absence of antibiotics without any reduction of recombinant protein yields [168].

- *Viral Vectors*

Viral vectors are an easy and effective way of delivering genes in host's cells. Some of them, such as retroviruses, promote stable integration of the viral genome into cell's chromosome, others are used for transient expression, but in the second case recombinant protein production occurs only during certain stages of the life cycle of the virus. The simplicity of virus-driven protein expression system makes it useful even for stable production in higher eukaryotes, like animal or plant cells, avoiding

longer cloning procedure. Instead, transient expression is often utilized for rapidly generating little amounts of protein for laboratory scale applications or for preliminary testing of drug candidates [169]. Viral expression systems may also be used for industrial protein production; for example, nowadays, Baculovirus Expression Vector System (BEVS) is utilized to commercially produce several recombinant proteins and especially vaccines [170]. It is mandatory that the products must be virus-free guaranteed, thus, special considerations are required during purification operations. The use of viral vectors involves a process with two different phases: first, cells are grown to a desired cell density, and then they are infected with the virus of interest. One of the most important limitations of expression systems based on viral vectors is the quality of the viral stock because of the necessary serial in vitro passaging can result in the appearance of so called Defective Interfering Particles (DIP), viruses that had several genome deletions that make their replication faster than that of intact ones. Therefore, DIP compete for the cellular machinery and can drastically reduce recombinant protein yields. As DIP replication requires a helper virus, in this case the complete intact virus, their accumulation can be avoided by using lower Multiplicities of Infection (MOI) in order to limit the probability that both an intact virus and a defective one will infect the same cell [171]. The Time of Infection (TOI) refers to the cell concentration at which virus is added to the culture and it is an important parameter: it should be late enough to allow for sufficient accumulation of cells, but should be early enough for nutrients to remain in an abundant concentration to sustain recombinant protein production. The MOI and TOI are closely related and should be selected carefully depending on the particular characteristics of the system of interest. In addition to MOI and TOI, culture conditions can also affect the infection process [169]. A direct relation between the amount of virus attached to cells and recombinant protein concentration has been observed; thus, infection strategies should be aimed at increasing virus attachment, which in turn depends on cell concentration, medium composition, temperature, viscosity, and amount of cell surface available for infection [169].

### ii. Post-translational Processing

As already mentioned, post-translation modifications are an essential step in protein biosynthesis. Most of mammalian proteins need a complex pattern of modifications (e.g. proteolitic cleavage, glycosylation, folding) in order to achieve their functional activity; but not all the expression systems are able to resemble their mammalian counterparts.

While microbial systems, have the advantage of low cost in establishing a production strain, quick production cycle, easy in-process control, and high productivity, the bulk of recent biologics have been produced in mammalian cell lines due to the requirement for post translational modification and the biosynthetic complexity of the target proteins. Nowadays mammalian cell expression has become the dominant recombinant protein production system for clinical applications because of its capacity for post-translational modification and human protein-like molecular structure assembly [172].

- *Folding, Aggregation, and Solubility*

When folding does not proceed adequately, maybe because of cell stress, it results in misfolded proteins that accumulate in intracellular aggregates known as inclusion bodies.[173]. Production of inactive proteins represents an energetic drain and metabolic load, while accumulation of inclusion bodies can cause structural strains to the cell. Accordingly, incorrect protein folding has adverse consequences. Even the overexpression of heterologous proteins often results in the formation of inclusion bodies, mainly because recombinant proteins often reach concentrations higher than physiological ones, which may promote aggregation [174]. During heterologous protein production, high rates of expression are required causing the saturation of the cellular folding machinery [175]; so, a rapid intracellular protein accumulation and the expression of large proteins increase the probability of aggregation [165].

However aggregation is not necessarily a trouble; indeed, it could be an useful way in order to protect proteins from proteolysis and speed their recovery by lysis of the cells and precipitation of the inclusion bodies by centrifugation [165]. In addition, when potential toxicity of the expressed protein to the host is suspected, this deleterious effect can be prevented by producing it as inclusion bodies [176]. Accordingly, in some circumstances, inclusion body formation can be promoted directly through molecular biology or operation strategies, such as the use of protease-deficient strains, or culturing at high temperatures, or designing suitable fusion peptides through protein engineering approaches [174]. In some cases, as for insulin and growth hormone, recovery and renaturation operations can be performed in an economically feasible manner [165]; however, refolding is an empirical process that could be very inefficient, with yields lower than 10%, thus, every case need a careful evaluation of advantages and disadvantages of a solubilisation of the proteins [177]. Moreover, it is almost impossible to predict whether a protein will aggregate or not in a particular expression system for sure, or how easily it will be solubilized and renaturated, thus, soluble proteins

production is generally preferred [165]. A strategy used in order to facilitate protein folding and reduce aggregation is cloning of chaperones and foldases into hosts [175]; however, since it is not possible to predict which chaperone will facilitate folding of a particular protein, the overexpression of more than one chaperone has been explored with satisfactory results [174,175]. Protein engineering can also reduce aggregation by changing the extent of hydrophobic regions or using fusion proteins[174]; both these strategies allow the production of soluble functional protein in the hosts [178].

- *Proteolytic Processing*

When recombinant proteins are expressed in host system, its sequence could include fragments, like the signal peptide or the pro-domain, that are needed in order to direct proteins to the various cellular compartments or to permit the correct folding, but that must be cleaved to obtain a functional protein. The signal peptide, for example, is removed upon membrane translocation, by a signal peptidase complex that is membrane-bound to the ER in eukaryotes or to the cellular membrane in prokaryotes [181]; an inefficient removal of it may result in protein aggregation and retention within incorrect compartments and, consequently, the yields of secreted proteins can be reduced. Study on overexpression of signal peptidase, the enzyme responsible for signal peptides removal, in bacteria, demonstrates an increasing of the release of mature recombinant secretory proteins [182]. Despite these promising results, signal peptidase overexpression in bacteria has rarely been used [157].

Other proteins, such as proteases or insulin, must be expressed as pro-proteins because their pro-domains act as folding catalysts [183]; in these cases, cells utilize endoproteases to produce a mature active protein [184]. Overexpression in mammalian and insect cells of the mammalian endoprotease furin increased the concentration of correctly folded product demonstrating itself a viable option to increase recombinant proteins production [185].

- *Glycosylation*

Glycosylation is a complex post-translational modification that requires several steps and enzymes and that determines protein stability, solubility, antigenicity, folding, localization, biological activity, and circulation half-life. As told before it usually occurs in the ER and cis-Golgi apparatus of eukaryotic cells [130]. Three types of glycosylation exist: N-(glycans linked to an Asn of an Asn-X-Ser, Asn-X-Thr, or Asn-X-Cys consensus sequence, where X is any amino acid except proline), O-(glycans linked to a Ser or Thr), and C (attached to a Tryptophan) linked [128]. Of these, N-linked

glycosylation is the most studied and is considered as the most relevant for recombinant protein production. Glycosylation profiles are protein-, tissue-, and animal-specific [186]; moreover, different glycosylation sites are often glycosylated in different ways [186]. Recombinant proteins may therefore present macroheterogeneous (differences in site occupancy) or microheterogeneous (differences in the structures of oligosaccharides between glycosylation sites) glycosylation [187], but nonauthentic glycosylation may trigger immune responses when present in proteins for human or animal use [188]. Accordingly, authentic glycosylation pattern is especially relevant for recombinant proteins to be utilized as drugs, but, given the complexity of the process, several bottlenecks are present.

For example, Dolichol-Phosphate is a molecule that functions as a membrane anchor for the formation of the oligosaccharides that will be transferred from the Dolichol donor to Asn residues of newly forming polypeptide chains[187]; So the synthesis of the dolichol-phosphate oligosaccharide can limit the extent of glycosylation. As an attempt to solve this, Dolichol-Phosphate has been fed to CHO cells producing recombinant proteins, but, although it was internalized, no increase in glycosylation site occupancy was observed[188]. Also sugar nucleotides, required for oligosaccharide synthesis, can limit the build-up of the polysaccharide precursors and reduce the glycosylation site occupancy[189]. Limitation of sugar nucleotide donors can occur upon prolonged glucose or glutamine starvation[189]; in order to alleviate such a problem, sugar nucleotide precursors are added to the culture medium[190]. Even the transport of sugar nucleotides to the ER and Golgi apparatus can affect glycosylation[190].

Moreover, protein glycosylation is a dynamic phenomenon that changes as culture time progresses[191]; glycosylation levels it has been demonstrated to increase with increasing culture time until the onset of cell death[188].

- *Other Posttranslational Modifications*

Other posttranslational modifications, such as myristoylation, palmitoylation, isoprenylation, phosphorylation, sulfation, C-terminal amidation, ββ-hydroxylation, and methylation, are less common than glycosylation, but may be important for certain recombinant proteins[192]. In general, the extent of modification depends on the host utilized, being the modifications performed by higher eukaryotic cells closer to those found in human proteins.

### iii. Transport and Localization

Recombinant proteins may be directed to different cellular compartments by signal peptides or through fusion proteins, and it is very important that characteristics of the protein were carefully considered when deciding the site of accumulation. For example, small proteins susceptible to proteolysis should be produced in bacteria as inclusion bodies. Apart from intra- or extracellular accumulation, certain applications may require recombinant proteins to be targeted to a specific compartment, like the cell membrane or the ER, through fusion proteins or by the addition of aminoacidic motif like the KDEL one[193]. The transport efficiency of the protein of interest depends on the signal peptide utilized, which should be chosen according to the host; indeed nonoptimal selection of the signal peptide may result in intracellular protein accumulation and aggregation. In some cases, in bacteria expression system, the accumulation of recombinant proteins in the periplasm often results in soluble and correctly folded proteins, whereas cytoplasmic localization yields an inactive and insoluble product[194].

### 1.3.2 Overview of Current Expression System

As said before, efficient strategies for the production of recombinant proteins are gaining increasing importance, as more applications that require high amounts of high-quality proteins reach the market. Higher production efficiencies and, consequently, lower costs of the final product are needed for obtaining a commercially viable process. Given size and complexity of proteins, chemical synthesis is not a viable option for this attempt; instead, living cells, like bacteria, yeast, or mammalian cells, with their cellular machinery are usually harnessed as factories to build and construct proteins based on supplied genetic templates. Indeed, unlike proteins, DNA is simple to construct and manipulate using well established recombinant DNA techniques; thus, DNA templates of specific genes, with or without add-on reporter or affinity tag sequences, can be constructed as templates for protein expression [157].

Traditional strategies for recombinant protein expression involve transfecting cells with a DNA vector that contains the template and then culturing the cells so that they transcribe and translate the desired protein. Typically, the cells are then lysed in order to extract the expressed protein for purification. Both prokaryotic and eukaryotic *in vivo* protein expression systems are widely used and the selection of a system might depend on the type of protein, the requirements for functional activity and the desired yield.

Nowadays biopharmaceutical products are currently made using micro-organisms[195], animal cell lines [196], plants[197] or animals[198] (see the resume in Table 3). Plants, particularly, have been proposed as an attractive alternative for pharmaceutical protein production to current mammalian or microbial cell-based systems. Eukaryotic protein processing coupled with reduced production costs and low risk for mammalian pathogen contamination and other impurities have led many to predict that agricultural systems may offer the next wave for pharmaceuticals production[199].

**Table 3: Overview of most common cellular expression systems and their main characteristics.**

| System | Overall Cost | Production Timescale | Scale-up Capacity | Product Quality | Glycosylation | Contamination Risks |
|---|---|---|---|---|---|---|
| Bacteria | Low | Short | High | Low | None | Endotoxins |
| Yeast | Medium | Medium | High | Medium | Not Correct | Low Risk |
| Mammalian Cell Culture | High | Long | Low | Very High | Correct | Viruses, prions |
| Transgenic Animals | High | Very Long | Low | Very High | Correct | Viruses, prions |
| Transgenic Plants | Very Low | Long | High | High | Minor Differences | Low Risk |

### i. Cell-free Protein Expression

Cell-free protein expression is the *in vitro* synthesis of protein using translation-compatible extracts of whole cells. In principle, whole cell extracts contain all the macromolecules components needed for transcription, translation and even post-translational modification, these components include RNA polymerase, regulatory protein factors, transcription factors, ribosomes, and tRNA. When supplemented with cofactors, nucleotides and the specific gene template, these extracts can synthesize proteins of interest in a few hours.

However this system is not sustainable for large scale production. Although the low yield of product, cell-free protein expression systems have some advantages over traditional *in vivo* systems, like enabling protein labelling with modified amino acids, as well as expression of proteins that undergo rapid proteolytic degradation by intracellular proteases [200]. Also, with the cell-free method, it is simpler to express many different proteins simultaneously [200].

### ii. Bacterial Expression System

Bacterial protein expression systems are popular because they are inexpensive, offer rapid culture times and may to achieve high biomass and high protein yields [201]. Today, one third of FDA approved protein therapeutics are produced in *Escherichia coli*, (*E. coli*) indicating that it is a major workhorse for recombinant therapeutic production [202].The Gram-negative bacterium *E. coli* was the first organism utilized for the production of recombinant human proteins, and is the most attractive because of its well-characterized genetics, versatile cloning tools and expression systems, and the fact that it has been successfully used to express vastly different proteins. It is also advantageous to use *E. coli* for industrial scale production because of its rapid growth, low-cost media, ease of scale-up and capability to produce therapeutics with high yield and quality [202]. However, despite these advantages there are also several limitations to using *E. coli* as an expression host; among the most constraining ones there are the inability to perform certain posttranslational modifications (such as glycosylation) and insufficiencies in proteolytic protein maturation and disulfide bond formation [203]. Such drawbacks prohibit *E. coli* from expressing some complex therapeutics, such as monoclonal antibodies (mAbs), where both correct folding and glycosylation play crucial roles in their biological activities. Also, as discussed in previous section, in *E. coli*, like in other bacteria, many proteins become insoluble as inclusion bodies that are very difficult to recover without harsh denaturants and subsequent protein-refolding procedures, and there is a problem of variability in the level of expression [159,173]. Another concern about the expression of recombinant proteins in bacteria include pyrogens and endotoxins, that must be totally eliminated from proteins to be injected in animals or humans. Despite all this drawbacks, in some specific conditions, like the case of a protein not requiring any of these modifications for obtaining an adequate product, bacteria can be host of choice [201].

Gram-positive bacteria have also been utilized for recombinant protein production, with the advantage that they can secrete large amounts of properly folded product and that contain lower concentrations of pyrogens [204]. Nowadays, *Bacillus subtilis (B. subtilis)* is generally recognized as safe and can be used for the production of proteins for the food industry [205]; however, recombinant plasmids are not stable in this system, so chromosomal integration is the only way to obtain a stable recombinant cell. Moreover, due to the high activity of endogenous proteases, yields are lower than those of Gram-negative bacteria [201].

### iii. Mammalian Expression System

Although animal cell cultures have been important at a laboratory scale for most of the last 100 years, it was the initial need for human viral vaccines (e.g. for poliomyelitis) in the 1950s that accelerated the design of large-scale bioprocesses for mammalian cells, but their complexity delayed their genetic manipulation to the time when manipulation of bacterial genomes was performed almost routinely [207]. Mammalian *in vivo* expression systems usually produce functional protein because of their proper protein folding, assembly, and post-translational modification, but the yield is low, cost of production is high and mammalian cell culturing is time-consuming [207]. In addition, subtle differences exist between glycans obtained in commonly used mammalian cell lines and those associated with glycoproteins synthesized in human cells [208]. For example, cell lines like CHO lack a α2,6-sialyltransferase enzyme, which exists in human cells; however, recently, cloning glycosyl-transferases into common mammalian cell lines has proved useful for the expression of humanized N-glycoproteins and O-glycoproteins [209]. Moreover, successful recombinant protein production in animal cells had to overcome many hurdles, such as the cellular fragility and the complex nutritional requirements of cells [157]. Animal cells require hormones and growth factors that were initially supplied by bovine serum, but possible contamination of the final product with virus or prions, and the difficulty of recovering extracellular proteins from serum-containing media, have resulted in the development of serum-free media that are used for large-scale production [210].

The development of mammalian cell culture methodologies included designing a variety of vector systems and of gene transfer methods [211]; a recent development in gene transfer is the use of Baculovirus vectors, that carry promoters that are efficiently transcribed in cultured mammalian cells.

Other issues that arise when expressing proteins in mammalian cells can be solved through cell engineering. For example, when large scale production is engaged, the cells suffer metabolic pressures, such as oxygen depletion and toxic metabolite accumulation, which affect final yields. Traditional strategies for productivity optimization involve manipulation of cell division as well as cell longevity, supported by the increasing knowledge of cell cycle control [212]. An interesting approach with CHO cells consisted of engineering their mRNA translation initiation machinery with the aim of leaving it on, despite stressful conditions derived from large-scale production [213].

Recently, the productivity of mammalian cells cultivated in bioreactors has reached high yield in recombinant protein production, which mainly resulted from improvements of cell line development through effective selection methods, media optimization, and process control.

Currently, more than 50% of therapeutic proteins approved and on the market are produced using mammalian cells [172,207].

### iv. Other Cell Expression System

The yeast *Saccharomyces cerevisiae* (*S. cerevisiae* ) was the first yeast species to be manipulated for recombinant protein expression, and many proteins have been produced in it [214]. *S. cerevisiae* grows rapidly in a relatively simple medium to high densities with a high level of product secretion and it also demonstrates a eukaryotic subcellular organization capable of some posttranslational modifications. Moreover, this yeast produces a reasonable amount of recombinant proteins and it also secretes the proteins outside the cell envelope, which reduces the cost of downstream processing[215]. However, because this yeast lacks the amylolytic activity necessary for starch utilization, the genes encoding Glucoamylase (GA) were cloned into this organism in order to extend its substrate utilization range and *S. cerevisiae* has been used as a host organism for the production of several recombinant proteins [216].

As other eukaryotes, yeasts are also capable of performing some posttranslational processing typical of mammalian cells; however, extracellular proteases and differences in glycosylation of proteins expressed in yeast compared to those of mammalian cells, limit their use. N-glycosylation of proteins produced by yeasts are high-mannose (with more than 3 mannose residues) or hyper-mannose (more than 6 mannose residues) types, with terminal α-1,3 linkages, and these forms are very immunogenic to mammals [217]. Even if cell engineering has been utilized for obtaining non-immunogenic glycoproteins from yeasts [218], unmodified yeasts are still suitable for the production of proteins that do not require mammalian-type glycosylation and are resistant to proteases. One of these proteins is insulin, which has been commercially produced in *S. Cerevisiae* after enhancing its folding and secretion capacities through genetic engineering [214]. Other yeasts, such as *Pichia pastoris*, and *Candida boidinii*, are hosts with various recombinant proteins, like the hepatitis B vaccine, human serum Albumin and insulin-like growth factor, within reach or already in the market [219].

*Filamentous fungi* have been utilized for the production of a variety of substances with various applications because of their ability of secrete large amounts of homologous and heterologous proteins [220]. Furthermore, recombinant protein concentration can be increased by fusing the gene of interest with genes of fungal origin [220]. A limitation for *fungi* utility for recombinant protein production is the presence of protease, even if promoting growth in pellets and controlling pH can

reduce their activity [221]. Furthermore, as in yeasts, filamentous fungi produce high-mannose-type glycans, easily recognized by mammalian lectins; therefore, recombinant proteins intended for therapeutic use and expressed in fungi can be rapidly and inconveniently cleared from blood [222].

### v. Transgenic Animals

The potential of using transgenic animals as bioreactors for production of complex and high-value proteins has been explored in recent years. Produce recombinant proteins in a transgenic animal system can be advantageous because of its ability to process highly modified proteins [223]. This system is also flexible: processing enzymes can be co-expressed with foreign products to obtain more recombinant proteins. The first transgenic livestock were produced in 1985 by microinjection of foreign DNA into zygotic pronuclei; this was the method of choice for more than 20 years, but more efficient protocols, including somatic cell nuclear transfer and lentiviral transgenesis, are available nowadays [224]. In last years, transgenic farm animal production for biomedical applications has found broad acceptance; the European Medicines Agency in 2006 approved the commercialization of the first recombinant pharmaceutical protein, Antithrombin, produced in the mammary gland of transgenic goats [225].

The production of recombinant pharmaceutically active human proteins in the mammary gland or blood of transgenic animals has advanced to the stage of commercial application [226]; mammary gland is the preferred production site mainly because it lactates a large quantity of protein in milk on a daily basis and because there are established methods for extraction and purification of the respective protein [223]. Guidelines developed by the FDA require, beside sequence validation of the gene construct and characterization of the isolated recombinant protein, also a strict monitoring of animals health in a specific pathogen free facility, in order to avoid the possibility of disease transmission from animal to human, that is one of major concern with transgenic animals [227]

Several products, including monoclonal antibodies derived from the mammary glands of transgenic goats, sheep and rabbits have progressed to advanced clinical trials [225]. An interesting new development is the generation of trans-chromosomal animals carrying the complete sequence of the human immunoglobulin heavy and light chain loci on a human artificial chromosome (HAC); trans-chromosomal bovine clones expressed human immunoglobulin in their blood [228]. This system is a significant step forward in the production of therapeutic human polyclonal antibodies [228]. Recently, the possibility to use transgenic chickens as bioreactors in order to synthesise therapeutic proteins as a component of egg yolk or white, has been investigated; the use of these animals may

have several advantages over mammalian expression systems, including a shorter timescale for setup and ease of scale-up to a transgenic flock, lower costs associated with husbandry, the potential to produce proteins that are toxic to mammalian cells, beneficial glycosylation profile, and reduced immunogenicity of the purified product [229].

### vi. Transgenic Plants

Starting from the end of '90s, many studies have shown that recombinant proteins production in plants has many practical, economic and safety advantages compared with more conventional systems, and so the use of plants for large-scale protein synthesis have started gaining wider acceptance [230,231]. Plant cells, as eukaryotic systems, possess the features to produce biologically active therapeutic proteins; indeed, different kind of proteins, such as enzymes (i.e. trypsin and lysozyme), monoclonal antibodies (i.e. IgA, IgG and IgM) and antigens for edible vaccines are currently produced in plants [232]. Soybean, tobacco and rice, are nowadays the most commonly used plants for transgenic protein production [230].

Transgenic plants are relatively cheap to grow on a large scale; there are established practices for effective harvesting, transporting, storing and processing of plants; other potential advantage of this system is the elimination of purification process when the plant tissue containing recombinant proteins can be used as food and the possibility that recombinant protein can be produced and accumulated in different organelles [231]. For example, recombinant proteins can accumulate in seeds, which can be harvested in dry state and then can be stored without concern of protein degradation [232]. Recombinant proteins can be also directed in a specific compartment of the cells in order to increase the level of final product, that can otherwise result insufficient without any kind of accumulation strategy. As a matter of fact, subcellular targeting is one of the most important factors governing the yield of recombinant proteins, but it is important to remember that the targeting affects also the interlinked processes of folding, assembly and post-translational modification [230].

## 1.4 Recombinant BMP-2 Production for Pharmaceutical Purpose

Since bone contain very low amounts of native BMPs, all methods to extract and purify biologically active BMPs from it are time-consuming, labour intensive, and most importantly, result in a very low yield [233,234]; for instance, beginning with kilogram quantities of bone, microgram quantities of

purified osteoinductive material eventually were obtained, and biochemical analysis indicated that despite the extensive purification, multiple proteins still were present [25,28]. Even more, the potential health risk associated with their isolation from allogeneic donor bone limits their clinical application [235]. For these reasons, when at the end of the 1980s the coding sequences for BMP family members were first cloned and expressed, recombinant human BMP-2 has started to be obtained from cellular expression systems, opening the possibilities for therapeutic application [29].

Over the '90s, several attempts to express active BMPs in different cellular systems (such as *E. coli, P. pastoris*, Baculovirus/insect cells, and mammalian cells) have been made [236], but, because of problems of glycosylation and folding that occur during the production of rhBMP-2 in bacteria and virus infected insect cells [237,238], none of these products reached the market. As a matter of fact, for therapeutic purposes it is necessary to express BMP-2 in large-scale and in a system that ensures correct folding, biological activity, and no immunogenicity, imposing the use of eukaryotic expression systems. One major advantage of recombinant expression of proteins in mammalian cells is that the yield is higher than that produced in the native cells because the level of gene expression in native cells is tightly controlled; conversely, recombinant DNA technology allows strategies to circumvent the natural control of the cells, hence increasing the level of protein production.

For all these reasons, it was only in 2001 that first FDA approved rhBMP obtained from CHO cell lines, reached the market [36,37]. Nowadays rhBMP-2 and rhBMP-7, successfully produced as active protein, have been shown to undergo proper post-translational modifications, are proteolytically processed from their precursor, and most importantly, have been shown to be biologically active, stimulating osteoblastic maturation *in vitro* and inducing bone formation *in vivo* [35,72].

These two BMPs are still the only ones to be approved and commercialized, even if recently, several papers describe the attempt of production of biologically active rhBMP-2 through in vitro refolding of engineered *Escherichia coli* produced inclusion bodies, with biological activity expected to be equivalent to CHO produced rhBMP-2 [238]. However, a disadvantage of these procedures is the requirement of several time-consuming purification and concentration steps before the refolding attempt [239], and the study of these products is still on a preclinical stage [240,241]

### 1.4.1 Recombinant BMP-2 Production in CHO Cell Line

Chinese Hamster Ovarian cell line has become the standard choice of mammalian host cells in recombinant protein production [201,242]. This cell line has been well characterized, allowing for

transfection, gene amplification, and selection of high-end clones. Even if non-human cells might produce proteins improperly glycosylated compared to human cell lines, the major concern with using human cell line for recombinant protein production is the possibility that pathogens infecting human host cells being present in the recombinant protein and thus being transmitted to patients; the chance of disease transmission from non-human cell line-produced recombinant proteins is lower than that from proteins produced by human cells since the disease spread easier in the same species [243].

