AUSTRIAN MARSHALL PLAN FOUNDATION

Marshall Plan Scholarship Research Report

BACHELOR DEGREE PROGRAM

Bio and Environmental Technology

The Mechanism of Nonsense mRNA Degradation

Bachelor Thesis submitted

for the attainment of the academic degree

Bachelor of Science in Engineering

by

Alexandra Wiesinger

February 2015

Supervised by

Dr. Manuel Selg

I faithfully declare, that I have written this thesis independently and without outside help. I have only used the sources which are quoted in the reference and I have marked the pages taken, either directly or as regards content from the sources.
The thesis was in this or similar style never shown to another board of examiners nor it was published.

Alexandra Wiesinger

.....
Alexandra Wiesinger

Schwertberg, 31 January 2015

ABSTRACT

The nonsense-mediated mRNA decay (NMD) is a quality mechanism that detects errors, which are produced during gene expression. These errors can appear as mutations created, for example, during replication of DNA. Moreover, mutations can also be the result of transcription, RNA processing, translation and many other conditions, which are explained at the beginning of the thesis. The creation of nonsense mutations, however, is not always due to defects in metabolic processes; it is also a way of gene expression regulation. Since eukaryotic cells own many different repair systems, many of those mutations can be eliminated before any harm is done. However, not all of them can be identified and, as a result, truncated proteins are synthesized.

Eukaryotic cells use the NMD to detect aberrant mRNAs that would further encode incomplete polypeptides. According to recent studies, the spatial relationship between termination codons and protein complexes, which assemble on exon-exon junctions, are decisive for the activation of the mechanism. Termination codons, so-called premature termination codons (PTCs), are recognized, if they are located 50-54 nucleotides upstream of an exon junction complex (EJC). After triggering of the mechanism, the PTC-containing mRNAs are rapidly degraded to avoid the production of non-functioning proteins. However, on the basis of new findings, there are also proteins that are able to suppress the surveillance pathway and cause severe diseases.

Beta-thalassemia and the Marfan syndrome are two examples, which are caused by a non-functional NMD. Both are genetic disorders that adversely affect the patient's lifestyle. In order to substitute medical therapy, research is engaged intensively in these areas, especially in the field of gene therapy.

TABLE OF CONTENT

LIST OF FIGURES

BACKGROUND

The Mammalian Cell

The mammalian cell is of eukaryotic origin and both share similar structural properties. However, there must be torn a distinction between eukaryotic and prokaryotic cells. The most decisive difference between eukaryotes and prokaryotes is the nucleus. Eukaryotes keep their genetic material in a membrane bound nucleus, whereas prokaryotes do not. Furthermore, eukaryotic cells possess organelles, which are surrounded by membranes and, therefore, allow separate compartments for metabolic reactions (Figure 1). Another big difference of eukaryotic cells is found within their chromosomes. The eukaryotic DNA is wrapped around proteins called histones, which again rises the complexity of eukaryotes. This complex structure is called chromatin (College of DuPage, 2004).

As mentioned before, the cell's genetic material (DNA) is residential in the nucleus. DNA is arranged in chromosomes. The number of chromosomes present in

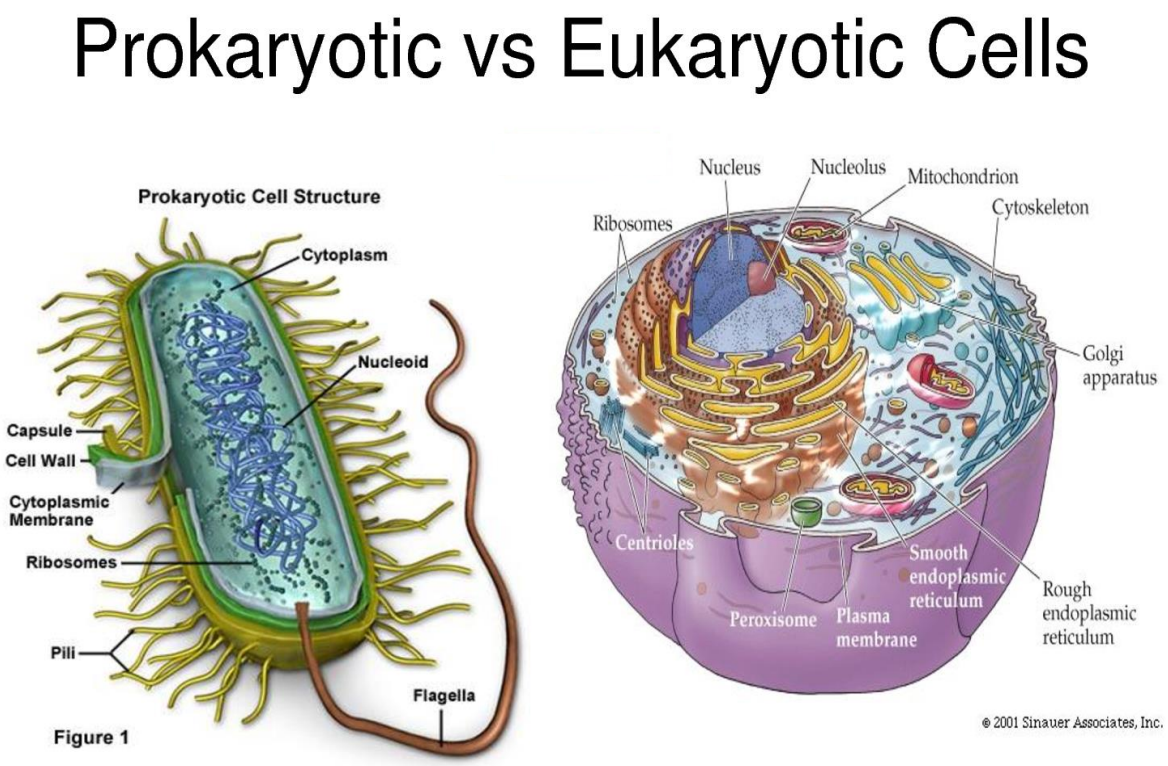

Figure 1: Comparison Prokaryotic Cell (left) vs. Eukaryotic Cell (right)

Source: cf: Sinauer Associates, Inc., 2001. Prokaryotic vs. Eukaryotic Cells. [image online] Available at: <https://illuminolist.wordpress.com/2013/03/> [Accessed 12 January 2015].

a mammalian cell varies with each species and can range from 36 to 60 diploid chromosomes. Each chromosome consists of a single, double stranded DNA, which is packed in a complex structure, the chromatin. There can be differentiated between two types of chromatin, euchromatin and heterochromatin. Euchromatin is less condensed than heterochromatin and, therefore, accessible for replication or transcription of DNA. Heterochromatin structure is required during cell division, when processes like replication or transcription are not required. The density of the packaging also gives some information about the activity of some genes and, therefore, determines how cells differentiate.

Depending on the stage of cell division, DNA fibers change between the state of euchromatin and heterochromatin, as these stages represent different levels of condensation of the genetic material (Arvind and Ram, 1987).

DNA Replication

In order to understand how alterations in DNA strands can occur, the process of DNA replication must be considered. During DNA replication, the whole genome of the cell is copied. In eukaryotic cells, this refers to lots of chromosomes. The initiation of DNA replication is always followed by cell division. The checkpoint, whether replication takes place or not, is the initiation of DNA amplification. Once, the signal for initiation is given, the whole genome is replicated. Due to the fact that eukaryotic cells consist of a huge amount of DNA present as chromatin, replication initiation does not start at one single point. Accelerating the process of replication, there are many so-called replicons, which allow replication to occur at many points simultaneously. As soon as the first replicon is activated, many other replicons on the chromosome are stimulated in an ordered pattern. Replicons adjacent to active genes are replicated earliest and replicons present in heterochromatin are replicate last. If all origins of replication functioned concurrent, the mammalian genome can be duplicated within one hour. Actually, the process takes several hours because not all origins of replication are initiated at once. This rate of replication refers to ~ 2000 base pairs per minute, which is, ,compared to bacterial replication, relatively slow (50,000 bp/min).

Chromosomal replication usually appears as bidirectional replication, which means the replication proceeds away from the origin in both directions. Moreover,

replication is started only once at a specific origin in a cell cycle. As soon as a replicon is activated, replication continues from its origin until the replication fork meets another fork from an adjacent replicon, so there is no termination sequence required. The origin of a replicon, however, includes a discrete sequence, which is rich in A-T base pairs.

DNA replication is a semiconservative, bidirectional and semidiscontinuous process. Unfortunately, the origins of replicons are not well defined yet in eukaryotic cells, except yeast. The process of replication is very complicating involving multiple enzymes working as a complex. The enzyme, which synthesizes the new DNA strand, is called DNA polymerase. In eukaryotic cells, several different DNA polymerases have been identified sharing the same fundamental type of synthetic activity, which means the antiparallel DNA synthesis from 5' to 3' (Bowman, O'Donnell and Kuriyan, 2004; McElhinny et al., 2008; Shiomi et al., 2000; Waga, et al., 2001; Zuo et al., 2000).

Before DNA replication is allowed to occur, the double stranded DNA must be processed to be accessible. As mentioned before, DNA is present either as euchromatin or as heterochromatin. For the amplification of the DNA strand, the euchromatin structure is needed, which is the less condensed pattern and, therefore, accessible for replication.

Prior to the initiation of replication, the duplex DNA strand must be unwound yielding two separated DNA strands, which further serve as templates for DNA synthesis. For the conversion of double-stranded DNA into single-stranded DNA, two types of proteins are required. First, an enzyme is needed to separate the double-stranded DNA. This type of enzyme, called helicase, encircles one strand of the DNA within a replicon. It is considered to have one conformation that attaches to the duplex DNA and another that binds the single-stranded DNA, which occurs within the replicon. The alternate system breaks open the double-stranded DNA hydrolyzing typically one ATP (adenosine triphospate) to break the hydrogen bonds between one base pair. Helicases can break open many kilobases of double-stranded DNA, generating two single strands.

In order to prevent the rebinding of the two single strands, the second function is activated. A single-strand binding protein (SSBP, RPA eukaryotic) binds to the single-stranded DNA as the replication fork advances. These proteins do not unwind DNA, they only stabilize the single strand and keep the two parental strands

separated. The entire single strand is covered with RPAs and maintains the appropriate condition to serve as template for replication.

A free 3'OH end is required allowing the DNA polymerase to bind to the single-stranded DNA. Therefore, a sequence of RNA is synthesized to the template providing the free 3'OH end. This short sequence is called primer and is generated by the primase enzyme, which is an RNA polymerase. Once the 3'OH end is provided, the DNA polymerase can start synthesizing the new DNA strand by adding nucleotide per nucleotide. As a matter of fact, DNA polymerase can only synthesize in 5' to 3' direction. That is why only one of the two single stands can be synthesized continuously. The continuously synthesized strand is referred to as leading strand and only one primer is required for initiation of replication. The lagging strand, however, requires multiple initiation events.

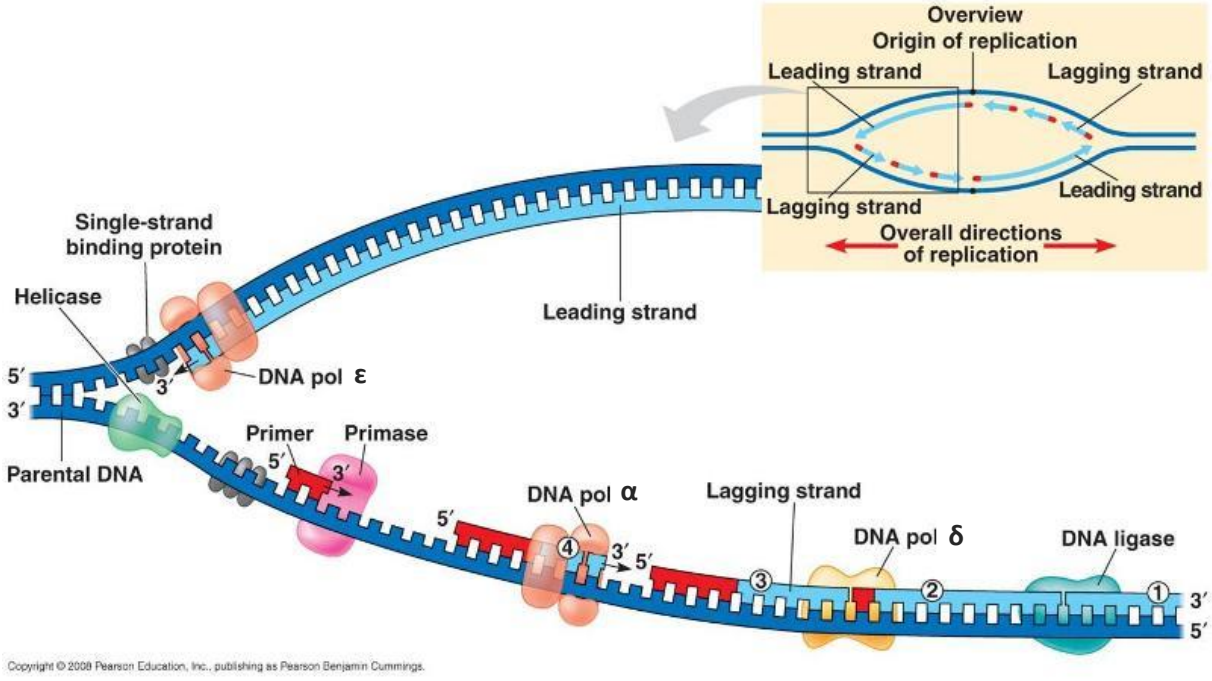

Figure 2: DNA Replication

Source: cf: Campbell, N. A., Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V. and Jackson, R. B.,2008. Biology. 8th ed. New York: Custom Publishing.

And so lots of shorter fragments are built onto the DNA. These short fragments are called Okazaki fragments and each of them requires its own primer. DNA polymerase, then, binds to the primer and elongates each Okazaki fragment, allowing the DNA stand to grow in the opposite direction. The entire process explained above, is shown in figure 2 (Bowman, O'Donnell and Kuriyan, 2004; McElhinny et al., 2008; Shiomi et al., 2000; Waga et al., 2001; Zuo et al., 2000).

A large number of different DNA polymerases carries out the synthesis of DNA. These can be distinguished, mainly, into two groups. First, there are DNA polymerases, which are only required for replication. The second DNA polymerase type is responsible for repair of damaged DNA. DNA polymerase α, δ and ε are responsible for the replication process. All other DNA polymerases in the nucleus are occupied in generating stretches of new DNA to replace damaged ones or in using incorrect DNA as template. DNA polymerases of the first type are large heterotetrameric enzymes with different subunits. One subunit is always responsible for the synthesis of the new strand and the other subunits are involved in additional functions, which include priming or processivity. These DNA polymerases work with high accuracy compared to the repair DNA polymerases. The latter have simpler structures although they often consist of a complex with other repair enzymes. The repair DNA polymerase itself often is composed of a single monomeric structure.

DNA polymerase α is responsible for the nuclear replication and belongs to the high fidelity replicases, as well as DNA polymerases ε, δ and γ. DNA polymerase α initiates the synthesis of a new strand and works as a primase, initiating both, the leading and the lagging strand. The enzyme is built up as a complex associated with three other subunits. There are two subunits, which are responsible for the primase activity and there is a so-called B subunit, which allows the formation of the complex. The leading strand is generated by DNA polymerase ε, whereas the lagging strand is elongated by DNA polymerase δ. The DNA polymerase γ is involved in mitochondrial replication. Base excision repair is carried out with high fidelity by the DNA polymerase β. All the other repair enzymes ζ, η, ι and κ show low accuracy and these are called error-prone polymerases (Bowman, O'Donnell and Kuriyan, 2004; McElhinny et al., 2008; Shiomi et al., 2000; Waga et al., 2001; Zuo et al., 2000).

As shown in figure 2, the polymerase α/primase associated with the initiation site of the DNA generates a short strand of RNA (primer) followed by 20 – 30 bases of DNA. This event is followed by a process called polymerase switch. The DNA polymerases ε and δ replace the DNA polymerase α and each of them catalyzes their specific strands, as mentioned before. The DNA synthesis on the leading strand needs the interaction with two other proteins, RFC clamp loader and trimeric PCNA processivity clamp. The RFC protein is a clamp loading ATPase, which is responsible for the binding of the PCNA onto DNA. The sliding clamp (PCNA)

associates with the 3’ end of the short DNA, which was catalyzed before by the DNA polymerase α. RFC hydrolyzes ATP to cut open the ring of PCNA and allows the encirclement of the DNA (not shown in figure 2). Moreover, PCNA is required for the maintenance of the polymerase δ activity because it connects the polymerase to the DNA template (Bowman, O’Donnell and Kuriyan, 2004; McElhinny et al., 2008; Shiomi et al., 2000; Waga et al., 2001; Zuo et al., 2000).

Hence, the DNA polymerase δ synthesizes many short DNA fragments, which are called Okazaki fragments, the new synthesized DNA strand is not continuous and the short fragments must be connected to each other. Termination of an Okazaki fragment occurs just before the start of an RNA primer. To receive a continuous strand of DNA, the RNA primers must be cut out and replaced with DNA followed by a sealing process. The connection of Okazaki fragments is carried out within two steps. During synthesis of one Okazaki fragment, the RNA primer of the previous fragment is replaced. The endonuclease activity of FEN1 is responsible for the RNA primer removal. Additionally, FEN1 has a 5’ to 3’ exonuclease activity, which is important within DNA repair actions. The main duty of FEN1 is to prevent the DNA from creating structures, like hairpins or misalignments of the template, which would further lead to deletions or duplications in the genome.

Now there are short adjacent Okazaki fragments, which must be connected to each other for receiving a continuous DNA strand. Therefore, the DNA ligase I is specifically needed in order to seal the successive Okazaki fragments. The 3’- OH end of one Okazaki fragment is spatially close to the 5’-phosphate end of the preceding fragment. DNA ligase I forms phosphodiester bonds between the two ends in complex with AMP (adenosine monophosphate). AMP binds to the 5’-phosphate end of the nick and forms the phosphodiester bond with the 3’OH end of the other fragment. After the bond formation, AMP is released, as well as the enzyme (Garg et al., 2004).

The bidirectional replication model states that the divergent replication forks move away from the replicon until they run up against a fork traveling from an adjacent replicon. When the two forks meet, they fuse at the replication fork barrier (RFB) (Gerber et al., 1997). The replication barrier in human and mouse can be found within a region, which is also necessary for RNA polymerase I transcription termination. This is not the only factor that is used by both, transcription termination and replication termination. TTF-I, usually intended to be a transcription termination

factor, could be shown to be essential, as well as in replication termination. The factor consists of two functional domains, one is responsible for interactions with the DNA and the second for transcription termination or replication termination. The replication fork is stopped by a subdomain of the factor, which is also involved in transcription termination. Replication termination is not the result of the TTF-I alone. There are other sequence-specific DNA binding proteins, which are located to neighboring sequences. The proteins that specifically associate to the DNA have not been identified yet. It is proposed that interactions between DNA and proteins, as well as protein-protein interactions are responsible for the termination of the fork movement. The exact mechanism that stops the replication fork is not known yet (Gerber et al., 1997).

As mentioned before, DNA polymerases do not only have the function of synthesizing DNA strands. Many polymerases provide activities that allow to repair damaged nuclear DNA or the translesion replication of incorrect DNA, if the repair of DNA is not possible anymore. The precision of replication relies on the level of specificity of complementary base pairing. The errors that are caused by DNA polymerases can be distinguished into two types.

In the first case, a wrong nucleotide can be added to the growing DNA strand, which leads to improperly base pairing. The mispairing of the base is recognized by an action called proofreading, in which the polymerase examines the growing strand, identifies the wrong-paired base and removes the base.

The second type of error involves an additional or omitted nucleotide and is referred to as frameshifts. The processivity of the enzyme affects the fidelity. So, every DNA polymerase has a special error rate that is limited by its own proofreading activity. Due to the proofreading activity, the error rate in replication can be reduced from 10^{-5} to 10^{-7} per base pair per replication. Moreover, there are more systems that recognize errors, which happen during the amplification of DNA and, so, the number of errors eliminated during replication can be estimated at $< 10^{-9}$ per replicated base pair (Wood et al., 2001).

Eukaryotic Repair Systems

The survival of a cell often depends on the repair systems and these systems are as complicated as the replication itself, during which errors may be created. Repair systems are able to recognize changes in the DNA and it is their job to

remove them. If the repair system fails to replace alterations in the DNA, mutations with severe consequences may occur. In the human genome, more than 130 repair genes has been identified, which reveals the importance of DNA repair. So there are several different systems, which help to avoid detrimental mutations (Wood et al., 2001).

In eukaryotic cells, the packaging of the DNA leads to an even more complex process for DNA repair mechanisms. To allow the process to occur, the nucleosomes (DNA wrapped around histones) must be displaced before or during the repair process and afterwards, the normal state must be restored again. In mammalian cells, numerous histone modifications and chromatin-remodeling events are carried out. To receive access, the structure of chromatin must be altered, so that chromatin-binding proteins are able to tether. Nucleosomes are relocated using ATP. These chromatin-remodeling enzymes are required in all of the eukaryotic repair pathways (Tamburini and Tyler, 2005).

First of all, there are enzymes, which directly reverse specific sorts of damage to the DNA. However, direct repair is not that common. It simply reverses or removes the damage. For example, the removal of pyrimidine dimers can be carried out by light-dependent enzymes and is referred to as photoreactivation. In this type of damage, bases form covalent bonds between each other. However, this repair system is not present in placental mammals.

Another way to repair incorrect or damaged DNA is by cutting out the damage, which is referred to as excision repair. Thereby the aberrant DNA is recognized, the damage is removed and followed by repair synthesis. Usually, there are many different systems of excision repair within one cell. It can be differentiated between the base excision repair, in which the damaged or incorrect base is replaced, and the nucleotide excision repair, in which a sequence is replaced containing the incorrect base or several incorrect bases.

The mismatch repair system is triggered, if a mismatch between two bases within a DNA strand is detected. If base pairs do not properly match or if there are loops, in which sequences are only present in one strand but lack in the other strand, the mechanism of mismatch repair is activated.

If there are problems during replication or due to the influence of radiation and different cell metabolites, double-strand breaks can be created. Double-strand breaks (DSB) are able to result in mutations, including huge chromosomal areas.

The repair of DSBs is carried out via recombination repair using homologous sequences or by linking non-homologous DNA ends.

Two different classes of damage activate repair systems. In the case of single-base changes, they can affect the sequence of the DNA but do not distort the entire DNA. They are caused by alterations of a base *in situ* or can also be caused by replication errors. For example, the deamination of cytosine to uracil causes a mismatched U-G pair. This type of damage does not affect replication or transcription, because the DNA strands are separated and the damage is not detectable anymore. However, if the damage is not removed before these processes occur, the alterations are passed on to the next generation. This common feature that the mismatch is present only until the next replication, only limited time is available to fix the damage.

If alteration within the DNA distort the complete structure of the DNA, replication, as well as transcription, are impeded. Covalent links between bases on the same strand or on the opposite strand do not allow these processes to occur. The most common covalent bond is formed between two thymines. The common feature of this type of damage is, that the damaged adduct is kept in the DNA and results in structural problems and create mutations until it is replaced. Without this repair mechanism, cells are extremely sensitive to evolve DNA damage (Klungland and Lindahl, 1997; Matsumoto and Kim, 1995; Reardon and Sancar, 2003).

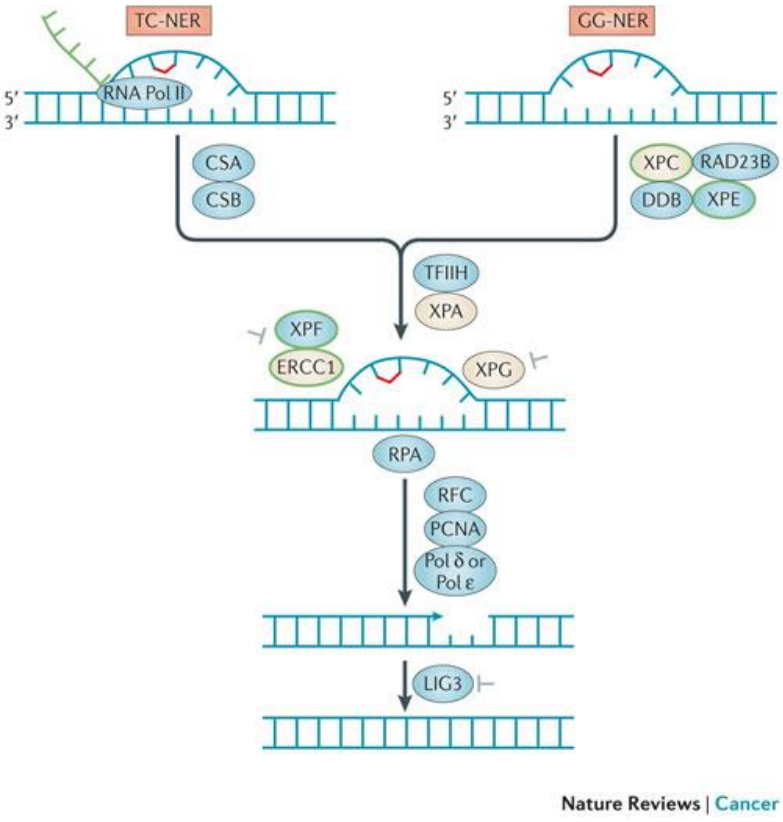

Figure 3: Nucleotide Excision Repair

Source: cf: Curtin, N. J., 2012. Nucleotide excision repair. [image online] Available at: Nature Reviews Cancer < http://www.nature.com/nrc/journal/v12/n12/fig_tab/nrc3399_F3.html#figure-title> [Accessed 13 January 2015].

In mammalian cells, the nucleotide excision repair (Figure 3) has a crucial role to avoid many hereditary diseases. In eukaryotic cells, two different pathways of nucleotide excision repair exist.

The main difference between the pathways is the recognition of the damage.

In the global genome repair (GG-NER), the Xeroderma pigmentosum group C-complementing protein (XPC) identifies together with RAD23B and DNA damage-binding protein (DDB) the damage and activates the repair mechanism. After the recognition of the damage, the lesion-sensing complex assembles. The complex includes, additionally to XPC, HR23B and centrin2 (not shown in figure). However, some types of damage, like UV-induced cyclobutane pyrimidine dimers, are not well detected by XPC. Therefore, a DNA damage-binding complex recruits XPC to the damage.

The second pathway of nucleotide excision repair is triggered by RNA polymerase II during the process of transcription that is why it is called transcription-coupled repair. The polymerase stops transcribing when it detects a bulky lesion and the large subunit is degraded. Cockayne syndrome WD repeat protein A (CSA) and CSB facilitate the recognition of the damage. The transcription factor $TF_{II}H$ acts as a helicase and unwound the DNA strands (approximately 20 nucleotides). $TF_{II}H$ and RPA (replication protein A) work in both pathways. Moreover, XPG and ERCC1-XPF provide endonuclease activity and cleave the oligonucleotide. XPG is related to the endonuclease FEN1, which acts during the base excision repair pathway. Generally, 25-30 nucleotides are cut out and a single-stranded stretch is replaced via synthesis by DNA polymerase δ/ε supported by proliferating cell nuclear antigen (PCNA) and the replication factor C (RFC). The remaining nicks are connected with a complex consisting of ligase III and XRCC1 (not shown in

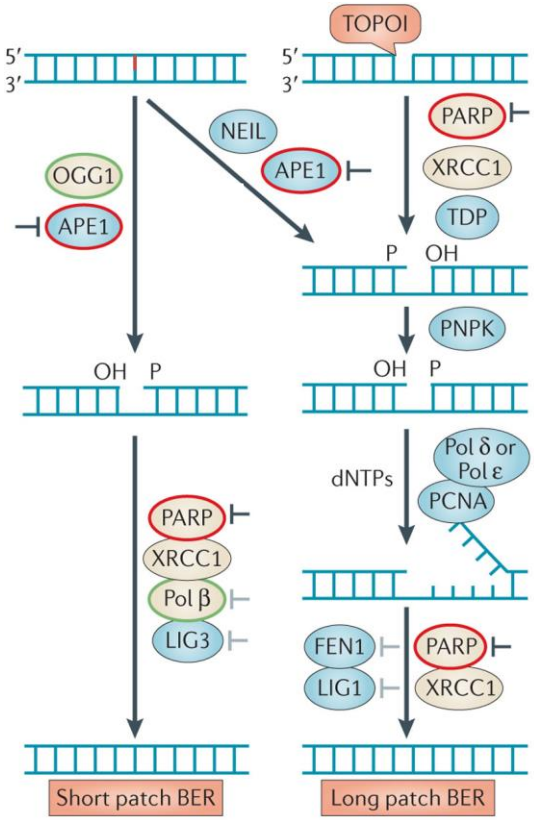

Figure 4: Base excision repair: short patch (left), long patch (right)

Source: cf: Curtin, J. C., 2012. Base excision repair. [image online] Available at: Nature Reviews Cancer <http://www.nature.com/nrc/journal/v12/n12/fig_tab/nrc3399_F2.html#figure-title> [Accessed 13 January 2015].

figure) (Klungland and Lindahl, 1997; Matsumoto and Kim, 1995; Reardon and Sancar, 2003).

The base excision repair mechanism (Figure 4) is similar to the nucleotide excision repair, except the recognition is different because only an individual damaged base must be removed. These enzymes that carry out the removal of the damaged base are called glycosylases and lyases. In the long-patch pathway, the glycosylase separates the bond between the mutated base and the deoxyribose. There are also some glycosylases, which provide lyase activity that can use an amino group to cleave the deoxyribose ring. This results in a nick of the polynucleotide chain. In figure 4, two different glycosylases are depicted, the 8-oxoguanine DNA glycosylase (OGG1) and a member of the Nei-like protein family (NEIL), which are both examples. After glycosylase activity, an AP endonuclease (APE1) cuts the polynucleotide chain on the 5'side. Polynucleotide kinase phosphatase (PNKP; a 3′ DNA phosphatase and 5′ DNA kinase) is sometimes needed to modify the ends for further repair. Within the long pathway, the DNA polymerase δ/ε is recruited to the nick and starts synthesis of the new strand. The replaced nucleotides are finally removed by the endonuclease FEN1 and the proliferating cell nuclear antigen (PCNA). The ligase I seals the gap. The short-patch pathway includes the removal of the base by lyase action and the AP endonuclease APE1, which creates the nick again. However, APE1 recruits the polymerase β, which replaces a single nucleotide, and the sealing of the nick is carried out again by the ligase XRCC1/ligase-3. The repair system is facilitated by Poly(ADP-ribose) polymerase 1 (PARP1) and XRCC1 because they are able to activate additional repair enzymes. If the damage is caused due to a stalled topoisomerase I (TOPO I), the removal of TOPO I must be carried out by tyrosyl-DNA phosphodiesterase 1 (TDP1) (Lau et al., 2000).

