

Review

# Increasing the Efficiency of the Accumulation of Recombinant Proteins in Plant Cells: The Role of Transport Signal Peptides

Sergey M. Rozov \* and Elena V. Deineko \*

Federal Research Center Institute of Cytology and Genetics, Siberian Branch of Russian Academy of Sciences, Pr. Akad. Lavrentieva 10, Novosibirsk 630090, Russia

\* Correspondence: rozov@bionet.nsc.ru (S.M.R.); deineko@bionet.nsc.ru (E.V.D.)

**Abstract:** The problem with increasing the yield of recombinant proteins is resolvable using different approaches, including the transport of a target protein to cell compartments with a low protease activity. In the cell, protein targeting involves short-signal peptide sequences recognized by intracellular protein transport systems. The main systems of the protein transport across membranes of the endoplasmic reticulum and endosymbiotic organelles are reviewed here, as are the major types and structure of the signal sequences targeting proteins to the endoplasmic reticulum and its derivatives, to plastids, and to mitochondria. The role of protein targeting to certain cell organelles depending on specific features of recombinant proteins and the effect of this targeting on the protein yield are discussed, in addition to the main directions of the search for signal sequences based on their primary structure. This knowledge makes it possible not only to predict a protein localization in the cell but also to reveal the most efficient sequences with potential biotechnological utility.

**Keywords:** recombinant protein; transport signal peptide; plant expression system



**Citation:** Rozov, S.M.; Deineko, E.V. Increasing the Efficiency of the Accumulation of Recombinant Proteins in Plant Cells: The Role of Transport Signal Peptides. *Plants* **2022**, *11*, 2561. <https://doi.org/10.3390/plants11192561>

Academic Editor: Rene Kizek

Received: 5 September 2022

Accepted: 26 September 2022

Published: 28 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The latest advances in molecular biology and genetic engineering allow for the production of a large number of diverse recombinant proteins for medical applications and vaccines—for example, antibodies and commercial enzymes—using different expression platforms [1–3]. These proteins are synthesized with the help of prokaryotic and eukaryotic expression systems, such as *Escherichia coli*, yeasts, and cultured cells of insects and mammals. Over half of pharmaceutically valuable proteins are produced in mammalian cells [2,3] because these cells provide correct post-translational modifications (PTMs) of the eukaryotic type [4,5]. Incorrect PTMs—for example, in yeasts—or their absence in prokaryotes may worsen the properties of a synthesized recombinant protein—in particular, by altering its biological activity and pharmacokinetics. That is why prokaryotic expression systems are mainly used for synthesizing relatively simple therapeutic proteins. As for more complex proteins, they are usually synthesized in expression systems involving mammalian cell cultures [5]; however, these systems also have their flaws. First and foremost, there is a high cost of cultivation, difficulties with the upscaling of the processes, and a potential virus infection. Plant cell cultures are free of these shortcomings, thereby representing the most promising alternative for the synthesis of recombinant proteins. Plant cells combine the ability to perform PTMs according to the eukaryotic pattern and to rapidly grow, being both simple and inexpensive, similar to bacterial expression systems. In addition, plant systems exclude any potential contamination by animal viruses and bacterial toxins [2,6,7]. These advantages have formed the basis for so-called green pharmaceuticals, in which plants or plant cell cultures act as biofactories producing recombinant proteins [7,8]. Unfortunately, plant expression systems are not free of certain flaws either; the main one is a low yield of recombinant proteins despite all attempts to increase it [9,10]. Several reviews describe the different methods used for enhancing the biosynthesis of transgene-encoded proteins [10–12].

The plant cell, as a representative of eukaryotes, has an extremely intricate pattern of intracellular compartmentalization and contains a complex system of subcellular membranes. When considering the plant cell in terms of an experimental tool for the biosynthesis of recombinant proteins, it should be kept in mind that the final yield of a target recombinant protein is determined not only by the intensity of its biosynthesis in the cytosol but also by the rate of its degradation in a given cell compartment. The cytosol has a high protease activity; moreover, in the cytosol, proteins can undergo undesirable modifications altering their folding [13]. However, all improperly folded proteins are degraded via the ubiquitin–proteasome proteolytic pathway [14], which leads to an even greater drop in the yield of a recombinant protein. Usually, the accumulation of recombinant proteins in the cytosol is rather low [15]; accordingly, this compartment is regarded as unsuitable for the storage of target recombinant proteins. The difference in the number of accumulated recombinant proteins between the cytosol and apoplast may reach several-thousand-fold [16]. Thus, it is reasonable to immediately transport the synthesized proteins to the apoplast or some other cell compartments. Cell organelles contain different types and numbers of proteases, and various recombinant proteins are, to various degrees, prone to degradation by individual proteases; the lowest protease activity is present in the apoplast and plastids [17,18]. Therefore, the selection of a cell compartment for the accumulation of a specific recombinant protein is not subject to simple universal rules.

Most of proteins in the plant cell are synthesized by the ribosomes in the cytosol, including secreted proteins and proteins of cell organelles (mitochondria, the nucleus, and chloroplasts). That is why the cell, during its evolution, has acquired the efficient mechanisms for transporting newly synthesized proteins to suitable cell organelles and even through the endoplasmic reticulum (ER) and Golgi apparatus to vacuoles and/or further to the apoplast. Specialized signal peptide (SP) sequences with a length of 20–80 amino acid residues (aa) are used for the transport of proteins to the target compartments; most frequently, these signal sequences are localized to the N terminus of proteins and, as a rule, are cleaved by specific proteases after reaching the target compartment [19]. For example, over 3000 cytosol-synthesized proteins of different types are targeted to chloroplasts by transit peptides [20,21]. These peptide signal sequences can be successfully employed to target recombinant proteins synthesized in plant cells to the compartments with a low protease activity. For this purpose, researchers use transfection with genetic constructs carrying the gene of a target protein fused in a head-to-tail manner to the sequence coding for the corresponding signal transport peptide [19,22,23]. This subcellular addressing of recombinant proteins is key to the boosting of protein accumulation because the cell compartment housing proteins directly influences their folding, assembly, and PTMs and prevents the direct degradation and interference of the proteins with the cell metabolism, which often takes place in the cytosol. The attachment of SPs for the transport of recombinant proteins to “protected” cell compartments or the apoplast usually guarantees higher levels of their production. In addition, some proteins may be cytotoxic in the cytosol but nontoxic when shielded by the double membrane of the chloroplast [16,24]. The apoplast, ER, and chloroplasts [22,23], as well as storage protein bodies (PBs), characteristic for plant seeds [25], are the most promising compartments for the accumulation of recombinant proteins. The SPs targeting different proteins to the same organelles differ not only in their amino acid sequences but also in the efficiency of delivery; accordingly, the yield of recombinant proteins carrying different SPs can vary considerably [23,26,27]. Thus, the expression efficiency of recombinant proteins in plant systems depends, to a considerable degree, on the selection of their subcellular targeting and the most efficient transport SPs.

## 2. Intracellular Transport of Proteins to the ER and Its Derivatives

Eukaryotes possess two types of ribosomes, namely, heavy 80S ribosomes (residing in the cytosol), lighter 70S ribosomes (residing in the endosymbiotic organelles), mitochondria, and plastids in plants. Nonetheless, 99% of mitochondrial proteins [28] and 95% of plastid proteins [29] are encoded by nuclear genes and synthesized on ribosomes of the cytosol. In

addition, over one-third of all proteins synthesized by the eukaryotic cell have to, in part or in full, start making their way (to a destination) in the ER, where they undergo PTMs, are folded and verified for correct conformation, interact with chaperones, and form complexes of subunits [30–33]. The overall intracellular transport of proteins in the plant cell falls into two major routes, the first one comprising the translocation of proteins to the ER and its derivatives (Golgi complex, vacuoles, PBs, and apoplast). The second route includes the translocation of proteins to the nucleus and endosymbiotic organelles (mitochondria and plastids). Commonly, two types of protein transport to all eukaryotic cell compartments are distinguished, namely, cotranslational and post-translational. In the former, a protein is synthesized directly on the membrane of the ER or an organelle and cotranslationally transferred through the membrane; in the latter, the protein completely synthesized in the cytosol is delivered one way or another to its destination. It is not entirely clear what determines the choice of a pathway.

### 2.1. Cotranslational Protein Targeting to the ER and Its Derivatives

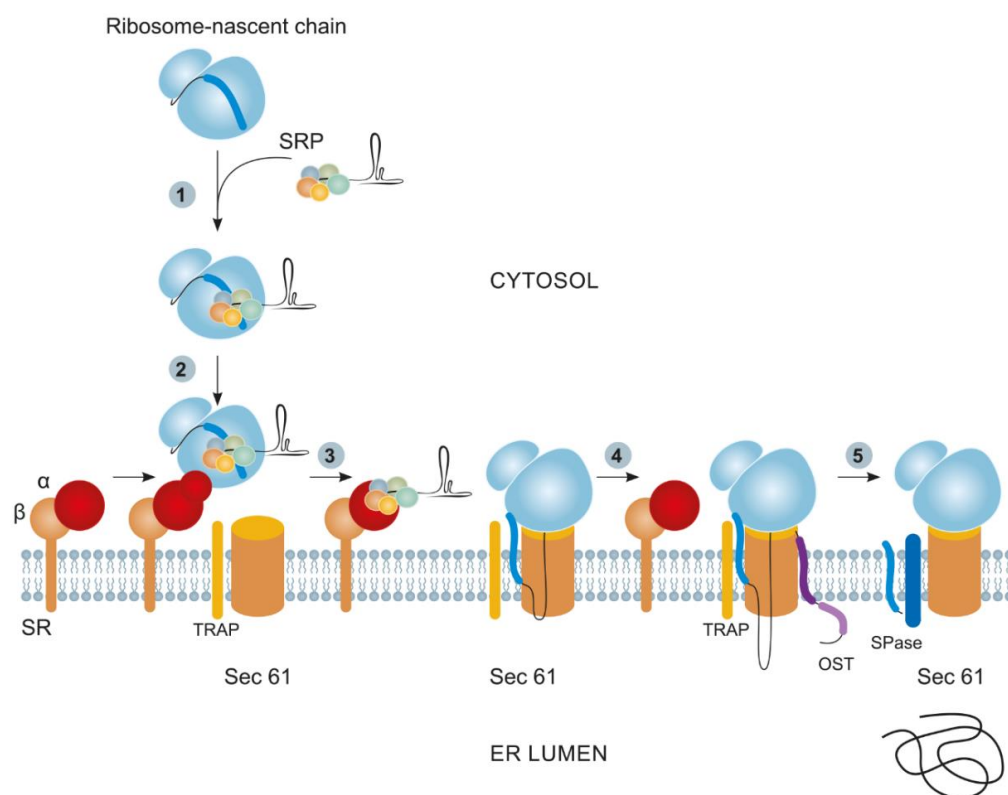
Many proteins targeted to the ER follow the signal recognition particle (SRP) route. The SRP is an RNA–protein complex of 7S RNA and six small proteins [34]. As a rule, proteins targeted via the SRP pathway carry either a hydrophobic SP at their N terminus, which is subsequently cleaved by a specific signal peptidase (SPase) in the ER, or an internal uncleaved SP acting as a signal anchor [35,36]. The cotranslational protein transport commences with the SRP binding to a ribosome with the subsequent scanning of the nascent chain for the presence of signal sequences. As soon as the SP or signal anchor leaves the ribosome, it binds to the SRP and activates it [37]. The activated SRP, together with the ribosome, the mRNA, and a nascent polypeptide chain (ribosome–nascent chain complex, RNC), temporarily arrests translation to perform the GTP-dependent translocation of the entire complex to the cognate SRP receptor on the outer ER side. The interaction with the receptor weakens the association of the SRP with the RNC, and the ribosome with mRNA and the nascent polypeptide chain remains on the outer ER membrane, where it interacts with heterotrimeric translocon Sec61. After this event, the translation restarts, and the synthesized protein is gradually translocated across the Sec61 transmembrane channel into the ER lumen (Figure 1) [35,36,38].

Several additional components involved in the maturation of newly synthesized proteins and the implementation of PTMs and chaperone-like folding functions are associated with the Sec61 translocon. One of these components is the translocon-associated protein complex (TRAP). The TRAP is a heterotetrameric complex of membrane proteins [39], assists Sec61 with protein translocation across the membrane, and controls the correct protein topogenesis in the ER lumen [40]. Another extra component of the Sec61 translocon is the oligosaccharyltransferase (OST) complex, which catalyzes protein N-glycosylation. This is a multisubunit complex comprising at least seven membrane proteins [41].

### 2.2. Post-Translational Targeting to the ER

Unlike cotranslational protein targeting, post-translational targeting commences only after a protein is completely synthesized and leaves the ribosome for the cytosol. Important specific features of post-translationally translocated polypeptides are their relatively small size (not exceeding 100 aa) and the C-terminal position of the transmembrane helix (TMH), acting as a translocation SP or, as in the case of yeasts, a “weak” N-terminal SP required for targeting [42]. These specific features interfere with the effective recognition of these proteins by the cotranslational targeting factor SRP and require the presence of a post-translational targeting factor—for example, GET3 (GET, guided entry of tail-anchored proteins). GET3 is involved in the targeting of tail-anchored (TA) proteins, which carry one region of the TMH (at the C terminus), acting as a guiding signal. In addition, GET3 enhances the targeting of short polypeptides with a cleavable N-terminal SP [43–45]. Two factors—SGTA [46] and heterotrimeric BAG6 [47]—participate in the translocation of a targeted protein to the GET3 complex [48]. Recently, one more SRP-independent targeting

route, SND, was discovered; presumably, it acts as a reserve system for the translocation of certain protein types and overlaps with GET and SRP routes [49,50]. Many cytosol chaperones enhance post-translational protein targeting [51,52].



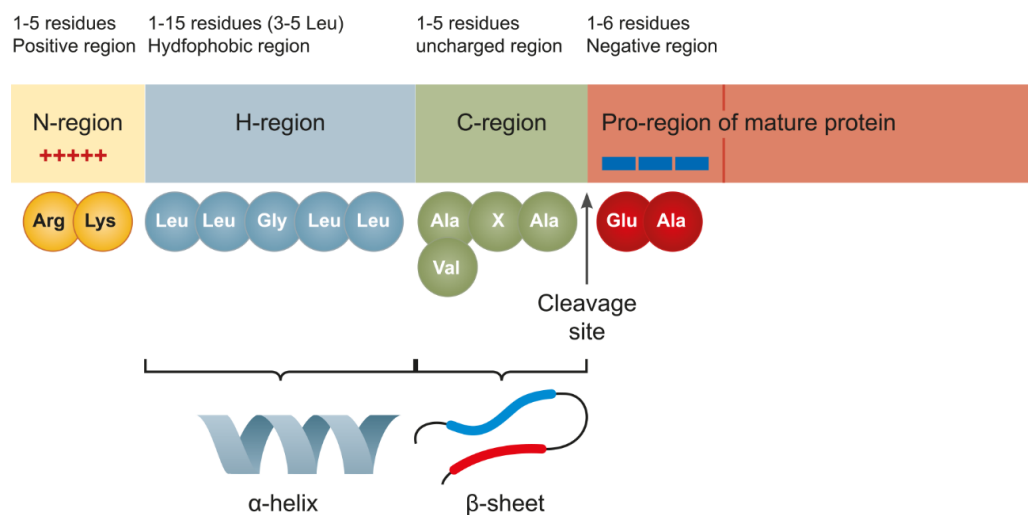
**Figure 1.** The cotranslational protein targeting pathway: (1) The SRP binds a signal sequence (dark blue) as it emerges from a ribosome to form the RNC–SRP complex; (2) the RNC–SRP complex docks with the ER membrane by binding to a cognate SRP receptor (SR, consisting of  $\alpha$  and  $\beta$  subunits); (3) subsequently, the RNC is transferred from SR–SRP to translocon Sec61, resulting in the intercalation of the signal sequence into the translocon pore and its opening; (4) the nascent chain is translocated through the pore, and the SRP disengages from the SR; (5) concomitantly with translocation, the signal peptidase (SPase) and oligosaccharyltransferase (OST) enzyme complexes are recruited to the translocon to cleave the SP and add N-linked glycans to the nascent chain, respectively. The termination of the protein synthesis releases the nascent chain from the ribosome; the translocation is completed; and the protein folds in the ER lumen (adapted from [38] with permission).