To establish a cell line, the cDNA encoding BMP-2 protein is inserted into a plasmid with a strong promoter, which drives the expression of the recombinant gene, then this plasmid is transfected into host cells [244]. Non-viral gene transfer methods including calcium phosphate transfection, electroporation, lipofection and polymer-mediated gene transfer, are preferred methods for manufacturing and regulatory purposes [243]. In addition to the recombinant gene, a second gene is incorporated into the plasmid, to give host cells a selective advantage, ensuring that only those cells that express the selector gene survive. Following selection, survivors are transferred as single cells to a second cultivation vessel, and the cultures are expanded to produce a clonal population; individual clones are then evaluated for recombinant protein expression, with the highest producers being retained for further cultivation and analysis [245]. From these candidates, one cell line with the appropriate growth and productivity is chosen for production of recombinant protein and expanded in order to obtain an homogeneous population of cells capable of producing the correct recombinant protein at high level. Then, from a single batch of rhBMP-2 production, the cells are grown and distributed into several small vials, forming a "cell bank", that is a source for all future production.

To promote faster cellular replication and production of rhBMP-2, the cells are transferred to a bioreactor, a computer-controlled, closed-system environment where large-scale production of the protein begins. After a growth period of about three days, the cells are filtered away from the rhBMP-2 containing medium and discarded; the rhBMP-2 moves on to the purification process, that involves a series of four chromatography columns, a virus-retaining filtration step, an ultrafiltration step and a further final nanofiltration step. All these steps are performed in order to assess the total viral safety of the product, even though no human or animal components are added during the recombinant production process. Throughout the production process, quality control testing is done to assess the safety, consistency and purity of all materials by a large number of tests that are completed during the manufacture of rhBMP-2. In addition, viral and microbial agent

evaluations were conducted in accordance with guidance document regarding viral safety evaluation of human and animal cell lines.

### i. Productivity and Efficiency of this Production System

The productivity of recombinant cell lines has increased in the past 20 years. In the 1980s, mammalian cells typically reached a density of about 2 x $10^6$ cells/ml with a batch process production phase of about a week and a specific productivity slightly below 10 pg/cell/day [243]; during the same process performed in 2004, the culture was started at a low cell density of about 100,000 cells/ml and rapidly grew into a density of more than 10 x $10^6$ cells/ml; a high level of cell viability was maintained for almost 3 weeks with a specific productivity up to approximately 90 pg/cell/day [243]. As explained in previous sections, this higher yield obtained in today's processes is the result of years of research that led to a better understanding of gene expression, metabolism, growth and apoptosis in mammalian cells. Overall efforts have led to improvements in vectors, host cell engineering, medium development, screening methods, and process engineering and development.

Despite the success of producing active rhBMP-2, the main challenge of a mammalian expression system is its yield, which has been reported to be lower than other systems, in a μg/L range [35]. From the pioneering work of Israel, assuming a yield of 30 μg/L and a cell density of $10^6$ cell/, the specific productivity is merely 0.03 pg/cell/day after a 24-hr production period, while recombinant antibodies produced by CHO cells the same cell density in a batch system for 3 weeks had a specific productivity of 90 pg/cell/day [243]. Comparing with the specific productivities, the rhBMP-2 production results lower than that of other recombinant products, although considering that the rhBMP-2 was, in 1992, produced in a lab scale setup whereas the recombinant antibodies were produced on an industrial scale, with an optimized manufacturing process.

Efforts made to improve rhBMP-2 productivity in mammalian cell systems were suggested and introduced; they can be divided into two areas. The first involves increasing or maintaining transcription of the transgenes by methods such as improving vector design, plasmid integration and optimizing the chromosomal environment. The second works on maximizing the translational or secretory capacity of the host cells through methods such as host cell engineering, media optimization, and improved bioreactor design and feeding [246].

Although all correct post-translational modifications in rhBMP-2 production can be obtained by the co-expression of soluble exogenous enzyme in CHO cells, the issue of low productivity has not been totally resolved [247].

## 1.5 Rationale of Using Plants for Producing rhBMPs

As told before, despite initial industry conservatism, plants have emerged as one of the most promising general production platforms for biologics, and transgenic plants are nowadays accepted as a safe, effective and affordable system to produce recombinant proteins for medical purposes, including mammalian antibodies, blood substitutes, vaccines and growth factors [248]. Plants allow the cost-effective production of recombinant proteins on an agricultural scale, while eliminating risks of product contamination with endotoxins or human pathogens [249].

Main limitations encountered during plant production systems development include the low yields achieved for some proteins, often caused by poor protein stability, and the difficulties with downstream processing that lead to insufficient product quality [250]; nevertheless, many technical advances achieved over the past years have helped to address these limitations, thus making the prospect of affordable, plant-derived biologics possible [249].

Tobacco plants are the system of choice in which many of the early plant-derived recombinant proteins were produced and extracted, mainly directly from harvested leaves. Tobacco is considered a well-established expression host, with strong transformation procedures available and well-characterized regulatory elements for the control of transgene expression [251]. Moreover, high biomass yields and rapid scalability make tobacco a very suitable host for commercial molecular farming, and it is also a non-food, non-feed crop, carrying a reduced risk of transgenic material or recombinant proteins contaminating feed and human food chains [249], even if it should be considered that it has an high content of nicotine and other toxic alkaloids, which must be removed completely during downstream processing steps [249]. Tobacco has been already adopted as a platform system by several biotech companies that have plant-derived pharmaceuticals undergoing clinical trials, including “Planet Biotechnology Inc.” and “Meristem Therapeutics”.

Knowing the variability of recombinant protein yield obtained from tobacco plant, a strategy that allows to obtain the greater quantity (and quality) of the product is required. As a matter of fact, the level of accumulation of a recombinant protein in transgenic plants is protein specific and strongly influenced by the subcellular compartment of destination [231]; thus, search for the best

subcellular compartment for targeting the protein of interest represents a major issue in the effort to maximize production [252].

The subcellular compartments influence levels of accumulation of proteins, mainly for two reasons [231]. Former, the subcellular environment affect the folding of proteins: the more similar the environment to the native location is, the more probably the protein will be able to acquire the native conformation. Latter, the proteolysis of the protein by host proteases, that occur in the cytoplasm, and the interferences with host molecular function may lead to production drawbacks.

As mentioned in chapter 1.2, ER is a compartment of transit for secretory proteins; it allows the entry into the secretory pathway of newly synthesized proteins and ensures their correct folding, moreover this compartment demonstrated to possess a very low hydrolytic activity and to be very plastic (Figure 6) [253]. When abnormal proteins accumulation in the ER occurs in mammalian cells, it results in large aggregates that in the long term are detrimental and toxic to cellular functions, causing pathologic conditions [254]. Conversely, the ER of plant cells can tolerate high accumulation

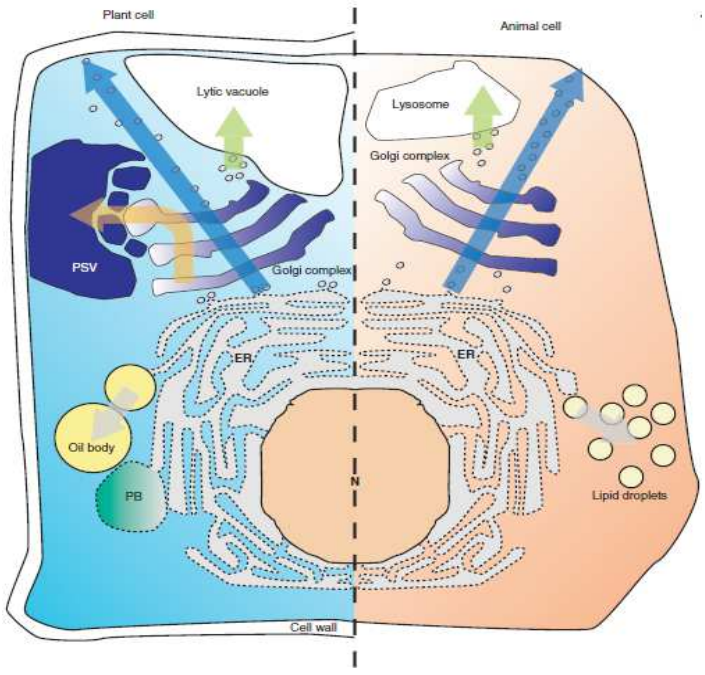



**Figure 6: The secretory pathways of plant cells versus animal cells.** Secretory proteins enter the ER (delimited here by the broken line) in both cell types and travel to the Golgi complex via vesicular trafficking. In the absence of ER-retention signals, soluble proteins are then secreted (blue arrows). Vesicular sorting to the inner hydrolytic compartments (lytic vacuoles in plants, lysosomes in animals; green arrows) requires signals that differ in plants and animals. In addition to lytic vacuoles, plant cells may also have protein storage vacuoles (PSV), which develop according to specific signals and protein sorting mechanisms (orange arrow). Vast amounts of proteins of cereals (e.g., the prolamin storage proteins) form protein bodies (PB) within the ER in aggregation processes that may be useful in the biotechnological production of recombinant pharmaceutical proteins.
Figure taken from Vitale & Pedrazzini, Molecular Interventions 2005.

of proteins without compromising plant development [255]; indeed, plants naturally use ER of determined tissue to accumulate vast amounts of specific storage protein in form of large aggregates, as part of the developmental process of seed maturation [256] (Figure 6).

Prolamins, a class of storage proteins which has nutritional importance, are some of the major storage proteins of cereals' seeds, not found in organisms other than plants [257]. It is important to know that prolamins do not have any known ER-localization signals and that they remains in the ER lumen by forming large aggregates termed protein bodies (PB) [257]. Because the purification of recombinant proteins is a fundamental step both in terms of costs and final productivity, the accumulation of proteins in PB provides an advantage, as their insolubility makes them easy to purify by centrifugation [258].

The 27kD γ-zein (here on termed γ-zein, for simplicity) is among the most ancient maize prolamins [259]. After the co-translational removal of the signal peptide, mature γ-zein consists of two major regions [260,261]: the N-terminal region is characterized by eight repeats of the hexapeptide VHLPPP and by seven Cysteine residues involved in inter-chain bond; the C-terminal region is homologous to 2S albumins, a class of proteins characterized by three domains named A, B, and C, containing eight Cys that form four intra-chain disulfide bonds. [262] (Figure 7).

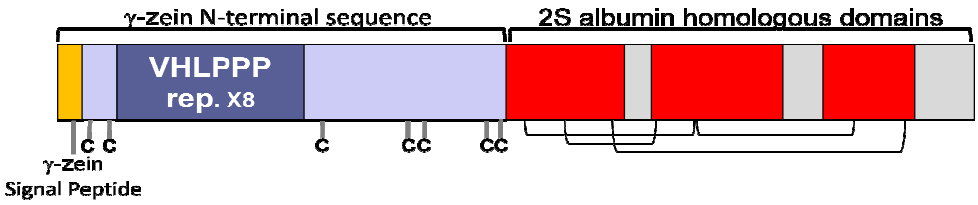



**Figure 7: Schematic drawing of the γ-zein protein.** :signal peptide(yellow); Cys residues in the N-terminal region, repeated domain (blue, 8X PPPVHL), 2S albumin homologous domains (red), and the four putative intra-chain disulfide bonds. (Mainieri et al., 2014) [261]

γ-zein is able to form protein bodies when transiently expressed alone in animal cell lines [263] and when it is stably expressed in vegetative non-seed tissues of transgenic plants [264]. It has been demonstrated that the N-terminal region is retained in the ER even if the 2S albumin-like region is deleted, whereas efficient secretion occurs when most of the N-terminal region is deleted [264,265]. Therefore, a synthetic peptide of the N-terminal proline rich repetitive sequence of γ-zein was synthesized and tested in vitro and, surprisingly, it demonstrated to be able to self assemble and form cylindrical micelles, held together by hydrophobic interactions between the monomers, where the hydrophobic regions of the peptide are clumped in the centre of the aggregate [258,266].

Most important, when the N-terminal fragment, that includes the eight Prolin-rich repeats and the first six Cystein residues, was fused to the C-terminus of phaseolin, a vacuolar storage globulin of common bean, the chimeric construct forms polymers having the main features of γ-zein: insolubility unless reduction and accumulation as PB in the ER [258]. Collectively, these data indicate that the N-terminal region of γ-zein forms inter-chain disulfide bonds and contains information sufficient for ER retention, independent of any specific machinery. Consequently, a fusion protein-based system designed to accumulate recombinant proteins in ER-derived PBs in plant was developed: the approach is based on fusion proteins that contain aminoacids 1–112 of γ-zein sequence, named Zera® sequence (developed by ERA Biotech, Barcelona, Spain) fused to the protein of interest [267,268]. The formation γ-zein-induced protein bodies and an enhanced recombinant proteins accumulation were confirmed both in vegetative tissue and in non plant eukaryotic systems (fungal, insect and mammalian cells) [267,268]; PB formation enables high local concentrations of heterologous proteins to exist within the limited space of the cell, while preserving the protein from normal cellular protein degradation mechanisms, and without subjecting the ER to an intolerable level of stress [252,269].

These results demonstrate that this portion of γ-zein sequences can be used as an alternative to the KDEL signal when accumulation of recombinant proteins in the ER is useful for their production and recovery [270].

## 1.6 Objectives

In an effort to obtain cost-effective production of functional rhBMP2, the purpose of this study is to explore tobacco plants as bioreactors.

Thus, in order to increase the accumulation of rhBMP-2 in these plants, a fusion-protein (named ER-BMP2) between the C-terminal active peptide of human BMP2 and the N-terminal domain of maize γ -zein was obtained. The accumulation level of ER-BMP2 was then compared with that of the native form of human BMP-2 (N-BMP2).

# 2 Materials and Methods

## 2.1 Plant Material

### 2.1.1 Protoplasts Isolation from Leaf Tissue

Protoplasts were isolated from tobacco leaves and subjected to polyethylene glycol-mediated transfection, as described by Pedrazzini et al. (1997) [271].

Briefly, leaves of 3-5 weeks old from axenic *Nicotiana Tabacum* (L. var. Petit Havana SR1) plants were cut underside every 1-2mm and laid down overnight on a digestion enzyme solution, which consisted of a combination of 0,17% (w/v) macerozyme and 0,33% (w/v) cellulase prepared in K3 medium (Gamborg's B5 basal media with minimal organics (Duchefa Biochemie) supplemented with 400mM sucrose, 1,5M xylose, 3mM NH4NO3, 1 mg/l α-naphtalenacetic acid (NAA), 1 mg/l 6-benzylaminopurine (BAP) and 5mM $CaCl_2$). Protoplasts were filtered through a 85μm filter and washed twice with W5 buffer (154mM NaCl, 125mM $CaCl_2$, 5mM KCl, 5mM Glucose) by adding 4 volumes of buffer, resuspending the protoplasts and centrifuging at 60*g* for 10 min (brake off). Protoplasts were then resuspended in 10ml of W5 and incubated 30 min in the dark. An aliquot of protoplasts preparation was diluted 10 times in K3 buffer containing 10 μg/ml fluorescein diacetate (FDA) and 1μl of this diluted preparation was observed at microscope to count fluorescent cells. After a new centrifugation (60*g* for 10 min) protoplasts were resuspended at a concentration of $10^6$ cell/ml in K3 medium.

### 2.1.2 Protoplast Transformation

The protoplasts, isolated as described in the previous section, were centrifuged at 60*g* for 10 min (brake off) and resuspended in MaCa buffer (0.5mM mannitol, 20mM $CaCl_2$ and 0.1% MES pH 5.7) at a concentration of $10^6$ cell/ml. After heat shock (5 min at 45°C) 40 μg of pDHA or pGreen plasmid without insert (as a negative control) or with inserted recombinant coding sequences were used in each transfection at a concentration of $10^6$ cells/ml. Equal volume of 40% PEG solution (0,4M mannitol, 0,1M $Ca(NO_3)_2$, 40% (w/v) PEG) was added dropwise. After 30 min of incubation at room temperature, protoplasts were washed with 10 ml of W5 buffer, resuspended in 1ml of K3 buffer. Protoplasts were then allowed to recover overnight in the dark at 25 °C in K3 medium before performing protein extraction.

### 2.1.3 Wild-Type and Transgenic Plants

Wild-type or transgenic tobacco plants (Nicotiana Tabacum L. var. Petit Havana SR1) were grown in sterile conditions on half-concentrated Murashige and Skoog Basal Medium (½MS, Duchefa Biochemie) supplemented with 10g/l sucrose and 0,8% (w/v) phyto agar (Duchefa Biochemie) at 25°C under 16/8h light/dark cycle.

Transgenic Tobacco plants expressing ER-BMP2 and N-BMP2 were generated by *Agrobacterium*-mediated transformation of plant tissues (leaf-disk). The plants were transformed with *Agrobacterium tumefaciens* strain GV3101 harbouring binary vectors containing expression cassettes with the sequences of the Cauliflower Mosaic Virus (CaMV) S35 promoter and the sequences of the proteins of interest.

Fully expanded tobacco leaves were excised from axenically grown wild-type plants and cut with sterile tools into 1 $cm^2$ leaf discs. The explants were submerged for 10 min in a suspension of *A. tumefaciens* carrying the plasmid of interest, and then blotted on to sterile filter paper to remove the excess *Agrobacterium* culture. Leaf discs were incubated underside down at 25 °C in the dark on shoot-inducing medium (SIM) agar (MS salts containing 30 g/l sucrose and 0.8% phyto agar (Duchefa Biochemie)). After 2 days the leaf discs were transferred to shooter generation medium (SGM) (½MS salts supplemented with 30 g/l sucrose, 0,1 mg/l NAA, 1 mg/l BAP, 50 μg/ml hygromycin, 100 μg/ml carbenicillin and 250 μg/ml cefotaxime) in order to promote callus growth, select transformed plants and to prevent further growth of *Agrobacterium*. Leaf discs were transferred to new media at least every 2 weeks until the growth of calli. When calli were present they were excised and placed on new plates to promote shoots formation. Elongated shoots were excised from calli and transferred on ½MS supplemented with 30 g/l sucrose, 0,1 mg/l indole-3-acetic acid (IAA), 50 μg/ml hygromycin and 100 μg/ml carbenicillin to promote lateral root formation.

After 2 – 3 weeks, regenerated plants were tested for recombinant protein expression and then grown in soil.

## 2.2 Protein Methods

### 2.2.1 Protein Extraction from Protoplasts and Plant Tissue

#### i. Protoplasts

Frozen samples of protoplasts were homogenised with 2 volumes of ice-cold 1,5X protoplasts homogenisation buffer (150 mM NaCl, 1,5 mM EDTA, 1,5% (v/v) Triton X-100, 150 mM Tris-Cl pH 7.5) supplemented with Complete protease inhibitors (Roche) in presence or absence of reducing agent (4% (v/v) 2-mercaptoethanol (2-ME).

Homogenization of protoplasts incubation medium was performed similarly using homogenization buffer supplemented or not with 2-ME.

The homogenates were then centrifuged at 1500*g* for 10 min at 4°C and supernatants and pellets were adjusted for SDS-PAGE.

#### ii. Tobacco Leaves

For total protein extraction from plant tissues, Tobacco leaves were homogenised in ice-cold homogenisation buffer (100mM Tris-Cl pH 7.8, 200mM NaCl, 1mM EDTA, 0,2% (v/v) Triton X-100, supplemented with Complete (Roche)) in presence or absence of 4% (v/v) 2-ME. Ratio volume of buffer/weight of the leaves was 5-7/1.

The homogenates were then centrifuged at 1500*g* for 10 min at 4°C and supernatants and pellets were adjusted for SDS-PAGE experiments.

#### iii. Tobacco Seeds

For total protein extraction from Tobacco seeds, 0.05g of such tissue were homogenised in SDS-PAGE loading buffer (20mM Tris-Cl pH 8.6, 1% (w/v) SDS, 10% (v/v) glycerol, 0,17% (w/v) bromophenol blue and 4% (v/v) 2-ME) at 95°C. the volume of buffer/weight of the seeds ratio was 10/1.

The homogenates were then centrifuged at 14000*g* for 5 min at 4°C. Supernatants were then collected by paying care to avoiding the upper oil phase and 5µl were analyzed with SDS-PAGE.

### 2.2.2 Protein Quantification

Total proteins were quantified using the Bradford protein assays [272]. Briefly, 1 to 10µl of protein samples were mixed with Bradford Reagent (Sigma) up to 1 ml volume; after an incubation of 15 min at room temperature (RT) , absorbance was read at 595 nm. A standard curve was obtained with Bovine Serum Albumin (BSA, Sigma) solution (from 0 to 1000 µg/ml) in Bradford Reagent.

### 2.2.3 Protein Extraction for Transgenic Tobacco Plant Screening

Tobacco leaves were homogenised in SDS-PAGE loading (denaturing) buffer (20mM Tris-Cl pH 8.6, 1% (w/v) SDS, 10% (v/v) glycerol, 0,17% (w/v) bromophenol blue and 4% (v/v) 2-ME) at 95°C. The volume of buffer/weight of the leaves ratio was 5/1.

The homogenates were then centrifuged at 1500*g* for 10 min at 4°C and the supernatants were collected, diluted 1/10 in loading buffer and analyzed with SDS-PAGE and protein blot (used for the screening of F0 positive plants) or directly spotted (5µl) on Protran nitrocellulose membrane (Perkin Elmer) for direct immunodecoration (used for screening of positive F1 plants).

### 2.2.4 Endoglycosidase H Treatment

For Endoglycosidase H (Endo H, New England BioLabs) treatment, protein samples were incubated with 10x Glycoprotein Denaturing Buffer (0,5% (w/v) SDS, 40mM DTT (New England Biolabs, Beverly, MA, USA)) for 10 min at 100°C. Samples were brought to a volume of at least 20µl by adding of 10x G5 Reaction Buffer (50mM sodium citrate pH 5.5) and divided in two equal aliquots. One of them was treated with 2,000 Units of Endo H enzyme (New England Biolabs), the other one with an equal volume of water. After 2 hours incubation at 37°C the samples were supplemented with Laemmli denaturation buffer and analyzed by either SDS-PAGE and protein blotting.

### 2.2.5 Thrombin Cleavage

About 20µg of total protein extracts from transgenic and control plants in reducing condition were incubated with 10 units of thrombin or with same quantity of PBS (as control) at 22°C for 20 h under gentle agitation. After digestion, to verify the occurrence and extent of thrombin digestion, samples were analysed by SDS-PAGE and protein blot with anti-FLAG antiserum.

### 2.2.6 Velocity Centrifugation on Sucrose Gradient

Tobacco leaves or protoplasts were homogenised in ice cold homogenisation buffer in non-reducing conditions; homogenate was centrifuged at 1500*g* for 10 min at 4°C and supernatant was recovered. Extracted proteins were layered on a 11.5 ml linear 5-25% (w/v) sucrose gradient made in 150mM NaCl, 1mM EDTA, 0,1% (v/v) Triton X-100, 50mM Tris-Cl pH 7.5. An additional gradient was loaded with a mixture of markers containing 200µg each of cytochrome C (12.4 kDa), ovalbumin (43 kDa), BSA (66 kDa), aldolase (161 kDa) and catalase (232 kDa). After centrifugation at 39.000 rpm for 20 h at 4°C in a Beckman SW40 rotor (200,000*g* average, Beckman Instruments), gradients were divided into 18 fractions of about 650µl each, and the precipitate at the bottom of the tube was solubilized with 650µl of SDS-PAGE loading buffer. Equal volume (usually 40µl) of each fraction and the solubilized precipitate were analysed on SDS-PAGE.

### 2.2.7 SDS-PAGE and Protein Blot

For protein blot, the protein samples, extracted in presence or absence of reducing agent, were denaturated in 3X SDS-PAGE loading buffer (20mM Tris-Cl pH 8.6, 1% (w/v) SDS, 10% (v/v) glycerol, 0,17% (w/v) bromophenol blue and 4% (v/v) 2-ME) and incubated 5 min at 95°C. Proteins were separated by SDS-PAGE, using a 15% (v/v) acrylamide gel and using an unstained protein molecular weight markers (Unstained Protein Molecular Weight Marker, Thermo Scientific). Proteins were then transferred to nitrocellulose membrane (Protran Nitrocellulose Transfer Membranes, Perkin Elmer), stained with Ponceau-S staining solution (0,1% (w/v) Ponceau-S in 5% (w/v) trichloroacetic acid) to check correct transfer and quantity of proteins and washed in water. For immunoblotting experiments, the membrane was successively incubated in T-TBS (137mM NaCl, 20mM Tris-Cl and 0,2% (v/v) Triton X-100), supplemented with 1% (w/v) Fetal Bovine Albumin (FBS) and 0.02% $NaN_3$, containing primary antibody anti-FLAG (1:2.000 (Sigma)). After incubation, the membrane was washed with T-TBS and incubated with T-TBS supplemented with 5% (w/v) powder-milk containing secondary anti-Rabbit antibody (horseradish peroxidase (HRP) conjugated (1:18.000 (Pierce)) and washed twice with T-TBS and once with TBS (137mM NaCl, 20mM Tris-Cl). The membrane was then treated with enhanced chemiluminescent (ECL) HRP substrate (SuperSignal West Pico chemiluminescence substrate, Thermo Scientific) according to manufacturers' protocol. The signal was detected with Amersham

Hyperfilm (GE HealthCare) or with ChemiDOC Digital Instrument (Bio-Rad). Relative and absolute signal intensity for bands were quantified using Image-Lab (Bio-Rad).

## 2.3 DNA Methods

### 2.3.1 Polymerase Chain Reaction (PCR)

Typical PCR [273] mixture was composed of 1ng DNA matrix, 50 nmol dNTP mix, 40pmol of each primers, 1 U of DNA polymerase (DreamTaq DNA Polymerase, Thermo Scientific) and 2,5µl of 10X polymerase buffer in a total volume of 25µl. Reactions were incubated 5 min at 94°C, before being subjected to 35 cycles of amplification: 30 sec at annealing temperature (according to primers) and 1 min/amplified kb at 72°C. A final cycle of 5 min at 72°C was performed to promote complete elongation.