In order to repair double-stand breaks (DSBs) (Figure 5), again two different mechanisms can be triggered. The favored mechanism for repairing DSBs is the usage of recombination-repair. This type of repair assure no loss of important genetic information. This mechanism is conserved in all eukaryotes and carried out by so called RAD genes. There is the RAD3 group, which is involved in excision repair, there is the RAD4 group, which is needed for postreplication repair and at least the

RAD52 group, which is involved in recombination-like processes. The RAD proteins are induced at different stages of the repair mechanism of DSB. During meiotic recombination the MRN complex, consisting of MRE11, RAD50 and Nbs1, tether the free DNA ends of the DSB and link the ends. The MRN complex is recruited by BRCA1 and poly(ADP) ribose polymerase 1 (PARP1). Together with CtBP-interacting protein (CtIP) and exonuclease 1 (EXO1), the MRN complex generates single-stranded ends with free 3'-OH ends. Thus, the cell division is stopped until the damage is repaired. Furthermore, the ataxia-telangiectasia mutated (ATM) is recruited by the MRN complex, which phosphorylates MRE11, NBS1, CtIP, EXO1 and histone H2AX. The phosphorylations, further, lead to the activation of p53 binding protein 1 (53BP1) and BRCA1. The replication protein A (RPA) binds to the single-stranded DNA preventing it from being degraded. A nucleoprotein filament is formed, which is required for strand invasion of a homologous sequence. A homolog donor DNA is searched and followed by strand invasion. All these processes are carried out by different RAD proteins. RAD51 is recruited by BRCA2 to remove the RPA in order to allow the invasion of the complementary strand. DNA polymerase then synthesizes the new strand and the ends of the DSB are joined together. Chromatin-remodeling enzymes are needed then for reconfiguration of the chromatin structure at both the damage site and the donor DNA. After the repair of the DSB, the DNAs are resolved and cell division can carry on (Wolner et al., 2003).

If the replication fork stopped, the Fanconi anaemia (FANC) proteins are recruited by interstrand crosslinks (ICLs). Afterwards, repair proteins, like RPA, BRCA1, FANCN and BRCA2 are activated and HHR is carried out.

However, if no sister chromatid or homologous chromosome is available to use them as template for the repair, DSBs can also be restored using error-prone mechanisms. This mechanism is also called nonhomologous end-joining (NHEJ) and refers to the ligation of the ends. First of all, the DSB is recognized by the Ku complex, consisting of Ku70, Ku80 and the DNA-dependent protein kinase catalytic subunit (DNA-PKcs). The binding of the Ku complex recruits the MRN complex, which helps to link the two broken ends. DNA-PKcs then phosphorylates histone H2AX to allow disassociation. The two ends, which are modified by Artemis, are then joined together by the DNA ligase IV and the protein XRCC4. XLF (XRCC4–XRCC4-like factor) helps stabilizing the complex. However, during NHEJ process, mutations can be generated through the deletion and insertion of nucleotides before the two

strands are ligated. Such mutations causes more susceptibility for eukaryotic cells for radiation (Ramsden and Gellert, 1998; Ma et al., 2002).

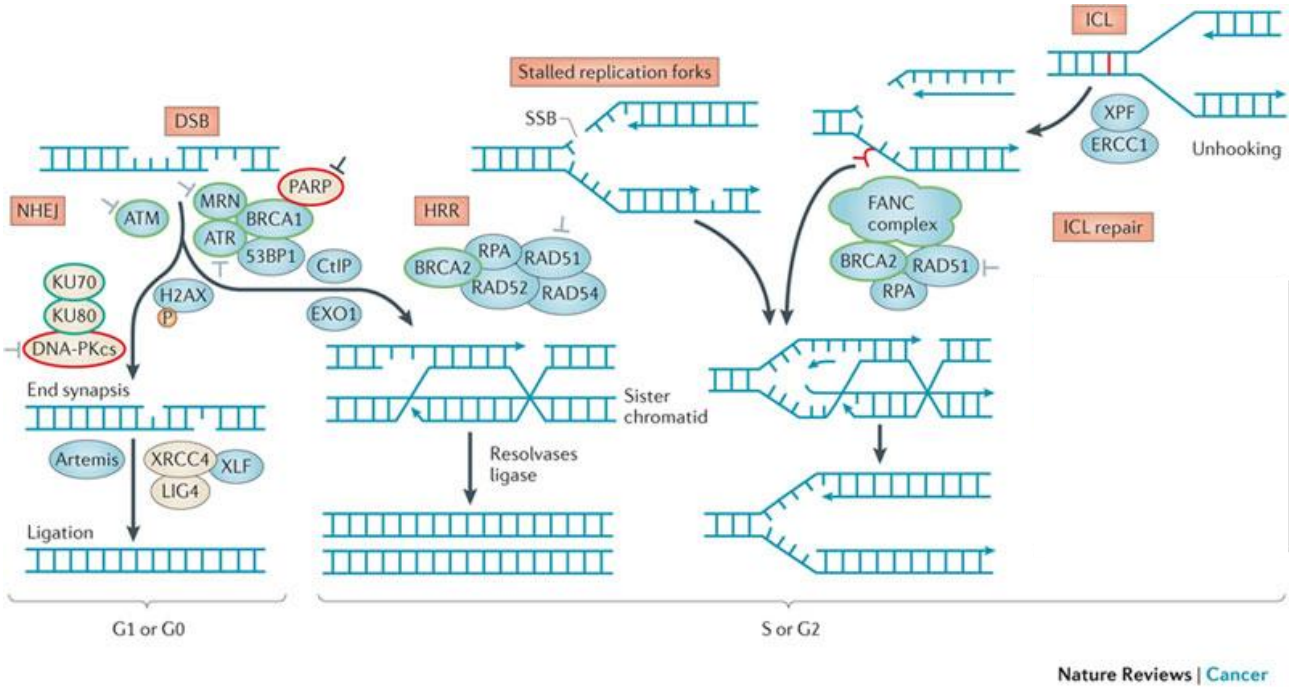

Figure 5: Repair of Double Strand Breaks

Source: cf: Curtin, J. C., 2012. DNA double-strand break and interstrand crosslink repair. [image online] Available at: Nature Reviews Cancer <http://www.nature.com/nrc/journal/v12/n12/fig_tab/nrc3399_F5.html#figure-title> [Accessed 13 January 2015].

If the process of DNA replication is finished and damages of DNA are detected and repaired, the next process of the central dogma can be discussed. In order to produce a functional protein, a copy of the DNA must be produced. The DNA itself never leaves the nucleus, therefore, a copy is required, which then serves as template for the protein translation in the cytoplasm.

Transcription

The first aspect that must be considered is that there is a huge difference between transcription of eukaryotic RNAs and prokaryotic RNAs. In bacteria, a DNA serves as template, whereas in eukaryotes, transcription is carried out on a chromatin template. This important fact must be considered in every step of transcription. First of all, the chromatin must provide an open structure, otherwise the DNA is not accessible. Moreover, nucleosome octamers must be removed from the promoter, otherwise transcription cannot start because the RNA polymerase cannot bind. The eukaryotic RNA polymerase is not able to read the DNA and the initiation of transcription includes several factors that bind to the promoter in advance. These factors are able to tether cis-acting elements on the DNA or recognize other factors that are already bound. When all these factors are bound, which is referred to as basal transcription factors, the RNA polymerase can tether to the basal transcription factor/DNA complex. This binding site is called the basal core promoter. These factors, which build the core promoter are mostly only needed for initiation of transcription, but not for transcription itself.

RNA Polymerases

In eukaryotic cells, three different RNA polymerases can be distinguished, which all have different functions. The RNA polymerase I only transcribes 18S/28S ribosomal RNAs (rRNAs), whereas the RNA polymerase II produces mRNAs and some small RNAs. The RNA polymerase III is required for transcription of tRNA, 5S rRNA and for some small RNAs.

The transcription factors involved in transcription with RNA polymerase I and III are quite simple, but in the case of RNA polymerase II, the factors form a huge complex that defines the start point of transcription. Together with the RNA polymerase II, it is called the basal transcription apparatus. The promoters itself contain sequence elements near to the start point and for RNA polymerase II, many different promoters are available. Furthermore, there are other sequences, which can be far away from the initiation site, which are called enhancer. Enhancers are cell specific and they are responsible for the expression of a promoter. The sequences are located at different distances from the core promoter, but looping of the DNA allows interaction between core promoter and enhancer, which often requires

coactivators. Moreover, enhancers can be located upstream, inside a gene, beyond the end of a gene and in either orientation. Enhancers also bind transcription factors, which can act in both a positive and negative way. Mostly, the activity of a promoter is enormously increased by the presence of an enhancer. Enhancers start working through the binding of a combination of transcription factors, which then regulate the promoter and gene expression. The enhancer, then, interacts with the basal apparatus at the core promoter element and functions to raise the concentration of activator near to the promoter. The more activators are tethered to the promoter, the higher is the level of gene expression. Limitations of the activity of enhancers are not understood completely yet, but it could be shown that the enhancer normally activates the nearest promoter and when located between two promoters, only one of them is activated. The function itself may be limited by so-called insulators, which are elements in DNA that prevent the enhancer from interacting with promoters beyond the insulator (Mueller-Storm, Sogo and Schaffner, 1989).

If the transcription factors contribute transcription, it is referred to as enhancer, but if more negative regulators than positive ones bind to the sequence to control transcription, then they are called silencer. Depending on the gene and in what way the gene is regulated, the specification of a cell is determined. These special types of promoter can be turned on and off, which requires either activators or repressors. However, there are also promoters, which are expressed constitutively. This type of promoter regulates the expression of genes that are required in every cell, also called housekeeping genes.

The main focus of this paper is put on the RNA polymerase II, which is responsible for the transcription of the messenger RNA. The RNA polymerase II can be found in the nucleoplasm of the nucleus. Most of the cellular activity is accredited to this RNA polymerase. Additionally to the premature messenger RNA, it is also involved in the production of the heterogeneous nuclear RNA (hnRNA) and many more. hnRNA means everything but rRNA and tRNA in the nucleus. Mature mRNA can only be found in the cytoplasm and is the least abundant one of the three main RNAs. mRNA accounts only for 2-5 %, whereas rRNA refers up to 90%.

Eukaryotic RNA polymerases are large proteins that are composed of several subunits. The carboxy-terminal domain (CTD) is the largest subunit of the RNA polymerase II. The CTD includes many repetitions of a consensus sequence, which consists of seven amino acids. In mammals, there are about 50 repeats, which are

important for regulating the initiation of transcription, elongation of transcription, posttranscriptional processing of mRNA and transport of mRNA into the cytoplasm (Burke and Kadonaga, 1996; Singer, Wobbe and Struhl, 1990; Smale and Baltimore, 1989).

Transcription Initiation

As mentioned before, RNA polymerase II is dependent on the basal transcription factors that bind to the promoter sequence for initiating transcription. The general factors are described as TFIIX, where X defines the individual factor. As soon as the basal transcription apparatus is formed, which means the assembly of the transcription factors at the start point, the process of transcription can start. A typical core promoter consists of the three major elements, which are a TATA-box, an initiator (Inr) and a downstream promoter element (DPE). However, there are also several minor elements that also define the promoter. The initiator often starts with an adenosine (A), which is flanked on either side by pyrimidines. About 25 nucleotides upstream the initiator, the TATA-box is located. As its name indicates, the consensus sequence is TATAA and has a relatively contant location upstream of the start point. The TATA-box is normally followed by three more A-T base pairs and there is a tendency for the TATA-box to be surrounded by G-C rich sequences. However, it is assumed that about 50% of the promoters in eukaryotes do not have a TATA-box, but consist of another element, the downstream promoter element (DPE). So, a core promoter compromise usually of either TATA-box and Inr or DPE and Inr, although other combinations with minor elements are also functional (Burke and Kadonaga, 1996; Singer, Wobbe and Struhl, 1990; Smale and Baltimore, 1989; Bregman et al., 2011).

Each type of RNA polymerase requires a TATA-binding protein (TBP), which is a positioning factor that tethers to the promoter. However, the name of TBP is misleading because the protein also binds promoters without the TATA-box. For the RNA polymerase II, the TBP and approximately 14 TATA box binding protein associated factors (TAFs) assemble and form a complex called TFIID. This complex is only responsible for recognition of the promoter, so that the RNA polymerase II can bind subsequently. If the promoter provides a TATA-box, the TBP tethers the TATA-box, but if there is no TATA-box available, the TAFs are able to recognize other promoter elements, like the Inr and DPE.

Unlike the majority of DNA-binding proteins, the TBP always tethers the minor groove of the DNA. Moreover, the protein bends the DNA by ~80°, so that the minor groove is accessible. The bending of the DNA strand promotes the binding of transcription factors and RNA polymerase because a closer association is possible (Martinez et al., 1994; Verrijzer et al., 1995; Wu and Parkhurst, 2001).

In eukaryotic cells, there are three different basic types of gene promoters. One of them is an inactive gene in closed chromatin. Then there is one that is a potentially active gene on open chromatin with RNA polymerase bound, referred to a poised gene. This type of promoter requires a second signal to start transcription, which can be carried out by heat-shock genes. The third type is a gene that is turned on in open chromatin.

Due to the fact that RNA polymerase II promoters differ in their structure, a model is described, that summarizes the major events. Once, the basal transcription factors are bound to the DNA and the RNA polymerase II is recruited to the apparatus, the activity of the polymerase is dependent on enhancer-binding transcription factors. The efficiency and specificity of the recognition of a promoter is controlled by short sequences farther upstream, which are again detected by a different group of transcription factors, called activators. Activators have an influence on the formation of the initiation complex.

The first step of initiating a TATA-box-containing promoter in open chromatin is started when the TBP subunit of the TFIID binds to the TATA-box in the minor groove of DNA. Then TFIIB tether downstream of the TATA-box, near the TBP. TFIIB provides the binding site for RNA polymerase and is responsible for the directionality of the polymerase binding. The

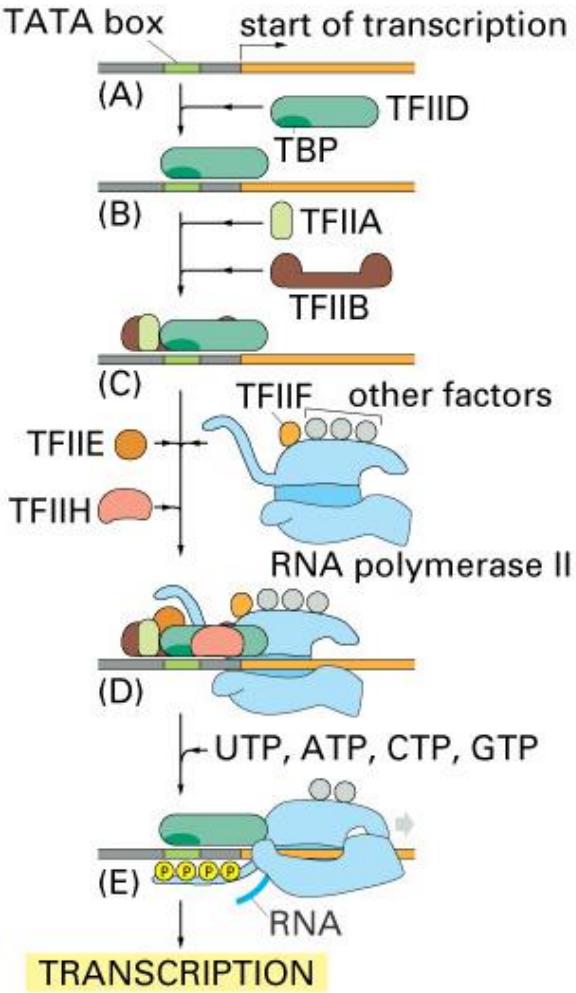

Figure 6: Transcription initiation complex assembly

Source: Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P., 2004. Essential Cell Biology. [e-book] New York: Garland Science. Available at: Scribd <http://de.scribd.com/doc/202451270/Biology-Essential-Cell-Biology-2nd-Edition-Bruce-Alberts-Dennis-Bray-Garland-Science-2004> [Accessed 12 January 2015].

transformation from the closed to the open complex means the recruitment of the RNA polymerase II to the TFIID/TFIIA/promoter DNA complex and the selection of the transcription start site (TTS). The factor TFIIF is another factor, which is required for the preinitiation complex (PIC) formation and, together with TFIIB, selects the TTS. The complex of TBP and TAFs interact further with the CTD tail of RNA polymerase (Figure 6).

TATA-less promoter contain the same general transcription factors, including TFIID. However, it is not the TATA-box that defines the binding site. The Inr is the positioning element. The TFIID directly detects and binds the Inr. Other TAFs in TFIID detect the DPE element downstream from the initiation point (Buratowski et al., 1989; Burke and Kadonaga, 1996; Kostrewa et al., 2009; Lui et al., 2011).

In cells, as mentioned before, three basic types of promoters can be found, which differ due to their activity. If the chromatin is closed, it is referred to as an inactive gene. A poised gene is a gene that is potentially active and an RNA polymerase is tethered. The complete basal apparatus is formed, but to initiate transcription another signal is required. For instance, heat shock proteins are transcribed from poised genes because a rise in heat activates transcription. The third type of promoters is a gene that is turned on in open chromatin. Gene activation furthermore involves so called noncoding RNA transcripts, which regulate transcription of adjacent or overlapping proteincoding genes. These cryptic unstable transcripts, also called CUTs, are much more common due to recent studies. Several active promoters have transcripts produced upstream of the promoters (PROMPTs). PROMPTs are transcribed in sense and antisense directions relative to the downstream promoter. These transcripts are believed to have a significant role within the transcriptional regulation.

Once the basal transcription apparatus has assembled, RNA polymerase binds to the apparatus and transcription is initiated. The activity of the polymerase is then regulated by enhancer- binding transcription factors. However, RNA polymerase II promoters can differ in structure, as mentioned before. The promoter is recognized with high efficiency and specificity depending upon short sequences upstream that are noticed by another group of transcription factors, sometimes referred to as activators. Many different elements can contribute to the function of the promoter, but there is no element that is essential for all promoters (Buratowski et al., 1989; Burke and Kadonaga, 1996; Kostrewa et al., 2009; Lui et al., 2011).

Furthermore, in order to activate the promoter, demethylation at the 5' end is required for transcription. Methylation at the promoter hampers the process of transcription, so the methyl groups must be removed in advance. The methylations at promoters occur on the 5' end of C at CG doublets, also referred to as "CpG" doublets and are carried out by two different DNA methyltransferases. In mammals, an enzyme called TET (ten eleven translocation) carries out the demethylation of DNA converting 5mC to 5-hydroxymethylcytosine as first step of DNA damage excision repair mechanism. Another way to get rid of the methylations is the usage of restriction enzymes that cleave at restrictions sites involving the CpG doublet.

However, there are also genes that are expressed, even though they are methylated. So the need for demethylation is not universal, but the general rule is that methyl groups at CpG doublets prevent transcription and the removal of these groups is necessary for gene expression (Zemach et al., 2010)

Originally, DNA methylation served as a mechanism to avoid the insertion of unwanted sequences. Most of the methylations are located at so-called CpG islands in the 5' regions of genes. These islands are characterized by an increased occurrence of the dinucleotide CG. An average C-G content of about 60% is present in these islands compared to the normal C-G content of DNA that is about 20%. The human genome contains between 29,000 and 45,000 of these CpG islands (Antequera and Bird, 1993; Boyes and Bird, 1991).

Transcription Elongation

The clearance of the promoter is an essential step that determines if a poised gene or an active gene is copied. This important key step is regulated by the so-called enhancers. The transcription factors that bind enhancers do not share binding sites with the promoter itself, but rather bind to a coactivator that tethers promoter elements. The Mediator, a coactivatior, is the most common one in eukaryotic cells. It can involve 30 subunits or more and they can be type- or gene specific referring to their form. The Mediator is conserved from yeast to humans and both, poised and active genes need the interaction of the transcription factors, which are bound to enhancers, with the promoter through looping the DNA. The loop is formed with the Mediator as the intermediate of DNA and transcription factors.

$TF_{II}E$ and $TF_{II}H$ are the last two factors that join the initiation complex. The assembly of $TF_{II}E$ causes the extension of the boundary of the region protected

downstream by another turn of the double helix. $TF_{II}H$ is known as the only general transcription factor that has several independent enzymatic activities including an ATPase, helicases of both polarities and a kinase activity, which is responsible for the phosphorylation of the CTD tail of RNA polymerase II. $TF_{II}H$ is an additional factor that also has a crucial role in elongation of transcription. The interaction between $TF_{II}H$ and the DNA downstream of the initiation apparatus allows the RNA polymerase to leave the promoter. For the polymerase to move on a linear template, ATP hydrolysis, $TF_{II}E$ and the helicase activity of $TF_{II}H$ are necessary. With a supercoiled template, this requirement is bypassed. This further means that both, $TF_{II}E$ and $TF_{II}H$ are needed for the melting of the DNA so that the polymerase is able to move along. However, $TF_{II}H$ subunits are responsible for the actual melting (Goodrich and Tjian, 1994).

Before RNA polymerase II starts the process of transcription, it stutters. This means that the RNA polymerase starts transcribing, but stops after a brief distance. The same mechanism occurs if a false nucleotide is inserted or if the 3' end is improperly base paired, which is an important part of the fidelity mechanism. The small oligonucleotides, which are produced during this process are unstable and are degraded very fast. This process is explained as a type of proofreading of the promoter (Cheung and Cramer 2011; Hendrix et al., 2008). In order to start the elongation, a kinase complex, called positive elongation factor (P-TEFb) is needed, which contains CDK9 kinase. P-TEFb interacts with the CTD of the polymerase and phosphorylate it further on serine 2 of the heptapeptide repeat. This effect is only required at some promoters but not at others, which is not fully understood yet. The phosphorylation step is an important step to release RNA polymerase II from the initiation apparatus, so that the transition to the elongation step is possible. The phosphorylation step is a dynamic process, which is catalyzed and regulated by several protein kinases like P-TEFb and phosphatases. Moreover, most of the basic transcription factors are released from the promoter during this process (Montanuy et al., 2008; Luse and Spangler and Ujvari, 2011).

The CTD is also important because it is involved in processing mRNA while it is synthesized and afterwards, when the mRNA has been release by RNA polymerase II. Furthermore, the phosphorylated CTD serves as an anchor point for other proteins, like the cap-binding enzyme (guanylyl transferase), which produces the G residues on the 5' end of newly synthesized mRNA. This allows quick modification of

the 5' end of the mRNA as soon as it is formed. Subsequently, serine 2 phosphorylation by P-TEFb causes the recruitment of a bunch of proteins, called SCAFs, to the CTD, which in turn bind to splicing factors. Even some components of the cleavage/polyadenylation apparatus used for transcription termination also bind to the CTD phosphorylated at serine 2. The binding also occurs at the beginning of transcription initiation, so that, as soon as transcription terminates, the processing of the 3'end can be carried out. Additionally, the transport of the mRNA from the nucleus through the nuclear pore is controlled by the CTD (Fong and Bentley, 2001).

As soon as the promoter clearance, the release of the promoter, has occurred and initiation factors are released, the transition to the elongation phase is induced. The transition complex itself comprises the RNA polymerase II, the basal factors TFIIE and TFIIH, and all of the enzymes and factors bound to the CTD. Elongation factors, like TFIIF and TFIIS and others to prevent inappropriate pausing, may be present in another big complex, which is called super—elongation complex (SEC) (Figure 7). The super elongation complex (SEC) is activated by the mediator complex and the RNA polymerase II (Pol II)-associated factor 1 (PAF1) complex, which are both Pol II transcriptional co-activators. The bromodomain-containing protein 4 (BRD4) interacts with P-TEFb and is known to be necessary for elongation, however, the precise function is not known yet. This whole complex, including the RNA polymerase II and general transcription factors (GTFs), now has to transcribe the chromatin template. Meanwhile, the RNA polymerase functions as a DNA repair mechanism. If the polymerase encounters DNA damage in the template strand, it stalls because it is not able to use the damaged DNA. The stalling of the RNA polymerase activates a pair of proteins, CSA and CSB (see Eukaryotic Repair Systems). TFIIH is also required and responsible for the repair of the DNA damage. Furthermore, the repair function may need alterations or degradation of the stalled RNA polymerase, so the large subunit of the RNA polymerase is often

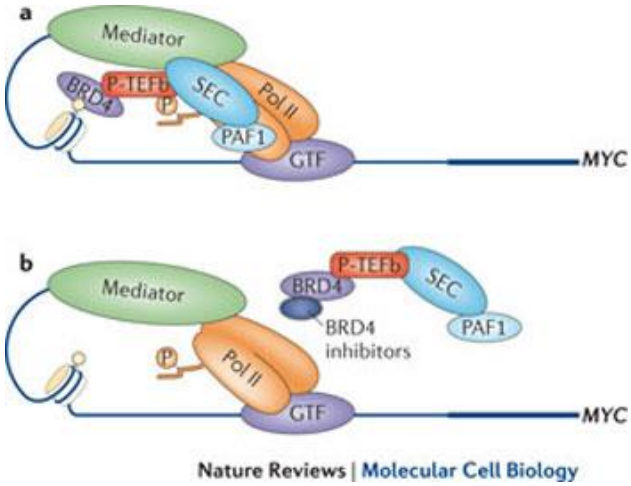

Figure 7: Transcription Elongation

Source: cf: Luo, Z., Lin, C. and Shilatifard, A., 2012. The SEC in cancer progression. [image online] Available at: Nature Reviews Molecular Cell Biology <http://www.nature.com/nrm/journal/v13/n9/box/nrm3417_BX1.html> [Accessed 13 January 2015].

removed by the ubiquitylation pathway. However, the connection between the transcription/repair apparatus and the degradation of RNA polymerase is not yet understood completely (Lans et al., 2010).

RNA Splicing And Processing

As soon as the transcription of RNA is finished, it requires further processing to become mature and functional. Genes can be very different in aspect to the number of introns they include. A typical mammalian gene contains about 7 to 8 exons spread out over about 16 kb. Exons are relatively short (100-200 bb), compared to introns (> 1 kb).

The primary transcript is an exact copy of the DNA, with introns and exons. However, to become a functional mRNA, introns are cleaved out and exons are joined together. In all eukaryotes, the removal of the introns is the major step of processing. This specific process is called RNA splicing. Such interrupted genes encode relatively abundant proteins and so splicing is involved in the production of a huge number of the total mRNA. This amount is estimated at around 50%.

Splicing as well as other processing steps are carried out in the nucleus including the capping of the 5’ end, the removal of the introns and the polyadenylation of the 3’ end. The mature mRNA is then transported via nuclear pores to the cytoplasm, where translation into proteins is performed.

Capping of the 5’ end

Transcription generally starts with a nucleoside triphosphate, normally a purine (A or G). The first nucleotide keeps the 5’ triphosphate group and forms a normal phosphodiester bond from the 3’ position to the 5’ position of the following nucleotide. After transcription, a guanine base is added to the end of the mRNA molecule. The terminal guanine is added by an enzyme called guanylyl-transferase, which, in mammals, has two catalytic activities. Firstly, it works as a triphosphase, which removes the two phosphates in GTP. Secondly, it functions as a guanylyl-transferase, which is responsible for the fusion of the guanine to the original 5’ triphosphate terminus of the mRNA. This new G residue added is in reverse orientation from all the other nucleotides and is often the substrate of many methylation events. This specific structure is referred to as a cap.

The addition of a methyl group at the 7 position of the terminal guanine is a crucial step that is carried out by guanine-7-methyltrasferase. After transcription initiation, RNA polymerase II halts after about 30 nucleotides to allow the addition of the cap by the capping enzymes. This is one of the most important events because without the cap, the nascent mRNA is vulnerable to 5' to 3' exonucleases, which would further lead to the degradation of the mRNA and the RNA polymerase II would disassemble. This mechanism correspond to the first checkpoint for transcription reinitiation from the pausing event. Eukaryotic mRNAs usually inheritate only one methyl group in the terminal guanine, which is called monomethylated cap. Noncoding RNAs, like those who help carrying out the splicing, are often methylated three times, which is referred to as trimethylated cap.

The major function of the cap is the protection of the mRNA from being degraded. mRNA turnover is therefore regulated by enzymes, which remove the cap. In the nucleus, the cap is detected and bound by cap binding proteins (CBP). The heterodimer CAP80/20 stimulates splicing of the first intron and interacts with the export machinery (TREX complex) to help transporting the mature mRNA into the cytoplasm. In the cytoplasm, eIF4FA tethers the cap and activates the translation process (Mandal et al., 2004; McCracken et al., 1997)

Splice Sites

Comparing the nucleotide sequence of a mature mRNA with that of the original gene, the junctions between exons and introns can be detected. The so-called splice sites contain well-conserved consensus sequences, so the recognition of exon-intron junction is possible. The presumed junction of the intron is highly conserved. This identifies the sequence of a generic intron as: GU……AG. The intron starts with the dinucleotide GU and terminates with AG. It is often referred to as GU-AG rule, which represent directionally defined ends. This ends in the intron are named as the 5' splice site (left or donor site) and the 3' splice site (right or acceptor site). This consensus sequences are recognized during splicing. However, there is also a small number of introns, which do not follow the GU-AG rule and provide different consensus sequences. They, then, follow an AU-AC role because of the conserved AU-AC dinucleotides at both ends of each intron (Graveley, 2005; Krainer et al., 1984).

Pre-mRNA Splicing

The general process of splicing could be shown not to require subsequent translation. Moreover, splicing can also be performed without the presence of a 5' cap or 3' poly(A)tail. The first step of the process is a nucleophilic attack by the 2'-OH on the 5' splice site. The first exon is linearized and the following intron-exon molecule forms a loop structure, which is also called lariat. Within this step, the 5' terminus generated at the end of the intron transesterificates and attaches by a 2' to 5' bond to an adenine, which is located in a branch site. Secondly, the free 3'-OH of the exon that has been released now attacks the bond at the 3' splice site. Before the splicing event, two 5' to 3' bonds were present at the exon-intron splice sites. One bond has been substituted by the 5' to 3' bond between the exons and one has been replaced by the lariat, which means the 2' to 5' bond. Afterwards, the lariat is debranched in order for the intron to be rapidly degradable.