### 2.3. Structure of SPs Targeting Proteins to the ER and Its Derivatives

The SPs of different proteins, mediating their delivery to the ER and its derivatives, considerably differ in their primary structure but have a common structural design. They consist of 20–30 aa and comprise three domains: a positively charged N-terminal domain (1–7 aa), a central hydrophobic domain (7–15 aa), and a polar C-terminal one (3–7 aa), carrying an SPase recognition site [53]. Longer SPs are more abundant in eukaryotes but are also detectable in viral and bacterial proteins. Often, long SPs are not cleaved during protein maturation and may perform some other functions along with targeting [54].

The N-terminal domain is responsible for the interaction with negatively charged phosphate groups of the double lipid membrane and is important for the protein translocation across it [55] and for the interaction with the phosphate backbone of the SRP complex [56]. A positive charge of the N-terminal domain also determines the protein orientation favorable for its translocation across the membrane [57]. The high hydrophobicity of the central SP domain results from a large number of leucine residues, which stabilize its  $\alpha$ -helical structure [57,58]. The hydrophobic domain determines the SP orientation relative to the ER membrane, subsequent processing (for example, N-glycosylation), and the rate and

efficiency of protein translocation across the membrane. A decrease in the hydrophobicity of the domain lowers the rate of protein translocation to the ER [59]. The helix-breaker residues Gly, Pro, and Ser, situated in the central part of the hydrophobic domain, give rise to a hairpin-like structure, which enhances SP insertion into the membrane and its subsequent cleavage by SPase [60]. A change in the position of helix-breaker residues in the hydrophobic domain may substantially change the level of protein accumulation [61]. In particular, the insertion of Gly into the center of the hydrophobic SP domain of recombinant cyclodextrin glucanotransferase doubles the level of protein secretion [62]. The C-terminal domain of SPs of different eukaryotic proteins somewhat varies in its length (3–7 aa) and amino acid composition and is composed of neutral or negatively charged amino acids that form a  $\beta$ -sheet, representing an SPase recognition site. This site is rather conserved among almost all translocated eukaryotic proteins but considerably differs among only a few lipoproteins. Positions  $-1$  and  $-3$  from the cleavage site (most frequently containing an Ala residue and forming the so-called AXA motif) are of paramount importance. A substitution with another amino acid, especially at position  $-1$ , causes a shift in the cleavage site [53,63] (Figure 2).



**Figure 2.** General structure of an SP. It is composed of three main parts: (1) an N-terminal region (a positively charged domain); (2) H-region (the hydrophobic core forming an  $\alpha$ -helix); and (3) a C-terminal region (a cleavage site forming a  $\beta$ -sheet). The initial part of the protein, important to protein secretion, is referred to as the pro-region. Cleavage occurs within the AXA or VXA motif (adapted from [19], with permission).

Gly, Ser, and Cys occur infrequently at positions  $-1$  and  $-3$  [64], whereas Ala, at position  $-3$ , can be replaced with Val, Leu, or Ile [65]. The amino acid composition at other positions of the C-domain, except for the AXA motif, may vary widely and has almost no effect on the interaction with SPase [66]. The amino acid residues located immediately after the cleavage site are referred to as the pro-region; its length can reach 30 aa. The first six residues of this region are the most important, being involved in the interaction with SPase [61,67]. The pro-region mainly consists of neutral and acidic residues. A substitution of negatively charged residues with basic ones deteriorates protein secretion and changes the cleavage site [53] (for SP sequences, see the Signal Peptide Website at <http://www.signalpeptide.com> (accessed on 25 September 2022), UniProt Database (accessed on 25 September 2022), and SPSED at <http://www.spсед.com> for their efficiency (accessed on 25 September 2022).

#### 2.4. Using SPs Targeting to the ER and Its Derivatives to Raise the Accumulation Efficiency of Recombinant Proteins

The ER is the beginning of the secretory pathway for many proteins. Upon entering the ER, followed by a trip to the Golgi complex, proteins are subject to PTMs, folding, and quality control and can then be translocated to vacuoles or the apoplast with the help of COPII-coated vesicles [68]. Furthermore, plants have functionally specialized, ER-derived vesicles, including storage PBs [69,70]. Unlike COPII, ER-derived vesicles do not follow the standard route through the Golgi complex but are rather formed in the ER lumen, where the PBs either become independent storage entities or find themselves in storage vacuoles [71].

##### 2.4.1. Retention of Recombinant Proteins in the ER

Many therapeutic proteins, such as monoclonal antibodies or various enzymes, require correct glycosylation to preserve their pharmacological properties and optimal pharmacodynamics, as well as to guarantee the absence of adverse effects. Complex N-glycans, considerably differing between plants and mammals, are synthesized in the Golgi cis- and trans-cisternae. In plants, they contain xylose and fucose residues, which are absent in animal glycans. It is believed that these glycans can induce allergic reactions. However, the high-mannose N-glycans formed in the ER at initial glycosylation stages are identical between plants and humans [72]; thus, it is reasonable to prevent the protein translocation to Golgi cisternae by accumulating it in the ER instead. For this purpose, a short C-terminal sequence, (SE)KDEL/HDEL, acting as a signal for retrieving the synthesized protein to the ER, is attached to the protein in addition to the N-terminal SP [73,74]. Along with the prevention of undesirable glycosylation of the plant type, the retention of the protein in the ER enhances its accumulation because the ER contains only a small number of protease types [75]. Despite the concerns that high-mannose glycans have a negative effect on the pharmacological polypeptides of synthesized recombinant proteins, the method of protein retention in the ER is widely used because it allows for a rapid increase in the yield of a target protein [73]. The better accumulation and improved stability of many recombinant proteins in the cell have been shown for many species of transgenic plants as a result of targeting and retention in the ER [76–81]. Unfortunately, this translocation of recombinant proteins fails in certain tissues sometimes and has undesirable consequences. In particular, some recombinant antibodies targeted to the ER in *Arabidopsis* and tobacco seeds partially enter the Golgi complex, followed by the storage vacuoles and the apoplast, where they undergo glycosylation of an undesirable type or are degraded proteolytically [82–85].

##### 2.4.2. Targeting of Recombinant Proteins to PB-like Structures

Cells of the cereal seed endosperm possess two types of structures containing storage proteins, namely, ER-derived PBs (mainly composed of prolamins) and post-Golgi protein storage vacuoles (containing globulins and glutelins). Both organelles have a low amount of water and a low protease activity; accordingly, they are most suitable for the accumulation and long-term storage of recombinant proteins [73,86,87]. One of the PB types in the maize (*Zea mays*) endosperm is formed from proteins called zeins, which account for over half of the total endosperm protein [88]. The 27 kDa  $\gamma$ -zein is localized to the periphery of PBs formed in the ER, where it encompasses the aggregates of other zeins, stabilizes them, and enhances their sequential assembly into PBs [89,90]. Heterologous  $\gamma$ -zein expression in the leaf epidermis of *Arabidopsis* and tobacco leads to the formation of membrane-encompassed structures closely resembling PBs of the cereal endosperm, suggesting that one or several  $\gamma$ -zein structural motifs are responsible for the de novo formation of these bodies [91].  $\gamma$ -Zein contains a proline-rich domain at its N terminus; the domain is represented by the PPPVHL sequence repeated eight times. An artificial peptide, Zera, is used to assemble the PBs of a target recombinant protein. This polypeptide comprises 112 aa and contains the signal  $\gamma$ -zein peptide (which targets it to the ER) and its proline-rich domain [92,93] (with six cysteine residues capable of forming disulfide bonds between chains, thereby ensuring the oligomerization of Zera-containing molecules and the formation of PBs) [94,95].

The sequence coding for the C-terminal end of the Zera peptide is directly fused to the gene of a target protein. The expression of this hybrid protein leads to PB formation not only in vegetative plant tissues but also in insect and mammalian cells [92,93]. The isolation and purification of the proteins fused to Zera facilitate the centrifugation and separation of the induced PBs in density gradients [95,96]. A number of therapeutic proteins, such as calcitonin, epidermal growth factor, and human growth hormone, have been successfully synthesized in plant cells via the fusion to Zera and PB induction [93]. A protein recombinant vaccine against human papillomavirus, with its Zera-fused E7 protein as the major component, has been successfully synthesized in *Nicotiana benthamiana* with a high yield and specific immunogenicity [97].

Along with Zera, two more peptides are successfully used for the translocation of target proteins to induced PBs; these peptides are elastin-like polypeptides (ELPs) and hydrophobins (HFBs). ELPs are synthetic peptides comprising pentapeptide VPGXG (X: any nonproline amino acid) repeated 5 to 160 times, first discovered in the elastin of mammals [98]. ELPs share structural characteristics with intrinsically disordered proteins and are capable of reversible transition from a soluble protein to insoluble aggregates and back depending on specific transition temperatures [99]. This property of ELPs can be exploited for the rapid purification of a protein fused to an ELP by an inverse transition cycling procedure [98]. The ELP of 30–40 VPGXG repeats considerably increases the accumulation of spider silk proteins, interleukins, and monoclonal antibodies such as PBs in the tobacco leaves and cells [100–102]. A GFP–ELP fusion considerably increases GFP (green fluorescent protein) accumulation in tobacco leaves (to 40% of the total soluble protein, TSP) because of the induction of the formation of GFP-containing PBs [103]. HFBs are a family of small secreted proteins synthesized by filamentous fungi—in particular, *Trichoderma reesei* [104]. HFBs are globular proteins with four disulfide bonds and a hydrophobic layer on the surface, which underlie their hydrophobic and surfactant characteristics. Proteins fused to HFBs acquire hydrophobic properties and can be utilized to purify the former via aqueous two-phase separation [105]. The transient expression of GFP fused to an HFB in *N. benthamiana* leaves gives rise to numerous PBs and increases GFP accumulation up to 51% of TSP. The fusion of other target proteins to HFBs also elevates their yield [106].

Of note, all three peptide agents induce PB assembly not only in plants but also in the cells of all three eukaryotic kingdoms, implying the high conservation of the mechanism underlying PB formation. The induced PBs coated with a membrane are protected from proteolytic degradation by cytosolic enzymes, while the cell itself is protected from undesirable effects of foreign proteins. In addition, PBs have a high density and are easily separable by centrifugation, which considerably facilitates the purification of target proteins. Thus, these properties are suggestive of the utility of induced PBs for the production of recombinant proteins.

#### 2.4.3. Targeting of Recombinant Proteins to Vacuoles

In seeds and other plant storage tissues, vacuoles are specialized for long-term protein storage; accordingly, the targeted translocation of recombinant proteins to vacuoles seems worthwhile [107]. On the contrary, the accumulation in vacuoles of leaves or undifferentiated cells in suspension culture is rather undesirable for many recombinant proteins because the conditions in these compartments cannot guarantee their stability. Nonetheless, there are some examples of a high accumulation of certain proteins (endolysin, avidin, and cellulolytic enzymes) in the central vacuoles of leaves [108]. The targeting of monoclonal antibodies to vacuoles has also been successfully implemented in *N. benthamiana* [109]. DLLVDTM, a vacuolar SP from chitinase A of tobacco, has been used to synthesize human  $\beta$ -glucocerebrosidase, a therapeutic for Gaucher's disease. The glucocerebrosidase synthesized in carrot cell culture is the first pharmaceutical produced in a plant and approved for human use. To be pharmacologically active, this enzyme has to be glycosylated and contain exposed mannose residues. This particular type of glycosylation takes place in plant vacuoles and yields paucimannosidic-type N-glycan structures [110]. These specific features of the PTMs in vacuoles make them preferable for the translocation of some recom-

binant proteins. *Aspergillus niger* polygalacturonase, transiently expressed in *N. benthamiana* and targeted to vacuoles, accumulates there in a lower amount as compared with the enzyme targeted to the ER and apoplast but manifests a considerably higher activity, which is most likely explained by the specific PTM features in the vacuole [111]. One of the *Triticum aestivum* enzymes involved in fructan synthesis, 6-SFT, carries a 26-aa SP (at its N terminus) that targets it to vacuoles. GFP has been fused to this SP to obtain transgenic sugarcane plants with the vacuoles of stem parenchyma cells containing a large amount of GFP. A similar picture is observed for transient SP–GFP expression in celery and maize stigmata [112]. Consequently, the SPs targeting recombinant proteins to vacuoles can be effective at increasing the yield of a target product, especially if these proteins require specific PTMs implementable in vacuoles.

#### 2.4.4. Targeting of Recombinant Proteins to the Apoplast

Most of secreted proteins are transported via the route conserved among all eukaryotes, which starts from the translocation to the ER lumen through Golgi cisternae to be then delivered by vesicles to the cytoplasmic membrane and—with the help of exocytosis—to the apoplast [113,114]. To enter the ER, where a protein is folded and subjected to initial PTMs, the protein requires a specialized secretory SP at its N terminus [115,116]. Proteins are further post-translationally modified in the trans-cisternae, and some of them are translocated to vacuoles [117]. The apoplast targeting of recombinant proteins is widely used in the cultivation of transgenic plant cells because a protein of interest in this case is directly released into the culture medium to be subsequently extracted easily [118]. Most of secreted proteins carry a short (20–40 aa) peptide targeting them to the outside of the cell. The SPs of this type have no conserved regions, and individual secreted proteins may carry completely different SPs. On the other hand, the secretion mechanisms, which are highly conserved among eukaryotes, allow SPs of a heterologous origin—for example, animal SPs in a plant expression system—to be employed for targeting recombinant proteins to the apoplast. In particular, the authentic SP of human growth hormone ensures the correct processing and translocation of human growth hormone to the apoplast of *N. benthamiana* [16,119], as does the authentic SP of human acid  $\alpha$ -glucosidase in *A. thaliana* cell culture [120]. Various SPs yield different secretion efficiency levels; hence, these peptides should be individually selected for each expression system. For example, the 33KDsp peptide has emerged as the most efficient in the rice suspension-cultured cells of three tested variants: it provides the targeting of up to 92% of a recombinant protein to the apoplast [121].