### 2.3.2 Digestion, Ligation, DNA Purification and Quantification

Restriction enzymes were used for both preparative digestion (prior to ligation) and analytical digestion (after mini or maxi DNA preparation). Typically, plasmidic DNA, from 100ng (for analytical test digestion) up to 500ng (for preparative digestion), was incubated with specific restriction enzymes with corresponding buffer (supplemented with BSA when required) in a total volume of 20 or 50µl. Digestion mixes were incubated for 1 hour or more at the proper temperature.

Digested plasmids were ligated using a 5:1 insert/plasmid ratio. DNA was mixed with 1 U of T4 ligase (New England BioLabs) in corresponding buffer for a total volume of 20µl and incubated at least 24 h at room temperature. Usually, half of the ligation product was used for *E. coli* transformation.

Agarose gels were prepared with 0,8 to 1,5% (w/v) agarose melted in TBE buffer (89mM Tris-Cl pH 8.0, 89mM $H_3BO_3$, 20mM EDTA) containing 0.05% (w/v) ethidium bromide. DNA samples were mixed with 6X DNA loading buffer (40% (v/v) glycerol, 0.25% (w/v) bromophenol blue) and loaded on gels, besides a DNA ladder mix (GelRuler™, Fermentas). DNA was visualised on an ultraviolet transilluminator (UviDoc, Uvitec Cambridge).

Quantification was performed by visual comparison with DNA ladder mix or by measurement with a spectrophotometer (NanoDrop 2000c UV/IV Spectrophotometer ,Thermo Scientific).

For purification of DNA fragments from agarose gels, DNA bands were cut out from gel and DNA was extracted and purified using QIAquick Gel Extraction Kit (Qiagen) according to manufacturers' instructions. Briefly, agarose bands were mixed with proper quantity (2 volume of buffer to 1 volume of gel) of QC (1.0M NaCl, 50mM MOPS, pH 7.0, 15% (v/v) isopropanol) buffer and incubated at 50°C until complete agarose melting. Resulting solution was filtrated through the column by centrifugation (1 min at 11.000g) and washed twice. DNA was then eluted with 30 – 50µl of distilled water. The eluted was finally checked on agarose gel and quantified.

### 2.3.3 Chemical Competent *Escherichia coli* DH5-α Cells: Preparation and Transformation

A 2ml pre-culture of E. coli DH5-α competent cells, grown over night in LB medium (Duchefa Biochemie) at 37°C, was used to inoculate 100ml of LB medium. Cells were grown in a shaking incubator to an $OD_{600}$ of 0,4 and then cooled down for 30 min on ice. Cells were collected by centrifugation at 4000rpm for 10 min at 4°C and resuspended in 50ml of 100mM $CaCl_2$, 70mM $MgCl_2$ and 40mM NaOAc pH 5.5. After incubation for 45 min on ice, cells were pelleted by centrifugation at 4000rpm for 10 min at 4°C and resuspended with 5ml of 100mM $CaCl_2$, 70mM $MgCl_2$, 15% (v/v) glycerol and 40mM NaOAc pH 5.5. Cell preparation was divided into 100 - 200µl aliquots that were incubated on ice for 1 h, then frozen in liquid nitrogen and stored at -80°C.

For transformation, aliquots of competent cells were mixed with half of the ligation product and incubated 20 min on ice. Cells were heat shocked 2 min at 42°C and then diluted in 1 ml LB medium. After 1 h at 37°C different quantity of cells were plated on LB Agar medium (Duchefa Biochemie) supplemented with selective antibiotic (100 µg/ml of ampicillin for pDHA vector, 100 µg/ml of kanamycin for pGreen vector). Cells were grown overnight at 37°C.

### 2.3.4 Competent *Agrobacterium tumefaciens* GV3101 Cells: Preparation and Transformation

*Agrobacterium tumefaciens* is a plant pathogen which causes the formation of crown galls or tumours in tissues infected by the bacterium [274]. Briefly, the determinants for establishing and sustaining tumours are located mostly on large Tumour-Inducing (Ti) plasmids. The T-DNA is a discrete section of the Ti plasmid bounded by 25 bp imperfect repeats termed the right (RB) and left borders (LB). The T-DNA and the virulence (*vir*) region are two distinct regions of all Ti plasmids which are essential for *Agrobacterium*-mediated plant transformation. The T-DNA is transferred to, and integrated in, the host cell nuclear genome at the on-set of infection. The processing of the T-DNA and its transfer to the host plant cell nucleus is achieved primarily by the concerted action of about 20 *vir* gene products. Ti plasmid-encoded *vir* genes can function in trans to promote the transfer of T-DNAs from co-resident plasmids to recipient plant cells. Such T-DNA-containing plasmids are termed Ti vectors. Genes and sequences to be transformed into plants are inserted between the LB and RB of the Ti vector T-DNA.

*Agrobacterium* GV3101 strain carries a disarmed Ti plasmid that possesses the *vir* genes needed for T-DNA transfer, but has no functional T-DNA region of its own. A single colony was inoculated in 3 ml of LB completed with the appropriate antibiotic selection in a 15 ml Falcon tube at 28°C overnight on a roller drum. 0.5 - 1 ml of the overnight culture was then inoculated in 100 ml of LB and grown at 28°C with vigorous shaking until reaching an OD600 of 0.8 and then cooled down for 30 min on ice. Cells were collected in Corex tube and centrifuged at 4000rpm for 10 min at 4°C in a JA-20 rotor. Cells pellet was resuspended in ice cod 10% (v/v) glycerol. After 3 cycles of resuspension and spin of the cells, the pellet was resuspended in 2.5ml of ice cold 10% (v/v) glycerol and divided in 50µl aliquots that were frozen in liquid nitrogen and stored at -80°C.

*A. tumefaciens* transformation was performed by electroporation. Competent cells were transferred into electroporation cuvettes and 50 - 200ng of DNA were added; the cuvette was placed in the GenePulser unit (Bio-Rad), that was set on 2.4 kV, 25 µFD and 400 Ohm for the resistance, and a single pulse of 9 msec was given. Cells were transferred in a 2ml Eppendorf tube with 1ml of LB. After 6 h at 28°C different dilutions of the cells were plated on LB Agar medium (Duchefa Biochemie) supplemented with selective antibiotic (50 µg/ml gentamicin for pHelper positive selection, 25 µg/ml rifampicin for GV3101 strain selection, 5 µg/ml tetracycline for pSoup vector

selection and 100 μg/ml kanamycin for pGreen selection). Cells were grown for at least 36h at 28°C.

### 2.3.5 DNA-Mini and Maxi Preparation

DNA mini-preparations were performed using the QIAfilter Plasmid Mini Kit (Qiagen) and following the manufacturers' instructions. *E. coli* cells from a 2 ml culture in LB (Duchefa Biochemie) were collected by centrifugation at maximal speed for 1 min. Cells were resuspended in 250μl of P1 buffer (10mM EDTA, 100 μg/ml RNase A, 50mM Tris-Cl pH 8.0) and lysed by addition of 250μl of P2 buffer (200mM NaOH, 1% (w/v) SDS). After 5 min of incubation protein were precipitated by addition of 350μl of N3 buffer (3M potassium acetate pH 5.0) and a centrifugation of 1 min at maximal speed. Supernatant was then filtrated through a QIAfilter column and washed twice with PE buffer; DNA was then eluted with 35 – 50μl of distilled water and quantified.

For DNA maxi-preparation QIAfilter Plasmid Maxi Kit (Qiagen) was used, following the manufacturers' instructions. Starting from a pre cultured 400ml LB culture, cells were collected by centrifugation at 6000*g* for 15 min, resuspended in 10 ml of P1 buffer and lysed by adding 10 ml of P2 buffer. After an incubation of 5 min, proteins were precipitated by adding 10 ml of P3 buffer (3M potassium acetate pH 5.0) and filtered through a QIAfilter Maxi Cartridge. Cell lysate was then loaded into a QIAfilter column, previously equilibrated with QBT buffer (750mM NaCl, 50mM MOPS pH7.0, 15% (v/v) isopropanol, 0.15% (v/v) Triton X-100). After washing twice the resin with 30 ml of QC buffer (1.0M NaCl, 50mM MOPS, pH 7.0, 15% (v/v) isopropanol), DNA was eluted with 15 ml of QF elution buffer (1.25M NaCl, 50mM Tris-Cl, pH 8.5, 15% (v/v) isopropanol) and precipitated by addition of 10 ml isopropanol and a centrifugation at 15.000 g for 30 min at 4°C. The resulting pellet was washed with 70% (v/v) ethanol, precipitated and dried. DNA was finally resuspended in distilled water and quantified.

### 2.3.6 Colony Screen by Cracking Procedure

20μl of *E. coli* cells cultured overnight in LB (Duchefa Biochemie) were added to 20μl of 2X cracking solution (100mM NaOH, 10mM EDTA, 1% (w/v) SDS, 10% (v/v) glycerol, 0.5% (w/v) bromocresol green) and mixed vigorously. After 5 min at room temperature the samples were

mixed with 6X DNA loading buffer (40% (v/v) glycerol, 0.25% (w/v) bromophenol blue) and loaded on gels. DNA was visualised on an ultraviolet transilluminator (UviDoc).

## 2.4 Plasmids and Constructs

### 2.4.1 Plasmids

#### i. pUC57

pUC57 is a cloning vector commonly used in *E. coli*. The vector length is 2.710 bp and is isolated from *E. coli* strain. ER-BMP2/N-BMP2 fusion construct was synthesized and cloned in this vector from GeneCust (*GeneCust* Europe, Laboratoire de Biotechnologie du Luxembourg S.A.).

#### ii. pDHA

The plasmid pDHA was derived from the Cauliflower Mosaic Virus (CaMV) 35S gene expression cassette pDH51 [275] as follows: An EcoRV – BamHI fragment from pDH51 incorporating the 3' proximal nucleotides of the CaMV 35S promoter was replaced with a similar fragment from pAMV9 [276] in which the 3' end of the CaMV 35S promoter was fused to the 5' untranslated region of the coat protein mRNA from alfalfa mosaic virus. This sequence has been reported to enhance the translatability of messenger RNA [277]. The expression vector pDHA [271,278] was used for cloning experiment and for transient expression in tobacco protoplasts.

#### iii. pGreen

For replication in *A. tumefaciens*, the *pSa* replication locus has been used in the pGreen vector. Sequencing of this fragment showed that the pSa replication locus consists of an *Ori* and *RepA* gene. Some experiments established that the *pSa-RepA* gene can act in trans on the *Ori* and is not needed on the pGreen vector, offering a substantial saving on the size of the binary vector. Therefore, *pSa-RepA* was subcloned replacing the T-DNA region. The gene coding for bacterial resistance to kanamycin, was replaced with the tetracycline resistance gene (*Tet*) to create pSoup (Figure 8) [274].

The pGreen vector (Figure 8) is based on the general-purpose pBluescript cloning vector [279] and therefore contains a *Ori* for replication in *E. coli*. The plasmid's ampicillin resistance gene was

replaced with the gene encoding for kanamycin resistance, the *pSa-Ori* was inserted, the f1 *Ori* and Lac Z' region deleted and a Bgl II site was left for the introduction of a T-DNA cassette.

A T-DNA cassette containing the pBluescript SKII LacZ' gene and multiple cloning sites, flanked by synthetic LB and RB sequences and with an additional T-DNA transfer enhancer motif immediately adjacent and external to the RB, was inserted to create pGreen.

The expression vector pGreen [272,273] was used for cloning experiment and for stable expression of ER-BMP2 and N-BMP2 in tobacco plants.

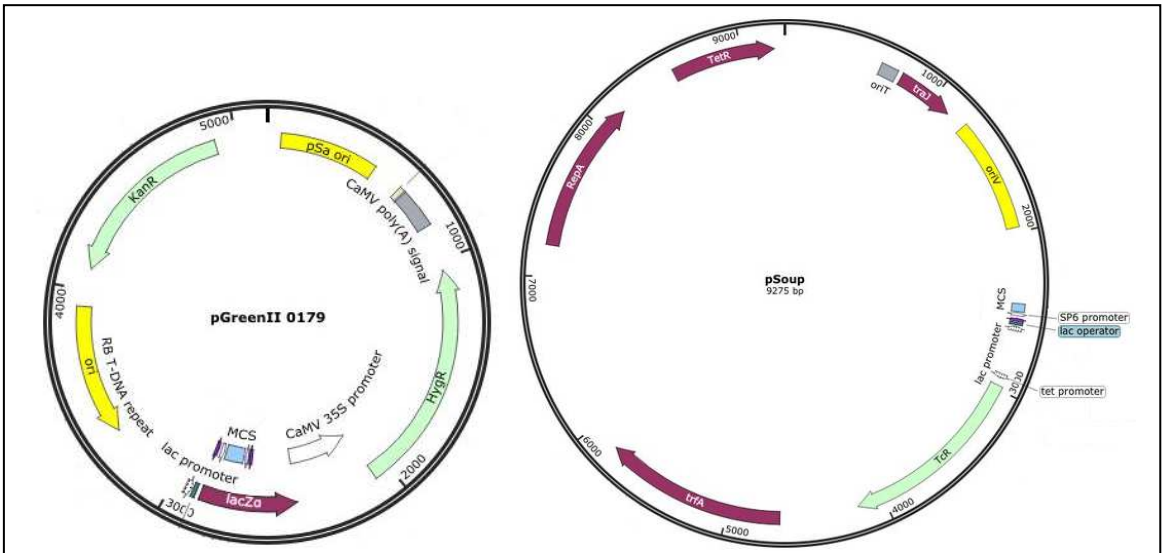
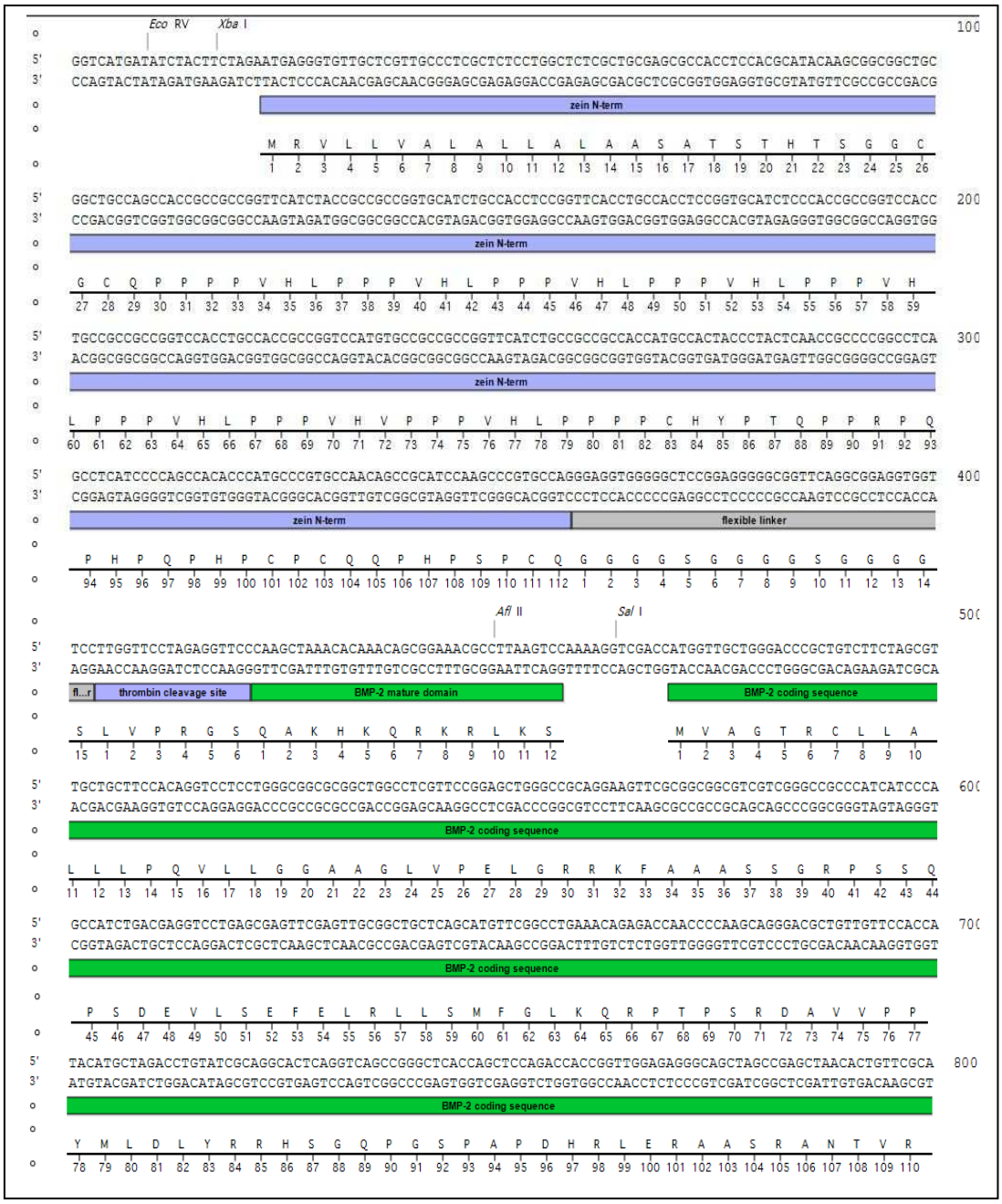



**Figure 8: pSoup and pGreen binary vector system structure.**

## 2.4.2 ER-BMP2 and N-BMP2 constructs

A single chimeric construct was ordered to GeneCust (*GeneCust* Europe, Laboratoire de Biotechnologie du. Luxembourg S.A.) in order to obtain the two ER-BMP2 and N-BMP2 final constructs with specific restriction enzymes (Figure 9). This was possible thanks to an Afl II enzyme site, that is present into the BMP-2 mature domain, that permitted a subcloning procedure. The FLAG epitope was added to both constructs, to allow easy immunodetection. The synthetic chimeric construct (1707 bp long) was optimized for tobacco expression and cloned from GeneCust (*GeneCust* Europe, Laboratoire de Biotechnologie du Luxembourg S.A.) into a pUC57 vector with EcoR V restriction enzyme.

The two constructs were finally inserted into pDHA vector for cloning experiment and for transient expression in tobacco protoplasts and in pGreen vector for constitutive expression under the 35S promoter in plants.



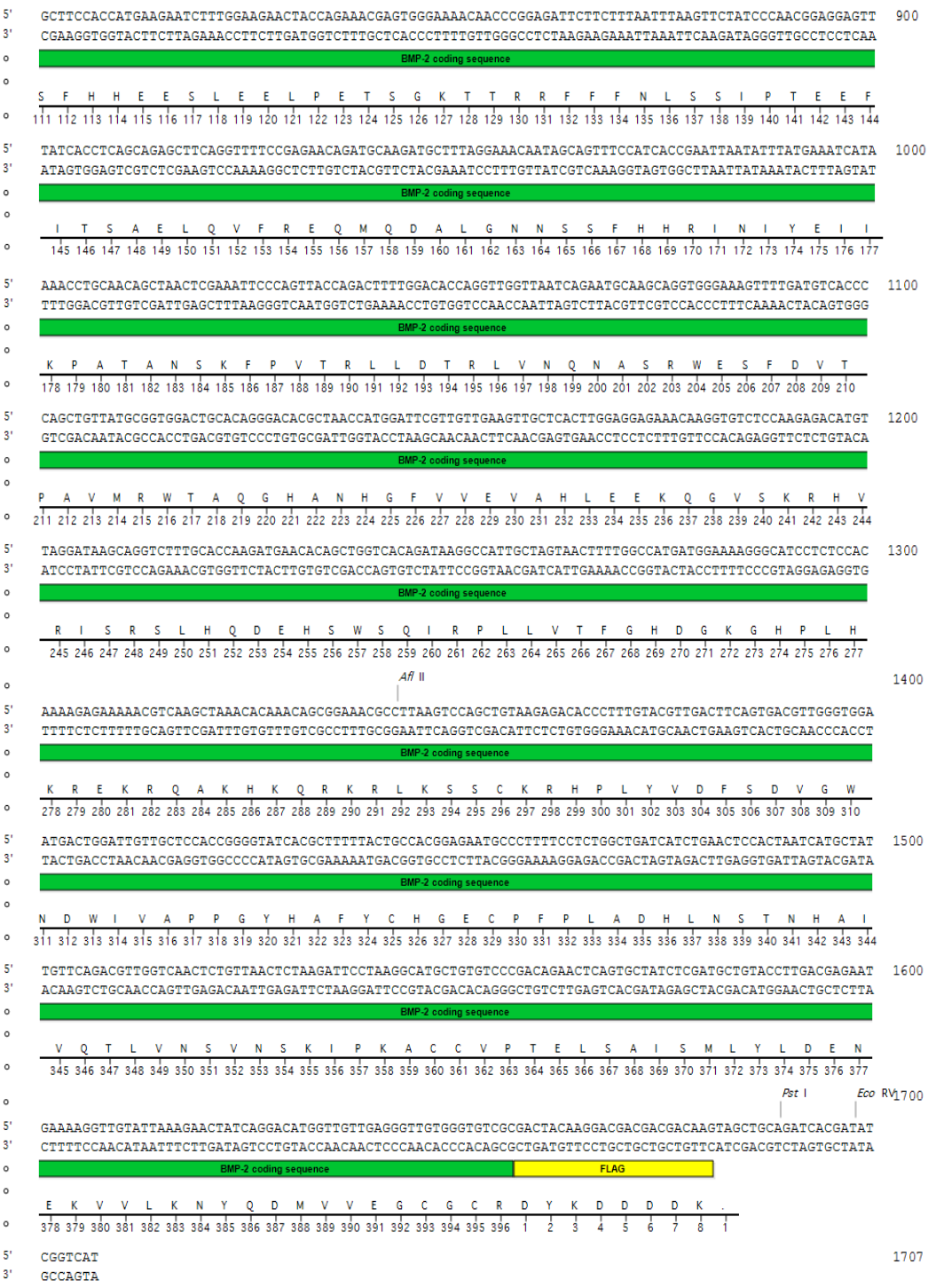



**Figure 9: nucleotide and aminoacidic sequences of the chimeric construct** inserted in the pUC57 vector by GeneCust and used to obtain the final constructs used for transformation. γ-zein N-terminal is represented in light blue, BMP-2 coding sequence is represented in green and FLAG epitope is yellow. The restriction enzymes used for cloning experiments are written above respective sites.

### i. ER-BMP2

To obtain high accumulation of rhBMP-2 in transgenic tobacco plants, the C-terminal mature peptide of hBMP-2 cDNA was cloned downstream the N-terminal domain of the maize seed storage protein γ-zein (Figure 10), which has been previously demonstrated to promote extensive polymerization of heterologous proteins within the ER lumen, leading to assembly into protein bodies [258]. The pUC57 vector containing the chimeric construct, after expansion in *E. coli*, was digested with *Xba I* and *Pst I* and subcloned in pDHA vector. To reach the correct sequence this vector was digested with *Afl II* in order to excise the signal peptide and the pro-peptide of BMP-2.

The expected molecular masses of the encoded ER-BMP2 polypeptide was around 25 kDa.

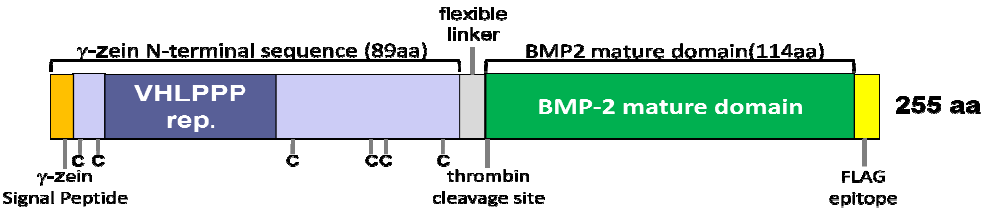



**Figure 10: schematic representation of ER-BMP2 construct final structure.** γ-zein N-terminal is represented in light blue, BMP-2 coding sequence is represented in green and FLAG epitope is yellow. Between the γ-zein polypeptide and BMP-2 mature domain a flexible linker (represented with gray colour) and a thrombin cleavage site were added in order to recover only the BMP-2 portion.

### ii. N-BMP2

Full-length hBMP-2 (N-BMP2, Figure 11) was also tested. The pUC57 vector containing the chimeric construct, after expansion in *E. coli*, was digested with *Sal I* and *Pst I* and subcloned in pDHA vector; in this way the complete CDS of hBMP-2, from its own signal peptide to the stop codon, was isolated.

The expected molecular mass of the encoded N-BMP2 polypeptide was around 43 kDa (N-BMP2).

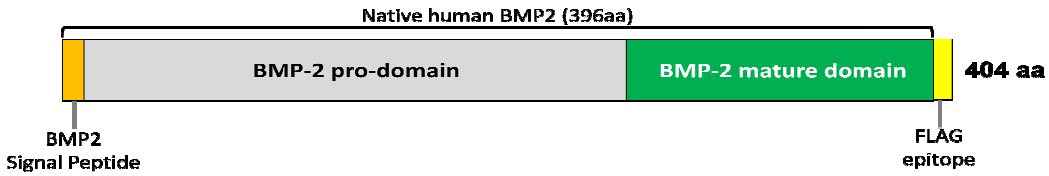



**Figure 11: schematic representation of N-BMP2 construct.** It is composed by the 396aa of human BMP-2 (including Signal Peptide, pro-domain and mature domain) with a C-terminal FLAG epitope (yellow)

# 3 Results

## 3.1 Transient Expression of Recombinant Human BMP-2 in Tobacco Protoplasts

The first part of this project consisted of the designing and cloning of the two recombinant BMP-2 constructs. Tobacco protoplast were then transiently transformed with the two constructs inserted in the pDHA vector and after 24h total protein were extracted and analyzed in order to detect the presence and the main characteristics of the two recombinant proteins.

