Theses of the doctoral (Ph.D.) dissertation

Farkas Kerényi

Gödöllő

2017

Szent István University

The quality control systems of the gene expression:

the regulatory elements of the Nonsense-Mediated mRNA Decay system

Farkas Kerényi

Gödöllő

2017

Agricultural Biotechnology Institute

The name of the doctoral school: **Doctoral School of Biological Sciences**

tudományága: **molekular biology**

Head of the Ph.D. school: **Dr. Zoltán Nagy**

Head of Department,

Professor of Hungarian Academy of Sciences

Institute of Botany and Ecophysiology,

Szent Istvan University

Supervisor: **Dr. Dániel Silhavy**

Scientific advisor,

Professor of Hungarian Academy of Sciences

NAIK Agricultural Biotechnology Institute

Departement of Genetics, Group of Plant RNA Biology

....

Approved by the doctoral school leader   Approved by the supervisor

Background and objectives

The gene expression in eukaryotic cells is highly regulated. The quality control is at least as important as the control of the volume of the gene expression. The quality control systems are responsible for not presenting faulty proteins in the cells. All steps of the gene expression are controlled by different quality control systems. The most diversified quality control systems are checking the mRNA maturation and editing because this process is very complicated therefore there are lots of possibilities of making mistakes. The different errors are identified by different quality control systems. The common in these systems that they do not repair but induce the degradation of the incorrect mRNAs. The degradation process of the incorrect mRNAs is made through those degradation pathways that eliminate the wild type mRNAs, also. One of the evolved quality control systems recognizes those incorrect mRNAs that harbor premature termination codon (PTC). This quality control system is called Nonsense Mediated mRNA Decay (NMD). From mRNAs containing PTC, truncated and potebtially harmful proteins could be translated that can cause serious deseases. The NMD, in addition to the degradation of incorrect mRNAs, involved in the regulation of the expression of several wild type genes. The most important core factors of NMD, the *trans* factors UPF1, UPF2 and UPF3 (up frameshift 1, -2, -3), are conserved from yeasts to mammals. But the *cis* elements, that are the signals for the NMD to recognize the PTC containing mRNAs, are different. In yeasts and invertebrates, the NMD recognizes those stop codons as PTCs that are followed by unusually long 3'UTR. This is called 'long 3'UTR-based NMD'. In vertebrates, the NMD recognizes stop codons as PTCs that harbor at least one intron in their 3'UTR and the distance between the stop codon and the intron is at least 50 nucleotides and an exon junction complex (EJC) is presented in the 3'UTR after the splicing (intron excision). This is called 'intron based NMD'. The UPF1-UPF2-UPF3 complex (also called as NMD complex) is set up during the

process of recognizing the incorrect mRNA. The NMD complex induces the degradation of the mRNA. In animals, the key step that connect the recognition and degradation of the mRNA is the phosphorylation of the UPF1 by the SMG1 (suppressor with morphogenetic effect on genitalia 1) because the phosphorylated UPF1 can serve as a binding platform for the SMG6, for the SMG5/SMG7 heterodimer and for the PNRC2 (proline-rich nuclear receptor coregulatory protein 2). mRNAs having protein-free, unprotected 5' and/or 3' ends ae produced by these proteins: the SMG6 is an endonuclease, the SMG5/SMG7 complex causes deadenilation and decapping, while the PNRC2 causes decapping. These unprotected mRNAs can be degraded quickly by exonucleases. Therefore the NMD is a very important mRNA quality control system and have been intensively studied and well known process in yeasts, *Drosophila* and in mammalian cells. Ennek megfelelően élesztőben. Our research group have already developed an effective transient gene expression system to identify and examine the *cis* elements and *trans* factors of the plant NMD. Our research group have explored the basics of the mechanism of the plant NMD using this system. It have been proved that both long 3'UTR based and intron based NMD is working in plants. The main *trans* factors of plant NMD, the genes of UPF1, UPF2, UPF3 and SMG7 have been identified. It has been proved that these three genes are involved in both type of NMD but two orthologs of components of mammalian EJC (Y14 and Mago) are needed only in intron based NMD in plants. It has been proved that, like in mammals, introns located close to the stop codon fail to induce NMD in plants. The UPF1 can be phosphoregulated in plants, also. It has been shown that the N- and C-terminal region of plant UPF1 are functionally redundant and both region can induce NMD and both region are phosphorylated.

