Nonsense-mediated mRNA decay: defense mechanism against viral infections and its interference with HIV-1

Nonsense-mediated mRNA decay: Abwehrmechanismus vor viralen Infektionen und die Wechselwirkung mit HIV-1

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Lara Andrea Walotka
aus Duisburg

Düsseldorf, Oktober 2021

aus dem Institut für Virologie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Dr. Heiner Schaal

2. Prof. Dr. Michael Feldbrügge

Tag der mündlichen Prüfung: 11.02.2022

"All we have to decide is what to do with the time that is given us."
Gandalf the Grey

Abstract

The cellular nonsense-mediated decay (NMD) pathway detects and degrades mRNAs with unusual structural features, such as long 3' UTRs or overlapping reading frames. Many of which are found in viral transcripts, shedding light on NMD as a possible antiviral defense mechanism. This work focused on the interplay of the human immunodeficiency virus 1 (HIV-1) and the NMD machinery, including the analysis of HIV-1 transcripts and their sensitivity to NMD-mediated degradation, as well as potential infection-based modulations of the NMD machinery.

The results of this work showed that HIV-1 transcripts were indeed targeted for degradation by NMD, specifically the mRNAs of the 2kb class. This NMD-sensitivity was independent of the number of splicing events and thus presumably the deposited exon-junction complexes (EJCs). Furthermore, transcript modifications, like shortening of the 3' UTR, or re-introduction of the rev-responsive element (RRE), did not resolve this sensitivity. However, insertion of the well-described *cis*-acting Rous-sarcoma RNA stability element (RSE) into the 3' UTR stabilized HIV-1 transcripts and confirmed the initially observed NMD-sensitivity of these viral mRNAs.

Furthermore, in addition to the observed NMD-sensitivity of HIV-1 transcripts, the overall translation dependency of NMD-mediated target degradation was investigated, as some viral mRNAs with mutated translational start codon showed sensitivity to NMD-mediated degradation. Using a dual-luciferase NMD reporter, the results demonstrated that the intrinsic strength of the translational start codon not only had an impact on the total amount of mRNA but also affected the NMD-sensitivity of a given reporter transcript. Therefore, NMD, in general, remains a translation-dependent process, but transcripts with NMD-triggering features might escape recognition and degradation by maintaining low levels of translation via inefficient translation initiation. Therefore, in contrast to the *tat* mRNA, some viral transcripts with suboptimal Kozak sequence surroundings, like HIV-1 *rev, vpu,* and *vpr* might escape NMD.

As HIV-1 transcripts were targeted by NMD, subsequent analysis of potential infection-based interference with the cellular NMD machinery illustrated the complexity of simple-looking NMD reporter systems. Based on alternative splicing events, it was shown that the widely used triosephosphate isomerase (TPI) NMD reporter was already a target of the NMD pathway even without an inserted pre-mature stop codon. Therefore, the expression of this reference had to be rendered NMD independent, which was achieved by modulations of the splice site strengths, generating a reliable system with a single splice isoform. However, neither transient transfection experiments with the modified TPI NMD reporters nor lentiviral-transduced cells expressing the dual-luciferase NMD reporter system revealed an impact of HIV-1 infection on the NMD activity.

Even though it is known that viral infections induce an interferon (IFN) response, and the results of this work suggested that recombinant IFNs can inhibit the cellular NMD machinery, HIV-1 infection neither resulted in enhanced expression of interferon-stimulated genes (ISGs) nor the accumulation of cellular NMD targets, again suggesting no impact of HIV-1 infection on the cellular NMD activity.

Finally, the expression of endogenous genes with alternative-NMD-targeted splice isoforms was investigated during HIV-1 infection. Even though an impaired degradation of NMD-sensitive splice isoforms was observed for SRS7, this could not be confirmed for hnRNPDL, indicating rather a transcript-specific interference than a global inhibition of the cellular NMD machinery by HIV-1 infection.

Zusammenfassung

Der zelluläre nonsense-mediated decay (NMD) pathway erkennt mRNAs mit ungewöhnlichen strukturellen Merkmalen, wie z. B. langen 3'-UTRs oder überlappende Leserahmen und baut diese ab. Viele dieser Merkmale finden sich in viralen Transkripten, was NMD im Licht eines möglichen antiviralen Abwehrmechanismus erscheinen lässt. Diese Arbeit konzentrierte sich auf das Zusammenspiel zwischen dem Humanen Immundefizienz-Virus 1 (HIV-1) und der NMD-Maschinerie. Dies beinhaltete die Analyse von HIV-1 Transkripten und ihrer Empfindlichkeit gegenüber dem NMD-vermittelten Abbau, sowie die Analyse der potenziellen infektionsbedingten Modulationen der NMD-Maschinerie.

Die Ergebnisse dieser Arbeit zeigten, dass HIV-1 Transkripte tatsächlich durch NMD abgebaut werden, insbesondere die mRNAs der 2kb Klasse. Diese NMD-Empfindlichkeit war unabhängig von der Anzahl der Spleißereignisse und damit vermutlich von der Anzahl der Exon-Junction-Komplexe. Weiterhin konnten Transkriptmodifikationen, wie die Verkürzung des 3'-UTRs oder die Wiedereinführung des *rev*-responsiven Elements (RRE) diese Empfindlichkeit nicht aufheben. Nur das Einfügen des gut bekannten *cis*-wirksamen Rous-Sarkoma RNA-Stabilitätselements (RSE) in die Sequenz des 3'-UTR, konnte die HIV-1-Transkripte stabilisieren, was die ursprünglich beobachtete NMD-Empfindlichkeit dieser viralen mRNAs bestätigte.

Zusätzlich zur beobachteten NMD-Empfindlichkeit von HIV-1 Transkripten wurde die allgemeine Translationsabhängigkeit des NMD-vermittelten mRNA Abbaus untersucht, da einige virale mRNAs mit mutiertem translationalem Startcodon eine Empfindlichkeit gegenüber dem NMD-vermittelten Abbau zeigten. Mit Hilfe eines dualen Luziferase-NMD-Reporters konnte gezeigt werden, dass die intrinsische Stärke des Startcodons nicht nur einen Einfluss auf die Gesamtmenge der mRNA hat, sondern auch die NMD-Empfindlichkeit eines bestimmten Reportertranskripts beeinflusst. Allgemein kann NMD daher weiter als ein von der Translation abhängiger Prozess angesehen werden, aber Transkripte mit NMD-auslösenden Merkmalen könnten der Erkennung und dem Abbau entgehen, indem sie durch ineffiziente Translationsinitiierung niedrige Translationsraten aufrechterhalten. Im Gegensatz zur *tat*-mRNA könnten daher einige virale Transkripte mit suboptimaler Kozak-Sequenzumgebung, wie HIV-1 *rev*, *vpu* und *vpr*, NMD entgehen.

Auf Grundlage der NMD-Sensitivität mancher HIV-1 Transkripte, wurde die mögliche infektionsbedingte Modellierung der NMD Aktivität untersucht, was die die Komplexität von einfach aussehenden NMD-Reportersystemen verdeutlichte. Auf Grundlage alternativer Spleißereignisse konnte gezeigt werden, dass der weit verbreitete Triosephosphat-Isomerase (TPI) NMD Reporter auch bereits ohne vorzeitiges Stoppcodon ein Ziel des NMD-pathways ist. Um die Expression dieser Referenz unabhängig von NMD zu machen, wurde ein zuverlässiges System mit einer einzigen Spleißisoform, durch Modulationen der Spleißstellenstärken geschaffen. Allerdings zeigten weder transiente Transfektionsexperimente mit den modifizierten TPI NMD Reportern, noch lentiviral-transduzierte Zellen, welche das Dual-Luziferase NMD Reportersystem exprimierten, einen Einfluss der HIV-1 Infektion auf die NMD Aktivität.

Obwohl bekannt ist, dass Virusinfektionen eine Interferonantwort auslösen können, und die Ergebnisse dieser Arbeit darauf hindeuten, dass rekombinante Interferone (IFNs) die Aktivität der zellulären NMD-Maschinerie hemmen können, führte eine HIV-1 Infektion weder zu einer verstärkten Expression von Interferon-stimulierten Genen (ISGs) noch zu einer Anhäufung von

zellulären NMD-Zielen, was wiederum keinen Einfluss einer HIV-1-Infektion auf die zelluläre NMD-Aktivität vermuten lässt.

Abschließend wurde die Expression von zellulären Genen mit alternativen Spleißisoformen, von denen manche durch NMD abgebaut werden, während einer HIV-1 Infektion untersucht. Obwohl für SRS7 ein beeinträchtigter Abbau von NMD-empfindlichen Spleißisoformen beobachtet wurde, konnte dies für hnRNPDL nicht bestätigt werden, was eher auf einen transkriptspezifischen Einfluss, als auf eine globale Beeinflussung der zellulären NMD-Maschinerie durch eine HIV-1 Infektion hinweist.

Table of contents

1 Introduction

In this work, the interaction of the nonsense-mediated decay (NMD) pathway on the cellular translation apparatus, as well on HIV-1 transcripts was investigated. Therefore, the eukaryotic gene expression, the NMD pathway, and the HIV-1 replication cycle are introduced in this chapter.

1.1 Eukaryotic gene expression

Eukaryotic gene expression mediates the translation and usage of genetic information, which is encoded on our genomic DNA. This includes not only the protein-coding sequences but also non-coding regulatory elements. The correct regulation of gene expression in all eukaryotic cells of a given organism is essential and the disruption might lead to disease (Levine and Tjian 2003; Herz, Hu, and Shilatifard 2014).

1.1.1 Transcription

Gene expression of protein-coding genes, as well as many non-coding genes, starts with the transcription of a genomic DNA sequence into the complementary RNA transcript by the RNA polymerase II (Pol II). Transcription initiates at a defined position at the 5' end of a given gene, which is called the transcriptional start site (TSS). The TSS is embedded in a core promoter region, which serves as binding platform for the transcription machinery and is sufficient to promote transcription initiation, however only to a basal level (Kadonaga 2012). The transcription machinery consists not only of Pol II but also of its associated general transcription factors (GTFs) (Hampsey 1998).

Basal usage of a TSS can be further enhanced by the presence of regulatory elements called enhancers, which can be bound by transcription factors as well as transcription co-factors (Banerji, Rusconi, and Schaffner 1981; Shlyueva, Stampfel, and Stark 2014). In general, there a three different types of core promoters, which were identified by mapping endogenous transcription initiation sites (Lenhard, Sandelin, and Carninci 2012; Haberle and Stark 2018). First, the best-described core promoters with sharp initiation patterns, TATA-box and initiator (Inr) motifs, which are mostly active in terminally differentiated cells in adult tissues. Second, core promoters of broadly and ubiquitous expressed housekeeping genes, which are associated with dispersed transcription initiation, and third, developmentally regulated core promoters of key transcription factors which are involved in embryo patterning and morphogenesis.

For example, the conserved TATA-box motif (TATAWAAR, W = A or T, R = A or G) is found approximately 30 nucleotides upstream of a single dominant TSS (Ponjavic et al. 2006) and gets recognized by the TATA-box-binding-protein (TBP). TATA-driven promoters are strongly associated with genes that show a tissue-specific expression (Carninci et al. 2006). Binding of TBP, functioning as GTF, to the TATA-box results in recruitment of RNA polymerase II and

mediates assembly of the pre-initiation complex, DNA duplex melting and pre-messenger RNA (pre-mRNA) synthesis (Burley and Roeder 1996; Patikoglou et al. 1999; Louder et al. 2016).

1.1.2 Splicing

Almost twenty years ago, the human genome project revealed that the human genome encodes approximately 20,000 protein-coding genes (International Human Genome Sequencing 2004). On average these genes contain seven to eight introns, with a median length of 1.3 kb (Hong, Scofield, and Lynch 2006). Those introns must be removed from the pre-mRNA to form the protein-coding messenger RNA (mRNA), a process that gets mediated by the essential splicing machinery. Furthermore, approximately 95% of all human multiexon genes, not only encode a single isoform but are alternatively spliced to produce multiple protein isoforms (Pan et al. 2008; Wang et al. 2008). Therefore, the usage of alternative splicing might be one reason for the complexity of higher organisms, as it can expand the proteome that can be derived from a limited number of genes (Ule and Blencowe 2019).