Short consensus sequences at the 5' and 3' splice sites and at the branch site are mostly present, however, the deletion of most of the sequence of an intron does not impede the splicing process. So the necessity of a specific conformation is not decisive. However, the branch site is required for the identification of the 3' splice site. This specific site is located about 18 to 40 nucleotides upstream of the 3' splice site. If the branch site is deleted or mutated in multicellular eukaryotes, related sequences are used. In yeast, for example, the splicing process fails to occur, if the branch site is eradicated. The crucial role of the branch site is to identify the closest 3' splice site as the connection point for the 5' splice site. This can be explained due to the fact that protein complexes are located at these two sites, which allow interactions (Reed and Maniatis, 1985).

The protein complex is a component of the splicing apparatus. This complex brings the 5' and 3' splice site in close proximity. The complex assembles sequentially on the pre-mRNA and passes through several stages before forming the final active complex, which is referred to as spliceosome. The spliceosome does not only involve proteins, but also RNAs are included. These RNAs are included in ribonucleoprotein particles, which are located in the nucleus and the cytoplasm. In the nucleus, they are called small nuclear RNAs (snRNAs) and in the cytoplasm, they are called small cytoplasmic RNAs (scRNAs). Naturally, they are only found complexed as ribonucleoprotein particles (snRNPs), which are often called snurps or scyrps. Moreover, in the nucleus, another type of RNA is located, which are called

small nucleolar RNAs (snoRNAs), which work in processing ribosomal RNA. However, the snRNPs, which are involved in splicing together with many additional proteins, form the spliceosome, which is a large complex. Five snRNAs account for more than a quarter of its mass. Furthermore, there are 41 additional proteins, which make up almost half of its mass. Additionally, there are about 70 other proteins, which are referred to as splicing factors and these are proteins, which are needed for the assembly of the spliceosome, for binding to the RNA and for providing a platform for transesterification reactions.

The spliceosome assembles on the pre-mRNA and migrates through an intermediate state, which contains the linear exon molecule and the intron-exon lariat. The snRNPs are responsible for the formation of the structure of the spliceosome, which means that they provide protein-protein interactions, as well as protein-RNA interactions and also RNA-RNA interactions. All of the snRNAs involved in the splicing process are present in conserved forms in all eukaryotes. Each snRNP comprises of a single snRNAs and many proteins. The snRNPs responsible for splicing are U1, U2, U4, U5 and U6. U4 and U6 often form a U4/U6 particle.

Some of the proteins included in the snRNPs are directly involved in the splicing process, other have structural roles or are just necessary for interaction or assembly of the snRNP particles, like the splicing factors. Therefore, the splicing factors do not have a crucial role within the splicing process itself, but they are needed to maintain the structure and help the assembly of the spliceosome (Grabowski, Seiler and Sharp, 1985; Krainer and Maniatis, 1985; Black, Chabot and Steitz, 1985).

First of all, the U1 snRNP binds to the 5’ splice site by base pairing between its single-stranded 5’ terminus and a stretch of four to six bases of the 5’ splice site. U1 snRNA and the 5’ splice site must be complementary, which is needed for splicing. Afterwards, the so-called commitment complex is formed, which means the interaction of the branch point binding protein (BBP or SF1) with the branch point sequence, and the heterodimer U2AF interacts with the polypyrimidine tract between the branch point sequence and the constant AG dinucleotide at the end of each intron. In mammalian cells, this complex is also called E complex, which stands for early complex and does not need ATP for assembling.

In mammalian genes, the consensus sequences are only loosely conserved and additional protein factors are required for the formation of the E complex. These

factors are called SR proteins, which constitute a family of splicing factors. They contain an RNA-recognition motif and they are also able to perform protein-protein interactions. SR proteins help to stabilize the base pairing between U1 snRNP and the 5’ splice site. However, SR proteins can also bind to 3’ splice site bound U2AF. These proteins, therefore, are referred to as splicing initiators in multicellular eukaryotes.

Introns are long and highly variable in length in multicellular eukaryotic genomes and there are many sequences that resemble real splice sites in them. This makes the paired recognition of the 5’ and 3’ splice sites inefficient or even impossible. The solution of this problem is the process of exon definition, which takes advantage of normally small exons. During exon definition, the U2AF heterodimer binds to the 3’ splice site and U1 snRNP binds to the 5’ splice site downstream from the exon sequence. These complexes, which are formed across the exon, then link the 3’ splice site to the upstream 5’ splice site and the downstream 5’ splice site to the next downstream 3’ splice site across introns. There, the spliceosome assembly steps are then able to occur. Finally, the exon definition mechanism provides a mechanism to only allow adjacent 5’ and 3’ splice sites to be paired and linked by splicing (Abovich and Rosbash, 1997; Kohtz et al., 1994; Robberson and Berget, 1990).

As soon as the E complex is formed, other snRNPs and factors assemble at the complex in a certain order. The first ATP dependent step involves the association of U2 snRNP to U1 snRNP on the pre-mRNA by binding to the branch point sequence, which requires ATP. The pre-spliceosome complex is then referred to as A complex. Afterwards, the B1 complex is formed, in which U5 assembles with U4/U6 to a trimer and binds to the A complex. This complex, then, contains all components needed for splicing and can be regarded as spliceosome. However, after the release of U1, the complex is converted into the B2 complex and this step is necessary because this allows U6 to tether to the 5’splice site and therefore replaces U1. As soon as U4 is released, the catalytic reaction is activated, which also takes place in the conversion from B1 complex to B2 complex. The release of U4 further allows U6, which formed a dimer before, to interact with U2 and another part of U6 forms an intramolecular hairpin. U4, therefore, regulates the ability of the spliceosome to be activated.

During the splicing process, the snRNAs strongly interact with the substrate RNA. The snRNPs interact with each other, but also with the pre-mRNA substrate. The activated spliceosome complex, also called C complex, run through transesterification reactions and the formation of the lariat at the branch site is responsible for determining the use of the 3' splice site because the 3' splice site is transesterificated secondly (Figure 8) (Konarska and Sharp, 1987; Tseng and Cheng, 2008; Zhuang and Weiner, 1986)

The splicing process itself can occur after transcription, but also during the transcription process. In general, introns near the 5' end are cleaved during transcription, whereas introns near the 3' end are removed afterwards. The machineries for 5' capping, intron removal and 3' polyadenylation reveal physical interactions with the core machinery for transcription. This means that transcription and splicing are not only temporally linked, but also in a functional way. The CTD again plays a crucial role because it provides something like a loading pad for various RNA-processing factors.

However, also the transcriptional process could been shown to benefit from the splicing event. The initial transcriptional pausing event adjacent the promoter seems to be regulated by 5' capping enzymes. Moreover, the 5' capping enzymes play an important part during transcriptional elongation and the 3' end formation of mRNA. RNA processing is, furthermore, linked to downstream processes, like mRNA export and stability control. mRNA containing introns, cannot be exported. Moreover, mRNA that has been spliced successfully can be exported more efficiently, than the RNA that is derived from cDNA. This again leads to the assumption that the splicing process helps exporting mature mRNA.

After the splicing process has succeeded, a complex called exon junction complex (EJC), is formed on the exon—exon junction. This complex recruits several RNA-binding proteins, which further help to transport the mRNA into the cytoplasm. The cap binding CBP20/80 complex seems to directly bind to the mRNA export machinery (TREX complex), which helps to cleave the first intron near the 5' end and facilitates the export. Another key factor, that helps exporting the mRNA, is Aly/REF export factor (Aly/REF). Aly/REF is part of the EJC and has the ability to directly interact with the transporter TAP (Transporter associated with antigen processing).

The EJC also plays a decisive role in recruiting other proteins, which for example promote decapping in the case of aberrant mRNA. Normally, the EJCs are

removed during the first round of translation by the translating ribosome, however, if there is for instance a premature stop codon (PTC), caused by point mutations or alternative splicing, the ribosome disassociates before reaching the natural termination codon. The natural termination codon is located at the last exon. If the ribosome is not able to remove the last EJC or even several EJCs, the recruitment of decapping enzyme is the consequence. This leads to rapid degradation of the nonsense mRNA. The process is called nonsense-mediated mRNA decay, which represents a quality control mechanism that prevents the production of shortened proteins of mRNA, which contain a PTC (Cheng et al., 2006; Das et al., 2007; Le Hir et al., 2000; Zhou et al., 2000).

The transcription of an interrupted gene into an RNA that further is processes to a single type of spliced mRNA, offers no certainty about the assignment of exons and introns. However, most of the transcripts of mammalian genes follow a structure of alternative splicing, which means that one RNA can give rise to many different mRNA sequences. Alternative splicing is, therefore, not only the result of mistakes that are carried out by the splicing machinery, it is even more a way to generate many different gene products from one single genes. Such a single gene can undergo several different modes of alternative splicing, including intron retention, alternative 5’ splice sites, alternative 3’ spice sites, exon inclusion or skipping, mutually exclusive exons, combinatorial exon selection, alternative promoter/splicing and alternative polyadenylation/splicing. Mutually exclusive exons are normally regulated in a tissue-specific connection.

Gene expression itself can be affected in two different ways by splicing. One way is to mingle exons and, therefore, create a huge structural diversity by including or leaving out exons. This type of processing can alter or modify the function of the produced protein. However, it can also result in the opposite function of the spliced product.

The second way influences significantly the half-life of mRNAs. In many cases, the main reason for alternative splicing may be the production of several primary transcripts, which carry a so-called premature stop codon (PTC). The recognition of these PTCs leads to rapid degradation of the aberrant mRNAs and serves as regulation of gene expression (Ge and Manley, 1990; Krainer, Conway and Kozak, 1990; Wang et al., 2008).

Alternative splicing represents a mechanism that might not only be the result from mistakes occurring during splicing, it is also seen as a mechanism that is evolutionary conserved to preserve gene expression regulation at RNA processing level. The process of alternative splicing regulation is a complex mechanism that is carried out by several RNA-binding trans-acting splicing regulators. These regulators bind to specific RNA elements in both exons and introns adjacent to the alternative splice site. The binding of these factors can have various influence on the alternative splice site. Positive splicing regulators bind to exons and improve to the selection. The corresponding cis-acting elements are referred to as exonic splicing enhancers (ESEs). One example of ESEs are SR proteins, which are the best characterized ESE-binding regulators until now. In opposition to the ESEs, there are RNA-binding proteins, like hnRNP A and B, which bind also to exonic sequences to hamper splice site selection. These cis-acting elements are thus called exonic splicing silencers (ESSs).

Similar to the exon splice site regulator, there are also regulators present, which bind to intron sequences. In contrast, these are known as intronic splicing enhancers (ISEs) and intronic splicing silencers (ISSs).

Depending on the location they bind to, the Nova and Fox families of RNA-binding splicing regulators can act as enhancer or silencer of splicing. Moreover, the level of expression of both regulators is tissue specific, but how such specific alternative splicing processes are controlled has not been figured out clearly. It is only known that these splicing

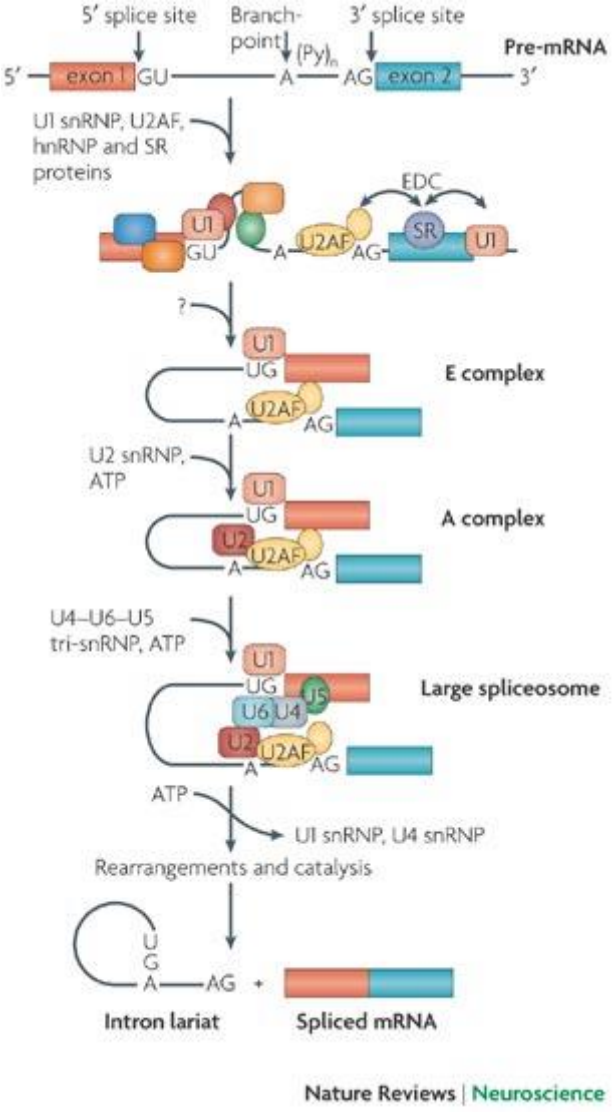

Figure 8: Pre-mRNA Splicing Process

Source: Li, Q., Lee J.-A. and Black, D. L., 2007. The pathway of spliceosome assembly in vitro. [image online] Available at: Nature Reviews Neuroscience <http://www.nature.com/nrn/journal/v8/n11/box/nrn2237_BX1.html> [Accessed 13 January 2015].

regulators work in a way that improves or hampers the recognition of specific splicing signals.

One of the best-known enhancers and silencers are the SR proteins, as well as the hnRNA A and B proteins. SR proteins have the ability to facilitate the assembly of the spliceosome through binding of the SR proteins to ESEs. This, further, improves and stabilizes the interaction between U1 and the 5' splice site and U2AF binding to the 3' splice site. However, the function of the SR proteins contributes to both, constitutive and alternative splicing. In contrast, if hnRNP A and B proteins bind to their recognitions sites, SR proteins and other core spliceosome components are not able to bind anymore, which reveals that hnRNA A and B proteins are so-called silencers.

Due to the close linkage between transcription and the splicing process, promoters can also affect alternative splicing, as well as the rate of transcription during the elongation event. In the case of specific promoters, different transcriptions factors may be recruited, which have an effect on transcription elongation and, therefore, the process of alternative splicing can be influenced (Cramer et al., 1999; De la Mata et al., 2003; Fairbrother et al., 2002; Sharma, Falick and Black, 2005; Wang et al., 2004; Zhu et al., 2001).

Polyadenylation of the 3' end

Until now, it is not completely clear, whether the termination event is at a specific site or not. Generally, termination takes place more than 1000 bp (base pairs) downstream of the site, which refers to the mature 3' end of the mRNA. However, often non-specific terminator sequences, which are located in long terminator regions, also lead to the halt of the ribosome. So, the actual nature of an individual termination site is unknown.

To produce a mature mRNA, first of all, a cleaving step must be performed followed by the polyadenylation. The poly(A)tail is an important feature of the mRNA, which is needed for the transportation into the cytoplasm. The prevention of the poly(A)tail addition leads to the result that no mRNA is present in the cytoplasm, which could be shown by the addition of cordycepin. Cordycepin is a 3'-deoxyadenosine, which does not impair the transcription process itself, but the mRNA, then, cannot be exported from the nucleus. Moreover, the poly(A)tail protects the mRNA from rapid degradation by 3' to 5' exonucleases.

As soon as the RNA polymerase migrates over the site corresponding to the 3' end and sequences are detected that are targets of the endonucleolytic cleavage, endonucleases are recruited. After the cleavage, the polymerase continues transcription and produces an unprotected 3' end. This 3' end signals transcription termination. At the site of cleavage, two cis-acting signals are present. There is an AAUAAA motif, about 11-30 nucleotides upstream from the site and a U-rich or GU-rich element downstream of the site. The upstream motif is required for the cutting step and polyadenylation because mutations or deletion within this site prevents the formation of the poly(A)tail.

The synthesis of the cleavage, as well as the poly(A)tail, are dependent on cleavage and polyadenylation specific factors (CPSF). These factors, including many subunits, bind directly to the AAUAAA motif and to the cleavage stimulatory factor (CstF) complex. One of its components also bind directly to the downstream motif. CPSF and CstF can strengthen each other through recognizing the polyadenylation signals.

The polyadenylation process requires two specific enzymes, which are an endonuclease (subunit of CPSF) to cleave the RNA and a poly(A) polymerase (PAP) to generate the poly(A)tail. PAP gains specificity for the reaction when bound to other components. The polymerase must go through two stages, which must be distinguished. First of all, a quite short oligo (A) sequence is added to the 3' end. This reaction requires the AAUAAA motif and the poly(A) polymerase. Secondly, the nuclear poly(A) binding-protein (PABP II) tethers the oligo(A) tail and allows its extension to a length of approximately 200 residues. Normally, the poly(A) polymerase dissociates after each nucleotide, however when bound to CPSF and PABP II, it works processively to extend an individual poly(A)tail. After the generation of the poly(A) tail, PABP II binds stoichiometrically to the poly(A)tail, which then limits the activity of the PAP. The cytoplasmic poly(A) binding protein (PABP I) also tethers the poly(A)tail followed by the transportation of the mature mRNA into the cytoplasm. PABP I functions as protection from 3' to 5' exonucleases, as well as binding site for the translation initiation factor eIF4G. In the cytoplasm, the mRNA forms a loop, which is formed of a protein complex containing both the 5' and the 3' ends of the mRNA. The addition of the poly(A) tail, therefore, provides not only stability, but also helps to facilitate initiation of translation (Cornway and Wickens, 1985; Fox, Sheets and Wickens, 1989; Gil and Proudfoot, 1987; Takagaki, Ryner and Manley, 1988).

Transcription Termination

There is no discrete termination site, which is recognized by RNA polymerase II (Pol II), so it continues transcription about 1.5 kb past the 3'end, where termination is triggered. Termination is activated by the event of cleavage. However, there are two different models for Pol II termination.

The first model suggests that RNA cleavage at the polyadenylation site may trigger some conformational alterations in the Pol II complex and also the chromatin structure. This model is referred to as allosteric model.

The second model, the torpedo model, assumes that a specific exonuclease binds to the 5' end of the RNA. This exonuclease degrades the RNA faster than it can be synthesized, so that it catches up with RNA polymerase. It then interacts with proteins bound to the carboxy-terminus of the RNA polymerase, which further leads to the termination of transcription. The second model could serve as explanation, why the RNA polymerase II does not have well defined termination sites. There is experimental evidence for the torpedo model, which was carried out with the nuclear 5' to 3' exonuclease Xrn2 in mammals. Deletion of the Xrn2 gene causes the read-through to the next gene without termination. However, mutations within the AAUAAA motif does not necessarily lead to the dissociation of the RNA polymerase II and also prevents termination. This evidence thus favors the allosteric model. By either or both mechanisms, it is clear that transcriptional termination by RNA polymerase II is tightly coupled with the 3' end formation for most mRNAs in eukaryotic cells (Dye and Proudfoot, 1999; Kim et al., 2004; Luo, Johnson and Bentley, 2006).

Translation

The process of translation in mammals (Figure 9), again, must be distinguished from the bacterial translation. The mammalian ribosome is a large ribonucleoprotein particle that contains more RNA than protein and consists of a large (60S) and a small subunit (40S). The ribosomal proteins are also called r-proteins. Each ribosome subunit contains a major rRNA with catalytic activity and several small proteins. The large subunit can also consist of many smaller RNAs. The ribosomes in the cytoplasm are larger than those of bacteria, which means the

content of RNA and protein is greater in eukaryotes. Furthermore, the ribosome provides many active centers, which are each build out of proteins that are bound to a specific part of a ribosomal RNA. The rRNA is the decisive part of the ribosome because they do not only provide structural stability, furthermore, they act as ribozyme and carry out catalytic activity. Proteins alone are not able to function catalytically, but some of them are needed to support rRNA, so that they can work as ribozyme.

The ribosome provides three tRNA-binding sites. The tRNA functions as connecter between ribosome and amino acid. First of all, the amino acid is transported to the ribosome by an aminoacyl- tRNA. So, the incoming aminoacyl-tRNA binds to the so called A site via codon-anticodon interactions. The A site provides the codon, so that the correct amino acid can be added to the growing polypeptide chain. Then the ribosome slides to the next codon downstream. The ribosome always moves one triplet along the messenger RNA. The most recent added tRNA is now located at the P site. This site is bound by the peptidyl-tRNA, which is a tRNA that carries the nascent polypeptide chain and the A site can be bound by the following amino acid-bound tRNA. Both sites extend across ribosomal subunits. The liking step between the two amino acids of site A and site P is catalyzed by the large subunit of the ribosome. Therefore, the aminoacyl-ends of the tRNAs must be positioned correctly. Next, the polypeptide chain is transferred to the amino acid on the A site. This means that at this stage, the P site only contains the deacetylated tRNA, and the nascent polypeptide chain is located at the A site. However, a translocation step through sliding of the ribosome creates a free A site, so that a new amino acid- bound tRNA can bind. The polypeptide chain-bound tRNA is then located at the P site through the sliding of the ribosome. The tRNA that was located at the P site before and is now deacetlyated, is then positioned on the so called E site, which allows the tRNA to leave the ribosome. Thus, the flow of tRNA is into the A site, through the P site and leaves the ribosome at the E site.

Translation initiation includes the step, in which the polypeptide bond between the first two amino acids is formed. The ribosome must have assembled at the messenger RNA, which also includes the first aminoacyl-tRNA. During initiation, the small ribosomal subunit first binds to the mRNA, which is followed by the large subunit. First of all, the small subunit indentifies the 5’ end of the mRNA and, thus, detects to the initiation site. There, the large subunit also assembles. The methylated

cap is, therefore, the first feature that is recognized, which serves as a mark for the 5’ end. The small subunit needs many initiation factors, which detect the structure of the cap. Moreover, any secondary structure in the mRNA must be removed in advance, otherwise the availability could be impeded. The removal of the secondary structure is also carried out by several accessory factors. So, the cap is an important feature for the initiation codon, AUG, to be recognized. The small subunit of the ribosome first detects the cap, then it migrates until it reaches the initiation codon, where it stops, when the AUG is in the right content. An AUG codon, which is not in the right content, could be an AUG within the 5’ untranslated region. The ribosome, then, is able to pass past the noninitiating codon due to leaky scanning. If it does recognize the AUG as initiation codon, it starts translation but terminates before the proper AUG codon, after which it starts scanning again. Most of the eukaryotic initiation events start scanning at the cap, however, there is also an alternative, which includes the internal site, called internal ribosome entry site (IRES). In this case, every AUG in the untranslated 5’end is bypassed entirely because the ribosome is looking for IRES elements. There are three different types of interactions between the IRES elements and the 40S subunit. The most common type involves the AUG initiation codon at its upstream boundaries. The small subunit binds directly to this AUG, using the same factors, which are present at the initiation site at 5’ends. Another type is posed about 100 nucleotides upstream of the initiation codon, requiring a small subunit to migrate, again probably by a scanning mechanism. The last type of IRES can bind the subunit directly, without any initiation factors. The use of IRES is mostly common within different viruses because some viruses inhibit host translation by destroying cap structures, which are decisive for initiation without IRES elements. As soon as the 40S subunit is joined by the 60S subunit, the binding is stabilized and translation elongation can be started (Kaminski, Howell and Jackson, 1990; Pestova et al., 1998; Pestova, Hellen and Shatsky, 1996).

For initiation, however, a complex of initiation factors is required before elongation may occur. Currently, there are 12 factors known that are necessary for initiation. The different factors are working throughout the whole process, including the formation of the initiation complex at the 5’end of the mRNA, forming and binding the Met-tRNA complex, enabling the ribosome to migrate the mRNA, detecting tethering of initiator tRNA to AUG at the start site and mediating joining of the 60S subunit. First of all, the 43S preinitiation complex must be formed, which includes

eIF2, Met-tRNAi, eIF3, eIF1 and eIF1A. This complex binds to the 40S subunit. eIF4A, eIF4B, eIF4E and eIF4G bind to the 5' end and form the cap-binding complex. This complex then associates with the 3' end of the mRNA via eIF4G, which is bound to the poly(A)tail of the mRNA through interaction with the poly(A)-binding protein. The 43S complex assembles with the initiation factors at the 5' end and detects the AUG initiation codon. This is then called 48S initiation complex. The eIF2 is a monomeric GTP-binding protein consisting of an α, β and γ subunit. It is activated by binding to GTP and is inactivated by hydrolysis to GDP. eIF2-GTP tethers to Met-tRNA, which is referred to as ternary complex. This complex, then, places Met-tRNA onto the 40S subunit. eIF3 is required to maintain 40S subunits in a dissociated state. eIF1 and eIF1A seem to promote the eIF3 activity. eIF4F is a protein complex that consists of three initiation factors. It is not known yet, whether the complex assembles before binding to mRNA or directly on the mRNA. The complex, however, includes the cap-binding subunit eIF4E, the helicase eIF4A and the scarfolding subunit eIF4G. After eIF4E tethers the cap, eIF4A unwinds secondary structures, which are present within the mRNA. For the unwinding process, energy is given through ATP hydrolysis and is supported by eIF4B. eIF4G is mainly involved in binding other components of the initiation complex, which means the interaction with PABP, which further leads to the circulation of the mRNA with both the 5' end and the 3' end together in one complex. This special interaction leads, then, to the recruitment of the large subunit. All of the remaining factors likely are released when the complete 80S (40S + 60S; S… Svedberg sedimentation coefficient) ribosome is formed (Asano et al., 2000; Kahvejian et al., 2005; Pestova and Kolupaeva, 2002; Pestova et al., 2000; Tarun and Sachs, 1996).

Translation elongation represents all reactions, started with the formation of the first peptide bond to the addition of the last amino acid. One by one amino acid is added by each reaction. The bond formation step is the fastest process during translation. The ribosome migrates along the mRNA, which actually means that the mRNA moves through the ribosome and not the other way round. The movement always includes one triplet, which codes for one amino acid. The first triplet always encodes for the same amino acid, which is methionine, but the corresponding tRNA is already added during formation of the initiation complex. There are two different types of tRNAs, which recognize the AUG codon. On is used for initiation, the other one for detecting AUG codons during elongation, however, none of them becomes

formylated, like in bacteria. As soon as the first aminoacyl-tRNA enters the A site, the P site is already occupied by a peptidyl-tRNA, which encodes methionine, as far as it is the first tRNA that enters the P site. Furthermore, every aminoacyl-tRNA is able to enter the A site, except the initiator tRNA. The entry of the tRNAs is facilitated by an elongation factor. As soon as the aminoacyl-tRNA is placed at the A site, the elongation factor (eEF1α) leaves the ribosome and mediates the next tRNA to the ribosome. eEF1α is a GTP-binding protein that is active when tethered to GTP (guanine triphhophate) and inactive when tethered to GDP (guanine diphosphate). As far as the aminoacyl-tRNA and peptidyl-tRNA are positioned correctly, the peptide bond formation can be carried out. The interaction between tRNA and eEF1α also plays an important role in quality control. tRNAs are brought to the A site without regard for whether their anticodons will fit the codon. The hydrolysis of eEF1α -GTP is relatively slow, so it takes more time than the dissociation of incorrect aminoacyl-tRNA from the A site. So most of the non-complementary tRNAs are removed at this point. If both tRNAs are on the right place, the peptide bond is formed by an enzyme called peptidyl transferase, which is included in the large ribosomal subunit. Within this process, the nascent peptide chain, which is tethered to the tRNA in the P site, is transferred to the A site, where the aminoacyl-tRNA is located. This reaction is activated as soon as the eEF1α leaves the ribosome. (Krebs, Goldstein and Kilpatrick, 2013) Once the addition of the amino acid to the growing polypeptide chain is finished, a process called translocation is performed. The mRNA is pulled through the ribosome for three nucleotides, so that an empty A site is provided to be able to add a new amino acid. The hybrid state model is a model that describes the process of translocation within two steps in bacteria, however, it is thought to be the same in eukaryotes. First of all, the large subunit shifts relative to the small subunit, which means that the tRNA on the large subunit change from the P site to the E site, but the contacts on the small subunit does not change. Afterwards, to restore the original conformation, the small subunit of the ribosome migrates and shifts after the large subunit. It is interpreted that the aminoacyl ends of the tRNAs, which are located in the large subunit, first get assigned to the new site, while the anticodons remain bound to the small subunit. Then the small subunit also moves, so that the anticodon-codon pairing region find itself in the right position. (Moazed and Noller, 1989).

The process of translocation, however, needs energy, which is provided as GTP and another elongation factor, eEF2. Both elongation factors, eEF2 and eEF1α are GTP-binding proteins, which share the same binding-site on the ribosome. This means that if one of the proteins is bound to the ribosome, the binding of the other one is not possible. eEF2 binds to the ribosome to facilitate the process of translocation and at this stage of translation, eEF1α binding is not required because there is no free A site provided yet. The hydrolysis of GTP releases eEF2 after the process of translocation, which means a free A site is available (Nissen et al., 1995; Stark et al., 2000).

There are three different codons, which encode a termination signal. These three triplets, also known as nonsense codons or stop codons, end translation. The UAG codon is called the amber codon, the UAA codon is referred to as ochre codon and UGA is the opal codon. These codons were identified by genetic tests that make a distinction between two types of point mutations. A missense mutation means that the point mutation changes the codon, so that it encodes a different amino acid. A point mutation that alters a codon in a way, so that it encodes one of the three termination codons, is called nonsense mutation. It leads to the premature termination of the translational process, which causes truncated proteins, which can abolish the proteins function depending on where the mutation is located. Normally, stop codons are located at the end of an open reading frame (ORF). If a stop codon is detected, the produced nascent polypeptide chain is released into the cytoplasm and the ribosome disassembles from the mRNA. There are again two stages that occur during translation termination. The termination reaction itself means the release of the polypeptide chain into the cytoplasm, whereas the posttermination reaction involves the release of the tRNA and the dissociation of the ribosome. In contrast, a termination codon is not detected by tRNAs. The stop codons are detected by class 1 release factors (eRF). The class 1 eRFs are assisted by class 2 eRFs, which are not codon specific. The class 2 release factors are proteins that bind GTP. Class 1 release factor, eRF1, detects the stop codon and triggers the ribosome to hydrolyze the peptidyl tRNA. Subsequently, the polypeptide chain is cleaved from the peptidyl tRNA, which is similar to the usual peptidyl transfer reaction, except H_2O and not an aminoacyl-tRNA is the acceptor. Then, eRF3 releases eRF1 from the ribosome. eRF3-GDP, therefore, binds to the ribosome before the termination reaction takes place and GDP is replaced by GTP. This

activates eRF3 and it is able to bind to the GTPase center, which leads to the release of eRF1. However, in yeast, additionally to eRF1, a second release factor, eRF2, is essential. In other eukaryotes, eRF1 is able to recognize all three termination codons as single protein (Freistroffer et al., 2000; Ito et al., 1996; Klaholz, Myasnikov and van Heel, 2004; Mikuni et al., 1994; Milman et al., 1969; Scolnick et al., 1968; Song et al., 2000).

