PLANT SCIENCE TODAY
ISSN 2348-1900 (online)
Vol 12(2): 1-9
https://doi.org/10.14719/pst.6416

HORIZON
e-Publishing Group
HePG

REVIEW ARTICLE

Insights into Microprotein A novel tool to unravel crop improvement

N Premalatha[1], A Thanga Hemavathy[2*], Kumar Mahadevan[2*], D Kavithamani[3], S Kavitha[4], V Vijaya Geetha[5], M Kanimoli Mathivathana[6], S M Samyuktha[7]

[1]Department of Cotton, Tamil Nadu Agricultural University, Coimbatore 641 006, Tamil Nadu, India

[2]Department of Pulses, Tamil Nadu Agricultural University, Coimbatore 641 006, Tamil Nadu, India

[3]Department of Millets, Tamil Nadu Agricultural University, Coimbatore 641 006, Tamil Nadu, India

[4]Centre For Students Welfare , Tamil Nadu Agricultural University, Coimbatore 641 006, Tamil Nadu, India

[5]Krishi Vigyan Kendra, Tamil Nadu Agricultural University, Tindivanam 604 001, Tamil Nadu, India

[6]College of Agricultural Technology, Theni 625 531, Tamil Nadu, India

[7]Department of Genetics and Plant Breeding, Dhanalakshmi Srinivasan Agriculture College, Perambalur 606 302, Tamil Nadu, India

Corresponding authors *Email: hemavathytnau@gmail.com, kumarm@tnau.ac.in

OPEN ACCESS

ARTICLE HISTORY
Received: 29 November 2024
Accepted: 22 January 2025
Available online
Version 1.0 : 04 March 2025
Version 2.0 : 01 April 2025

Check for updates

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

CITE THIS ARTICLE
Premalatha N, Hemavathy AT, Mahadevan K, Kavithamani D, Kavitha S, Geetha VV, Mathivathana MK, Samyuktha SM.Insights into Microprotein A novel tool to unravel crop improvement.Plant Science Today. 2025; 12(2): 1-9. https://doi.org/10.14719/pst.6416

Abstract

Small regulatory proteins with a size range of 5 to 20 kilodaltons (kDa) are known as microproteins (miPs). They are connected to bigger, frequently multi-domain proteins and typically include a single protein domain. Through their interactions with other proteins, these microproteins modify the post-translational gene expression level. Numerous microproteins that are essential for controlling transcription factor activity have been discovered in both plants and animals in recent years. Microproteins are necessary for several phases of plant development, such as seed germination, seedling growth, stomatal regulation, root formation, pigment synthesis, blooming and floral development. Certain microproteins *viz.,* viral protein U (Vpu) microProtein**,** negatively regulates the K+ ion channel TASK1 in humans, LITTLE ZIPPER proteins found in arabidopsis which regulate transcription factor and mitochondrial microprotein BRAWNI are conserved only among vertebrates are exclusive to a given species, whilst others have evolved to be conserved since they first appeared early in evolutionary history. Food security is being challenged by the cumulative consequences of climate change and unsustainable agricultural methods, which increases the need for sustainable and innovative solutions since microproteins are essential regulators of several physiological processes in plants. They are excellent candidates for creating synthetic miPs that can be employed to support plant stress resilience leading to increased productivity. Understanding the microproteins' regulatory mechanisms is a crucial step in developing microproteins into useful biotechnological tools for crop bioengineering. There is a theory that target proteins and microproteins have similar evolutionary histories. Microproteins work at the molecular level by obstructing the assembly of higher-order protein complexes. Their potential for biotechnological applications is further enhanced by their ability to function as dominant regulators in a focused and precise manner. In addition to exploring the processes of microproteins and their functional roles in plant biology, this study intends to provide the groundwork for future investigations by helping scientists identify, characterize and map these proteins.

Keywords

Gene regulation; Microprotein; MiP Finder; Protein-protein interaction

Introduction

Microproteins are small proteins that typically range in size from 5 to 20 kilodaltons (kDa). They are also known as alternative proteins or short open reading frame-encoded peptides (SEPs). The only domain in them is in charge of protein-protein interactions (PPI). Evolutionarily, microproteins and the larger multidomain proteins they interact with are thought to be linked. Food security is being challenged by the cumulative consequences of climate change and unsustainable agricultural methods, which increases the need for sustainable and innovative solutions since microproteins are essential regulators of several physiological processes in plants. They are excellent candidates for creating synthetic miPs that can be employed to support plant stress resilience leading to increased productivity (1). According to (2), they are believed to have started when gene duplications and deletions resulted in the creation of single-domain proteins that kept their capacity to interact with their ancestral counterparts. Over the past few decades, nanotechnology has become a blessing for humanity (3). Microproteins are not directly involved in controlling transcription since they do not have DNA-binding or activation domains, in contrast to traditional transcriptional regulators. Nonetheless, they can engage in post-translational interactions with homologous proteins because of their PPI domains. Usually, they work by joining these proteins to form heterodimers, which inactivates their partners (4). A variety of microproteins have been discovered based on their interactions with functional proteins, frequently acting as regulators of transcriptional activity. Much like micronutrients, which are necessary in small amounts but are crucial for various biochemical and physiological functions, microproteins are required in limited quantities yet are essential for particular regulatory roles within cellular processes (5).

The DNA-binding Inhibitor (Id) protein was the first known microprotein (miP), having been found in living things about thirty years ago. This inhibitor was isolated from a cDNA library derived from murine erythroleukemia cells. The protein in question is a tiny 16 kDa helix-loop-helix (HLH) domain protein that interacts with the transcriptional regulator MyoD (Myoblast Determination Protein 1), which is involved in the differentiation of muscles (6). The LITTLE ZIPPER (ZPR) family of proteins, namely ZPR1-4, was the first miP family identified in plants (7). Because ZPR proteins have a leucine zipper domain, they can interact with other leucine zipper-containing proteins and control their post-translational modifications, thereby affecting important developmental processes. ZPR peptides interact with HD-ZIPIII/REV transcription factors to control leaf development and stem cell maintenance in the shoot apical meristem (SAM).