### 3. Protein Translocation to Endosymbiotic Organelles

Mitochondria and plastids are endosymbiotic organelles originating from ancestors of extant  $\alpha$ -proteobacteria and cyanobacteria. The symbiosis commenced at different time points; the progenitors of mitochondria were the first to be taken up by a proto-eukaryotic cell, followed by plastids [122]. The transformation of endosymbionts into organelles has been accompanied by the massive translocation of their genetic material to the nucleus. Given that most of mitochondrial and plastid proteins are synthesized in the cytosol on nuclear transcripts [123–125], these proteins must be targeted from the cytosol to the corresponding organelles. The majority of proteins targeted to plastids and mitochondria carry at their N terminus an SP named the “transit peptide” for chloroplasts and “presequence” for mitochondrial targeting; it is recognized by import mechanisms of these organelles and then cleaved by specific signal peptidases. One of the most intriguing aspects is that the protein-targeting mechanisms of plastids and mitochondria are very similar [126]. Along with the specific sequences targeting proteins to either plastids or mitochondria, some SPs are recognizable by the import machineries of both organelles [127]. Recently, an ever-increasing volume of data revealed that proteins are targeted to mitochondria or to the ER not only by means of SPs but also at the level of the targeted transport of the mRNAs carrying zip codes for the interaction with RNA-binding proteins at their 3' end [128,129]. As a result, the mRNAs and even mRNA complexes with ribosomes are localized near



the mitochondrial outer membrane and even on its surface, thereby directly interacting with its import machinery. Note that protein targeting to mitochondria can be not only post-translational, as previously believed, but also cotranslational. The role of mRNA targeting in the case of plastids is rather vague; however, some indirect evidence suggests that it occurs there too [130,131]. This issue is beyond the scope of our review.

### 3.1. Major Mechanisms Underlying Protein Import into Mitochondria

Most proteins targeted from the cytosol to mitochondria carry an N-terminal SP referred to as presequence, which is subject to cleavage later. The main part of these precursor proteins is transferred to the translocon on the external side of the outer mitochondrial membrane (TOM complex, comprising several proteins), which is the main portal for the protein import into a mitochondrion [123,132]. Tom20 and Tom22 recognize the presequence of a transported precursor and directly bind to it, allowing for its passage through the pore formed by Tom40 to the intermembrane space of a mitochondrion. Small proteins, Tom5–Tom7, are involved in the functioning of the TOM complex [133]. Then, the presequence of the precursor protein binds to Tim50 of the TIM transmembrane complex (of the inner mitochondria membrane), which sorts the precursors and targets them to either the matrix or inner membrane or leaves them in the intermembrane space [132,134]. The matrix-targeted precursors pass through the pore formed by TIM23 coupled with an ATP-dependent translocon-associated motor. Upon entering the stroma, the presequences of precursors are cleaved by mitochondrial processing peptidase, and the protein is folded by the chaperones Hsp60 and Hsp10 [132].

Most of the mitochondrial proteins synthesized on eukaryotic 80S ribosomes are post-translationally transferred to these organelles [135,136]. However, the mechanisms of this protein traffic are studied insufficiently. It is known that molecular chaperones, such as Hsp70 and Hsp90, and their cochaperones play important roles in this process. Hsp70 binds to the short hydrophobic presequence of the protein, which explains its broad substrate specificity. Hsp70 maintains the protein in an unfolded state, which is necessary for its passing through the TOM complex to the intermembrane space [135,136]. Tom70 associated with Tom40 contains a TPR (tetratricopeptide repeat): the domain recruiting and binding molecular chaperones and preventing the aggregation of precursor proteins [137]. It is not clear how Hsp70 specifically mediates the targeting of mitochondrial proteins because the same chaperone targets proteins to the ER as well [138] (Figure 3A).

Hsp70 cochaperones and J proteins also take part in the transport of precursor proteins to mitochondria. The latter stimulate the Hsp70 ATPase activity, contain the domain-binding precursors, and are able to deliver them to Hsp70 [139,140]. Some precursors of hydrophobic proteins of the inner mitochondria membrane are first delivered to the outer ER surface but do not get integrated into the membrane. After that, one of the J proteins transfers these precursors from the ER surface to the outer mitochondrial membrane. This is the so-called ER–SURF pathway [141].

Some data suggest that the protein import into a mitochondrion follows the cotranslational pattern. First, translation-arrested ribosomes were observed by cryo-tomography on the mitochondrial surface in the region of TOM [142]. Second, many proteins of the inner mitochondrial membrane were shown to be synthesized on mitochondria-associated ribosomes [143]. Third, an mRNA coding for mitochondrial proteins was observed on the surface of mitochondria [144,145]. This mitochondrial localization of mRNA is associated with the presence of zip codes in its 3' untranslated region and an interaction with the Puf3 mRNA-binding protein [146]. The transport of mRNAs of mitochondrial proteins, along with translating ribosomes to the surface of mitochondria, is not well understood. The cleaved presequence at the precursor's N terminus is the first to leave the ribosome exit tunnel and to be able to interact with Tom20, thereby initiating the import of the nascent precursor [129,147]. The initiation of cotranslational import into a mitochondrion requires that the complex of a ribosome, mRNA, and nascent polypeptide chain (RNC) be in direct proximity to the TOM complex. Om14, an outer mitochondrial membrane protein, can

serve as a receptor for RNC-associated complexes of this type [148,149]. The molecular mechanism behind the translocation of a nascent protein chain of the overall RNC complex from Om14 to the TOM complex has yet to be studied.

### 3.2. Main Mechanisms of Protein Import into Plastids

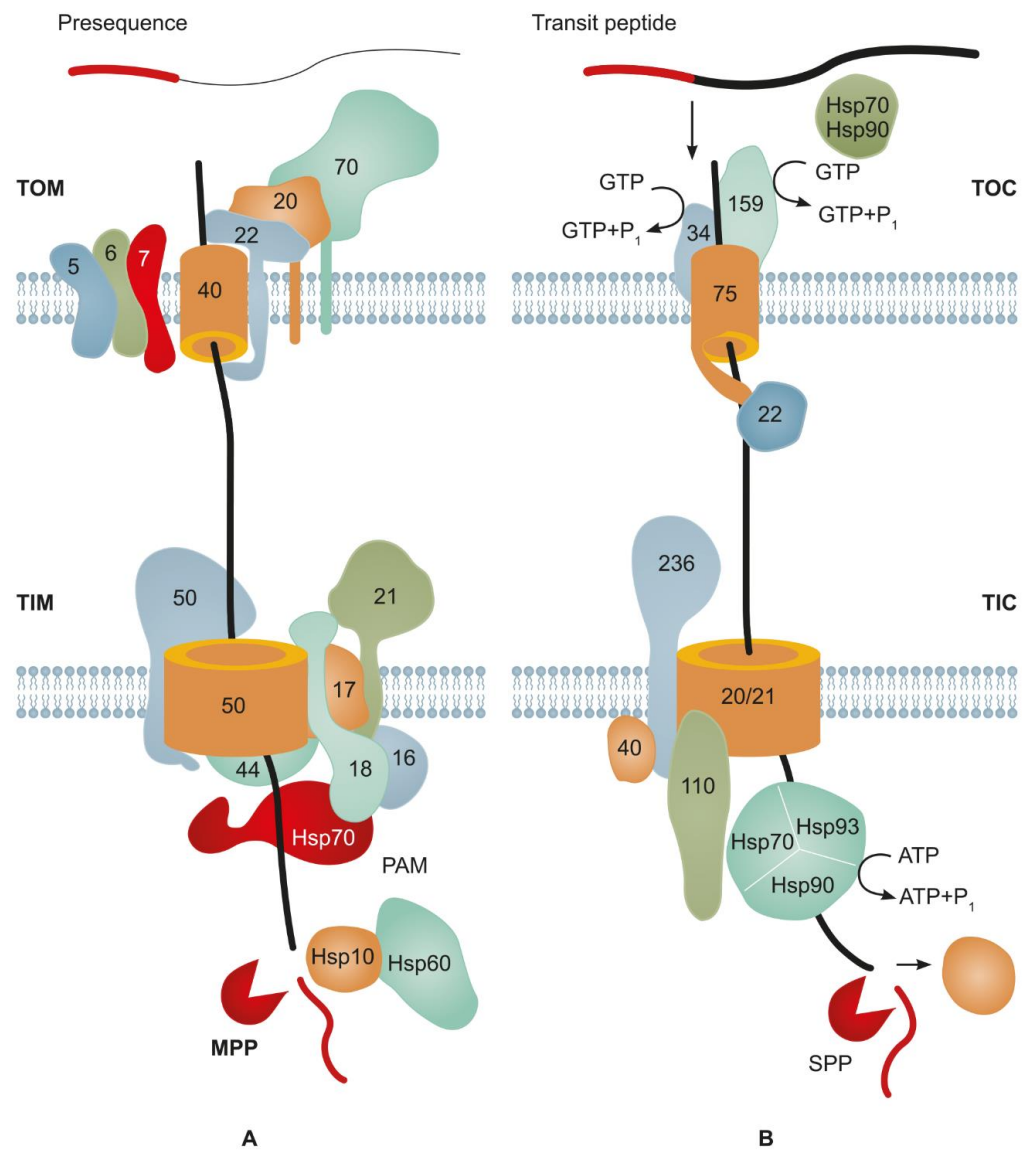
Most of the proteins transferred to plastids carry an SP (or transit peptide, determining their target) at their N terminus. The transit peptides targeting proteins to plastids are most different in their primary structure and length (30–150 aa) [150]. As a rule, transit peptides are cleaved from a proprotein in the plastid stroma by stromal processing peptidases (SPPs) [151]. Most proteins are transported to plastids in a post-translational manner and enter the TOC/TIC protein translocation system (Figure 3). The TOC translocon resides on the outer plastid membrane, while the TIC is responsible for the translocation across the inner membrane [152,153]. Numerous chaperone complexes of different families, including Hsp70 and Hsp90, interact with the proteins targeted to plastids and interfere with their folding, misfolding, and aggregation [154]. The transit peptides of plastid preproteins bind to the core of the TOC complex assembled from Toc34, Toc159, and Toc75 [155]. Toc34 and Toc159 are membrane-bound GTPases acting as primary import receptors of the outer membrane [156]. Their GTPase activity triggers the transit of a TOC-bound preprotein [157]. In this process, Toc75 is the major component forming the protein import channel of the outer membrane [158].

The proteins targeted to the plastid stroma have to pass through the inner membrane too. For this reason, the TOC complex is tightly associated with the second translocation system situated in the inner membrane, the TIC, forming a supercomplex for transferring proteins from the cytosol to the plastid stroma [159]. The joint work of the TOC and TIC complexes is coordinated by Tic236, which mediates physical contacts among the TOC, the TIC membrane channel, and components of the ATP-dependent molecular motor responsible for the preprotein translocation across two membranes of the plastid envelope [160]. Small Tic22 chaperons are involved in the transit of preproteins across the intermembrane space [161]. Tic20 and Tic110 form the main transmembrane channels across the inner membrane [162]. In the stroma, the imported protein interacts with the complex of stromal chaperons, comprising Hsp90C, cpHsp70, and ClpC/Hsp93, which provide the folding of the imported protein and its interaction with the stromal processing peptidase cleaving the transit sequence [151,163] (Figure 3B).

The main protein import routes to plastids are studied much worse as compared with mitochondrial ones; however, it is clear that the post-translational mechanism is not the only one. Proteins are also imported into plastids in a cotranslational manner via the targeting of plastid protein mRNAs in a complex with cytosol ribosomes onto the external surface of the plastid outer membrane. As shown by the in situ hybridization of *Chlamydomonas reinhardtii*, the mRNA coding for the LHCI chloroplast protein mainly accumulates in its basal region. Puromycin treatment interferes with this localization of the LHCI mRNA on the chloroplast membrane. It is noteworthy that puromycin causes a premature release of the nascent polypeptide chain from the ribosome; accordingly, the LHCI import into the chloroplast is explainable by the targeted transport of mRNA to the chloroplast membrane coupled with translation, that is, in a cotranslational manner [164]. Recent studies based on high-resolution electron tomography indicate that the cytosol ribosomes associated with the mRNAs coding for subunits of chloroplast proteins called LHCs (light-harvesting complexes) and RBCs (Rubisco small subunits) reside on the external surface of the basal part of the *C. reinhardtii* outer chloroplast membrane and are translationally active. Therefore, these proteins are imported into the chloroplast via TOC/TIC translocons in a cotranslational manner. Once cytosol ribosomes emerge on the chloroplast membrane, the membrane must have receptors for these ribosomes; the fact that their dissociation requires high ionic strength supports this theory [165].

Some proteins transferred to plastids require glycosylation for their functionality. At their N terminus, they carry a peptide targeting them to the ER. They are then imported in

a cotranslational manner to go to the Golgi complex, where they are glycosylated before reaching the target plastids. The mechanism underlying the import from the Golgi complex is still obscure [166,167].



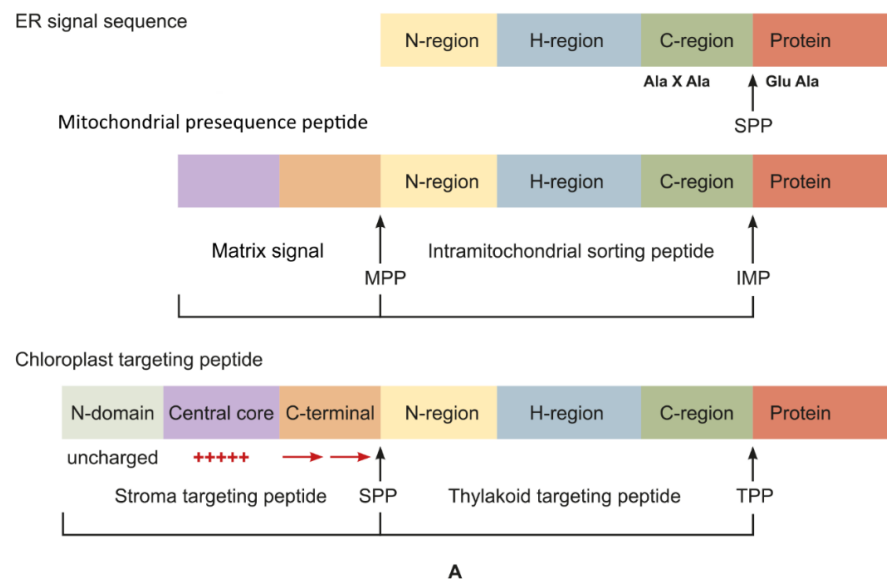
**Figure 3.** Mitochondrial and chloroplast protein import machineries. (A) Mitochondrial transmembrane protein transport. TOM, translocon of the outer mitochondrial membrane; TIM, translocase of the inner mitochondrial membrane; PAM, presequence translocon-associated motor; and MPP, mitochondrial processing peptidase. Approximate molecular weights (kDa) are indicated on individual TOM, TIM, and PAM components. See the main text for mechanistic details (adapted from [135], with permission). (B) Core components of TOC–TIC import machinery of chloroplasts. A newly synthesized preprotein is targeted to the TOC complex at the outer membrane by the binding of its transit peptide to receptors Toc34 and Toc159 with GTPase activities. The targeting is aided by the cytosolic chaperone complexes of Hsp70 and Hsp90. The GTPase activity of the receptors is required for the transport of the preprotein through the TOC and TIC transmembrane channels formed by TOC75 and TIC20/21, respectively. The TOC–TIC supercomplex assembled through the binding of Tic236 to Toc75 and of Tic110 to the TIC complex facilitates the direct transport of the preprotein from the cytosol to the chloroplast stroma. Tic110 and Tic40 (40), together with the chloroplast chaperones Hsp70, Hsp93, and Hsp90, form the ATP-dependent protein motor performing import into the chloroplast stroma. After the import, the transit peptide is cleaved by stromal processing peptidase (SPP) (adapted from [159], with permission).