The empty tRNAs are also released into the cytoplasm and can be reused. This means that they get loaded again with their specific amino acid and work together with another ribosome to carry out the process of translation.

Each step is supported by different accessory factors. Moreover, energy is needed at many stages, which is provided by the hydrolysis of guanosine triphosphate (GTP). Additionally, each step is controlled for accuracy by specific mechanisms. The error rate of translation is very low, however, one error for every 10^4 to 10^5 amino acids incorporated is generally thought to occur. It is not completely clear, how such a low error rate can be achieved and how one particular member can be distinguished from the entire set, which almost share the same general features. It is thought that any substrate can hit the active center by a random-hit process, but then the wrong substrates are rejected and only the appropriate substrate are accepted or are able to bind properly. An error within the mRNA, which are very rare, would be even worse because one mRNA is translated into many proteins. An error during translation must not always have tremendous effects on the phenotype of the cell.

The ribosome can create two different types of errors. Firstly, it may cause a frameshift, which means the skipping of a base when it reads the mRNA, or it can even read a base twice. Secondly, the ribosome could allow incorrect base pairing between the aminoacyl-tRNA and the codon provided by the mRNA. This type of failure is probably the most common one.

Errors can also occur during loading of tRNA, which means that the aminoacyl-tRNA synthetase places the wrong amino acid on its tRNA. However, the aminoacyl-tRNA synthetases have again a specific mechanism that recognizes errors and corrects them before the mischarged tRNA is released (Giege, Sissler and Florentz, 1998; Ibba and Söll, 2000; Perona and Hou, 2007;. Berg, Tymoczko and Stryer, 2002)

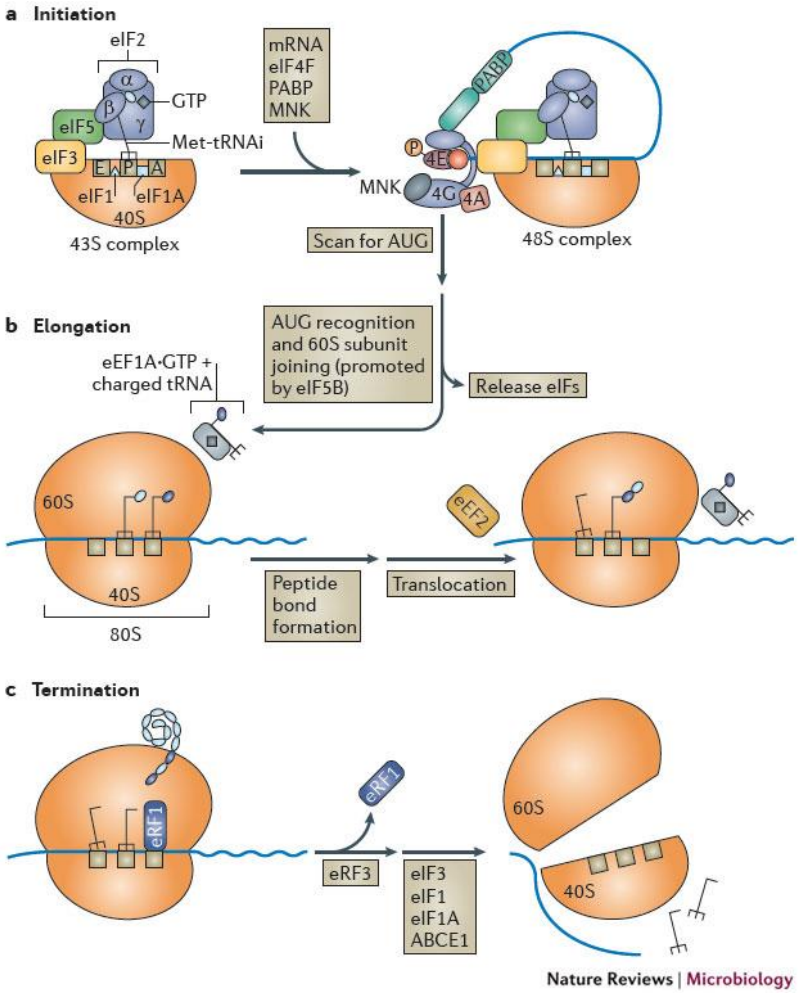

Figure 9: Process Of Translation

Source: cf: Walsh, D. and Mohr, I., 2011. Overview of mRNA translation in eukaryotes. [image online] Available at: Nature Reviews Microbiology <http://www.nature.com/nrmicro/journal/v9/n12/fig_tab/nrmicro2655_F1.html> [Accessed 13 January 2015].

INTRODUCTION

As it can be seen now, eukaryotic gene expression, involving multiple steps and series of highly regulated and interconnected events, has a high level of complexity in which the mRNA plays a crucial role. Due to the complex pathway, the probability of errors increases. To assure a viable system, the quality and fidelity of mRNA metabolism must be given. In order to maintain the correct function of mRNAs, there are quality controls that recognize and degrade defective and improperly-processed mRNAs. The mechanism that performs these quality control steps is called nonsense-mediated mRNA decay (NMD). Premature termination codons (PTCs) are detected by this mechanism and labeled for rapid degradation. So, synthesis of C-terminal truncated polypeptides is avoided, that would further lead to deleterious dominant-negative or gain-of-function effects within the cell.

However, not only mRNAs with PTCs are targeted by NMD. It could be shown that the process of alternative splicing produces a high range of nonsense mRNAs in order to regulate gene expression. The results of bioinformatics analysis revealed that the production of premature stop codons by alternative splice events is a common method for mammalian cells to maintain homeostasis of gene expression. So the production of PTCs is not always an accidently occurring error, but also allows to keep balance between producing nonsense mRNA and avoiding consequences due to aberrant mRNA translation (Mendell et al., 2004).

The significance of this specific quality control pathway is given by the fact that about one third of inherited genetic disorders are caused by PTCs, most of which are targeted for NMD. Disorders, which occur due to a failure in degradation of the nonsense mRNA, are, among other diseases, ß-thalassemia and the Marfan syndrome. Both diseases are caused by nonsense-containing mRNAs, which should be target of NMD. However, sometimes the mechanism is not triggered or does not work properly.

Formation of Premature Termination Codons

Before the mechanism of nonsense-mediated mRNA degradation can be dedicated, the reason why this mechanism is triggered must be taken into consideration. Mostly, mRNAs containing premature stop codons (PTCs) are target of the degradation mechanism. If the quality control would not be triggered, truncated proteins are the consequence, which would further lead to genetically related disorders.

mRNAs, which contain PTCs, can result as a consequence of alterations in the DNA, errors within the transcription process or afterwards during mRNA processing (Mendell et al., 2004). However, it is not always the error of mRNA metabolism. PTCs can also be created from somatic rearrangements in the DNA or intentionally through alternative splicing. Wild type transcripts can also provide UGA codons, which encode for selenocysteine, upstream open reading frames (uORFs), which lead to be recognized as PTCs. The utilization of alternative AUG initiation sites can also be recognized as a nonsense codon and trigger NMD. Furthermore, an intron within the 3'UTR or also AU-rich sequences in the 3'UTR are revealed to activate rapid degradation of mRNAs (Li et al., 1999; Lewis, Green and Brenner, 2003; Grundner-Culemann et al., 2001; Wang et al., 2002; Bicknell et al., 2012; Shaw and Kamen, 1986).

These natural NMD targets have been identified to pose 5-10% of the transcriptome of human cells, which are targeted by NMD degradation (Silvia and Romão, 2009).

Somatic rearrangements in DNA are generally used to create a certain diversity of proteins, for example, during production of immunoglobulin and T-cell receptors, nucleotides that are not template-directed are inserted into the strand using terminal transferase. Another way to create diversity is transferring single nucleotides from one template to another. If the rearrangement works well, the diversity and, therefore, increased selectivity of receptors is given. However, two-third of these recombination events lead to modifications within the reading frame. These alterations can generate PTCs and their translation into proteins would result, in this case, in non-functional receptors (Li et al., 1999).

The mechanism of alternative splicing is generally used to create isoforms of different proteins, which again contribute to diversity and specificity. However, it could be demonstrated that up to 35% of all alternative splice events lead to degradation through NMD (Lewis et al., 2003). Wollerton et al. (2004) suggested that not only the productions of isoforms of proteins from an individual gene are the aim of alternative splicing. Bioinformatic analysis showed that alternative splicing is used to quantitatively supervise gene expression through introduction of premature termination codons (Wollerton et al., 2004).

Moreover, mRNAs that encode selenoproteins contain UGA codons, which would encode a functional protein. However, the UGA codon is typically known as a termination indicator and, therefore, termination of translation may occur. Whether translation is stopped or the tRNA with that specific anticodon is added to form the functional protein is dependent on the overexpression of eRF1 or eRF3. eRF1 is a translation termination release factor, which recognizes stop codons and eRF3 is an eRF1- and ribosome-dependent GTPase. Grundner-Culemann et al. (2001) revealed that overexpression of eRF1 leads to a decrease in termination at the UGA codon, whereas an overexpression of eRF3 raises the rate of termination events (Grundner-Culemann et al., 2001). Consistent with the results of Hoshino et al. (1999), eRF3 seems to be an important factor that is part of NMD machinery activation (Hoshino et al., 1999). Additionally, the concentration of Sec-tRNA is decisive whether NMD is activated or not (Weiss and Sunde, 1998).

Nonsense codons, which trigger NMD, are often defined from AUG codons, which are generated during splicing. This means that through the event of splicing of an intron, downstream of the exon, results in an AUG codon. So, the process of splicing has created an AUG codon within the frame and this further is defined as a nonsense codon in the exon. It is hypothesized that the recognition of this nonsense mutation occurs due to the presence of the internal ribosome entry site (IRES) upstream of the alternative initiator AUG in the exon (Wang et al., 2002).

Triggering of NMD can also result due to the presence of introns in 3'-untranslated regions (UTRs). Recent studies could demonstrate, that introns within the 3'UTR are often target of NMD as a consequence of gene expression regulation. 3'UTRs containing introns are regarded as decisive cis-regulatory elements that influence regulation of gene expression at multiple stages (Bicknell et al., 2012). Consistently with the results of a microarray, Mendell et al. (2004) could demonstrate

that the presence of an intron more than 50 nucleotides downstream of a termination codon is able to trigger NMD and, therefore, lead to the degradation of the mRNA (Mendell et al., 2004).

Furthermore, the presence of AU-rich sequences within the 3'UTR is also assumed to trigger rapid mRNA degradation. Experimental trials demonstrated that the insertion of a 51 nucleotide AT sequence into the 3' untranslated region of a rabbit ß-globulin gene resulted in high instability of the transcribed mRNA *in vivo*. The ß-globulin gene was used because it is generally considered stable. Through the insertion of the AT sequence into the gene, the transcript of the gene received an AU – rich sequence in the 3'UTR. These adenylate/uridylate- rich elements (AREs) within the 3'UTR seem to be a common RNA stability determinant in mammalian cells. However, it is not clear yet, which exact sequence is decisive for the triggering of NMD. The AU sequence bears similarities to downstream polyadenylation sites and is assumed to result in a PTC or an aberrant 3'terminus generation. So, the degradation of these NMD targets are coupled with translation and, therefore, occur in the cytoplasm (Shaw and Kamen, 1986).

The mechanism of nonsense-mediated mRNA, does not only play a crucial role in degrading nonsense mutations or frameshift mutations, which are detected as premature termination codons, but also transcripts that contain upstream open reading frames in the 5' untranslated region. Moreover, it could also be revealed that mRNA transcribed from transposons or endogenous retroviruses are target of NMD. The mariner 2 transposon, as well as the human endogenous retrovirus H, has been identified to contain a premature stop codon, which has been evolved during evolution and is now target of nonsense-mediated degradation in mammalian cells, when expressed (Mendell et al., 2004).

These results again lead to the assumption that NMD is not only evolutionary maintained due to the degradation of nonsense or frameshift mutations, but is also important in the means of regulating gene expression.

PTCs are not always consequences of errors within mRNA metabolism, whereas a bunch of PTC are created because of gene expression regulation. However, the process of translation seems to be decisive in order to detect PTCs, as well as a splicing-dependent signal is also necessary for PTC recognition in mammalian cells.

Splicing and the role of the EJC(s)

First of all, premature termination codons must be positioned more than 50-54 nucleotides upstream of an exon junction complex (EJC) in order to trigger nonsense-mediated decay (Kashima et al., 2006).

The connection with the recognition of the EJC and the activation of NMD suggests that the splicing process in mammalian NMD is a decisive event. However, the NMD mechanism is not triggered by nonsense transcripts that are created from naturally intronless genes, like histone H4, heat shock protein 70 or melanocortin 4-receptor genes (Maquat and Li, 2001), which reveals the importance of the splicing process within mammals.

It is suggested that a mark is assembled during splicing adjacent to the exon-exon junctions. This mark would communicate the location of the translation termination event to the NMD machinery. It is known that there is a protein complex, called exon junction complex (EJC), which assembles 20-24 nucleotides upstream of each exon-exon junction (Le Hir et al., 2000). The EJC could be identified to serve as a connection platform for different factors, which are part of the NMD machinery. Some of the NMD effectors are directly involved in the following degradation mechanism. EJCs are essential for the definition of premature termination codons. After recognition of a PTC, NMD is activated through binding of many different factors.

One of the most important factors in humans seems to be the eukaryotic initiation factor 4AIII (eIF4A3). As mRNA splicing is needed to produce functional mRNAs, the process of splicing can have profound influence on following events. Created EJCs are, therefore, the most responsible part for determining downstream processes. After the process of pre-mRNA splicing in the nucleus, eIF4A3 was shown to be part of the RNA-binding platform and recruits other EJC components to the spliced mRNA. The EJC is formed of a stable core that is tethered on the 5’ exon within the second step of splicing. This core remains bound to the spliced mRNA until it is transported into the cytoplasm. This specific core further serves as binding platform for many other factors, which are part of the following mRNA metabolism. (Tange et al., 2004). eIF4A3, as well as the human Metastatic Lymph Node 51 (MLN51), has been revealed to directly interact with spliced mRNA at the EJC deposition site. eIF4A3 belongs to the DEAD-box family and is characterized as an RNA-dependent ATPase. The factor also depicts many similarities to general

translation initiation factors, like eIF4AI and eIF4AII (Li et al., 1999). DEAD-box proteins,normally, function in a sequence independent manner and due to the fact that eIF4A3 contains a nucleic acid binding site (Kim et al., 1998) of the same size like the EJC footprint on spliced mRNA (Le Hir et al., 2000), it is assumed that eIF4A3 is the first factor, which binds to the EJC and that all other factors follow. The other factors, which form the core of the EJC complex, are MLN51, Y14 (RNA binding protein) and MAGOH (mago nashi homolog) (Tange, Nott and Moore, 2004). However, Shibuya et al. (2006) could reveal that there are specific motifs, which are required for EJC formation and NMD. Mutations within the motifs Ia, VI and motif H, which is specific for eIF4A3, caused the inhibition of EJC formation *in vitro* and no interaction was visible between eIF4A3 and MAGOH *in vivo* (Shibuya et al., 2006).

They stated that mutations in motif Ia and VI result in a lower affinity of eIF4A3 for RNA. The low affinity of eIF4A3 leads to an unstable EJC and, therefore, NMD cannot be triggered. Thus, the corollary seems to be that the interaction of eIF4A3 and MAGOH needs to be stable and this can only be reached when the EJC can be assembled properly (Shibuya et al., 2006).

The motif H of eIF4A3 is assumed an important binding site for one of the other core factors, but it is not clear yet. MLN51 is out of the question because this factor is also tethered to eIF4A3, containing mutations within the H motif. It is suggested that either MAGOH or Y14 may bind to the H motif. They furthermore hypothesize that mutations within any of these motifs could lead to the disruption of a key contact that is necessary for forming a stable EJC complex. So, an unstable EJC would further result in not triggering NMD (Shibuya et al., 2006).

To explore the binding site of MLN51, Shibuya et al. (2006) could reveal that the motif C of eIF4A3 and the N-terminal site of MLN51 interact with each other *in vitro*, but further experiments showed that mutations within the motif C did not inhibit this interaction. However, they conjecture that the formation of the EJC complex is stabilized through the binding of the MAGOH and Y14 (Shibuya et al., 2006) because interactions between eIF4A3 and MLN51 was greatly enhanced when MAGOH and Y14 were present (Ballut et al., 2005; Palacois et al., 2004).

So now that the core exon junction complex is defined (Figure 10), it remains open how a premature termination codon can be identified.

Mechanism of PTC definition

In mammalian cells, a splicing-dependent signal is considered a crucial part of the PTC recognition. The model of PTC definition in mammals suggests the involvement of both, splicing and EJC components. However, EJCs seem to be the labels that distinguish normal from premature termination codons. The formation of EJCs occurs next to each exon-exon junction and serve as Upf2 and Upf3 (up-frameshift protein 2/3) binding sites, additionally to the core exon junction complex proteins (Kadlec, Izaurralde and Cusack, 2004; Mendell, Rhys and Dietz, 2002). It could also been showed that the quality mechanism is activated in the first round of translation during which the nuclear cap-binding complex is still linked with the mRNA. (Ishigaki et al., 2001). If the PTC is located more than 50-54 nucleotides before the last exon-exon junction, at least one EJC is maintained. If the terminating ribosome detects a PTC upstream of one or more EJCs, it interacts with these EJCs or with the single EJC downstream of the PTC, which correspond to the interaction of the Upf1 and SMG-1 (Serine/threonine-protein kinase) (Hoshino et al., 1999) with the proteins of the EJC. Upf1 and SMG-1, therefore, associate with translation termination factors (eRF3 and eRF1), which assemble on the ribosome, and finally tether to Upf2 and Upf3, which are part of the EJC (Figure 10) (Le Hir et al., 2001). The assembly of the proteins leads to the activation of NMD and further causes the triggering of the degradation mechanism of the nonsense mRNA (Ishigaki et al., 2001).

Normally, if an mRNA does not contain a PTC, which leads to the production of a full length protein, or if the PTC is located downstream or adjacent to the last EJC, the termination ribosome removes all EJCs present on the mRNA and the degradation pathway is not activated. This means that the distance of the PTCs to the 3’end of the mRNA is decisive and of prime importance, whether the degradation machinery is activated or not (Ishigaki et al., 2001).

mRNA decay rates have depicted that upregulated transcripts that were substrates of the degradation machinery, have always been depleted tethering Upf1. Moreover, the trans-effector Upf2 could also be determined to be part of the NMD machinery (Mendell et al., 2004). So, the Upf proteins are of high importance for the triggering of nonsense-mediated mRNA decay. The next part focuses on each Upf protein and how they contribute to the NMD.

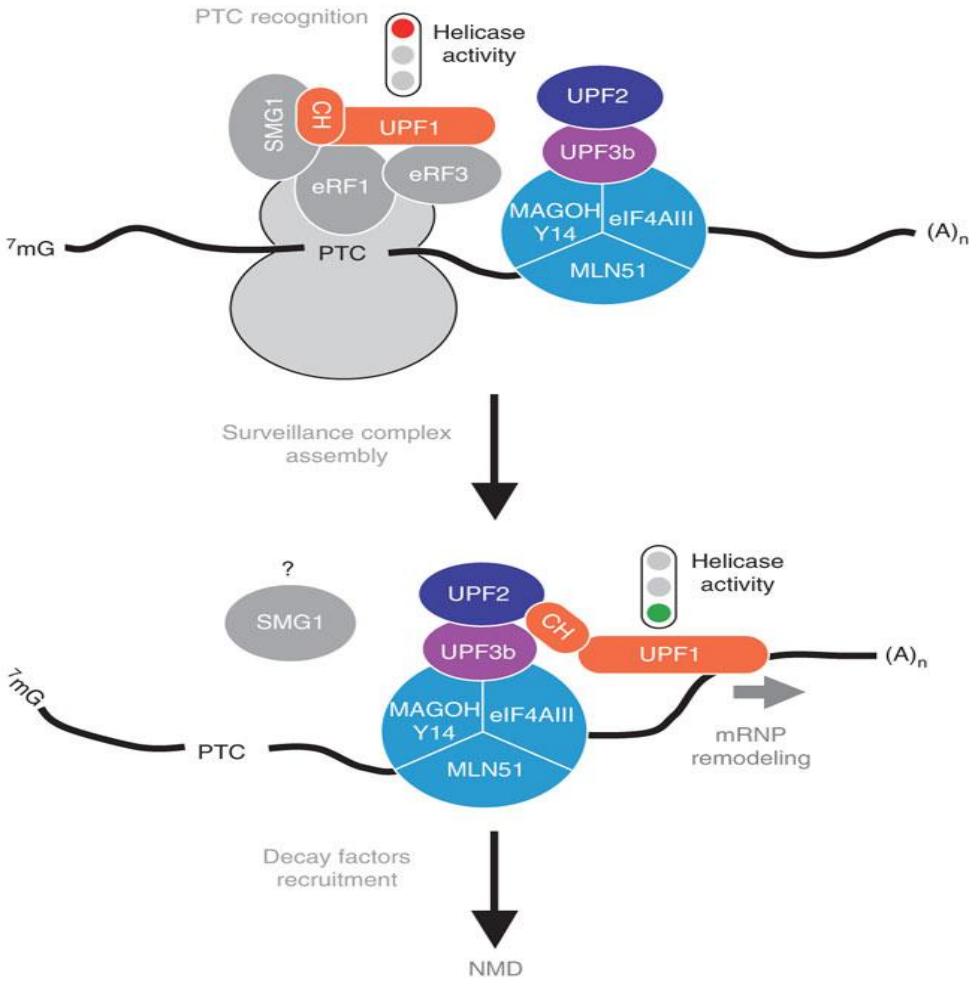

Figure 10: Model showing PTC recognition and assembly of protein complex for NMD.

On the picture, the assembly of the core exon-junction-complex (EJC) can be seen. The core EJC consists of four proteins: Y14, MAGOH, eIF4AIII and MLN51. The core EJC binds Upf3b and the adapter Upf2. If a PTC is detected by the ribosome, Upf1 and SMG1 associate with the translation termination factors (eRF3 and eRF1) on the ribosome and NMD is triggered. ,

Source: cf: Chamieh, H., Ballut, L., Bonneau, F. and Le Hir, H., 2008. Model illustrating the modulation of UPF1 activity during the human NMD pathway. [image online] Available at: Nature Structural and Molecular Biology <http://www.nature.com/nsmb/journal/v15/n1/fig_tab/nsmb1330_F8.html> [Accessed 10 January 2015].

The NMD effectors: UPF proteins

The human Upf1 protein is necessary for the NMD pathway to be activated. The protein is an ATP-dependent 5' to 3' helicase and an RNA-dependent ATPase. For the NMD to be triggered, the Upf1 is essential. Therefore, it must be phosphorylated by SMG-1. Moreover, Upf1 is also needed to recruit the decapping assembly to P-bodies for translation repression and degradation of NMD substrates (Sheth and Parker, 2006).

Upf2 is another important protein that is required in some pathways, but it could be shown that there are two different pathways of NMD. One is an Upf2 dependent pathway and the other pathway has been depicted Upf2 independent. Within the Upf2 protein, many nuclear localization sequences (NLS) were found, as well as the Upf2 provides three MIF4G (middle domain of eukaryotic initiation factor 4G) domains, which are also necessary for mRNA processing and maturating. Moreover, the Upf2 consists of two domains, which serve as binding sites for the Upf1 and Upf3 proteins. So, it can be considered an adapter, which provides interaction between these two factors and, therefore, provides the possibility of interaction of Upf1 with the exon-junction complex (Kadlec, Izaurralde and Cusack, 2004).

In humans, two different Upf3 proteins, labeled Upf3a and Upf3b, could be detected, which reveal RNA recognition motifs (Lykke-Andersen, Shu and Steitz, 2000). However, none of them showed direct binding to the RNA, (Kadlec, Izaurralde and Cusack, 2004) so this indicates that the Upf3 proteins must contain other binding sites, which tether any of the exon-junction proteins.

The two homologs Upf3a and Upf3b work differently within triggering NMD. Upf3b revealed strong NMD activation when bound downstream to a normal stop codon. In comparison, the activity of Upf3a was moderate (Lykke-Andersen, Shu and Steitz, 2000).

The importance of the Upf proteins was tested in many trials. NMD activation could be shown to be impaired by the expression of a dominant-negative hUpf1, which contained a point mutation (Sun et al., 1998). Moreover, the inhibition of hUpf1, hUpf2 and hUpf3b by RNAi, hindered the mechanism of NMD (Mendell, Rhys and Dietz, 2002; Kim et al., 2005).

In mammalian cells, co-immunoprecipitation assays depicted that the Upf proteins do not require mRNA to be able to interact with each other. Although it was considered that these three factors are required for triggering NMD, the insertion of a mutated Upf3b, which had no domain for Upf2 binding, also triggered NMD. This results in the support of the suggestion that there is an Upf2 dependent and Upf2 independent pathway in mammalian cells. Nevertheless, Upf1 was definitely required. It can be concluded that the NMD is a transcript- or cell type- specific process (Gehring et al., 2005). However, the triggering of NMD by Upf1 seems to be conserved within eukaryotic cells, which is linked with many phosphorylation steps. In the following parts, a detail explanation of the activation process is provided.

Upf1 and NMD activation

The phosphoprotein Upf1 (Yamashita et al., 2001) is required for the activation of NMD, as mentioned before. Therefore, the Upf1 must be phosphorylated and dephosphorylated, which is regulated by four proteins (Rehwinkel et al., 2005). The protein that is responsible for the phosphorylation of Upf1 is SMG-1, the human ortholog is called hSMG-1 and shows similar activity. SMG-1 is a phosphatidylinositol 3-kinase-related protein kinase, which phosphorylates specific serine residues (Yamashita et al., 2001). The other three proteins, which are involved in the phosphorylation/dephosphorylation cycle, do not reveal phosphatase activity. However, these proteins, SMG-5, SMG-6 and SMG-7, contribute to the dephosphorylation of Upf1 and it is suggested that these proteins help to guide the PP2A (protein phosphatase 2A) to Upf1 and carry out the phosphorylation (Rehwinkel et al., 2005).

All three SMG proteins (5, 6 and 7) contain a 14-3-3 domain, which is an N-terminal motif that provides a binding site for phosphoserine residues (Fukuhara et al., 2005). These three proteins are considered as adapters between the phosphorylated Upf1 and the PP2A, causing the dephosphorylation of Upf1 (Figure 11) (Ohnishi et al., 2003; Chiu et al., 2003; Kashima et al., 2006).

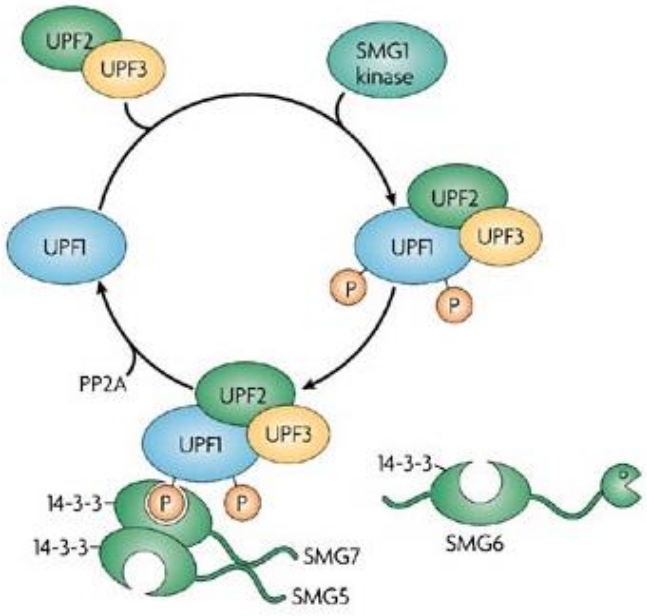

Figure 11: Upf1 phosphorylation – dephosphorylation cycle

The phosphoprotein Upf1 is phosphorylated by SMG1. Dephosphorylation of Upf1 is catalyzed by SMG5, SMG6 and SMG7 by recruiting PP2A.

Source:cf: Eulalio, A., Behm-Ansmant, I. and Izaurralde E., 2007. Molecular link between NMD and P bodies. [image online] Available at: Nature Reviews Molecular Cell Biology <http://www.nature.com/nrm/journal/v8/n1/fig_tab/nrm2080_F2.html> [Accessed 10 January 2015].

However it was shown, that the SMG-7 is able to activate mRNA degradation, even if it lacks the phosphoserine binding site and independently of its position to the stop codon. So mRNA degradation is assumed to be independent of the interaction of Upf1 and the PTC recognition. So, the SMG-7 must

provide two different domains, one that allows the binding to the surveillance complex and another that is able to trigger NMD. The latter function was shown to be related to the activity of Dcp2 (mRNA-decapping enzyme 2) and Xrn1 (5'-3' exoribonuclease 1), which are enzymes involved in decapping and exonucleolytic 5’ to 3’ mRNA degradation (Unterholzner and Izaurralde, 2004).

SMG-6 was shown to have a degrading function *in vitro* in single stranded RNA. This protein provides an essential nuclease activity that may contribute to a rapid degradation of PTC-containing mRNAs in mammalian cells. However, it remains unclear whether the SMG-6 protein is part of an alternative decay pathway or works together with SMG-5 and SMG-7 in the ordinary pathway (Huntzinger et al., 2008).

Due to analysis of co-immunoprecipitation of human cell extracts, it was shown that Upf1 binds translation termination factors eRF1 and eRF3, as well as SMG-1. This assembly is called SURF complex (SMG-1/UPF1/eRFs) and it is formed independently from binding to the EJC (Figure 12). It is put up on a terminating ribosome before interaction with Upf2 in the EJC occurs. The interaction of the SURF complex and the Upf2 and, therefore, the EJC, is considered to form another complex. This complex is termed DECID (decay-inducing complex) and its assembly is performed according to the interaction between Upf1 and Upf2, which each occurs within another complex. Furthermore, it is suggested that the formation of the DECID complex activates SMG-1, which lead to the phosphorylation of Upf1. In human cells, the decrease of translation initiation of NMD targets could be observed, if hUpf1 was phosphorylated. Although the process of phosphorylation of Upf1 is definitely required for NMD, the DECID formation, as well as Upf2 and Upf3b interaction are questioned to be necessary for the Upf1 phosphorylation (Isken et al., 2008; Kashima et al., 2006). Moreover, the necessity of Upf2 for NMD is not completely clear now. As mentioned before, Gehring et al. (2003) suggest that there also exists an Upf2 independent pathway.