Translation of tiny open reading frames (smORFs) on long non-coding RNAs (lncRNAs) produces a large number of microproteins. Numerous smORFs, which can make up between 5 and 10% of an organism's entire genome, have been found in a variety of kingdoms, including bacteria, plants and animals (8).

2. Mechanism

Research on the regulatory feedback mechanism of class III homeodomain-leucine zipper (HD-ZIPIII) transcription factors, mediated by Little Zipper (ZPR) microproteins, provided the first experimental proof of the existence and function of microproteins (6). The ZPR microprotein family in *Arabidopsis* has four members, designated ZPR1-4, which solely include a leucine zipper domain. LITTLE ZIPPER (ZPR) proteins were the first microproteins characterized in plants (36). ZPR proteins contain a leucine zipper domain but lack other domains required for DNA binding and transcriptional activation. ZPR proteins thus function in analogy to Id-type proteins and physically interact with class III homeodomain-leucine zipper (HD-ZIPIII) transcription factors to control developmental processes such as stem cell maintenance in shoot apical meristem (SAM) formation and leaf development. Many ZPR genes have their transcription directly increased by the HD-ZIPIII transcription factor REVOLUTA. ZPR proteins then physically engage with their HD-ZIPIII targets, preventing them from binding DNA (9). As a result, a direct negative feedback loop is created that modifies plant development by controlling the activity of the shoot apical meristem. HD-ZIPs support the auxin biosynthesis, transport and response of auxin, also an important mediator of embryo development in Arabidopsis (Fig.1).

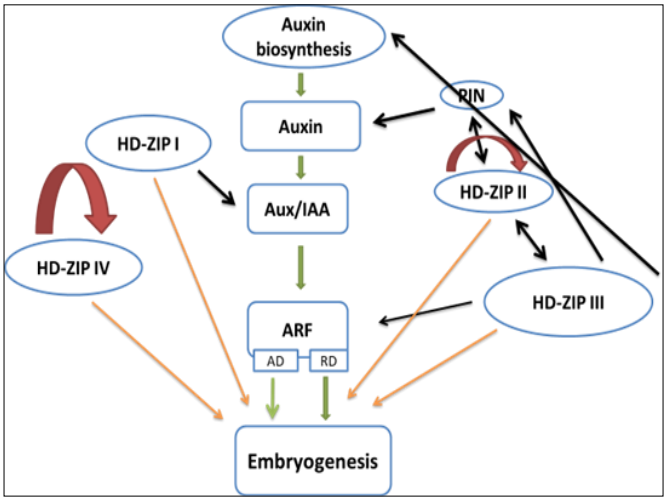

Fig. 1. Regulatory mechanism of HD-ZIPIII transcription factors in *Arabidopsis thaliana* embryogenesis.

Auxin pathways are represented in green, HD-ZIP-dependent regulation of auxin pathways in black and auxin-independent regulation of embryogenesis in orange. Regulation within a class is shown in red (PIN - (PIN-FORMED proteins, ARFs - Auxin Response Factors, AD - Activating Domains and RD - Repressing Domains)

***2.1 Classification of Microprotein* :** Microproteins were categorised as microproteins according to their size and function (10). Depending on where they come from, microproteins are categorised as either trans- or cis-miPs. Trans-miPs are evolved through genome amplification and subsequent domain loss events. They are paralogous to their targets, they come from genes that have been

duplicated. For example, the LITTLE ZIPPER protein possesses a leucine zipper domain and engages in homotypic interactions, regulating class III homeodomain-leucine zipper (HD-ZIPIII) transcription factors. They play a significant role in maintaining stem cells in the shoot apical meristem and influencing leaf development through interaction with the HD-ZIPIIIs. This interaction establishes a negative feedback loop. Cis-miP are produced when mRNA isoforms encoding microproteins are produced by methods, such as alternative splicing and the use of alternative translation start and stop sites. They are derived from the same gene as their target protein. For example, In Arabidopsis, 5′ paired-end analysis of transcription start site sequencing (5'PEAT-Seq) was employed to identify the novel *cis*-miP. One microprotein candidate, named ATHB2miP, showed it comprises a single leucine zipper domain and mimics the phenotype of an *athb2* knockout when over-expressed. Gene expression analysis in plants either lacking ATHB2 or overexpressing *ATHB2miP* revealed common mis regulation patterns, supporting that ATHB2miP functions like a canonical. Furthermore, smaller protein products produced by post-translational processing, such as proteolytic cleavage, can interfere with their bigger, uncleaved precursor proteins. Individual transcription units that are evolutionarily connected to bigger genes encoding multidomain proteins give rise to trans-microproteins. They developed as single-domain inhibitory small proteins as a result of genome amplification events and the subsequent loss of certain domains.

2.2 Identification of microproteins: This is one of the most efficient ways to find microproteins because they are usually found by examining small proteins. The size of microproteins is defined by their range of 7 to 20 kDa, which is the approximate size of one protein domain. It's crucial to remember that not all tiny proteins fall into the microprotein category. Small proteins need to be sorted according to recognised microprotein traits to distinguish real microproteins. They are distinguished from other small proteins by their size, protein domain and evolutionary relationship to their targets. It can interfere with the formation of higher-order protein complexes and alter the biological function of their targets It has a single protein domain, often a protein-protein interaction (PPI) domain. It is difficult to detect in total protein mass, isolate and track with conventional bioengineering tools.

A technology called MiPFinder was created to identify microproteins. miPFinder described a program to identify microProteins, novel regulators of protein function. MiPFinder starts with a set of protein sequences and considers information about protein size, sequence similarity and domain composition to create a list of miP candidates. Small proteins are ubiquitous in all kingdoms of life. MicroProteins, initially characterized as small proteins with protein interaction domains that enable them to interact with larger multidomain proteins, frequently modulate the function of these proteins. To identify microproteins and evaluate their functional potential, it makes use of data on protein size, domain organisation, known protein interactions and evolutionary origin. The ability to identify microproteins across different genomes using this freely available program will help uncover new microproteins in both plants and mammals.