In general, initiation of the splicing process relies on the recognition of three different elements, which are schematically illustrated in Figure 1.1. First, the 5' splice site (splice donor (SD), or 5' ss), which is located at the 3' end of the upstream exon and is marked by an 11-nucleotide sequence with an invariant GU dinucleotide at position +1 and +2 (CAG/GURAGUNN (R = purine (A or G), N = any nucleotide (A, G, U or C), "/" = exon-intron border)). Second, the 3' splice site (splice acceptor (SA), or 3' ss), which is located at the 5' end of the downstream exon and is marked by an invariant AG dinucleotide, and third, the branch point sequence (BPS), which is located approximately up to 100-180 nucleotides upstream of the SA and is characterized by the recognition sequence YNCUR$\underline{A}$C (Y = pyrimidine (U or C), R = purine (A or G)). Additionally, in higher eukaryotes, a polypyrimidine tract (PPT) is located between the BPS and the SA (Senapathy, Shapiro, and Harris 1990; Smith, Chu, and Nadal-Ginard 1993; Lopez and Seraphin 1999; Clark and Thanaraj 2002; Cartegni, Chew, and Krainer 2002; Sheth et al. 2006; Wahl, Will, and Lührmann 2009; Will and Lührmann 2011).

Figure 1.1: Schematic architecture of a human splice site. The conserved sequences at the 5' splice site and the 3' splice site, as well as the branch point sequence (BPS) and the polypyrimidine tract (PPT). Grey boxes indicate the two flanking exons, while the black line in between indicates the intron. R = purine (A or G), N = any nucleotide (A, G, U or C), Y = pyrimidine (U or C). Figure adapted from (Will and Lührmann 2011).

From a chemical point of view, the splicing reaction is based on two transesterifications, which are schematically shown in Figure 1.2. At first, the free OH group of the BP adenine attacks the phosphate group of the 5' SD guanine, resulting in a 2'-5' phosphodiester bond at the lariat

intron and 3' exon, as well as a free hydroxyl group at the 5' exon. This is followed by the second transesterification reaction, in which the free hydroxyl group of the 5' exon performs a nucleophilic attack on the guanine phosphate of the 3' SA site, resulting in the linkage of the two exons by a phosphodiester bond and the release and degradation of the intron lariat structure (Grabowski, Padgett, and Sharp 1984; Padgett et al. 1984; Ruskin et al. 1984). These two transesterification reactions are catalyzed by the spliceosome.
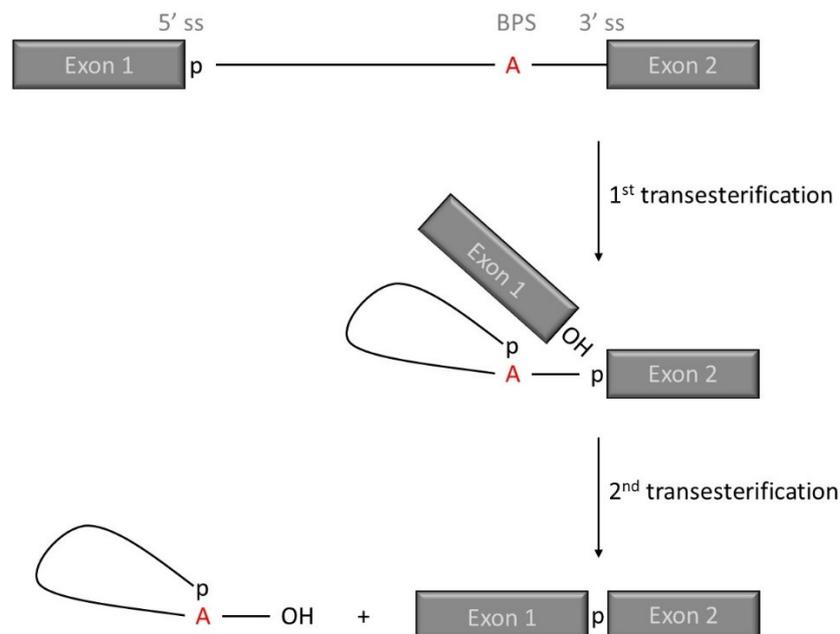
Figure 1.2: Schematic representation of the two-step transesterification reactions of pre-mRNA splicing. In the first step, the OH-group of the branch point adenosine (A) attacks the phosphate group (p) of the 5' splice site. Then the free OH-group of the 5' ss attacks the phosphate group of the 3' ss resulting in exon-exon fusion and lariat intron. Grey boxes represent exons, the black line represents the intron. Figure adapted from (Will and Lührmann 2011).

The major spliceosome consists of five different uridine-rich small nuclear RNAs (U snRNAs; U1, U2, U4/U6, and U5), which are associated with a specific set of additional proteins to form small nuclear ribonucleoprotein (U snRNP) particles (Lerner and Steitz 1979; Bringmann and Lührmann 1986; Fica et al. 2013). The U-rich sequence of the U snRNAs is known as the Sm site and is located at the 3' end of the U1, U2, U4, and U5 snRNAs. Seven homologous Sm proteins, or paralogous like-Sm proteins (LSm) in the case of the U6 snRNA, assemble into a ring around these sequences (Lerner and Steitz 1979; Bringmann and Lührmann 1986; Seraphin 1995; Achsel et al. 1999; Leung, Nagai, and Li 2011).

The first complex formed in the process of spliceosome assembly is called the spliceosomal **E-complex** (early, E) and depends on the correct recognition of the splice sites. It is initiated by recognition and binding of U1 snRNP to the 5' splice site through base pairing between the SD and the 5' end of the U1 snRNA (Zhuang and Weiner 1986; Matera and Wang 2014; Wong, Kinney, and Krainer 2018). In addition to the U1 snRNP : SD interaction, E-complex formation is

mediated by non-snRNP factors such as splicing factor 1 (SF1) binding to the BPS, and the two subunits of U2 snRNP auxiliary factor (U2AF), where U2AF65 interacts with the PPT and U2AF35 with the AG dinucleotide of the 3' SA. A schematic overview of the spliceosome assembly and the splicing reaction is depicted in Figure 1.3.

Following binding of the U1 snRNP to the 5' SD, the DEAD-box helicases Prp5 and Sub2 displace SF1 from the BPS and recruit the U2 snRNP, resulting in the formation of the pre-spliceosome, or **A-complex** (Kistler and Guthrie 2001; Perriman and Ares 2010; Liang and Cheng 2015). Here, U1 and U2 snRNP-associated components interact with each other and thereby pair the respective splice sites and form the exon recognition complex (De Conti, Baralle, and Buratti 2013). Then, upon stable integration of the U2 snRNP into the pre-spliceosome, the U2 snRNA pairs with the branch point region of the pre-mRNA to form the branch helix (Parker, Siliciano, and Guthrie 1987; Zhuang, Goldstein, and Weiner 1989; Wu and Manley 1989). The branch point adenosine base is flipped out from the branch helix, but its 2'OH group, which acts as a nucleophile for the first reaction, is not yet accessible (Plaschka, Lin, and Nagai 2017).

Now, the pre-B-complex which is also called the inactive spliceosome gets formed by association of the pre-spliceosome with the largest pre-assembled spliceosomal complex, the tri-snRNP U4/U6.U5 (Nguyen et al. 2015; Agafonov et al. 2016; Charenton, Wilkinson, and Nagai 2019). Here, the substrates for the first catalytic reaction (the 5' SD and the BPS), as well as the RNA elements in the U2, U5, and U6 snRNAs form a complex, which remains catalytically inactive as long as the U6 snRNA is chaperoned by the U4 snRNA and the 5' SD is bound by the U1 snRNP (Bringmann et al. 1984; Hashimoto and Steitz 1984). It was shown that the U5 snRNA loop 1 is important for tethering the 5' exon during branching and later for alignment of the 5' and 3' exons during exon ligation (Newman and Norman 1992; Sontheimer and Steitz 1993). Then, the initiator of the spliceosome activation, the DEAD-box helicase Prp28, facilitates the transfer of the 5' ss from the U1 snRNP to the flexible U6 loop within the U6 snRNA (Staley and Guthrie 1999; Charenton, Wilkinson, and Nagai 2019). Formation of this U6/5' ss duplex induces an extensive remodeling of the spliceosome, including the release of the U1 snRNP, and result in the formation of the **B-complex** (Zhan et al. 2018). Furthermore, the RNA helicase Brr2 separates U4 snRNP from U6 snRNP (Raghunathan and Guthrie 1998) what subsequently results in the active site formation of the spliceosome, the B^{act}-complex, by folding and association of the U6 snRNA sequence with part of the U2 snRNA (Hang et al. 2015; Charenton, Wilkinson, and Nagai 2019). This active site snRNP structure remains remarkably static from the formation of the B^{act}-complex until the last step, the intron-lariat spliceosome (ILS) complex (Fica and Nagai 2017).

Even though the B^{act}-complex is the first spliceosome intermediate with a fully formed catalytic core, this core is still catalytically inactive, as the distance to the 2' OH group of the BPS adenosine is still too large for a nucleophilic attack (Haselbach et al. 2018). Hence, the two

reaction partners needed for branching cannot reach each other at the active site and the spliceosome is maintained in its inhibited conformation, which is stabilized by RNA helicase Brr2 (Raghunathan and Guthrie 1998). To disrupt the inhibitory conformation and enable branching, DEAH-box ATPase Prp2 is needed for the remodeling and activation of the spliceosome, so that the BP adenosine can get into the active site and form the catalytically active **B*-complex** (Will and Lührmann 2011).

Now, the activated spliceosome performs the first catalytic step, which is called branching. The 2' hydroxyl group of the BP adenosine, which is now located in the active site, attacks the phosphodiester group at the 5' ss. This attack results in a cleaved 5' exon and a lariat-intron-3' exon intermediate, in which the 5' phosphate of the first intronic guanine nucleotide is linked to the 2' oxygen of the BP adenosine (Galej et al. 2016). Then, upon the performance of the first catalytic step, the spliceosome is called the **C-complex**, which transforms to the C*-complex upon remodeling of the active site and docking of the 3' ss (Fica et al. 2017). In the following second catalytic step, which is called exon ligation, the newly exposed 3' hydroxyl group of the 5' exon attacks the phosphodiester group at the 3' ss, resulting in ligation of the 5' and the 3' exons, the release of the lariat intron and dissociation of the spliceosome (Wahl, Will, and Lührmann 2009). Upon exon ligation, the spliceosome forms the post-catalytic spliceosome, or the P-complex, where the release of the newly ligated exons gets catalyzed by the DEAH-box ATPase Prp22 (Company, Arenas, and Abelson 1991; Schwer 2008).
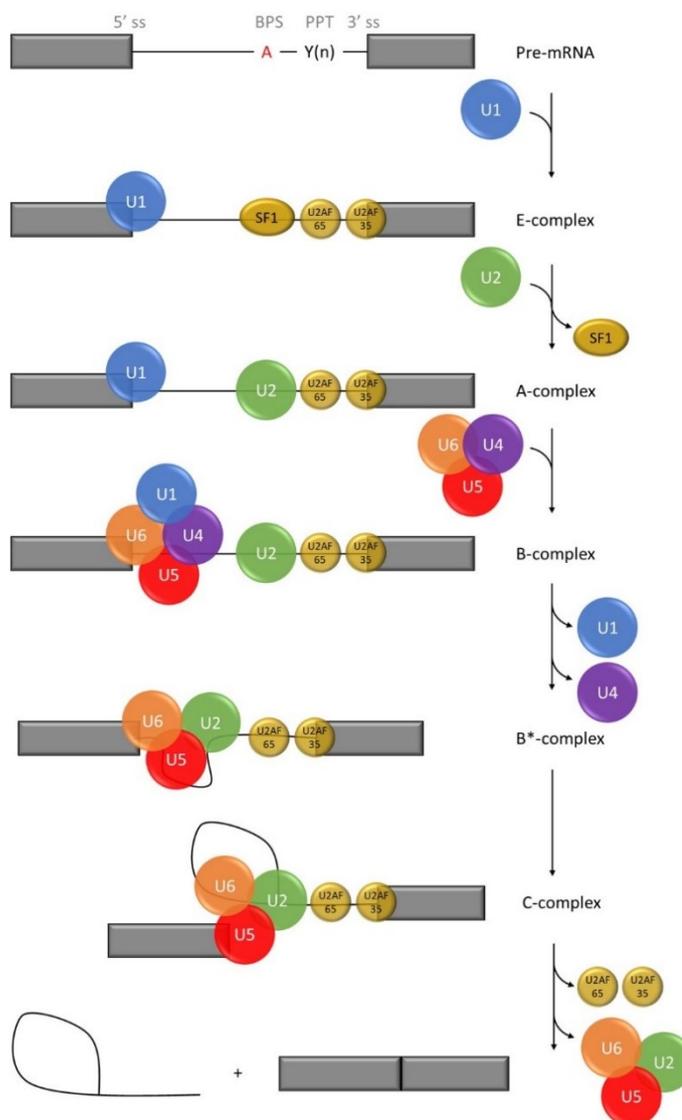
Figure 1.3: Schematic sketch of the spliceosome assembly during the pre-mRNA splicing process. The splicing process gets initiated by E-complex formation which is characterized by the interaction of the U1 snRNP with the 5' ss, SF1 to the BPS, U2AF65 to the PPT, and U2AF35 to the 3' ss. A-complex gets formed by the replacement of SF1 from the BPS by U2 snRNP. Joining of the pre-assembled U4/U6.U5 tri-snRNP results in the formation of B-complex, which gets activated by conformational re-arranges and release of U1 and U4 snRNPs to form B*-complex. The following first transesterification reaction forms the C-complex which consists of an exon intermediate and one lariat-intron-exon intermediate. Now, upon the second transesterification, the exons are ligated, the spliceosomal components dissociate and the lariat-intron gets degraded. The figure is adapted from (Will and Lührmann 2011).