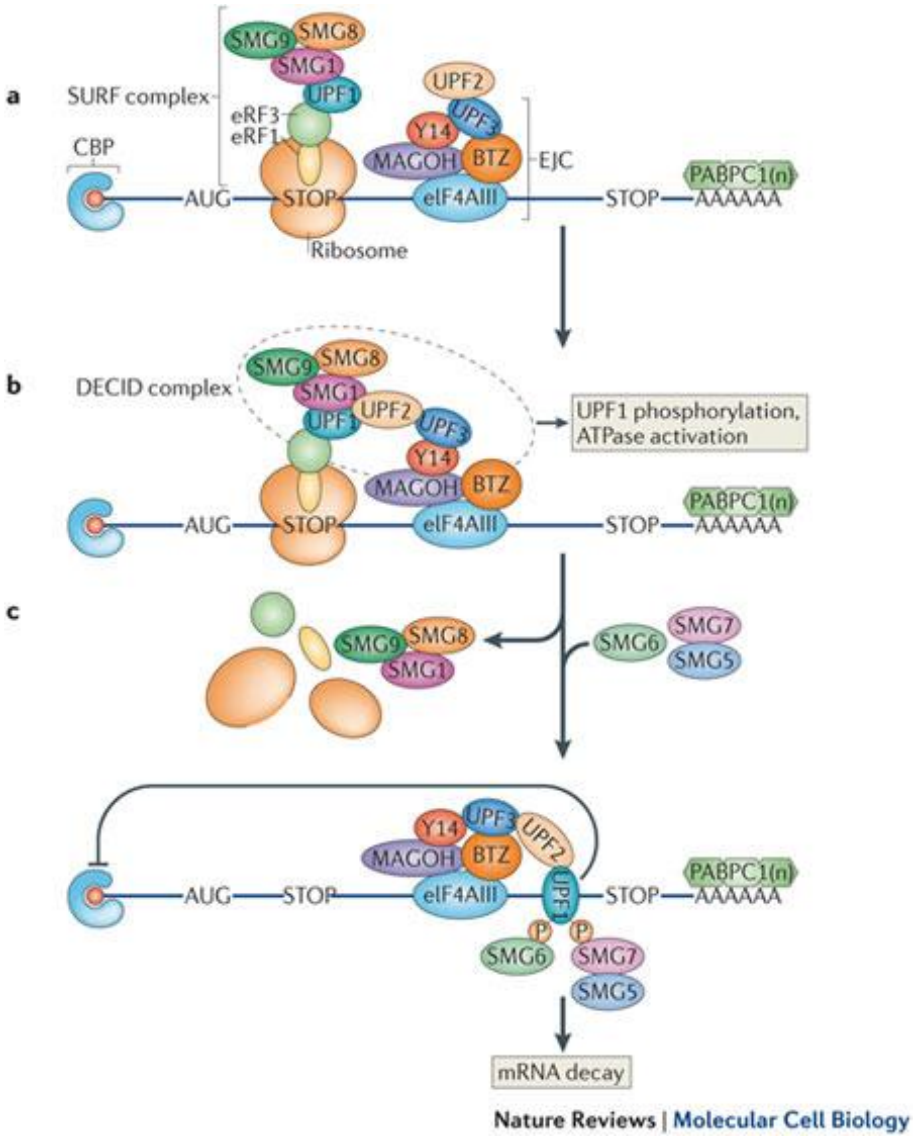

Figure 12: Model illustrating the assembly of the SURF complex and the DECID complex

On the picture, the assembly of the SURF complex and the DECID complex can be seen.

The SURF complex constitutes SMG-1, Upf1 and eRFs, which assemble on the ribosome. The EJC is dynamic but always consists of the core proteins Y14, MAGOH, eIF4AIII and BTZ. Upf3 and Upf2 also bind to the EJC.

The DECID complex is formed through interaction of Upf1 and Upf2.

The phosphorylation of Upf1 allows interaction with SMG6, SMG5 and SMG7, which further leads to mRNA deadenylation and decapping

Source: cf: Kervestin, S. and Jacobson, A., 2012. Activation of metazoan NMD by EJC-dependent interactions.[image online] Available at: Nature Reviews Molecular Cell Biology <http://www.nature.com/nrm/journal/v13/n11/fig_tab/nrm3454_F2.html> [Accessed 05 January 2015].

THE MECHANISM

Two branches of NMD pathway

The mammalian NMD pathway does not follow a linear path, but it can split in different pathways, which are defined by EJC components. Different cofactors define the direction of how a nonsense mRNA is degraded. Gehring et al. (2005) found that two different pathways can be selected, of which one is Upf2 dependent and one is Upf2 independent. Both pathways have in common that these are Upf1 dependent, which could be shown based on the depletion of Upf1 using RNAi. Moreover, a depletion of all other NMD factors showed that some of the components are definitely responsible for NMD triggering.

To examine, which of the EJC components is required to define the distinct branches of the NMD pathways, one after another cofactor or component of EJC was downregulated using small interfering RNA. As mentioned before, NMD is dependent on the presence of Upf1. Depletion of Upf1 resulted in recapitulation of NMD activity and, therefore, NMD can be characterized by its Upf1 dependence.

The linear pathway is expected that the EJC is bound by Upf3. Upf2, then, acts as an adapter between Upf3 and Upf1. Therefore, depletion of Upf2 is expected to inactivate the mechanism of NMD. The degradation of nonsense mRNA tethering RNPS1 requires Upf2, in contrast the complex of Y14, MAGOH, Upf3b and eIF4A3 are not effected by depletion of Upf2. NMD was activated under normal conditions when Upf2 is inhibited and this leads to the assumption that NMD can also be triggered without the presence of Upf2 (Figure 13).

Within these two pathways, it is revealed that Barentsz (BTZ) is also one of the proteins decisive for NMD. If this protein was depleted in mammalian cells, the Y14-MAGOH- induced NMD was impaired by 2.1 – 3.1 fold. However, the activity of RNPS1-induced NMD did not show any alterations considering the BTZ depletion. So, BTZ-deficient cells are not influenced and can activate NMD normally. This leads to the assumption that the BTZ is not involved in the RNPS1-mediated pathway, but is exquisitely required in the Y14-MAGOH-induced pathway, which further confirm the theory of non-linear, branched pathways that require different cofactors for eliciting nonsense mRNA degradation (Gehring et al. 2005).

RNPS1 (RNA-binding protein with serine-rich domain 1), the essential protein for activating the Upf2-dependent pathway, provides four domains. To identify, which of the four domains are important for mRNA degradation, the protein was changed through different mutations within these domains. The four domains, N terminus, S domain, RRM domain and C terminus, has been deleted one after another and compared with the normal length RNPS1. A tethering assay demonstrated that only the removal of the S domain inhibited NMD activity. There was no loss in activating NMD, even if all three other domains were deleted. Now it can be assumed that the S domain is a binding site of Upf2 (Gehring et al. 2005).

Lykke-Andersen, Shu and Steitz (2001) showed, that Upf2 was immunoprecipitated by RNPS1, as well as all the other human Upf proteins. (Lykke-Andersen, Shu and Steitz, 2001). Therefore, interactions between the S domain of RNPS1 and Upf2 could be verified. (Gehring et al. 2005)

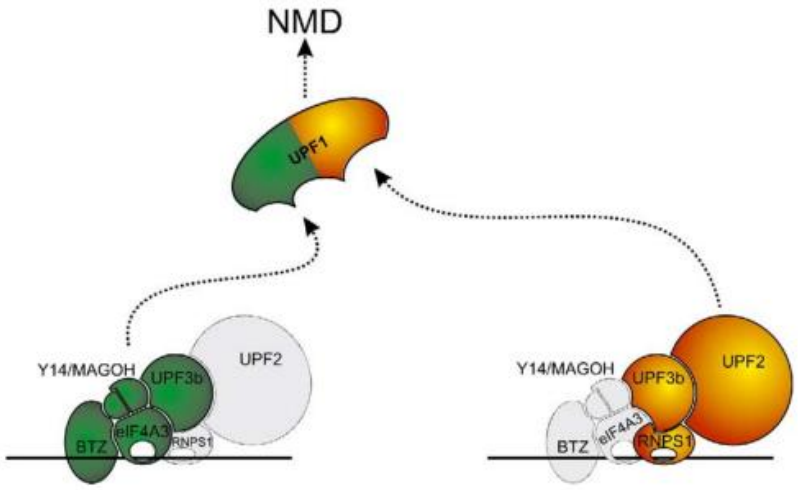

Figure 13: Illustration Of The Branched Model For Mammalian NMD

Proteins that are important actors in the specific complex are colored in green (NMD triggered by eIF4A3, Y14 and MAGOH) or in orange (NMD activated by RNPS1). The proteins colored in grey indicated proteins, which are not decisive for complex to activate NMD.

Source: cf. Gehring, N. H., Kunz, J. B., Neu-Yilik, G., Breit, S., Viegas,, M. H., Hentze, M. W. and Kulozik, A. E., 2005. Branched Model for Mammalian NMD.[image online] Available at: Molecular Cell <http://www.sciencedirect.com/science/article/pii/S1097276505015546> [Accessed 05 January 2015].

Gaining more insight into the Upf2-independent pathway, the complex formation was analyzed. In a tethering assay both eIF4A3 and Upf3b showed to be decisive. Although Upf3b is also necessary within the RNPS1-mediated pathway, it is also required to trigger Y14-MAGOH-induced pathway, but it is less essential than in the Upf2-dependent pathway. eIF4A3 and BTZ, when bound to Y14 and MAGOH, depict to be an important requirement for active NMD. However, Upf3b could be identified to directly bind to Y14 and MAGOH. eIF4A3 and BTZ could be shown to be needed for the formation of a stable complex. The Y14-MAGOH complex is directly bound by eIF4A3 and as a secondary reaction, BTZ connects to the complex through binding to the N terminus of eIF4A3. This terminus is also important for the connecting to the Y14-MAGOH complex. The linking between Y14 and eIF4A3 can occur independent of Y14-MAGOH-complex, but the formation of the whole complex stabilizes the connection between eIF4A3 and BTZ. NMD can only be triggered, if the whole complex is correctly build up, but without the necessity of Upf2 (Gehring et al. 2005).

RNPS1 communicate the position of exon-exon junctions to the mRNA surveillance machinery

Splice junctions are important as they communicate the selective transport of mature mRNA into the cytoplasm. Moreover, junctions produced during splicing serve as a mark for being recognized by the nonsense-mediated decay machinery. The postsplicing junctions consist of several subunits, which form the junction complex (Lykke-Andersen, Shu and Steitz, 2001).

Using immunoprecipitation, the postsplicing junction complex could be identified to contain five components. The splicing-associated factors SRm160 (Ser/Arg-related nuclear matrix protein), DEK and RNPS1, an mRNA shuttling protein Y14 and the mRNA export factor (REF) revealed to be part of the complex and play a crucial role in triggering NMD (Le Hir et al., 2000).

The three components, which are directly associated with the splicing event, are SRm160, DEK and RNPS1. The splicing process is activated by the formation of a complex between SRm160 and SRm300, which are both nuclear matrix components and belong to the SR splicing factor family. This complex interacts with SR proteins and U1 snRNP. DEK, a phosphoprotein, was also found to interact with

the splicing complex of pre-mRNA and it was shown that DEK also shows chromatin remodeling activity. However, the function of DEK during pre-mRNA splicing is not yet revealed (Le Hir et al., 2000).

The significance of RNPS1 is given due to the supply of an RNA recognition motif (RRM), which allows the tethering of the complex to the RNA and, therefore, the activation of the splicing process. Consequently, the assembly of SRm160, DEK and RNPS1 with the spliced mRNA is the result of their involvement in the splicing event. This leads to the assumption that the complex binds to the pre-mRNA during an early stage of spliceosome assembly (Le Hir et al., 2000). RNPS1 is demonstrated to be one of the most important subunits triggering nonsense-mediated decay when bound to the 3'UTR (untranslated region) of ß-globulin mRNA. RNPS1 is positioned 5' to the exon-exon junctions and serves as interaction platform for the human UPF complex (Lykke-Andersen, Shu and Steitz, 2001). The interaction of human Upf proteins, hUpf1, hUpf2 and hUpf3 with the 3' UTR of ß- globulin mRNA causes rapid degradation of nonsense mRNA (Lykke-Andersen, Shu and Steitz, 2000).

The other two proteins, Y14 and REF, are not part of the spliceosome per se, but they both function as mRNA transporters from the nucleus to the cytoplasm. However, both proteins showed to contain the same RNA recognition motif like RNPS1 and thus, both are able to directly interact with mRNA (Le Hir et al., 2000).

hUpf3 binds to the postsplicing complex through interactions with the RNPS1, which can further tether to the export protein TAP. TAP is an export receptor, which triggers the export of the spliced mRNA from the nucleus to the cytoplasm. During this process, not all of the bound proteins of the complex remain associated. REF and Aly, which help to form the link between TAP and the spliced mRNA, dissociate before transportation, as well as SRm160 (Kim et al., 2001).

During the pioneer round of translation, the remaining bound proteins are wiped off the mRNA and return to the nucleus. If there were PTCs produced during the splicing event and the premature codon is recognized, the NMD is elicited. Interactions between the eukaryotic release factors eRF1 and eRF3 and the downstream complex, which is located adjacent to the exon-exon junction, trigger decapping of the mRNA. (Lykke-Andersen, Shu and Steitz, 2001).

Y14 and hUPF3b form an NMD-activating complex

To gain better insight into the Upf2 independent pathway, Gehring et al. (2003) demonstrated that the human Upf3b is the activator for the Upf2 independent NMD when bound to the Y14 (Gehring et al., 2003). It is assumed that hUpf3b is the first factor that binds to the EJC in the nucleus after the process of splicing. This factor is thought to be followed by the human Upf2 factor, which is a perinuclear factor and tethers to the N-terminus of both hUpf3a and hUpf3b (Lykke-Andersen, Shu and Steitz, 2000,2001). However, as mentioned before, the Upf2 is not of crucial necessity for elucidating NMD. The Upf2 independent pathway could been shown to be dependent on the EJC protein Y14, which is bound to the mRNA in both the nucleus and the cytoplasm (Le Hir et al., 2001). Y14 was found to establish ties with MAGOH (Le Hir et al., 2001) and it could also been demonstrated that Y14 remains tethered on the mRNA until the mRNA enters the polysome, where it is removed (Dostie and Dreyfuss, 2002). These results are consistent with the assumption that Y14 is an important activator for NMD.

Using homologous analyses, a stretch of 14 amino acids (position 421-434 in hUpf3b) could been shown to be present in hUpf3a and hUpf3b, but also in Upf3 homologous of other eukaryotes, except yeast. The insertion of point mutations within this stretch of amino acids showed the reduction of the surveillance mechanism. The amino acids 421-434 are identified to be a crucial domain for triggering NMD. Moreover, the arginine at position 423 has been revealed to be a decisive part for NMD activation of bound hUpf3b. The C-terminus of hUpf3b was established to function as a bridge between the EJC and the NMD mechanism (Gehring et al., 2003).

Coimmunoprecipitation of hUpf3b was used to identify which factors bind to this specific amino acid region and, therefore, being essential for the activation of the degradation mechanism. Y14 and hUpf2 were tested and they could show that these proteins bind to different regions on the hUpf3b. The mutant hUpf3b co-precipitate Y14, even though the Upf2 interaction domain was removed and the nonsense-mediated decay mechanism was elucidated. The results confirm that hUpf3b requires the interaction with Y14 to activate NMD, but not Upf2, which is consistent with the Upf2-independent pathway theory. So, it could be demonstrated that tethering of Y14 to the specific region is crucial for the activation of NMD and, therefore, Y14 is a bona fide NMD factor (Gehring et al., 2003).

The dependence of hUpf3b-induced NMD on Y14 was confirmed by using interference RNA. The knockdown of Y14 resulted in an impairment of the degradations mechanism and verifies the importance of Y14 as NMD protein. Furthermore, Gehring et al. (2003) tested the requirement of hUpf1 and hUpf2 using RNAi and knocking down one after the other. They showed that both hUpf proteins are necessary in downstream processes, whereby inhibition of Upf1 was much more efficient in impairing NMD than Upf2. Upf2 only plays a crucial role in triggering NMD through binding to hUpf3b, if hUpf3b is not bound to the mRNA. The unbound hUpf3b then needs the assistance of Upf2, which leads to the assumption that Upf2 acts as a stabilizer, but if hUpf3b is already bound to the mRNA, NMD triggering does not require Upf2 (Gehring et al., 2003).

Until now, it was thought that an EJC is the crucial mark that labels a stop codon to be premature. However, in mammalian cells, another protein could be identified to trigger NMD.

Stau1-mediated mRNA decay (SMD)

Until now, it was assumed that the splicing event is important for triggering NMD and an EJC is necessary. However, in mammals, a protein was detected, which does not depend on the presence of an EJC and, therefore, is not the result of splicing. The protein, which elicits NMD without the complex is called mammalian Staufen 1 (Stau1). Stau1 has been shown to be an RNA binding protein, which functions in mRNA transport, as well as control of translation. Containing double-stranded RNA binding domains, Stau1 tethers to an extended RNA secondary structure. Moreover, Stau1 is omnipresent in mammalian cells (Marión et al., 1999; Wickham et al., 1999) and involved in many different processes within the cell, like DNA replication, cell division, RNP assembly or localization (Figure 14) (Le, Sternglanz and Greider, 2000).

Kim et al. (2005) demonstrated that Stau1 could bind Upf1 in mammals, as well as the 3'UTR, and trigger the degradation of mRNA without dependence on splicing. Using sequence analysis, the binding of Stau1 to Upf1 could be identified. Stau1 has several double-stranded RNA binding domains, which were shown to interact not only with RNA but also with the Upf1 protein. The results of

immunopurification revealed that the interaction between Stau1 and Upf1 are not dependent on the presence of RNA (Kim et al., 2005).

The downregulation of Stau1, using small interfering RNA, showed no effect on the EJC-dependent decay mechanism, which further lead to the assumption that Stau1 is not required for the EJC-dependent NMD. However, the downregulation of Upf1 resulted in not triggering NMD. The outcome of the experiment shows that EJC-dependent NMD requires Upf1 bound to Upf2 and Upf3, but not Stau1 (Kim et al., 2005).

To demonstrate the independence of EJC, Kim et al. (2005) used Arf1 (ADP-ribosylation factor 1) mRNA, which was identified to be a ligand of Stau1, and produced cDNA to avoid EJCs. The lack of EJCs shows whether Stau1 is dependent on the splicing process or not. The findings suggest that Arf1 mRNA was reduced through the presence of Stau1 and Upf1, but the mechanism does not require Upf2 or Upf3. Moreover, the deletion of 3'UTR showed an increase of Arf1 mRNA and, therefore, confirms the binding of Stau1 to 3'UTR. To sum up, the Stau1-mediated mRNA decay is dependent on a 3'UTR, as well as the presence of Upf1, but not Upf2 or Upf3. Additionally, it was shown that the PTC must be located upstream of the Stau1 binding site (SBS) to trigger the Stau1-mediated decay of nonsense mRNA. (Kim et al., 2005).

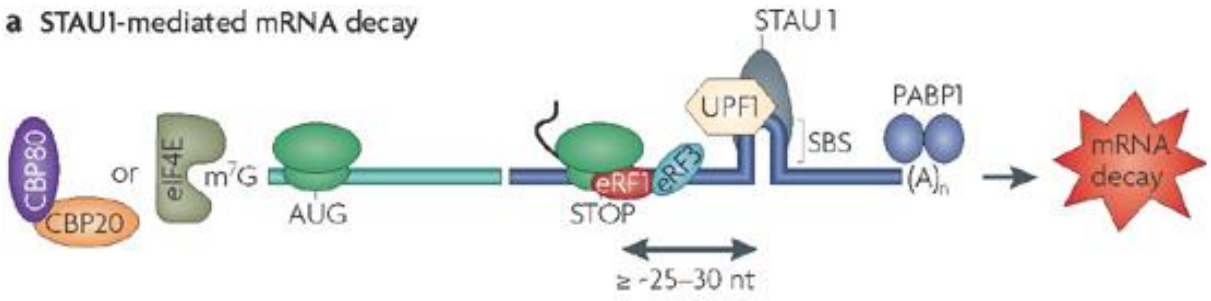

Figure 14: Stau1-mediated mRNA decay

Source: cf: Isken, O. and Maquat, L. E., 2008. UPF1 function in specialized mRNA-decay pathways.[image online] Available at: Nature Reviews Genetics <http://www.nature.com/nrg/journal/v9/n9/fig_tab/nrg2402_F2.html> [Accessed 05 January 2015].

Nucleus associated NMD

In mammalian cells, NMD targets mRNAs mostly before the release from the nucleus into the cytoplasm. However, there are still some nonsense mRNAs present

in the cytoplasm, which are degraded after the transport out of the nucleus (Thermann et. al., 1998).

Nevertheless, NMD is dependent on translation, which occurs in the cytoplasm. The nucleus-associated NMD must, therefore, take place during the export of mRNA from the nucleus to the cytoplasm. It is assumed that during the export of mRNA, 5' end first, the cytoplasmic ribosome binds to the 5'cap of the mRNA before the 3' poly(A)tail is released from the nucleus pore complex (Mehlin, Daneholt and Skoglund, 1992). However, it cannot be excluded that there is a translation-like process inside the nucleoplasm. To gain more insight, proteins, which are associated with NMD mechanism, were identified and their role within the degradation determined (Ishigaki et al., 2001).

As mentioned before, the 5'end of the mRNA first leaves the nuclear pore complex and ribosomes bind to the mRNA. The 5' cap of mRNA contains cap binding proteins (CBP), which functions as a nucleocytoplasmic shuttling protein in mammalian cells (Visa et al., 1996). After the export of the mRNA into the cytoplasm, eIF4E (eukaryotic initiation factor 4E), which is another cap binding protein, interacts with CBP80 (McKendrick et al., 2001) and replaces the CBP80 at an, until now, undefined stage (Ishigaki et al., 2001). Another cap binding protein, which is expected to be part of the NMD mechanism, is called CBP20, which associate with CBP80 to form the cap binding complex, as well as with eIF4G (Fortes et al., 2000)
On the basis of quantitative analysis of RNAs, proteins associated with RNAs are identified and, therefore, the determination of NMD associated proteins is possible. The results of Western blotting, using anti-CBP80 antibody, showed the specific immunopurification for CBP80-containing complexes. The quantitation of CBP80-associated mRNA revealed that the amount of nonsense mRNA was decreased to 19 % ±4 % of the amount of normal mRNA, which does not contain nonsense mutations. These results lead to the assumption that NMD takes place, if mRNA is associated with CBP80 (Ishigaki et al., 2001).

Additional insight was given, using real time PCR (qPCR), to demonstrate a possible linkage between CBP80, eIF4E and NMD. The results demonstrated that more than 80 % of CBP80-associated mRNAs, which contain a nonsense mutation, are targeted by NMD. However, the results did not state clearly the importance of eIF4E. There is no evidence, whether eIF4E does not activate NMD at all or whether

eIF4E-associated mRNA is also degraded due to the detection of NMD (Ishigaki et al., 2001).

Using Western Blotting and anti-CBP80 antibody, four additional proteins were identified, which are associated with CBP80-bound mRNA. CBP80 was found in complex with CBP20, which together form the cap binding complex. Moreover, CBP80 was found tethered to poly(A)binding protein II (PABP2), which is bound to pre-mRNA or mRNA. Upf2 and Upf3 are known to function within the NMD mechanism and could also be shown to be bound to CBP80 on mRNAs, but not in eIF4E-containing complexes. These results lead to the assumption that NMD mechanism targets mRNAs when bound to CBP80, but mRNAs are not detected when associated with eIF4E (Ishigaki et al., 2001).

Three of the four proteins (CBP20, PABP2, Upf3) are especially found in the nucleus, which reveals the possibility of NMD in the nucleus. However, PABP2 and Upf3 are part of the mRNA export shuttling and, therefore, NMD could take place outside of the nuclear envelope in the cytoplasm (Ishigaki et al., 2001).

Surprisingly, Upf1 was not found bound with neither CBP80 nor eIF4E associated mRNAs. The absence of this protein, which interacts with Upf2 and Upf3, proceed on the assumption that Upf1 is only required in cytoplasmic NMD after translation termination (Ishigaki et al., 2001).

Cytoplasmic NMD

Cytoplasmic NMD takes place in the cytoplasm of the cell on cytoplasmic ribosomes. Due to results of an mRNA immunopurification with anti-CBP80 antibody and anti-eIF4E antibody, it was shown that mRNA is no longer appreciably susceptible for degradation if CBP80 is replaced by eIF4G. It could also be demonstrated that CBP80-associated mRNA, but not eIF4G-associated mRNA, is tethered to CBP20, PABP2, Upf2 and Upf3. Moreover, none of them was detectably bound to Upf1 protein (Ishigaki et al., 2001).

In mammalian cells, most mRNAs containing a premature termination codon (PTC) are target of the nucleus-associated NMD. However, there was ß-globulin mRNA found in the cytoplasm, which contained PTCs. This leads to the assumption that these mRNAs escaped the nucleus-associated NMD (Baserga and Benz, 1992). There is also evidence that the 3’ poly(A)tail can communicate with the translation

termination codon and it could be shown that the mammalian translation releasing factor 3 (eRF3) can bind to the poly(A)binding protein (PABP). This leads to the further assumption that the eRF3 bound to the poly(A)tail helps to recognize termination codons and is, therefore, involved in mammalian nonsense-mediated decay (Hoshino et al., 1999).

To illustrate the deadenylation process of PTC containing mRNAs, deadenylation kinetics were used by extracting cytoplasmic mRNA from transfected cells at different time points. So, a clear deadenylation determination *in vivo* was able to be displayed. Chen and Shyu (2003) revealed that PTC containing mRNA shows rapid deadenylation prior to the decay of the mRNA in the cytoplasm (Chen and Shyu, 2003). In comparison with a wild-type mRNA, the deadenylation process occurred much slower without being followed by the process of decay. They confirmed the existence of an active NMD pathway in the cytoplasm that is distinct to the pathway in the nucleus (Hoshino et al., 1999).

Using a step-loop structure, which efficiently blocks translation, they displayed that the accelerated deadenylation of the mRNA, involving a premature termination codon, is dependent on an active translation of the mRNA. The blocking of the translation of the PTC containing mRNA obstruct the deadenylation process and, therefore, NMD cannot occur (Hoshino et al., 1999).

Moreover, to verify a bona fide NMD in the cytoplasm, Chen and Shyu (2003) tested the correlation between bona fide NMD and hUpf1 by inserting a dominant-negative mutant of hUpf1. Monitoring the kinetics of PTC-containing mRNA in the presence of the mutated hUpf1 protein, the results showed that mutations in hUpf1 hamper NMD in proliferating mammalian cells. The dominant-negative mutant of hUpf1 considerably slows down deadenylation of the target mRNA, although after 2 hours a subset of PTC-containing mRNA was still degraded. However, a significant portion of PTC-containing mRNAs still existed without being target of the decay mechanism. These results suggest that hUpf1 has a considerable role in the rapid deadenylation in the cytoplasm triggered by PTC containing mRNA (Hoshino et al., 1999).

To verify the importance of deadenylation within the cytoplasmic NMD, RNA-binding proteins, UNR and NSAP1, were used to impede deadenylation. These two proteins disturb the function of the poly(A) nuclease and, therefore, impede cytoplasmic mRNA deadenylation. It could be shown that deadenylation was

completely arrested and no degradation of the PTC - containing mRNA occurred. These results show again that deadenylation of PTC-containing mRNA is necessary for cytoplasmic nonsense-mediated decay in mammalian cells (Hoshino et al., 1999).

In conclusion, the cytoplasmic NMD pathway could be substantiated to be dependent on the removal of the poly(A)tail of PTC-containing mRNA before being target of the NMD in the cytoplasm. Furthermore, the accelerated deadenylation of PTC-containing mRNA is dependent on an active translation process and the presence of hUpf1. The results of the examination showed that interfering with deadenylation has effect on the decay mechanism and that the prevention of deadenylation always lead to stabilization of nonsense mRNA (Hoshino et al., 1999).

Degradation of mRNA

mRNA decay rates can vary substantially. In mammalian cells the average mRNA half-life is approximately 8 hours. For instance, C-fos mRNA degrades very fast within 15 minutes, whereas ß-globulin mRNA needs about 17 hours for degradation (Parker, 2013).

One feature of mRNA decay rates that controls mRNA decay are stability elements, which are short sequence elements within the mRNA that control the mRNA turnover. Stability elements are features that determine the mRNA turnover. The stability elements serve as binding sites for trans-acting factors that control mRNA decay. mRNA caps and poly(A)tails play crucial roles in translation and mRNA degradation. One of the striking features of mRNA decay and translation is how they interact with each other. The distinct role of the cap and the poly(A)tail are that, in translation, the cap and the poly(A)tail serves as binding sites for RNA binding proteins, which then assemble as complex that promotes ribosome loading. In Contrast, during degradation, the poly(A)tail is removed by the CCR4 or PARN2 complex. This automatically leads to decapping, which is carried out by the decapping enzyme and once the cap is removed, the 5’ to 3’ exonuclease degrades the mRNA extremely rapidly. However, this is only one pathway that can be initiated (Parker, 2013).

In general, translation and mRNA decapping are inversely related. During translation, there is a complex that contributes ribosome loading, during degradation the complex is lost and degradation is promoted (Parker, 2013).

Moreover, there are two mRNP complexes which must be differentiated in the cell. The translation mRNP complex consists of eIF4E, eIF4G, eIF4A and eIF4B. This complex is also known as the eIF4F complex and it can only assemble, if the decapping mRNP complex is not bound to the mRNA. The decapping mRNP complex consists of DHH1, PAT1, DC1, DC2 and Dcp1 and Dcp2. Dcp1 and Dcp2 cleave the cap. Furthermore, the mRNA decapping factors are not associated with the polysome and accumulate in P-bodies. Polysomes are the locations where ribosome mRNA complexes are produced (Parker, 2013).

In mammalian cells, many proteins can be found in the P-bodies, which are among others Dcp1 and Dcp2 and also Xrn1, which are involved in decapping (Parker, 2013).

After detection of an aberrant mRNA through the mechanisms described above, the process of degradation remains open. Despite information about the factors and mRNP proteins, which are required for the triggering of the mechanism are known, the polarity and the enzymology of the degradation mechanism are not understood completely yet. Upf factors are known that they act as indicators, whether mRNAs are target of NMD or not. However, the process after detecting of the nonsense mRNAs is not completely solved yet (Lykke-Andersen, 2002).

In eukaryotic cells the first step of degradation is the shortening of the 3' poly(A)tail by a variety of nucleases. The two most important nucleases are the CCR4 complex and the less important PAN2 and PAN3 complex, which seems to play a more minor role. Following the deadenylation, the mRNAs can be degraded from the 3' end from a complex called the exosome, which takes advantage of different cofactors. Within the more common degradation pathway, mRNA is decapped by decapping enzymes followed by 5' to 3' degradation of the mRNA. The distribution between these two mRNA decay pathways is not solved yet (Parker, 2013).