### 3.1.1 ER-BMP2 and N-BMP2 are Soluble, Not Secreted Proteins

Due to the presence of γ-zein N-terminal – which is able to induce the formation of insoluble PB even when fused to other proteins of the secretory pathway [258] – in the chimeric protein, ER-BMP2 was expected to be insoluble in absence of reducing agent.

For this reason, to determine the presence and the molecular weight of ER-BMP2 and N-BMP2 produced by transient transformation, after protoplasts homogenisation under both reducing and non-reducing conditions, samples were centrifuged at 4000*g* for 10 min and both the supernatants (S) – were the soluble proteins are collected – and pellets (P) – were the insoluble polymers are collected – were analyzed by SDS-PAGE and protein blot using anti-FLAG antiserum (Figure 12a). Since N-BMP2 is a secreted protein, the medium where protoplast were incubated were also analyzed, according to the same procedure (Figure 12b). Protoplasts transformed with the empty vector were used as control (CTRL).

In absence of reducing agent ER-BMP2 migrates mainly as a soluble dimer of about 55 kDa (Figure 12a, band D), that is disassembled in presence of 2-ME, when only the monomer was observed (Figure 12a, bands M).

N-BMP2 was poorly solubilized in the absence of 2-ME and migrated at 60 kDa (Figure 12a, empty circle), slower than expected; conversely, in the presence of 2-ME, a band of the expected monomeric mobility, about 40 kDa, was detected (Figure 12a, band M).

Both proteins appear to be soluble even in the absence of 2-ME, as demonstrated by the low intensity of the bands detected in the pellet samples (Figure 12a, P lanes), which are probably due a residual of supernatants after centrifugation. ER-BMP2 was detected in higher amount than N-BMP2 inside protoplasts (Figure 12a, compare ER-BMP2 with N-BMP2 lanes).

The band at 25 kDa (dark arrowhead) detected in every lane is present also in the control samples and it is probably an unrelated polypeptide that reacts with the secondary antibody.

Neither ER-BMP2 nor N-BMP2 were detected in the incubation medium, indicating that both proteins was not secreted (Figure 12b).

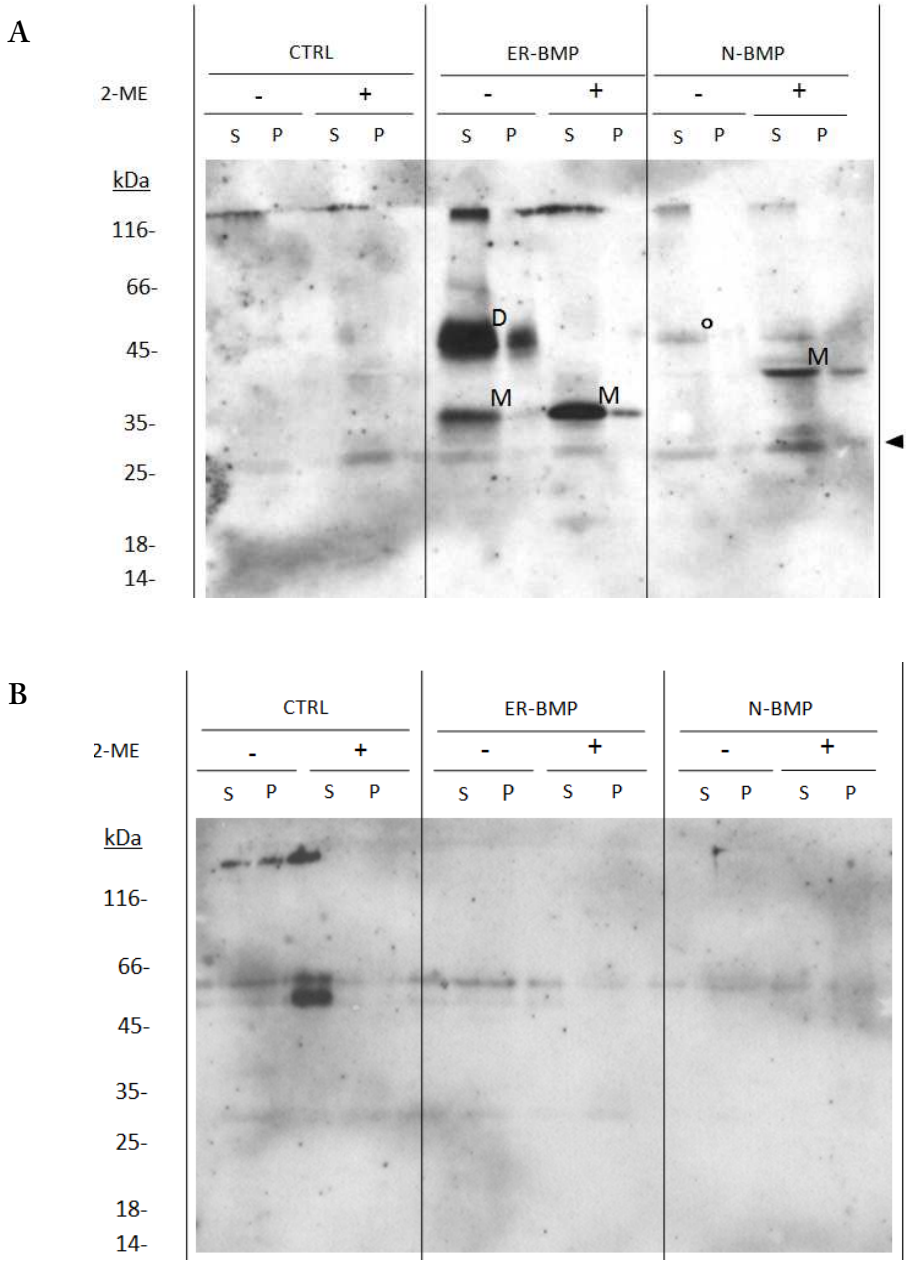



**Figure 12: ER-BMP2 and N-BMP2 are soluble proteins and are not secreted in the protoplast medium.** Total protein from 20.000 transformed protoplasts (**A**) and their medium (**B**) were extracted in the presence (+) or absence (-) of the reducing agent 2-mercaptoethanol (2ME) in the homogenization buffer. After centrifugation, the supernatants (S), which contains soluble proteins, and pellets (P), which contains insoluble polymers, were analyzed by SDS-PAGE and protein blot. The migration of molecular weight standards (kDa) is indicated.

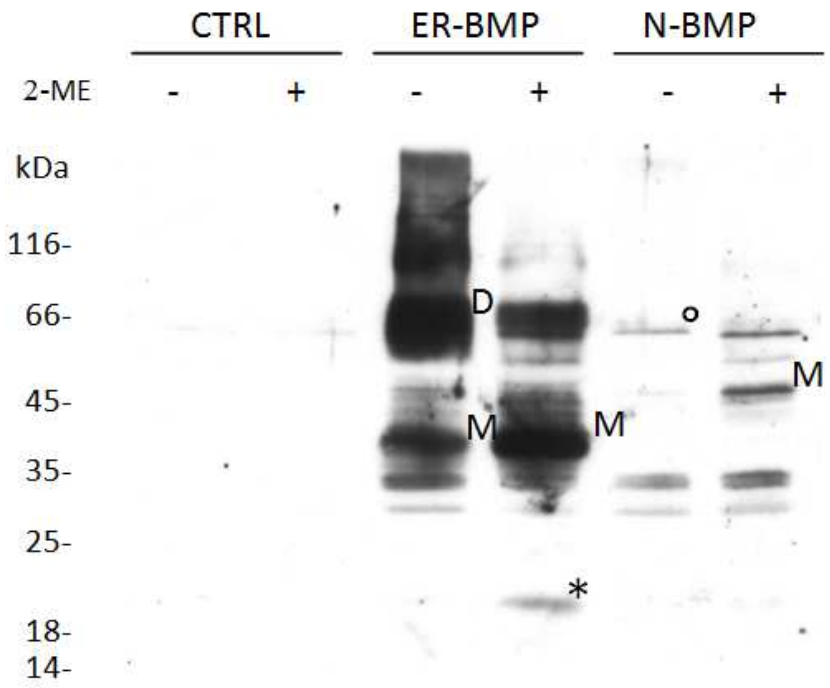



**Figure 13: ER-BMP2 and N-BMP2 are soluble proteins.** Protein blot performed on an independent experiment of transient transformation. Total protein of 20.000 protoplasts were extracted in the presence (+) or absence (-) of the reducing agent 2-mercaptoethanol (2ME) in the homogenization buffer; after centrifugation to separate insoluble polymers, the supernatants were analyzed by SDS-PAGE and protein blot. The migration of molecular weight standards (kDa) is indicated.

In order to confirm the obtained results an independent transformation experiment (Figure 13) was performed.

ER-BMP2 dimers (D) migrated mainly at 55 kDa, while their monomers (M) migrate at approximately 35 kDa. When protein extraction was performed in reducing condition a band of about 20 kDa (Figure 13, asterisk), that may correspond to a degradation product, was detected.

N-BMP2 is confirmed to have a lower level of expression than ER-BMP2 and a monomeric form of about 45 kDa detectable in the presence of reducing agent (M).

### 3.1.2 ER-BMP2 is N-Glycosylated and Retained in the ER

ER-BMP2 has a predicted N-glycosylation site in the BMP-2 mature domain (see $N_{338}ST$ sequence in Figure 9). However, not all glycosylation consensus sites are really glycosylated, because they can be masked by folding [280].

Endoglycosidase H enzyme is able to remove N-linked mannose-rich oligosaccharides from a polypeptidic chain unless these are modified by Golgi enzymes and it is therefore used for *in vitro* experiments in order to demonstrate traffic through the Golgi complex [281]. Thus, to verify if ER-BMP2 was correctly glycosylated and to assess whether the presence of the γ-zein portion induced ER retention and accumulation of the chimeric protein, an Endoglycosidase H assay was performed.

Total proteins were extracted from transfected protoplasts in presence of reducing agent and incubated (Figure 14 – lanes +) or not (Figure 14 – lanes -) with EndoH enzyme. Samples were analyzed with SDS-PAGE and immunoblot.

The removal of the N-glycan caused a shift in ER-BMP2 mobility (Figure 14, compare electrophoresis mobility in lanes - and + (black arrowhead)) indicating that most probably the fusion protein does not traffic through the Golgi complex but instead remains in the ER, as expected.

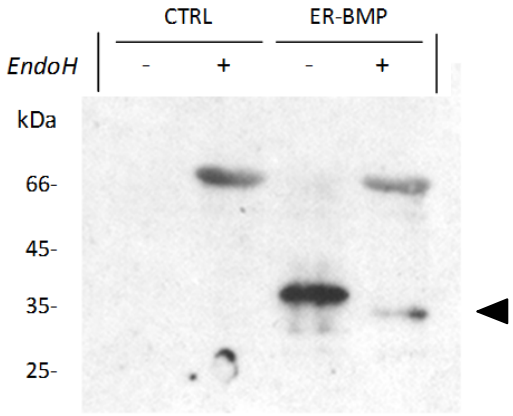



**Figure 14: Endoglycosidase H assay on ER-BMP2 transformed protoplasts.** Proteins extracted from 20.000 wild-type (CTRL) and ER-BMP2 transformed protoplasts were treated (+) or not (-) with Endoglycosidase H (EndoH) and analyzed by SDS-PAGE and protein blot. The migration of molecular weight standards (kDa) is indicated.

### 3.1.3 ER-BMP2 Does Not Form Large Polymers in the ER

It is demonstrated that γ-zein N-terminal domain is able to induce the formation of large insoluble polymers (PB) which accumulate within the ER, even when fused to other protein, as demonstrated for zeolin [258,282]. When these proteins were subjected to a sucrose velocity gradient centrifugation they were completely recovered in the pellets fraction.

Total proteins from 400.000 protoplasts transformed with ER-BMP2 construct (Figure 15a) were extracted in the absence of reducing agent and then subjected to sedimentation velocity centrifugation on a continuous 5-25% (w/v) sucrose gradient. Aliquots of gradient fractions and of material precipitated at the bottom of tube (P) were analyzed by SDS-PAGE and protein blot using anti-FLAG antiserum.

ER-BMP2, unlike γ-zein, forms polymers in the range of 50-150 kDa (Figure 14a) that don't reach the pellet fraction.

### 3.1.4 N-BMP2 is a Monomer

The same experiment was performed on proteins extracted from protoplasts transformed with N-BMP2 (Figure 15b) construct.

N-BMP2 migrated as a monomeric peptide of about 60 kDa (Figure 15b - see also empty circle in Figure 12a).

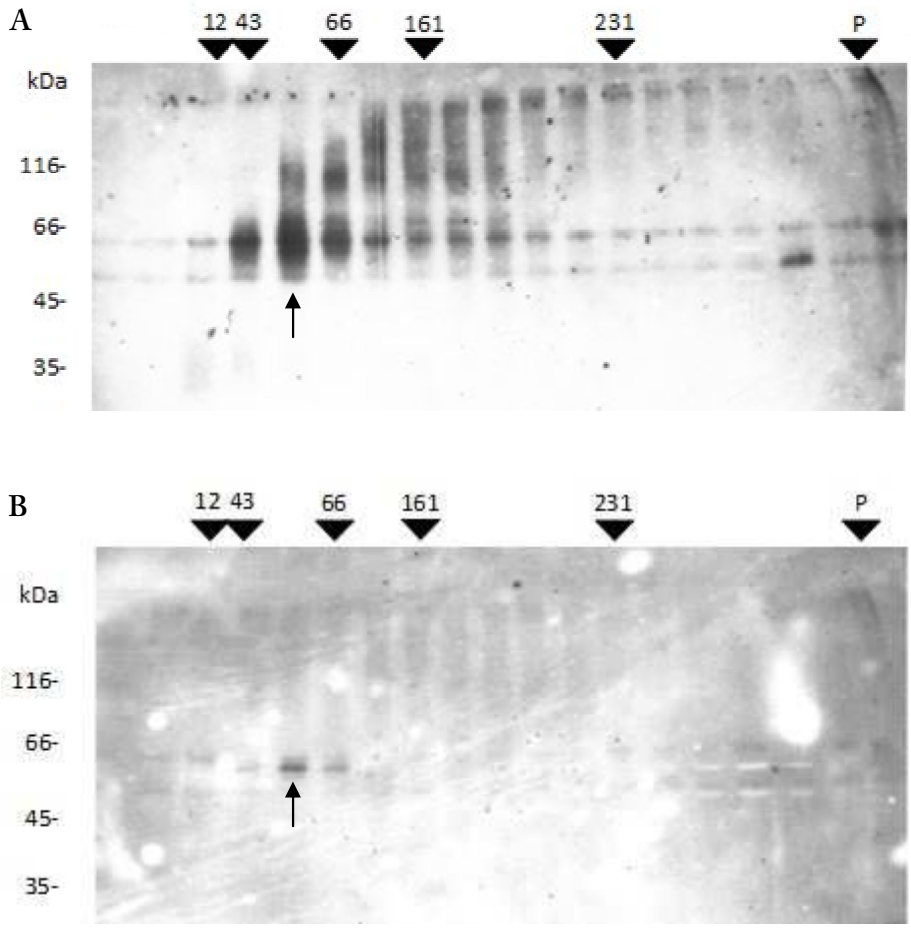



**Figure 15: Velocity gradient on protoplasts extracts.** 400.000 protoplasts transformed with ER-BMP2 (A) and N-BMP2 (B) constructs were homogenised in absence of reducing agent and then subjected to sedimentation velocity centrifugation on a continuous 5-25% (w/v) sucrose gradient. Aliquots of gradient fractions and of material precipitates at the bottom of tube (P) were analyzed by SDS-PAGE and protein blot using anti-FLAG antiserum. Top of gradient is on the left, numbers on top indicate the molecular mass of markers in kDa. The migration of molecular weight standards (kDa) is indicated on the left.

## 3.2 Transgenic Tobacco Plants Expressing rhBMP-2

On the basis of the promising results obtained with the experiments of transient expression in protoplasts, transgenic Tobacco plants expressing ER-BMP2 and N-BMP2 were generated by *Agrobacterium*-mediated transformation of plant tissues (leaf-discs). As described in chapter 2.1.3, plants were transformed with *Agrobacterium tumefaciens* strain GV3101 carrying the plasmid of interest.

Leaf disk were incubated in a suspension of *A. tumefaciens* and then placed on MS medium supplemented with the correct combination of hormones – to induce shoot and roots generation – and antibiotics – to select transformed cells and remove remaining bacteria. After 4-6 weeks on selective medium, plants were tested for protein expression and transferred in soil. Plants with high expression levels were propagated and new generations were screened and selected for high and stable expression of recombinant proteins.

### 3.2.1 Transgenic Tobacco Plants Production and Transformants Selection

ER-BMP2 and N-BMP2 were expressed in transgenic tobacco under the control of the constitutive CaMV 35S promoter via *Agrobacterium tumefaciens*-mediated transformation. Expression levels and patterns of transgene inheritance generally show wide variation between independently transformed plants carrying the same construct and many factors can contribute to these variations; these include integration site, transgene copy number and transgenic locus configuration [283]. Thus, several putative transgenic lines, selected on the basis of their hygromycin resistance, were tested for ER-BMP2 and N-BMP2 accumulation by protein blot analysis of equal amounts of leaf extracts.

ER-BMPs accumulation was not homogeneous in the analyzed plants, however, considering an equal volume of leaf analyzed, ER-BMP2 showed a clear tendency to accumulate at higher levels than N-BMP2 (Compare Figure 16a and 16b).

Five lines with the highest protein accumulation levels were selected for each construct from a population of 13 N-BMP2 (Figure 16a) and 20 ER-BMP2 (Figure 16b) primary transformants and used in all subsequent experiments to further characterize protein expression in tobacco leaves. When the shoots were about 4 weeks old they were transferred and grown in soil. The seeds produced from every transgenic plant were collected and stored; while seeds of 3 selected ER-

BMP2 transformants (ER7, ER23 and ER27 – Figure 16b, black arrows), selected according to their high accumulation levels, were propagated in order to generate F1 transgenic plants.

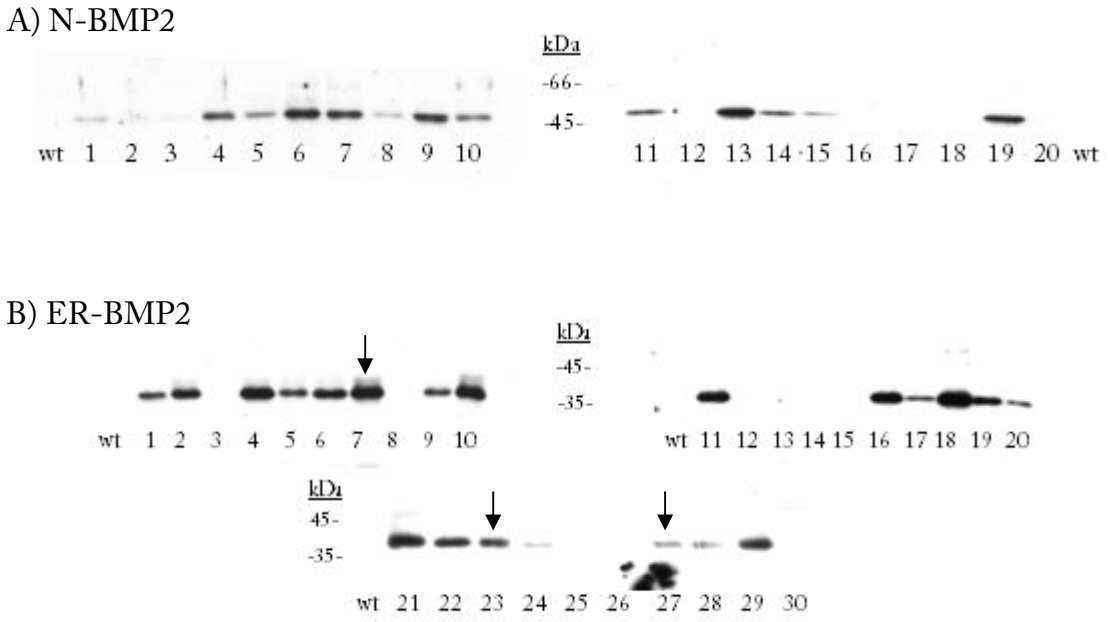



**Figure 16: Screening of tobacco transgenic F0 plants.** Young leaves of transformed shoots were homogenised in the presence of the reducing agent in SDS-PAGE loading buffer at 95°C. After centrifugation the supernatants were diluted 1:10 for N-BMP2 transformed shoots (A) or 1:20 for ER-BMP2 transformed shoots (B) and analyzed by SDS-PAGE and protein blot with anti-FLAG antiserum.

### 3.2.2 Screening of F1 Tobacco Plants

The lines with the highest ER-BMP2 accumulation levels were selected from the population of primary transformants, and their seeds were collected and propagated in soil for further investigations. After *Agrobacterium tumefaciens* transformation we obtained heterozygous F0 transformants with unknown copy-number of the gene. Since tobacco is a self-pollinating plant, a Mendelian segregation was expected and therefore not all the seeds generated from F0 expressing plants would be themselves transgenic.

ER-BMP2 demonstrated very high level of accumulation, then, in order to select F1 ER-BMP2 expressing plants, a dot-blot screening protocol was adopted. Young leaves were homogenised in SDS-PAGE loading buffer at 95°C and, after a centrifugation, 5µl of the supernatants were spotted on nitrocellulose membrane for direct immunodecoration (Figure 17). Seven positive plants were found and grown in soil until new seeds were produced. Seeds of F2 generation were then collected and stored at 4°C.

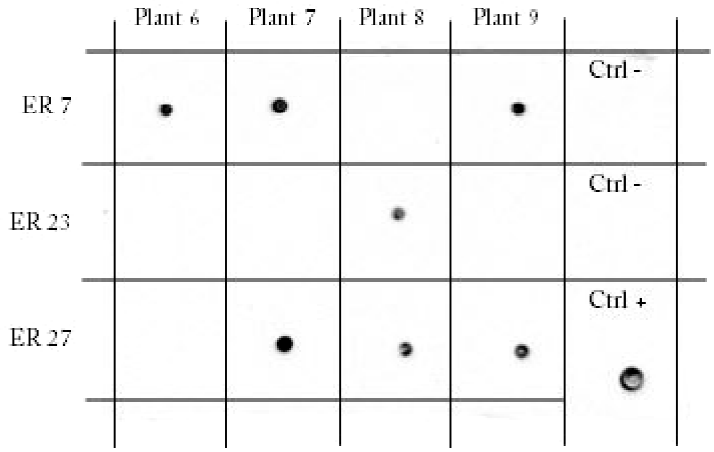



**Figure 17: Dot-Blot screening of F1 plants.** Young leaves of transgenic F1 plants were homogenised in presence of the reducing agent in SDS-PAGE loading buffer at 95°C. Volume of buffer/weight of the leaves Ratio was 5/1. After centrifugation, 5 µl of the supernatants were spotted on membrane for direct immunodetection with anti-FLAG antiserum. Leaves collected from wild type plant were used as negative control (Ctrl -) and leaf from ER23 F0 plant was used as positive control (Ctrl +).

### 3.2.3 ER-BMP2 Accumulates to Higher Level than N-BMP2 in Transgenic Tobacco

In order to determine the expression levels and the main characteristics of ER-BMP2 and N-BMP2 in transgenic plants, total protein extracts were analyzed by protein blot.

Total proteins were extracted in ice-cold homogenisation buffer supplemented with protease inhibitor in presence or absence of 2-ME. The homogenates were then centrifuged and supernatants were collected. Total proteins in supernatants were quantified using the Bradford protein assays and diluted for SDS-PAGE and immunoblotting analysis.

N-BMP2 plants in previous experiments demonstrated a lower expression level than ER-BMP2 plants, hence a 10 time higher amount of total proteins were used (Figure 18).

In ER-BMP2 expressing plants, in non reducing conditions one main band was detected, corresponding to the dimeric form of the protein (Figure 18, ER-BMP2 section, lanes -), while in presence of reducing agent, monomers (M) of approximately 35 kDa were detected (Figure 18 – ER-BMP2 section, lanes +). Even in stable transformation when 2-ME was added to the homogenisation buffer a band of about 20 kDa (Figure 18 – ER-BMP2 section, asterisks), that may correspond to a degradation product, was detected (compared to Figure 13, asterisk). Almost all ER-BMP2 tested transgenic lines demonstrated a good level of protein expression and accumulation, even if it was highly variable (data not shown).

N-BMP2 plants showed a clear tendency to accumulate lower level of protein than ER-BMP2 plants, thus, a 10 fold amount of total protein extracts were analyzed; for this reason an increased detection of non-specific bands could be expected. Even if in absence of reducing agent a putative monomeric band of about 60 kDa (Figure 18 - N-BMP2 section, white circle), that change its mobility in reducing conditions (Figure 18 - N-BMP2 section, M), was detected. The intricate pattern of bands of non-expected molecular weight, that could be due not only to non-specific detection but also to a degradation of the protein, makes the interpretation of the results uncertain.