### 3.3. Structure of SPs Targeting Proteins to Endosymbiotic Organelles

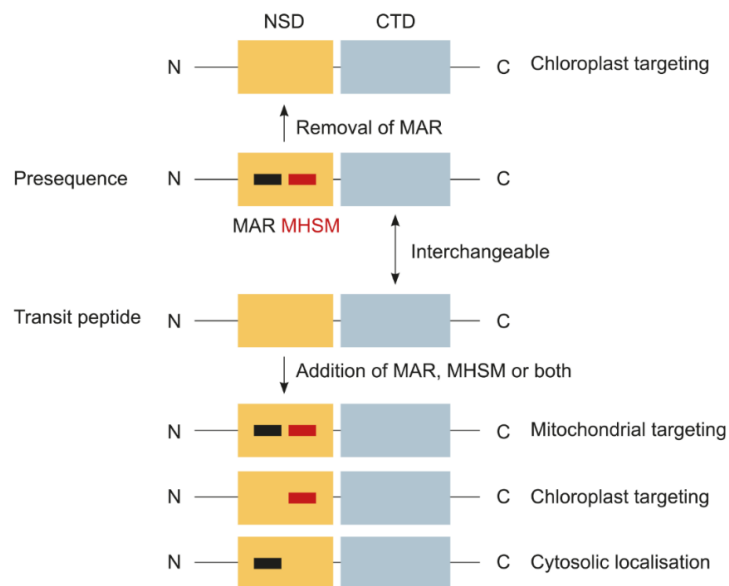
The analysis of the structure of many transit peptides in chloroplast proteins and presequences of mitochondrial proteins suggests that their signal sequences are rather similar in their amino acid composition despite tremendous differences in primary structures [168]. Neither the former nor latter contain any conserved consensus sequences. The differences between various transit peptides and presequences are the same as those between these groups of SPs. The analysis of the primary structure of transit peptides and presequences indicates that they contain numerous dispersed short motifs responsible for different stages of import into the respective organelles (cytosol navigation, interaction with TOC/TOM and TIC/TIM complexes, and interaction with chaperones) [169,170] (Figure 4A).

Two functional domains are distinguishable in the transit peptides and presequences, namely, an N-terminal specificity domain (NSD) and a C-terminal translocation domain (CTD). Although the transit peptides and presequences are very similar, they have a difference in their structure: the N-terminal region of transit peptides is mostly hydrophobic and forms random coils, in contrast to presequences, where it is less hydrophobic, contains multiple arginine residues (MAR) and the moderately hydrophobic sequence motif (MHSM), and forms an amphiphilic  $\alpha$ -helix [169,171]. The charge and amino acid composition of the N-terminal domain are determinants of the target (chloroplast, mitochondrion, or both organelles) of a signal sequence [127,171,172]. The removal of the MAR from a presequence switches the targeting from mitochondria to chloroplasts [169]; the addition of the MAR to the NSD of transit peptides blocks the targeting to chloroplasts, and the protein remains in the cytosol; an extra copy of the MHSM retargets the protein to mitochondria. Nevertheless, the addition of the MHSM alone fails to retarget the transit peptide to chloroplasts [169,173]. Thus, the presence or absence of MAR in SPs is the key determinant of the targeting to mitochondria or chloroplasts. The C-terminal domain (CTD) of transit peptides and presequences is the signal for the transit through TOC/TIC and TOM/TIM import systems [19,169,174] (Figure 4B). The CTDs of both systems of transmembrane translocation are almost interchangeable, i.e., the mitochondrial system recognizes a chloroplast CTD signal and vice versa [169]. See the UniProt Database (<https://www.uniprot.org>, accessed on 25 September 2022) for the sequences of chloroplast transit peptides and presequences of mitochondria.

Consequently, the transit peptides and presequences targeting proteins to plastids and mitochondria, respectively, are rather alike and share similar structures but lack any consensus sequences. A multitude of SPs are capable of concomitant targeting to both organelles, i.e., dual targeting. These signal sequences guide the proteins functioning in both compartments [169,175,176]. Accordingly, the question arises of whether and to what degree the presequences are conserved among eukaryotes and whether plant presequences differ from animal and fungal ones, which do not need any segregation between mitochondrial and plastid targeting. Amazingly, the presequences of animal and fungal mitochondrial proteins are capable of mitochondrial targeting in plant cells, whereas plant presequences can target to animal mitochondria. The replacement of the MAR with alanine residues in a nonplant presequence retargets the protein to plant cell chloroplasts. Moreover, MAR and MHSM insertion into an NSD of a plant transit peptide guarantees the protein's targeting to mitochondria in animal cells [177]. Most likely, the need for transit signals appeared at the early stages of organogenesis, when the nucleus started capturing the genetic material of proto-organelles. It is unlikely that this signal mechanism arose *de novo*; rather, one of the pre-existing cell systems formed its basis. According to one of the hypotheses, the NSD of SPs of plastids and mitochondria has originated from the bacterial twin-arginine translocation (TAT) SP, which can substitute for the presequence in both plant and animal cells [177]. Another hypothesis postulates that the SPs for protein targeting to endosymbiotic organelles have originated from the antimicrobial peptides of the host cell, thus destabilizing the cell membrane of prokaryotes, and have been imported by them for detoxification by proteolysis in the cytoplasm [178,179].



A



B

**Figure 4.** Specific features of the structure of mitochondrial and chloroplast SPs. (A) Eukaryotic signal sequences depending on the target organelle: the ER SP has a common tripartite structure; the mitochondrial-targeting peptide is composed of a matrix signal and intramitochondrial sorting signal; and the chloroplast transit peptide comprises the stroma-targeting peptide and thylakoid-targeting peptide. The intramitochondrial sorting signal and thylakoid-sorting signal share a common tripartite structure. Upward arrows denote the cleavage site, and the horizontal arrow denotes a  $\beta$ -sheet (adapted from [19], with permission). (B) The relation between the chloroplast transit peptide and mitochondrial presequence. The N-terminal specificity domain (NSD) of mitochondrial presequences contains multiple arginine residues (MAR) and a moderately hydrophobic sequence motif (MHSM). The removal of the MAR is sufficient to switch the targeting specificity from mitochondria to chloroplasts. Conversely, the incorporation of both the MAR and MHSM into the NSD of chloroplast transit peptides changes the targeting specificity from chloroplasts to mitochondria. The insertion of the MAR or MHSM alone into a transit peptide results in cytosolic localization or chloroplast targeting, respectively. C-terminal translocation domains (CTDs) are interchangeable between these targeting signals (adapted from [170], with permission).

### 3.4. Targeting of Recombinant Proteins to Endosymbiotic Organelles

Plastids possess their own genome and are the most attractive for the biosynthesis of recombinant proteins because transplastomic plants have a number of advantages over nuclear transformants. However, the creation of homoplastomic plants or cell cultures entails certain difficulties, making the transformation of the nuclear genome a much easier task [180,181]. On the other hand, chloroplasts perhaps have the lowest level of proteolytic activity among all cell compartments [182] and are readily separable from the remaining cell components; therefore, they are perfectly suited for the storage of produced recombinant proteins.

For the targeting of recombinant proteins to the chloroplast, they are fused to a suitable N-terminal peptide—most often, to the transit peptide of the Rubisco small subunit (RbcS). This approach frequently exerts a considerable positive effect on the protein accumulation and is thus widely used, especially in the design of plant-made vaccines for humans and animals [183]. Over the last 30 years, more than 10 recombinant proteins synthesized from a nuclear transcript have been targeted to chloroplasts. The transit peptides of, e.g., RbcS, chlorophyll a/b-binding protein, and granule-bound starch synthase have been utilized for the targeting [184]. Such proteins as phosphoenolpyruvate synthetase,  $\beta$ -glucuronidase, and xylanase have been synthesized and accumulated in different plant species, such as petunia, potato, tobacco, and rice [185–188]. The level of protein accumulation has varied considerably. In particular, the resulting xylanase constitutes almost 5% of TSP, in contrast to phosphoenolpyruvate synthetase, which constitutes only 0.1% of TSP [185,188]. The amount of GFP synthesized in rice cells reaches 10% of TSP [189]. Of interest is that the RbcS transit protein has been used for the chloroplast targeting of all three proteins [184]. The Cry1Ac endotoxin of *Bacillus thuringiensis* and the Cel5A endoglucanase have been produced in rice and tobacco cells in an amount of 2% and 5% of TSP, respectively [190,191]. In *N. benthamiana*, the transient expression of fused p17/p24 HIV-1 (human immunodeficiency virus type 1) proteins carrying the RbcS transit peptide (ensuring chloroplast targeting) raises the expression eightfold (to ~4 mg/kg) as compared with the cytosol and ER localizations (~0.5 mg/kg) [192]. The yield of a human papillomavirus type 16 (HPV-16) capsid protein, L1, targeted to chloroplasts, reaches 11% of TSP in the case of nuclear expression and 17% in the case of transient expression [193]. As shown later, the L1 protein targeted to chloroplasts forms virus-like particles, enabling the production of a commercial virus-like-particle-based vaccine [194]. The HPV-16 oncoprotein, E7, when fused to an anti-lipopolysaccharide factor fragment (LALF32–51) transiently expressed in *N. benthamiana* leaves and targeted to chloroplasts, shows a 27-fold increase in its yield as compared with the cytosol localization [195].

The mitochondrial targeting of recombinant proteins is not as widely used for chloroplast targeting, except for the targeting of proteins, peptides, and other therapeutics and nanoparticles to mitochondria in the case of mitochondrial diseases, cancer, and many energy generation problems and other metabolic disorders [196,197]. This is a large research field beyond the scope of our review. Nonetheless, the mitochondrial targeting of recombinant proteins in plants makes sense for modulating certain functions of mitochondria, such as energy generation or responses to biotic and abiotic stressors. A low-oxygen mitochondrial environment is suitable for metabolic engineering based on oxygen-sensitive enzymes [198,199]. The mitochondrion is an isolated compartment surrounded by a double membrane and containing only a small number of proteases [182]. Baysal et al. [200] have studied the efficiency of six presequence peptides in the targeting to rice mitochondria. ATPA and COX4 (*Saccharomyces cerevisiae*), SU9 (*Neurospora crassa*), pFA (*A. thaliana*), and OsSCSb (*Oryza sativa*) successfully targeted the eGFP protein to mitochondria, whereas MTS2 (*Nicotiana glauca*) almost completely failed despite its plant origin [200].

Nitrogenase cofactor maturase Nif comprises four polypeptide chains and serves as a cofactor of all nitrogenases (key enzymes in the bacterial nitrogen fixation pathway) of diazotrophic bacteria and archaea. Its components have been expressed in *N. benthamiana* leaves and successfully targeted to mitochondria with the help of COX4 (NifB) and SU9 (NifU, NifS, and FdxN) presequences. The attainment of high levels of a soluble and func-

tional Nif cofactor in plant mitochondria is of paramount importance for the subsequent construction of a nitrogen fixation pathway in plants [201].

#### 4. Prediction of Signal Sequences That Determine Protein Localization in the Cell

Specific intracellular sorting signals considerably differ in their sequence, structure, and length between individual proteins targeted to the same compartment and between proteins localized to different organelles. The N-terminal targeting peptides guiding proteins to the secretory route (SPs), mitochondria (presequences), chloroplasts (transit peptides), and specific compartments in mitochondria and chloroplasts are the most widespread. Because these signals target the protein transport in the cell, it is important that a researcher can precisely detect them in sequences of protein-coding genes. Accordingly, a large number of computer programs utilizing various machine learning algorithms have been designed over the last quarter century, including Grammatical Restrained Hidden Conditional Random Fields, support vector machines, N-to-1 extreme learning machines, Markov chains, profile-hidden Markov models, and neural networks [202–207]. See the review by Nielsen et al. [208] for a brief history of the creation and application of these methods for the prediction of the sorting signals in proteins. The very first software predicting signal sequences in proteins was SignalP 1.0, designed in 1996 and based on machine learning principles; its version 6.0 appeared in 2021 (<https://services.healthtech.dtu.dk/service.php?SignalP-6.0> (accessed on 25 September 2022)) [209]. Unfortunately, SignalP only predicts SPs and is unable to predict presequences and transit peptides for mitochondrial and chloroplast targeting. Among the best-known software packages for the prediction of signal sequences is TargetP 2.0, based on neural networks. TargetP utilizes feed-forward networks and position weight matrices to analyze an amino acid sequence and predict SPs of the secretory route, mitochondrial presequences, chloroplast transit peptides, and thylakoid luminal transit peptides as well as the positions of their cleavage sites. TargetP predicts SPs with a probability of 97% and transit peptides and presequences with a probability of 90% [210]. TargetP 2.0 is available at <https://services.healthtech.dtu.dk/service.php?TargetP-2.0> (accessed on 25 September 2022). The TargetP predictions of SPs match, with 90% agreement, the peptides annotated in the UniProt database, whereas its predictions of presequences and transit peptides yield only 80% agreement. Notably, TargetP 2.0 predicts twice as many mitochondrial proteins in plant proteomes as compared with animal proteomes. For grape and rice chloroplasts, this program outputs these numbers at 1125 and 2049, respectively [210]. Another useful software application for the prediction of the signal sequences of all targeting types in nine subcellular locations, DeepLoc 2.0, is based on convolutional neural networks (and is available at <https://services.healthtech.dtu.dk/service.php?DeepLoc-2.0>, accessed on 25 September 2022) [211]. The review by Jiang et al. [207] gives an impressive list of computer programs predicting SPs and shows the algorithms used, the types of signal sequences, and the server addresses.

#### 5. Conclusions

The main problem with plant systems for the expression of recombinant proteins is a low yield of a target protein. This problem can be addressed via various approaches, including the translocation of a target protein to the cell compartments featuring a low protease activity. In addition, many therapeutic proteins require correct mammalian-type glycosylation for their functionality. All these problems can be solved by fusing target protein genes to signal sequences that determine the transport of the synthesized or nascent protein to certain cell compartments with either a low protease activity or appropriate glycosylation machinery. In their Golgi apparatus, plant cells form complex N-glycans with terminal fucose and xylose residues, absent in mammalian glycans. However, the initial glycosylation stages taking place in the ER and associated with the biosynthesis of polymannose glycans are similar between plants and animals. That is why it is worthwhile to target the recombinant therapeutic proteins requiring glycosylation to the ER with the

help of an N-terminal SP and keep them there by means of a C-terminal ER retrieval peptide, thereby preventing their translocation to the Golgi complex. Furthermore, the protease localization in the ER is much better compared with the cytosol, thus also enhancing the accumulation of a target protein. If a recombinant protein requires glycosylation of the paucimannosidic type, it is reasonable to target it using an SP to vacuoles, where plants implement this type of complex glycosylation. In some cases, it is beneficial to target a recombinant protein to the apoplast (where the protease activity is weak) through the secretory route. Moreover, in the case of cultured cells, it is the culture medium that allows the target protein to be extracted rather easily.