It could be revealed that NMD includes a 5'-end degradation, which was shown on the basis of coimmunopurification of Dcp1 with Upf1. Dcp1 is known as a component of the decapping complex. Dcp1 coimmunopurifies with Upf1 independently of the presence of the other Upf proteins (Upf2, Upf3a, Upf3b). Dcp2 is another decapping protein, which communicates decapping and can be found in both the nucleus and the cytoplasm (Lykke-Andersen, 2002).

To verify that mammalian cells use both, the 5’ to 3’ mRNA decay pathway and the 3’ to 5’ mRNA decay pathway, downregulation of either Dcp2, PM/Scl100 and PARN (poly(A)ribonuclease) was carried out. By reducing the amount of Dcp2 and PM/Scl100 to 35% of normal, using RNAi, the level of nonsense - haboring mRNAs increased to 3-fold to 5-fold. So the presence of Dcp2 RNAi and PM/Scl100 RNAi decrease the level of nonsense - containing mRNA. The higher amount of nonsense mRNA was proven not to be result of decreased translation activity because the downregulation of the proteins did not affect the cap-dependent luciferase activity produced by a transiently introduced plasmid. Due to the results of immunofluorescence and confocal microscopy, Dcp2, as well as PM/Scl100, are present in both the nucleus and the cytoplasm. These results lead to the assumption that both nucleus-associated and cytoplasmic-associated NMD carries out decapping, as well as 3’ to 5’ exonucleolytic decay. The downregulation of PARN should confirm that the NMD also includes a 3’ to 5’ exonucleolytic decay. The downregulation of PARN to 15% of normal amount lead to a rise of nonsense mRNA level in both nucleus - associated and cytoplasmic - associated NMD. These results suggest that deadenylation is followed by 3’ to 5’ exonucleolytic degradation in both the nucleus and the cytoplasm (Lejeune, Li and Maquat, 2003).

Xrn1 and Xrn2 (Rat1) are 5’ to 3’ exonucleases, although only Xrn1 could be shown to be active within the NMD. Xrn1 was detected in the cytoplasm using immunofluorescence and purification of Xrn1 functions as a 5’ to 3’ exoribonuclease (Bashkirov et al., 1997). Xrn2 is found in the nuclear fraction, consistent with the findings of the Rat1 yeast ortholog, which is also found adjacent to the nuclear envelope using indirect immunofluorescence. Purified protein fractions of Xrn2 has shown double-stranded DNA degradation as a 5’ to 3’ exonuclease (Lejeune, Li and Maquat, 2003).

Using immunodepletion studies, PARN, a poly(A) ribonuclease, could be identified to show deadenylase activity and the protein was also detected in the nucleus of the cell (Körner et al., 1998).

However, the process of degradation is not only carried out by 5’ to 3’ exonucleases. In the human exosome also 3’ to 5’ exonucleases could be detected. These components, which are Rrp4, Rrp40, Rrp41, Rrp46, Csl4, Dis3p, PM/Scl75 and PM/Scl100 could be identified to be present in both the nucleus and the cytoplasm (Allmang et al., 1999; Brouwer et al., 2001; Raijmakers et al., 2002).

In order to corroborate the results that NMD involves both 5' to 3' and 3' to 5' mRNA decay an assessment of factors that bind to specific NMD factors was carried out. The NMD factors used are the three Upf factors, which have been identified to be necessary for NMD. All three Upf factors were immunopurified from Cos cells under preservative conditions to maintain mRNP. The efficacy of each immunopurification was confirmed using Western blotting. This method was also used to detect the presence of Dcp2, the assumed 5' to 3' exoribonuclease Xrn2, the known 5' to 3' exoribonuclease Xrn1, PM/Scl100, the exosomal components Rrp4 and Rrp41, and PARN. The results of the coimmunopurification showed that not only Dcp2, PM/Scl100 and PARN coimmunopurify with Upf factors, but also Xrn2 and Xrn1 exoribonucleases, which were assumed to be part of the degradation machinery, could be shown to immunopurify. Thus, both directions of exoribonuclease activity are part of the NMD pathway. Moreover, Rrp4 and Rrp41, which are exosomal components, could be immunopurified with all Upf factors. Both exosomal components act tethered to PM/Scl100 and after PARN (Lejeune, Li and Maquat, 2003).

For strengthening the results, an anti-Dcp2 antibody was used and the antibody immunopurified Upf1, Upf2, Upf3b, Xrn2, PM/Scl100, Rrp41 and PARN in presence and absence of RNase, which leads to the conclusion that, as soon as the complex is formed, the presence of RNA is not necessary (Lejeune et al., 2003).

Lejeune, Li and Maquat (2003), furthermore, examined, whether 5' exoribonuclease is degrading faster than 3' exoribonuclease and determining, where the NMD initiates primarily. Although they were not able to show at which end the degradation starts first, they could reveal that each end of nonsense mRNA was degraded faster than corresponding ends of normal mRNAs. However, a direct comparison between the degradation rates revealed that the 5'end of a nonsense-containing end degrades 1.5 fold quicker than the 3'end. Nevertheless, they concluded that the specific end, at which degradation starts first, remains unclear and that NMD probably initiates at both ends (Lejeune, Li and Maquat, 2003).

Depending on which end (5' end or 3' end) the degradation starts, it can be distinguished between two pathways. The 5' to 3' decay is initiated by the Dcp2 decapping enzyme and subsequently Xrn2, or Xrn1, which are 5' to 3' exonucleases, degrade the mRNA. The 3' to 5' decay is initiated by PARN and, at least the

PM/Scl100, a 3' to 5' exonuclease, which probably but not certainly in combination with the exosome, degrade the mRNA (Figure 15) (Lejeune, Li and Maquat, 2003).

These results corroborate that Upf NMD factors initiate the NMD machinery, which involves decapping, deadenylation and exonucleolytic enzymes, to trigger nonsense mRNA degradation in mammalian cells. However, the most important protein still seems to remain Upf1 because this protein interacts with decapping proteins in a direct or indirect manner.

The mRNA decay pathways are well conserved pathways in all eukaryotes, from yeast to humans, and it could be shown that in mammalian cells, the degradation can occur in either 5' to 3' or 3' to 5' direction. If mRNA is degraded normally, two different major pathways are described. In both pathways, the degradation in initiated with the shortening of the poly(A)tail of the mRNA.

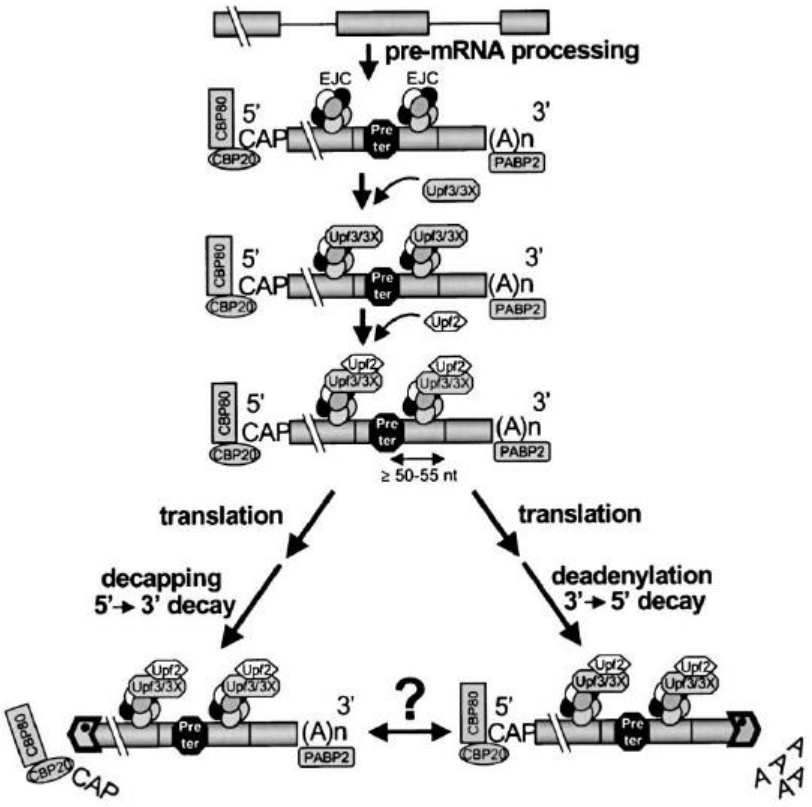

Figure 15: Two Pathways Of Degradation

Source: Lejeune, F., Li, X. and Maquat, L. E., 2003. Model for the Polarity and Enzymology of NMD in Mammalian Cells. [image online] Available at: Molecular Cell <http://www.cell.com/action/showFullTextImages?pii=S1097-2765%2803%2900349-6> [Accessed 05 January 2015].

Shortening of poly(A)tail

The initiation of degradation of mRNA is triggered by the shortening of the poly(A)tail (Wang et al., 1999). The degradation from the 3' end to the 5'end was shown to be probably 5-fold higher than a degradation starting at the 5'cap. In mammalian cells, the 3' to 5' decay pathway seems to be the major contributor to mRNA degradation. Moreover, it could be revealed that the deadenylation is always the first step in mRNA degradation, independent of the presence or absence of a cap at the 5' end (Wang and Kiledjian, 2001).

The first protein that is recruited to the mRNA for deadenylation is called CELF1. The protein is known to play a crucial role in the posttranscriptional gene expression regulation by tethering mRNA at GU-rich elements (GREs). GREs are highly conserved elements within the 3'UTR of mRNA in eukaryotic cells. CELF1 is an RNA-binding protein with three highly conserved RNA-recognition motifs (RRMs). In humans and mice, six members of the CELF family have been detected and they are shown to have significant functions in both the nucleus and the cytoplasm. The human CELF1 could be shown to contribute to mRNA deadenylation. Moreover, in HeLa cell extracts, CELF1 is the protein that first binds near to the poly(A)tail and recruits the poly(A) ribonuclease PARN. However, it is not known yet if CELF1 triggers the binding of other deadenylases in mammalian cells (Moraes, Wilusz and Wilusz, 2006).

In primary human T cells, GREs could be found in the 3'UTR of mRNAs and the CELF1 was detected to specifically bind to these conserved regions. CELF1 regulate the decay of GRE containing mRNAs. Once the CELF1 is bound to the 3'UTR of an mRNA, deadenylation is triggered and subsequent decay follows by the exosome (Figure 17) (Beisang et al., 2012).

However, next to GU-rich elements, also AU-rich elements (AREs) can trigger rapid deadenylation when present in the 3'UTR. ARE binding proteins (AUBPs) were identified to activate ARE- containing mRNAs through the recruitment of RHAU, an RNA helicase. RHAU is shown to interact with the deadenylase PARN, like CELF1. Furthermore, PARN triggers the assembly of the exosome, which is a complex of several exonucleases. These exonucleases degrade the mRNA from the 3'end (Figure 16) (Tran et al., 2004).

In mammalian cells, the 3' to 5' seems to be the predominantly pathway for nonsense-mediated mRNA decay. After the recruitment of the exosome, the

decapping is activated. This so-called scavenger decapping activity always follows deadenylation. The degradation-depending decapping activity of mRNA is considered to be complexed to a subset of exosome components, which are specific for 3' to 5' degradation (Wang and Kiledjian, 2001). However, a link between the RNA helicase and degradation-dependent mRNA decapping has not been drawn but it is assumed that RHAU is combined within cellular exosomes dedicated to the 3' to 5' degradation pathway (Tran et al., 2004).

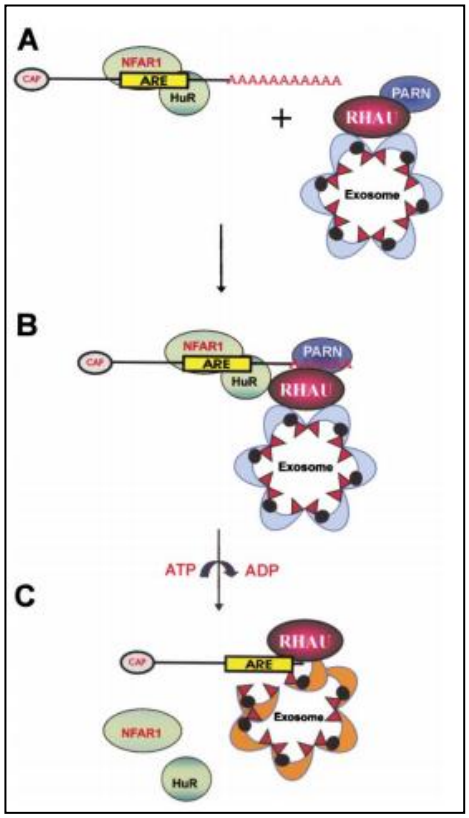

Figure 16: RHAU- mediated decay

Source:cf: Tran, H., Schilling, M., Wirbelauer, C. Hess, D. and Nagamine, Y., 2004. Model for RHAU-Mediated Recognition and Enhancement of AREuPA-mRNA Decay. [image online] Available at: Molecular Cell <http://www.sciencedirect.com/science/article/pii/S1097276503004817> [Accessed 04 January 2015].

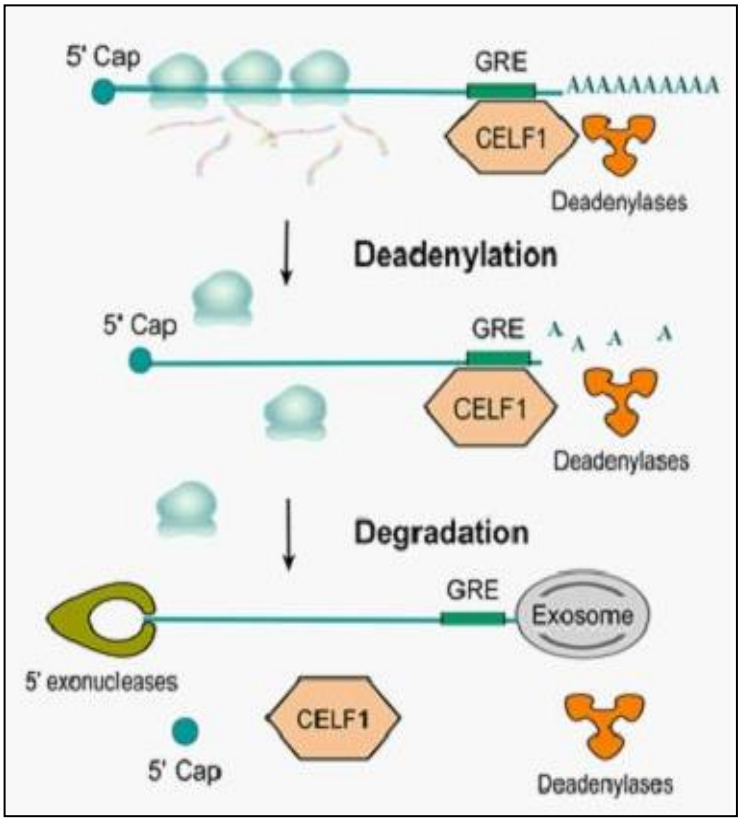

Figure 17: CELF1- mediated decay

Source: cf: Beisang, D., Bohjanen, P. R. and Vlasova-St. Louis, I. A., 2012. Evolutionary conservation of deadenylation by CELF1 protein and GU-rich sequences. [image online] Available at: Intech <http://www.intechopen.com/books/binding-protein/celf1-a-multifunctional-regulator-of-posttranscriptional-networks#SEC1> [Accessed 04 January 2015].

Decapping

Generally, the capping of mRNA provides protection of degradation, which could be shown comparing half-lives of capped and uncapped RNAs. The half-life of uncapped RNA was approximately 50 % less than the stability provided through a

cap. These results suggest that a 5’ to 3’ exoribonuclease activity enhances the turnover of mRNA (Wang and Kiledjian, 2001).

In eukaryotic cells, mRNAs are produced in the nucleus and for protein translation they must be transported out of the nucleus into the cytoplasm. For this process to happen, the mRNAs are equipped with several features. One of these features is a cap that is bound with an unusual 5’-5’ linkage to the mature mRNA. This cap is composed of a guanine nucleoside, which is methylated at the N-7 position (m7G) and covalently bound to the terminal nucleoside of the RNA. This structure is not only important for the transport into the cytoplasm, it is also necessary to protect the mRNA from exoribonucleolytic degradation and plays a crucial role in translation of the mRNA into a protein. Moreover, the 5’cap is involved in the process of splicing. In the nucleus, the 5’cap is bound by two cap binding proteins, which are CBP80 and CBP20. In the cytoplasm, the cap binding protein eIF4E also interacts with the heterodimeric cap binding protein complex (Fortes et al., 2000).

In eukaryotic cells, it can be distinguished between two different decapping pathways (Figure 18). In one of the two pathways, Dcp1 hydrolyzes the cap of an mRNA and subsequently releases an m7GDP. A decapping enzyme, known as Dcp1, which is a pyrophosphatase that removes the m7GDP of the mRNA, triggers the degradation of the cap. This enzyme was shown to favor hydrolyzing methylated caps of RNAs, which are longer than 25 nucleotides (LaGrandeur and Parker, 1998). Before the decapping of mRNA is recruited by Dcp1, the poly(A)tail is deadenylated to an oligo(A). The deadenylation is a crucial step, which makes the mRNA susceptible for decapping (Beelman et al., 1996). The Dcp1 does not only initiate and accelerate the process of decapping, it also allows the binding of another decapping protein Dcp2. The interaction of these two decapping proteins recruits more proteins, which act as activators, including Dhh1/RCK/DDX6, Edc3, Edc4/Hedls/Ge-1 and Pat1. These activator proteins are conserved during evolution from yeast to mammals (She et al., 2008).

The second pathway, which could also be detected in mammalian cells, also carries out decapping, but this type of protein releases m7GMP and functions on the residual cap, which is followed by degradation of the mRNA. The two decapping proteins are assumed to be different. Moreover, the Dcp1-like protein only hydrolyzes

capped oligonucleotides with a length less than 10 nucleotides, but does not affect mRNAs with more nucleotides (Nuss et al., 1975).

In an *in vitro* decay assay, the pyrophosphatase activity in mammalian cells could be revealed to be present during the decapping process. Wang and Kiledjian (2001) demonstrated that m7GMP is the primary cleavage product of decapping and not only m7GDP, which is cleaved also in yeast. Moreover, the pyrophosphatase activity is specific for methylated N-7 positions on the terminal guanosine of the 5'cap and mRNAs lacking the N-7 methyl group were prevented of decapping. However, decapping is still preceded by the process of deadenylation. The poly(A)tail on the 3' end has even the ability to inhibit the decapping process. Deadenylation of the 3' end leads to a 5-fold higher degradation of the 5'cap than the presence of a poly(A) tail (Wang and Kiledjian, 2001).

In mammalian cells, the essential function of the 5' to 3' decapping mechanism, however, has not been figured out. Furthermore, it is not a major

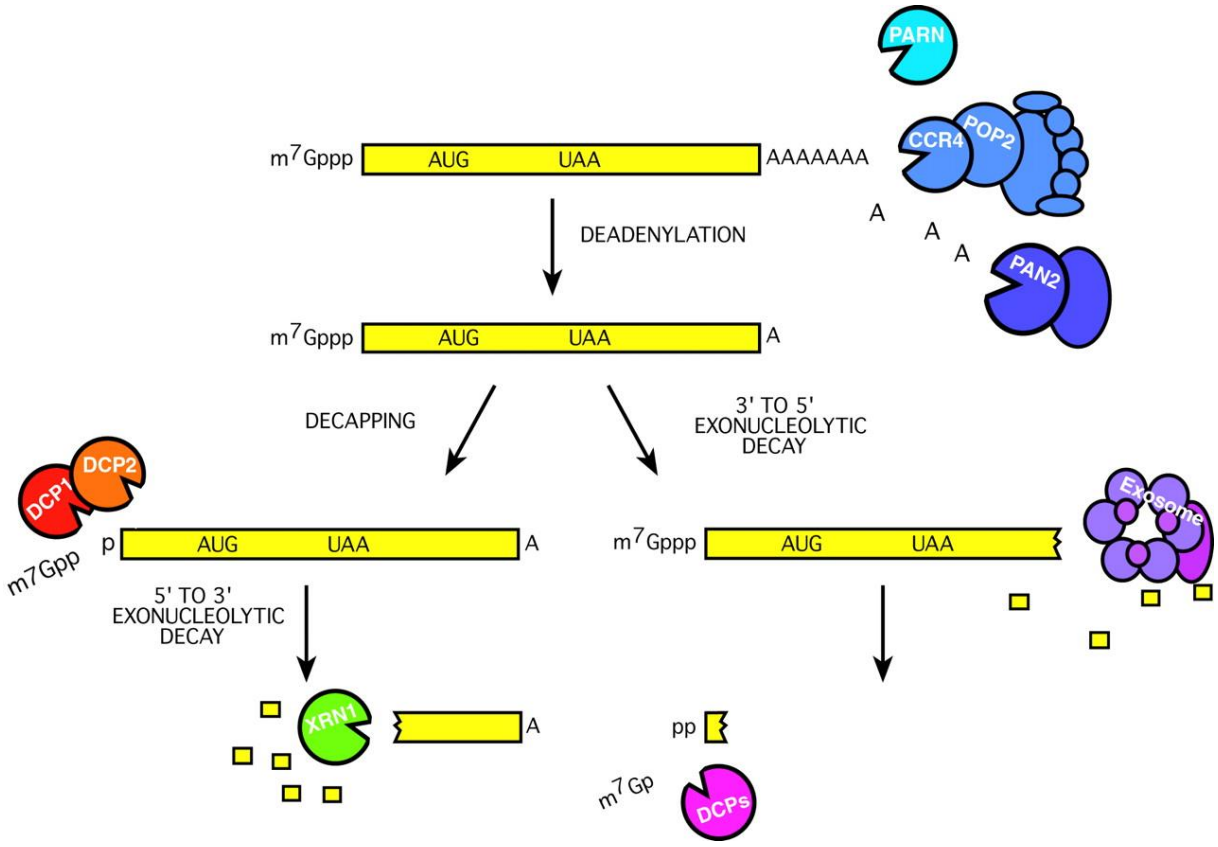

Figure 18: Degradation pathway: deadenylation followed by decapping (left), deadenylation followed by 3' to 5' exonucleolytic decay (right)

Source: cf: Decker, C. J. and Parker, R., 2002. Eukaryotic mRNA decay mechanisms and enzymes. [image online] Available at: Proceedings of the National Academy of Sciences <http://www.pnas.org/content/99/20/12512.full> [Accessed 04 January 2015].

promotion of the NMD to start the degradation at the cap end (Wang and Kiledjian, 2001).

After decapping, the 5'end is exposed and, subsequently, the degradation of the mRNA is initiated by Xrn1. Xrn1 is, like mentioned before, a 5' to 3' exoribonuclease, which degrades the uncapped mRNA quickly (Wang and Kiledjian, 2001).

However, like in every system of the body, the mechanism of nonsense-mediated mRNA decay does not always work properly. Proteins could be identified, which are able to suppress the mechanism, which further can have severe consequences.

Suppression of NMD

One of these proteins could be identified to be the eukaryotic initiation factor 4G (eIF4G), which is able to suppress the NMD by two genetically separable mechanisms. For initiating NMD, many different proteins are assembled to a complex, which then triggers deadenylation, as well as decapping of the mRNA and the mRNA gets degraded, as mentioned above. So, the kinetic NMD model implies that proper translation termination depends on specific termination encouraging signals.

Furthermore, Joncourt et al. (2014) were able to depict that NMD cannot be activated, if the cytoplasmic poly(A)binding protein (PABPC1) is tethered to the mRNA. The activation fails because the proteins, which trigger the NMD are not able to bind because the binding site on the mRNA is occupied by PABPC1.

To gain more insight into the molecular mechanism, Joncourt et al. (2014) found out that, if the PABPC1 is bound adjacent to a premature termination codon, NMD is suppressed. The PABPC1 is a major cytoplasmic poly(A)binding protein, providing four RNA recognition motifs (RRMs), a linker region and a eRF3-binding domain (Figure 19). The four different RRMs consist of conserved domains for RNPs. These domains are needed for specific and efficient binding of RNA molecules. RRM1 and RRM2 provide high affinity for binding the poly(A)tail of the mRNA, if it consists of 12 to 25 nucleotides. Moreover, eIF4G and the PABP-interacting proteins (PAIP1 and PAIP2) also have the ability to tether RRM1 and RRM2 (Joncourt et al., 2014).

In order to allow multimerization of PABPC1, the four RRM domains are followed by a linker region, which is rich in proline and glutamine. Many translational factors are transported by the C-terminal domain of PABPC1 to the poly(A)tail of the mRNA possessing the PABP-interacting motif 2 (PAM2). PAM 2 consists of eRF3, eIF4B, PAIP1 and PAIP2.

To identify which domains of the PABPC1 are required for suppressing NMD, fusion proteins are created and transiently expressed in HeLa cells. The fusion proteins consisted of different domains of the PABPC1 and the reporter protein MS2. So, the HeLa cell expressed the fusion proteins PABPC1::MS2 and the mRNAs with different MS2 binding sites. The mRNAs were expressed in various ways, in order to provide binding sites for the MS2 protein. The different constructions of mRNA contain binding sites located adjacent to the PTC or adjacent to the normal termination codon (Picture 10). On the basis of real-time PCR, they could show that the binding of the fusion protein lead to a 9-fold abundance of mRNA if the PABPC1 is bound adjacent to the PTC. Whereas the binding of the MS2 protein alone did not affect NMD, which served as a control (Joncourt et al., 2014).

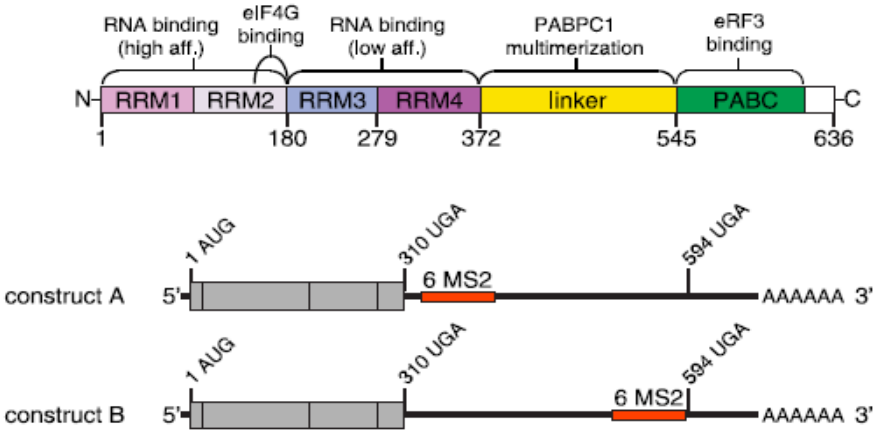

Figure 19: Schematic representation of PABPC1 and mRNAs with MS2 binding sites

PABPC1 consists of four RRM domains, linker domain and C- terminal domain (above), Construct A and Construct B present two different mRNAs providing binding sites for MS2 proteins, located adjacent to the PTC (Construct A) and adjacent to the normal termination codon (Contruct B).

Source: cf: Joncourt, R., Eberle, A. B., Rufener, S. C. and Mühlemann, O., 2014. The C-terminal domain of PABPC1 is dispensable for inhibition of NMD, RRMs1-3 are sufficient, and RRMs1-2 are necessary. [image online] Available at: PLOS ONE <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0104391> [Accessed 04 January 2015].

The deletion of the C-terminal part of PABPC1, which involves the eRF3 binding site, rose the mRNA level with the same efficiency as the full length PABPC1. The reporter mRNA level increased similarly, when PABPC1 without RRM1 and RRM2 were bound. However, RRM1 and RRM2 are required for poly(A)tail binding of the mRNA, so these results cannot be possible. To explain these results, it is assumed that the linker region of the PABPC1 polymerized with an endogenous PABPC1, which then bound the mRNA. So the results were obscured by an endogenous PABPC1, which is always present within a cell. In further experiments, Joncourt et al. (2014) were able to show that RRM1 and RRM2 are necessary for NMD suppression by deleting RRM1, RRM2 and the linker region at once. Moreover, RRM1-3 could be identified to be able to sufficiently suppress NMD.

As shown in figure 19, RRM1-3 contain binding sites for RNA and eIF4G, which is located at RRM2. The comparison of proteins containing mutations within the RRM2 domain and non-mutated proteins resulted in a reduced ability to suppress NMD. In order to verify that eIF4G binding to PABPC1 is needed for suppressing NMD, the eIF4G was knocked down by expressing a shRNA for eIF4G mRNA. Full-length protein with the knockdown eIF4G has extremely decreased the activity of NMD, which means that the mRNA was not degraded. These results show that eIF4G is needed for the suppression of degradation of mRNA. eIF4G is known to be involved in protein synthesis. NMD depends on translation, so the depletion of eIF4G results in an increase of mRNA level. So, it could be shown again, that eIF4G is necessary for the suppression of NMD when bound to PABPC1.

The next step involves the identification of the specific part of eIF4G (Figure 20), which is needed for the suppression. Therefore, three parts are tested. Fusion proteins were created. The first one involved the N-terminal of eIF4G, which consists of the PABP and eIF4E interaction domains. The second fusion protein was only built up of the central MIF4G (middle domain of eIF4G) and the third fusion protein comprised the C-terminal part. A fourth fusion protein was tested, which did not contain the central MIF4G domain. The results show that without the core domain of eIF4G, no rise in mRNA level is caused (Joncourt et al., 2014).

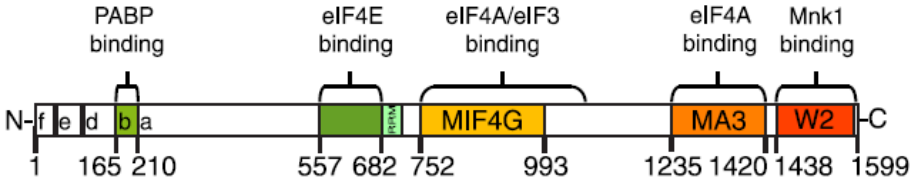

Figure 20: Schematic representation of eIF4G

Source: cf: Joncourt, R., Eberle, A. B., Rufener, S. C. and Mühlemann, O., 2014. The C-terminal domain of PABPC1 is dispensable for inhibition of NMD, RRMs1-3 are sufficient, and RRMs1-2 are necessary. [image online] Available at: PLOS ONE <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0104391> [Accessed 04 January 2015].