2.3 Interaction with domains of the proteins : Homotopic interaction is the word used to describe the interaction between a microprotein and a protein that shares a similar domain. One instance of this is the regulation of blooming by miP1a/b (also called BBX31/BBX30), which interacts with the proteins TOPLESS and CONSTANS through the B-box motif (11). On the other hand, a heterotypic contact occurs when a microprotein interacts with a protein that has a different but compatible domain. For example, the B -box domain of miP1a/b interacts with the C-terminal DNA -binding domain of ETHYLENE INSENSITIVE 3 (EIN3), the bHLH domain of PHYTOCHROME INTERACTING FACTORs (PIFs) and both to promote photomorphogenic growth (10). Microproteins' regulatory scope is broadened by their capacity to create heterotypic interactions, which enables them to control targets unrelated to their evolutionary history. This widens the functional responsibilities of microproteins.

Until they interact with bigger homologous proteins of greater physiological significance, microproteins (miPs) are usually functionally inert. miPs are essential for regulating protein activity within signalling pathways at the post-translational stage (12). Furthermore, the majority of well-known miPs control transcription factors using homologous domains they share (13) .

Nevertheless, in some organisms, miPs can also interact with and control non-transcription factor proteins. For example, the HIV accessory protein Vpu (Viral Protein U) interacts with the mammalian K^+-channel TASK-1, sequestering and deactivating it (14). Similarly, a long non-coding RNA (lncRNA) unique to skeletal muscle called MYOREGULIN (MLN) generates a microprotein that obstructs a kind of calcium ATPase called SERCA (Sarco/ Endoplasmic Reticulum Ca^{2+}-ATPase), which in turn prevents Ca^{2}+ transport in the sarcoplasmic reticulum (15). Moreover, microproteins can control proteins by encouraging their cytoplasmic retention (16). Microproteins exhibit considerable variability in both their form and mechanisms of action, highlighting their diverse roles in cellular regulation.

2.4 Mapping of microproteins and short peptides in plants : Microproteins (miPs) are biologically significant and versatile, which emphasises the increasing need for techniques to detect and characterise novel miPs. Microproteins are small proteins, but not all small proteins meet the criteria to be called microproteins (16). This means that prospective miPs must be distinguished from other small proteins. Putative miPs and their target proteins can be found using a variety of methods. Since miP genes are small, it's possible that complicated genomes do not have annotations for them, which makes early detection and characterisation difficult.

However, the ability to find miP genes using modifiable sets of criteria has been made possible by advances in computational techniques; these will be covered

in more detail in the following parts. To find putative miPs, multiomics approaches-which include information from transcriptomics, proteomics and genomics-are frequently employed. However, these methods have drawbacks and need more experimental verification (16). Currently, it is expected that high-throughput mass spectrometry, sophisticated computational methods and predictive protein tools will expedite the identification and development of novel microproteins.

A popular method for predicting and identifying microproteins is proteomics, which is based on mass spectrometry. The combination of mass spectrometry and chromatography methods presents a promising avenue for the characterisation of tiny proteins. To find tiny candidate peptides, high-throughput proteomics is widely used, especially using liquid chromatography-tandem mass spectrometry (LC-MS/MS). Protein lysates are commonly employed in this MS methodology to identify proteins using top-down and shotgun proteomics techniques. Careful sample preparation is essential for tiny protein detection, with a focus on efficiently denaturing, reducing and digesting protein samples (17).

Target protein concentration and purification can occasionally be achieved by protein precipitation using organic solvents such as acetone and methanol. SDS-PAGE can also be used to separate small proteins, but special precautions must be taken to ensure that small proteins are not lost during sample processing (18). These methods make it possible to extract patterns of peptide fragmentation, which reveal information about the structure of proteins. To further enrich tiny proteins, the top-down strategy makes use of size exclusion and ion-exchange chromatography.

However, despite these advances, there are still challenges in distinguishing microproteins from small proteins. Several limitations were listed, such as the absence of standardised processes, challenges in extracting particular proteins, optimisation of protein digestion procedures, steps in liquid chromatography, data gathering and the requirement for intricate data processing. Reliability in distinguishing microproteins from other tiny proteins is hampered by these characteristics (15).

The detection of tiny proteins, the analysis of mass-to-charge (m/z) fragment patterns and the construction of protein complex structures are all made possible by mass spectrometry (MS) techniques. Furthermore, MS can be combined with several excellent computational techniques to learn more about the roles played by microproteins (miPs). When mass spectrometry and matrix -assisted laser desorption ionisation (MALDI) are combined, information regarding tiny proteins and their molecular weight can also be directly examined (19).

The development of next-generation sequencing (NGS) and modern algorithms have made it possible to do thorough genomic research to find novel microproteins. Even with these developments, accurately classifying miPs in silico is still difficult. To separate miPs from small proteins, a computational tool called miPFinder needs information from databases about things like genome annotation, protein size, domain details, evolutionary origin, protein organisation and evidence of protein-protein interactions (2).

For sequence alignment and annotation, miPFinder depends on many computational tools, including Python, BLAST, ClustalW2, hmmbuild, Pfam and iPfam. To functionally annotate protein classes based on evolutionary links in plant systems, Gene Ontology (GO) functions are obtained from databases such as Plant GO, agriGO, GOSlimViewer and PANTHER (20).

Another sophisticated technique for identifying short open reading frames (smORFs), many of which might encode miPs, is ribosomal profiling, or ribo-seq (20). To anticipate functional smORFs, new databases and technologies including sORF finder, CRITICA, CPC, PhastCons, PhyloCSF and Micropeptide Detection Pipeline (micPDP) are frequently utilised. Many of these instruments have been used on animals (humans, mice, fruit flies, etc.), but they can also be used to detect microproteins in plants. PEAT, a bioinformatics tool, offers transcription start site (TSS) information for Arabidopsis, which may help with microprotein prediction (22).