In terms of the preservation and accumulation of recombinant proteins, the most attractive approach is to construct artificial PBs that are actually membrane-enveloped microvacuoles with a high content of the target protein. It is rather easy to generate such PBs by fusing the target gene to the sequences coding for either the N-terminal domain of  $\gamma$ -zein or elastin-like/hydrophobin peptides. Additionally, these PBs have a high density and are readily extractable by density gradient centrifugation. In the same way, the accumulation of target proteins in plastids, mainly chloroplasts, is the most promising because these organelles possess the lowest protease activity of all cell compartments and are separated from the cytosol by a double membrane. Plastids have their own genome and are transformable themselves; these properties also offer a number of advantages over nuclear transformation. Nonetheless, several difficulties arise there that are linked with the construction of homoplastidic and homoplastomic plant cells. Consequently, it is much easier to produce nuclear transformants and equip the protein of interest with a transit peptide targeting it to chloroplasts. Thus, the yield of the recombinant protein can be strongly increased via the purposeful use of translocation signal sequences. However, it should be made clear that different transport peptides have dissimilar efficiency levels; therefore, each recombinant protein requires an individual approach.

**Author Contributions:** All the coauthors participated in the review of the literature, the analysis, and the interpretation of the data. E.V.D. initiated this review to clarify one of the major bottlenecks in the problem faced by the scientific community working on the expression of therapeutic recombinant proteins in plant systems and coordinated all the work. S.M.R. wrote the manuscript. Both authors were fully involved in preparing and critically revising the manuscript to its current state. The authors approved the final version for publishing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Russian Science Foundation (project No. 21-14-00091).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Burnett, M.J.B.; Burnett, A.C. Therapeutic recombinant protein production in plants: Challenges and opportunities. *Plants People Planet* **2020**, *2*, 121–132. [[CrossRef](#)]
2. Huebbers, J.W.; Buyel, J.F. On the verge of the market—Plant factories for the automated and standardized production of biopharmaceuticals. *Biotechnol. Adv.* **2021**, *46*, 107681. [[CrossRef](#)] [[PubMed](#)]
3. Schillberg, S.; Finfern, R. Plant molecular farming for the production of valuable proteins—Critical evaluation of achievements and future challenges. *J. Plant Physiol.* **2021**, *258–259*, 153359. [[CrossRef](#)] [[PubMed](#)]
4. Walsh, G.; Jefferis, R. Post-translational modifications in the context of therapeutic proteins. *Nat. Biotechnol.* **2006**, *24*, 1241–1252. [[CrossRef](#)]
5. Zhu, J. Mammalian cell protein expression for biopharmaceutical production. *Biotechnol. Adv.* **2012**, *30*, 1158–1170. [[CrossRef](#)]
6. Fischer, R.; Buyel, J.F. Molecular farming—The slope of enlightenment. *Biotechnol. Adv.* **2020**, *40*, 107519. [[CrossRef](#)]
7. Egelkrou, E.; Rajan, V.; Howard, J.A. Overproduction of recombinant proteins in plants. *Plant Sci.* **2012**, *184*, 83–101. [[CrossRef](#)]
8. Mirzaee, M.; Osmani, Z.; Frébortová, J.; Frébort, I. Recent advances in molecular farming using monocot plants. *Biotechnol. Adv.* **2022**, *58*, 107913. [[CrossRef](#)]



9. Karki, U.; Fang, H.; Guo, W.; Unnold-Cofre, C.; Xu, J. Cellular engineering of plant cells for improved therapeutic protein production. *Plant Cell Rep.* **2021**, *40*, 1087–1099. [[CrossRef](#)]
10. Rozov, S.M.; Deineko, E.V. Strategies for optimizing recombinant protein synthesis in plant cells: Classical approaches and new directions. *Mol. Biol.* **2019**, *53*, 157–175. [[CrossRef](#)]
11. Buyel, J.F.; Stöger, E.; Bortesi, L. Targeted genome editing of plants and plant cells for biomanufacturing. *Transgenic Res.* **2021**, *30*, 401–426. [[CrossRef](#)] [[PubMed](#)]
12. Rozov, S.M.; Permyakova, N.V.; Sidorchuk, Y.V.; Deineko, E.V. Optimization of genome knock-in method: Search for the most efficient genome regions for transgene expression in plants. *Int. J. Mol. Sci.* **2022**, *23*, 4416. [[CrossRef](#)] [[PubMed](#)]
13. Faye, L.; Boulaflous, A.; Benchabane, M.; Gomord, V.; Michaud, D. Protein modifications in the plant secretory pathway: Current status and practical implications in molecular pharming. *Vaccine* **2005**, *23*, 1770–1778. [[CrossRef](#)]
14. Smalle, J.; Vierstra, R.D. The ubiquitin 26S proteasome proteolytic pathway. *Annu. Rev. Plant Biol.* **2004**, *55*, 555–590. [[CrossRef](#)] [[PubMed](#)]
15. Conrad, U.; Fiedler, U. Compartment-specific accumulation of recombinant immunoglobulins in plant cells: An essential tool for antibody production and immunomodulation of physiological functions and pathogen activity. *Plant Mol. Biol.* **1998**, *38*, 101–109. [[CrossRef](#)] [[PubMed](#)]
16. Gils, M.; Kandzia, R.; Marillonnet, S.; Klimyuk, V.; Gleba, Y. High-yield production of authentic human growth hormone using a plant virus-based expression system. *Plant Biotechnol. J.* **2005**, *3*, 613–620. [[CrossRef](#)] [[PubMed](#)]
17. Powers, J.C.; Asgian, J.L.; Ekici, O.D.; James, K.E. Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem. Rev.* **2002**, *102*, 4639–4750. [[CrossRef](#)]
18. Pillay, P.; Schlüter, U.; Van Wyk, S.; Kunert, K.J.; Vorster, B.J. Proteolysis of recombinant proteins in bioengineered plant cells. *Bioengineered* **2014**, *5*, 15–20. [[CrossRef](#)]
19. Owji, H.; Nezafat, N.; Negahdaripour, M.; Hajiebrahimi, A.; Ghasemi, Y. A comprehensive review of signal peptides: Structure, roles, and applications. *Eur. J. Cell Biol.* **2018**, *97*, 422–441. [[CrossRef](#)]
20. Jarvis, P. Targeting of nucleus-encoded proteins to chloroplasts in plants. *New Phytol.* **2008**, *179*, 257–285. [[CrossRef](#)]
21. Lee, D.W.; Jung, C.; Hwang, I. Cytosolic events involved in chloroplast protein targeting. *Biochim. Biophys. Acta* **2013**, *1833*, 245–252. [[CrossRef](#)] [[PubMed](#)]
22. da Cunha, N.B.; Leite, M.L.; Dias, S.C.; Vianna, G.R.; Rech Filho, E.L. Plant genetic engineering: Basic concepts and strategies for boosting the accumulation of recombinant proteins in crops. *Int. J. Eng. Sci.* **2019**, *8*, 7–13.
23. Muthamilselvan, T.; Kim, J.S.; Cheong, G.; Hwang, I. Production of recombinant proteins through sequestration in chloroplasts: A strategy based on nuclear transformation and post-translational protein import. *Plant Cell Rep.* **2019**, *38*, 825–833. [[CrossRef](#)] [[PubMed](#)]
24. Wu, T.; Zhu, J. Recent development and optimization of pseudomonas aeruginosa exotoxin immunotoxins in cancer therapeutic applications. *Int. Immunopharmacol.* **2021**, *96*, 107759. [[CrossRef](#)] [[PubMed](#)]
25. Khan, M.S.; Joyia, F.A.; Mustafa, G. Seeds as economical production platform for recombinant proteins. *Protein Pept. Lett.* **2020**, *27*, 89–104. [[CrossRef](#)]
26. Lee, D.W.; Kim, J.K.; Lee, S.; Choi, S.; Kim, S.; Hwang, I. Arabidopsis nuclear-encoded plastid transit peptides contain multiple sequence subgroups with distinctive chloroplast-targeting sequence motifs. *Plant Cell* **2008**, *20*, 1603–1622. [[CrossRef](#)]
27. Lee, D.W.; Yoo, Y.-J.; Razzak, M.A.; Hwang, I. Prolines in transit peptides are crucial for efficient preprotein translocation into chloroplasts. *Plant Physiol.* **2018**, *176*, 663–677. [[CrossRef](#)]
28. Adams, K.L.; Palmer, J.D. Evolution of mitochondrial gene content: Gene loss and transfer to the nucleus. *Mol. Phylogenet. Evol.* **2003**, *29*, 380–395. [[CrossRef](#)]
29. Shi, L.X.; Theg, S.M. The chloroplast protein import system: From algae to trees. *Biochim. Biophys. Acta* **2013**, *1833*, 314–331. [[CrossRef](#)]
30. Aviram, N.; Schuldiner, M. Targeting and translocation of proteins to the endoplasmic reticulum at a glance. *J. Cell Sci.* **2017**, *130*, 4079–4085. [[CrossRef](#)]
31. Guerriero, C.J.; Brodsky, J.L. The delicate balance between secreted protein folding and endoplasmic reticulum-associated degradation in human physiology. *Physiol. Rev.* **2012**, *92*, 537–576. [[CrossRef](#)] [[PubMed](#)]
32. Pobre, K.F.R.; Poet, G.J.; Hendershot, L.M. The endoplasmic reticulum (ER) chaperone BiP is a master regulator of ER functions: Getting by with a little help from ERdj friends. *J. Biol. Chem.* **2019**, *294*, 2098–2108. [[CrossRef](#)] [[PubMed](#)]
33. Araki, K.; Nagata, K. Protein folding and quality control in the ER. *Cold Spring Harbor Perspect. Biol.* **2011**, *3*, a007526. [[CrossRef](#)] [[PubMed](#)]
34. Luirink, J.; Dobberstein, B. Mammalian and *Escherichia coli* signal recognition particles. *Mol. Microbiol.* **1994**, *11*, 9–13. [[CrossRef](#)] [[PubMed](#)]
35. Egea, P.F.; Stroud, R.M.; Walter, P. Targeting proteins to membranes: Structure of the signal recognition particle. *Curr. Opin. Struct. Biol.* **2005**, *15*, 213–220. [[CrossRef](#)]
36. Alberts, B.; Hopkin, K.; Johnson, A.; Morgan, D.; Raff, M.; Roberts, K.; Walter, P. *Essential Cell Biology*, 5th ed.; W.W. Norton & Company: New York, NY, USA, 2018.
37. Voorhees, R.M.; Fernández, I.S.; Scheres, S.H.; Hegde, R.S. Structure of the mammalian ribosome–Sec61 complex to 3.4 Å resolution. *Cell* **2014**, *157*, 1632–1643. [[CrossRef](#)]

38. Nyathi, Y.; Wilkinson, B.M.; Pool, M.R. Co-translational targeting and translocation of proteins to the endoplasmic reticulum. *Biochim. Biophys. Acta (BBA)-Mol. Cell Res.* **2013**, *1833*, 2392–2402. [[CrossRef](#)]
39. Pfeffer, S.; Dudek, J.; Zimmermann, R.; Förster, F. Organization of the native ribosome–translocon complex at the mammalian endoplasmic reticulum membrane. *Biochim. Biophys. Acta Gen. Subj.* **2016**, *1860*, 2122–2129. [[CrossRef](#)]
40. Sommer, N.; Junne, T.; Kalies, K.U.; Spiess, M.; Hartmann, E. TRAP assists membrane protein topogenesis at the mammalian ER membrane. *Biochim. Biophys. Acta (BBA)-Mol. Cell Res.* **2013**, *1833*, 3104–3111. [[CrossRef](#)]
41. Mohorko, E.; Glockshuber, R.; Aebi, M. Oligosaccharyltransferase: The central enzyme of N-linked protein glycosylation. *J. Inherit. Metab. Dis.* **2011**, *34*, 869–878. [[CrossRef](#)]
42. Johnson, N.; Powis, K.; High, S. Post-translational translocation into the endoplasmic reticulum. *Biochim. Biophys. Acta (BBA)-Mol. Cell Res.* **2013**, *1833*, 2403–2409. [[CrossRef](#)] [[PubMed](#)]
43. Haßdenteufel, S.; Nguyen, D.; Helms, V.; Lang, S.; Zimmermann, R. ER import of small human presecretory proteins: Components and mechanisms. *FEBS Lett.* **2019**, *593*, 2506–2524. [[CrossRef](#)] [[PubMed](#)]
44. Schuldiner, M.; Metz, J.; Schmid, V.; Denic, V.; Rakwalska, M.; Schmitt, H.D.; Schwappach, B.; Weissman, J.S. The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell* **2008**, *134*, 634–645. [[CrossRef](#)]
45. Johnson, N.; Vilardi, F.; Lang, S.; Leznicki, P.; Zimmermann, R.; High, S. TRC40 can deliver short secretory proteins to the Sec61 translocon. *J. Cell Sci.* **2012**, *125*, 3612–3620. [[CrossRef](#)] [[PubMed](#)]
46. Chartron, J.W.; VanderVelde, D.G.; Clemons Jr, W.M. Structures of the Sgt2/SGTA dimerization domain with the Get5/UBL4A UBL domain reveal an interaction that forms a conserved dynamic interface. *Cell Rep.* **2012**, *2*, 1620–1632. [[CrossRef](#)] [[PubMed](#)]
47. Mock, J.Y.; Chartron, J.W.; Zaslaver, M.A.; Xu, Y.; Ye, Y.; Clemons, W.M. Bag6 complex contains a minimal tail-anchor–targeting module and a mock BAG domain. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 106–111. [[CrossRef](#)]
48. Shao, S.; Rodrigo-Brenni, M.C.; Kivlen, M.H.; Hegde, R.S. Mechanistic basis for a molecular triage reaction. *Science* **2017**, *355*, 298–302. [[CrossRef](#)]
49. Aviram, N.; Ast, T.; Costa, E.A.; Arakel, E.C.; Chuartzman, S.G.; Jan, C.H.; Haßdenteufel, S.; Dudek, J.; Jung, M.; Schorr, S.; et al. The SND proteins constitute an alternative targeting route to the endoplasmic reticulum. *Nature* **2016**, *540*, 134–138. [[CrossRef](#)]
50. Casson, J.; McKenna, M.; Haßdenteufel, S.; Aviram, N.; Zimmerman, R.; High, S. Multiple pathways facilitate the biogenesis of mammalian tail-anchored proteins. *J. Cell Sci.* **2017**, *130*, 3851–3861. [[CrossRef](#)]
51. Shao, S.; Hegde, R.S. A calmodulin-dependent translocation pathway for small secretory proteins. *Cell* **2011**, *147*, 1576–1588. [[CrossRef](#)]
52. Haßdenteufel, S.; Schäuble, N.; Cassella, P.; Leznicki, P.; Müller, A.; High, S.; Jung, M.; Zimmermann, R. Ca<sup>2+</sup>-calmodulin inhibits tail-anchored protein insertion into the mammalian endoplasmic reticulum membrane. *FEBS Lett.* **2011**, *585*, 3485–3490. [[CrossRef](#)] [[PubMed](#)]
53. Von Heijne, G. The signal peptide. *J. Membr. Biol.* **1990**, *115*, 195–201. [[CrossRef](#)] [[PubMed](#)]
54. Hiss, J.A.; Schneider, G. Architecture, function and prediction of long signal peptides. *Brief. Bioinform.* **2009**, *10*, 569–578. [[CrossRef](#)]
55. Van Voorst, F.; De Kruijff, B. Role of lipids in the translocation of proteins across membranes. *Biochem. J.* **2000**, *347*, 601–612. [[CrossRef](#)]
56. Batey, R.T.; Rambo, R.P.; Lucast, L.; Rha, B.; Doudna, J.A. Crystal structure of the ribonucleoprotein core of the signal recognition particle. *Science* **2000**, *287*, 1232–1239. [[CrossRef](#)] [[PubMed](#)]
57. Nilsson, I.; Lara, P.; Hessa, T.; Johnson, A.E.; von Heijne, G.; Karamyshev, A.L. The code for directing proteins for translocation across ER membrane: SRP cotranslationally recognizes specific features of a signal sequence. *J. Mol. Biol.* **2015**, *427*, 1191–1201. [[CrossRef](#)]
58. Hegde, R.S.; Bernstein, H.D. The surprising complexity of signal sequences. *Trends Biochem. Sci.* **2006**, *31*, 563–571. [[CrossRef](#)]
59. Mordkovich, N.; Okorokova, N.; Veiko, V. Structural and functional organization of the signal peptide of pro-enterotoxin B from *Staphylococcus aureus*. *Appl. Biochem. Microbiol.* **2015**, *51*, 641–648. [[CrossRef](#)]
60. Driessen, A.J.; Van der Does, C. Protein export in bacteria. In *Protein Targeting, Transport and Translocation*; Dalbey, R., von Heijne, G., Eds.; Academic Press: San Diego, CA, USA, 2002; pp. 47–73. [[CrossRef](#)]
61. Low, K.O.; Mahadi, N.M.; Illias, R.M. Optimisation of signal peptide for recombinant protein secretion in bacterial hosts. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 3811–3826. [[CrossRef](#)]
62. Jonet, M.A.; Mahadi, N.M.; Murad, A.M.A.; Rabu, A.; Bakar, F.D.A.; Rahim, R.A.; Low, K.O.; Illias, R.M. Optimization of a heterologous signal peptide by site-directed mutagenesis for improved secretion of recombinant proteins in *Escherichia coli*. *J. Mol. Microbiol. Biotechnol.* **2012**, *22*, 48–58. [[CrossRef](#)]
63. Paetzel, M.; Karla, A.; Strynadka, N.C.; Dalbey, R.E. Signal peptidases. *Chem. Rev.* **2002**, *102*, 4549–4580. [[CrossRef](#)] [[PubMed](#)]
64. Karamyshev, A.L.; Karamysheva, Z.N.; Kajava, A.V.; Ksenzenko, V.N.; Nesmeyanova, M.A. Processing of *Escherichia coli* alkaline phosphatase: Role of the primary structure of the signal peptide cleavage region. *J. Mol. Biol.* **1998**, *277*, 859–870. [[CrossRef](#)]
65. Choo, K.H.; Tan, T.W.; Ranganathan, S. SPdb—a signal peptide database. *BMC Bioinform.* **2005**, *6*, 249. [[CrossRef](#)] [[PubMed](#)]
66. Ting, Y.T.; Harris, P.W.; Batot, G.; Brimble, M.A.; Baker, E.N.; Young, P.G. Peptide binding to a bacterial signal peptidase visualized by peptide tethering and carrier-driven crystallization. *IUCr* **2016**, *3*, 10–19. [[CrossRef](#)] [[PubMed](#)]
67. Voss, M.; Schröder, B.; Fluhrer, R. Mechanism, specificity, and physiology of signal peptide peptidase (SPP) and SPP-like proteases. *Biochim. Biophys. Acta (BBA)-Biomembr.* **2013**, *1828*, 2828–2839. [[CrossRef](#)]