In conclusion, the eIF4G core domain can suppress NMD and this shows that there is a second mechanism that inhibits NMD, which is independent of PABPC1. The pathway that is independent of PABPC1 is detected to be related to the eukaryotic initiation factor 3 (eIF3) because the MIF4G domain provides a binding site of eIF3. For tethering, eIF3f and eIF3h (subunits of eIF3) are crucial. The depletion of either of the two eIF3 subunits had no effect on the PABPC1 dependent mechanism. However, depleting eIF3f and eIF3h in the PABPC1 independent pathway, the mRNA levels were reduced to 40-50 %. So, these results show that eIF3 is involved in the PABPC1 independent pathway, but is not needed in the PABPC1 dependent pathway (Joncourt et al., 2014).

A third way of preventing NMD could be revealed to be related to the Y14 protein, which is part of the exon junction core complex.

The RNA-binding protein Y14 was also identify to be able to inhibit the process of NMD by suppressing decapping and altering processing body (P-body) formation (Chuang et al., 2013). Y14 is a protein conserved in all eukaryotic cells and consists of an RNA-recognition motif (RRM) in the central domain, to be able to bind to mRNA. Moreover, two arginine-serine (RS) dipeptides are located at the C-terminal end and these are important for phosphorylation of the Y14 (Hsu et al., 2005). As mentioned before, the Y14/MAGOH heterodimer is an important part of the EJC and recruits several other factors for NMD to the nonsense-mRNA. The Y14 protein is tethered to the mRNA and is removed during the first round of translation (Dostie and Dreyfuss, 2002).

The Y14 has been shown to directly interact with the mRNA-decapping factors, Dcp1 and Dcp2 (Chuang et al., 2013). These decapping factors are found within, so called P-bodies. So, if an mRNA is targeted for degradation, it is recruited to the cytoplasmic P-bodies, where also the deadenylation factors and many other factors involved in NMD are located (Sheth and Parker, 2006).

Chuang et al. (2013) could depict that Y14 can interact with both the Dcp1 and Dcp2. Since the Dcp2 is the catalytic unit of the decapping complex, they focused on the interaction between Y14 and Dcp2 and revealed that interaction between these proteins inhibits the decapping process. However, Y14 can only suppress decapping when it is not bound to MAGOH. The heterodimer Y14/MAGOH failed to bind Dcp2, so possible competition of binding sites could be the reason. So, the disassociation of Y14 and MAGOH is required, that Y14 can interact with Dcp2 and, therefore, inhibit decapping. Furthermore, the dephosphorylation of Y14 is needed for the interaction of Y14 with Dcp2. However, their findings only provide a hint that Y14 can prevent mRNAs from being decapped. The overexpression of Y14, however, confirmed that Y14 is able to inhibit nonsense-mediated mRNA decay, when bound to Dcp2. Moreover, the prevention of decapping can also occur, even after deadenylation of the mRNA (Chuang et al., 2013).

It is known that many mRNA-decapping factors are located within the P-bodies and whenever an mRNA is targeted by NMD, the mRNA is recruited to P-bodies. There are many factors within the P-bodies, which directly interact with the decapping-complex and, therefore, the P-bodies have a crucial role in NMD. However, it could be shown that Y14- depletion in cells leads to a fewer detection of P-bodies. This results indicate, that Y14 is also involved in the proper formation of P-bodies, whereas depletion of eIF4A3 did not affect the number of P-bodies (Chuang et al., 2013)

The overexpression of Y14 in P-bodies showed that the phosphorylation of Y14 plays an important role to the formation of P-bodies. The overexpression of phosphorylated Y14 caused an increase of P-body formation, however, non-phosphorylated Y14 did not alter the level of P-bodies (Chuang et al., 2013).

If the mechanism of nonsense-mediated mRNA decay does not work properly due to different reasons, diseases, like ß-thalassemia or Marfan syndrome, can be the result. This thesis will go into more detail about this two diseases is the next part.

Diseases caused due to a non-functional NMD

ß-Thalassemia

Thalassemia is an inherited genetic blood disorder that affects red blood cells, which contain the hemoglobin molecules. In thalassemia, one of the constituents of the hemoglobin molecule is produced at an amount that is not sufficient or is not produced at all. Beta thalassemia patients suffer from the lack of beta chain production, which means inadequate or absent synthesis of the ß-globulin chains, due to a mutation in the HBB gene. The hemoglobin molecule consists of four subunits, normally two ß- globulin chains and two α-globulin chains. In the case of ß-thalassemia, the ß-globulin chain is not produced appropriately, which leads to the accumulation of α-globulin. Red blood cells, then, cannot develop normally and this further leads to an inappropriate oxygen supply of the body (U.S. National Library of Medicine, 2014).

As explained in the background, the conversion of genetic information is necessary to produce proteins. Therefore, a copy of the DNA, which is called RNA, is produced and transported out of the cell. Before the export, these pre-messenger RNAs undergo a processing step called splicing, which produces functional mRNA. This means that the introns of the immature mRNA are removed and the exons are joined together and form a chain (Arvind, 2010).

The ß - globulin chain constitutes three exons and two introns. During the splicing mechanism, the three exons are linked with each other and the introns are rejected. Beta-thalassemia is caused by nucleotide mutations and alterations, which lead to an aberrant splicing process. The splicing mechanism involves an RNA sequence and a spliceosome that consists of small ribonucleoprotein particles, called SNURPs. For the splicing mechanism to occur, signals are needed, which label the introns. A 5’ splice site, a branch site and a 3’splice site are the signals, which exist within an intron and mark it. In the normal splicing mechanism, the SNURPs form base pairs at the 5' end of the intron and the branch site A. Consequently, the RNA forms a loop structure and other SNURPs are able to tether. The intron is, then, cleaved at 5’ end and 3’ end. The 5’ end of the intron forms a lariat with the branch

site A and the two ends of the exons are ligated. At least, the SNURPs disassemble and the splicing process is finished (Arvind, 2010).

How is ß-thalassemia caused?

There are different forms of abnormal splicing mechanisms, which result in ß-thalassemia. Abnormal splicing can result in the retention of a part of the intron. This condition is caused by a G to A mutation in intron one, which activates a cryptic splice site AG, similar to the actual 3' splice site. The cryptic splice site is mistaken by the spliceosome as the actual 3' splice site. As a consequence, the intron is cleaved at 5' splice site and the cryptic 3' splice site, which is located within the intron and not at the actual 3' end. A part of intron, then, becomes integrated into the exon chain, which are used for the production of the protein.

Abnormal splicing can also result in the exclusion of an exon. In this case, a G to A mutation destroys the 5' splice site of the second intron, which means that it is not detected as 5' splice site anymore. Consequently, the 3' splice sites of intron two and intron one compete with each other for the only available 5' splice site of the first intron. Since the signal strength of 3' splice of intron two is stronger than the signal strength of 3' splice of intron one, the probability of splicing at 3' splice site of intron two and the 5' splice site of intron one is much higher. If this case occurs, it results in the exclusion of exon two.

Abnormal splicing, furthermore, can result in unsplicing of an intron. Therefore, 25 nucleotides are deleted in intron one, which destroys the signal sequence of the 3' splice site. Due to this, the 5' splice signals of intron one and intron two compete with each other for the 3' splice site of intron two. Since 5' splice site of intron two is stronger that the 5' splice site of intron one, the splicing occurs at 5' splice site and 3' splice site of intron two. This process results in not cutting out intron one, which means, intron one serves as template for the protein production.

The insertion of a new exon is also a form of abnormal splicing. In this case, a cryptic 5' splice site (GT) is created due to a G to T mutation in intron two. Additionally, a cryptic 3' splice site, AG, exists along the upstream of intron two. Because of the two new created cryptic splice sites, a new exon is produced. Consequently, the splicing occurs at the actual 5' splice site, cryptic 3' splice site, cryptic 5'splice site and actual 3' splice site of intron two. Splicing also occurs at the 5' and 3' splice sites of intron one as normal. This newly created spurious exon now

becomes part of the exon chain, which is used as template for the production of proteins (Arvind, 2010).

Consequences of ß-thalassemia

The genetic defect affects erythropoiesis (process, which produces red blood cells) through the entire process of maturation and proliferation of all erythrocytes because alpha2 chains accumulate in the erythroid precursors. This leads to early death of the red blood cells and patients suffering from this disorder need blood transfusions. Without a transfusion, the patient experience an enormous expansion of the erythroid, which is followed by severe hemolytic anemia. Hemolytic anemia is a form of anemia caused by the abnormal degradation of red blood cells (RBCs) in the blood vessels (intravascular hemolysis) or anywhere in the body (extravascular). It has many possible effects, ranging from relatively harmless to life threatening. Symptoms of hemolytic anemia are equal to other forms of anemia, but in addition, the degradation of red cells may cause jaundice and raises the risk of some long-term consequences, such as gallstones and pulmonary hypertension (U.S. National Library of Medicine, 2014).

Gene therapy

Beta thalassemia is cause by a malfunction of the gene that makes ß-globin. ß - globin is part of the hemoglobin molecule and decisive for red blood cell to transport oxygen throughout the body. People who suffer from ß-thalassemia can not produce enough or any ß-globin. Currently, patients have to receive frequent and life-long blood transfusions. So, red blood cells are delivered from donors to the patient to even out the anemia. However, blood transfusion have a severe side effect. Iron that is contained in the blood is able to build up in the body. The accumulation of iron can cause severe damage to organs, abdominal pain, weakness, fatigue and joint pain. In order to get rid of the excess iron, patients, who receive blood transfusions, must take additional medicines.

In order to circumvent those side effects, scientists are looking for an investigational gene therapy approach as treatment for ß-thalassemia. Therefore, a functional copy of the ß-globulin gene is put into a lentiviral vector. This vector inherits some parts of the human immunodeficiency virus (HIV) to allow the entering of the vector into cells. The vector has been changed so it cannot grow or cause HIV

infection. Then, blood stem cells are collected from the patient. In the laboratory, the lentiviral vector is used to insert the functional copy of the ß-globin gene into the DNA of the blood stem cells of the patient. To remove stem cells in the body, which carry the non-functional gene, chemotherapy medicine is used. Now there is room for the new modified blood stem cells and they are able to multiply when returned into the patient's body. These new blood cells, now, should become a permanent source for new blood cells, which then also contain a functional copy for the ß-globin protein. This investigational gene therapy approach may potentially help patients to produce healthy red blood cells and decrease or eliminate the need for blood transfusions (Bluebird Bio, 2014)

Marfan Syndrome

The Marfan syndrome results from a genetic disorder affecting the connective tissue of the whole body. The connective tissue plays a crucial role holding together cells, organs and other tissue. Moreover, it is decisive for a proper body growth and development. The syndrome is caused by a mutation in the gene that is responsible for the production of fibrillin-1, an important protein that constitutes the connective tissue. Due to different mutations within this gene, located on chromosome 15q21.1, the production of another protein, called transforming growth factor beta (TGF-ß) increases. The rise of the growth factor affects the connective tissue of the whole body and causes many different problems. As the connective tissue is present throughout the whole body, various difficulties may occur at different parts of the body (Marfan Foundation, 2014; Pees, n.d.)

The mutation of the gene affects, among others, the binding of calcium at fibrillin. This condition leads to alterations of the elasticity of the fibers, which further affects the extracellular matrix and apoptosis (cell death) of the vascular smooth muscle cells (VSMCs). Furthermore, cystic structure may accumulate and, as mentioned before, the release of TGF-ß increases (Marfan Foundation, 2014; Pees, n.d.).

The life expectancy of patients suffering from Marfan syndrome without any treatment is substantially lowered. A therapy that perfectly fits to one patient raises not only the living standard but also life expectancy to up to 60 years. However, it

always depends on the degree of severity and, so, to give a forecast is not easy and cannot be generalized (Gumpert, 2002).

Consequences of Marfan Syndrome

Since the connective tissue is decisive for the entire body, many parts of the body are affected, including the cardiovascular system, the skeletal system, the ocular system, lungs and other body systems. However, Marfan syndrome does not affect the intelligence of patients (Marfan Foundation, 2014).

The alterations within the cardiovascular system are mainly present as an enlarged or bulging aorta, which is the main vessel that transports blood from the heart to the entire body. Another big problem can occur, if the layers of the aorta are separated and further cause it to tear. The affection of the skeletal system is more visible. Patients suffering from Marfan syndrome often have long arms and legs and generally have a tall and thin body. Another conspicuous characteristic of people with Marfan syndrome are long, thin fingers and very flexible joints. Moreover, distortions of the spine or chest may occur.

Marfan syndrome patients often suffer from serious nearsightedness and sometimes the lenses are dislocated within the eye. Another severe problem could be the peeling of the retina. Furthermore, the chance to suffer from glaucoma or cataracts is much higher with patients. Another organ that is affected by the syndrome are the lungs. They tend to suddenly collapse. Asthma is also quite common with Marfan syndrome patients (Marfan Foundation, 2014).

Therapy

First of all, after diagnosis of Marfan syndrome, an adaption of the lifestyle of the patient must be carried out. Whiplash or collisions with others, which may occur during playing football or basketball, should be avoided because the chance of an aorta injury is at high risk. The same danger can be found with other disciplines, in which elevated blood pressure is caused, like bodybuilding.

The affected vessels need to be checked regularly. An annual control is sufficient, if the diameter of the aorta is below 40 mm, however, if the risk of a rupture of the aorta is given, an operation cannot be inevitable. If the operation is well planned, the lethality rate is around 1%, however, the mortality rate increases to 27 %, if an emergency operation is necessary. During the procedure, the extended

part of the aorta is removed and the aortic valve, as well as the ascending aorta, are substituted by a valve-carrying prosthesis. Nonetheless, after an operation, the patient must take medication, like coagulation inhibitors, for the entire life. To avoid lifelong medication, a valve-sparing technique is used, in which only the aorta is replaced by a prosthesis. This may, however, result in degeneration of the received valve, which further requires another operation. If the valve of the left part of the heart is weak, the valve should also be replaced by a prosthesis to avoid further expansion of the ascending aorta. In this case, lifelong medication is not avoidable (Gumpert, 2002).

If the case of lens dislocation occurs, normally, the lens is removed and substituted by an appropriately refractive spectacles design. If the lens can be maintained, the excessive lens fibers are eliminated. The short sightedness can be treated, in some cases, with glasses or contact lens.

To avoid a deterioration of the skeletal system, children suffering from Marfan syndrome have to wear a corset, so that the deformation of the spinal column during the growth stage can be limited. In the worst case, a straightening of the spine in an operative way must be considered, to prevent lung problems and backache. Funnel chests and pigeon chests do not have to be treated because only in rare cases, patients suffer from compression of the lungs, the heart or the aorta (Gumpert, 2002).

Prophylaxis

As prophylaxis, patients suffering from Marfan syndrome are treated with beta-blocker, to stop an expansion of the aorta and blood pressure reducing medication. Especially for young patients without an operation, the treatment is effective. Elderly or patients after an operation do not show any effectiveness.

The prescription of antibiotics as prophylaxis is appropriate in advance of an operative procedure or in the case of severe injuries, in order to avoid inflammation of the heart’s inner membrane (Gumpert, 2002).

References

Abovich, N. and Rosbash, M., 1997. Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. Cell, [pdf] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC363691/pdf/molcellb00060-0009.pdf> [Accessed 03 October 2014].

Allmang, C., Petfalski, E., Podtelejnikov, A., Mann, M., Tollervey, D. and Mitchell, P., 1999. The yeast exosome and human PM–Scl are related complexes of 3′ → 5′ exonucleases. Genes and Development, [online] Available at: <http://genesdev.cshlp.org/content/13/16/2148.long> [Accessed 08 December 2014].

Antequera, F. and Bird, A., 1993. Number of CpG islands and genes in human and mouse. Proceedings of the National Academy of Sciences, [online] Available at: <http://www.pnas.org/content/90/24/11995.long> [Accessed 25 September 2014].

Arvind, B. and Ram, S. V., 1987. International Review of Cytology. [e-book] Orlando: Academic Press. Available at: Google Books <https://books.google.at/> [Accessed 08 January 2015].

Arvind, P., 2010. Bioinformatics Animation - Beta Thalassemia. [video online] Available at: <https://www.youtube.com/watch?v=Qw27dySE7wA> [Accessed 31 December 2014].

Asano, K., Clayton, J., Shalev, A. and Hinnebusch, A. G., 2000. A multifactor complex of eukaryotic initiation factors, eIF1, eIF2, eIF3, eIF5 and initiator tRNA (Met) is an important translation initiation intermediate *in vitro*. Genes and Development, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC316978/> [Accessed 31 October 2014].

Ballut, L., Marchadier, B., Baguet, A., Tomasetto, C., Séraphin, B. and Le Hir, H., 2005. The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. Nature Structural & Molecular Biology, [online] Available at: <http://www.nature.com/nsmb/journal/v12/n10/full/nsmb990.html> [Accessed 24 October 2014].

Baserga, S. J. and Benz, E. J., 1992. β-Globin Nonsense Mutation: Deficient Accumulation of mRNA Occurs Despite Normal Cytoplasmic Stability. Proceedings of the National Academy of Sciences, [online] Available at: <http://www.jstor.org/discover/2359008?sid=21105529909973&uid=3737528&uid=4&uid=2> [Accessed 14 October 2014].

Beelman, C. A., Stevens, A., Caponigro, G., LaGrandeur, T. E.,Hatfield, L.,Fortner, D. M. and Parker, R., 1996. An essential component of the decapping enzyme required for normal rates of mRNA turnover. Nature, [pdf] Available at: <http://www.nature.com/nature/journal/v382/n6592/pdf/382642a0.pdf> [Accessed 06 November 2014].

Beisang, D., Rattenbacher, B., Vlasova-St Louis, I. A. and Bohjanen, P. R., 2012. Regulation of CUG-binding protein 1 (CUGBP1) binding to target transcripts upon T cell activation. The Journal of Biological Chemistry, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3256867/> [Accessed 07 November 2014].

Berg, J. M., Tymoczko, J. L. and Stryer, L., 2002. Biochemistry. [e-book] New York: W H Freeman. Available at: Bookshelf NCBI <http://www.ncbi.nlm.nih.gov/books/NBK22356/> [Accessed 12.01.2015].

Bicknell, A. A., Cenik, C., Chua, H. N., Roth, F. P. and Moore, M. J., 2012. Introns in UTRs: why we should stop ignoring them. BioEssays, [pdf] Available at: <http://onlinelibrary.wiley.com/doi/10.1002/bies.201200073/pdf> [Accessed 20 October 2014].

Black, D. L., Chabot, B. and Steitz, J. A., 1985. U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. Cell, [online] 42(3), 737-50. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2996775> [Accessed 29 September 2014].

Bluebird Bio, 2014. Gene Therapy for Beta Thalassemia. [video online] Available at: <https://www.youtube.com/watch?v=MUi94tvclxY> [Accessed 01 January 2015].

Bowman, G.D., O'Donnell, M. and Kuriyan, J., 2004. Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. Nature, [online] 429(6993), 724-30. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15201901> [Accessed 29 October 2014].

Boyes, J. and Bird, A., 1991. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. Cell, [online] 64(6), 1123-34. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2004419> [Accessed 27 September 2014].

Bregman, A., Avraham-Kelber, M., Barkai, O., Duek, L., Gutman, A., and Choder, M., 2011. Promoter elements regulate cytoplasmic mRNA decay. Cell, [pdf] Available at: <http://ac.els-cdn.com/S0092867411015030/1-s2.0-S0092867411015030-main.pdf?_tid=ccad7c86-98b8-11e4-a10b-00000aacb35f&acdnat=1420888200_f1aed056a895dc507b16e207f5291263> [Accessed 14 September 2014].

Brouwer, R., Allmang, C., Raijmakers, R., Can Aarssen, Y., Egberts, W. V., Petfalski, E., Van Venrooij, W. J., Tollervey, D. and Pruijn, G. J. M., 2001. Three Novel Components of the Human Exosome. The Journal of Biological Chemistry, [online] Available at: <http://www.jbc.org/content/276/9/6177.full> [Accessed 09 December 2014].

Buratowski, S., Hahn, S., Guarente, L., and Sharp, P. A., 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. Cell, [online] 56(4), 549-61. Abstract only. Available at: <http://www.cell.com/cell/abstract/0092-8674(89)90578-3> [Accessed 12 September 2014].

Burke, T. W., and Kadonaga, J.T., 1996. Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes and Development, [online] Available at: <http://genesdev.cshlp.org/content/10/6/711.long> [Accessed 12 September 2014].

Chen, C. Y.,Shyu, A. B., 2003. Rapid deadenylation triggered by a nonsense codon precedes decay of the RNA body in a mammalian cytoplasmic nonsense-mediated decay pathway. Molecular and Cellular Biology, [pdf] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC162215/pdf/0337.pdf> [Accessed 14 October 2014].

Cheng, H., Dufu, K., Lee, C. S., Hsu, J. L., Dias, A. and Reed, R., 2006. Human mRNA export Machinery recruited to the 5' end of mRNA. Cell, [pdf] Available at: <http://ac.els-cdn.com/S0092867406015261/1-s2.0-S0092867406015261-main.pdf?_tid=7b91fc44-9a1d-11e4-8b82-00000aacb362&acdnat=1421041394_0cebf35f9fea8b7af108ffd3c8b4c7d9> [Accessed 06 October 2014].

Cheung, A. C., and Cramer, P., 2011. Structural basis of RNA polymerase backtracking, arrest and reactivation. Nature, [online] 471(7337), 249-53. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21346759> [Accessed 15 September 2014].

Chiu, S. Y.,Serin, G.,Ohara, O. and Maquat, L. E., 2003. Characterization of human Smg5/7a: a protein with similarities to Caenorhabditis elegans SMG5 and SMG7 that functions in the dephosphorylation of Upf1. RNA, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1370372/> [Accessed 08 November 2014].

Chuang, T. W., Chang, W. L., Lee, K. M. and Tarn, W. Y., 2013. The RNA-binding protein Y14 inhibits mRNA decapping and modulates processing body formation. Molecular Cell, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3530774/> [Accessed 11 November 2014].

College of DuPage, 2004. Prokaryotic and Eukaryotic Cells. [online] Available at: <http://www.cod.edu/people/faculty/fancher/ProkEuk.htm> [Accessed 08 January 2015].

Cornway, L., and Wickens, M., 1985. A sequence downstream of AAUAAA is required for formation of SV40 late mRNA 3' termini in frog oocytes. Proceedings of the National Academy of Sciences, [pdf] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC397911/pdf/pnas00352-0008.pdf> [Accessed 26 October 2014].

Cramer, P., Cáceres, J. F., Cazalla, D., Kadener, S., Muro, A. F., Baralle, F. E. and Kornblihtt, A. R., 1999. Coupling of transcription with alternative splicing: RNA Pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. Molecular Cell, [online] Available at: <http://www.sciencedirect.com/science/article/pii/S109727650080372X> [Accessed 21 October 2014].

Das, R., Yu. J,, Zhang, Z., Gygi, M. P., Krainer, A. R., Gygi, S. P. and Reed, R., 2007. SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. Molecular Cell, [online] Available at: <http://www.sciencedirect.com/science/article/pii/S1097276507003607> [Accessed 06 October 2014].

De la Mata, M., Alonso, C. R., Kadener, S., Fededa, J. P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D. and Kornblihtt, A. R., 2003. A slow RNA polymerase II affects alternative splicing *in vivo* . Molecular Cell, [online] 12(2), 525-32. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14536091?dopt=Abstract&holding=f1000,f1000m> [Accessed 20 October 2014].

Dostie, J. and Dreyfuss, G., 2002. Translation Is Required to Remove Y14 from mRNAs in the Cytoplasm. Current Biology, [pdf] Available at: <http://www.med.upenn.edu/dreyfusslab/publications/2002DostieCurrBio.pdf> [Accessed 26 October 2014].

Dye, M. J. and Proudfoot, N. J., 1999. Terminal exon definition occurs cotranscriptionally and promotes transcription termination of RNA polymerase II. Molecular Cell, [online] Available at: <http://www.sciencedirect.com/science/article/pii/S1097276500804645> [Accessed 27 October 2014].

Fairbrother, W. G., Yeh, R. F., Sharp, P. A. and Burge, C. B., 2002. Predictive identification of exonic splicing enhancer in human genes. Science, [online] 297(5583), 1007-13. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12114529> [Accessed 22 October 2014].

Fong, N. and Bentley, D. L., 2001. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. Genes and

Development, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC312735/> [Accessed 15 September 2014].

Fortes, P., Inada, T., Preiss, T., Hentze, M. W., Mattaj, I. W. and Sachs, A. B., 2000. The yeast nuclear cap binding complex can interact with translation factor eIF4G and mediate translation initiation. Molecular Cell, [pdf] Available at: <http://ac.els-cdn.com/S1097276505000031/1-s2.0-S1097276505000031-main.pdf?_tid=c81f1a08-904d-11e4-a942-00000aacb361&acdnat=1419962627_68ca39d99e16ac5e913da63ae02c9a2a> [Accessed 12 October 2014].

Fox, C. A., Sheets, M. D. and Wickens, M. P., 1989. Poly(A) addition during maturation of frog oocytes: distinct nuclear and cytoplasmic activities and regulation by the sequence UUUUUAU. Genes and Development, [pdf] Available at: <http://genesdev.cshlp.org/content/3/12b/2151.full.pdf> [Accessed 26 October 2014].

Freistroffer, D. V., Kwiatkowski, M., Buckingham, R. H. and Ehrenberg, M., 2000. The accuracy of codon recognition by polypeptide release factors. Proceedings of the National Academy of Sciences, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC15751/> [Accessed 10 November 2014].

Fukuhara, N., Ebert, J., Unterholzner, L., Lindner, D., Izaurralde, E. and Conti, E., 2005. SMG7 is a 14-3-3-like adaptor in the nonsense-mediated mRNA decay pathway. Molecular Cell, [e-journal] 17 (4), 537-547. Abstract only. Available through: <http://www.ncbi.nlm.nih.gov/pubmed/15721257> [Accessed 08 November 2014].

Garg, P., Stith, C. M., Sabouri, N., Johansson, E. and Burgers, P. M., 2004. Idling by DNA polymerase δ maintains a ligatable nick during lagging-strand DNA replication. Genes and Development, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC528896/> [Accessed 31 October 2014].

Ge, H. and Manley, J. L., 1990. A protein, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA *in vitro*. Cell, [online] 4(90), 90236-8. Abstract only. Available at: <http://www.cell.com/abstract/0092-8674%2890%2990236-8> [Accessed 20 October 2014].

Gehring, N. H., Kunz, J. B., Neu-Yilik, G., Breit, S., Viegas,, M. H., Hentze, M. W. and Kulozik, A. E., 2005. Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. Molecular Cell, [online] Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16209946> [Accessed 18 October 2014].

Gehring, N. H., Neu-Yilik, G., Schell, T., Hentze, M. W. and Kulozik, A. E., 2003. Y14 and hUpf3b form an NMD-activating complex. Molecular Cell, [online] Available at: <http://www.sciencedirect.com/science/article/pii/S1097276503001424> [Accessed 26 October 2014].

Gerber, J. K., Gögel, E., Berger, C., Wallisch, M., Müller, F., Grummt, I., and Grummt, F., 1997. Termination of mammalian rDNA replication: polar arrest of replication fork movement by transcription termination factor TTF-I. Cell, [online] Available at: <http://www.sciencedirect.com/science/article/pii/S0092867400805152> [Accessed 09 September 2014].

Giege, R., Sissler, M. and Florentz, C., 1998. Universal rules and idiosyncratic features in tRNA identity. Nucleic Acids Research, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC147952/> [Accessed 05 December 2014].

Gil, A. and Proudfoot, N., 1987. Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit b-globulin mRNA 3'end formation. Cell, [online] 49(3), 399-406. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/3568131> [Accessed 27 October 2014].

Goodrich, J. A. and Tjian, R., 1994. Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. Cell, [online] 77(1), 145-56. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8156590> [Accessed 16 September 2014].

Grabowski, P. J., Seiler, S. P., and Sharp, P. A., 1985. A multicomponent complex is involved in the splicing of messenger RNA precursors. Cell, [pdf] Available at: <http://www.pitt.edu/~biohome/Dept/pdf/650.pdf> [Accessed 29 September 2014].

Graveley, B. R., 2005. Mutually exclusive splicing of the insect Dscam pre-mRNA directed by competing intronic RNA secondary structures. Cell, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2366815/> [Accessed 28 September 2014].

Grundner-Culemann, E., Martin, G. W., Tujebajeva, R., Harney, J. W. and Berry, M. J., 2001. Interplay between termination and translation machinery in eukaryotic selenoprotein synthesis. Journal of Molecular Biology, [online] Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11453681> [Accessed 20 October 2014].

Gumpert, N., 2002. Marfan-Syndrom. [online] (19 January 2015) Available at: <http://www.dr-gumpert.de/html/marfan_syndrom.html> [Accessed 22 January 2015].

Hendrix, D. A., Hong, J. W., Zeitlinger, J., Rokhsar, D. S. and Levine, M. S., 2008. Promoter elements associated with RNA polymerase II stalling in the Drosophila embryo. Proceedings of the National Academy of Sciences, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2396556/> [Accessed 20 September 2014].

Hoshino, S., Imai, M., Kobayashi, T., Uchida, N., Katada, T., 1999. The eukaryotic polypeptide chain releasing factor (eRF3/GSPT) carrying the translation termination signal to the 3'-Poly(A) tail of mRNA. Direct association of erf3/GSPT with polyadenylate-binding protein. The Journal of Biological Chemistry, [online] Available at: <http://www.jbc.org/content/274/24/16677.long> [Accessed 14 October 2014].

Hsu, I. W., Hsu, M., Li, C., Chuang, T. W., Lin, R. I. and Tarn, W. Y., 2005. Phosphorylation of Y14 modulates its interaction with proteins involved in mRNA metabolism and influences its methylation. The Journal of Biological Chemistry, [online] Available at: <http://www.jbc.org/content/280/41/34507.long> [Accessed 11 November 2014].

Huntzinger, E., Kashima, I., Fauser, M., Saulière, J. and Izaurralde, E., 2008. SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. RNA, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2590965/> [Accessed 08 November 2014].

Ibba, M. and Söll, D., 2000. Aminoacyl-tRNA synthesis. Annual Review of Biochemistry, [online] 69, 617-50. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10966471> [Accessed 05 December 2014].

Ishigaki, Y., Li, X., Serin, G. and Maquat, L. E., 2001. Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. Cell, [online] Available at: <http://www.cell.com/cell/abstract/S0092-8674(01)00475-5> [Accessed 12 October 2014].

Isken, O., Kim, Y. K., Hosoda, N., Mayeur, G. L., Hershey, J. W. and Maquat, L. E., 2008. Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay. Cell, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4193665/> [Accessed 08 November 2014].