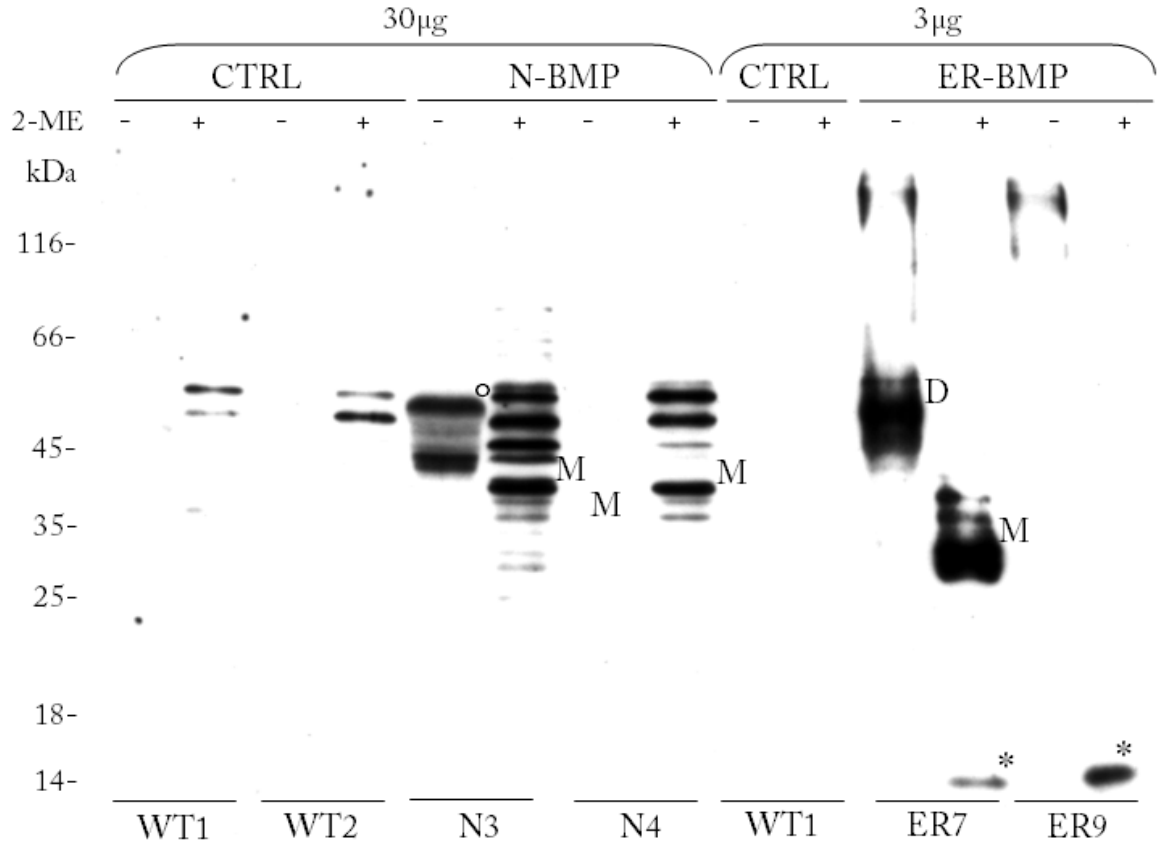



**Figure 18: ER-BMP2 is Soluble in Transgenic Tobacco Plants.** Total protein from two N-BMP2 transformed plants, two ER-BMP2 transformed plants and two wild type plants (CTRL) were extracted from young leaves in the presence (+) or absence (-) of the reducing agent 2-mercaptoethanol (2-ME) in the homogenization buffer. After centrifugation to separate insoluble polymers, the supernatants were analyzed by SDS-PAGE and protein blot with anti-FLAG antiserum: 30µg of total protein of N-BMP2 samples and 3µg of ER-BMP2 samples were loaded on gel with the same quantity of control samples. The migration of molecular weight standards (kDa) is indicated.

### 3.2.1 ER-BMP2 Is N-Glycosylated and Retained in the ER

As described before, ER-BMP2 has a predicted site for N-glycosylation in the BMP-2 mature domain that in transient expression experiments demonstrated to be sensible to Endoglycosidase H enzyme. Total proteins extracts obtained from transgenic plants in presence of 2-ME were incubated (+) or not (-) with EndoH to demonstrate ER-BMP2 sensitivity to the enzyme and thus its retention in the ER.

The removal of the single ER-BMP2 N-glycan by Endoglycosidase H (Figure 22, compare electrophoresis mobility in lanes - and +) indicated that most probably the fusion protein did not traffic through the Golgi complex, but remained in the ER, as expected.

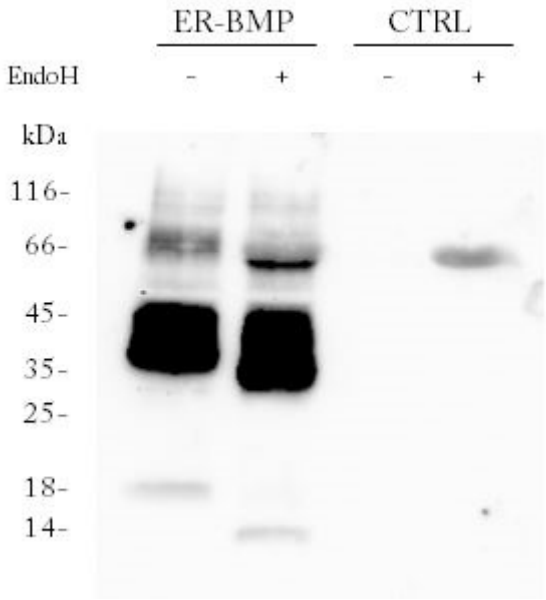



**Figure 22: ER-BMP2 is retained in the ER of transgenic leaves.** Total proteins from leaves collected from a ER-BMP2 transgenic plant and a wild-type plant (CTRL) were extracted in the presence of reducing agent 2-ME and 8 μg of samples were treated (+) or not (-) with Endoglycosidase H (EndoH). Samples were then analyzed by SDS-PAGE and protein blot with anti-FLAG antiserum. The migration of molecular weight standards (kDa) is indicated.

### 3.2.1 ER-BMP2 Is a Soluble Protein That Forms Large Oligomers

Experiments conducted on transiently transformed tobacco protoplasts demonstrated that ER-BMP2, unlike γ-zein, is a mainly soluble protein present mostly in a dimeric form, even considering the formation of higher polymers.

To investigate its assembly grade in stable transformation, total protein from young leaves harvested from one transgenic ER-BMP2 plant were extracted in absence of reducing agent and then subjected to sedimentation velocity centrifugation on a continuous 5-25% (w/v) sucrose gradient. Aliquots of gradient fractions and of precipitates at the bottom of tube were analyzed by SDS-PAGE and protein blot.

ER-BMP2 expressed in transgenic plants confirmed to be a soluble protein that presents a dimeric form (Figure 21, peak around 66 kDa); nevertheless, ER-BMP2 was able to form high polymers that exceeded the 231 kDa marker. As hypothesized in section 3.1.3, the different solubility of ER-BMP2 and γ-zein is possibly due to a kind of interference between the 6 cysteines of γ-zein and the 7 ones of BMP-2 mature domain that formed unexpected inter- and intra-chain bonds; despite this, the formation of large oligomers is not precluded.

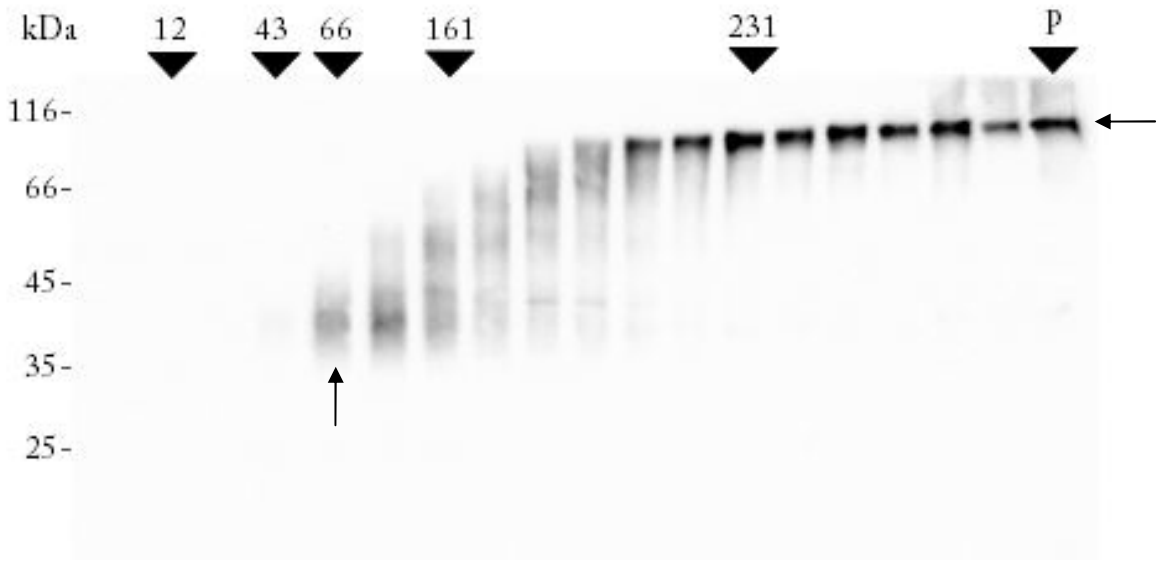



**Figure 21: ER-BMP2 forms large oligomers.** Leaves collected from one ER-BMP2 plant (ER23) were homogenised in absence of reducing agent and 800μg of total protein extract were analyzed with a sedimentation velocity centrifugation on a continuous 5-25% (w/v) sucrose gradient. Aliquots of gradient fractions and of material precipitated at the bottom of tube (P) were analyzed by SDS-PAGE and protein blot using anti-FLAG antiserum. Top of gradient is on the left, numbers on top of the gel indicate the molecular mass of markers in kDa. The migration of molecular weight standards (kDa) is indicated at left.

### 3.2.2 ER-BMP2 Accumulates in Seeds

Zeins are a class of prolamin proteins account for more than 70% of maize endosperm protein [284]. Because ER-BMP2 is a fusion protein that includes the N-terminal of γ-zein prolamin, the presence of recombinant protein in seeds was investigated. Tobacco seeds produced by Wild-Type tobacco, N-BMP2 and ER-BMP2 F0 transformed plants were homogenised in SDS-PAGE Laemmli buffer at 95°C and centrifuged in order to remove seeds debris. Supernatants were then collected and analyzed with SDS-PAGE and immunoblot. Specific signal were detected in 3 of 4 ER-BMP2 analyzed samples (Figure 19 – ER-BMP2, ER5, ER7 and ER27, dark arrowhead). ER4 did not show a detectable signal, probably due to a too low level of expression.

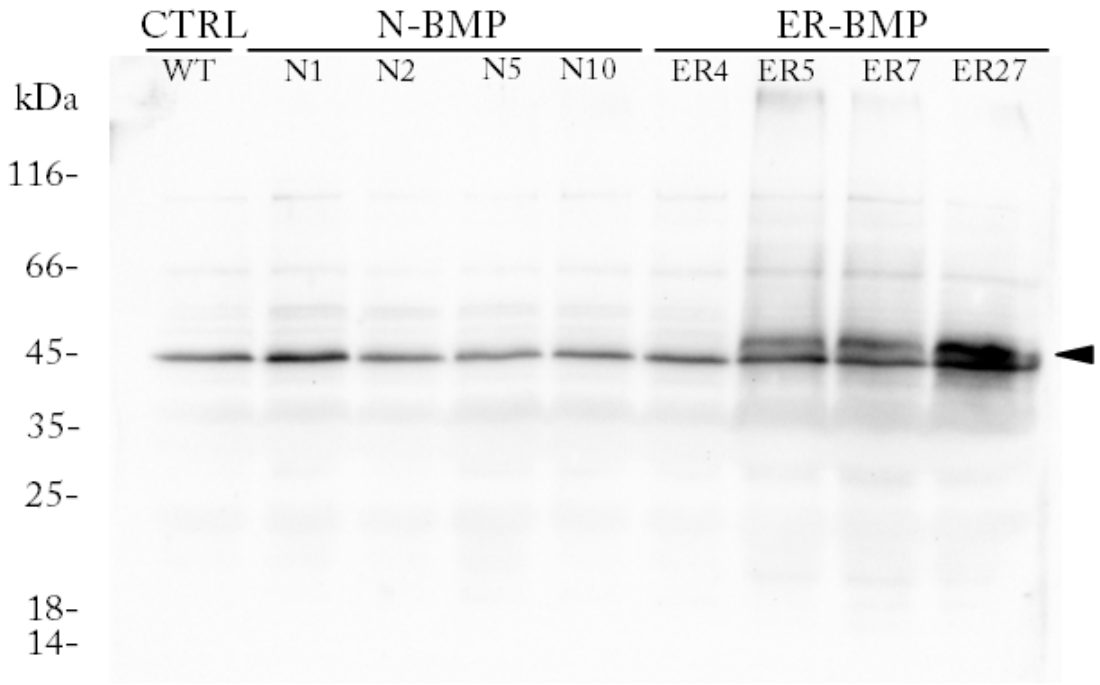
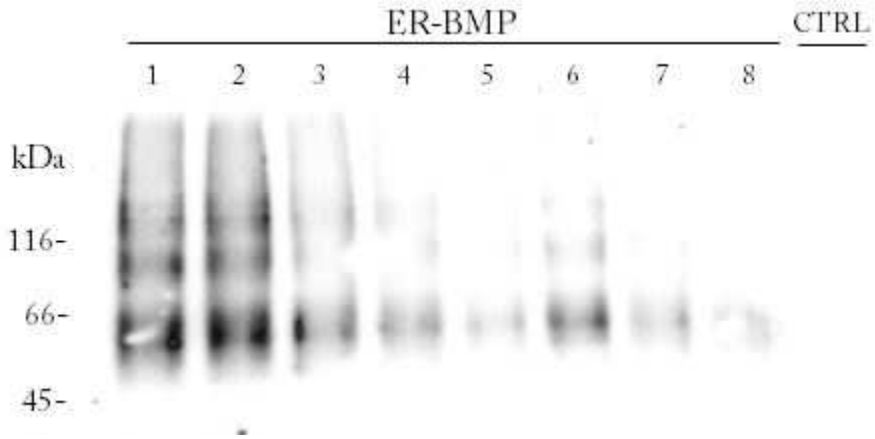



**Figure 19: ER-BMP2 accumulates in tobacco seeds** This protein blot was performed on 0.05g of seeds collected from 1 wild type (CTRL), 4 N-BMP2 and 4 ER-BMP2 plants. seeds were homogenised in SDS-PAGE loading buffer at 95°C. Ratio volume of buffer/weight of the seeds was 10/1. After centrifugation 5µl of the supernatants were collected – avoiding the oil phase – and analyzed with SDS-PAGE and protein blot with anti-FLAG antiserum. The migration of molecular weight standards (kDa) is indicated.

### 3.2.3 Young Leaves Produces Higher Level of ER-BMP2 Compared to Older Ones.

In many cases the accumulation of recombinant protein demonstrated to be sensitive to the age and the dimension of the leaf, with younger leaves showing better performance than the older ones [285,286]. Thus a protein blot experiment was conducted in order to investigate if the levels of expression of ER-BMP2 could change with respect to the age and the dimension of the leaf. Total proteins were extracted from 8 leaves of an ER-BMP2 plant (ER23), chosen from those who demonstrated higher levels of expression during previous experiments, and a SDS-PAGE experiment was performed using anti-FLAG antibody. The dimensions of each harvested leaf are shown in the table below. The proteins were extracted in ice-cold homogenisation in absence of 2-ME and, after protein quantification, 4µg of each samples were analyzed.

The gel showed an overall downward trend. Youngest leaves (Figure 20 – lanes 1and 2) showed a significantly higher level of expression compared to the older ones (Figure 20 – compare lane 7and 8 with lane 1 and 2), that were larger, thinner and that presented barely detectable bands.




| Lane n° | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Size (cm) | 3.0x1.5 | 3.2x1.8 | 4.2x2.2 | 4.8x3.5 | 7.0x5.5 | 9.3x6.0 | 11.7x8.2 | 12.4x9.3 |

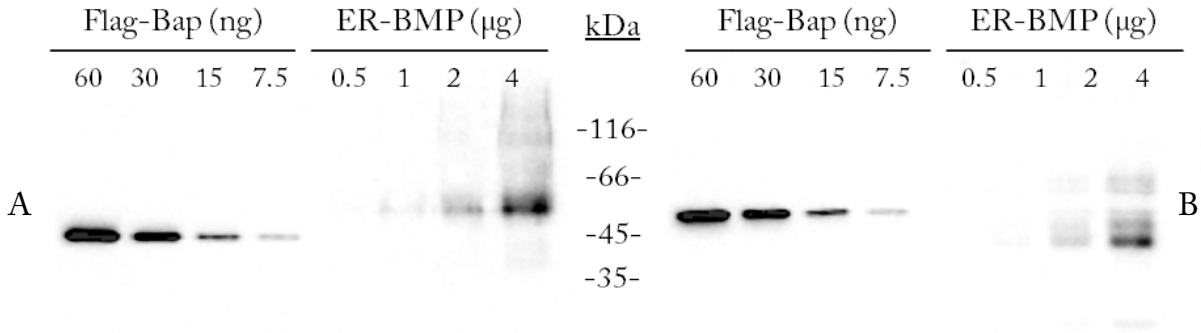
**Figure 20: Young leaves produces higher level of protein compared to older ones.** The protein blot was performed on leaves collected from one F0 ER-BMP2 plant. Total protein of 8 leaves of different age and size (see the table) were extracted in absence of 2-ME and 4µg of total protein of each sample were loaded on acrylamide gel together with wild-type leaf extract (CTRL) and analyzed by protein blot. Numbers on top of the lane indicate the position of the leaf: n°1 was the smallest leaf closest to the top, n°8 was instead the biggest one in the lower part of the plant. The migration of molecular weight standards (kDa) is indicated.

### 3.2.4 Quantification of ER-BMP2

Absolute quantification of produced ER-BMP2 was performed by comparison of progressive dilutions of leaf extracts with known amounts of the commercial standard protein Flag–BAP, a 49 kDa fusion protein commercialized by Sigma-Aldrich (Figure 23a and b). To set-up the experiment, a first blot was done to compare the results obtained from quantifications performed on samples extracted in the presence or absence of 2-ME, in order to check if the presence of reducing agent leads to a better results.

Total proteins were extracted in ice-cold homogenisation buffer in the absence (Figure 23a) or presence (Figure 23b) of 2-ME; the homogenates were then centrifuged and supernatants were collected. Total proteins in supernatants were quantified using the Bradford protein assays and diluted and different amount (from 0.5 to 4µg) were analyzed for SDS-PAGE and immunoblotting with anti-FLAG antiserum. Absolute signal intensity of bands were quantified using Image-Lab (see Table below).

The calibration curves results, normalized for the different molecular weight of Flag-BAP standard (49 kDa) and of ER-BMP2 bands, indicated that ER-BMP2 appeared to be about 40ng on 4µg of total protein, representing around 1% of total soluble proteins (TSP), regardless the presence of 2-mercaptoethanol in the homogenization buffer. The higher absolute quantification (table) obtained on sample extracted in absence of reducing agent is probably due to the smeared signal, which was resolved in more discrete bands in the presence of 2-ME.




| Absence of 2-ME | Presence of 2-ME |
|---|---|
| Total: 48 ng | Total: 41 ng |

**Figure 23: Absolute quantification of ER-BMP2 in transgenic leaves.** Total proteins from leaves collected from one ER-BMP2 transgenic plant(ER23) were extracted in absence (Left blot, A) or presence (Right blot, B) of reducing agent 2-ME and progressive dilutions of leaf extracts were analyzed by SDS-PAGE and protein blot with anti-FLAG antiserum together with known amounts of the standard protein Flag–BAP. numbers on the top of the lanes indicate the amount of total proteins analyzed. The migration of molecular weight standards (kDa) is indicated.

Another protein blot experiment was performed in order to assess the amount of ER-BMP2 produced by different plants. Four micrograms of leaf extracts from 7 different ER-BMP2 expressing plants were analyzed on SDS-PAGE and protein blot. Signal intensities were compared to known amounts of the commercial standard protein Flag–BAP. Absolute quantification results demonstrate an high variability between each plant production levels with the amount of detected ER-BMP2 ranging from 0.5 to 2.5% of total proteins (see the table below).

### 3.2.5 rhBMP-2 Active Domain (rhBMP2-AD) is Efficiently Recovered on Thrombin Cleavage

To use rhBMP2 as a pharmaceutical, it will be necessary to minimize the extra sequences that are present in ER-BMP2; for this purpose, the accessibility of the thrombin cleavage site was investigated. Total protein were extracted in the presence of the reducing agent 2-mercaptoethanol in the homogenization buffer and, after incubation (+) or not (-) with thrombin, samples were analyzed by SDS-PAGE and protein blot using anti-FLAG antiserum.

The fusion protein ER-BMP2 (Figure 24 – ER-BMP2 lanes) was efficiently cleaved by thrombin (Figure 24 – lanes +), and the rhBMP2-AD (asterisk) was, in most part, released by thrombin cleavage in a 20 kDa form. Internal relative quantifications, performed with Image-Lab software, indicated that the released polypeptide includes about 50% of the non digested lane, indicating a sufficiently good accessibility to thrombin.

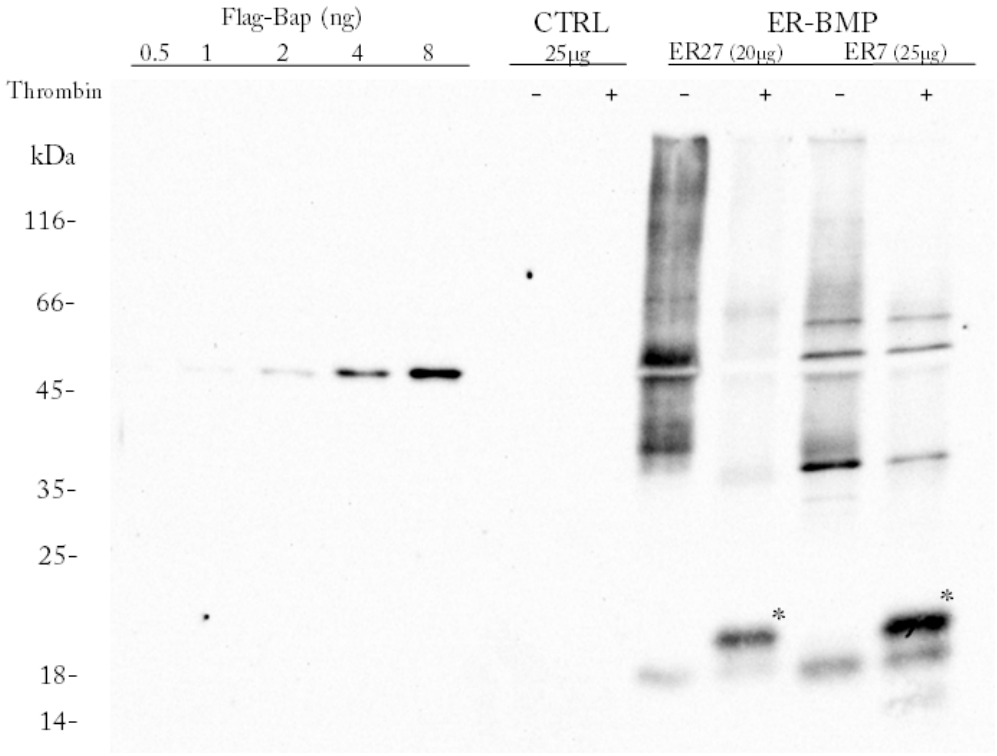
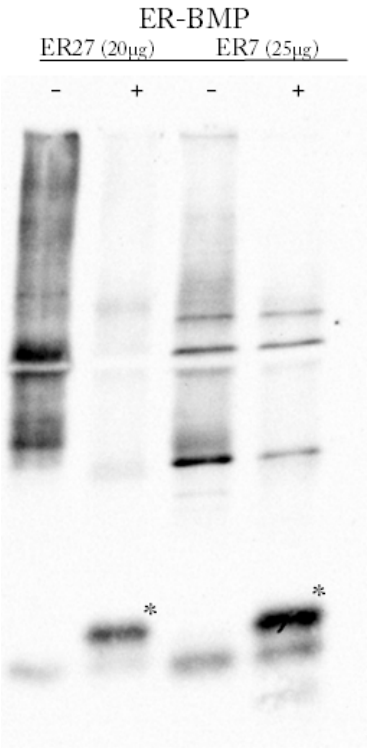



**Figure 24: rhBMP2-AD is efficiently recover by thrombin cleavage.** Total proteins from leaves, harvested from two ER-BMP2 transgenic plants (ER27 and ER7) and one Wild-Type plant (CTRL), were extracted in presence of reducing agent 2-ME and 20 or 25µg of samples were incubated (+) or not (-) with thrombin (Thr, 8U). Samples were then analyzed by SDS-PAGE and protein blot with anti-FLAG antiserum. Numbers on the top of the lanes indicate the amount of total proteins analyzed. The migration of molecular weight standards (kDa) is indicated.

### 3.2.6 Quantification of rhBMP2-AD




Absolute quantification of the rhBMP2-AD bands (Figure 25, asterisks) was performed using Image-Lab. Results were normalized for the different molecular weight of Flag-BAP standard compared to the molecular weight of the BMP-2 mature peptide. The software indicated that the fragments appeared to be 3.5ng/20µg for ER27 lane and 5.5ng/25µg for ER7 lane.
These results indicate that the amount of mature BMP-2 peptide in these plants was about 0.02% of total soluble protein.

**Figure 25: Absolute quantification of rhBMP2 active domain.** Detail of Figure 24, total proteins of two transgenic plant, ER27 (first 2 lanes, 20µg) and ER7(second part of the blot, 25µg) were treated (+) or not (-) with thrombin and analyzed with SDS-PAGE and immunoblotting.

An assessment of the yield of rhBMP2 for every g of fresh leaves was performed for ER27 and ER7 plants:

ER27: 3.5µg of rhBMP2 mature domain obtained for 1g of fresh leaves harvested.

ER7: 6,4 µg of rhBMP2 mature domain obtained for 1g of fresh leaves harvested.

This is a good results, however the high variability of the expression levels of each plant must be considered: choose plants with a higher level of expression may lead to an increase of the amount of recovered protein and to even better results.

# 4 Discussion

## 4.1 Overview

Bone morphogenetic proteins (BMPs) have been demonstrated to be effective for the treatment of bony defects, fracture repair and non-unions, but, due to the low yield, the limited amounts of source material and the potential risk of disease transmission, purified native BMPs are not suitable for large scale use. Furthermore, the yield of BMPs produced by recombinant expression in mammalian cells is relatively low compared to other recombinant protein production systems, and this make their use in common clinical applications extremely costly.