68. Adams, B.M.; Oster, M.E.; Hebert, D.N. Protein quality control in the endoplasmic reticulum. *Protein J.* **2019**, *38*, 317–329. [[CrossRef](#)]
69. Chrispeels, M.J.; Herman, E.M. Endoplasmic reticulum-derived compartments function in storage and as mediators of vacuolar remodeling via a new type of organelle, precursor protease vesicles. *Plant Physiol.* **2000**, *123*, 1227–1234. [[CrossRef](#)]
70. Cheung, A.Y.; Cosgrove, D.J.; Hara-Nishimura, I.; Jurgens, G.; Lloyd, C.; Robinson, D.G.; Staehelin, L.A.; Weijers, D. A rich and bountiful harvest: Key discoveries in plant cell biology. *Plant Cell* **2022**, *34*, 53–71. [[CrossRef](#)]
71. Hara-Nishimura, I.; Matsushima, R.; Shimada, T.; Nishimura, M. Diversity and formation of endoplasmic reticulum-derived compartments in plants. Are these compartments specific to plant cells? *Plant Physiol.* **2004**, *136*, 3435–3439. [[CrossRef](#)]
72. Rozov, S.M.; Permyakova, N.V.; Deineko, E.V. Main strategies of plant expression system glycoengineering for producing humanized recombinant pharmaceutical proteins. *Biochemistry* **2018**, *83*, 215–232. [[CrossRef](#)]
73. Sabalza, M.; Christou, P.; Capell, T. Recombinant plant-derived pharmaceutical proteins: Current technical and economic bottlenecks. *Biotechnol. Lett.* **2014**, *36*, 2367–2379. [[CrossRef](#)]
74. Newstead, S.; Barr, F. Molecular basis for KDEL-mediated retrieval of escaped ER-resident proteins—SWEET talking the COPs. *J. Cell Sci.* **2020**, *133*, jcs250100. [[CrossRef](#)] [[PubMed](#)]
75. Strasser, R. Protein quality control in the endoplasmic reticulum of plants. *Annu. Rev. Plant Biol.* **2018**, *69*, 147–172. [[CrossRef](#)] [[PubMed](#)]
76. Ma, S.; Huang, Y.; Davis, A.; Yin, Z.; Mi, Q.; Menassa, R.; Brandle, J.E.; Jevnikar, A.M. Production of biologically active human interleukin-4 in transgenic tobacco and potato. *Plant Biotechnol. J.* **2005**, *3*, 309–318. [[CrossRef](#)] [[PubMed](#)]
77. Wang, M.L.; Goldstein, C.; Su, W.; Moore, P.H.; Albert, H.H. Production of biologically active GM-CSF in sugarcane: A secure biofactory. *Transgenic Res.* **2005**, *14*, 167–178. [[CrossRef](#)]
78. Sharma, M.K.; Singh, N.K.; Jani, D.; Sisodia, R.; Thungapathra, M.; Gautam, J.K.; Meena, L.S.; Singh, Y.; Ghosh, A.; Tyagi, A.K.; et al. Expression of toxin co-regulated pilus subunit A (TCPA) of *Vibrio cholerae* and its immunogenic epitopes fused to cholera toxin B subunit in transgenic tomato (*Solanum lycopersicum*). *Plant Cell Rep.* **2008**, *27*, 307–318. [[CrossRef](#)]
79. Viridi, V.; Coddens, A.; De Buck, S.; Millet, S.; Goddeeris, B.M.; Cox, E.; De Greve, H.; Depicker, A. Orally fed seeds producing designer IgAs protect weaned piglets against enterotoxigenic *Escherichia coli* infection. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 11809–11814. [[CrossRef](#)]
80. Choi, J.; Diao, H.; Feng, Z.C.; Lau, A.; Wang, R.; Jevnikar, A.M.; Ma, S. A fusion protein derived from plants holds promising potential as a new oral therapy for type 2 diabetes. *Plant Biotechnol. J.* **2014**, *12*, 425–435. [[CrossRef](#)]
81. Lee, G.; Na, Y.J.; Yang, B.G.; Choi, J.P.; Seo, Y.B.; Hong, C.P.; Yun, C.H.; Kim, D.H.; Sohn, E.J.; Kim, J.H.; et al. Oral immunization of haemagglutinin H5 expressed in plant endoplasmic reticulum with adjuvant saponin protects mice against highly pathogenic avian influenza A virus infection. *Plant Biotechnol. J.* **2015**, *13*, 62–72. [[CrossRef](#)]
82. Petruccielli, S.; Otegui, M.S.; Lareu, F.; Tran Dinh, O.; Fitchette, A.C.; Circosta, A.; Rumbo, M.; Bardor, M.; Carcamo, R.; Gomord, V.; et al. A KDEL-tagged monoclonal antibody is efficiently retained in the endoplasmic reticulum in leaves, but is both partially secreted and sorted to protein storage vacuoles in seeds. *Plant Biotechnol. J.* **2006**, *4*, 511–527. [[CrossRef](#)]
83. Van Droogenbroeck, B.; Cao, J.; Stadlmann, J.; Altmann, F.; Colanesi, S.; Hillmer, S.; Robinson, D.G.; Van Lerberge, E.; Terryn, N.; Van Montagu, M.; et al. Aberrant localization and underglycosylation of highly accumulating single-chain Fv-Fc antibodies in transgenic *Arabidopsis* seeds. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 1430–1435. [[CrossRef](#)] [[PubMed](#)]
84. Loos, A.; Van Droogenbroeck, B.; Hillmer, S.; Grass, J.; Kunert, R.; Cao, J.; Robinson, D.G.; Depicker, A.; Steinkellner, H. Production of monoclonal antibodies with a controlled N-glycosylation pattern in seeds of *Arabidopsis thaliana*. *Plant Biotechnol. J.* **2011**, *9*, 179–192. [[CrossRef](#)] [[PubMed](#)]
85. Sukenik, S.C.; Karuppanan, K.; Li, Q.; Lebrilla, C.B.; Nandi, S.; McDonald, K.A. Transient recombinant protein production in glycoengineered *Nicotiana benthamiana* cell suspension culture. *Int. J. Mol. Sci.* **2018**, *19*, 1205. [[CrossRef](#)]
86. Robinson, D.G.; Oliviusson, P.; Hinz, G. Protein sorting to the storage vacuoles of plants: A critical appraisal. *Traffic* **2005**, *6*, 615–625. [[CrossRef](#)] [[PubMed](#)]
87. Wakasa, Y.; Takaiwa, F. The use of rice seeds to produce human pharmaceuticals for oral therapy. *Biotechnol. J.* **2013**, *8*, 1133–1143. [[CrossRef](#)] [[PubMed](#)]
88. Larkins, B.A.; Hurkman, W.J. Synthesis and deposition of zein in protein bodies of maize endosperm. *Plant Physiol.* **1978**, *62*, 256–263. [[CrossRef](#)]
89. Lending, C.R.; Larkins, B. Changes in the zein composition of protein bodies during endosperm development. *Plant Cell* **1989**, *1*, 1011–1023. [[CrossRef](#)]
90. Coleman, C.E.; Herman, E.M.; Takasaki, K.; Larkins, B.A. The maize  $\gamma$ -zein sequesters  $\alpha$  zeins and stabilizes its accumulation in protein bodies of transgenic tobacco endosperm. *Plant Cell* **1996**, *8*, 2335–2345. [[CrossRef](#)]
91. Kogan, M.J.; Dalcol, I.; Gorostiza, P.; Lopez-Iglesias, C.; Pons, R.; Pons, M.; Sanz, F.; Giralt, E. Supramolecular properties of the proline-rich gamma-zein N-terminal domain. *Biophys. J.* **2002**, *83*, 1194–1204. [[CrossRef](#)]
92. Torrent, M.; Llop-Tous, I.; Ludevid, M.D. Protein body induction: A new tool to produce and recover recombinant proteins in plants. In *Recombinant Proteins from Plants. Methods in Molecular Biology*; Gomord, V., Faye, L., Eds.; Springer: Heidelberg/Berlin, Germany; Humana Press: Totowa, NJ, USA, 2009; Volume 483, pp. 193–208. [[CrossRef](#)]