Upon successful intron excision and exon-exon junction, the mRNA is released from the spliceosome and exported into the cytoplasm for translation and protein production. However, one important feature of this new mRNA, which was not mentioned so far is the deposition of exon-junction complexes (EJCs) upstream of the exon-exon junctions. The deposition of EJCs is important for the recognition of pre-mature termination codons (PTCs) within the transcript and subsequent degradation of these faulty transcripts by the nonsense-mediated mRNA decay

(NMD) (Le Hir et al. 2000). During the splicing process, the proteins of the EJC are recruited to the 5' exon by the spliceosome component Cwc22, whose position on the RNA serves as a marker for the EJC deposition which is approximately 20-24 nucleotides upstream of the exon-exon junction (Steckelberg et al. 2012).

1.1.2.1 Alternative splicing

In addition to the described constitutive splicing, a variety of gene expression is obtained by alternative splicing (Black 2003). Different forms of alternative splicing have been described, which either include skipping of an exon, retention of an intron, as well as usage of alternative 5', or 3' splice sites. A schematic overview of the major forms of alternative splicing and the obtained mRNA structures is shown in Figure 1.4.
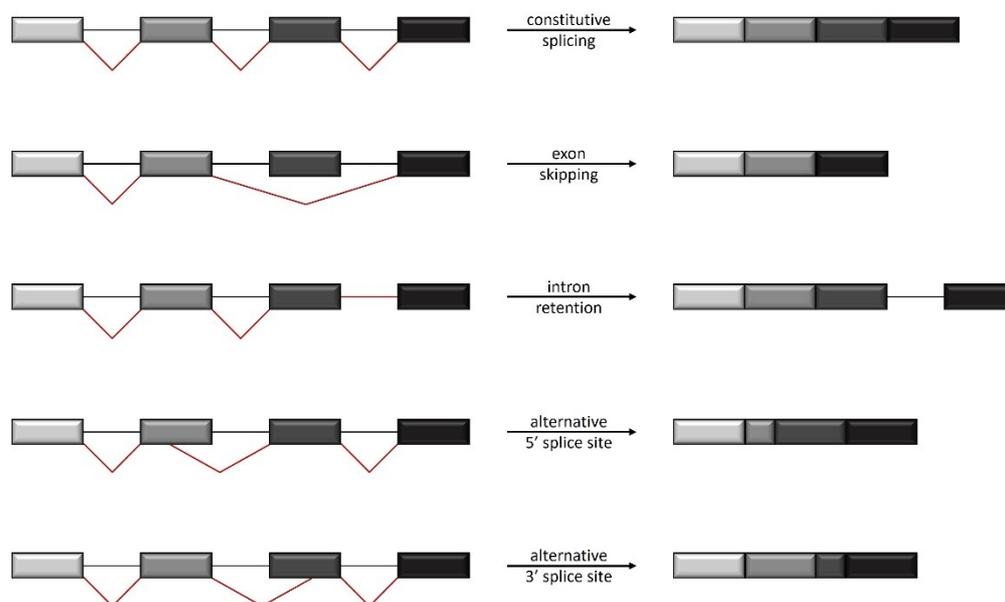
Figure 1.4: Different forms of alternative splicing. While constitutive splicing results in an mRNA that includes all designated exons, alternative splicing allows modification of the final mRNA. Alternative splicing can be achieved by skipping one or more exons, intron retention, or usage of alternative 5' or 3' splice sites.

Importantly, it is to mention, that the majority of alternative splicing events do not result from the inaccuracy of the spliceosome, as these events are regulated with high precision (Fox-Walsh and Hertel 2009). Furthermore, it was estimated that 95% of human genes are alternatively spliced, meaning that they have at least one alternatively used splice site and thereby produce at least two different mRNA isoforms (Pan et al. 2008; Wang et al. 2008). Furthermore, numerous genes produce many mRNA isoforms by selecting different combinations of exons from one primary transcript (Wang et al. 2008).

1.1.2.2 Splice site strength and bioinformatic splicing tools

As mentioned before, the initiation of the splicing reaction relies on the correct recognition of the splice sites (cf. chapter 1.1.2). The strength of a given splice site is, however, not only

dependent on the intrinsic strength, but also on the presence of cellular splicing factors. These *trans*-acting splicing factors are called splicing regulatory proteins (SRPs) which can bind to *cis*-acting splicing regulatory elements (SREs) – together referred to as 'the splicing code' (Wang and Burge 2008; Barash et al. 2010; Matera and Wang 2014).

Depending on their location within the pre-RNA, SREs can act as splicing enhancers or silencers from both exonic (ESE, ESS) and intronic (ISE, ISS) positions (Ptok et al. 2019). There are two main types of *trans*-acting splicing factors, the serine and arginine-rich splicing factors (SRSF) proteins, and the heterogeneous ribonuclear protein (hnRNP) family (Manley and Krainer 2010; Geuens, Bouhy, and Timmerman 2016). It has been shown that SREs function in a position-dependent manner, meaning that the position of an SRE relative to a splice site determines whether it acts as splicing enhancer or silencer. For example, while SR proteins were shown to enhance splicing from an exonic position, they repress splicing from an intronic position, a situation that is reversed for hnRNP proteins, which activate splicing from an intronic position and repress splicing from an exonic position (Erkelenz, Mueller, et al. 2013).

The strength of a given splice site as well as the potential influence of SRPs which might bind to SREs in the vicinity of a splice site can be estimated and calculated by the joint use of different bioinformatic tools, which are described below. Using the HBond score, the intrinsic strength of the 5' splice site can be calculated, while the 3' splice strength can be calculated by use of the MaxEntScore. The HEXplorer score can ultimately be used to examine the splice site surrounding and to find potential binding sites for SRPs.

1.1.2.2.1 HBond score

As mentioned before, the intrinsic strength of a given 5' ss is defined by the complementarity between the single-stranded 5' end of the spliceosomal U1 snRNA and the 11-nucleotide long sequence of the splice donor (CAG/GURAGUNN (R = purine, N = any nucleotide, "/" = exon-intron border)). The HBond score (HBS) calculates this intrinsic strength based on the predicted hydrogen bond formation of the U1 snRNA and the 11-nucleotides of the splice donor (Kammler et al. 2001; Freund et al. 2003).

The website for HBond score calculation is freely accessible and can be reached through the following link: https://www2.hhu.de/rna/html/hbond_score.php (Freund et al. 2003).

1.1.2.2.2 MaxEntScore

The intrinsic strength of a given 3' splice site is dependent on the pyrimidine content of the PPT, the distance of the BPS and the 3' ss, and the complementarity between the BPS and the U2 snRNA. (Wahl, Will, and Lührmann 2009). This intrinsic strength can be estimated using the 3' maximum entropy score (MaxEntScore). The 3' MaxEntScore is calculated from 23-nucleotide

long sequences of the intron-exon border, starting at the intronic position -20, ending with the invariant AG dinucleotide, up to the exonic position +3 from the downstream exon.

The 3' MaxEntScore website is freely accessible and can be reached through the following link: http://hollywood.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq_acc.html (Yeo and Burge 2004).

1.1.2.2.3 HEXplorer

For analysis of SRE motifs in the vicinity of splice sites, the HEXplorer score is a valuable tool. It can be used to analyze potential SRE motifs, as mutations within SREs can result in aberrant splicing, frameshift mutations, and pre-mature translation termination. In general, the HEXplorer score is based on hexamer weights which were calculated on a RESCUE (Relative Enhancer and Silencer Classification by Unanimous Enrichment) type approach (Fairbrother et al. 2002) from a dataset of 43,464 constitutively spliced canonical annotated human exons. Hexamer distributions in the vicinity of splice sites were calculated using 100 nucleotides of the exonic and intronic sequences, which were located up and downstream of 5' splice sites (10,407 "strong" 5' ss, HBS ≥ 17.0, and 10,359 "weak" 5' ss, HBS ≤ 13.5). For this calculation, the 11-nucleotide long splice donor was excluded.

Taking the intrinsic splice site strength into account, differential hexamer frequencies were calculated for exonic and intronic sequences (Z_{EI}). Hexamers that were enriched in the surrounding of weak splice sites are thought to be more dependent on splicing enhancing SREs and therefore probably contain more exonic splicing enhancers than hexamers which are enriched in the surrounding of strong splice sites (Z_{WS}). Using the position-dependent Z_{EI} score, an average score for each index nucleotide was calculated from the overlapping hexamers (HZ_{EI}) (Erkelenz et al. 2014). The resulting HEXplorer plot illustrates potential splicing enhancing and silencing properties of the splice site neighborhood. Furthermore, the HZ_{EI} score can be used to monitor the impact of a mutation, by calculating the differences between the HZ_{EI} scores of a given wild-type sequence and its mutated version resulting in the ΔHZ_{EI}.

The HEXplorer score website is freely accessible and can be reached through the following link: https://www2.hhu.de/rna/html/hexplorer_score.php (Erkelenz et al. 2014).

1.1.3 Nuclear export

In general, the genetic information in eukaryotic cells is stored as DNA within the nucleus, which is surrounded by the nuclear envelope. As the transcription takes place in the nucleus and the translation in the cytoplasm, the newly synthesized and spliced mRNAs must be transferred to the cytoplasm, what is called nuclear export. This export is mainly mediated through the nuclear pore complexes, which consist of approximately 30 distinct proteins, known as nucleoporins, which perforate the nuclear envelope (Wente and Rout 2010). The nuclear export of mRNAs

requires the evolutionary conserved non-karyoprotein export receptor, the heterodimer NXF1/NXT1, which is also called Tap-p15 (Katahira et al. 1999; Köhler and Hurt 2007). The actual RNA binding, however, gets mediated by another complex, which is called the transcription-export (TREX) complex, which gets recruited in a splicing-dependent manner to the most 5'-located EJC on the mRNA (Masuda et al. 2005; Cheng et al. 2006; Meinel et al. 2013). Upon mRNA binding, the TREX complex directly interacts with the export receptor NXF1/NXT1 heterodimers, thereby functioning as an mRNA adaptor (Köhler and Hurt 2007; Katahira 2012). Finally, the translocation of the bound mRNA from the nucleus to the cytoplasm through the nuclear pore complex is mediated by interaction with the phenylalanine-glycine (FG) peptide repeat sequences of nucleoporins (Strässer, Bassler, and Hurt 2000; Terry and Wente 2009).