In this study, in order to improve rhBMP2 accumulation levels using a plant-based system, a γ-zein/BMP2 chimeric construct – expected to be retained and accumulated in the endoplasmic reticulum – was synthesized by transient and stable transformation experiments and characterized. The level of accumulation and the main characteristics of this construct were then compared to the native form of the protein.

## 4.2 ER-BMP2 and N-BMP2 Expression and Characterization

SDS-PAGE and protein blot experiments performed both on transiently transfected protoplasts and on transgenic plants tissues, showed a clear tendency of ER-BMP2 to accumulate at high levels. This result confirmed the ability of γ-zein to induce protein accumulation in cells and proved the non-recognition of the chimeric protein by the ER quality control system. The ER, indeed, has the function of controlling successful folding and assembling of the newly synthesized proteins and of delivering to the degradation pathways those polypeptides that do not meet expected requirements. Consequently, this property of the reticulum may be a serious drawback when the aim is to increase ER accumulation of secretory proteins. Previous results showed that when γ-zein N-terminal was fused to cytosolic protein Nef, it was unable to prevent the degradation of this protein; instead, when a fusion on Nef to the complete zeolin sequence (consisting of the N-terminal of γ-zein fused to the C-terminal of phaseolin [258]) was performed, the protein avoided the degradation and formed PBs [287]. All these data suggest that besides the necessary role of the hydrophobic regions and the Cysteine residues of zein [265] another kind of stabilization may be needed in order to avoid degradation early after synthesis and provide enough time to start PB assembly [271,287]. In the ER-BMP2 case, this mechanism can be accomplished by the 7 Cysteines of BMP-2 mature domain, which may form inter-chain bonds able to stabilize the structure and promote the formation of polymers.

During immunoblot experiments, ER-BMP2 demonstrated to be a soluble N-glycosylated protein that, in non reducing conditions, migrates mainly in a dimeric form which is disassembled in monomers of approximately 35 kDa in the presence of reducing agent. Moreover, the presence of 2-ME in the homogenisation buffer led to detection of a further band of about 20 kDa; this particular band demonstrated to be N-glycosylated and sensitive to EndoH treatment, but its molecular weight was not affected by thrombine treatment. Such findings suggest that most probably this band corresponds to a specific C-terminal product of degradation on ER-BMP2, but it is unknown if it will conserve its activity after purification.

Accordingly to the original design of ER-BMP, in order to obtain a correct glycosylated rhBMP2-AD that could be used for pharmaceutical and clinical purposes, a final ER localization mediated by PBs formation was expected. Indeed, locating a recombinant protein in the ER of a plant based system has two main advantages: the first one is protecting the protein from protease degradation, the second one is to obtain a correct N-glycosylated protein. As described before, plant glycosylation presents some differences respect to the human counterpart. In plants the N-glycosylation starts in the endoplasmic reticulum by the co-translational transfer of an oligosaccharide precursor to specific Asparagine residues of the nascent polypeptide chain; the processing of this oligosaccharide into high-mannose-type, paucimannosidic-type, hybrid-type or complex-type N-glycans occurs in the secretory pathway as the glycoprotein moves out from the ER to its final destination. At the end of their maturation, some plant N-glycans have typical structures that differ from those found in mammalian by the absence of sialic acid [288] and the presence of β(1,2)-xylose and α(1,3)-fucose residues[289]. These plant glyco-epitopes are of major interest in the context of the production of proteins dedicated to human therapy because they could be immunogenic in mammals [290] and could be involved in allergic reactions [291]. The localization of ER-BMP2, which has a glycosylation site in its mature domain, is therefore an important issue for its future applications. Even if the carbohydrate moiety does not appear to be essential for biological activity of BMPs, since the deglycosylated proteins are still capable of inducing bone formation in vivo [141], the lack of such modification leads to a decrease in solubility of proteins and may lead to misfolding of the final product [143]. Given the results on ER-BMP2 solubility, the localization of the protein was investigated by Endoglycosidase H assay. This experiment demonstrated that the protein has an unmodified N-glycan that is sensitive to the enzyme and,

accordingly, ER-BMP2 is not only properly post-translationally modified, but is also located within the endoplasmic reticulum.

As described before, γ-zein N-terminal fragment induces the formation of insoluble PBs even when it is fused to other proteins of the secretory pathway, like the C-terminal of phaseolin [258,267]. For this reason the solubility of ER-BMP2 in the absence of reducing agent was an unexpected result and its ability to form large polymers was further investigated using sucrose velocity gradient centrifugation. Nevertheless, even if ER-BMP2 is mainly present in the dimeric form, the protein is able to form larger oligomers (up to 250 kDa). This different behavior of the chimeric protein respect to γ-zein and zeolin proteins may be due to a kind of unexpected interaction between the two portions of ER-BMP2. Indeed, it is possible that the 6 cysteines of the γ-zein portion and the 7 ones of BMP-2 mature domain interfere each other in spite of the presence of the DNA-linker, that fails its function of ensuring independent folding of the peptides.
The final recovery of the protein from the expression host is one of the most significant drawbacks, both in terms of costs and final productivity, in non-secreted recombinant protein production. To accumulate recombinant proteins in insoluble PBs provides an advantage because, by virtue of their high density, they can be easily isolated from crude cell or tissue extracts by sucrose density gradients fractioning or by low-speed direct centrifugation of the homogenates [265]. Because of the solubility of ER-BMP2 this methodology is not a feasible option and a different strategy is required. One viable option could be an affinity chromatography purification performed with specific antibodies; for example, if the protein will be purified using a anti-γ-zein antibody, the recovery of the BMP-2 active domain will be easier, since it will be released in the supernatant while the γ-zein portion will remain bound to the antibody [252].

Seed storage proteins have unique characteristics that allow very high accumulation during seed development [261]. Based on their specific features, like assembly, solubility, and intracellular localization, they can be divided into three major classes: albumins, globulins and prolamins [292]. Prolamins form large polymers that rapidly become insoluble and do not proceed along the secretory pathway, accumulating as protein bodies within the ER lumen of endosperm cells and consequently in seeds [261]. When γ-zein, which is one of the most ancient maize prolamins, is individually expressed in vegetative tissues of transgenic plants it forms homotypic PBs within the ER, indicating that it contains all the necessary information to form an ER-retained polymer [264].

For these reasons, the presence of ER-BMP2 – which contains the N-terminal portion of γ-zein and forms oligomers in the ER – in the seeds was investigated. Even if the accumulation level was not as high as that of γ-zein in maize seeds, probably due to a competition between the chimeric protein and the endogenous tobacco storage proteins, the accumulation of ER-BMP2 in seeds *per se* is a promising result. Indeed, there are significant advantages in directing the expression of a recombinant protein into the plant seed. In fact, unlike leaf tissues, which consists mostly of water and fiber, seeds have a low content of water, offer high stability and natural microencapsulation, and are capable to high accumulation of non-endogenous proteins. Even if seeds have evolved to store precise amounts of proteins at maximum density within storage cells and consequently leaving little space for additional products resulting from transgene expression, there are some strategies that lead to an increased degree of storage substance plasticity. For example, it is possible to redirect a significant part of the protein synthesis capacity from the production of intrinsic seed proteins to the synthesis of foreign proteins, by silencing one of more genes of endogenous storage proteins. In this way the recombinant protein of interest could accumulate at higher levels in seeds, exceeding the 8% of the total soluble proteins of seeds [293], without compromising their ability to germinate [293].

Compared to ER-BMP2, N-BMP2 demonstrated a lower level of accumulation with barely detectable bands in immunoblot performed on transgenic plants extracts.

When immunoblotting experiments were performed on transiently transformed protoplast, a monomeric form of N-BMP2, that changed its mobility according to the presence of reducing agent, could be detected. The mobility shift of the monomer was probably due to a rearrangement of the intra-chain disulphide bonds caused by the presence of 2-ME, or maybe, the presence of BMP-2 pro-peptide prevented the correct folding of the protein.

Conversely, immunoblotting experiments performed on transgenic plants extracts resulted in an intricate pattern of bands of non-expected molecular weight. This could be due to non-specific detection caused by an high amount of total proteins analyzed and/or to a degradation of the protein, that creates a number of bands of different molecular weight.

These results suggest a combination of adverse factors that led to a low level of transcription. One factor that could be involved in the low expression levels might be the presence of the native human signal peptide of the protein, that may be not effective in a vegetable system. Another possibility is

the addressing of N-BMP2 towards the degradation pathway, due to misfolding or to an incorrect processing during protein synthesis.
All the collected data indicate that without a further strategy that leads to an increased level of accumulation and, in case of secretion, bypasses the issues of the modified glycosylation out of the ER, the production of a native form of the BMP-2 protein in transgenic plants is not a viable option.

## 4.3 ER-BMP2 Quantification

Absolute quantification of produced ER-BMP2 was performed on different plant's extracts and results indicate a highly variable yield. On the average, ER-BMP2 represents about the 1% of total leaf soluble proteins, a value that is usually considered as high accumulation in vegetative tissues of transgenic plants, even if it is not as high as the 3.5% obtained from zeolin transgenic plants [252]. The first quantification experiment, performed on total proteins extracted in the presence and in the absence of the reducing agent in order to make a comparison on the amount of protein obtained, indicated that there was no significant difference in protein amount between the extraction procedures. For this reason further quantification analyses were performed on protein samples extracted in the absence of 2-ME. The data obtained from these plants were highly variable, ranging from 0.5% to more than 2.5%. However, it is known that the expression levels of recombinant proteins may show wide variations between independently transformed plants carrying the same construct, and that several factors, including integration site, transgene copy number and transgenic locus configuration, can contribute to these variations [283]. For this reason, besides the accurate screening and selection of the plants with the highest level of expression, it is also possible to apply strategies that may improve these levels, for example selecting plants with multiple copies of the gene of interest.

However, plant-derived recombinant proteins, produced in transgenic tobacco plants and extracted directly from leaves, are generally produced at low levels, typically about the 0.1% of the total soluble protein [230]. This low level of production reflects a combination of factors, with poor protein folding and stability among the most relevant, specific for every recombinant protein [294,269]. Nevertheless, in the recent years, other strategies have permitted an higher accumulation of recombinant proteins; for example, a tobacco chloroplast system of expression has been used to

produce human proteins, like Human Serum Albumin, getting very solid results [295]. Even if our results do not reach such levels, they still represent a good starting point that can be further improved with the implementation of the strategies adopted.

After demonstrating the accessibility of the thrombin cleavage, the amount of rhBMP2 active domain recovered from 1g of fresh leaves was assessed in two different plants. The best performance achieved was 6,4µg of rhBMP2-AD every g of fresh leaves. Considering the high variability of the expression levels, this is a good result which can still be improved by selecting plants with even better levels of accumulation or maybe trying to direct the protein to other cellular compartments, given the excellent results recently achieved, for example, in tobacco chloroplasts [295].

Because plant-based techniques do not require the same expensive investments as other production methods, once the initial costs due to the F0 transgenic production have been overcome, the propagation of plants implies a reduced financial impact. Rough calculations based on our results suggest that, considering the size and biomass of the greenhouse-produced tobacco plants, in order to obtain the amount of rhBMP2-AD which is contained in a single commercial Infuse® kit, only one ER-BMP2 tobacco plant is required.

## 4.4 Future Perspectives

The outbreak of Ebola virus disease has highlighted the potential advantages of plant-based production systems. Indeed, to develop a vaccine that would be appropriate for human use, three different mouse monoclonal antibodies (mAbs) directed against three distinct Ebola virus glycoprotein epitopes were chimerized with human constant regions of antibodies and, in order to evaluate a more cost-effective and scalable alternative to production in (CHO) cells, they were produced in a *Nicotiana benthamiana*-based antibody manufacturing platform [296]. This system allows rapid, scalable production of mAbs in less than a month. Moreover, via the use of a transgenic strain of *N. benthamiana* in which plant-specific glycosyltransferases (α(1,3)-fucosyltransferase and β(1,2)-xylosyltransferase) are inhibited by RNA interference (RNAi), the derived mAbs have homogenous mammalian glycans [297]. Due to more and more encouraging results, plant based pharmaceuticals are finally reaching a wide acceptance, and clinical trials investigating their effectiveness have multiplied (Table 4). It is important to think about a great advantage of plant-made pharmaceuticals: their cost-benefits advantages can allow more capital to

be invested in research and development of new therapeutics, bringing an overall advantage to the society.

Nowadays the diversity of plant-based systems contrasts with the small number of microbial and animal-cell platforms regarded as industry gold standards; this can be considered as a strength in terms of innovation for product-specific requirements, but also as a drawback in terms of standardization and regulatory harmonization. Indeed, the limited number of approved production facilities correlates with the relatively small number of plant-derived pharmaceutical products that reached the market or are undergoing clinical development.

**Table 4: plant-derived pharmaceutical products assessed in clinical studies**

| Company | Products | Main Application | Current Status |
|---|---|---|---|
| Protalix | Elelyso | Gaucher's ERT | FDA-approved in US, not in Europe |
| | α-Galactosidase | Fabry ERT | Phase I/II |
| | Oral Glucocerebrosidase | Gaucher's ERT | Phase II |
| Ventria | Lactoferrin | Antibiotic-associated diarrhea, anti-inflammatory | Phase II |
| | VEN120 | Inflammatory bowel disease | Phase I |
| | VEN130 | Osteoporosis | Phase I |
| Biolex | Locteron™ | HCV | Phase II/IIb |
| Icon Genetics | NHL vaccine | HSV/HIV | Phase I/II |
| | MAPP66 | Monoclonal antibodies | Phase I |
| Medicago | H5 | Pandemic influenza vaccine | Phase II/III, approved for emergency use |
| | H5 intradermal | Pandemic influenza vaccine | Phase I |
| | Seasonal influenza vaccine | Seasonal influenza vaccine | Phase I |
| Planet Biotechnology | CaroRX | Anti-caries antibody | Approved as medical device |
| Fraunhofer IME | HIV antibody | Microbicide | Phase I |
| Fraunhofer CMB | HA vaccine | Vaccine | Phase I |
| VAXX/Arizona State University | NoroVAXX | Vaccine | Phase I |
| MAPP | ZMapp | Ebola antibody cocktail | Emergency use, Phase I expected soon |

*Data extracted from Sack et al., 2015 [298]

Nevertheless, finally something is changing. Elelyso is a transgenic suspended carrot cells expressed form of the glucocerebrosidase enzyme that has been approved from FDA for injection and indicated for long-term enzyme replacement therapy for adults with a confirmed diagnosis of Gaucher disease [299].

Moreover, in order to avoid the lengthy and costly regulatory procedures for FDA approval, several companies have established commercial platforms for the production of non-pharmaceutical products, even in order to generate a revenue [298]. For example TrypZean™ is a recombinant bovine trypsin, sold by Sigma-Aldrich company, that is produced in corn and is sold for research purposes[300]. Another example is ORF Genetics, that in 2007 have introduced a next generation of human growth factors and cytokines produced in the endosperm tissue of the barley grain and marketed under the brand name ISOkine™.

After the necessary tests that assess an *in vitro* effectiveness of the protein, also rhBMP-2 protein produced in Tobacco transgenic plants might find its primary market as a non-pharmaceutical products. This market strategy would avoid the time required for regulatory procedures and guarantee time necessary for a first assessment of its effectiveness after a scale up of protein production, creating a framework for its possible further use in clinical practice. Moreover, in the last years BMP-2 have been demonstrated to play key roles as neurotrophic factor, by regulating the development of the nervous system and its maintenance in adulthood, drawing attention to its potential use in the treatment of neurodegenerative diseases [301,302]. For these reasons, the possibility to use the recombinant BMP-2 protein at a reduced cost, not only in the studies on osteogenic differentiation, but also for those on innovative therapies, could have a positive impact on its market. Because Italy don't allow the cultivation of transgenic plants for commercial purpose, the first step required must be the recruitment of a foreign company that has the possibility to perform a scale-up of the protein production, in order to obtain it in quantity that allow a large scale effectiveness tests.

# 5 Conclusions

The poor yield of recombinant human BMPs in mammalian cell culture systems led to the investigation of a method to enhance the production of these proteins. The addition of γ-zein N-terminal domain to the BMP-2 mature domain sequence resulted in a significant increase of accumulation of the protein in plant tissues while showing no negative effects on cell viability and plant growth.

Total protein extracts from transgenic plants showed high levels of ER-BMP2 accumulation, which were even higher if compared to those of N-BMP2 plants.

Cleavage of mature BMP-2 domain using the thrombine enzyme also demonstrated promising results: mature BMP-2 proteins were observed after incubation with the enzyme; however, the biological activity of the rhBMP2-AD has to be further investigated.

In summary, producing and accumulating recombinant human BMP-2 in the ER of transgenic tobacco plants demonstrated to be a safe, rapid and simple strategy that significantly enhances its recovery during the extraction procedure. Nevertheless, further investigations on the biological activity of this BMP-2 on different cell cultures and in vivo tests are needed before a potential commercialization of this production strategy might take place.

# 6 Bibliography

1. Bessa, P. C., Casal, M. & Reis, R. L. Bone morphogenetic proteins in tissue engineering: the road from the laboratory to the clinic, part I (basic concepts). *J. Tissue Eng. Regen. Med.* 2, 1–13 (2008).
2. Lewandrowski, K. U., Gresser, J. D., Wise, D. L. & Trantol, D. J. Bioresorbable bone graft substitutes of different osteoconductivities: a histologic evaluation of osteointegration of poly(propylene glycol-co-fumaric acid)-based cement implants in rats. *Biomaterials* 21, 757–764 (2000).
3. Giannoudis, P. V., Dinopoulos, H. & Tsiridis, E. Bone substitutes: an update. *Injury* 36 Suppl 3, S20–27 (2005).
4. Baroli, B. From natural bone grafts to tissue engineering therapeutics: Brainstorming on pharmaceutical formulative requirements and challenges. *J. Pharm. Sci.* 98, 1317–1375 (2009).
5. U.S. Markets for Fracture Fixation Products | Medical Device Market Research. *Life Science Intelligence* (2011). at <http://lifescienceintelligence.com/market-reports-page.php?id=A326>
6. the Bone and Joint Decade. *The Bone and Joint Decade*. (2009). at <http://bjdonline.org/>
7. Dimitriou, R., Jones, E., McGonagle, D. & Giannoudis, P. V. Bone regeneration: current concepts and future directions. *BMC Med.* 9, 66 (2011).
8. Nazirkar, G., Singh, S., Dole, V. & Nikam, A. Effortless effort in bone regeneration: a review. *J. Int. Oral Health JIOH* 6, 120–124 (2014).
9. Delloye, C., Cornu, O., Druez, V. & Barbier, O. Bone allografts: What they can offer and what they cannot. *J. Bone Joint Surg. Br.* 89, 574–579 (2007).
10. Wozney, J. M. Overview of bone morphogenetic proteins. *Spine* 27, S2–8 (2002).
11. St John, T. A. *et al.* Physical and monetary costs associated with autogenous bone graft harvesting. *Am. J. Orthop. Belle Mead NJ* 32, 18–23 (2003).
12. Dimitriou, R., Mataliotakis, G. I., Angoules, A. G., Kanakaris, N. K. & Giannoudis, P. V. Complications following autologous bone graft harvesting from the iliac crest and using the RIA: a systematic review. *Injury* 42 Suppl 2, S3–15 (2011).
13. Garrison, K. R. *et al.* Clinical effectiveness and cost-effectiveness of bone morphogenetic proteins in the non-healing of fractures and spinal fusion: a systematic review. *Health Technol. Assess. Winch. Engl.* 11, 1–150, iii–iv (2007).
14. Senn N. Senn on the Healing of Aseptic Bone Cavities by Implantation of Antiseptic Decalcified Bone. *Ann. Surg.* 10, 352–368 (1889).
15. Levander, G. On the formation of new bone in bone transplantation. *Acta Chir. Scand.* 74, 425–426 (1934).
16. Levander, G. A study of bone regeneration. *J. Surg. Gynecol. Obstet.* 67, 705–714 (1938).
17. Urist, M. R. Bone: formation by autoinduction. *Science* 150, 893–899 (1965).
18. Urist, M. R. & Strates, B. S. Bone morphogenetic protein. *J. Dent. Res.* 50, 1392–1406 (1971).
19. Reddi, A. H. & Huggins, C. Biochemical sequences in the transformation of normal

fibroblasts in adolescent rats. *Proc. Natl. Acad. Sci. U. S. A.* 69, 1601–1605 (1972).
20. Reddi, A. H. Cell biology and biochemistry of endochondral bone development. *Coll. Relat. Res.* 1, 209–226 (1981).
21. Sampath, T. K. & Reddi, A. H. Dissociative extraction and reconstitution of extracellular matrix components involved in local bone differentiation. *Proc. Natl. Acad. Sci. U. S. A.* 78, 7599–7603 (1981).
22. Sampath, T. K. & Reddi, A. H. Homology of bone-inductive proteins from human, monkey, bovine, and rat extracellular matrix. *Proc. Natl. Acad. Sci. U. S. A.* 80, 6591–6595 (1983).
23. Reddi, A. H. Morphogenesis and tissue engineering of bone and cartilage: inductive signals, stem cells, and biomimetic biomaterials. *Tissue Eng.* 6, 351–359 (2000).
24. Reddi, A. H. Bone and cartilage differentiation. *Curr. Opin. Genet. Dev.* 4, 737–744 (1994).
25. Wozney, J. M. *et al.* Growth factors influencing bone development. *J. Cell Sci. Suppl.* 13, 149–156 (1990).
26. Wang, E. A. *et al.* Recombinant human bone morphogenetic protein induces bone formation. *Proc. Natl. Acad. Sci. U. S. A.* 87, 2220–2224 (1990).
27. Sampath, T. K. & Reddi, A. H. Distribution of bone inductive proteins in mineralized and demineralized extracellular matrix. *Biochem. Biophys. Res. Commun.* 119, 949–954 (1984).
28. Wang, E. A. *et al.* Purification and characterization of other distinct bone-inducing factors. *Proc. Natl. Acad. Sci. U. S. A.* 85, 9484–9488 (1988).
29. Wozney, J. M. *et al.* Novel regulators of bone formation: molecular clones and activities. *Science* 242, 1528–1534 (1988).
30. Boyne, P. J. *et al.* A feasibility study evaluating rhBMP-2/absorbable collagen sponge for maxillary sinus floor augmentation. *Int. J. Periodontics Restorative Dent.* 17, 11–25 (1997).
31. McKay, B. & Sandhu, H. S. Use of recombinant human bone morphogenetic protein-2 in spinal fusion applications. *Spine* 27, S66–85 (2002).
32. Nakashima, M. & Reddi, A. H. The application of bone morphogenetic proteins to dental tissue engineering. *Nat. Biotechnol.* 21, 1025–1032 (2003).
33. Seeherman, H. & Wozney, J. M. Delivery of bone morphogenetic proteins for orthopedic tissue regeneration. *Cytokine Growth Factor Rev.* 16, 329–345 (2005).
34. Sampath, T. K. *et al.* Recombinant human osteogenic protein-1 (hOP-1) induces new bone formation in vivo with a specific activity comparable with natural bovine osteogenic protein and stimulates osteoblast proliferation and differentiation in vitro. *J. Biol. Chem.* 267, 20352–20362 (1992).
35. Israel, D. I., Nove, J., Kerns, K. M., Moutsatsos, I. K. & Kaufman, R. J. Expression and characterization of bone morphogenetic protein-2 in Chinese hamster ovary cells. *Growth Factors Chur Switz.* 7, 139–150 (1992).
36. Food and Drug Administration. OP-1™ - H010002. (2001). at <http://www.fda.gov/medicaldevices/productsandmedicalprocedures/deviceapprovalsandclearances/recently-approveddevices/ucm085026.htm>
37. Food and Drug Administration. INFUSE® Bone Graft - P000054. (2002). at <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cftopic/pma/pma.cfm?num=p000054>
38. Food and Drug Administration. InFUSE™ Bone Graft/LT-CAGE™ - P000058. (2002). at