Even while smORFs and possible miPs can be predicted using computational and MS techniques, more experimental validation is frequently needed to clarify their regulatory roles. Gain-of-function and loss-of-function mutants are two genetic techniques that are useful for figuring out physiological functions and comprehending miP-target interactions (10,6). Biochemically, immunoprecipitation of FLAG-tagged miPs is a popular technique to identify interacting partners, albeit the tiny size of miPs can lead to nonspecific interactions. Engineered ascorbate peroxidase 2 (APEX) in situ tagging has shown to be a better option, yielding more precise interaction data while utilising less nonspecific binders.

Furthermore, structural analysis, immunoprecipitation, imaging and yeast two-hybrid screenings are examples of genetic, molecular and biochemical approaches that can aid in the prediction of the possible targets and activities of microproteins. In the future, synthetic biology techniques might also shed light on the regulatory roles of miPs (23).

3. Key features of microprotein candidates

Microproteins (miPs) found thus far are typically small, with sizes ranging from 7 to 17 kDa, having less than 120 amino acids (10). To sequester their targets, these proteins only need one functional domain to operate as a platform for protein-protein interactions. Protein-interaction domains range widely in size, although they typically have a maximum length of about 100 amino acids. This means that the majority of known microproteins have lengths under 120 amino acids, while the prediction of novel microproteins frequently starts at 140 amino acids.

An important criterion for finding putative microprotein candidates is the protein organisation of the putative target or

its ancestor protein. As a result, the candidates are guaranteed to possess the structural characteristics required for functional interactions, making them potential miPs with the ability to alter biological processes. This method improves prediction accuracy for novel microproteins in a variety of organisms by taking into account both protein organisation and size.

Experts can monitor the evolutionary history of trans-microproteins since they function as separate transcription units. The ZPR (Little Zipper) proteins found only in plants are a prominent illustration of this; they descended from a larger homeodomain protein known as leucine-zipper (HD-ZIP). ZPR proteins were created by a sequence of changes to this ancestral protein, including gene duplication, degeneration and truncation (24)

With its leucine zipper domain, the original ZPR-ancestor protein could form dimers with itself. It was a multi-domain protein that was capable of homodimerization. ZPR proteins were able to interact and regulate other leucine zipper-containing proteins since they only kept the leucine zipper domain after these alterations. Trans-microproteins can develop from multi-domain proteins into more specialised, single-domain regulators while retaining crucial interaction capabilities with their ancestral counterparts, as demonstrated by this evolutionary shortening process.

The probable ancestor proteins of potential microproteins must be big enough to have two or more functional domains. Because of this, it is common practice to specify a minimum ancestral protein size of 250 amino acids. By removing the possibility of tiny proteins from families whose members are only marginally larger, this criterion guarantees that the ancestral proteins are suitably complicated. Furthermore, the sensitivity of finding distantly related sequences is greatly increased when a consensus sequence of related microprotein candidates is used instead of individual protein sequences (25). The accuracy of detecting putative microprotein precursors is improved by this method. This approach is used by the microProtein-finder tool, which starts by obtaining consensus sequences from every small protein family. This makes it possible for the identification process to be more reliable, especially when looking for evolutionary relationships between small proteins and large protein databases. This improves the discovery of new microproteins and their progenitors (26).

4. Functional role in plants

Numerous microProteins from various families have been found in plants, such as Arabidopsis. Numerous physiological processes, including epidermal cell patterning during root hair growth, trichome formation, light-mediated responses, leaf development, pigment production and floral development, have been linked to these miPs (9) (Table 1). Role of microproteins in light signalling responses: Light is one of the leading environmental elements affecting a plant's development. It aids in photosynthesis, detects seasonal and diurnal changes and serves as a cue for changing a plant's developmental pattern (6). Recent research has highlighted the critical roles of microproteins (miPs) in regulating light signalling, which enables plants to thrive in different environmental conditions. The HLH miP type in light signalling is the most researched miP type.

Table 1. List of microProteins characterized and their function in plants

Species	Microproteins	Functions	References
Arabidopsis	TRYPTYCHON (TRY)	Trichrome architecture	13
Solanum lycopersicum	SlPTS	Leaf Development	17
Oryza sativa	ILI1/PRE1	Leaf inclination, cell elongation	24
Solanum lycopersicum	SlTRY	Trichome and root hair development, Anthocyanin biosynthesis	25
Gossypium spp.	GhCPC	Cotton fiber elongation	9
Solanum lycopersicum	SlMIA	Floral development, stem cell homeostasis	26
Solanum lycopersicum	SlDTM	Stem cell homeostasis	13
Brachypodium distachion	LNJ	Shoot architecture	12
Oryza sativa	Hd1miP	Flowering	9
Populus tometosa	$PtRD26^{IR}$	Leaf Senescence	19

5. Trichome and root hair development

Trichomes are epidermal appendages that are present on the leaves, stems and fruits of plants. They are essential to structural plants emphasize the practical applications of these findings, such as their role in improving crop resistance or productivity because they shield plants from biotic (such as herbivores) and abiotic (such as UV radiation and extremely high or low temperatures) stresses. As a naturally occurring defense mechanism that aids plants in overcoming environmental obstacles, trichomes have been the subject of much research. Trichomes mutants from a population of ethyl methanesulfonate (EMS) mutants were screened to perform the first functional characterisation of microproteins (miPs) in plants. One gene was identified as TRYPTYCHON (TRY) out of the 21 genes found during this screening. The 106 amino acid (13 kDa) protein that TRY encodes lacks a transcriptional activation domain but has an R3 single-repeat MYB domain-characteristics of a microprotein. This finding shed light on the crucial role miPs play in controlling the growth of trichomes and offered an understanding of their functional traits in plants (27).

Try mutants had clusters (or nests) of up to four larger and aberrantly branching trichomes in place of a single trichome. Defects in epidermal cell patterning, a procedure necessary for proper trichome creation, are connected to this aberrant trichome development. The subsequent investigation demonstrated that GL1 (GLABRA1) and TTG (TRANSPARENT TESTA GLABRA), two important regulators of trichome initiation, are acted upon by the TRYPTYCHON (TRY) gene (28). Moreover, TRY acts upstream of genes necessary for appropriate epidermal cell patterning, such as STI (STICHEL), DIS1 (DISTORTED TRICHOMES 1), DIS2 and GL3 (GLABRA3). TRY may have a regulatory function in regulating the ratio of cell division to differentiation during trichome development, based on its placement in the genetic hierarchy. Excessive results from a malfunctioning TRY function.