1.1.4 Translation

Once the mRNA has been transported from the nucleus to the cytoplasm translation can be initiated. In general, the eukaryotic translation process, hence, the protein synthesis, can be divided into four steps: initiation, elongation, termination, and ribosome recycling.

The small (40S) ribosomal subunit gets pre-loaded with the initiator methionyl tRNA (Met-tRNA$_i$), a process mediated by the GTP-bound form of eukaryotic initiation factor 2 (eIF2) and interaction with eIF1, eIF1A, eIF5, and eIF3, resulting in the formation of the 43S pre-initiation complex (Sokabe, Fraser, and Hershey 2012). The 43S pre-initiation complex then attaches to the 5' end of the mRNA, a process that is mediated by the m^{7}G cap and the cap-binding complex eIF4F. Then, the 43S pre-initiation complex scans the 5' UTR until it encounters a translational start AUG codon in a suitable sequence context through complementary base pairing between the mRNA and the Met-tRNA$_i$ anticodon (Kozak 1979; Sherman, McKnight, and Stewart 1980; Kozak 1984a; Kolitz, Takacs, and Lorsch 2009). However, not every AUG codon is recognized with the same efficiency as the intrinsic strength of a given translational start codon is dependent on the surrounding nucleotides, especially positions -3 and +1. The optimal Kozak sequence in mammals was defined as 5' GCCRCC <u>AUG</u> G 3' with a purine (R = A or G) in position -3 and a guanine in position +1 (Kozak 1986).

The encounter of the 43S complex with the start AUG codon results in the arrest of the scanning pre-initiation complex, dissociation of several eIFs, conformational rearrangements, and joining of the large (60S) ribosomal subunit, which gets catalyzed by eIF5B (Pestova et al. 2000). That interaction results in the formation of the 80S translation initiation complex with the Met-tRNA$_i$ base-paired with its anticodon to the AUG-codon on the mRNA in the P-site of the ribosome, thereby creating a complex which is now competent for the elongation phase of protein synthesis. During translation elongation, the 80S ribosomal complex moves over the mRNA in 5' to 3' direction, three nucleotides at a time, extending the growing protein chain coordinated by multiple eukaryotic elongation factors (eEFs) and aminoacyl-tRNAs (aa-tRNAs). The incoming

aa-tRNAs bind to the aminoacyl site (A-site) of the ribosome by complementary base pairing, thereby allowing the formation of a new peptide-bond. This peptide-bond is formed within the peptidyl site (P-site) of the ribosome, where the elongating peptide chain gets attached to the new aa-tRNA. Once the used tRNAs are deacylated and the ribosomal subunit translocates, the unloaded tRNA is transferred to the exit-site (E-site) of the ribosome before its dissociation. During that elongation process, the translocating ribosome removes proteins that are bound to the mRNA, like for example the exon-junction-complexes which are described in more detail in the following chapter (cf. chapter 1.2.1). A schematic overview of the elongation process is shown in Figure 1.5.
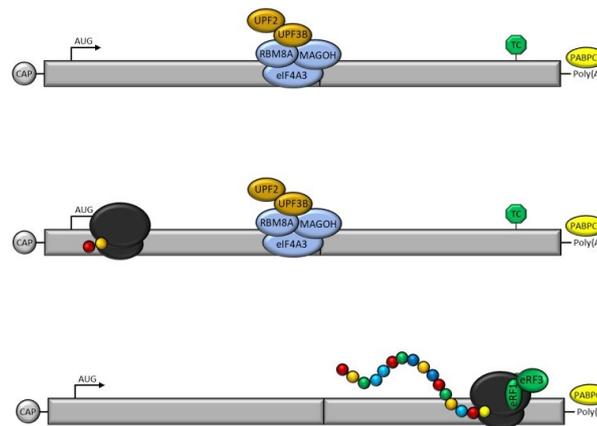
Figure 1.5: Schematic overview of the translation elongation process. An mRNA consisting of two exons (grey boxes) is shown. The 5' end of the mRNA is capped and the 3' end is poly adenylated. The EJC was splicing-dependently deposited on the mRNA, upstream of the exon-exon junction. The core EJC consists of the three proteins eIF4A3, RBM8A, and MAGOH (light blue) and can be extended by UPF3B and UPF2 (brown). The ORF is indicated by the AUG arrow and the green translational stop codon (TC) sign. The poly(A) tail is bound by the cytoplasmic poly-A binding protein 1 (PABPC1, yellow). The small ribosomal subunit begins scanning of the mRNA at the 5' cap structure. The big ribosomal subunit joints at the translational start codon (AUG) and initiates protein synthesis. During translation, the ribosome moves over the mRNA in a 5' to 3' direction and removes proteins that are bound to the mRNA, like the EJC. Upon encounter of the translational stop codon (UAA, UAG, or UGA) the eukaryotic release factors 1 and 3 (eRF) facilitate translation termination, the release of the peptide chain (depicted as colored dots), and ribosome recycling. Efficient translation termination is enhanced by the interaction of eRF3 with PABPC1.

The elongation process is repeated until the translocating ribosome encounters the end of the designated open reading frame (ORF), which is marked by one of three translational stop-codons (TC; UAA, UAG, UGA). All these TCs lack a complementary tRNA anticodon and are instead recognized by the eukaryotic release factor 1 (eRF1), which gets bound by eRF3, a ribosome-dependent GTPase (Frolova et al. 1994). Together this ternary complex (eRF1, eRF3, and GTP) assembles in the A-site of the terminating ribosome, GTP gets hydrolyzed by eRF3, thereby inducing a conformational change in eRF1 which activates the hydrolysis of the nascent polypeptide from the polypeptidyl-tRNA in the ribosomal P-site (Dever and Green 2012; Schuller and Green 2018). The process of translation termination is enhanced by binding of the cytoplasmic poly-A binding protein 1 (PABPC1) to the amino-terminal region of eRF3 (Hoshino et al. 1999; Ivanov et al. 2008). Finally, the release of the polypeptide, eRF3, and GDP is coupled

via eRF1 to the recruitment of the ATP-binding cassette subfamily E member 1 (ABCE1). ABCE1 facilitates the separation of the post-termination 80S ribosome complex into the 40S and 60S ribosomal subunits and efficient ribosome recycling (Pisarev et al. 2010).

1.2 The nonsense-mediated mRNA decay (NMD) pathway

The nonsense-mediated mRNA decay (NMD) pathway was first described as an mRNA surveillance mechanism that recognizes and degrades transcripts containing a pre-mature termination codon (PTC). As these transcripts might otherwise encode non-functional or harmful proteins, NMD emerged as an important regulator of eukaryotic gene expression. For example, it was shown that NMD is essential for cellular homeostasis, cell cycle progression, cellular stress response, cell development, and differentiation, as well as neural activity and immunity (Lykke-Andersen and Jensen 2015; Ottens and Gehring 2016). In general, the NMD pathway is an evolutionarily conserved process, which requires translation (Carter, Li, and Wilkinson 1996; Thermann et al. 1998; Karousis, Nasif, and Mühlemann 2016). Originally discovered in the budding yeast *Saccharomyces cerevisiae*, it was first described in humans in the context of a disease, called β_0-thalassemia (Losson and Lacroute 1979; Leeds et al. 1991; Chang and Kan 1979; Kugler et al. 1995).

In all tested eukaryotes NMD gets mediated by the RNA-dependent helicase and ATPase up frameshift 1 (UPF1), which was described as the core factor of NMD. Due to its ATPase and helicase activities, UPF1 can bind sequence-independently to any accessible transcript, independent of that transcript being an NMD target or not (Hogg and Goff 2010; Kurosaki and Maquat 2013; Zünd et al. 2013; Kurosaki et al. 2014; Lee et al. 2015). UPF1 has two evolutionary conserved functional domains: an amino-terminal cysteine- and histidine-rich (CH) domain, as well as a carboxy-terminal RNA helicase domain, which enables UPF1 to translocate over mRNAs by hydrolyzing ATP (Czaplinski et al. 1995; Bhattacharya et al. 2000). Without the interaction with other proteins, UPF1 forms a closed intramolecular structure which gets mediated through the interaction of its terminal domains (Fiorini, Boudvillain, and Le Hir 2013). This closed structure can be relaxed by interaction of the CH-domain of UPF1 with UPF2 (Chamieh et al. 2008). Furthermore, the interaction of UPF1 with UPF2, as well as UPF3B, further enhances the ATPase and helicase activities of UPF1 (Chamieh et al. 2008; Chakrabarti et al. 2011).

So far, different NMD models have been described, all of which require UPF1, but different compositions of the other NMD factors (Gehring et al. 2005; Bühler et al. 2006; Chan et al. 2007; Aznarez et al. 2018). That might be one reason why the cellular abundance of UPF1 protein is higher than the abundancies of other NMD factors, such as UPF2 or UPF3B (Maderazo et al. 2000). Using transcriptome-wide analysis it was shown that UPF1 binds to both, NMD targets and not-NMD targets, therefore the question of how the cell discriminates between target mRNAs and mRNAs without NMD triggering features is still discussed. Currently, there are two

different models for NMD target recognition: the exon junction complex (EJC)-dependent NMD model, as well as the long 3' UTR-mediated NMD model.

1.2.1 The EJC-dependent NMD model

Recognition of target mRNAs by the EJC-dependent NMD model depends on the deposition of exon junction complexes on the mRNA and the retention of one or more EJCs downstream of the translational termination codon (TC) upon translation termination. During the splicing process, EJCs are sequence-independently deposited 20-24 nucleotides upstream of exon-exon junctions, with an estimated efficiency of about 80% (Le Hir et al. 2000; Le Hir et al. 2001; Bono et al. 2006). During translation, the translocating ribosome removes proteins that are bound to the mRNA, such as the EJCs, until the ribosome encounters a translational stop codon and dissociates (cf. Figure 1.5). Proteins that are associated with the 3' UTR, however, remain associated with the mRNA (Dostie and Dreyfuss 2002). Hence, in the case of a pre-mature termination codon (PTC) upstream of the last exon-exon junction, the downstream EJCs will remain on the mRNA and facilitate the recruitment of the NMD machinery (Lykke-Andersen, Shu, and Steitz 2000; Gehring et al. 2005). Here, the general rule was implied, that PTCs very efficiently trigger NMD if they are located more than 50-55 nucleotides upstream of the last exon-exon junction (Thermann et al. 1998). A schematic overview of the EJC-dependent NMD model is depicted in Figure 1.6.
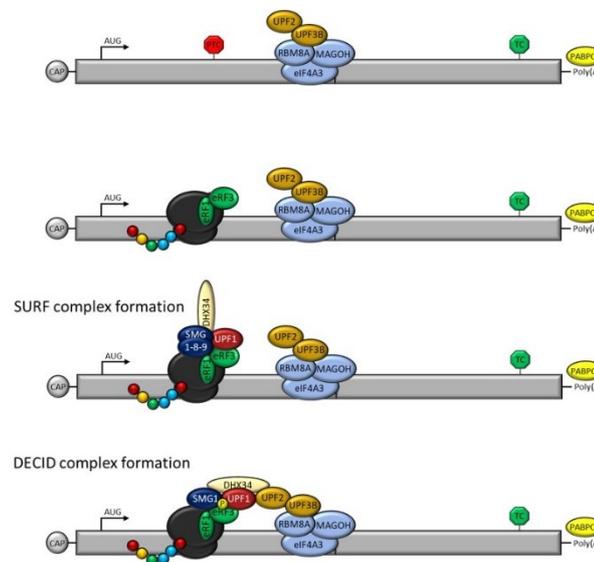
Figure 1.6: Schematic overview of the EJC-dependent NMD model. When the translation termination occurs at a pre-mature termination codon (PTC, red stop sign), the ribosome does not translocate further and proteins which are associated with the mRNA downstream of the PTC remain bound to the mRNA, like the EJC core (light blue; eIF4A3, RBM8A and MAGOH) and its associated proteins UPF3B and UPF2 (brown). Upon pre-mature translation termination, UPF1 (red) and SMG1 (dark blue) are recruited to the terminating ribosome and their interaction with the eukaryotic release factors forms the SMG1-UPF1-eRFs (SURF) complex. Within the SURF complex SMG1 is kept in its inactive form by SMG8 and SMG9. DEAH-box polypeptide 34 (DHX34, light yellow) mediates the interaction of UPF1 with the EJC-bound UPF2, which promotes SMG8/9 dissociation, SMG1 activation, and UPF1 phosphorylation. The SURF

complex transforms into the decay-inducing (DECID) complex upon UPF1 phosphorylation, resulting in the recruitment of mRNA decay factors. This figure was adapted from (Kurosaki, Popp, and Maquat 2019).