Ito, K., Ebihara, K., Uni, M. and Nakamura, Y., 1996. Conserved motifs in prokaryotic and eukaryotic polypeptide release factors: tRNA-protein mimicry hypothesis. Proceedings of the National Academy of Sciences, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC39265/> [Accessed 10 November 2014].

Joncourt, R., Eberle, A. B., Rufener, S. C. and Mühlemann, O., 2014. Eukaryotic initiation factor 4G suppresses nonsense-mediated mRNA decay by two genetically separable mechanisms. PLOS ONE, [pdf] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4141738/pdf/pone.0104391.pdf> [Accessed 08 November 2014].

Kadlec, J., Izaurralde, E. and Cusack, S., 2004. The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3. Nature Structural and Molecular Biology, [online] Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15004547> [Accessed 08 November 2014].

Kahvejian, A., Svitkin, Y. V., Sukarieh, R., M'Boutchou, M.-N. and Sonenberg, N., 2005. Mammalian poly(A)-binding is a eukaryotic translation initiation factor, which acts via multiple mechanism. Genes and Development, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC540229/> [Accessed 04 November 2014].

Kaminski, A., Howell, M. T. and Jackson, R. J., 1990. Initiation of encephalomyocarditis virus RNA translation: the authentic initiation site isnot selected by a scanning mechanism. The EMBO Jounal, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC552132/> [Accessed 31 October 2014].

Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G. and Ohno, S., 2006. Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. Genes & Development, [pdf] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1361706/pdf/0355.pdf> [Accessed 08 November 2014].

Kim, J. L., Morgenstern, K. A., Griffith, J. P., Dwyer, M. D., Thomson, J. A., Murcko, M. A., Lin, C. and Caron, P. R., 1998. Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. Structure, [pdf] Available at: <http://ac.els-cdn.com/S0969212698000100/1-s2.0-S0969212698000100-main.pdf?_tid=393d7892-9049-11e4-a7f7-00000aacb360&acdnat=1419960669_cd7ea5e37ddcb512af594bf0a54d6f59> [Accessed 24 October 2014].

Kim, V. N., Yong, J., Kataoka, N., Abel, L., Diem, M. D. and Dreyfuss, G., 2001. The Y14 protein communicates to the cytoplasm the position of exon–exon junctions. The EMBO Journal,

[online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC125236/> [Accessed 16 October 2014].

Kim, M., Krogan, N. J., Vasiljeva, L., Rando, O. J., Nedea, E., Greenblatt, J. F. and Buratowski, S., 2004. The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature, [online] 433(7026), 661. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15565157> [Accessed 22 October 2014].

Kim, Y. K., Furic, L., DesGroseillers, L. and Maquat, L. E., 2005. Mammalian Staufen1 Recruits Upf1 to Specific mRNA 3 UTRs so as to Elicit mRNA Decay. Cell, [pdf] Available at: <http://ac.els-cdn.com/S0092867404011572/1-s2.0-S0092867404011572-main.pdf?_tid=886cd27e-8f99-11e4-85d1-00000aacb362&acdnat=1419885211_f3212f9e900294c5536c25bec415e344> [Accessed 18 October 2014].

Klaholz, B. P., Myasnikov, A. G., Van Heel, M., 2004. Visualization of release factor 3 on the ribosome during termination of protein synthesis. Nature, [online] 427(6977), 862-5. Abstract only. Available through: < http://www.ncbi.nlm.nih.gov/pubmed/14985767> [Accessed 10 November 2014].

Klungland, A. and Lindahl, T., 1997. Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). The EMBO Jounal, [online] Available at: <http://emboj.embopress.org/content/16/11/3341.long> [Accessed 08 September 2014].

Kohtz, J. D., Jamison, S. F., Will, C. L., Zuo, P., Lührmann, R., Garcia-Blanco, M. A. and Manley, J. L., 1994. Protein-proteininteractions and 5' splice-site recognition in mammalian mRNA precursors. Nature, [online] 368(6467), 119-124. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8139654> [Accessed 04 October 2014].

Konarska, M. M., and Sharp, P. A., 1987. Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. Cell, [online] 8674(87), 90614-3. Abstract only. Available at: <http://www.cell.com/cell/abstract/0092-8674(87)90614-3> [Accessed 04 October 2014].

Körner, C. G., Wormington, M., Muckenthaler, M., Schneider, S., Dehlin, E. and Wahle, E., 1998. The deadenylating nuclease (DAN) is involved in poly(A) tail removal during the meiotic maturation of Xenopus oocytes. The EMBO Journal, [online] Available at: <http://emboj.embopress.org/content/17/18/5427.long> [Accessed 10 December 2014].

Kostrewa, K., Zeller, M. E., Armache, K. J., Seiz, M., Leike, K., Thomm, M., and Cramer, P., 2009. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. Nature, [online]

62(7271), 323-30. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19820686> [Accessed 13 September 2014].

Krainer, A. R., and Maniatis, T., 1985. Multiple components including the small nuclear ribonucleoproteins U1 and U2 are required for pre-mRNA splicing *in vitro*. Molecular and Cellular Biology, [pdf] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC363691/pdf/molcellb00060-0009.pdf> [Accessed 29 September 2014].

Krainer, A. R., Conway, G. C., Kozak, D., 1990. The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites. Cell, [online] 62(1), 35-42.. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2364434> [Accessed 20 October 2014].

Krainer, A. R., Maniatis, T., Ruskin, B. and Green, M. R., 1984. Normal and mutant human b-globulin pre-mRNAs are accurately and efficiently spliced *in vitro*. Cell, [online] 36(4), 993-1005. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/6323033> [Accessed 28 September 2014].

Krebs, J. E., Goldstein, E. S. and Kilpatrick, S. T., 2013. Lewin's Genes XI. 11th ed. Burlington: Jones & Barlett Learning.

LaGrandeur, T. E. and Parker R., 1998. Isolation and characterization of Dcp1p, the yeast mRNA decapping enzyme. The EMBO Journal, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1170496/ > [Accessed 11 November 2014].

Lans, H., Marteijn, J. A., Schumacher, B., Hieijmakers, J. H. J., Lansen, G. and Vermeulen, W., 2010. Involvement of global genome repair, transcription coupled repair and chromosome remodeling in UV damage response changes during development. PLoS Genetics, [online] Available at: <http://www.plosgenetics.org/article/info%3Adoi%2F10.1371%2Fjournal.pgen.1000941> [Accessed 20 September 2014].

Lau, A. Y., Glassner, B.J., Samson, L.D., and Ellenberger, T., 2000. Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAg. Proceedings of the National Academy of Sciences, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC17617/> [Accessed 08 September 2014].

Le Hir, H., Izaurralde, E., Maquat, L. E. and Moore, M. J., 2000. The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. The EMBO Journal, [online] Available at: <http://emboj.embopress.org/content/19/24/6860.long> [Accessed 12 October 2014].

Le Hir, H., Gatfield, D., Izaurralde, E. and Moore, M. J., 2001. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. The EMBO Journal, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC125616/> [Accessed 08 November 2014].

Le, S., Sternglanz, R. and Greider, C.W., 2000. Identification of two RNA-binding proteins associated with human telomerase RNA. Molecular and Cellular Biology, [online] Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10712515> [Accessed 18 October 2014].

Lejeune, F., Li, X. and Maquat, E., 2003. Nonsense-Mediated mRNA Decay in Mammalian Cells Involves Decapping, Deadenylating, and Exonucleolytic Activities. Molecular Cell, [online] Available at: <http://www.sciencedirect.com/science/article/pii/S1097276503003496> [Accessed 08 December 2014].

Lewis, B. P., Green, R. E. and Brenner, S. E., 2003. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proceedings of the National Academy of Sciences, [online] Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12502788> [Accessed 20 October 2014].

Li, Q., Imataka, H., Morino, S., Rogers, G. W., Richter-Cook, N. J., Merrick, W. C. and Sonenberg, N., 1999. Eukaryotic Translation Initiation Factor 4AIII (eIF4AIII) Is Functionally Distinct from eIF4AI and eIF4AII. Molecular and Cellular Biology, [pdf] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC84727/pdf/mb007336.pdf> [Accessed 24 October 2014].

Liu, X., Bushnell, D.A., Wang, D., Calero, G., and Kornberg, R.D., 2011. Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. Science, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2813267/> [Accessed 13 September 2014].

Luo, W., Johnson, A. W., and Bentley, D. L., 2006. The role of Rat 1 in coupling mRNA 3' end processing to transcription termination: implication for a unified allosteric-torpedo model. Genes and Development, [online] Available at: <http://genesdev.cshlp.org/content/20/8/954.long> [Accessed 29 October 2014].

Luse, D. S., Spangler, L. C. and Ujvari, A., 2011. Efficient and rapid nucleosome traversal by RNA polymerase II depends on a combination of transcription elongation factors. The Journal of Biological Chemistry, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3057798/> [Accessed 21 September 2014].

Lykke-Andersen, J., 2002. Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay. Molecular and Cellular Biology, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC134073/> [Accessed 08 December 2014].

Lykke-Andersen, J., Shu, M. D. and Steitz, J. A., 2000. Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. Cell, [online] Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11163187> [Accessed 26 October 2014].

Lykke-Andersen, J., Shu, M. D. and Steitz, J. A., 2001. Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1. Science, [online] Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11546874> [Accessed 16 October 2014].

Ma, Y., Pannicke, U., Schwarz, K., and Lieber, M. R., 2002. Hairpin opening and overhang processing by an Artemis/DNA-Dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. Cell, [online] Available at: <http://www.sciencedirect.com/science/article/pii/S0092867402006712> [Accessed 08 September 2014].

Mandal, S. S., Chu, C., Wada, T., Handa, H., Shatkin, A. J. and Reinberg, D., 2004. Functional interaction of RNA-capping enzyme with factors that positively and negatively regulated promoter escape by RNA polymerase II. Proceedings of the National Academy of Sciences, [online] Available at: <http://www.pnas.org/content/101/20/7572.full> [Accessed 27 September 2014].

Maquat, L. E. and Li, X., 2001. Mammalian heat shock p70 and histone H4 transcripts, which derive from naturally intronless genes, are immune to nonsense-mediated decay. RNA, [online] Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11333024> [Accessed 24 October 2014].

Marfan Foundation, 2014. What is Marfan Syndrome? [online] Available at: <http://www.marfan.org/about/marfan> [Accessed 07 January 2015].

Marión, R. M., Fortes, P., Beloso, A., Dotti, C. and Ortín, J., 1999. A human sequence homologue of Staufen is an RNA-binding protein that is associated with polysomes and localizes to the rough endoplasmic reticulum. Molecular and Cellular Biology, [online] Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10022908> [Accessed 18 October 2014].

Martinez, E., Chiang, C. M., Ge, H., and Roeder, R. G., 1994. TATA-binding protein-associated factors in TFIID function through the initiator to direct basal transcription from a TATA-less class II promoter. The EMBO Journal, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC395202/> [Accessed 11 September 2014].

Matsumoto, Y. and Kim, K., 1995. Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair. Science, [online] 269(5224), 699-702. Abstract only. Available through: <http://www.ncbi.nlm.nih.gov/pubmed/7624801> [Accessed 08 September 2014].

McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Program, A. E., Shuman, S. and Bentley, D. L., 1997. 5'-capping enzymes are targented to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. Genes and Development, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC316822/> [Accessed 28 September 2014].

McElhinny, S.A., Gordenin, D.A., Stith, C. M., Burger, P.M. and Kunkel, T.A., 2008. Division of labor at the eukaryotic replication fork. Molecular Cell, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2654179/> [Accessed 29 October 2014].

McKendrick, L., Thompson, E., Ferreira, J., Morley, S. J. and Lewis, J. D., 2001. Interaction of eukaryotic translation initiation factor 4G with the nuclear cap-binding complex provides a link between nuclear and cytoplasmic functions of the m(7) guanosine cap. Molecular Cell Biology, [online] Available at : <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC86986/> [Accessed 12 October 2014].

Mehlin, H., Daneholt, B. and Skoglund, U., 1992. Translocation of a specific premessenger ribonucleoprotein particle through the nuclear pore studied with electron microscope tomography. Cell, [pdf] Available at: <http://www.cell.com/cell/pdf/0092-8674%2892%2990224-Z.pdf> [Accessed 12 October 2014].

Mendell, J. T., ap Rhys, C.M. and Dietz, H. C., 2002. Separable roles for rent1/hUpf1 in altered splicing and decay of nonsense transcripts. Science, [online] Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12228722> [Accessed 08 November 2014].

Mendell, J. T., Sharifi, N.A., Meyers, J. L., Martinez-Murillo, F. and Dietz, H.C., 2004. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nature Genetics, [online] Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15448691> [Accessed 19 October 2014].

Mikuni, O., Ito, K., Moffat, J., Matsumura, K., McCaughan, K., Nobukuni, T., Tate, W. and Nakamura, Y., 1994. Identification of the prfC gene, which encodes peptide-chain release factor 3 of E. coli. Proceedings of the National Academy of Sciences, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC44084/> [Accessed 10 November 2014].

Milman, G., Goldstein, J., Scolnick, E. and Caskey, T., 1969. Peptide chain termination. 3. Stimulation of *in vitro* termination. Proceedings of the National Academy of Sciences, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC534020/> [Accessed 15 November 2014].

Moazed, D. and Noller, H. F., 1989. Intermediate states in the movement of tRNA in the ribosome. Nature, [pdf] Available at: <http://www.nature.com/nature/journal/v342/n6246/pdf/342142a0.pdf> [Accessed 04 November 2014].

Montanuy, I., Torremocha, R., Hernandez-Munain, C. and Suné, C., 2008. Promoter influences transcription elongation: TATA-BOX element mediates the assembly of processive transcription complexes responsive to cyclin-dependent kinase 9. The Journal of Biological Chemistry, [online] Available at: <http://www.jbc.org/content/283/12/7368.long> [Accessed 21 September 2014].

Moraes, K. C., Wilusz, C. J. and Wilusz, J., 2006. CUG-BP binds to RNA substrates and recruits PARN deadenylase. RNA, [pdf] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1464848/pdf/1084.pdf> [Accessed 07 November 2014].

Mueller-Storm, H. P., Sogo, J. M. and Schaffner, W., 1989. An enhancer stimulates transcription in trans when attached to the promoter via a protein bridge. Cell, [online] 59(2), 405. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2548735> [Accessed 22 September 2014].

Nissen, P., Kjeldgaard, M., Thirup, S., Polekhina, G., Reshetnikova, L. Clark, B. F. and Nyborg, J., 1995. Crystal structure of the ternary complex of Phe-tRNAPhe, EF-Tu, and a GTP analog. Science, [online] 270(5241), 1464-72. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7491491> [Accessed 06 November 2014].

Nuss, D. L., Furuichi, Y., Koch, G. and Shatkin, A. J., 1975. Detection in HeLa cell extracts of a 7-methyl guanosine specific enzyme activity that cleaves m7GpppNm. Cell, [pdf] Available at:<http://www.cell.com/cell/pdf/0092-8674(75)90069-0.pdf> [Accessed 06 November 2014].

Ohnishi, T., Yamashita, A., Kashima, I., Schell, T., Anders, K. R., Grimson, A., Hachiya, T., Hentze, M. W., Anderson, P. and Ohno, S., 2003. Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7. Molecular Cell, [online] Available at:<http://www.ncbi.nlm.nih.gov/pubmed/14636577> [Accessed 08 November 2014].

Palacios, I. M., Gatfield, D., St Johnston, D. and Izaurralde, E., 2004. An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. Nature, [online] Available at:<http://www.ncbi.nlm.nih.gov/pubmed/14973490l> [Accessed 24 October 2014].

Parker, R., 2013. mRNA degradation - Roy Parker (Boulder/HHMI). [video online] Available at: <https://www.youtube.com/watch?v=d7nmojex01c> [Accessed 08 November 2014].

Pee, C., n.d. Das Marfan-Syndrom Medikamentöse Therapie einer genetischen Erkrankung. [online] Available at: <http://www.meduniwien.ac.at/hp/fileadmin/kinderherzzentrum/pdf/Das_Marfan_Syndrom.pdf> [Accessed 07 January 2015].

Perona, J. J. and Hou, Y-M., 2007. Indirect read-out of tRNA for aminoacylation. Biochemistry, [online] 46(37), 10419-32. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17718520> [Accessed 05 December 2014].

Pestova, T. V., Hellen, C. U. and Shatsky, I. N., 1996. Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. Molecular and Cellular Biology, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC231689/> [Accessed 31 October 2014].

Pestova, T. V., Shatsky, I. N., Fletcher, S. P., Jackson, R. J. and Hellen, C. U., 1998. A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. Genes and Development, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC316404/> [Accessed 31 October 2014].

Pestova, T. V., Lomakin, I. B., Lee, J. H., Choi, S. K., Dever, T. E. and Hellen, C. U., 2000. The joining of ribosomal subunits in eukaryotes require eIF5B. Nature, [pdf] Available at: <http://tetrad.ucsf.edu/sites/tetrad.ucsf.edu/files/general/Wk8Pestova2-29-12Disc.pdf> [Accessed 04 November 2014].

Pestova, T. V. and Kolupaeva, V. G., 2002. The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. Genes and Development, [online] Available at: <http://genesdev.cshlp.org/content/16/22/2906.long> [Accessed 04 November 2014].

Raijmakers, R., Noordman, Y. E., Van Venrooij, W. J. and Pruijn, G. J. M., 2002. Protein-protein interactions of hcsl4p with other human exosome subunits. The Journal of Molecular Biology, [online]315(4), 809-18. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11812149> [Accessed 10 December 2014].

Ramsden, D. A., and Gellert, M., 1998. Ku protein stimulates DNA end joining by mammalian DNA ligases: a direct role for Ku in repair of DNA double-strand breaks. The EMBO Journal, [online] Available at: <http://emboj.embopress.org/content/17/2/609.long> [Accessed 08 September 2014].

Reardon, J.T., and Sancar, A., 2003. Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by the human excision nuclease. Genes and Development, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC218148/> [Accessed 08 September 2014].

Reed, R. and Maniatis, T., 1985. Intron sequences involved in lariat formation during pre-mRNA splicing. Cell, [online] 41(1), 95-105. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/3888410> [Accessed 27 September 2014].

Rehwinkel, J., Letunic, I., Raes, J., Bork, P. and Izaurralde, E., 2005. Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. RNA, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1370837/> [Accessed 08 November 2014].

Robberson, B. L. and Berget, S. M., 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. Molecular and Cellular Biology, [pdf] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC360715/pdf/molcellb00037-0104.pdf> [Accessed 04 October 2014].

Scolnick, E., Tompkins, R., Caskey, T. and Nirenberg, M., 1968. Release factors differing in specificity for terminator codons. Proceedings of the National Academy of Sciences, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC225226/> [Accessed 15 November 2014].

Sharma, S., Falick, A. M. and Black, D. L., 2005. Polypyrimidine tract binding protein blocks the 5' splice site-dependent assembly of U2AF and the prespliceosome E complex. Molecular Cell, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1635971/> [Accessed 22 October 2014].

Shaw, G. and Kamen, R., 1986. A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell, [pdf] Available at:<http://www.cell.com/cell/pdf/0092-8674%2886%2990341-7.pdf> [Accessed 24 October 2014].

She, M., Decker, C. J., Svergun, D. I., Round, A., Chen, N., Muhlrad, D., Parker, R. and Song, H., 2008. Structural basis of dcp2 recognition and activation by dcp1. Molecular Cell, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2323275/> [Accessed 11 November 2014].

Sheth, U. and Parker, R., 2006. Targeting of aberrant mRNAs to cytoplasmic processing bodies. Cell, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1858659/> [Accessed 11 November 2014].

Shibuya, T., Tange, T. Ø., Stroupe, M. E. and Moore, M. J., 2006. Mutational analysis of human eIF4AIII identifies regions necessary for exon junction complex formation and nonsense-

mediated mRNA decay. RNA, [online] Available at: <http://rnajournal.cshlp.org/content/12/3/360.long> [Accessed 24 October 2014].

Shiomi, Y., Usukura, J., Masamura, Y., Takeyasu, K., Nakayama, Y., Obuse, C., Yoshikawa, H. and Tsurimoto, T., 2000. ATP- dependent structural change of the eukaryotic clamp-loader protein, replication factor C. Proceedings of the National Academy of Sciences, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC18882/> [Accessed 30 October 2014].

Silva, A. L. and Romão, L., 2009. The mammalian nonsense-mediated mRNA decay pathway: to decay or not to decay! Which players make the decision? FEBS Letters, [online] Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19162024> [Accessed 24 October 2014].

Singer, V. L., Wobbe, C.R., and Struhl, K., 1990. A wide variety of DNA sequences can functionally replace a yeast TATA element for transcriptional activation. Genes and Development, [online] Available at: <http://genesdev.cshlp.org/content/4/4/636.long> [Accessed 10 September 2014].

Smale, S. T., and Baltimore, D., 1989. The “initiator” as a transcription control element. Cell, [online] 57(1),103-13. Abstract only. Available through: <http://www.ncbi.nlm.nih.gov/pubmed/2467742> [Accessed 11 September 2014].

Song, H., Mugnier, P., Das, A. K., Webb, H. M., Evans, D. R., Tuite, M. F., Hemmings, B. A. and Barford, D., 2000. The crystal structure of human eukaryotic release factor eRF1- mechanism of stop codon recognition and peptidyl-tRNA hydrolysis. Cell, [online] Available at: <http://www.sciencedirect.com/science/article/pii/S0092867400806674> [Accessed 17 November 2014].

Stark, H., Rodnina, M. V., Wieden H. J., van Heel, M. and Wintermeyer, W., 2000. Large-scale movement of elongation factor G and extensive conformational change of the ribosome during translocation. Cell, [online] Available at: <http://www.sciencedirect.com/science/article/pii/S0092867400806662> [Accessed 06 November 2014].

Sun, X., Perlick, H. A., Dietz, H. C. and Maquat, L. E., 1998. A mutated human homologue to yeast Upf1 protein has a dominant-negative effect on the decay of nonsense-containing mRNAs in mammalian cells. Proceedings of the National Academy of Sciences, [pdf] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC21452/pdf/pq010009.pdf> [Accessed 08 November 2014].

Takagaki, Y., Ryner, L. C., Manley, J. L., 1988. Separation and characterization of a poly(A) polymerase and a cleavage/specificity factor required for pre-mRNA polyadenylation. Cell, [online] 52(5), 731-42. Abstract only. Available through: <http://www.ncbi.nlm.nih.gov/pubmed/2830992> [Accessed 27 October 2014].

Tamburini, B. A., and Tyler, J.K., 2005. Localized histone acetylation and deacetylation triggered by the homologous recombination patheways of double-strand DNA repair. Molecular and Cellular Biology, [online] Available at: <http://www.sciencedirect.com/science/article/pii/S0092867400805152> [Accessed 09 September 2014].

Tange, T. Ø., Nott, A. and Moore, M. J., 2004. The ever-increasing complexities of the exon junction complex. Current Opinion in Cell Biology, [pdf] Available at: <http://www.bio.brandeis.edu/publications/tange_nott_moore.pdf> [Accessed 24 October 2014].

Tarun, S. Z. and Sachs, A. B., 1996. Association of the yeast poly (A) tail binding protein with translation initiation factor eIF-4G. The EMBO Journal, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC452544/> [Accessed 04 November 2014].

Thermann, R., Neu-Yilik, G., Deters, A., Frede, U., Wehr, K., Hagemeier, C., Hentze, M. W. and Kulozik, A. E., 1998. Binary specification of nonsense codons by splicing and cytoplasmic translation. The EMBO Journal, [pdf] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1170685/pdf/003484.pdf> [Accessed 12 October 2014].

Tran, H., Schilling, M., Wirbelauer, C., Hess, D. and Nagamine, Y., 2004. Facilitation of mRNA deadenylation and decay by the exosome-bound, DExH protein RHAU. Molecular Cell, [pdf] Available at: <http://ac.els-cdn.com/S1097276503004817/1-s2.0-S1097276503004817-main.pdf?_tid=7d8612ac-91bb-11e4-8671-00000aacb360&acdnat=1420119697_b4a5d644282face5c46c8f49ae33d8cb> [Accessed 07 November 2014].

Tseng, C. K. and Cheng, S. C., 2008. Both catalytic steps of nuclear pre-mRNA splicing are reversible. Science, [pdf] Available at: <http://aao.sinica.edu.tw/download/publication_list/en/159.pdf> [Accessed 04 October 2014].

U.S. National Library of Medicine, 2014. Beta thalassemia. [online] Available at: <http://ghr.nlm.nih.gov/condition/beta-thalassemia> [Accessed 03 January 2015].

Unterholzner, L. and Izaurralde, E., 2004. SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. Molecular Cell, [pdf] Available at: <http://ac.els-cdn.com/S1097276504006392/1-s2.0-S1097276504006392-main.pdf?_tid=98f274be-931d-11e4-bcac-00000aacb35e&acdnat=1420271785_43e77a31f889ccda8f5dbe23939e537b> [Accessed 08 November 2014].

Verrijzer, C. P., Chen, J. L., Yokomori, K., Tjian, R., 1995. Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. Cell, [pdf] Available at: <http://ac.els-cdn.com/S0092867405800169/1-s2.0-S0092867405800169-main.pdf?_tid=d51d4f56-98b6-11e4-80ba-00000aacb360&acdnat=1420887355_68291f84e7a929dbf9ee20d1e7d1286b> [Accessed 12 September 2014].

Visa, N., Izaurralde, E., Ferreira, J., Daneholt, B. and Mattaj, I. W., 1996. A nuclear cap-binding complex binds Balbiani ring pre-mRNA cotranscriptionally and accompanies the ribonucleoprotein particle during nuclear export. The Journal of Cell Biology, [online] Available at:<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2120770/> [Accessed 12 October 2014].

Waga, S., Masuda, T., Takisawa, H. and Sugino, A., 2001. NA polymerase ε is required for coordinated and efficient chromosomal DNA replication in Xenopus egg extracts. Proceedings of the National Academy of Sciences, [online] Available at: <http://www.pnas.org/content/98/9/4978.full> [Accessed 31 October 2014].

Wang, Z., Day, N., Trifillis, P. and Kiledjian, M., 1999. An mRNA stability complex functions with poly(A)-binding protein to stabilize mRNA *in vitro*. Molecular and Cellular Biology, [pdf] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC84253/pdf/mb004552.pdf> [Accessed 06 November 2014].

Wang, Z. and Kiledjian, M., 2001. Functional link between the mammalian exosome and mRNA decapping. Cell, [pdf] Available at: <http://ac.els-cdn.com/S009286740100592X/1-s2.0-S009286740100592X-main.pdf?_tid=5c9b4014-91b9-11e4-8a96-00000aab0f01&acdnat=1420118783_9fe5c044033f1b1f966cd3b01475fb3d> [Accessed 06 November 2014].

Wang, J., Chang, Y. F., Hamilton, J. I. and Wilkinson, M. F., 2002. Nonsense-associated altered splicing: a frame-dependent response distinct from nonsense-mediated decay. Molecular Cell, [online] Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12419238> [Accessed 20 October 2014].

Wang, Z., Rolish, M. E., Yeo, G., Tung, V., Mawson, M. and Burge, C.B., 2004. Systematic identification and analysis of exonic splicing silencers. Cell, [pdf] Available at: <http://web.mit.edu/merolish/Public/wang-ess.pdf> [Accessed 22 October 2014]

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P. and Burge, C. B., 2008. Alternative isoform regulation in human tissue transcriptomes. Nature, [online] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2593745/> [Accessed 21 October 2014].

Weiss, S. L. and Sunde, R. A., 1998. Cis-acting elements are required for selenium regulation of glutathione peroxidase-1 mRNA levels. RNA, [pdf] Available at: <http://rnajournal.cshlp.org/content/4/7/816.full.pdf> [Accessed 24 October 2014].

Wickham, L., Duchaîne, T., Luo, M., Nabi, I. R. and DesGroseillers, L., 1999. Mammalian staufen is a double-stranded-RNA- and tubulin-binding protein which localizes to the rough endoplasmic reticulum. Molecular and Cellular Biology, [online] Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10022909> [Accessed 18 October 2014].

Wollerton, M.C., Gooding, C., Wagner, E. J., Garcia-Blanco, M. A. and Smith, C. W., 2004. Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. Nature Genetics, [pdf] Available at:<http://ac.els-cdn.com/S1097276503005021/1-s2.0-S1097276503005021-main.pdf?_tid=151f3456-903e-11e4-967b-00000aab0f01&acdnat=1419955884_7853ff944dbdf7762112c7725e02030b> [Accessed 20 October 2014].

Wolner, B., Van Komen, S., Sung, P., and Peterson, C. L., 2003. Recruitment oft he recombinational repair machinery to a DNA double-strand break in yeast. Molecular Cell, [online] Available at: <http://www.sciencedirect.com/science/article/pii/S1097276503002429> [Accessed 08 September 2014].

Wood, R. D., Mitchell, M., Sgouros, J. and Lindahl, T., 2001. Human DNA Repair Genes. Science, [online] Available at: <http://www.sciencemag.org/content/291/5507/1284.full> [Accessed 08 January 2015].

Wu, J. and Parkhurst, L. J., 2001. DNA bends in TATA-binding protein-TATA complexes in solution are DNA sequence-dependent. The Journal of Biological Chemistry, [online] Available at: <http://www.jbc.org/content/276/18/14614.long> [Accessed 12 September 2014].

Yamashita, A., Ohnishi, T., Kashima, I., Taya, Y. and Ohno, S., 2001. Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay. Genes and Development, [pdf] Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC312771/pdf/X2.pdf> [Accessed 08 November 2014].

Zemach, A., McDaniel, I. E., Silvia, P. and Zilberman, D., 2010. Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science, [online] 328(5980), 916-9. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20395474> [Accessed 22 September 2014].

Zhou, Z., Luo, M. J., Straesser, K., Katahira, J., Hurt, E. and Reed, R., 2000. The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. Nature, [pdf] Available at: <file:///C:/Users/Alexandra/Downloads/00b4952b86b1928a01000000.pdf> [Accessed 13 October 2014].

Zhu, J., Mayeda, A., Krainer, A. R., 2001. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. Molecular Cell, [online] Available at: <http://www.sciencedirect.com/science/article/pii/S1097276501004099> [Accessed 22 October 2014].

Zhuang, Y. and Weiner, A. M., 1986. A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. Cell, [online] 46(6), 827-35. Abstract only. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/3757028> [Accessed 04 October 2014].

Zuo, S., Bermudez, V., Zhang, G., Kelma, Z. and Hurwitz, J., 2000. Structure and activity associated with multiple forms of S. pombe DNA polymerase δ. The Journal of Biological Chemistry, [online] Available at: <http://www.jbc.org/content/275/7/5153.long> [Accessed 31 October 2014].