Further research revealed an antagonistic link between TRY and GL1/TTG in controlling the endoreduplication of trichome and epidermal cells (29). Later

research revealed that TRY prevents GL1 from connecting to GL3, which prevents the MBW complex (MYB-bHLH-WDR) from forming. As an activator, the MBW complex, which is made up of TTG1, GL1, GL3 and EGL3, encourages the activation of important downstream genes including TTG2 and GL2. These genes are essential for the growth of trichomes. Furthermore, NTL8 (NTM1-LIKE 8) and the MBW complex both support the expression of TRY itself (30). By controlling the MBW complex's formation, TRY can precisely control the trichome and epidermal cell development processes. This is made possible by a feedback mechanism.

Later, an 11 kDa MYB microprotein called CAPRICE (CPC) was discovered while screening transfer DNA-tagged Arabidopsis lines (31). The plants with CPC mutations had sporadic root hairs and did not differ significantly in size or shape from the wild-type plants. It is interesting to note that trichome formation in the plant's aerial sections was adversely affected by ectopic expression of CPC, in addition to increasing the amount of root hairs. The complicated regulatory role of CPC in epidermal patterning is highlighted by its dual effect in limiting the growth of trichomes and stimulating the production of root hairs.

The epidermal cell fate regulator WEREWOLF (WER) and the root-hair cell differentiation-promoting protein CAPRICE (CPC) are thought to have diverged due to CPC's loss of DNA binding capacity and truncation of its activation domain. The WER-GL3/TTG1-EGL3 transcriptional complex, which raises GL2 expression in non-hair cells, inhibits the growth of root hair. The transfer of CPC from non-hair cells to hair cells, where it competes with WER for binding to the same transcriptional complex, is what gives birth to CPC's role in root-hair specification. This process inhibits GL2 expression and facilitates the development of root hairs (30). Furthermore, by binding directly to the CPC promoter, WER functions as a negative regulator of CPC in addition to adversely regulating the development of root hair. It has been determined that ENHANCER OF TRY AND CPC 1 (ETC1), a significant MYB microprotein that is 9 kDa, functions somewhat redundantly with CPC and TRY in epidermal patterning (31). Five R3-MYB microProtein homologs (9–13 kDa) were analysed and the results showed that they all have similar and redundant roles in root epidermal cell patterning, which together control the destiny of epidermal cells in Arabidopsis.

5.1 Light-mediated responses in plants **:** Plants can orient themselves towards or away from light through a directional growth response called phototropism, which is crucial for their survival and ability to compete. When seedlings are not exposed to light, they go through a process called skoto-morphogenesis, which is marked by extended hypocotyls, closed cotyledons and the development of an apical hook. The seedlings transition to photo-morphogenesis-a process that includes inhibiting cotyledon opening, hypocotyl elongation and greening-when they are exposed to light. ELONGATED HYPOCOTYL 5 (HY5), a crucial regulator of photo-morphogenesis, regulates several aspects of seedling growth and is essential for light signalling. Under both white and UV light, HY5 directly controls the transcription of MICROPROTEIN 1A (miP1a), also referred to as B-BOX DOMAIN PROTEIN 31 (BBX31). While HY5 regulates BBX31, it also acts on its own to affect photo-morphogenesis. The BBX protein family includes the proteins miP1a and miP1b. These proteins lack the CCT domain necessary for DNA binding, but they do have a single B-box domain for protein-protein interactions.

ELONGATED HYPOCOTYL 5 (HY5), a crucial regulator of photo-morphogenesis, regulates several aspects of seedling growth and is essential for light signalling. Under both white and UV light, HY5 directly controls the transcription of MICROPROTEIN 1A (miP1a), also referred to as B-BOX DOMAIN PROTEIN 31 (BBX31). While HY5 regulates BBX31, it also acts on its own to affect photo-morphogenesis.

The BBX protein family includes the proteins miP1a and miP1b. These proteins have a single B-box domain for protein-protein interactions, but they lack the CCT domain required for DNA binding. During the transition from darkness to light, MiP1a/b proteins accumulate rapidly, interact with key transcription factors including PIF3 (PHYTOCHROME INTERACTING FACTOR 3) and EIN3 (ETHYLENE INSENSITIVE 3) and impede their oligomerisation. The plant's growth response to light is regulated and the shift to photo-morphogenesis is facilitated by this interaction, which inhibits the transcriptional activities of PIF3 and EIN3.

5.2 Leaf development **:** The main organs of plants that carry out photosynthesis are the leaves. With a clear dorsoventral asymmetry, they arise from the pluripotent cells of the shoot apical meristem. The leaf surfaces' orientation is referred to as this asymmetry: the abaxial surface develops on the opposite side, facing away from the meristem, while the adaxial surface stays near it (31). As it affects both the cellular specialisation of leaf tissues and the formation of axillary meristems, which give rise to lateral shoots and branches, proper adaxial/abaxial polarity is essential for proper leaf development. HD-ZIPIII transcription factors are essential in regulating leaf development and are largely responsible for the regulation of this polarity and the appropriate differentiation of leaf surfaces. These proteins support the proper construction and function of the leaf by coordinating the creation of the adaxial and abaxial domains.

ZPR1, ZPR2, ZPR3 and ZPR4 are LITTLE ZIPPER (ZPR) proteins; they belong to a family of microproteins (miP) that range in size from 7 to 16 kDa. This information was discovered through transcriptome profiling of seedlings with inducible HD-ZIPIII activity (7). By directly interacting with HD-ZIPIII transcription factors, especially REVOLUTA (REV) and blocking their ability to bind DNA, these ZPR proteins provide a regulatory function. ZPR protein overexpression suppresses HD-ZIPIII/REV activity, which throws off leaves' typical adaxial-abaxial polarity. Furthermore, ZPR genes are transcriptionally regulated by HD-ZIPIII/REV, which creates a negative feedback loop in which ZPR proteins are expressed and then suppress HD-ZIPIII/REV activity (32). Through the creation of adaxial-abaxial polarity in the leaf tissues and appropriate leaf

growth, this feedback mechanism aids in fine-tuning the balance of HD-ZIPIII activity.