The EJC core is composed of three different proteins: The eukaryotic initiation factor 4A3 (eIF4A3), a helicase that anchors the EJC on the mRNA. The RNA-binding protein 8A (RBM8A) as well as mago nashi homolog (MAGOH) (Mabin et al. 2018). This core EJC is extended by other proteins, such as UPF3B and RNA-binding protein with serine-rich domain 1 (RNPS1) (Singh et al. 2012; Mabin et al. 2018). During the transport from the nucleus to the cytoplasm, RNPS1 gets replaced by metastatic lymph node 51 (MLN51) (Singh et al. 2012; Mabin et al. 2018).

When the first round of translation terminates and one or more EJCs remain bound to the mRNA, the EJC-dependent NMD gets efficiently activated. Translation termination involves eRF1 and eRF3 binding to the A-site of the 80S ribosome, where they promote the release of the nascent protein. During NMD-triggering translation termination UPF1 and SMG1 (suppressor with morphogenetic effect on genitalia 1) are recruited to the eRF1-eRF3 translation termination complex to form the SMG1-UPF1-eRFs (SURF) complex (Kashima et al. 2006).

Within the SURF complex, before SMG1 and UPF1 interact with the downstream EJC, SMG1 is kept in an inactive form by complexing with SMG8 and SMG9 (Kashima et al. 2006; Yamashita et al. 2009; Arias-Palomo et al. 2011; Deniaud et al. 2015). Associated with this SMG1-SMG8/SMG9 complex is the RNA helicase DEAH-box polypeptide 34 (DHX34), which mediates the association of UPF1 with the EJC-bound UPF2 (Hug and Cáceres 2014). While UPF2 is bound to the EJC by interaction with UPF3B, the interaction of UPF2 with UPF1 stimulates the helicase activity of UPF1 (Chamieh et al. 2008; Chakrabarti et al. 2011).

The formation of the SURF complex and the bridging to the EJC promote the dissociation of SMG8/SMG9, the activation of SMG1, and the subsequent phosphorylation of UPF1 by SMG1 (Yamashita et al. 2001; Kashima et al. 2006). SMG1 is a phosphatidylinositol 3-kinase related kinase, that preferentially phosphorylates serine and threonine residues upstream of glutamine residues (S/TQ motives), which are enriched in the amino- and carboxy terminus of UPF1 (Yamashita et al. 2001; Durand, Franks, and Lykke-Andersen 2016). The phosphorylation of UPF1 ultimately results in the formation of the decay-inducing (DECID) complex, translation repression, and the recruitment of mRNA decay factors (Kashima et al. 2006; Kurosaki, Popp, and Maquat 2019).

In General, the EJC-mediated NMD occurs near the nucleus, immediately after nuclear export of the mRNA and the so-called pioneer round of translation, while the 5' end of the mRNA is still bound by the cap-binding protein (CBP) heterodimer CBP20-CBP80 (Cheng and Maquat 1993; Ishigaki et al. 2001; Hosoda et al. 2005). It was shown that CBP80 interacts with UPF1 and SMG1 at the terminating ribosome to promote SURF complex formation and promotes bridging to the EJC (Hwang et al. 2010). However, once the CBP20-CBP80 heterodimer gets replaced by the

steady-state cytoplasmic CBP eIF4E, the EJC-mediated NMD target mRNAs become relatively insensitive to NMD, as their half-lives equal those of not target mRNAs (Trcek et al. 2013).

1.2.2 The long 3' UTR-mediated NMD model

The NMD target recognition without a downstream located EJC is less well understood, but with the long 3' UTR-mediated NMD model target mRNAs degradation can be explained beyond the pioneer round of translation, therefore targeting CBP20-CBP80 bound, as well as eIF4E bound mRNAs (Hosoda et al. 2005). The efficient translation termination and ribosome recycling is enhanced by the binding of polyadenylate-binding protein 1 (PABPC1) to the poly(A) tail of the mRNA, as PABPC1 stimulates the recruitment of eRF1 and eRF3 to the terminating ribosome (Ivanov et al. 2016). For NMD target recognition a competition between UPF1 and PABPC1 for the interaction with eRF3 was shown (Silva et al. 2008; Kervestin et al. 2012).

Mainly, the EJC-independent NMD target recognition seems to be due to the physical distance between the translational stop codon and the poly(A) tail of an mRNA, as the shortening of that distance reduced UPF1 binding to the 3' UTR and the sensitivity of the mRNA to NMD (Eberle et al. 2008; Hogg and Goff 2010; Kurosaki and Maquat 2013). However, it was shown by transcriptome-wide analysis that phosphorylated UPF1, which is the activated from of UPF1, is not only enriched on 3' UTRs of EJC-mediated NMD target mRNAs, but also on 3' UTRs of EJC-independent NMD targets (Kurosaki et al. 2014; Imamachi et al. 2017). As bound UPF1 is actively phosphorylated by SMG1 in the DECID complex it was postulated, that for the EJC-independent form of NMD, UPF1 accumulates in the long 3' UTR regions and will be phosphorylated by chance (Imamachi et al. 2017). A schematic overview of the long 3' UTR-mediated NMD model is shown below in Figure 1.7.
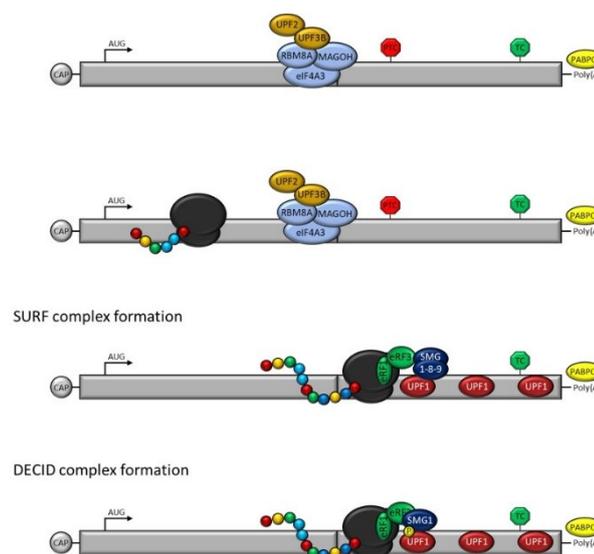
Figure 1.7: Schematic sketch of the long 3' UTR-mediated NMD model. When the translation termination occurs at a pre-mature termination codon (PTC, red stop sign) which is located within the last exon, the translocating ribosome removes all upstream-located proteins, like the exon-junction complexes (EJCs; light blue; eIF4A3, RBM8A and

MAGOH) and the EJC associated proteins UPF3B and UPF2 (brown). Therefore, upon encounter of the PTC, no more EJCs are located downstream of the terminating ribosome. The presence of UPF1 protein (red) in the 3' untranslated region (UTR) depends on the length of the UTR and the distance between the PTC to the poly(A) tail and the cytoplasmic poly-A binding protein 1 (PABPC1, yellow), as the interaction of PABPC1 and eRF3 enhances the efficient termination. Here, the SURF complex can form upon delayed translation termination, UPF1 gets phosphorylated by SMG1 resulting in the formation of the DECID complex and the recruitment of mRNA decay factors. The figure was adapted from (Kurosaki, Popp, and Maquat 2019).

In mammals the long 3' UTR-mediated NMD was observed for endogenous, non-mutated transcripts that harbor 3' UTRs which are longer than 1 kb (Singh, Rebbapragada, and Lykke-Andersen 2008). Even though a positive correlation between the 3' UTR length and UPF1 binding was observed (Hogg and Goff 2010; Kurosaki and Maquat 2013; Hurt, Robertson, and Burge 2013), it was shown by transcriptome-wide analysis which measured mRNA half-lives and abundancies, that the 3' UTR length did not directly correlate with NMD susceptibility (Tani et al. 2012; Hurt, Robertson, and Burge 2013). That effect is due to *cis*-acting regulatory elements that inhibit NMD (Toma et al. 2015). Examples for these NMD inhibiting factors are the polypyrimidine tract-binding protein 1 (PTBP1) and the heterogeneous nuclear ribonucleoprotein L (hnRNPL), which were shown to inhibit NMD if they are bound to the mRNA immediately downstream of the termination codon (Ge et al. 2016; Kishor, Ge, and Hogg 2019).

1.2.3 Degradation of target mRNAs by NMD

As mentioned before, the most important step for mRNA degradation through NMD is the phosphorylation of UPF1 at multiple residues within its amino-terminal and carboxy-terminal regions by SMG1. That phosphorylation promotes the interaction of UPF1 with eRF3 (Isken et al. 2008) and is necessary for recruitment of RNA degradation enzymes (Ohnishi et al. 2003; Okada-Katsuhata et al. 2012). In humans there are two well-described ways of NMD-mediated RNA degradation, which are mediated by SMG6 or SMG5-SMG7 (Loh, Jonas, and Izaurralde 2013; Colombo et al. 2017).

The SMG6-mediated RNA degradation is initiated by SMG6 recruitment to the mRNA by phosphorylated UPF1. SMG6 is an endonuclease, which cleaves target mRNAs in the vicinity of the termination codon, generating a 5' and a 3' cleavage product (Huntzinger et al. 2008; Eberle et al. 2009; Lykke-Andersen et al. 2014). The 5' cleavage product gets degraded in 3' to 5' direction by the exosome, or the DIS3-like exonuclease 2 (DIS3L2) (Malecki et al. 2013; Kurosaki et al. 2018). The 3' cleavage product, however, is still bound by proteins, which have to be removed by the UPF1 helicase activity to provide excess to the mRNA by the exoribonuclease XRN1 for 5' to 3' RNA degradation (Huntzinger et al. 2008; Eberle et al. 2009).

The other way of NMD-induced RNA degradation involves the recruitment of the SMG5-SMG7 heterodimer by phosphorylated UPF1, resulting in exoribonucleolytic decay by deadenylation and decapping. The SMG5-SMG7 heterodimer recruits the CCR4-NOT deadenylation complex through the interaction of SMG7 and the catalytic POP2 subunit (Unterholzner and Izaurralde

2004; Yamashita et al. 2005; Loh, Jonas, and Izaurralde 2013). The poly(A) tail gets removed by CCR4-NOT, followed by exosome-mediated 3' to 5' degradation and recruitment of the decapping complex member mRNA decapping enzyme 2 (DCP2) (Unterholzner and Izaurralde 2004; Yamashita et al. 2005; Loh, Jonas, and Izaurralde 2013; Cho et al. 2013). The interaction of phosphorylated UPF1 with SMG5, however, can additionally directly recruit the decapping complex and the decapping enhancer PNRC2 (Lykke-Andersen 2002; Lejeune, Li, and Maquat 2003; Cho, Kim, and Kim 2009; Cho et al. 2013; Nicholson et al. 2018). The obtained decapped and deadenylated RNA is then degraded in either 5' to 3', or in 3' to 5' direction. Following RNA degradation, the SMG5-SMG7 heterodimer recruits the protein phosphatase 2A (PP2A) which dephosphorylates UPF1 (Ohnishi et al. 2003; Kurosaki et al. 2014).