Based on these similarities between the mammal and the plant intron based NMD we speculated that the intron based NMD in plants also needs at least 50 nucleotides between the stop codon and the intron excision site to be effective. To prove this theory and to clarify the question that the NMD is also a 'yes or no'

answer in plants or there is an increase in NMD effectivity while increasing the distance between the stop codon and the intron we designed and cloned different NMD reporter GFP constructs harboring an intron in their 3'UTR in different distances from the stop codon.

The UPF1 is a highly conserved protein. The N- and C-terminal egion of plant UPF1, like in mammals, also contains several S/TQ potential phosphorylation sites. But it has not been shown yet that these S/TQ sites are needed for plant NMD, that these sites are phosphorylated and that if it exists, this phosphorylations are needed for the plant NMD. To examine these questions we designed and cloned several point and deletion mutants of plant UPF1. With these constructs we could examine and functionally map the S/TQ sites in both N- and C-terminal egions of plant UPF1.

The aims of our study:

1. To specify that the introns in 3'UTR of plant mRNAs should be at least 50 nucleotides from the stop codon to induce NMD and, that the efficiency of the NMD is changing while increasing the distance of the STOP-EJC.

2. To bring to light if the S/TQ sites in the N- and C-terminal regions of UPF1 are needed for plant NMD, and if so, which ones and, whether these S/TQ sites are phosphorylated, and if so, whether these phosphorylations are needed for the NMD.

Materials and methods

Expression vectors and NMD test constructs

The genes examined were cloned between the 35S promoter and 35S terminator regions in a HA-tagged binary vector pBin61S. The BinP14 (P14), the TRV-PDS,

TRV-PDS-UPF1 vectors, the BinU1DN (U1DN) NMD construct and the G-95 and the G-95I NMD reporters have already been prepared earlier by my colleagues.

To clone GFP reporter constructs containing intron in their 3'UTR (G-30I, G-40I, G-50I, G-60I, G-70I and G-80I), the pFF2R reverse and the 30cF, 40cF, 50cF, 60cF 70cF and the 80cF forward primers were used for the PCR with the plasmid G-95I as template. The fragments were cloned into the pBin-GFP vector.

The UPF1 point mutant constructs were made using the linker scanning mutagenesis technique.

Agroinfiltration and virus induced gene silencing (VIGS)

For agroinfiltration and VIGS experiments 3 weeks old wild type *N. benthamiana* plants grown in greenhouse were used. After VIGS or agroinfiltration the plants were kept in plant growth chamber (24°C, 16 hours light/8 hours dark, 7000 lux). To agroinfiltrate kanamycin resistant binary vectors *Agrobacterium tumefaciens* C58C1 strain was used.

The bacterium suspensions were diluted to OD_{600}=0,4 except in case of P14 which was diluted to OD_{600}=0,2. The suspensions to be infiltrated together were mixed and the infiltrated to upper leaves of wild type (3-4 weeks old) or VIGS induced (4-6 weeks old) *N. benthamiana* plants. Photographs were made and RNA and/or protein samples were taken 3 days post-infiltration.

RNS extraction and Northern blot

Total RNA was extracted from cca. 1 cm^2 infiltrated patches smashed to powder under liquid nitrogen followed by extraction buffer (0.1 M glycine, 10 mM EDTA, 0.1 M NaCl, 2% SDS) teatment. After this, phenol-chloroform extraction method was used.

2 µg of RNA was loaded into agarose gel and after running the samples were blotted to nitrocellulose membrane using capillary technique. For hybridization P14 and GFP probes labelled with α-32P isotope wee used. The hibridizition was carried out in Church-buffer at 65 °C, overnight. After washing, the radioactive signals were detected in a phosphoimager.

Protein extaction and Western blot

Protein was extracted from infiltrated patches smashed to powder under liquid nitrogen followed by icecold extraction buffer (25 mM Tris puffer (pH 7.6), 150 mM NaCl, 10% glycerol, 2% PVPP, 0.15% Igepal, 5 mM DTT, 1% Plant Protease Inhibitor Cocktail) treatment. 2x Laemmli buffer was added followed by 5 minutes boiling. Samples were stored at -70°C until further experiments.

Protein samples were separated in SDS-PAGE and transferred to nitrocellulose membrane using a semi-dry blotter. After blocking, the membrane was incubated with HRP-conjugated anti-HA and anti-P14 antibodies for 1 hour. The chemiluminescent signals generated with an ECL-kit was detected using roentgen-film.