<http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cftopic/pma/pma.cfm?num=p000058>
39. Food and Drug Administration. OP-1 Putty - H020008. (2004). at <http://www.fda.gov/medicaldevices/productsandmedicalprocedures/deviceapprovalsandclearances/recently-approveddevices/ucm081181.htm>
40. Hogan, B. L. Bone morphogenetic proteins: multifunctional regulators of vertebrate development. *Genes Dev.* 10, 1580–1594 (1996).
41. Reddi, A. H. Role of morphogenetic proteins in skeletal tissue engineering and regeneration. *Nat. Biotechnol.* 16, 247–252 (1998).
42. Yamashita, H., Ten Dijke, P., Heldin, C. H. & Miyazono, K. Bone morphogenetic protein receptors. *Bone* 19, 569–574 (1996).
43. Hoffmann, A., Weich, H. A., Gross, G. & Hillmann, G. Perspectives in the biological function, the technical and therapeutic application of bone morphogenetic proteins. *Appl. Microbiol. Biotechnol.* 57, 294–308 (2001).
44. Ozkaynak, E. *et al.* Osteogenic protein-2. A new member of the transforming growth factor-beta superfamily expressed early in embryogenesis. *J. Biol. Chem.* 267, 25220–25227 (1992).
45. Lyons, K. M., Jones, C. M. & Hogan, B. L. The DVR gene family in embryonic development. *Trends Genet. TIG* 7, 408–412 (1991).
46. Matzuk, M. M. Functional analysis of mammalian members of the transforming growth factor-beta superfamily. *Trends Endocrinol. Metab. TEM* 6, 120–127 (1995).
47. Karsenty, G., Luo, G., Hofmann, C. & Bradley, A. BMP 7 is required for nephrogenesis, eye development, and skeletal patterning. *Ann. N. Y. Acad. Sci.* 785, 98–107 (1996).
48. Ten Dijke, P., Miyazono, K. & Heldin, C. H. Signaling via hetero-oligomeric complexes of type I and type II serine/threonine kinase receptors. *Curr. Opin. Cell Biol.* 8, 139–145 (1996).
49. Cook, S. D., Wolfe, M. W., Salkeld, S. L. & Rueger, D. C. Effect of recombinant human osteogenic protein-1 on healing of segmental defects in non-human primates. *J. Bone Joint Surg. Am.* 77, 734–750 (1995).
50. Simic, P. & Vukicevic, S. Bone morphogenetic proteins: from developmental signals to tissue regeneration. Conference on bone morphogenetic proteins. *EMBO Rep.* 8, 327–331 (2007).
51. Koenig, B. B. *et al.* Characterization and cloning of a receptor for BMP-2 and BMP-4 from NIH 3T3 cells. *Mol. Cell. Biol.* 14, 5961–5974 (1994).
52. Imamura, T. *et al.* Smad6 inhibits signalling by the TGF-beta superfamily. *Nature* 389, 622–626 (1997).
53. Yamashita, H., ten Dijke, P., Franzén, P., Miyazono, K. & Heldin, C. H. Formation of hetero-oligomeric complexes of type I and type II receptors for transforming growth factor-beta. *J. Biol. Chem.* 269, 20172–20178 (1994).
54. Liu, F., Ventura, F., Doody, J. & Massagué, J. Human type II receptor for bone morphogenic proteins (BMPs): extension of the two-kinase receptor model to the BMPs. *Mol. Cell. Biol.* 15, 3479–3486 (1995).
55. Ebara, S. & Nakayama, K. Mechanism for the action of bone morphogenetic proteins and regulation of their activity. *Spine* 27, S10–15 (2002).
56. Rosenzweig, B. L. *et al.* Cloning and characterization of a human type II receptor for

bone morphogenetic proteins. *Proc. Natl. Acad. Sci. U. S. A.* 92, 7632–7636 (1995).
57. Hoodless, P. A. *et al.* MADR1, a MAD-related protein that functions in BMP2 signaling pathways. *Cell* 85, 489–500 (1996).
58. Komori, T. *et al.* Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. *Cell* 89, 755–764 (1997).
59. Komori, T. Regulation of osteoblast differentiation by transcription factors. *J. Cell. Biochem.* 99, 1233–1239 (2006).
60. Yamaguchi, A., Komori, T. & Suda, T. Regulation of osteoblast differentiation mediated by bone morphogenetic proteins, hedgehogs, and Cbfa1. *Endocr. Rev.* 21, 393–411 (2000).
61. Nishimura, R., Hata, K., Harris, S. E., Ikeda, F. & Yoneda, T. Core-binding factor alpha 1 (Cbfa1) induces osteoblastic differentiation of C2C12 cells without interactions with Smad1 and Smad5. *Bone* 31, 303–312 (2002).
62. Liu, T. *et al.* BMP-2 promotes differentiation of osteoblasts and chondroblasts in Runx2-deficient cell lines. *J. Cell. Physiol.* 211, 728–735 (2007).
63. Ulsamer, A. *et al.* BMP-2 induces Osterix expression through up-regulation of Dlx5 and its phosphorylation by p38. *J. Biol. Chem.* 283, 3816–3826 (2008).
64. Hayashi, H. *et al.* The MAD-related protein Smad7 associates with the TGFbeta receptor and functions as an antagonist of TGFbeta signaling. *Cell* 89, 1165–1173 (1997).
65. Piccolo, S., Sasai, Y., Lu, B. & De Robertis, E. M. Dorsoventral patterning in Xenopus: inhibition of ventral signals by direct binding of chordin to BMP-4. *Cell* 86, 589–598 (1996).
66. Zimmerman, L. B., De Jesús-Escobar, J. M. & Harland, R. M. The Spemann organizer signal noggin binds and inactivates bone morphogenetic protein 4. *Cell* 86, 599–606 (1996).
67. Yamashita, H. *et al.* Osteogenic protein-1 binds to activin type II receptors and induces certain activin-like effects. *J. Cell Biol.* 130, 217–226 (1995).
68. Blader, P., Rastegar, S., Fischer, N. & Strähle, U. Cleavage of the BMP-4 antagonist chordin by zebrafish tolloid. *Science* 278, 1937–1940 (1997).
69. Kusanagi, K. *et al.* Characterization of a bone morphogenetic protein-responsive Smad-binding element. *Mol. Biol. Cell* 11, 555–565 (2000).
70. Chen, D., Zhao, M. & Mundy, G. R. Bone morphogenetic proteins. *Growth Factors Chur Switz.* 22, 233–241 (2004).
71. Daluiski, A. *et al.* Bone morphogenetic protein-3 is a negative regulator of bone density. *Nat. Genet.* 27, 84–88 (2001).
72. Yamaguchi, A. *et al.* Recombinant human bone morphogenetic protein-2 stimulates osteoblastic maturation and inhibits myogenic differentiation in vitro. *J. Cell Biol.* 113, 681–687 (1991).
73. Thies, R. S. *et al.* Recombinant human bone morphogenetic protein-2 induces osteoblastic differentiation in W-20-17 stromal cells. *Endocrinology* 130, 1318–1324 (1992).
74. Katagiri, T. *et al.* Bone morphogenetic protein-2 converts the differentiation pathway of C2C12 myoblasts into the osteoblast lineage. *J. Cell Biol.* 127, 1755–1766 (1994).
75. Moses, H. L. & Serra, R. Regulation of differentiation by TGF-beta. *Curr. Opin. Genet. Dev.* 6, 581–586 (1996).
76. Chen, D. *et al.* Bone morphogenetic protein 2 (BMP-2) enhances BMP-3, BMP-4,

and bone cell differentiation marker gene expression during the induction of mineralized bone matrix formation in cultures of fetal rat calvarial osteoblasts. *Calcif. Tissue Int.* 60, 283–290 (1997).
77. Chen, D. *et al.* Differential roles for bone morphogenetic protein (BMP) receptor type IB and IA in differentiation and specification of mesenchymal precursor cells to osteoblast and adipocyte lineages. *J. Cell Biol.* 142, 295–305 (1998).
78. Zhao, M. *et al.* Bone morphogenetic protein receptor signaling is necessary for normal murine postnatal bone formation. *J. Cell Biol.* 157, 1049–1060 (2002).
79. Bostrom, M. P. *et al.* Immunolocalization and expression of bone morphogenetic proteins 2 and 4 in fracture healing. *J. Orthop. Res. Off. Publ. Orthop. Res. Soc.* 13, 357–367 (1995).
80. Champagne, C. M., Takebe, J., Offenbacher, S. & Cooper, L. F. Macrophage cell lines produce osteoinductive signals that include bone morphogenetic protein-2. *Bone* 30, 26–31 (2002).
81. Yang, L. J. & Jin, Y. Immunohistochemical observations on bone morphogenetic protein in normal and abnormal conditions. *Clin. Orthop.* 249–256 (1990).
82. Baur, S. T., Mai, J. J. & Dymecki, S. M. Combinatorial signaling through BMP receptor IB and GDF5: shaping of the distal mouse limb and the genetics of distal limb diversity. *Dev. Camb. Engl.* 127, 605–619 (2000).
83. Jena, N., Martín-Seisdedos, C., McCue, P. & Croce, C. M. BMP7 null mutation in mice: developmental defects in skeleton, kidney, and eye. *Exp. Cell Res.* 230, 28–37 (1997).
84. Zhang, H. & Bradley, A. Mice deficient for BMP2 are nonviable and have defects in amnion/chorion and cardiac development. *Dev. Camb. Engl.* 122, 2977–2986 (1996).
85. Boden, S. D. Bioactive factors for bone tissue engineering. *Clin. Orthop.* S84–94 (1999).
86. Poynton, A. R. & Lane, J. M. Safety profile for the clinical use of bone morphogenetic proteins in the spine. *Spine* 27, S40–48 (2002).
87. Seeherman, H. J. *et al.* Recombinant human bone morphogenetic protein-2 delivered in an injectable calcium phosphate paste accelerates osteotomy-site healing in a nonhuman primate model. *J. Bone Joint Surg. Am.* 86-A, 1961–1972 (2004).
88. Boden, S. D., Kang, J., Sandhu, H. & Heller, J. G. Use of recombinant human bone morphogenetic protein-2 to achieve posterolateral lumbar spine fusion in humans: a prospective, randomized clinical pilot trial: 2002 Volvo Award in clinical studies. *Spine* 27, 2662–2673 (2002).
89. Dimar, J. R., Glassman, S. D., Burkus, K. J. & Carreon, L. Y. Clinical outcomes and fusion success at 2 years of single-level instrumented posterolateral fusions with recombinant human bone morphogenetic protein-2/compression resistant matrix versus iliac crest bone graft. *Spine* 31, 2534–2539; discussion 2540 (2006).
90. Glassman, S. D. *et al.* Initial fusion rates with recombinant human bone morphogenetic protein-2/compression resistant matrix and a hydroxyapatite and tricalcium phosphate/collagen carrier in posterolateral spinal fusion. *Spine* 30, 1694–1698 (2005).
91. Lanman, T. H. & Hopkins, T. J. Lumbar interbody fusion after treatment with recombinant human bone morphogenetic protein-2 added to poly(L-lactide-co-D,L-lactide) bioresorbable implants. *Neurosurg. Focus* 16, E9 (2004).
92. Boakye, M., Mummaneni, P. V., Garrett, M., Rodts, G. & Haid, R. Anterior cervical

discectomy and fusion involving a polyetheretherketone spacer and bone morphogenetic protein. *J. Neurosurg. Spine* 2, 521–525 (2005).
93. Lo, K. W.-H., Ulery, B. D., Ashe, K. M. & Laurencin, C. T. Studies of bone morphogenetic protein-based surgical repair. *Adv. Drug Deliv. Rev.* 64, 1277–1291 (2012).
94. McKay, W. F., Peckham, S. M. & Badura, J. M. A comprehensive clinical review of recombinant human bone morphogenetic protein-2 (INFUSE Bone Graft). *Int. Orthop.* 31, 729–734 (2007).
95. White, A. P. *et al.* Clinical applications of BMP-7/OP-1 in fractures, nonunions and spinal fusion. *Int. Orthop.* 31, 735–741 (2007).
96. Chen, Z., Ba, G., Shen, T. & Fu, Q. Recombinant human bone morphogenetic protein-2 versus autogenous iliac crest bone graft for lumbar fusion: a meta-analysis of ten randomized controlled trials. *Arch. Orthop. Trauma Surg.* 132, 1725–1740 (2012).
97. Schwartz, C. E. *et al.* Prospective evaluation of chronic pain associated with posterior autologous iliac crest bone graft harvest and its effect on postoperative outcome. *Health Qual. Life Outcomes* 7, 49 (2009).
98. Burkus, J. K., Gornet, M. F., Dickman, C. A. & Zdeblick, T. A. Anterior lumbar interbody fusion using rhBMP-2 with tapered interbody cages. *J. Spinal Disord. Tech.* 15, 337–349 (2002).
99. Burkus, J. K., Sandhu, H. S., Gornet, M. F. & Longley, M. C. Use of rhBMP-2 in combination with structural cortical allografts: clinical and radiographic outcomes in anterior lumbar spinal surgery. *J. Bone Joint Surg. Am.* 87, 1205–1212 (2005).
100. Haid, R. W., Branch, C. L., Alexander, J. T. & Burkus, J. K. Posterior lumbar interbody fusion using recombinant human bone morphogenetic protein type 2 with cylindrical interbody cages. *Spine J. Off. J. North Am. Spine Soc.* 4, 527–538; discussion 538–539 (2004).
101. Mummaneni, P. V., Pan, J., Haid, R. W. & Rodts, G. E. Contribution of recombinant human bone morphogenetic protein-2 to the rapid creation of interbody fusion when used in transforaminal lumbar interbody fusion: a preliminary report. Invited submission from the Joint Section Meeting on Disorders of the Spine and Peripheral Nerves, March 2004. *J. Neurosurg. Spine* 1, 19–23 (2004).
102. Glassman, S. D. *et al.* Posterolateral lumbar spine fusion with INFUSE bone graft. *Spine J. Off. J. North Am. Spine Soc.* 7, 44–49 (2007).
103. Carreon, L. Y. *et al.* RhBMP-2 versus iliac crest bone graft for lumbar spine fusion in patients over 60 years of age: a cost-utility study. *Spine* 34, 238–243 (2009).
104. Vaccaro, A. R. *et al.* Comparison of OP-1 Putty (rhBMP-7) to iliac crest autograft for posterolateral lumbar arthrodesis: a minimum 2-year follow-up pilot study. *Spine* 30, 2709–2716 (2005).
105. Rengachary, S. S. Bone morphogenetic proteins: basic concepts. *Neurosurg. Focus* 13, e2 (2002).
106. Govender, S. *et al.* Recombinant human bone morphogenetic protein-2 for treatment of open tibial fractures: a prospective, controlled, randomized study of four hundred and fifty patients. *J. Bone Joint Surg. Am.* 84-A, 2123–2134 (2002).
107. Swiontkowski, M. F. *et al.* Recombinant human bone morphogenetic protein-2 in open tibial fractures. A subgroup analysis of data combined from two prospective randomized studies. *J. Bone Joint Surg. Am.* 88, 1258–1265 (2006).

108. Alt, V. *et al.* A health economic analysis of the use of rhBMP-2 in Gustilo-Anderson grade III open tibial fractures for the UK, Germany, and France. *Injury* 40, 1269–1275 (2009).
109. Jones, A. L. *et al.* Recombinant human BMP-2 and allograft compared with autogenous bone graft for reconstruction of diaphyseal tibial fractures with cortical defects. A randomized, controlled trial. *J. Bone Joint Surg. Am.* 88, 1431–1441 (2006).
110. Friedlaender, G. E. *et al.* Osteogenic protein-1 (bone morphogenetic protein-7) in the treatment of tibial nonunions. *J. Bone Joint Surg. Am.* 83-A Suppl 1, S151–158 (2001).
111. Calori, G. M. *et al.* Application of rhBMP-7 and platelet-rich plasma in the treatment of long bone non-unions: a prospective randomised clinical study on 120 patients. *Injury* 39, 1391–1402 (2008).
112. Rabie, A. B. & Chay, S. H. Clinical applications of composite intramembranous bone grafts. *Am. J. Orthod. Dentofac. Orthop. Off. Publ. Am. Assoc. Orthod. Its Const. Soc. Am. Board Orthod.* 117, 375–383 (2000).
113. Tasoulis, G., Yao, S. G. & Fine, J. B. The maxillary sinus: challenges and treatments for implant placement. *Compend. Contin. Educ. Dent. Jamesburg NJ 1995* 32, 10–14, 16, 18–19; quiz 20, 34 (2011).
114. Wikesjö, U. M. E. *et al.* Bone morphogenetic proteins for periodontal and alveolar indications; biological observations - clinical implications. *Orthod. Craniofac. Res.* 12, 263–270 (2009).
115. Sykaras, N., Triplett, R. G., Nunn, M. E., Iacopino, A. M. & Opperman, L. A. Effect of recombinant human bone morphogenetic protein-2 on bone regeneration and osseointegration of dental implants. *Clin. Oral Implants Res.* 12, 339–349 (2001).
116. Boyne, P. J. Application of bone morphogenetic proteins in the treatment of clinical oral and maxillofacial osseous defects. *J. Bone Joint Surg. Am.* 83-A Suppl 1, S146–150 (2001).
117. Boyne, P. J. *et al.* De novo bone induction by recombinant human bone morphogenetic protein-2 (rhBMP-2) in maxillary sinus floor augmentation. *J. Oral Maxillofac. Surg. Off. J. Am. Assoc. Oral Maxillofac. Surg.* 63, 1693–1707 (2005).
118. Triplett, R. G. *et al.* Pivotal, randomized, parallel evaluation of recombinant human bone morphogenetic protein-2/absorbable collagen sponge and autogenous bone graft for maxillary sinus floor augmentation. *J. Oral Maxillofac. Surg. Off. J. Am. Assoc. Oral Maxillofac. Surg.* 67, 1947–1960 (2009).
119. Fiorellini, J. P. *et al.* Randomized study evaluating recombinant human bone morphogenetic protein-2 for extraction socket augmentation. *J. Periodontol.* 76, 605–613 (2005).
120. Wikesjö, U. M., Sorensen, R. G. & Wozney, J. M. Augmentation of alveolar bone and dental implant osseointegration: clinical implications of studies with rhBMP-2. *J. Bone Joint Surg. Am.* 83-A Suppl 1, S136–145 (2001).
121. Alberts, B. *et al. Molecular Biology of the Cell*. (Garland Science, 2002).
122. Watson, J. D. Involvement of RNA in the synthesis of proteins. *Science* 140, 17–26 (1963).
123. Brimacombe, R. RNA-protein interactions in the Escherichia coli ribosome. *Biochimie* 73, 927–936 (1991).
124. Nyathi, Y., Wilkinson, B. M. & Pool, M. R. Co-translational targeting and translocation of proteins to the endoplasmic reticulum. *Biochim. Biophys. Acta* 1833, 2392–

2402 (2013).
125. Hong, W. & Tang, B. L. Protein trafficking along the exocytotic pathway. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 15, 231–238 (1993).
126. Palade, G. Intracellular aspects of the process of protein synthesis. *Science* 189, 347–358 (1975).
127. Rothman, J. E. & Orci, L. Molecular dissection of the secretory pathway. *Nature* 355, 409–415 (1992).
128. Kornfeld, R. & Kornfeld, S. Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* 54, 631–664 (1985).
129. Klausner, R. D., Donaldson, J. G. & Lippincott-Schwartz, J. Brefeldin A: insights into the control of membrane traffic and organelle structure. *J. Cell Biol.* 116, 1071–1080 (1992).
130. Mellman, I. & Simons, K. The Golgi complex: in vitro veritas? *Cell* 68, 829–840 (1992).
131. Saraste, J. & Svensson, K. Distribution of the intermediate elements operating in ER to Golgi transport. *J. Cell Sci.* 100 ( Pt 3), 415–430 (1991).
132. Griffiths, G. & Simons, K. The trans Golgi network: sorting at the exit site of the Golgi complex. *Science* 234, 438–443 (1986).
133. Mostov, K., Apodaca, G., Aroeti, B. & Okamoto, C. Plasma membrane protein sorting in polarized epithelial cells. *J. Cell Biol.* 116, 577–583 (1992).
134. Traub, L. M. & Kornfeld, S. The trans-Golgi network: a late secretory sorting station. *Curr. Opin. Cell Biol.* 9, 527–533 (1997).
135. Burgess, T. L. & Kelly, R. B. Constitutive and regulated secretion of proteins. *Annu. Rev. Cell Biol.* 3, 243–293 (1987).
136. Riley, E. H., Lane, J. M., Urist, M. R., Lyons, K. M. & Lieberman, J. R. Bone morphogenetic protein-2: biology and applications. *Clin. Orthop.* 39–46 (1996).
137. Ebendal, T., Bengtsson, H. & Söderström, S. Bone morphogenetic proteins and their receptors: potential functions in the brain. *J. Neurosci. Res.* 51, 139–146 (1998).
138. Brunner, A. M., Marquardt, H., Malacko, A. R., Lioubin, M. N. & Purchio, A. F. Site-directed mutagenesis of cysteine residues in the pro region of the transforming growth factor beta 1 precursor. Expression and characterization of mutant proteins. *J. Biol. Chem.* 264, 13660–13664 (1989).
139. Azari, K., Doll, B. A., Sfeir, C., Mu, Y. & Hollinger, J. O. Therapeutic potential of bone morphogenetic proteins. *Expert Opin. Investig. Drugs* 10, 1677–1686 (2001).
140. Israel, D. I. *et al.* Heterodimeric bone morphogenetic proteins show enhanced activity in vitro and in vivo. *Growth Factors Chur Switz.* 13, 291–300 (1996).
141. Sampath, T. K. *et al.* Bovine osteogenic protein is composed of dimers of OP-1 and BMP-2A, two members of the transforming growth factor-beta superfamily. *J. Biol. Chem.* 265, 13198–13205 (1990).
142. Celeste, A. J. *et al.* Identification of transforming growth factor beta family members present in bone-inductive protein purified from bovine bone. *Proc. Natl. Acad. Sci. U. S. A.* 87, 9843–9847 (1990).
143. Schmoekel, H. *et al.* Bone healing in the rat and dog with nonglycosylated BMP-2 demonstrating low solubility in fibrin matrices. *J. Orthop. Res. Off. Publ. Orthop. Res. Soc.* 22, 376–381 (2004).
144. Cui, Y., Jean, F., Thomas, G. & Christian, J. L. BMP-4 is proteolytically activated by

furin and/or PC6 during vertebrate embryonic development. *EMBO J.* 17, 4735–4743 (1998).
145. Jones, C. M., Armes, N. & Smith, J. C. Signalling by TGF-beta family members: short-range effects of Xnr-2 and BMP-4 contrast with the long-range effects of activin. *Curr. Biol. CB* 6, 1468–1475 (1996).
146. Seidah, N. G. & Chrétien, M. Eukaryotic protein processing: endoproteolysis of precursor proteins. *Curr. Opin. Biotechnol.* 8, 602–607 (1997).
147. Steiner, D. F. The proprotein convertases. *Curr. Opin. Chem. Biol.* 2, 31–39 (1998).
148. Gregory, K. E. *et al.* The prodomain of BMP-7 targets the BMP-7 complex to the extracellular matrix. *J. Biol. Chem.* 280, 27970–27980 (2005).
149. Cui, Y. *et al.* The activity and signaling range of mature BMP-4 is regulated by sequential cleavage at two sites within the prodomain of the precursor. *Genes Dev.* 15, 2797–2802 (2001).
150. Heng, S. *et al.* Posttranslational activation of bone morphogenetic protein 2 is mediated by proprotein convertase 6 during decidualization for pregnancy establishment. *Endocrinology* 151, 3909–3917 (2010).
151. Wakefield, L. M., Smith, D. M., Flanders, K. C. & Sporn, M. B. Latent transforming growth factor-beta from human platelets. A high molecular weight complex containing precursor sequences. *J. Biol. Chem.* 263, 7646–7654 (1988).
152. Liao, W. X., Moore, R. K., Otsuka, F. & Shimasaki, S. Effect of intracellular interactions on the processing and secretion of bone morphogenetic protein-15 (BMP-15) and growth and differentiation factor-9. Implication of the aberrant ovarian phenotype of BMP-15 mutant sheep. *J. Biol. Chem.* 278, 3713–3719 (2003).
153. Hosaka, M. *et al.* Arg-X-Lys/Arg-Arg motif as a signal for precursor cleavage catalyzed by furin within the constitutive secretory pathway. *J. Biol. Chem.* 266, 12127–12130 (1991).
154. Leader, B., Baca, Q. J. & Golan, D. E. Protein therapeutics: a summary and pharmacological classification. *Nat. Rev. Drug Discov.* 7, 21–39 (2008).
155. Woodcock, J. Testimony - Follow-on Protein Products. *http://www.fda.gov* (2007). at <http://www.fda.gov/newsevents/testimony/ucm154070.htm>
156. Reichert, J. M. Trends in development and approval times for new therapeutics in the United States. *Nat. Rev. Drug Discov.* 2, 695–702 (2003).
157. Palomares, L. A., Estrada-Moncada, S. & Ramírez, O. T. in *Recombinant Gene Expression* (eds. Balbás, P. & Lorence, A.) 15–51 (Humana Press, 2004). at <http://link.springer.com/protocol/10.1385/1-59259-774-2%3A015>
158. Ozgur, A. & Tutar, Y. Therapeutic proteins: A to Z. *Protein Pept. Lett.* 20, 1365–1372 (2013).
159. Birnbaum, S. & Bailey, J. E. Plasmid presence changes the relative levels of many host cell proteins and ribosome components in recombinant Escherichia coli. *Biotechnol. Bioeng.* 37, 736–745 (1991).
160. O'Kennedy, R., Houghton, C. J. & Patching, J. W. Effects of growth environment on recombinant plasmid stability in Saccharomyces cerevisiae grown in continuous culture. *Appl. Microbiol. Biotechnol.* 44, 126–132 (1995).
161. Paulsson, J. & Ehrenberg, M. Noise in a minimal regulatory network: plasmid copy number control. *Q. Rev. Biophys.* 34, 1–59 (2001).
162. Corchero, J. L. & Villaverde, A. Plasmid maintenance in Escherichia coli recombinant

cultures is dramatically, steadily, and specifically influenced by features of the encoded proteins. *Biotechnol. Bioeng.* 58, 625–632 (1998).
163. Summers, D. Timing, self-control and a sense of direction are the secrets of multicopy plasmid stability. *Mol. Microbiol.* 29, 1137–1145 (1998).
164. Baneyx, F. Recombinant protein expression in Escherichia coli. *Curr. Opin. Biotechnol.* 10, 411–421 (1999).
165. Swartz, J. R. Advances in Escherichia coli production of therapeutic proteins. *Curr. Opin. Biotechnol.* 12, 195–201 (2001).
166. Twyman, R. M. & Whitelaw, B. in *Encyclopedia of Cell Technology* (John Wiley & Sons, Inc., 2003). at <http://onlinelibrary.wiley.com/doi/10.1002/0471250570.spi066/abstract>
167. Balbás, P. Understanding the art of producing protein and nonprotein molecules in Escherichia coli. *Mol. Biotechnol.* 19, 251–267 (2001).
168. Olson, P. *et al.* High-level expression of eukaryotic polypeptides from bacterial chromosomes. *Protein Expr. Purif.* 14, 160–166 (1998).
169. Petricevich, V. L., Palomares, L. A., González, M. & Ramírez, O. T. Parameters that determine virus adsorption kinetics: toward the design of better infection strategies for the insect cell - baculovirus expression system. *Enzyme Microb. Technol.* 29, 52–61 (2001).
170. Chen, W. *et al.* Expression and purification of human interleukin-1beta converting enzyme from Trichoplusia ni insect cells using a baculovirus expression system. *Protein Expr. Purif.* 9, 69–75 (1997).
171. Ozturk, S. & Hu, W.-S. *Cell Culture Technology for Pharmaceutical and Cell-Based Therapies*. (CRC Press, 2014).
172. Zhang, j. in *Manual of Industrial Microbiology and Biotechnology.* 145–156 (Baltz Richard H, editor, 2010).
173. Kopito, R. R. Aggresomes, inclusion bodies and protein aggregation. *Trends Cell Biol.* 10, 524–530 (2000).
174. Carrió, M. M. & Villaverde, A. Construction and deconstruction of bacterial inclusion bodies. *J. Biotechnol.* 96, 3–12 (2002).
175. Schlieker, C., Bukau, B. & Mogk, A. Prevention and reversion of protein aggregation by molecular chaperones in the E. coli cytosol: implications for their applicability in biotechnology. *J. Biotechnol.* 96, 13–21 (2002).
176. Clark, E. D. Protein refolding for industrial processes. *Curr. Opin. Biotechnol.* 12, 202–207 (2001).
177. Datar, R. V., Cartwright, T. & Rosen, C. G. Process economics of animal cell and bacterial fermentations: a case study analysis of tissue plasminogen activator. *Biotechnol. Nat. Publ. Co.* 11, 349–357 (1993).
178. Bach, H. *et al.* Escherichia coli maltose-binding protein as a molecular chaperone for recombinant intracellular cytoplasmic single-chain antibodies. *J. Mol. Biol.* 312, 79–93 (2001).
179. Paetzel, M., Karla, A., Strynadka, N. C. J. & Dalbey, R. E. Signal peptidases. *Chem. Rev.* 102, 4549–4580 (2002).
180. Van Dijl, J. M., de Jong, A., Smith, H., Bron, S. & Venema, G. Signal peptidase I overproduction results in increased efficiencies of export and maturation of hybrid secretory proteins in Escherichia coli. *Mol. Gen. Genet. MGG* 227, 40–48 (1991).