1.2.4   Regulation of endogenous genes by NMD

Using transcriptome-wide analysis, it was calculated that about 30-35% of all pre-mRNA splicing events in humans routinely generate PTCs, which can trigger NMD, and that one-third of all inherited human diseases are caused by nonsense or frameshift mutations, resulting in a PTC within the transcript (Mendell and Dietz 2001; Lewis, Green, and Brenner 2003; Pan et al. 2006; Pan et al. 2008; Mort et al. 2008; Weischenfeldt et al. 2012). However, not only these "faulty" transcripts were shown to trigger NMD, as system-wide approaches on the human transcriptome showed that also 10-20% of "normal" transcripts are targeted by NMD (Mendell et al. 2004; Yepiskoposyan et al. 2011; Hurt, Robertson, and Burge 2013)

This physiological role of NMD includes recognition of transcripts with unusually long 3' UTRs (> 1 kb), spliced introns in 3' UTRs, or upstream open reading frames (uORFs) whose translational termination codon is located 50-55 nucleotides upstream of the next exon-exon junction (Lykke-Andersen and Jensen 2015; Nickless, Bailis, and You 2017). Another layer of complexity to NMD is introduced by tissue-specific or developmentally regulated alternative splicing, meaning that a PTC-containing transcript that is only produced in some tissues might escape NMD in other tissues (Yeo et al. 2004; Barberan-Soler, Lambert, and Zahler 2009).

Furthermore, alternative splicing coupled NMD (AS-NMD) is used as a homeostatic regulatory mechanism to regulate the expression of different mRNA isoforms produced from one gene. One prominent example are the members of the SR splicing factor family, which autoregulate their levels by inclusion or exclusion of a PTC-containing exon, a so-called "poison-cassette exon", by binding to an exonic splicing enhancer (ESE) within the poison exon. High levels of the SR protein promote binding to the ESE, inclusion of the PTC-containing exon, and NMD-dependent degradation of the transcript, which ultimately lowers the protein levels (Lewis, Green, and Brenner 2003; Lareau et al. 2007; Ni et al. 2007).

Finally, the UGA codon not only serves as a translational stop codon but can also function as a selenocysteine incorporation signal. However, if the intracellular selenocysteine levels are low,

that UGA codon might be used as a translational termination signal for the ribosome, thereby producing a PTC if located upstream of the last exon-exon junction (Moriarty, Reddy, and Maquat 1998).

In conclusion, due to the broad target range and the preference for "unusual" RNAs the nonsense-mediated decay pathway as a post-transcriptional mechanism, allows a fast response to various stimuli like stress or viral infections (Balistreri, Bognanni, and Mühlemann 2017; Contu et al. 2019).

1.2.5 Viruses and NMD

Viral genomes are tightly packed, storing a tremendous amount of information in a very limited genome. Therefore, viral mRNAs exhibit atypical features that might distinguish them from cellular transcripts. These features potentially mark these transcripts for degradation through NMD, making NMD an ancestral form of intrinsic antiviral immunity (Balistreri et al. 2014; Garcia, Garcia, and Voinnet 2014). These features include polycistronic RNAs, where the multiple ORFs might resemble uORFs, or very long 3' UTRs where the translation termination codon is in the middle of the transcript, far away from the poly-A tail (Rigby and Rehwinkel 2015; Balistreri, Bognanni, and Mühlemann 2017).

Based on their efficient replication and co-existence, viruses must have evolved mechanisms that protect their transcripts from recognition through the NMD machinery. However, NMD has been shown to target a broad range of viruses, from (+) stranded RNA viruses, which replicate in the cytoplasm, to retroviruses that stably integrate their DNA into the host genome.

For alphaviruses, for example, it was shown by genome-wide small interfering RNA (siRNA) screens, that the genomic RNA, with a 3' UTR of approximately 4000 nucleotides is targeted by NMD, even though that 3' UTR was not identified as the main NMD triggering feature they identified the NMD factors UPF1, SMG5, and SMG7 as restriction factors for Semliki forest virus (SFV) infection (Balistreri et al. 2014). For other viruses, like hepatitis C virus (HCV), it was shown that they cause a progressive inhibition of the cellular NMD activity. For HCV, the viral core protein interacts directly with the EJC recycling factor PYM and prevents its interaction with other components of the EJC, thereby inhibiting NMD (Ramage et al. 2015).

In the case of retroviruses, targeting of viral transcripts by NMD, as well as viral modulation of the hosts NMD machinery was observed (LeBlanc and Beemon 2004; Nakano et al. 2013; Mocquet et al. 2012). For example, the full-length unspliced transcript of Rous sarcoma virus (RSV), which gets translated into RSV Gag protein, has a 3' UTR of about 7000 nucleotides and might therefore be targeted by NMD. Indeed, the introduction of PTCs at different places of the *gag* ORF were shown to trigger NMD, suggesting that the natural TC is protected from NMD by specific *cis*-acting elements (Barker and Beemon 1994; LeBlanc and Beemon 2004). This NMD

antagonizing effect was shown to be mediated by the RNA stability element (RSE), a 400-nucleotide long sequence element which is located immediately downstream of the *gag* TC (Weil and Beemon 2006; Withers and Beemon 2011). The RSE forms a complex RNA secondary structure with several pyrimidine-rich stretches which serve as a binding platform for the polypyrimidine tract binding protein 1 (PTBP1) (Withers and Beemon 2011; Ge et al. 2016).

Another example of NMD inhibition is given by another retrovirus, the human T-lymphotropic virus type 1 (HTLV-1). Instead of encoding *cis*-acting elements within their mRNAs, like RSV, HTLV-1 was shown to interact in *trans* with proteins of the NMD machinery. Here, the HTLV-1 proteins TAX and REX were shown to bind UPF1 and the translation initiation complex component INT6/eIF3E, resulting in partial NMD inhibition and the stabilization of viral transcripts (Mocquet et al. 2012; Nakano et al. 2013).

In the case of human immunodeficiency virus type 1 (HIV-1) viral interference with the NMD machinery is under investigation. So far, UPF1 was shown to function less as a restriction factor and more as a positive regulator of HIV-1 gene expression, vRNA nuclear export, and infectivity of released progeny virions (Ajamian et al. 2008; Serquina et al. 2013; Ajamian et al. 2015).

Furthermore, ongoing HIV-1 research has shown that the virus interacts with the major NMD factor UPF1, potentially relieving it of its roles in the NMD degradation pathway. However, how the different mRNAs of HIV-1 are targeted by NMD and if the virus has established *cis*-working ways of stabilizing its transcriptome, and if this might be coupled to massive regulation via alternative splicing has to be investigated.

1.3   The human immunodeficiency virus 1 (HIV-1)

The human immunodeficiency virus 1 (HIV-1) was discovered in 1983, two years after the first cases of advanced and unexplained immunodeficiency, which is known today as the acquired immune deficiency syndrome (AIDS) (Gottlieb et al. 1981; Barre-Sinoussi et al. 1983; Gallo et al. 1983). Later, however, the first cases were traced back to 1959 (Zhu et al. 1998).

Until today HIV-1 remains a major health issue and the HIV-1 prevalence is highest in groups with certain risk factors. Those key populations include men who have sex with men, intravenous drug users, people in prisons and other closed settings, sex workers, and transgenders (Deeks et al. 2015). At the end of 2020, approximately 37.7 million people worldwide were infected with HIV-1 (WHO 2021), while in Germany approximately 90,700 people were infected with HIV-1 at the end of 2019 (RKI 2020).

HIV-1 is a complex retrovirus, that upon reverse transcription of its genomic RNA stably integrates into its host DNA, where it resides throughout the lifetime of the infected cells. Until today there is no sterile cure for HIV-1 infection, and infected individuals must be treated livelong with anti-retroviral therapy (ART) (Trono et al. 2010; Chun, Moir, and Fauci 2015). Even

though the present ART therapy has increased the life expectancy of HIV-1 infected individuals and enables the patients to control the disease by suppression of viral levels below the detection limit, the long-term use is still complicated due to adherence, severe side effects, and the emergency of drug resistance mutations (DRM) (Sertznig et al. 2018; Forsythe et al. 2019). One of the major reasons for therapy failure, especially in low-income countries are the acquired DRMs, as alternative therapies are limited for people with multi-drug resistant HIV-1 (Pennings 2013).

1.3.1 The HIV-1 replication cycle

HIV-1 virions have a size of around 100 nm and its surface, which consists of a host-cell derived lipid membrane is covered with viral envelope glycoprotein trimers (gp120 and gp41). The virus contains a conical-shaped capsid, which contains two copies of the viral RNA genome and retroviral enzymes, like the reverse transcriptase (RT), integrase (IN), and protease (Pro) (Ramdas et al. 2020). The HIV-1 replication cycle can be divided into the following steps: attachment, membrane fusion, reverse transcription, integration, transcription, splicing, translation, assembly, budding, and maturation. A schematic overview of the HIV-1 live cycle is depicted in Figure 1.8 and described below in the following chapters.
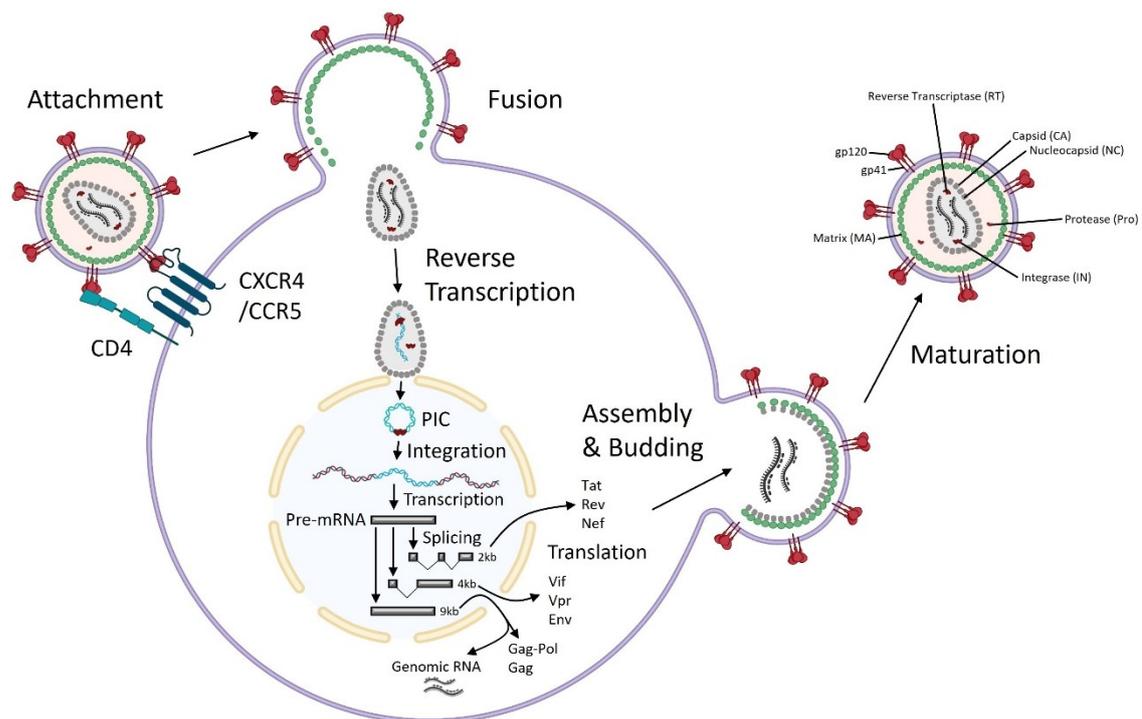
Figure 1.8: HIV-1 replication cycle. The virus attaches to the host cell via the interaction of gp120 (Env) and the cellular CD4 receptor and one of its co-receptors (CXCR4 or CCR5). Attachment and membrane fusion results in the release of the viral capsid (CA) into the cytoplasm of the host cell. During transport to the nucleus, the viral RNA gets reverse transcribed into double-stranded DNA (dsDNA). Uncoating of the dsDNA allows the formation of the pre-integration complex (PIC). The subsequent integration of the viral dsDNA into the host DNA gets mediated by the viral Integrase (IN), resulting in the formation of the provirus. Transcription and alternative splicing of the viral pre-mRNA generates three distinct mRNA classes: the fully spliced 2kb class, the intron-containing 4kb class, and the unspliced 9kb class.

The 4kb and 9kb mRNAs are translated into structural and enzymatic proteins, which build up the new virion, while the 9kb mRNA also serves as genomic RNA for newly assembling virions. Upon assembly new virions are released from the host cell by budding and mature into new infectious particles. The figure was created with BioRender.com and was adapted from (Ramdas et al. 2020).