Immunoprecipitation (IP) and phosphostaining

The UPF1 mutants were transiently expressed in the leaves of 3 weeks old wild type *N. benthamiana* plants. 3 days post-infiltration the leaves were smashed into powder and extracted in icecold IP buffer (25 mM Tris-HCl (pH 7.6), 150 mM NaCl, 10% glycerol, 2% PVPP, 0.15% Igepal, 5 mM DTT, 1 mM EDTA (pH 8.0), 1% plant protease inhibitor, 1 tablet of cOmplete Mini protease inhibitor). The samples were added to HA agarose beads and incubated for 1 hour with shaking (140 rpm). After washing, the samples were eluated from the beads with 2xLaemli buffer.

Half of the each sample was treated with alkaline phosphatase enzyme. The treated and untreated samples were separated on SDS-PAGE. To stain phosphorylated proteins ProQ Diamond phosphospecific stain was used. After phosphostaining, total protein was stained using PageBlue stain.

Results

Fine mapping of the introns in the 3'UTR of plant mRNAs

In mammals, introns located at least 50 nucleotides downstream from the stop codon induce NMD. It has been previously shown by our research group that 3′UTR located introns trigger plant NMD in a position-dependent manner. The potato Ls intron cloned 99 nt downstream from a stop codon induced strong NMD, whereas the same intron inserted 28 nt downstream from the stop did not trigger NMD. To further study the role of intron position in NMD activation, a series of NMD test constructs were generated by cloning the Ls intron with 30, 40, 50, 60, 70 or 80 nt long stuffer sequences between the GFP reporter gene and the 35S terminator (G-30I, G-40I, G-50I, G-60I, G-70I, G-80I). The construct that contains intron 95 nucleotides downstream from the stop codon (G-95I) has been previously made by my colleagues. We wanted to show what the critical distance is between the stop codon and the 3'UTR located intron that can induce NMD in plants. We wanted to clarify whether the efficiency of the NMD is increasing while the distance of intron from the stop codon is increased.

NMD sensitivity of the test constructs was studied in agroinfiltration-based UPF1 dominant-negative (U1DN) co-expression assays (overexpression of U1DN, a dominant negative mutant form of UPF1 inhibits NMD). The NMD test constructs were agroinfiltrated into *N. benthamiana* leaves only with a P14 RNA silencing suppressor or were co-infiltrated with P14 and U1DN (referred to as control and U1DN co-infiltrated samples, respectively). P14 was co-infiltrated in

these experiments for two reasons: to suppress agroinfiltration-induced intense RNA silencing and to serve as an internal control for RNA gel blot assays.

Expression of U1DN leads to the inactivation of NMD; therefore, NMD target transcripts accumulate to enhanced levels in U1DN co-infiltrated samples. As G-30I and G-40I transcripts accumulated to comparable levels in control and U1DN co-infiltrated, we concluded that these mRNAs are not targeted by NMD. By contrast, co-expression of U1DN significantly enhanced the transcript levels of G-50I, G-60I, G-70I, G-80I and G-95I mRNAs, suggesting that these transcripts are targeted by NMD. Moreover, we found that transcripts with longer stuffer sequences accumulated to lower levels, and that U1DN co-infiltration enhanced their expression more dramatically. Thus, we concluded that extension of the stuffer region from 50 to 80–95 nt gradually enhanced the efficiency of NMD.

Taken together, we concluded that the position of the introns located in the 3'UTR can decisively influence the NMD sensitivity of the mRNA in plants, because (i) only the introns located at least 50 nt downstream from the stop codon can induce NMD and (ii) the distance between the stop codon and the intron can gradually anhance the efficacy of the NMD.

Phosphoregulation of the plant UPF1

UPF1 consists of a highly conserved cysteine–histidine rich (CH) and an ATPase/helicase domain and less conserved N- and C-terminal S/TQ-rich regions. Previously, our research group has shown that UPF1 mutants lacking both terminal regions failed to complement UPF1-silenced plants, while mutants containing either the N- or C-terminal domain complemented UPF1-silenced plants. We also demonstrated that the N- and C-terminal region of UPF1 is phosphorylated. But it remained unclear which aminoacids are phosphorylated and whether these phosphorylations are needed for the NMD in plants.