5.3 Pigment biosynthesis : Previously identified as a critical regulator of the fate of root hair cells, the R3-MYB type microProtein (miP) CAPRICE (CPC) also affects other physiological processes, including those on the biosynthesis of anthocyanins. AtCPC was ectopically expressed in Nicotiana (tobacco) plants in a study by (25), which resulted in a variegated flowering pattern and changed trichome and root hair distribution. Differential amounts of anthocyanin, AtCPC mRNA and late flavonoid pathway genes were shown to correspond with this variation.

Furthermore, it was shown by a yeast two-hybrid (Y2H) study that AtCPC may bind to bHLH proteins, namely JAF13 in Petunia and ANTHOCYANIN 1 (AN1), which are both important regulators of anthocyanin biosynthesis. These interactions imply that AtCPC may suppress the transcriptional activity of AN1 and JAF13, which would negatively influence the formation of anthocyanins (25). Similar correlations between anthocyanin levels and CPC expression were shown in Arabidopsis, where anthocyanin content in CPC mutants correlated with CPC expression levels under various stress scenarios (33). These results suggest that CPC has a more extensive regulatory function that goes beyond determining the destiny of cells and includes modulating the production of anthocyanins, especially in stressful environmental conditions.

5.4 Flowering and floral development : MicroProteins have been discovered to be significant regulators in floral development and flowering time. One well-known example is the development of the gynoecium, which is the last structure to arise from the floral meristem. Three key genes-KNUCKLES (KNU), AGAMOUS (AG) and WUSCHEL (WUS)-are involved in mediating this developmental process. Specifically, AG increases the expression of the tomato ortholog SlIMA (INHIBITOR OF MERISTEM ACTIVITY) and about 10 kDa MIF1 miP. The MIF1 miP and SlIMA interact with KNU to recruit TOPLESS (TPL) and HISTONE DEACETYLASE19 (HDA19). They work together to create a repressor complex that transcriptionally controls WUS expression, which is essential for the upkeep and termination of the floral meristem (27). Moreover, two BBX (B-Box) domain proteins were found in Arabidopsis by an ab initio investigation of 44 putative microProteins: miP1a/BBX31 and miP1b/BBX30. These results highlight the variety of regulatory roles that miPs play in developmental processes.

Furthermore, CAPRICE-LIKE 3 (CPL3), often referred to as ENHANCER OF TRY AND CPC3 (ETC3), is another R3-MYB type miP that controls flowering. CPL3 mutants flower early, but 35S overexpression causes flowering to occur later. Important blooming genes like SUPPRESSOR OF OVEREXPRESSION OF CO 1 (SOC1), CONSTANS (CO) and blooming LOCUS T (FT) express differently in response to this delayed flowering (26, 34, 35) The precise method by which CPL3 controls flowering is unknown, although it interacts with GL3, EGL3 and AtMYC1. None of these interactions directly influence flowering.

6. Synthetic microprotein in plants

By engaging with certain target proteins through compatible protein-protein interaction (PPI) domains, microproteins (miPs) present a unique chance to fine-tune physiological processes. This makes it possible to create artificial miPs that are specifically targeted at important proteins in vital processes. To produce more specific or varied effects, synthetic miPs could be engineered with one or more functional domains, enabling more customised interventions in plant physiology. Transgenic Arabidopsis plants that expressed PPI domains of important proteins, such as CRYPTOCHROME1 (CRY1), BRASSINOSTEROID INSENSITIVE1 (BRI1) and DICER-LIKE1 (DCL1) was developed (23). Important physiological processes like light signalling, brassinosteroid response and miRNA synthesis are all regulated by these proteins. A single functional domain of a multi-domain protein that is capable of interacting with its target protein is expressed in a controlled manner to obtain the desired effects. In Arabidopsis, the overexpression of the PPI domain of the transcription factors SUPPRESSOR OF OVEREXPRESSOR OF CONSTANS 1 (SOC1), AGAMOUS (AG) and LATE ELONGATED HYPOCOTYL (LHY) resulted in phenotypes similar to the loss-of-function mutant of the respective transcription factors. This was a result of the ability of these PPI domains to heterodimerize with their source transcription factors and negatively regulate their function. This approach was also effective when the PPI domain of SOC1 was overexpressed in *Brachypodium distachyon* resulting in delayed heading. Synthetic microproteins have also been used to successfully modulate the flowering time of rice grown in long-day conditions. These studies reveal that the ectopic expression of the PPI domain of transcription factors can function as microproteins due to their ability to negatively regulate the larger transcription factors (37). A recent publication showed that synthetic microproteins are capable of regulating larger multi-domain proteins that are not transcription factor proteins (38). Proof of concept by using the technique on the commercially significant monocot model of rice was expanded by (10). Using HEADING DATE 1 (Hd1), the rice ortholog of CONSTANS (CO), they synthesised a miP that regulates flowering by controlling the expression of RICE FLOWERING LOCUS T 1 (RFT1) and HEADING DATE 3A (Hd3a). By truncating Hd1's CCT domain, they created Hd1miP, a synthetic microprotein that interacted with Hd1. Plants expressing Hd1miP showed a dose-dependent early flowering phenotype that was independent of photoperiod, highlighting its potential use in agriculture. Moreover, Hd1miP exhibited more profound effects on yield characteristics including grain length and width and blocked the Hd1-mediated regulation of Hd3a and RFT1.