1.3.1.1 Virus attachment and membrane fusion

Like all viruses, HIV-1 is dependent on its host for the replication and amplification of its genome and the assembly of new virions. Hence, the first step in the viral replication cycle is the attachment of the virus to its host cell and the subsequent fusion of the viral envelope with the target cell plasma membrane. This attachment gets mediated by the viral envelope glycoprotein gp120 (Env) and its interaction with the target cell cluster of differentiation receptor 4 (CD4) (cf. Figure 1.8). The human CD4 receptor is expressed on different immune cells, which include T-helper cells (CD4$^+$ T-cells), macrophages (MQ), and dendritic cells (DCs) (Pope et al. 1994). The interaction of gp120 with CD4 results in conformational changes, which allow the subsequent binding of gp120 to the cellular co-receptors CXCR4 or CCR5 (Wilen, Tilton, and Doms 2012). This binding again results in a conformational change, which enables the insertion of the hydrophobic gp41 fusion peptide into the cells membrane, ultimately resulting in the fusion of the viral membrane with the hosts' cell membrane (Doms and Moore 2000) (cf. Figure 1.8).

1.3.1.2 Uncoating, reverse transcription, and nuclear import

Upon successful fusion of the virus with the plasma membrane the viral capsid (CA) is released into the cytoplasm of the host cell (cf. Figure 1.8). The CA contains two molecules of the positive sense HIV-1 RNA genome, occupied by the nucleocapsid proteins (NC), as well as the viral encoded enzymes, reverse transcriptase (RT), integrase (IN), protease (Pro), and other cellular proteins (Miller, Farnet, and Bushman 1997). The CA remains intact while it gets transported from the viral entry site to the nuclear pores, a process for which the virus hijacks the hosts' microtubule network and migrates along the cytoskeleton of the cell (Fassati and Goff 2001; McDonald et al. 2002; Burdick et al. 2020; Selyutina et al. 2020). During this transport to the nucleus, the reverse transcription of the viral genome takes place within the reverse transcription complex (RTC), in which the viral RT enzyme transcribes the viral (+) RNA genome into a double-stranded DNA copy and simultaneously degrades the RNA template. A schematic overview of the reverse transcription process is shown in Figure 1.9 and described below. In general, the RTC consists not only of the viral RNA, but also of the host-derived transfer-ribonucleic acid Lysin 3 (tRNA^{Lys3}) which serves as primer for the reverse transcription, the eukaryotic translational elongation factor 1A (eEF1A), but also of newly synthesized DNA and several viral and host factors (Isel et al. 1995; Fassati and Goff 2001).
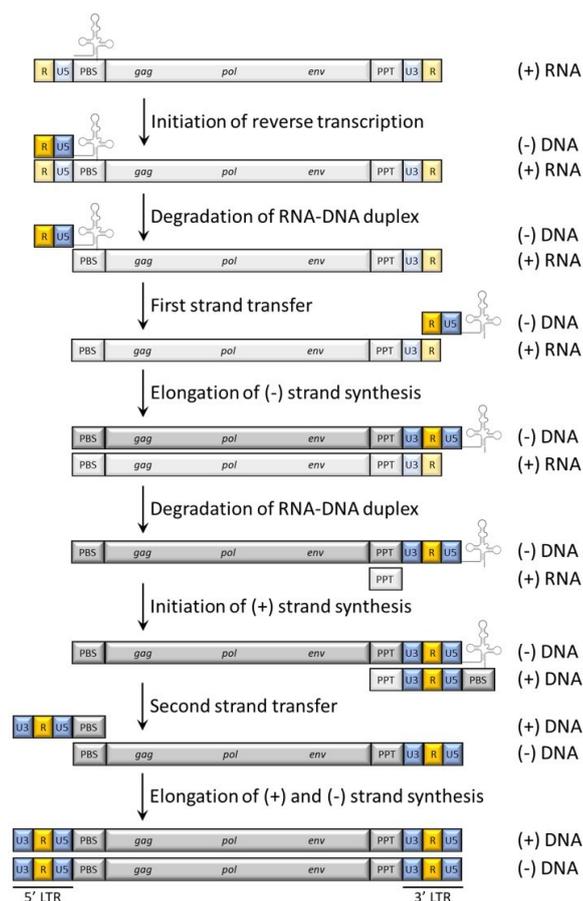
Figure 1.9: Schematic overview of the reverse transcription process. The viral (+) RNA is bound by the tRNA^Lys3 at the primer binding site (PBS), which serves as primer to initiate reverse transcription. First, the (-) strand of DNA is synthesized in 5' to 3' direction until the 5' end of the vRNA. The RNA-DNA duplex gets degraded by the RNaseH activity of the RT. Upon the first strand transfer, the complementary repeated region (R) of the newly synthesized (-) DNA strand and the (+) RNA pair at the 3' end of the (+) RNA and the elongation of the (-) DNA strand continues. Again, the RNA-DNA duplex gets degraded, except for the region of the polypurine tract (PPT). This PPT subsequently serves as primer for the (+) DNA strand synthesis. The second strand transfer allows the alignment of the complementary PBS sequences and the elongation of both DNA strands, the (+) and the (-) strand resulting in complementary DNA with long terminal repeats (LTRs) on both sides. This figure was adapted from (Hu and Hughes 2012), the tRNA was created with BioRender.com.

The tRNA^Lys3 is bound to the primer binding site (PBS), which is located at the 5' end of the viral genomic RNA (vRNA), downstream of the R and U5 region. Using the tRNA^Lys3 as a primer, the reverse transcription process is initiated by the viral RT with the help of several cellular factors (Isel et al. 1995; Isel, Ehresmann, and Marquet 2010). Starting at the PBS, the minus strand complementary DNA (cDNA) is synthesized in 5' to 3' direction until RT reaches the 5' end of the vRNA. The RNaseH activity of the RT cleaves the vRNA right after the synthesis of the new RNA-DNA-hybrid. Based on the complementarity of the repeated region (R) within the HIV-1 long terminal repeats (LTRs) the newly synthesized (-) strand cDNA transfers from the 5' to the 3' end of the vRNA, from where the (-) strand cDNA synthesis is continued. The RNA-DNA-hybrid is again resolved by the RNaseH activity of the HIV-1 reverse transcriptase, except for the region

of the polypurine tract (PPT), which is resistant to RNaseH cleavage and remains double-stranded. This PPT subsequently serves as primer for the reverse transcription of the new (+) strand cDNA, which finally results in double-stranded DNA (dsDNA) (Charneau, Alizon, and Clavel 1992). As for the synthesis of the first DNA strand, the reverse transcription of the new (+) DNA strand ends at the 5' end of the template, at the PBS. At this point, the second-strand transfer occurs, where the complementary PBS sequences align, and the RT is continued, on both DNA strands, to generate the full-length dsDNA (Arhel 2010).

As the newly generated double-stranded viral genomic DNA is still located within the viral capsid, it must be uncoated before integration into the host chromosome. Even though it was long thought that uncoating occurs at the nuclear pore complex (NPC), it was recently shown that whole capsids can be imported into the nucleus via nuclear pores (Zila et al. 2021). Capsid disassembly liberates the viral dsDNA, which is complexed with the host and viral proteins, known as the pre-integration complex (PIC) (Khiytani and Dimmock 2002).

1.3.1.3 Integration

Upon uncoating, the viral integrase (IN) enzyme directs the integration of the viral DNA into transcriptionally active sites within the host chromatin, specifically at AT-rich euchromatin regions (Marini et al. 2015; Ciuffi 2016). As the nuclear entry of the viral DNA depended on interactions with host proteins, so does the integration into the hosts' genome (Campbell and Hope 2015).

For the first step of integration, IN binds the newly synthesized dsDNA of the virus and orients the two ends of the LTRs into its catalytic domain (cf. Figure 1.8). Within the catalytic domain, two nucleotides are cleaved from the 3' end of each LTR, resulting in one free 3'OH end at each site. In the next step of integration, the host genome passes through the IN, while IN scans for a target site within the genome. For integration, a short palindromic cytidine-adenine sequence is preferred (Wu et al. 2005). Once IN found the target site, the viral dsDNA is integrated into the host genome by direct esterification, initiated by the free electrons of the 3' hydroxyl group. The integration results in 5' overhangs of the viral dsDNA on both sides, which are repaired by cellular ligases (Li et al. 2006). Now, the integrated viral dsDNA is called provirus, as viral proteins can now be expressed by transcription of the viral genome by cellular proteins.

1.3.1.4 Transcription and alternative splicing of HIV-1 mRNAs

As mentioned in chapter 1.1.1, transcription of eukaryotic genes by the cellular DNA-dependent-RNA-polymerase (RNA polymerase II) requires a promoter sequence. In the case of the integrated HIV-1 proviral DNA, this promoter sequence is encoded within the long terminal repeat (LTR). Even though the LTR is present on both sides of the viral genome, and both LTRs can function as promotors, the 5'-LTR-promoter has significantly higher transcriptional activity (Klaver and Berkhout 1994). Furthermore, HIV-1 encodes its own trans-activator protein (Tat),

which promotes transcription by the interaction of Tat with the transactivation response region (TAR) of HIV-1 (Kao et al. 1987). Transcription of the proviral DNA results in a full-length precursor mRNA (pre-mRNA), which contains the densely packed HIV-1 open reading frames (ORFs), encoding for at least eighteen protein and polyprotein-isoforms (Jäger et al. 2011; Karn and Stoltzfus 2012).

As the eukaryotic 43S ribosomal subunit starts scanning of the mRNA at the 5' CAP structure and initiates translation at the first efficient AUG codon, HIV-1 has to transfer each ORFs into proximity of the 5' CAP, a process that is mediated by extensive alternative splicing (Schwartz et al. 1990; Purcell and Martin 1993; Stoltzfus 2009). Here, one exception is given by the bicistronic *vpu*/*env* mRNA, which permits translation initiation at the *env* AUG by discontinuous ribosome scanning (Krummheuer et al. 2007; Anderson et al. 2007). However, in general, the excision of each upstream AUGs is achieved by alternative splicing of the HIV-1 RNAs and the usage of at least five different 5' splice sites (SD1, SD2, SD2b, SD3, SD4), and eight different 3' splice sites (SA1, SA2, SA3, SA4c, SA4a, SA4b, SA5, SA7), thereby producing over 50 different viral mRNA isoforms, which allow the balanced expression of all viral proteins (Schwartz et al. 1990; Purcell and Martin 1993; Ocwieja et al. 2012; Sertznig et al. 2018). A schematic overview of the overlapping HIV-1 ORFs, the encoded splice donors and acceptors, as well as the distinct mRNA classes is shown in Figure 1.10.
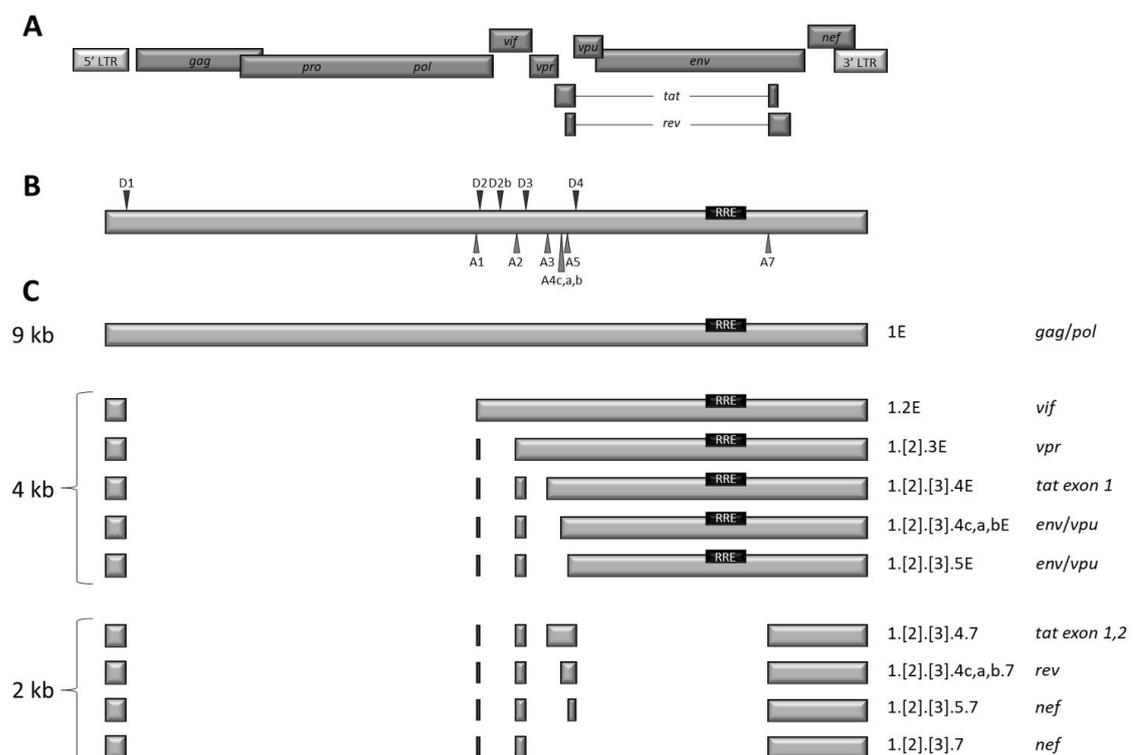
Figure 1.10 The HIV-1 genome and the mRNA classes. (A) Schematic overview of the HIV-1 genome structure with the 5' and 3' long terminal repeats (LTRs) and the overlapping open reading frames (ORFs). **(B)** The HIV-1 pre-mRNA