The S/TQ sites of the UPF1 N-terminal region are phosphorylated and involved in NMD in plants

To understand the function of UPF1 N-terminal phosphorylation in NMD, the potential phosphorylation target sites of the N-terminal domain of UPF1 were mutated, and then the NMD competency of these mutants was tested. As the N- and C-terminal domains are redundant, the function of the N-terminal phosphorylation sites was assessed in a UPF1 mutant (U1ΔC), in which the C-terminal domain was deleted. The NMD competency of the mutants was studied in the previously described virus-induced gene silencing (VIGS)-agroinfiltration depletion–complementation assay. As NMD does not function in UPF1-silenced leaves, if the Gc-I NMD-sensitive reporter is co-infiltrated with an NMD-incompetent UPF1 mutant (a UPF1 mutant that is unable to complement the NMD-deficiency of the UPF1-silenced leaf), the Gc-I transcript level will be high and the green fluorescence will be strong. In contrast, if Gc-I is co-infiltrated with an NMD-competent UPF1 mutant that complements the NMD deficiency of the UPF1-silenced leaf, the rescued NMD will efficiently target the Gc-I mRNAs resulting in a low Gc-I transcript level and weak green fluorescence.

In animals, UPF1 is phosphorylated by the SMG1 PIKK kinase, therefore we mutated the potential PIKK target (S/TQ) sites of the N-terminal domain. The N-terminal domain of UPF1 contains four S/TQ sites, S3, S13, T29 and S105 (referred to as P1, P2, P3 and P4 sites, respectively). To study the function of these S/TQ sites, three point and one deletion mutants were generated. In NP1-4A, all four S/TQ were changed to alanine. In the other two mutants the P1, P2 and P4 (S3, S13 and S105, respectively; NP1-2A4A) sites were mutated or only the P3 (T29; NP3A) was changed to alanine. In the NΔNn mutant, the extreme N-terminus (Nn region, 1–35 amino acids) was deleted.

The green fluorescence was strong and the Gc-I transcripts accumulated to high levels in the NP1-4A co-infiltrated sample. This result suggests that the N-

terminal S/TQ sites of UPF1 play an important role in NMD in plants. To identify which N-terminal S/TQ site of plant UPF1 is involved in NMD, the NMD competency of different point mutants was tested. Both the NP3A and NP1-2A4A mutants could complement the UPF1-silenced leaf. These results indicate that the N-terminal S/TQ sites are functionally redundant. The T29 site that corresponds to the mammalian T28 is sufficient, but is not essential for NMD. These data suggest that the three other SQ sites (or at least one of them) are also involved in NMD. Moreover, as the NΔNn construct failed to complement UPF1-silenced leaves, we concluded that the deleted Nn region plays an important role in NMD.

We hypothesized that the NMD-relevant S/TQ sites in the N-terminal domain are the phosphorylated residues. To test this assumption, phosphorylation of the N-terminal mutants was comparatively studied.

We found that, as shown in our research group's previous studies, the U1ΔC protein was heavily phosphorylated, while the ΔNΔC showed only a background phosphorylation level. Unexpectedly, all N-terminal UPF1 mutants, including the NP1-4A mutant, in which all four S/TQ sites were mutated, and the NΔNn mutant, in which the whole Nn region was deleted, were also heavily phosphorylated. These data indicate that the C-terminal region of the N-terminal domain of UPF1 (called the Nc region) is phosphorylated and that at least one non-S/TQ site is phosphorylated in the N-terminal domain.

We further tested whether phosphorylation of the Nc region is important for NMD. To address this, a new mutant (NΔNc) was created by deleting the Nc region from the U1ΔC construct. As the NΔNc protein efficiently complemented the UPF1-silenced leaves, we concluded that the phosphorylated Nc region is not important for NMD.

While the NΔNc protein that lacks the heavily phosphorylated Nc region complemented UPF1-silenced leaves, we wanted to study the phosphorylation of the NMD-relevant S/TQ sites in this mutant. The S/TQ sites were point-mutated

in the NΔNc construct and the obtained mutants were analyzed for their NMD competency and phosphorylation status. We found that mutations of all three remaining S/TQ sites (S3, S13 and T29) of the NΔNc construct to alanine impaired NMD function. In contrast, the UPF1-silenced leaf was efficiently complemented by the constructs in which the T29 site or the S3 and S13 residues were retained. These results further support that S3, S13 and T29 residues are involved in NMD.