Conclusion

Microproteins are small in size and singular functional domain, which can engage flexibly in networks of protein-protein interactions. miPs can modify target proteins in a variety of biological processes because they have lost their DNA-binding domains but kept their protein-protein interaction domains throughout evolution. It is noteworthy that miPs such as CAPRICE (CPC) exhibit functional variety

as they are implicated in many regulatory pathways. MicroProteins are small proteins that contain a single protein domain and are related to larger, often multi-domain proteins. At the molecular level, microproteins act by interfering with the formation of higher-order protein complexes. In the past years, several microproteins have been identified in plants and animals that strongly influence biological processes. Due to their ability to act as dominant regulators in a targeted manner, microproteins have a high potential for biotechnological use. Technological developments in analytical, molecular and computational approaches have increased the number of miPs found in various species. On the other hand, not much is known about their cellular specificities and dynamics. The majority of research conducted to date has been on the dominant-negative impact that miPs have on the proteins that they target; however, positive regulatory activities, such as stabilising or improving the function of multimeric complexes, may also be investigated. miPs offer an innovative and sustainable way to boost plant stress tolerance and productivity in the face of global climate change and unsustainable farming practices. They are ideal candidates for synthetic microprotein design because of their small size and flexibility, which might be used to strengthen plant stress tolerance. Studying miPs has a bright future thanks to CRISPR-Cas9 gene-editing technology, especially in light of the difficulties in transferring DNA insertional mutants for miP-coding genes in Arabidopsis and other organisms. Since miPs have a short sequence length, miRNA-induced gene silencing techniques have been favoured. However, CRISPR-Cas technology can offer a more precise and efficient way to knock out mutants, allowing for the functional characterisation of miPs.

Acknowledgements

The authors acknowledge the institution's provision of laboratory and field facilities for conducting the research.

Authors' contributions

ATH, KM and NP have contributed to writing the original draft, reviewing and editing, methodology, data curation and conceptualisation. SK has done review and editing and data curation. DK contributed through review and editing, data curation and methodology. VVG conducted formal analysis. MKM contributed to formal analysis, review and editing. SM Samyuktha also conducted formal analysis, review and editing. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest: Authors do not have any conflict of interest to declare.

Ethical issues: None

References

1. Harshitha BS, Manjunath KK, Bhargavi HA. Mysterious microproteins as a novel tool for crop improvement: A review. The Pharma Innov J. 2023;12(1):2317–22. https://doi.org/10.22271/tpi.2023.v12.i1z.18331.
2. Staub D, Wenkel S. Cross-species genome-wide identification of evolutionary conserved MicroProteins. Genome Biol Evolution. 2017;9:777–89. doi: 10.1093/gbe/evx041.
3. Saroha A, Kotiyal A, Nanoemulsion R. A futuristic approach to disease elimination for food safety. Ann Phytomed. 2023;12 (1):180–86. DOI:10.54085/ap.2023.12.1.50.
4. Makde DM, Baheti PJ, Thote L. Green synthesis of zinc oxide nanoparticles (Zno-Nps) *Ailanthus excelsa* Roxb. stem bark extract and its antibacterial activity. Ann Phytomed. 2022;11 (2):743–47. DOI: 10.54085/ap.2022.11.2.90.
5. Dang R. Role of antinutrient metabolites of plant on production of secondary metabolites and human health. Ann Phytomed. 2018;7(1):1–4. DOI:10.21276/AP.2018.7.1.1
6. Benezra R, Davis RL, Lockshon D, Turner DL, Weintraub H. The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. Cell. 1990;61:49–59. DOI: 10.1016/0092-8674(90)90214-y
7. Wenkel S, Emery J, Hou BH, Evans MMS, Barton MK. A feedback regulatory module formed by LITTLE ZIPPER and HD-ZIPIII genes. Plant Cell. 2007;19:3379–90. https://doi.org/10.1105/tpc.107.055772
8. Fesenko I, Shabalina SA, Mamaeva A, Knyazev A, Glushkevich A, Lyapina I, et al. A vast pool of line age specific microproteins encoded by long noncoding RNAs in plants. Nucleic Acids Res. 2021;49(18):10328–46. DOI: 10.1093/nar/gkab816
9. Mohideen AP. Green synthesis of silver nanoparticles (AgNPs) using of *Laurus nobilis* L. leaf extracts and evaluating its antiarthritic activity by *in vitro* protein denaturation and membrane stabilization assays. Ann Phytomed. 2022;10(2):67–71. DOI: http://dx.doi.org/10.21276/ap.2021.10.2.9
10. Eguen T, Straub D, Graeff M, Wenkel S. Micro proteins: small size - big impact. Trends Plant Sci. 2015;20:477–82. https://doi.org/10.1016/j.tplants.2015.05.011
11. Graeff M, Straub, D, Eguen T, Dolde U, Rodrigues V, Brandt R, Wenkel S. MicroProtein-mediated recruitment of CONSTANS into a TOPLESS trimeric complex represses flowering in arabidopsis. PLoS Genetics. 2016;12:1005959. https://doi.org/10.1371/journal.pgen.1005959
12. Bhati KK, Dolde U, Wenkel S. Micro proteins expanding functions and novel modes of regulation. Mol Plant Biol. 2021;14:705–07. DOI: 10.1016/j.molp.2021.01.006
13. Wu Q, Zhong S, Shi H. MicroProteins: dynamic and accurate regulation of protein activity. J Integr Plant Biol. 2022;64:812–20. https://doi.org/10.1111/jipb.13229
14. Hsu K, Seharaseyon J, Dong P, Bour S, Marba E. Mutual functional destruction of HIV-1 Vpu and host TASK-1 channel. Mol Cell. 2004;14:259–67. https://doi.org/10.1016/S1097-2765(04)00183-2
15. Ahrens CH, Wade JT, Champion MM, Langer JD. A practical guide to small protein discovery and characterization using mass spectrometry. J Bacteriol. 2022;204:e0035321. DOI: 10.1128/JB.00353-21
16. Bhati KK, Blaakmeer A, Paredes EB, Dolde U, Eguen T, Hong S, et al. Approaches to identify and characterize micro proteins and their potential uses in biotechnology. Cell Mol Life Sci. 2018;75:2529–36. DOI: 10.1007/s00018-018-2818-8
17. Donnelly DP, Rawlins CM, DeHart CJ, Fornelli L, Schachner LF, Lin Z, et al. Best practices and benchmarks for intact protein analysis for top-down mass spectrometry. Nat Methods. 2019;16:587–94. DOIhttps://doi.org/10.1038/s41592-019-0457-0.