181. Bryan, P. N. Prodomains and protein folding catalysis. *Chem. Rev.* 102, 4805–4816 (2002).
182. Nakayama, K. Furin: a mammalian subtilisin/Kex2p-like endoprotease involved in processing of a wide variety of precursor proteins. *Biochem. J.* 327 ( Pt 3), 625–635 (1997).
183. Laprise, M. H., Grondin, F. & Dubois, C. M. Enhanced TGFbeta1 maturation in high five cells coinfected with recombinant baculovirus encoding the convertase furin/pace: improved technology for the production of recombinant proproteins in insect cells. *Biotechnol. Bioeng.* 58, 85–91 (1998).
184. Gu, X., Harmon, B. J. & Wang, D. I. C. in *Advances in Bioprocess Engineering* (eds. Galindo, E. & Ramírez, O. T.) 1–24 (Springer Netherlands, 1998). at <http://link.springer.com/chapter/10.1007/978-94-017-0643-8_1>
185. Shelikoff, M., Sinskey, A. J. & Stephanopoulos, G. The effect of protein synthesis inhibitors on the glycosylation site occupancy of recombinant human prolactin. *Cytotechnology* 15, 195–208 (1994).
186. Lisowska, E. The role of glycosylation in protein antigenic properties. *Cell. Mol. Life Sci. CMLS* 59, 445–455 (2002).
187. Chojnacki, T. & Dallner, G. The biological role of dolichol. *Biochem. J.* 251, 1–9 (1988).
188. Yuk, I. H. Y. & Wang, D. I. C. Glycosylation by Chinese hamster ovary cells in dolichol phosphate-supplemented cultures. *Biotechnol. Appl. Biochem.* 36, 141–147 (2002).
189. Rearick, J. I., Chapman, A. & Kornfeld, S. Glucose starvation alters lipid-linked oligosaccharide biosynthesis in Chinese hamster ovary cells. *J. Biol. Chem.* 256, 6255–6261 (1981).
190. Gu, X. & Wang, D. I. Improvement of interferon-gamma sialylation in Chinese hamster ovary cell culture by feeding of N-acetylmannosamine. *Biotechnol. Bioeng.* 58, 642–648 (1998).
191. Andersen, D. C., Bridges, T., Gawlitzek, M. & Hoy, C. Multiple cell culture factors can affect the glycosylation of Asn-184 in CHO-produced tissue-type plasminogen activator. *Biotechnol. Bioeng.* 70, 25–31 (2000).
192. Tokmakov, A. A. *et al.* Multiple post-translational modifications affect heterologous protein synthesis. *J. Biol. Chem.* 287, 27106–27116 (2012).
193. Ma, W. F., Chen, H. Y., Du, J., Tan, Y. & Cai, S. H. A novel recombinant protein TAT-GFP-KDEL with dual-function of penetrating cell membrane and locating at endoplasm reticulum. *J. Drug Target.* 17, 329–333 (2009).
194. Wingfield, P. T. Overview of the purification of recombinant proteins produced in Escherichia coli. *Curr. Protoc. Protein Sci. Editor. Board John E Coligan Al* Chapter 6, Unit 6.1 (2003).
195. Gurramkonda, C. *et al.* Application of simple fed-batch technique to high-level secretory production of insulin precursor using Pichia pastoris with subsequent purification and conversion to human insulin. *Microb. Cell Factories* 9, 31 (2010).
196. Matasci, M., Hacker, D. L., Baldi, L. & Wurm, F. M. Recombinant therapeutic protein production in cultivated mammalian cells: current status and future prospects. *Drug Discov. Today Technol.* 5, e37–42 (2008).
197. Ruffoni, B., Pistelli, L., Bertoli, A. & Pistelli, L. Plant cell cultures: bioreactors for industrial production. *Adv. Exp. Med. Biol.* 698, 203–221 (2010).

198. Redwan, el-R. M. Animal-derived pharmaceutical proteins. *J. Immunoassay Immunochem.* 30, 262–290 (2009).
199. Pogue, G. P. *et al.* Production of pharmaceutical-grade recombinant aprotinin and a monoclonal antibody product using plant-based transient expression systems. *Plant Biotechnol. J.* 8, 638–654 (2010).
200. Casteleijn, M. G., Urtti, A. & Sarkhel, S. Expression without boundaries: cell-free protein synthesis in pharmaceutical research. *Int. J. Pharm.* 440, 39–47 (2013).
201. Assenberg, R., Wan, P. T., Geisse, S. & Mayr, L. M. Advances in recombinant protein expression for use in pharmaceutical research. *Curr. Opin. Struct. Biol.* 23, 393–402 (2013).
202. Huang, C.-J., Lin, H. & Yang, X. Industrial production of recombinant therapeutics in Escherichia coli and its recent advancements. *J. Ind. Microbiol. Biotechnol.* 39, 383–399 (2012).
203. Kamionka, M. Engineering of therapeutic proteins production in Escherichia coli. *Curr. Pharm. Biotechnol.* 12, 268–274 (2011).
204. Harwood, C. R. Bacillus subtilis and its relatives: molecular biological and industrial workhorses. *Trends Biotechnol.* 10, 247–256 (1992).
205. Sánchez, M., Prim, N., Rández-Gil, F., Pastor, F. I. J. & Diaz, P. Engineering of baker's yeasts, E. coli and Bacillus hosts for the production of Bacillus subtilis Lipase A. *Biotechnol. Bioeng.* 78, 339–345 (2002).
206. Kretzmer, G. Industrial processes with animal cells. *Appl. Microbiol. Biotechnol.* 59, 135–142 (2002).
207. Zhu, J. Mammalian cell protein expression for biopharmaceutical production. *Biotechnol. Adv.* 30, 1158–1170 (2012).
208. Jenkins, N., Parekh, R. B. & James, D. C. Getting the glycosylation right: implications for the biotechnology industry. *Nat. Biotechnol.* 14, 975–981 (1996).
209. Fussenegger, M. & Betenbaugh, M. J. Metabolic engineering II. Eukaryotic systems. *Biotechnol. Bioeng.* 79, 509–531 (2002).
210. Genzel, Y., Rödig, J., Rapp, E. & Reichl, U. Vaccine production: upstream processing with adherent or suspension cell lines. *Methods Mol. Biol. Clifton NJ* 1104, 371–393 (2014).
211. Colosimo, A. *et al.* Transfer and expression of foreign genes in mammalian cells. *BioTechniques* 29, 314–318, 320–322, 324 passim (2000).
212. Fussenegger, M. & Bailey, J. E. Molecular regulation of cell-cycle progression and apoptosis in mammalian cells: implications for biotechnology. *Biotechnol. Prog.* 14, 807–833 (1998).
213. Underhill, M. F. *et al.* Engineering mRNA translation initiation to enhance transient gene expression in chinese hamster ovary cells. *Biotechnol. Prog.* 19, 121–129 (2003).
214. Kjeldsen, T. Yeast secretory expression of insulin precursors. *Appl. Microbiol. Biotechnol.* 54, 277–286 (2000).
215. Kilonzo, P. M., Margaritis, A. & Bergougnou, M. A. Plasmid stability and kinetics of continuous production of glucoamylase by recombinant Saccharomyces cerevisiae in an airlift bioreactor. *J. Ind. Microbiol. Biotechnol.* 36, 1157–1169 (2009).
216. Gu, M. B., Park, M. H. & Kim, D. I. Growth rate control in fed-batch cultures of recombinant Saccharomyces cerevisiae producing hepatitis B surface antigen (HBsAg). *Appl. Microbiol. Biotechnol.* 35, 46–30 (1991).
217. Cereghino, G. P. L., Cereghino, J. L., Ilgen, C. & Cregg, J. M. Production of

recombinant proteins in fermenter cultures of the yeast Pichia pastoris. *Curr. Opin. Biotechnol.* 13, 329–332 (2002).
218. Chiba, Y. *et al.* Production of human compatible high mannose-type (Man5GlcNAc2) sugar chains in Saccharomyces cerevisiae. *J. Biol. Chem.* 273, 26298–26304 (1998).
219. Gellissen, G. Heterologous protein production in methylotrophic yeasts. *Appl. Microbiol. Biotechnol.* 54, 741–750 (2000).
220. Punt, P. J. *et al.* Filamentous fungi as cell factories for heterologous protein production. *Trends Biotechnol.* 20, 200–206 (2002).
221. Gyamerah, M., Merichetti, G., Adedayo, O., Scharer, J. M. & Moo-Young, M. Bioprocessing strategies for improving hen egg-white lysozyme (HEWL) production by recombinant Aspergillus niger HEWL WT-13-16. *Appl. Microbiol. Biotechnol.* 60, 403–407 (2002).
222. Maras, M. *et al.* In vivo synthesis of complex N-glycans by expression of human N-acetylglucosaminyltransferase I in the filamentous fungus Trichoderma reesei. *FEBS Lett.* 452, 365–370 (1999).
223. Pollock, D. P. *et al.* Transgenic milk as a method for the production of recombinant antibodies. *J. Immunol. Methods* 231, 147–157 (1999).
224. Lai, L. *et al.* Production of alpha-1,3-galactosyltransferase knockout pigs by nuclear transfer cloning. *Science* 295, 1089–1092 (2002).
225. Kues, W. A. & Niemann, H. Advances in farm animal transgenesis. *Prev. Vet. Med.* 102, 146–156 (2011).
226. Dyck, M. K., Lacroix, D., Pothier, F. & Sirard, M.-A. Making recombinant proteins in animals--different systems, different applications. *Trends Biotechnol.* 21, 394–399 (2003).
227. Niemann, H. & Kues, W. A. Transgenic farm animals: an update. *Reprod. Fertil. Dev.* 19, 762–770 (2007).
228. Kuroiwa, Y. *et al.* Antigen-specific human polyclonal antibodies from hyperimmunized cattle. *Nat. Biotechnol.* 27, 173–181 (2009).
229. Lillico, S. G. *et al.* Oviduct-specific expression of two therapeutic proteins in transgenic hens. *Proc. Natl. Acad. Sci. U. S. A.* 104, 1771–1776 (2007).
230. Ma, J. K.-C., Drake, P. M. W. & Christou, P. The production of recombinant pharmaceutical proteins in plants. *Nat. Rev. Genet.* 4, 794–805 (2003).
231. Vitale, A. & Pedrazzini, E. Recombinant pharmaceuticals from plants: the plant endomembrane system as bioreactor. *Mol. Interv.* 5, 216–225 (2005).
232. Horn, M. E., Woodard, S. L. & Howard, J. A. Plant molecular farming: systems and products. *Plant Cell Rep.* 22, 711–720 (2004).
233. Urist, M. R. *et al.* Preparation and bioassay of bone morphogenetic protein and polypeptide fragments. *Methods Enzymol.* 146, 294–312 (1987).
234. Hu, Z. M., Peel, S. A. F., Sandor, G. K. B. & Clokie, C. M. L. The osteoinductive activity of bone morphogenetic protein (BMP) purified by repeated extracts of bovine bone. *Growth Factors Chur Switz.* 22, 29–33 (2004).
235. Kirker-Head, C. A. Potential applications and delivery strategies for bone morphogenetic proteins. *Adv. Drug Deliv. Rev.* 43, 65–92 (2000).
236. Colin, C. *et al.* NUCEL (Cell and Molecular Therapy Center): a multidisciplinary center for translational research in Brazil. *Mol. Biotechnol.* 39, 89–95 (2008).
237. Maruoka, Y. *et al.* Production of functional human bone morphogenetic protein-2

using a baculovirus/Sf-9 insect cell system. *Biochem. Mol. Biol. Int.* 35, 957–963 (1995).
238. Vallejo, L. F. *et al.* Renaturation and purification of bone morphogenetic protein-2 produced as inclusion bodies in high-cell-density cultures of recombinant Escherichia coli. *J. Biotechnol.* 94, 185–194 (2002).
239. Carreira, A. C. *et al.* Bone morphogenetic proteins: facts, challenges, and future perspectives. *J. Dent. Res.* 93, 335–345 (2014).
240. Lee, J.-H. *et al.* The induction of bone formation in rat calvarial defects and subcutaneous tissues by recombinant human BMP-2, produced in Escherichia coli. *Biomaterials* 31, 3512–3519 (2010).
241. Kim, M.-S. *et al.* Prospective randomized, controlled trial of sinus grafting using Escherichia-coli-produced rhBMP-2 with a biphasic calcium phosphate carrier compared to deproteinized bovine bone. *Clin. Oral Implants Res.* (2014). doi:10.1111/clr.12471
242. Andersen, D. C. & Krummen, L. Recombinant protein expression for therapeutic applications. *Curr. Opin. Biotechnol.* 13, 117–123 (2002).
243. Wurm, F. M. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat. Biotechnol.* 22, 1393–1398 (2004).
244. Hammonds, G. R. & Mason, A. J. DNA constructs, mammalian expression vectors and host cells containing same. at <http://www.google.com/patents/US5168050>
245. Wang, E. A., Wozney, J. M. & Rosen, V. A. Proteins; bone disorders. (1991). at <http://www.google.com/patents/US5013649>
246. Barnes, L. M. & Dickson, A. J. Mammalian cell factories for efficient and stable protein expression. *Curr. Opin. Biotechnol.* 17, 381–386 (2006).
247. Meleady, P. *et al.* Proteomic profiling of CHO cells with enhanced rhBMP-2 productivity following co-expression of PACEsol. *Proteomics* 8, 2611–2624 (2008).
248. Giddings, G., Allison, G., Brooks, D. & Carter, A. Transgenic plants as factories for biopharmaceuticals. *Nat. Biotechnol.* 18, 1151–1155 (2000).
249. Fischer, R., Stoger, E., Schillberg, S., Christou, P. & Twyman, R. M. Plant-based production of biopharmaceuticals. *Curr. Opin. Plant Biol.* 7, 152–158 (2004).
250. Stein, K. E. & Webber, K. O. The regulation of biologic products derived from bioengineered plants. *Curr. Opin. Biotechnol.* 12, 308–311 (2001).
251. Fischer, R. & Emans, N. Molecular farming of pharmaceutical proteins. *Transgenic Res.* 9, 279–299; discussion 277 (2000).
252. Barbante, A. *et al.* Anchorage to the cytosolic face of the endoplasmic reticulum membrane: a new strategy to stabilize a cytosolic recombinant antigen in plants. *Plant Biotechnol. J.* 6, 560–575 (2008).
253. Vitale, null & Denecke, null. The endoplasmic reticulum-gateway of the secretory pathway. *Plant Cell* 11, 615–628 (1999).
254. Rutishauser, J. & Spiess, M. Endoplasmic reticulum storage diseases. *Swiss Med. Wkly.* 132, 211–222 (2002).
255. Wandelt, C. I. *et al.* Vicilin with carboxy-terminal KDEL is retained in the endoplasmic reticulum and accumulates to high levels in the leaves of transgenic plants. *Plant J. Cell Mol. Biol.* 2, 181–192 (1992).
256. Herman, null & Larkins, null. Protein storage bodies and vacuoles. *Plant Cell* 11, 601–614 (1999).
257. Shewry, P. R. & Halford, N. G. Cereal seed storage proteins: structures, properties and

role in grain utilization. *J. Exp. Bot.* 53, 947–958 (2002).
258. Mainieri, D. *et al.* Zeolin. A new recombinant storage protein constructed using maize gamma-zein and bean phaseolin. *Plant Physiol.* 136, 3447–3456 (2004).
259. Xu, J.-H. & Messing, J. Organization of the prolamin gene family provides insight into the evolution of the maize genome and gene duplications in grass species. *Proc. Natl. Acad. Sci. U. S. A.* 105, 14330–14335 (2008).
260. Prat, S., Cortadas, J., Puigdomènech, P. & Palau, J. Nucleic acid (cDNA) and amino acid sequences of the maize endosperm protein glutelin-2. *Nucleic Acids Res.* 13, 1493–1504 (1985).
261. Mainieri, D. *et al.* Protein body formation in the endoplasmic reticulum as an evolution of storage protein sorting to vacuoles: insights from maize γ-zein. *Front. Plant Sci.* 5, 331 (2014).
262. Shewry, P. R., Napier, J. A. & Tatham, A. S. Seed storage proteins: structures and biosynthesis. *Plant Cell* 7, 945–956 (1995).
263. Altschuler, Y. & Galili, G. Role of conserved cysteines of a wheat gliadin in its transport and assembly into protein bodies in Xenopus oocytes. *J. Biol. Chem.* 269, 6677–6682 (1994).
264. Geli, M. I., Torrent, M. & Ludevid, D. Two Structural Domains Mediate Two Sequential Events in [gamma]-Zein Targeting: Protein Endoplasmic Reticulum Retention and Protein Body Formation. *Plant Cell* 6, 1911–1922 (1994).
265. Llop-Tous, I. *et al.* Relevant elements of a maize gamma-zein domain involved in protein body biogenesis. *J. Biol. Chem.* 285, 35633–35644 (2010).
266. Kogan, M. J. *et al.* Supramolecular properties of the proline-rich gamma-Zein N-terminal domain. *Biophys. J.* 83, 1194–1204 (2002).
267. Torrent, M. *et al.* Eukaryotic protein production in designed storage organelles. *BMC Biol.* 7, 5 (2009).
268. Torrent, M., Llop-Tous, I. & Ludevid, M. D. Protein body induction: a new tool to produce and recover recombinant proteins in plants. *Methods Mol. Biol. Clifton NJ* 483, 193–208 (2009).
269. Conley, A. J., Joensuu, J. J., Richman, A. & Menassa, R. Protein body-inducing fusions for high-level production and purification of recombinant proteins in plants. *Plant Biotechnol. J.* 9, 419–433 (2011).
270. Hefferon, K. L. *Plant-derived Pharmaceuticals: Principles and Applications for Developing Countries*. (CABI, 2014).
271. Pedrazzini, E. *et al.* Protein quality control along the route to the plant vacuole. *Plant Cell* 9, 1869–1880 (1997).
272. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254 (1976).
273. Mullis, K. B. & Faloona, F. A. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* 155, 335–350 (1987).
274. Hellens, R. P., Edwards, E. A., Leyland, N. R., Bean, S. & Mullineaux, P. M. pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. *Plant Mol. Biol.* 42, 819–832 (2000).
275. Pietrzak, M., Shillito, R. D., Hohn, T. & Potrykus, I. Expression in plants of two

bacterial antibiotic resistance genes after protoplast transformation with a new plant expression vector. *Nucleic Acids Res.* 14, 5857–5868 (1986).
276. Lyon, B. R., Llewellyn, D. J., Huppatz, J. L., Dennis, E. S. & Peacock, W. J. Expression of a bacterial gene in transgenic tobacco plants confers resistance to the herbicide 2,4-dichlorophenoxyacetic acid. *Plant Mol. Biol.* 13, 533–540 (1989).
277. Jobling, S. A. & Gehrke, L. Enhanced translation of chimaeric messenger RNAs containing a plant viral untranslated leader sequence. *Nature* 325, 622–625 (1987).
278. Tabe, L. M. *et al.* A biotechnological approach to improving the nutritive value of alfalfa. *J. Anim. Sci.* 73, 2752–2759 (1995).
279. Alting-Mees, M. A. & Short, J. M. pBluescript II: gene mapping vectors. *Nucleic Acids Res.* 17, 9494 (1989).
280. Petrescu, A.-J., Milac, A.-L., Petrescu, S. M., Dwek, R. A. & Wormald, M. R. Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology* 14, 103–114 (2004).
281. Tai, T. *et al.* Structural studies of two ovalbumin glycopeptides in relation to the endo-beta-N-acetylglucosaminidase specificity. *J. Biol. Chem.* 250, 8569–8575 (1975).
282. Vitale, A. & Ceriotti, A. Protein quality control mechanisms and protein storage in the endoplasmic reticulum. A conflict of interests? *Plant Physiol.* 136, 3420–3426 (2004).
283. James, V. A., Avart, C., Worland, B., Snape, J. W. & Vain, P. The relationship between homozygous and hemizygous transgene expression levels over generations in populations of transgenic rice plants. *TAG Theor. Appl. Genet. Theor. Angew. Genet.* 104, 553–561 (2002).
284. Holding, D. R. Recent advances in the study of prolamin storage protein organization and function. *Plant Evol. Dev.* 5, 276 (2014).
285. Kempken, F. & Jung, C. *Genetic Modification of Plants: Agriculture, Horticulture and Forestry*. (Springer Science & Business Media, 2009).
286. McCabe, M. S. *et al.* Plastid transformation of high-biomass tobacco variety Maryland Mammoth for production of human immunodeficiency virus type 1 (HIV-1) p24 antigen. *Plant Biotechnol. J.* 6, 914–929 (2008).
287. De Virgilio, M. *et al.* The human immunodeficiency virus antigen Nef forms protein bodies in leaves of transgenic tobacco when fused to zeolin. *J. Exp. Bot.* 59, 2815–2829 (2008).
288. Zhang, S. *et al.* Comparative characterization of the glycosylation profiles of an influenza hemagglutinin produced in plant and insect hosts. *Proteomics* 12, 1269–1288 (2012).
289. Lerouge, P. *et al.* N-glycoprotein biosynthesis in plants: recent developments and future trends. *Plant Mol. Biol.* 38, 31–48 (1998).
290. Bardor, M. *et al.* Immunoreactivity in mammals of two typical plant glyco-epitopes, core alpha(1,3)-fucose and core xylose. *Glycobiology* 13, 427–434 (2003).
291. Altmann, F. The role of protein glycosylation in allergy. *Int. Arch. Allergy Immunol.* 142, 99–115 (2007).
292. Galili, G. ER-Derived Compartments Are Formed by Highly Regulated Processes and Have Special Functions in Plants. *Plant Physiol.* 136, 3411–3413 (2004).
293. Kinney, A. J., Jung, R. & Herman, E. M. Cosuppression of the alpha subunits of beta-conglycinin in transgenic soybean seeds induces the formation of endoplasmic reticulum-

derived protein bodies. *Plant Cell* 13, 1165–1178 (2001).
294. Joensuu, J. J., Niklander-Teeri, V. & Brandle, J. E. Transgenic plants for animal health: plant-made vaccine antigens for animal infectious disease control. *Phytochem. Rev.* 7, 553–577 (2008).
295. Fernández-San Millán, A., Mingo-Castel, A., Miller, M. & Daniell, H. A chloroplast transgenic approach to hyper-express and purify Human Serum Albumin, a protein highly susceptible to proteolytic degradation. *Plant Biotechnol. J.* 1, 71–79 (2003).
296. Olinger, G. G. *et al.* Delayed treatment of Ebola virus infection with plant-derived monoclonal antibodies provides protection in rhesus macaques. *Proc. Natl. Acad. Sci. U. S. A.* 109, 18030–18035 (2012).
297. Strasser, R. *et al.* Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol. J.* 6, 392–402 (2008).
298. Sack, M., Hofbauer, A., Fischer, R. & Stoger, E. The increasing value of plant-made proteins. *Curr. Opin. Biotechnol.* 32C, 163–170 (2015).
299. Shaaltiel, Y. *et al.* Production of glucocerebrosidase with terminal mannose glycans for enzyme replacement therapy of Gaucher's disease using a plant cell system. *Plant Biotechnol. J.* 5, 579–590 (2007).
300. Zhang, H. *et al.* Improved mass spectrometric characterization of protein glycosylation reveals unusual glycosylation of maize-derived bovine trypsin. *Anal. Chem.* 82, 10095–10101 (2010).
301. Liu, A. & Niswander, L. A. Bone morphogenetic protein signalling and vertebrate nervous system development. *Nat. Rev. Neurosci.* 6, 945–954 (2005).
302. Hegarty, S. V., Sullivan, A. M. & O'Keeffe, G. W. BMP2 and GDF5 induce neuronal differentiation through a Smad dependant pathway in a model of human midbrain dopaminergic neurons. *Mol. Cell. Neurosci.* 56, 263–271 (2013).