Next we studied the phosphorylation status of these mutants. We observed that the NΔNc protein was phosphorylated, although its phosphorylation was weaker than phosphorylation of the U1ΔC protein. Thus we conclude that both the Nn and Nc regions of the N-terminal domain of UPF1 are phosphorylated. Importantly, the construct in which all three phosphorylation sites were mutated was not phosphorylated indicating that the S/TQ sites are the phosphorylated residues in the Nn region. Consistently, we found that the constructs in which the T29 site or the S3 and S13 residues were retained showed reduced but detectable phosphorylation. These data suggest that the S3, S13 and perhaps T29 NMD-relevant S/TQ sites are phosphorylated. MS data carried out by my colleagues also support that the N-terminal S/TQ sites of UPF1 can be phosphorylated.

We have speculated that both phosphorylation and dephosphorylation of UPF1 are required for plant NMD. In line with this, we found that the phosphomimetic mutant of T29 was non-functional, as changing threonine29 to the frequently used phosphomimetic aspartate (T29D) impaired the NMD competency. We postulate that the phosphomimetic T29D mutant protein is non-functional because it fails to complete the phosphorylation cycle.

We found that co-expression of the NMD-incompetent UPF1 mutants act in a dominant-negative manner. We speculate that these mutants impair NMD because they are incorporated into the NMD complex, but fail to complete the function of UPF1.

Taken together, our mutation and phosphorylation studies revealed that the S3, S13 and T29 sites of UPF1 play an important role in NMD and that these NMD-relevant sites are phosphorylated. Notably, these NMD-relevant phosphorylated sites are in the S/TQ (potential PIKK target) context, whereas the phosphorylated residues of the Nc region that are not involved in NMD are not in the context of potential PIKK targets.

The C-terminal region of the plant UPF1 consists of functionally dominant segments and the C1 region contains phosphorylated S/TQ sites of different function

In mammals, the C-terminal domain of UPF1 contains several S/TQ sites; however, only S1096, the binding site for SMG7, is essential for NMD. The C-terminal domain of plant UPF1 also contains several S/TQ sites. To identify the S/TQ sites that play a role in plant NMD, we mapped the C-terminal domain in a UPF1 mutant lacking the redundant N-terminal domain (ΔNU1). The C-terminal domain was dissected into four regions (C1, C2, C3 and C4) and then the C-terminal domain of the ΔNU1 construct was replaced with one of the C regions (ΔNU1C1, ΔNU1C2, ΔNU1C3 and ΔNU1C4). Each C-terminal region contains four or five S/TQ sites. The NMD function of the four C-terminal deletion mutants was tested in NMD competency assays. All four mutants could complement the UPF1-silenced leaves. Thus, the C-terminal domain of the plant UPF1 consists of functionally redundant S/TQ rich regions. Therefore, we examined only the C1 section further.

The C1 region contains four S/TQ sites, S1013, T1056, S1076 and S1085 (referred to as C1P1, C1P2, C1P3 and C1P4). To study the role of these S/TQ sites in NMD, three point mutants were generated in which all four S/TQ sites were changed to alanine (C1P1-4A), the S1013 and T1056 (C1P1-2A) or the

S1076 and S1085 (C1P3-4A) residues were mutated to alanine. The NMD competency assay revealed that mutations of all four S/TQ residues impaired NMD function. Moreover, as C1P3-4A failed to complement UPF1-silenced leaves whereas C1P1-2A efficiently complemented UPF1-silenced leaves, we concluded that only the S1076 and/or S1085 sites of the C1 region play an important role in NMD. Further mutational studies showed that the mutant, in which S1013, T1056 and S1085 sites were changed to alanine, complemented the NMD deficiency of the UPF1-silenced leaf. In contrast, the C1P1-3A construct in which only the S1085 was retained failed to complement UPF1-silenced plants. Nevertheless, we speculate that S1085 might also be involved in NMD. It appears that the mutant in which both the S1076 and S1085 were present complemented the UPF1-deficiency more efficiently than the mutant in which S1076 was the only potential PIKK target site. Interestingly, the NMD-relevant S1076 and S1085 SQ sites of the C1 region are conserved.

We found that the UPF1 C-terminal NMD-incompetent mutants acted in a dominant-negative manner. This, like in case of the N-terminal region, shows that the phosphorylation of UF1 plays a role in the activity but not in the formation of the NMD complex.

The ΔNU1C1 was strongly phosphorylated whereas the C1P1-4A mutant was weakly phosphorylated. These data indicate that the S/TQ sites of the C1 region are phosphorylated.