93. Torrent, M.; Llompарт, B.; Lasserre-Ramassamy, S.; Llop-Tous, I.; Bastida, M.; Marzabal, P.; Westernholm-Parvinen, A.; Saloheimo, M.; Htietz, P.B.; Ludevid, M.D. Eukaryotic protein production in designed storage organelles. *BMC Biol.* **2009**, *7*, 5. [[CrossRef](#)]
94. De Virgilio, M.; Marchis, D.; Bellucci, F.; Mainieri, M.; Rossi, D.; Benvenuto, M.; Arcioni, S.; Vitale, A. The human immunodeficiency virus antigen Nef forms protein bodies in leaves of transgenic tobacco when fused to zeolin. *J. Exp. Bot.* **2008**, *59*, 2815–2829. [[CrossRef](#)] [[PubMed](#)]
95. Llop-Tous, I.; Madurga, S.; Giralt, E.; Marzabal, P.; Torrent, M.; Ludevid, M.D. Relevant elements of a maize gamma-zein domain involved in protein body biogenesis. *J. Biol. Chem.* **2010**, *285*, 35633–35644. [[CrossRef](#)] [[PubMed](#)]
96. Llop-Tous, I.; Ortiz, M.; Torrent, M.; Ludevid, M.D. The expression of a xylanase targeted to ER-protein bodies provides a simple strategy to produce active insoluble enzyme polymers in tobacco plants. *PLoS ONE* **2011**, *6*, e19474. [[CrossRef](#)] [[PubMed](#)]
97. Whitehead, M.; Ohlschlager, P.; Almajhdi, F.N.; Alloza, L.; Marzabal, P.; Meyers, A.E.; Hitzeroth, I.I.; Rybicki, E.P. Human papillomavirus (HPV) type 16 E7 protein bodies cause tumour regression in mice. *BMC Cancer* **2014**, *14*, 367. [[CrossRef](#)]
98. Conley, A.J.; Joensuu, J.J.; Jevnikar, A.M.; Menassa, R.; Brandle, J.E. Optimization of elastin-like polypeptide fusions for expression and purification of recombinant proteins in plants. *Biotechnol. Bioeng.* **2009**, *103*, 562–573. [[CrossRef](#)] [[PubMed](#)]
99. Roberts, S.; Dzuricky, M.; Chilkoti, A. Elastin-like polypeptides as models of intrinsically disordered proteins. *FEBS Lett.* **2015**, *589*, 2477–2486. [[CrossRef](#)] [[PubMed](#)]
100. Patel, J.; Zhu, H.; Menassa, R.; Gyenis, L.; Richman, A.; Brandle, J. Elastin-like polypeptide fusions enhance the accumulation of recombinant proteins in tobacco leaves. *Transgenic Res.* **2007**, *16*, 239–249. [[CrossRef](#)]
101. Kaldis, A.; Ahmad, A.; Reid, A.; McGarvey, B.; Brandle, J.; Ma, S.; Jevnikar, A.; Kohalmi, S.E.; Menassa, R. High-level production of human interleukin-10 fusions in tobacco cell suspension cultures. *Plant Biotechnol. J.* **2013**, *11*, 535–545. [[CrossRef](#)]
102. Floss, D.M.; Sack, M.; Stadlmann, J.; Rademacher, T.; Scheller, J.; Stoger, E.; Fischer, R.; Conrad, U. Biochemical and functional characterization of anti-HIV antibody-ELP fusion proteins from transgenic plants. *Plant Biotechnol. J.* **2008**, *6*, 379–391. [[CrossRef](#)]
103. Saberianfar, R.; Joensuu, J.J.; Conley, A.J.; Menassa, R. Protein body formation in leaves of *Nicotiana benthamiana*: A concentration-dependent mechanism influenced by the presence of fusion tags. *Plant Biotechnol. J.* **2015**, *13*, 927–937. [[CrossRef](#)]
104. Linder, M.B.; Szilvay, G.R.; Nakari-Setälä, T.; Penttilä, M.E. Hydrophobins: The protein-amphiphiles of filamentous fungi. *FEMS Microbiol. Rev.* **2005**, *29*, 877–896. [[CrossRef](#)] [[PubMed](#)]
105. Linder, M.B.; Qiao, M.; Laumen, F.; Selber, K.; Hyytiä, T.; Nakari-Setälä, T.; Penttilä, M.E. Efficient purification of recombinant proteins using hydrophobins as tags in surfactant-based two-phase systems. *Biochemistry* **2004**, *43*, 11873–11882. [[CrossRef](#)] [[PubMed](#)]
106. Joensuu, J.J.; Conley, A.J.; Lienemann, M.; Brandle, J.E.; Linder, M.B.; Menassa, R. Hydrophobin fusions for high-level transient protein expression and purification in *Nicotiana benthamiana*. *Plant Physiol.* **2010**, *152*, 622–633. [[CrossRef](#)]
107. Takaiwa, F.; Wakasa, Y.; Hayashi, S.; Kawakatsu, T. An overview on the strategies to exploit rice endosperm as production platform for biopharmaceuticals. *Plant Sci.* **2017**, *263*, 201–209. [[CrossRef](#)] [[PubMed](#)]
108. Viegas, V.S.M.; Ocampo, C.G.; Petruccioli, S. Vacuolar deposition of recombinant proteins in plant vegetative organs as a strategy to increase yields. *Bioengineered* **2017**, *8*, 203–211. [[CrossRef](#)]
109. Ocampo, C.G.; Lareu, J.F.; Viegas, V.S.M.; Mangano, S.; Loos, A.; Steinkellner, H.; Petruccioli, S. Vacuolar targeting of recombinant antibodies in *Nicotiana benthamiana*. *Plant Biotechnol. J.* **2016**, *14*, 2265–2275. [[CrossRef](#)] [[PubMed](#)]
110. Shaaltiel, Y.; Bartfeld, D.; Hashmueli, S.; Baum, G.; Brill-Almon, E.; Galili, G.; Dym, O.; Boldin-Adamsky, S.A.; Silman, I.; Sussman, J.L.; 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.* **2007**, *5*, 579–590. [[CrossRef](#)]
111. Pereira, E.O.; Kolotilin, I.; Conley, A.J.; Menassa, R. Production and characterization of in planta transiently produced polygalacturonase from *Aspergillus niger* and its fusions with hydrophobin or ELP tags. *BMC Biotechnol.* **2014**, *14*, 59. [[CrossRef](#)]
112. Palaniswamy, H.; Syamaladevi, D.P.; Mohan, C.; Philip, A.; Petchiyappan, A.; Narayanan, S. Vacuolar targeting of r-proteins in sugarcane leads to higher levels of purifiable commercially equivalent recombinant proteins in cane juice. *Plant Biotechnol. J.* **2016**, *14*, 791–807. [[CrossRef](#)]
113. Neumann, U.; Brandizzi, F.; Hawes, C. Protein transport in plant cells: In and out of the Golgi. *Ann. Bot.* **2003**, *92*, 167–180. [[CrossRef](#)]
114. Drakakaki, G.; Dandekar, A. Protein secretion: How many secretory routes does a plant cell have? *Plant Sci.* **2013**, *203*, 74–78. [[CrossRef](#)] [[PubMed](#)]
115. Park, E.; Rapoport, T.A. Mechanisms of Sec61/SecY-mediated protein translocation across membranes. *Annu. Rev. Biophys.* **2012**, *41*, 21–40. [[CrossRef](#)] [[PubMed](#)]
116. Alexandersson, E.; Ali, A.; Resjö, S.; Andreasson, E. Plant secretome proteomics. *Front. Plant Sci.* **2013**, *4*, 9. [[CrossRef](#)] [[PubMed](#)]
117. Hwang, I.; Robinson, D.G. Transport vesicle formation in plant cells. *Curr. Opin. Plant Biol.* **2009**, *12*, 660–669. [[CrossRef](#)] [[PubMed](#)]
118. Santos, R.B.; Abranches, R.; Fischer, R.; Sack, M.; Holland, T. Putting the spotlight back on plant suspension cultures. *Front. Plant Sci.* **2016**, *7*, 297. [[CrossRef](#)] [[PubMed](#)]
119. Klimyuk, V.; Pogue, G.; Herz, S.; Butler, J.; Haydon, H. Production of recombinant antigens and antibodies in *Nicotiana benthamiana* using ‘magniflection’ technology: GMP-compliant facilities for small-and large-scale manufacturing. In *Plant viral*

- vectors; Palmer, K., Gleba, Y., Eds.; Current Topics in Microbiology and Immunology; Springer: Heidelberg/Berlin, Germany, 2012; volume 375, pp. 127–154. [\[CrossRef\]](#)
120. Sariyatun, R.; Florence; Kajiura, H.; Ohashi, T.; Misaki, R.; Fujiyama, K. Production of human acid-alpha glucosidase with a paucimannose structure by glycoengineered arabidopsis cell culture. *Front. Plant Sci.* **2021**, *12*, 703020. [\[CrossRef\]](#)
121. Huang, L.F.; Tan, C.C.; Yeh, J.F.; Liu, H.Y.; Liu, Y.K.; Ho, S.L.; Lu, C.A. Efficient secretion of recombinant proteins from rice suspension-cultured cells modulated by the choice of signal peptide. *PLoS ONE* **2015**, *10*, e0140812. [\[CrossRef\]](#)
122. Margulis, L. *Symbiosis in Cell Evolution: Life and Its Environment on the Early Earth*; Freeman Press: San Francisco, CA, USA, 1981. [\[CrossRef\]](#)
123. Wiedemann, N.; Pfanner, N. Mitochondrial machineries for protein import and assembly. *Annu. Rev. Biochem.* **2017**, *86*, 685–714. [\[CrossRef\]](#)
124. Dobrogojski, J.; Adamiec, M.; Luciński, R. The chloroplast genome: A review. *Acta Physiol. Plant.* **2020**, *42*, 98. [\[CrossRef\]](#)
125. Zoschke, R.; Bock, R. Chloroplast translation: Structural and functional organization, operational control, and regulation. *Plant Cell* **2018**, *30*, 745–770. [\[CrossRef\]](#)
126. Schleiff, E.; Becker, T. Common ground for protein translocation: Access control for mitochondria and chloroplasts. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 48–59. [\[CrossRef\]](#) [\[PubMed\]](#)
127. Ge, C.; Spinning, E.; Glaser, E.; Wieslander, E. Import determinants of organelle-specific and dual targeting peptides of mitochondria and chloroplasts in *Arabidopsis thaliana*. *Mol. Plant.* **2014**, *7*, 121–136. [\[CrossRef\]](#)
128. Yogev, O.; Karniely, S.; Pines, O. Translation-coupled translocation of yeast fumarase into mitochondria in vivo. *J. Biol. Chem.* **2007**, *282*, 29222–29229. [\[CrossRef\]](#) [\[PubMed\]](#)
129. Eliyahu, E.; Pnueli, L.; Melamed, D.; Scherrer, T.; Gerber, A.P.; Pines, O.; Rapaport, D.; Arava, Y. Tom20 mediates localization of mRNAs to mitochondria in a translation-dependent manner. *Mol. Cell. Biol.* **2010**, *30*, 284–294. [\[CrossRef\]](#) [\[PubMed\]](#)
130. Weis, B.L.; Schleiff, E.; Zerges, W. Protein targeting to subcellular organelles via mRNA localization. *Biochim. Biophys. Acta (BBA)-Mol. Cell Res.* **2013**, *1833*, 260–273. [\[CrossRef\]](#)
131. Lashkevich, K.A.; Dmitriev, S.E. mRNA targeting, transport and local translation in eukaryotic cells: From the classical view to a diversity of new concepts. *Mol. Biol.* **2021**, *55*, 507–537. [\[CrossRef\]](#)
132. Maity, S.; Chakrabarti, O. Mitochondrial protein import as a quality control sensor. *Biol. Cell* **2021**, *113*, 375–400. [\[CrossRef\]](#)
133. Tucker, K.; Park, E. Cryo-EM structure of the mitochondrial protein-import channel TOM complex at near-atomic resolution. *Nat. Struct. Mol. Biol.* **2019**, *26*, 1158–1166. [\[CrossRef\]](#)
134. Edwards, R.; Gerlich, S.; Tokatlidis, K. The biogenesis of mitochondrial intermembrane space proteins. *Biol. Chem.* **2020**, *401*, 737–747. [\[CrossRef\]](#)
135. Becker, T.; Song, J.; Pfanner, N. Versatility of preprotein transfer from the cytosol to mitochondria. *Trends Cell. Biol.* **2019**, *29*, 534–548. [\[CrossRef\]](#)
136. Bykov, Y.S.; Rapaport, D.; Herrmann, J.M.; Schuldiner, M. Cytosolic events in the biogenesis of mitochondrial proteins. *Trends Biochem. Sci.* **2020**, *45*, 650–667. [\[CrossRef\]](#) [\[PubMed\]](#)
137. Backes, S.; Bykov, Y.S.; Flohr, T.; Räschele, M.; Zhou, J.; Lenhard, S.; Krämer, L.; Mühlhaus, T.; Bibi, C.; Jann, C.; et al. The chaperone-binding activity of the mitochondrial surface receptor Tom70 protects the cytosol against mitoprotein-induced stress. *Cell Rep.* **2021**, *35*, 108936. [\[CrossRef\]](#) [\[PubMed\]](#)
138. Balchin, D.; Hayer-Hartl, M.; Hartl, F.U. In vivo aspects of protein folding and quality control. *Science* **2016**, *353*, aac4354. [\[CrossRef\]](#) [\[PubMed\]](#)
139. Craig, E.A.; Marszalek, J. How do J-proteins get Hsp70 to do so many different things? *Trends Biochem. Sci.* **2017**, *42*, 355–368. [\[CrossRef\]](#)
140. Opalinski, L.; Song, J.; Priesnitz, C.; Wenz, L.S.; Oeljeklaus, S.; Warscheid, B.; Pfanner, N.; Becker, T. Recruitment of cytosolic J-proteins by TOM receptors promotes mitochondrial protein biogenesis. *Cell Rep.* **2018**, *25*, 2036–2043. [\[CrossRef\]](#)
141. Hansen, K.G.; Aviram, N.; Laborenz, J.; Bibi, C.; Meyer, M.; Spang, A.; Schuldiner, M.; Herrmann, J.M. An ER surface retrieval pathway safeguards the import of mitochondrial membrane proteins in yeast. *Science* **2018**, *361*, 1118–1122. [\[CrossRef\]](#)
142. Gold, V.A.; Chroscicki, P.; Bragoszewski, P.; Chacinska, A. Visualization of cytosolic ribosomes on the surface of mitochondria by electron cryo-tomography. *EMBO Rep.* **2017**, *18*, 1786–1800. [\[CrossRef\]](#)
143. Williams, C.C.; Jan, C.H.; Weissman, J.S. Targeting and plasticity of mitochondrial proteins revealed by proximity-specific ribosome profiling. *Science* **2014**, *346*, 748–751. [\[CrossRef\]](#)
144. Marc, P.; Margeot, A.; Devaux, F.; Blugeon, C.; Corral-Debrinski, M.; Jacq, C. Genome-wide analysis of mRNAs targeted to yeast mitochondria. *EMBO Rep.* **2002**, *3*, 159–164. [\[CrossRef\]](#)
145. Lapointe, C.P.; Stefely, J.A.; Jochem, A.; Hutchins, P.D.; Wilson, G.M.; Kwiecién, N.W.; Coon, J.J.; Wickens, M.; Pagliarini, D.J. Multi-omics reveal specific targets of the RNA-binding protein Puf3p and its orchestration of mitochondrial biogenesis. *Cell Syst.* **2018**, *6*, 125–135. [\[CrossRef\]](#)
146. Saint-Georges, Y.; Garcia, M.; Delaveau, T.; Jourden, L.; Le Crom, S.; Lemoine, S.; Tanty, V.; Devaux, F.; Jacq, C. Yeast mitochondrial biogenesis: A role for the PUF RNA-binding protein Puf3p in mRNA localization. *PLoS ONE* **2008**, *3*, e2293. [\[CrossRef\]](#) [\[PubMed\]](#)
147. Møller, I.M.; Rasmusson, A.G.; Van Aken, O. Plant mitochondria—past, present and future. *Plant J.* **2021**, *108*, 912–959. [\[CrossRef\]](#) [\[PubMed\]](#)

