Evolutionary tinkering: birth of a novel chloroplast protein

Tatjana KLEINE and Dario LEISTER*1
Lehrstuhl für Botanik, Department Biologie I, Ludwig-Maximilians-Universität, Menzinger Str. 67, 80638 München, Germany

The term ‘evolutionary tinkering’ refers to evolutionary innovation by recombination of functional units, and includes the creation of novel proteins from pre-existing modules. A novel instance of evolutionary tinkering was recently discovered in the flowering plant genus Nicotiana: the conversion of a nuclear transcription factor into the plastid-resident protein WIN4 (wound-induced clone 4) involved in environmental stress responses. In this issue of the Biochemical Journal, Kodama and Sano now show that two steps are necessary for the establishment of the novel plastid protein: the acquisition of an internal translation initiation site and the use of multiple transcription starts to produce short mRNA variants that encode the plastid-targeted protein form.

Key words: chloroplast, evolutionary tinkering, Nicotiana, protein evolution, transcription factor, transit peptide.

Chloroplasts are of endosymbiotic origin and derive from a cyanobacterium-like progenitor. Although the chloroplast has retained a genome of its own, most genes of cyanobacterial origin have been transferred to the nucleus over evolutionary time, as the endosymbiont was transformed into a specialized organelle [1–3]. Strikingly, less than half of the approx. 4500 proteins of cyanobacterial origin now encoded in the nuclear genomes of plants are predicted to be located in chloroplasts, implying that a massive redistribution of cyanobacterium-derived proteins to other cellular compartments has occurred during evolution of the ‘green’ lineage [1,4]. Thus the endosymbiotic acquisition of cyanobacterial genes has led not only to the successful establishment of a new organelle, but also to substantial changes in the composition of the proteomes of the other compartments in the plant cell. Furthermore, the chloroplast proteome itself now includes a substantial set of proteins not derived from the endosymbiont [2,4]. Comparative analyses of the chloroplast proteomes of different plant species indicate that, although a core set of chloroplast proteins has been maintained, substantial divergence of chloroplast proteomes has occurred during the evolution of flowering plants, implying that species-related differences in organelle function exist [3].

The classical view of endosymbiotic gene transfer assumes that organellar genes translocated into nuclear chromosomes must acquire appropriate promoter sequences and sequences encoding N-terminal targeting signals (transit peptides) before they can express proteins that can find their way back to their organelle of origin [5]. Similarly, recruitment of a transit peptide is thought to be a prerequisite for the transformation of proteins not derived from cyanobacteria into chloroplast-targeted proteins. Indeed, several cases are known in which nuclear genes of organellar origin have obtained new transit peptide-coding sequences by duplication, recombination with other nuclear sequences and exon shuffling [2,5]. In this issue of the Biochemical Journal, an alternative mechanism for creating a cTP (chloroplast transit peptide) is elucidated by Kodama and Sano [6].

These authors reported previously that, during the evolution of the genus Nicotiana, the nuclear bHLH (basic helix–loop–helix) transcription factor WIN4 (wound-induced clone 4) has become a chloroplast protein. The NtWIN4 (Nicotiana tabacum WIN4) gene was identified in a screen for transcripts that accumulate after wound stress. NtWIN4 encodes a protein with a predicted molecular mass of 28 kDa which is homologous with bHLH-type transcription factors. When the full-length NtWIN4 protein was expressed from appropriate constructs, it was shown [by GFP (green fluorescent protein) fusion experiments] to be located in the cytosol and the nucleus, and to repress the transcription of a reporter gene [7]. Surprisingly, cell fractionation experiments revealed that in planta NtWIN4 is smaller (17 kDa) than expected, and is found only in chloroplasts [7]. Inspection of the 5' segment of the NtWIN4 gene revealed the presence of an additional in-frame ATG codon downstream of the first. Use of this codon as the translation start site yields a 25 kDa protein. Subsequent analyses showed that an N-terminal segment of this form permits uptake into the chloroplast, and suggested that a cTP of 67 amino acids is removed after import into the organelle, resulting in the mature 17 kDa protein. The C-terminal portion of the predicted cTP overlaps with the DNA-binding domain of the bHLH motif. Removal of the cTP therefore presumably prevents the protein from serving as a transcription factor [7]. Transgenic tobacco plants that constitutively overexpress NtWIN4 show albinism and growth retardation, and they die prematurely. Plants in which NtWIN4 RNA levels were reduced by RNAi (RNA interference) appeared to be normal under non-stressed conditions, but showed delayed cell death when inoculated with pathogens. Based on these phenotypes and the observation that NtWIN4 transcripts are up-regulated in response to methyljasmonic acid, hydrogen peroxide, paraquat and pathogen attack, the authors concluded that NtWIN4 has been converted from a nuclear bHLH-type transcription repressor into a plastid-resident regulatory factor involved in defence against biotic and abiotic environmental stresses [7].

In this issue of the Biochemical Journal, Kodama and Sano elucidate the molecular alterations that led to this striking transformation in the genus Nicotiana [6]. N. tabacum is a natural amphidiploid (= allotetraploid) species that is thought to derive from two diploid progenitors, N. sylvestris and N. tomentosiformis. No WIN4 homologue could be amplified from N. tomentosiformis by PCR, whereas both WIN4 protein forms are immunologically detectable in N. sylvestris. Detailed inspection of WIN4 RNAs in N. sylvestris and N. tabacum revealed that

---

1 To whom correspondence should be addressed (email leister@lrz.uni-muenchen.de).
patterns of transcription site usage differ between the two species, even though the *N. sylvestris* genome is apparently the only source of WIN4 transcripts in *N. tabacum*. Furthermore, individual variants are translated with different efficiencies in vitro. The prevalence in *N. tabacum* of short WIN4 transcripts that possess only the internal start codon, and their higher translation efficiency compared with the long mRNA variants coding for the full-length protein, are together sufficient to account for the predominance of the plastid WIN4 variant in this species (Figure 1). This mechanism of establishing a cTP at the cost of another protein module, in this case the bHLH motif, represents a new mode of protein evolution.

How widespread might this mechanism be? The *Arabidopsis* WIN4 homologue, At5g43650, is located in the nucleus and cytosol, as shown by GFP fusion experiments [7]. Although the At5g43650 protein can be relocated to the chloroplast when the first 30 amino acids are artificially removed [7], this cannot happen *in planta* because a second ATG is missing in the *Arabidopsis thaliana* gene. The possibility cannot be excluded that translation of a shorter At5g43650 variant could be initiated by the thickness of the curly lines which represent mRNAs. Translation initiation frequency is symbolized by the thickness of the arrows, and differs between species and mRNA variants: the first AUG is utilized more often in RNAs from *N. sylvestris* than from *N. tabacum*, whereas the second AUG is equally efficient in both species. This accounts for the differences in the relative amounts of the 28 kDa (targeted to the cytosol and the nucleus) and 25 kDa (targeted to the chloroplast) WIN4 forms in the two plant species. cTP, chloroplast transit peptide; bHLH, basis helix–loop–helix motif.

---

**REFERENCES**