New scinentific results

1. In plants, introns located in the 3'UTR at least 50 nt downstream from the stop codon can induce NMD.

2. The plant intron-based NMD shows gradation: the greater the distance between the stop codon and the intron, the more effective the NMD.

3. The Nn and the Nc pat of the N-terminal region of the plant UPF1 are phosphorylated but their role in NMD is different. The S/TQ sites of the Nn part are phosphorylated and plays an important role in NMD while the Nc part does not involved in NMD and its strong phosphorylation is not coupled to S/TQ sites.

4. The C-terminal region of the plant UPF1 consists NMD-relevant functionally redundant sections.

5. The S/TQ sites in the C1 section of the C-teminal region of the plant UPF1 are phosphorylated and the S1076, and probably the S1085 also, are important for the plant NMD.

6. The examined NMD incompetent UPF1 mutants act in a dominant-negative manner.

Conclusions and recommendations

The role of the distance between the stop codon and the intron located in the 3'UTR in the plant NMD

Long 3′UTRs induce NMD in plants, fungi, invertebrates and (less efficiently) in vertebrates, whereas 3′UTR located introns act as efficient NMD *cis* elements only in vertebrates and higher plants. Two eukaryotic NMD evolution models have been proposed. One model assumes that the long 3′UTR-based NMD is the evolutionary ancient form, and intron-based (EJC coupled) NMD has evolved in vertebrates to efficiently eliminate the misproducts of alternative splicing. An alternative model suggests that in stem eukaryotes (the last common ancestors of extant eukaryotes), both types of NMD were functional, and that intron-based NMD was already mediated by EJC in stem eukaryotes. This latter model predicts that EJC also plays a critical role in intron-based NMD in plants.

Our data strongly support the second model, indicating that in plants, like in vertebrates, intron-based NMD is mediated by an EJC-like complex. (i) As vertebrate EJC is deposited 20–25 nt upstream from the exon–exon boundaries and ribosomes can still displace EJC from the 20–25 nt long region downstream of the stop codon, NMD inducing vertebrate introns are located >50 nt downstream of the stop codon. Similarly, we found that 3′UTR introns located at least 50 nt downstream from the stop induce NMD in plants, suggesting that plant NMD is also triggered by a protein complex placed 25 nt upstream of exon–exon junctions. (ii) Orthologs of all four EJC core components are required for intron-based but not for long 3′UTR-based plant NMD. (iii) In plants and in mammals, overexpression of the PYM EJC disassembling factor (or its N-terminal region) inhibits intron-based NMD but did not affect on long 3′UTR-based NMD. (iv) In vertebrates, EJC plays a critical role in intron-based NMD by forming a binding platform for UPF3 and UPF2. Our research group showed that UPF3 and UPF2 are also involved in intron-based plant NMD. It is possible that in plants, like in vertebrates, 3′UTR located EJC facilitates NMD by binding UPF2 and UPF3.

These results and other indirect evidence, for instance, co-localization of putative plant EJC factors and conserved interactions among the components (Y14-Mago, PYM-Y14-Mago), strongly suggest that splicing leads to the formation of an EJC in plants, and that the 3′UTR-bound plant EJCs can trigger NMD. Thus, we propose that intron-based NMD is evolutionary conserved, and that a similar EJC-mediated intron-based NMD system has already functioned in stem eukaryotes.

Although our data suggest that the plant and vertebrate intron-based NMD operates similarly, it is likely that the mechanistic details are different. For instance, vertebrate NMD is limited to the pioneer round of translation because the CBP80 component of the CBP20-80 pioneer cap binding complex stimulates at least two steps of NMD, the recruitment of UPF1 to terminating ribosome and formation of the functional NMD complex. By contrast, the CBP factors are not

required for plant NMD. These (and other) differences might lead to relevant biological consequences, for instance, that intron-based NMD acts more gradually in plants than in vertebrates, thus intron-based NMD might be more suitable for fine tuning of gene expression in plants.

The role of the phosphorylation of the UPF1 in the plant NMD

Based on our previous findings that (i) the S/TQ-rich N- and C-terminal domains of UPF1 play a redundant but essential role in plant NMD, (ii) the N-terminal domain of UPF1 is phosphorylated, (iii) the 14-3-3-phosphoserine-binding domain of SMG7 is essential for NMD and (iv) the SMG7 relocalizes UPF1 into the P-body, we hypothesized that phosphorylation of S/TQ motifs in the terminal domains of UPF1 connects the NMD complex formation and SMG7-mediated target transcript degradation steps of plant NMD. To test this hypothesis, we studied the role of the S/TQ sites of UPF1 in plant NMD, the phosphorylation status of these residues and their function in the formation of the UPF1–SMG7 complex. Consistent with our model, we found that certain S/TQ sites that are located either in the N- or C-terminal domains of UPF1 are phosphorylated and play an important role in plant NMD and that the S/TQ sites of the N-terminal region of UPF1, as well as the C terminal domain, are involved in the binding of the SMG7.