148. Lesnik, C.; Cohen, Y.; Atir-Lande, A.; Schuldiner, M.; Arava, Y. OM14 is a mitochondrial receptor for cytosolic ribosomes that supports co-translational import into mitochondria. *Nat. Commun.* **2014**, *5*, 5711. [[CrossRef](#)]
149. Ponce-Rojas, J.C.; Avendaño-Monsalve, M.C.; Yañez-Falcón, A.R.; Jaimes-Miranda, F.; Garay, E.; Torres-Quiroz, F.; DeLuna, A.; Funes, S. ab<sup>+</sup>-NAC cooperates with Sam37 to mediate early stages of mitochondrial protein import. *FEBS J.* **2017**, *284*, 814–830. [[CrossRef](#)]
150. Lee, D.W.; Hwang, I. Evolution and design principles of the diverse chloroplast transit peptides. *Mol. Cells* **2018**, *41*, 161–167. [[CrossRef](#)] [[PubMed](#)]
151. Park, M.H.; Zhong, R.; Lamppa, G. Chloroplast stromal processing peptidase activity is modulated by transit peptide determinants that include inhibitory roles for its N-terminal domain and initial Met. *Biochem. Biophys. Res. Commun.* **2018**, *503*, 3149–3154. [[CrossRef](#)]
152. Bölder, B. En route into chloroplasts: Preproteins' way home. *Photosyn. Res.* **2018**, *138*, 263–275. [[CrossRef](#)]
153. Day, P.M.; Theg, S.M. Evolution of protein transport to the chloroplast envelope membranes. *Photosyn. Res.* **2018**, *138*, 315–326. [[CrossRef](#)]
154. Chotewutmontri, P.; Bruce, B.D. Non-native, N-terminal Hsp70 molecular motor recognition elements in transit peptides support plastid protein translocation. *J. Biol. Chem.* **2015**, *290*, 7602–7621. [[CrossRef](#)]
155. Chotewutmontri, P.; Holbrook, K.; Bruce, B.D. Plastid protein targeting: Preprotein recognition and translocation. *Int. Rev. Cell Mol. Biol.* **2017**, *330*, 227–294. [[CrossRef](#)]
156. Chang, W.L.; Soll, J.; Bölder, B. The gateway to chloroplast: Re-defining the function of chloroplast receptor proteins. *Biol. Chem.* **2012**, *393*, 1263–1277. [[CrossRef](#)]
157. Wiesemann, K.; Simm, S.; Mirus, O.; Ladig, R.; Schleiff, E. Regulation of two GTPases Toc159 and Toc34 in the translocon of the outer envelope of chloroplasts. *Biochim. Biophys. Acta* **2019**, *1867*, 627–636. [[CrossRef](#)] [[PubMed](#)]
158. Ganesan, I.; Theg, S.M. Structural considerations of folded protein import through the chloroplast TOC/TIC translocons. *FEBS Lett.* **2019**, *593*, 565–572. [[CrossRef](#)] [[PubMed](#)]
159. Richardson, L.G.L.; Small, E.L.; Inoue, H.; Schnell, D.J. Molecular topology of the transit peptide during chloroplast protein import. *Plant Cell* **2018**, *30*, 1789–1806. [[CrossRef](#)] [[PubMed](#)]
160. Chen, Y.L.; Chen, L.J.; Chu, C.C.; Huang, P.K.; Wen, J.R.; Li, H.M. TIC236 links the outer and inner membrane translocons of the chloroplast. *Nature* **2018**, *564*, 125–129. [[CrossRef](#)]
161. Rudolf, M.; Machettira, A.B.; Gross, L.E.; Weber, R.L.; Bolte, K.; Bionda, T.; Sommer, M.S.; Maier, U.G.; Weber, A.P.M.; Shleiff, E.; et al. In vivo function of Tic22, a protein import component of the intermembrane space of chloroplasts. *Mol. Plant* **2013**, *6*, 817–829. [[CrossRef](#)]
162. Kasmati, A.R.; Topel, M.; Patel, R.; Murtaza, G.; Jarvis, P. Molecular and genetic analyses of Tic20 homologues in *Arabidopsis thaliana* chloroplasts. *Plant J.* **2011**, *66*, 877–889. [[CrossRef](#)]
163. Flores-Perez, U.; Jarvis, P. Molecular chaperone involvement in chloroplast protein import. *Biochim. Biophys. Acta* **2013**, *1833*, 332–340. [[CrossRef](#)]
164. Uniacke, J.; Zerges, W. Chloroplast protein targeting involves localized translation in *Chlamydomonas*. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 1439–1444. [[CrossRef](#)]
165. Sun, Y.; Bakhtiari, S.; Valente-Paterno, M.; Wu, Y.; Law, C.; Dai, D.; Dhaliwal, J.; Bui, K.H.; Zerges, W. Chloroplast-localized translation for protein targeting in *Chlamydomonas reinhardtii*. *bioRxiv* **2021**. [[CrossRef](#)]
166. Villarejo, A.; Buren, S.; Larsson, S.; Dejardin, A.; Monne, M.; Rudhe, C.; Karlsson, J.; Jansson, S.; Lerouge, P.; Rolland, N.; et al. Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. *Nat. Cell Biol.* **2005**, *7*, 1224–1231. [[CrossRef](#)]
167. Buren, S.; Ortega-Villasante, C.; Blanco-Rivero, A.; Martinez-Bernardini, A.; Shutova, T.; Shevela, D.; Messinger, J.; Bako, L.; Villarejo, A.; Samuelsson, G. Importance of post-translational modifications for functionality of a chloroplast-localized carbonic anhydrase (CAH1) in *Arabidopsis thaliana*. *PLoS ONE* **2011**, *6*, e21021. [[CrossRef](#)]
168. Bhushan, S.; Kuhn, C.; Berglund, A.K.; Roth, C.; Glaser, E. The role of the N-terminal domain of chloroplast targeting peptides in organellar protein import and miss-sorting. *FEBS Lett.* **2006**, *580*, 3966–3972. [[CrossRef](#)]
169. Lee, D.W.; Lee, S.; Lee, J.; Woo, S.; Razzak, M.A.; Vitale, A.; Hwang, I. Molecular mechanism of the specificity of protein import into chloroplasts and mitochondria in plant cells. *Mol. Plant* **2019**, *12*, 951–966. [[CrossRef](#)]
170. Lee, D.W.; Hwang, I. Understanding the evolution of endosymbiotic organelles based on the targeting sequences of organellar proteins. *New Phytol.* **2021**, *230*, 924–930. [[CrossRef](#)]
171. Garg, S.G.; Gould, S.B. The role of charge in protein targeting evolution. *Trends Cell Biol.* **2016**, *26*, 894–905. [[CrossRef](#)]
172. Sharma, M.; Bennewitz, B.; Klösgen, R.B. Rather rule than exception? How to evaluate the relevance of dual protein targeting to mitochondria and chloroplasts. *Photosynth. Res.* **2018**, *138*, 335–343. [[CrossRef](#)]
173. McKinnon, L.; Theg, S.M. Determinants of the specificity of protein targeting to chloroplasts or mitochondria. *Mol. Plant* **2019**, *12*, 893–895. [[CrossRef](#)]
174. Bruce, B.D. Chloroplast transit peptides: Structure, function and evolution. *Trends Cell Biol.* **2000**, *10*, 440–447. [[CrossRef](#)]
175. Daras, G.; Rigas, S.; Tsitsekian, D.; Zur, H.; Tuller, T.; Hatzopoulos, P. Alternative transcription initiation and the AUG context configuration control dual-organellar targeting and functional competence of *Arabidopsis* Lon1 protease. *Mol. Plant* **2014**, *7*, 989–1005. [[CrossRef](#)]

176. Ye, W.; Spanning, E.; Glaser, E.; Maler, L. Interaction of the dual targeting peptide of Thr-tRNA synthetase with the chloroplastic receptor Toc34 in *Arabidopsis thaliana*. *FEBS Open Biol.* **2015**, *5*, 405–412. [[CrossRef](#)]
177. Lee, D.W.; Lee, S.; Min, C.K.; Park, C.; Kim, J.M.; Hwang, C.S.; Park, S.K.; Cho, N.-H.; Hwang, I. Cross-species functional conservation and possible origin of the N-terminal specificity domain of mitochondrial presequences. *Front. Plant Sci.* **2020**, *11*, 64. [[CrossRef](#)]
178. Garrido, C.; Caspari, O.D.; Choquet, Y.; Wollman, F.A.; Lafontaine, I. Evidence supporting an antimicrobial origin of targeting peptides to endosymbiotic organelles. *Cells* **2020**, *9*, 1795. [[CrossRef](#)]
179. Caspari, O.D.; Lafontaine, I. The role of antimicrobial peptides in the evolution of endosymbiotic protein import. *PLoS Pathog.* **2021**, *17*, e1009466. [[CrossRef](#)]
180. Rascón-Cruz, Q.; González-Barriga, C.D.; Iglesias-Figueroa, B.F.; Trejo-Muñoz, J.C.; Siqueiros-Cendón, T.; Sinagawa-García, S.R.; Alevaro-Gallegos, S.; Espinoza-Sánchez, E.A. Plastid transformation: Advances and challenges for its implementation in agricultural crops. *Electron. J. Biotechnol.* **2021**, *51*, 95–109. [[CrossRef](#)]
181. Rozov, S.M.; Sidorchuk, Y.V.; Deineko, E.V. Transplastomic plants: Problems of production and their solution. *Russ. J. Plant Physiol.* **2022**, *69*, 20. [[CrossRef](#)]
182. Nishimura, K.; Kato, Y.; Sakamoto, W. Chloroplast proteases: Updates on proteolysis within and across suborganellar compartments. *Plant Physiol.* **2016**, *171*, 2280–2293. [[CrossRef](#)]
183. Rybicki, E.P. Plant-made vaccines for humans and animals. *Plant Biotechnol. J.* **2010**, *8*, 620–637. [[CrossRef](#)]
184. Sticklen, M. Plant genetic engineering to improve biomass characteristics for biofuels. *Curr. opinion biotechnol.* **2006**, *17*, 315–319. [[CrossRef](#)]
185. Hyunjong, B.; Lee, D.-S.; Hwang, I. Dual targeting of xylanase to chloroplasts and peroxisomes as a means to increase protein accumulation in plant cells. *J. Exp. Bot.* **2005**, *57*, 161–169. [[CrossRef](#)]
186. Kavanagh, T.A.; Jefferson, R.A.; Bevan, M.W. Targeting a foreign protein to chloroplasts using fusions to the transit peptide of a chlorophyll *a/b* protein. *Mol. Gen. Genet.* **1988**, *215*, 38–45. [[CrossRef](#)]
187. Köhler, R.H.; Cao, J.; Zipfel, W.R.; Webb, W.W.; Hanson, M.R. Exchange of protein molecules through connections between higher plant plastids. *Science* **1997**, *276*, 2039–2042. [[CrossRef](#)]
188. Panstruga, R.; Hippe-Sanwald, S.; Lee, Y.-K.; Lataster, M.; Lipka, V.; Fischer, R.; Liao, Y.C.; Häusler, R.E.; Kreuzaler, F.; Hirsch, H.-J. Expression and chloroplast-targeting of active phosphoenolpyruvate synthetase from *Escherichia coli* in *Solanum tuberosum*. *Plant Sci.* **1997**, *127*, 191–205. [[CrossRef](#)]
189. Jang, I.-C.; Nahm, B.H.; Kim, J.-K. Subcellular targeting of green fluorescent protein to plastids in transgenic rice plants provides a high-level expression system. *Mol. Breed.* **1999**, *5*, 453–461. [[CrossRef](#)]
190. Kim, E.H.; Suh, S.C.; Park, B.S.; Shin, K.S.; Kweon, S.J.; Han, E.J.; Park, S.H.; Kim, Y.S.; Kim, J.K. Chloroplast-targeted expression of synthetic cry1Ac in transgenic rice as an alternative strategy for increased pest protection. *Planta* **2009**, *230*, 397–405. [[CrossRef](#)]
191. Kim, S.; Lee, D.-S.; Choi, I.S.; Ahn, S.-J.; Kim, Y.-H.; Bae, H.-J. *Arabidopsis thaliana* Rubisco small subunit transit peptide increases the accumulation of *Thermotoga maritima* endoglucanase Cel5A in chloroplasts of transgenic tobacco plants. *Transgenic Res.* **2010**, *19*, 489–497. [[CrossRef](#)]
192. Meyers, A.; Chakauya, E.; Shephard, E.; Tanzer, F.L.; Maclean, J.; Lynch, A.; Williamson, A.-L.; Rybicki, E.P. Expression of HIV-1 antigens in plants as potential subunit vaccines. *BMC Biotechnol.* **2008**, *8*, 53. [[CrossRef](#)]
193. Maclean, J.; Koekemoer, M.; Olivier, A.J.; Stewart, D.; Hitzeroth, I.I.; Rademacher, T.; Fischer, R.; Williamson, A.L.; Rybicki, E.P. Optimization of human papillomavirus type 16 (HPV-16) L1 expression in plants: Comparison of the suitability of different HPV-16 L1 gene variants and different cell-compartment localization. *J. Gen. Virol.* **2007**, *88*, 1460–1469. [[CrossRef](#)]
194. Zahin, M.; Joh, J.; Khanal, S.; Husk, A.; Mason, H.; Warzecha, H.; Ghim, S.J.; Miller, D.M.; Matoba, N.; Jenson, A.B. Scalable production of HPV16 L1 protein and VLPs from tobacco leaves. *PLoS ONE* **2016**, *11*, e0160995. [[CrossRef](#)]
195. Yanez, R.J.R.; Lamprecht, R.; Granadillo, M.; Torrens, I.; Arcalis, E.; Stoger, E.; Rybicki, E.P.; Hitzeroth, I.I. LALF32-51-E7, a HPV-16 therapeutic vaccine candidate, forms protein body-like structures when expressed in *Nicotiana benthamiana* leaves. *Plant Biotechnol. J.* **2018**, *16*, 628–637. [[CrossRef](#)]
196. Buchke, S.; Sharma, M.; Bora, A.; Relekar, M.; Bhanu, P.; Kumar, J. Mitochondria-targeted, nanoparticle-based drug-delivery systems: Therapeutics for mitochondrial disorders. *Life* **2022**, *12*, 657. [[CrossRef](#)]
197. Palmer, C.S.; Anderson, A.J.; Stojanovski, D. Mitochondrial protein import dysfunction: Mitochondrial disease, neurodegenerative disease and cancer. *FEBS Lett.* **2021**, *595*, 1107–1131. [[CrossRef](#)]
198. Atkin, O.K.; Macherel, D. The crucial role of plant mitochondria in orchestrating drought tolerance. *Ann. Bot.* **2009**, *103*, 581–597. [[CrossRef](#)]
199. Lopez-Torrejón, G.; Jimenez-Vicente, E.; Buesa, J.M.; Hernandez, J.A.; Verma, H.K.; Rubio, L.M. Expression of a functional oxygen-labile nitrogenase component in the mitochondrial matrix of aerobically grown yeast. *Nat. Commun.* **2016**, *7*, 11426. [[CrossRef](#)]
200. Baysal, C.; Pérez-González, A.; Eseverri, Á.; Jiang, X.; Medina, V.; Caro, E.; Rubio, L.; Christou, P.; Zhu, C. Recognition motifs rather than phylogenetic origin influence the ability of targeting peptides to import nuclear-encoded recombinant proteins into rice mitochondria. *Transgenic Res.* **2020**, *29*, 37–52. [[CrossRef](#)]

201. Jiang, X.; Coroian, D.; Barahona, E.; Echavarri-Erasun, C.; Castellanos-Rueda, R.; Eneverri, Á.; Aznar-Moreno, J.A.; Buren, S.; Rubio, L.M. Functional nitrogenase cofactor maturase NifB in mitochondria and chloroplasts of *Nicotiana benthamiana*. *mBio* **2022**, *13*, e00268–22. [[CrossRef](#)]
202. Emanuelsson, O.; Nielsen, H.; Brunak, S.; von Heijne, G. Predicting subcellular localization of proteins based on their n-terminal amino acid sequence. *J. Mol. Biol.* **2000**, *300*, 1005–1016. [[CrossRef](#)]
203. Small, I.; Peeters, N.; Legeai, F.; Lurin, C. Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* **2004**, *4*, 1581–1590. [[CrossRef](#)]
204. Petsalaki, E.I.; Bagos, P.G.; Litou, Z.I.; Hamodrakas, S.J. PredSL: A tool for the N-terminal sequence-based prediction of protein subcellular localization. *Genom. Proteom. Bioinf.* **2006**, *4*, 48–55. [[CrossRef](#)]
205. Fukasawa, Y.; Tsuji, J.; Fu, S.-C.; Tomii, K.; Horton, P.; Imai, K. MitoFates: Improved prediction of mitochondrial targeting sequences and their cleavage sites. *Mol. Cell Proteom.* **2015**, *14*, 1113–1126. [[CrossRef](#)]
206. Savojardo, C.; Martelli, P.L.; Fariselli, P.; Casadio, R. TPpred3 detects and discriminates mitochondrial and chloroplastic targeting peptides in eukaryotic proteins. *Bioinformatics* **2015**, *31*, 3269–3275. [[CrossRef](#)]
207. Jiang, Y.; Wang, D.; Wang, W.; Xu, D. Computational methods for protein localization prediction. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 5834–5844. [[CrossRef](#)]
208. Nielsen, H.; Tsigirigos, K.D.; Brunak, S.; von Heijne, G. A brief history of protein sorting prediction. *Protein J.* **2019**, *38*, 200–216. [[CrossRef](#)]
209. Teufel, F.; Armenteros, J.J.A.; Johansen, A.R.; Gíslason, M.H.; Pihl, S.I.; Tsigirigos, K.D.; Winter, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 6.0 achieves signal peptide prediction across all types using protein language models. *bioRxiv* **2021**. [[CrossRef](#)]
210. Armenteros, J.J.A.; Salvatore, M.; Emanuelsson, O.; Winther, O.; Von Heijne, G.; Elofsson, A.; Nielsen, H. Detecting sequence signals in targeting peptides using deep learning. *Life Sci. Alliance* **2019**, *2*, e201900429. [[CrossRef](#)]
211. Thummuluri, V.; Armenteros, J.J.A.; Johansen, A.R.; Nielsen, H.; Winther, O. DeepLoc 2.0: Multi-label subcellular localization prediction using protein language models. *Nucleic Acids Res.* **2022**, *W1*, W228–W234. [[CrossRef](#)]