18. Kaulich PT, Cassidy L, Weidenbach K, Schmitz RA, Tholey. Complementarity of different SDS-PAGE gel staining methods for the identification of short open reading frame-encoded peptides. Proteomics. 2020;20. DOI: 10.1002/pmic.202000084

19. Zabret J, Bohn S, Schuller SK, Arnolds O, Langer JD. Structural insights into photosystem II assembly. Nat Plant. 2021;7:524–38. DOI: 10.1038/s41477-021-00895-0

20. Tian H, Wang X, Guo H, Cheng Y, Hou C, Chen JG, Wang S. NTL8 regulates trichome formation in *Arabidopsis* by directly activating R3 MYB genes TRY and TCL1. Plant Physiol. 2017a;174:2363–75. https://doi.org/10.1104/pp.17.00510

21. Olexiouk V, Crappe J, Verbruggen S, Verhegen K, Martens L, Menschaert G. sORFs.org: a repository of small ORFs identified by ribosome profiling. Nucleic Acids Res. 2016;44:324–29. DOI: 10.1093/nar/gkx1130

22. Morton T, Petricka J, Corcoran DL, Li S, Winter CM, Benfey PN, et al. Paired-end analysis of transcription start sites in *Arabidopsis* reveals plant-specific promoter signatures. Plant Cell. 2014;2746–60. https://doi.org/10.1105/tpc.114.125617.

23. Dolde U, Rodrigues V, Straub D, Bhati KK, Choi S, Yang SW, Wenkel S. Synthetic MicroProteins: versatile tools for posttranslational regulation of target proteins. Plant Physio. 2018;176:3136–45. https://doi.org/10.1104/pp.17.01743

24. Floyd SK, Ryan JG, Conway SJ, Brenner E, Burris KP, Burris JN, et al. Origin of a novel regulatory module by duplication and degeneration of an ancient plant transcription factor. Mol Phylogenetics Evolution. 2014;81:159–73. https://doi.org/10.1016/j.ympev.2014.06.017

25. Zhang W, Ning G, Lv H, Liao L, Bao M. Single MYB-type transcription factor At CAPRICE, a new efficient tool to engineer the production of anthocyanin in tobacco. Biochem Biophys Res Commun. 2009b; 388:742–47. DOI: 10.1016/j.bbrc.2009.08.092

26. Tominaga R, Iwata M, Sano R, Inoue K, Okada K, Wada T. *Arabidopsis* CAPRICE-LIKE MYB 3 (CPL3) controls endoreduplication and flowering development in addition to trichome and root hair formation. Development. 2008;135:1335–45. DOI: 10.1242/dev.017947

27. Bollier N, Sicard A, Leblond J, Latrasse D, Gonzalez N, At-MINI. zinc FINGER2 and Sl-INHIBITOR of meristem activity, a conserved missing link in the regulation of floral meristem termination in *Arabidopsis* and tomato. Plant Cell. 2018;30;83–100. https://doi.org/10.1105/tpc.17.00653

28. Schnittger A, Jurgens G, Hu lskamp M. Tissue layer and organ specificity of trichome formation are regulated by GLABRA1 and TRIPTYCHON in *Arabidopsis*. Development. 1998;125:2283–89. DOI: 10.1242/dev.125.12.2283

29. Tian T, Liu Y, Yan H, You Q, Yi X. AgriGO v2.0: a GO analysis toolkit for the agricultural community. Nucleic Acids Res. 2017b;45 W122–W129. https://doi.org/10.1093/nar/gkx382

30. Wada T, Tachibana T, Shimura Y, Okada K. Epidermal cell differentiation in *Arabidopsis* determined by a Myb homolog. CPC Sci. 1997;277:1113–16. DOI: 10.1126/science.277.5329.1113

31. Canales C, Grigg S, Tsiantis M. The formation and patterning of leaves: recent advances. Planta. 2005;221:752–56. DOI: 10.1007/s00425-005-1549-x

32. Brandt R, Xie Y, Musielak T, Graeff M, Stierhof YD, Huang H, et al. Control of stem cell homeostasis via interlocking microRNA and microProtein feedback loops. Mech Develop. 2013;130:25–33. http://dx.doi.org/10.1016/j.mod.2012.06.007

33. Zhu HF, Fitzsimmons K, Khandelwal A, Kranz RG. CPC, a single-repeat R3 MYB, is a negative regulator of anthocyanin biosynthesis in *Arabidopsis*. Mol Plant. 2009;2:790–802. DOI: 10.1093/mp/ssp030

34. Wada T, Tominaga-Wada R. CAPRICE family genes control flowering time through both promoting and repressing CONSTANS and FLOWERING LOCUS T expression. Plant Sci. 2015;241:260–65. https://doi.org/10.1016/j.plantsci.2015.10.015

35. Wada T, Kurata T, Tominaga R, Koshino- Kimura Y, Tachibana T, Goto K, et al. Role of a positive regulator of root hair development, CAPRICE, in *Arabidopsis* root epidermal cell differentiation. Develop. 2002;129:5409–19. ttps://doi.org/10.1242/dev.00111

36. Chory J, Chatterjee M, Cook RK, Elich T, Fankhauser C, Li J. From seed germination to flowering, light controls plant development via phytochrome. Proceed of the Nat Acad of Sci. 1996;93 (22):12066–71. DOI: 10.1073/pnas.93.22.12066

37. Bhati KK, Blaakmeer A, Paredes EB, Dolde U, Eguen T, Hong S, Rodrigues V. Approaches to identify and characterize microProteins and their potential uses in biotechnology. Cellular and Mol Life Sci. 2018;75:2529–36. doi: 10.1007/s00018-018-2818-8.

38. Seo PJ, Hong SY, Ryu JY, Jeong EY, Kim SG, Baldwin IT, Park CM. Targeted inactivation of transcription factors by overexpression of their truncated forms in plants. Plant J Cell Mol Biol. 2012;72 (1):162–72. https://doi.org/10.1111/j.1365-313X.2012.05069.x