We found that although both the Nn and Nc regions of the N-terminal domain of UPF1 are phosphorylated, the Nc region is dispensable for NMD, while the Nn region plays an important role in NMD. Moreover, we show that all three potential PIKK target (S3, S13 and T29) sites of the Nn region are involved in NMD and that these NMD-relevant S/TQ sites can be phosphorylated. Thus we propose that phosphorylation of S3, S13 and T29 S/TQ sites of UPF1 plays a critical role in plant NMD. Relevantly, the Nn regions including the three NMD-relevant S/TQ sites are highly conserved between monocots and dicots (the Nc

17

region shows limited conservation). The N-terminal S/TQ sites of UPF1 also play a critical role in the binding of SMG7, the key NMD degradation factor in plants. We suggest that the 14-3-3-like phosphoserine/threonine-binding domain of SMG7 binds to the phosphorylated S/TQ sites of the N-terminal domain of plant UPF1.

We have speculated that both phosphorylation and dephosphorylation of UPF1 are required for plant NMD. In line with this, we found that the phosphomimetic mutant of T29 was non-functional, as changing threonine29 to the frequently used phosphomimetic aspartate (T29D) impaired the NMD competency. We postulate that the phosphomimetic T29D mutant protein is non-functional because it fails to complete the phosphorylation cycle.

The C-terminal domains of plant and animal UPF1 proteins are mainly disordered and show only limited conservation. Although the C-terminal domains of plant and animal UPF1 proteins contain several S/TQ sites, the C-terminal domain of mammalian UPF1 harbors only one NMD-relevant S/TQ (S1096) phosphorylation target site, whereas the C-terminal domain of plant UPF1 consists of few functionally redundant S/TQ-rich segments including the C1 region. Our data suggest that the S/TQ sites in the C1 region of the C-terminal domain are functionally different; the S013 and T1056 sites are not involved in NMD, whereas the S1076 plays an important role in NMD and S1085 might also be involved. Phosphorylation assays showed that the S/TQ sites of the C1 region are phosphorylated. Moreover, co-localization and FRET–FLIM experiments unraveled that the C-terminal domain of UPF1 binds SMG7. We hypothesize that in the C-terminal domain of UPF1, as in the N-terminal domain, the phosphorylated NMD-relevant S/TQ sites are bound by SMG7.

Co-localization and FRET–FLIM assays revealed that plant SMG7 directly binds UPF1 and that the two proteins co-localize in the P-bodies. Both the N- and C-terminal regions of UPF1 are responsible for the formation of the UPF1-SMG7

complex, most probably in a phospho-dependent manner. We postulate that the N-terminal 14-3-3-like domain of SMG7 binds phosphorylated N- and C-terminal S/TQ sites of UPF1, and then the C-terminal domain of SMG7 relocalizes UPF1 and the UPF1-bound NMD target mRNA into the P-body RNA degradation foci, where SMG7 triggers rapid decay of the transcript. Recently it has been shown that UPF1 is also present in silencing bodies. It is possible that UPF1-bound RNPs can be alternatively moved into P-bodies or silencing bodies for degradation.

In mammals, SMG1 phosphorylates the NMD-relevant T28 and S1096 S/TQ sites of the N- and C-terminal domains of UPF1, and then these phospho-residues are bound by the 14-3-3 domains of the SMG6 and SMG5:SMG7 NMD target degradation factors, respectively. The N- and C-terminal domains of plant UPF1 might function similarly; phosphorylation of the NMD-relevant S/TQ sites (S3, S13, T29 S1076 and probably a few more in the C-terminal domain) allows the recruitment of the SMG7 NMD degradation factor to the UPF1. However, SMG7 can bind both terminal domains of plant UPF1, explaining the functional redundancy of the two terminal UPF1 domains.

As all NMD-relevant phosphorylated residues of plant UPF1 are in S/TQ contexts we propose that in plants, as in animals, a PIKK kinase phosphorylates UPF1. SMG1 is not present in *Arabidopsis* but potential orthologs of SMG1 are found in the rice and grapevine genomes, thus we suggest that plant PIKKs redundantly phosphorylate UPF1. Further studies are required to clarify which PIKKs phosphorylate plant UPF1 and how the phosphorylation step of plant NMD is regulated.

Publications related to the dissertation

Articles in peer reviewed journals

Kerenyi, F; Wawer, I; Sikorski, PJ; Kufel, J; Silhavy, D: Phosphorylation of the N- and C-terminal UPF1 domains plays a critical role in plant nonsense-mediated mRNA decay (2013) PLANT JOURNAL Volume: 76   Issue: 5   Pages: 836-848.

Nyiko, T; Kerenyi, F; Szabadkai, Levente; Benkovics A, Major P, Sonkoly B, MeraiZ, Barta E, Niemiec E, Kufel J and Silhavy D: Plant nonsense-mediated mRNA decay is controlled by different autoregulatory circuits and can be induced by an EJC-like complex (2013) NUCLEIC ACIDS RESEARCH Volume: 41   Issue: 13   Pages: 6715-6728.

Conference papers related to the dissertation

Kerényi F, Wawer I, Kufel J, Silhavy D: UPF1 phosphorylation is essential for Nonsense Mediated mRNA Decay (NMD) in plants. 'Our Future' conference, Keszthely, 2013.

Other publications not related to the dissertation

Articles in peer reviewed journals

Bakó J, Kerényi F, Hrubi E, Varga I, Daróczi L, Dienes B, Csernoch L, Gáll J, and Hegedűs C: Poly-γ-Glutamic Acid Nanoparticles Based Visible Light-Curable Hydrogel for Biomedical Application (2016) JOURNAL OF NANOMATERIALS Volume 2016 (2016), Article ID 7350516.

Kerényi, F; Tarapcsák, S; Hrubi, E; Baráthne, SÁ; Hegedüs, V; Balogh, S; Bágyi, K; Varga, G; Hegedüs, C: Comparison of sorting of fluorescently and magnetically labelled dental pulp stem cells [Fogbél eredetű őssejtek fluoreszcens és mágneses válogatásának összehasonlító vizsgálata] (2016) FOGORVOSI SZEMLE Volume 109, Issue 1, Pages 29-33.

Kuttor, A; Szaloki, M; Rente, T; Kerenyi, F; Bako, J; Fabian, I; Jenei, A; Lazar, I; Hegedus, C: Preparation and application of highly porous aerogel-based bioactive materials in dentistry (2014) FRONTIERS OF MATERIALS SCIENCE Volume: 8 Issue: 1 Pages: 46-52.

Deák, T; Hoffmann, S; Bodor, P; Kerényi, F; Bisztray, GD; Kozma, P: Marker assisted selection for Seedlessness in a Multiresistant table grape hybrid family (2014) ACTA HORTICULTURAE Volume 1046, Pages 485-492.

Conference papers not related to the dissertation

Kerényi, F; Tombácz, I; Lázár, I; Hegedűs, C: Silica based aerogel composites as potential bone regenerating scaffold materials (oral presentation); European Advanced Materials Congress 2016, Stockholm, Sweden, 2016

Kerényi, F; Kuttor, A; Tombácz, I; Baráthné, SÁ; Lázár, I; Hegedűs, C: Highly porous silica aerogel-based bioactive composite materials as potential scaffolds for bone regeneration (poster presentation); 27th European Conference on Biomaterials ESB2015, Kraków, Poland, 2015.

Kerényi, F; Kuttor, A; Tombácz, I; Baráthné, SÁ; Tarapcsák, S; Lázár, I; Hegedűs, C: Behaviour of STRO-1 sorted dental pulp stem cells on a highly porous silica aerogel-based bioactive composite material (oral presentation); BIT's 7th Annual World Congress of Regenerative Medicine and Stem Cells, Haikou, China, 2014.

Bakó, J; Szepesi, M; Kuttor, A; Kerényi, F; Jenei, A; Hegedűs, C: Drug release from visible-light curable biodegradable nanocomposite hydrogelsystems (poster presentation); 6th International Conference on Drug Discovery and Therapy, Dubai, UAE, 2014.

Kerényi, F; Hrubi, E; Szalóki, G; Jenei, A; Hegedűs, C: Diverse effect of BMP-2 on various cells able to differentiate to osteoblast (poster presentation); 3rd Biotechnology World Congress, Dubai, UAE, 2014.