with the respective location of the rev-responsive element (RRE) and the five major splice donors (D1, D2, D2b, D3, D4) and eight splice acceptors (A1, A2, A3, A4c, A4a, A4b, A5 and A7). **(C)** The distinct HIV-1 mRNA classes and their exon compositions. The unspliced 9kb class not only serves as the viral genome but also encodes for Gag and Gag/Pol. The intron-containing 4kb class encodes for Env, Vpu, Vif, and Vpr, as well as a Tat exon 1 variant. The fully spliced 2kb class encodes for Tat, Rev, and Nef. This figure was adapted from (Sertznig et al. 2018).

Based on their size, the HIV-1 transcripts are grouped into three distinct RNA classes: The 2kb class, or the fully spliced RNA class, encoding for the regulatory proteins Rev and Tat, as well as the accessory protein Nef. The intron-containing 4kb class, which encodes for the Env glycoprotein, as well as Vpu, Vif, and Vpr proteins. And finally, the unspliced 9kb RNA, which functions as genomic viral RNA, as well as mRNA for the Gag and Gag-Pol polyproteins (Purcell and Martin 1993). The expression of the three different mRNA classes during viral replication is temporally regulated. The 2kb class is mostly generated in the early phases of infection, whereas the 4kb class mRNAs are most abundant in the intermediate phase and the 9kb class mRNAs are primarily found in the late phase (Kim et al. 1989). Additionally, using next-generation sequencing (NGS) based transcriptome analysis an additional minor 1 kb mRNA class was identified (Ocwieja et al. 2012).

As described before, the success of a splicing reaction depends not only on the intrinsic strength of a given splice site but also on the presence of *cis*-, and *trans*-acting splicing regulatory elements. In general, most HIV-1 splice sites have a low intrinsic strength and are therefore used inefficiently. The splice donors with the highest intrinsic strength are SD1 and SD4, while donors SD2 and SD3 are relatively weak (O'Reilly, McNally, and Beemon 1995). The HIV-1 splice acceptor sites are also intrinsically weak, based on their suboptimal pyrimidine-rich regions with interspersed purines (Dyhr-Mikkelsen and Kjems 1995). Even though, the highest intrinsic strength of the HIV-1 acceptor sites are exhibited by SA2 and SA3 (Kammler et al. 2006). However, due to the overall low intrinsic strength of HIV-1 splice sites the splice site usage is mainly determined by the presence of splicing regulatory elements and splicing regulatory proteins.

1.3.1.4.1 The 2 kb class of HIV-1 RNAs

All transcripts of the 2kb class are spliced from the major splice donor 1 to one of the splice acceptors within the central 3' ss cluster, and additionally from SD4 to SA7, which removes the RRE. Generally, except for the 9kb class, all HIV-1 transcripts are spliced at the major SD1 (Purcell and Martin 1993; Chang et al. 2011). Furthermore, it was shown that the splicing process from SD4 to SA7 is dependent on the prior splicing of the SD1 to a downstream SA (Bohne, Wodrich, and Kräusslich 2005). Therefore, the fully spliced 2kb class of HIV-1 mRNAs is characterized by the absence of the Gag-Pol, as well as the Env coding region, and encodes for the regulatory proteins Tat and Rev, as well as the accessory protein Nef (Purcell and Martin 1993).

The full-length *tat* mRNAs are generated by splicing of SD1 to SA3 (*tat* exon 1), as well as SD4 to SA7 (*tat* exon 2). However, alternative splice variants have been described, differing in the presence of the two leader exons 2 and 3 (Purcell and Martin 1993). Tat protein consists of 101 amino acids which are encoded by two exons. The first *tat* exon encodes the nucleotides for amino acids 1-72, while the second *tat* exon encodes the nucleotides for the amino acids 73-101. There is, however, another *tat* mRNA isoform, encoding only for the first 72 amino acids. This truncated Tat protein is encoded by transcripts of the 4kb mRNA class and only contains the first *tat* exon (Purcell and Martin 1993) (cf. Figure 1.10C).

Rev is encoded by at least twelve alternative splice variants, which are mainly generated by the usage of one of the three splice acceptors SA4c, SA4a, or SA4b, which are located upstream of the rev translational start codon, as well as the potential presence of the two leader exons 2 and 3 (Schwartz et al. 1990; Purcell and Martin 1993; Stoltzfus 2009).

The accessory protein Nef is also encoded by different splice variants, which can again be distinguished by the presence of leader exons 2 and 3, as well as the presence of exon 5 (Purcell and Martin 1993). The *nef* ORF is located at the 3' end of the HIV-1 genome and *nef* transcripts are generated by the inclusion of exon 5 and the additional splicing from SD4 to SA7. The splice acceptor 5 is located within the central 3' ss cluster and is frequently recognized and spliced. It was estimated that approximately 80% of all 3' splicing reactions in HIV-1 occur at SA5 (Purcell and Martin 1993; Stoltzfus 2009; Ocwieja et al. 2012).

1.3.1.4.2 The 4kb class of HIV-1 RNAs

All transcripts of the 4kb class of HIV-1 mRNAs are spliced from SD1 to one of the splice acceptors within the central 3' ss cluster and are therefore characterized by the absence of the Gag-Pol ORF (Purcell and Martin 1993). The 4kb class retains intron 4 which contains the RRE, which gave this class the term "intron-containing" RNA class, whose nuclear export is Rev-dependent (Bohne, Wodrich, and Kräusslich 2005). The mRNAs of the 4kb class encode for the viral envelope protein Env as well as the accessory proteins Vif, Vpr, and Vpu (Purcell and Martin 1993).

Vpu and Env are encoded together on a 4kb class bicistronic mRNA, for which 16 alternative splice isoforms have been described. Those mRNAs differ in the presence of the leader exons 2 and 3, as well as exon 4cab and exon 5 (Purcell and Martin 1993) (cf. Figure 1.10C). For the generation of *vpu*/*env* mRNAs the binding of the U1 snRNP to SD4 is essential to activate the upstream splice site within the central cluster via cross exon interactions, however, the splicing at SD4 must be repressed to include the protein-coding sequences into the final transcript. Since *vpu* and *env* mRNAs require Rev for nuclear export, splicing at SD4 in the early phase of infection is crucial to synthesize Tat and Rev. In the course of infection, however, splicing at SD4 must be repressed to produce *vpu* and *env* mRNAs (Kammler et al. 2006; Asang, Hauber, and Schaal

2008). Once the bicistronic mRNA gets exported to the cytoplasm both ORFs are translated, resulting in high levels of Vpu and Env proteins (Krummheuer et al. 2007; Anderson et al. 2007).

The *vif* mRNA is generated by a single splicing event from SD1 to SA1 and encodes for a 23 kDa protein (Purcell and Martin 1993). The expression of Vif is tightly regulated, as high levels of Vif result in the inhibition of viral protein processing and replication, therefore it is not surprising that the *vif* mRNAs make up only 1% of the total 4 kb class mRNAs (Purcell and Martin 1993; Akari et al. 2004; Mandal et al. 2009; Widera et al. 2014). The recognition of SA1 is essential for the generation of *vif* mRNA, but as described before the downstream 5' ss must be rendered splice inactive to maintain the downstream located *vif* ATG (Purcell and Martin 1993; Widera et al. 2013; Widera et al. 2014). Therefore, the levels of *vif* mRNA depend on the activation of SA1 as well as the repression of SD2, a process which is tightly regulated by the presence of several splicing regulatory elements (Exline, Feng, and Stoltzfus 2008; Kammler et al. 2006; Widera et al. 2013; Widera et al. 2014; Brillen, Walotka, et al. 2017).

Finally, the *vpr* mRNAs are generated by splicing from SD1 to SA2, with one alternative splice isoform which can include exon 2 (Purcell and Martin 1993). As with *vpu*/*env* and *vif* mRNAs, the translational start codon of *vpr* is located in the downstream intron, hence, SA2 must be recognized, and the downstream SD3 is rendered splicing incompetent (Erkelenz, Poschmann, et al. 2013; Widera et al. 2013).

1.3.1.4.3 The 9kb class of HIV-1 RNAs

As mentioned before, the unspliced full-length 9 kb class of HIV-1 mRNAs not only serves as the viral genome but also encodes for the structural proteins (Gag) and the replication enzymes (Pol) of the virus. Upon translation, the *gag-pol* mRNA encodes for the Pr55 and Pr160 polyproteins (Bohne, Wodrich, and Kräusslich 2005).

1.3.1.5 Nuclear export of HIV-1 mRNAs and translation of structural proteins

As mentioned before, eukaryotic mRNAs are transported from the nucleus to the cytoplasm via the export proteins NXF1/Tap through the NPC (cf. chapter 1.1.3). For HIV-1, only the fully spliced mRNAs of the 2kb class leave the nucleus via this conserved non-karyoprotein export receptor pathway (Purcell and Martin 1993; Köhler and Hurt 2007). The intron-containing RNAs of the HIV-1 4kb and 9kb class are exported by a second group of mobile transport receptors, which belong to the family of karyopherins. Karyopherins can act as importin or exportin, depending on the cargo recognition sequence (Görlich and Kutay 1999). These recognition sequences are short peptide sequences with function either as nuclear localization signal (NLS) or as nuclear export signal (NES) (Pemberton, Blobel, and Rosenblum 1998; Görlich and Kutay 1999). For example, the viral proteins Tat and Rev both encode an NLS, which enables the proteins to be imported into the nucleus (Henderson and Percipalle 1997; Truant and Cullen 1999).

Once back imported to the nucleus, Tat promotes the transcription of the viral LTR promoter, while Rev binds to the Rev-responsive element (RRE), located within intron 4 of unspliced (9kb) and intron-containing (4kb) HIV-1 RNAs. The interaction of Rev with the RRE results in the multimerization of up to six Rev units at the RRE (Malim et al. 1989; Daugherty, D'Orso, and Frankel 2008). This RRE-Rev interaction further results in the recruitment of the hypermethylation enzyme PIMT, which modifies the 7-methylguanosine (m7G) cap to a trimethylguanosine (TMG) cap. That TMG cap allows the HIV-1 RNAs to get recognized by the chromosome maintenance region 1 (CRM1, also known as exportin-1), which belongs to the family of karyopherin transport receptors, and binds the C-terminal domain of Rev, which encodes the leucine-rich NES. That interaction with CRM1 then allows the export of the bound RNA-RRE-Rev complex from the nucleus to the cytoplasm and the subsequent translation of viral proteins (Fischer et al. 1995; Fischer et al. 1999; Yedavalli and Jeang 2010).

Translation of the HIV-1 glycoprotein Env from the bicistronic *vpu*/*env* mRNA takes place at the rough endoplasmic reticulum (RER). As mentioned before, the translation of Env is mediated by a leaky scanning mechanism, in which the upstream localized *vpu* translational start codon gets only poorly recognized due to a suboptimal Kozak surrounding, resulting in translation initiation at the downstream located *env* ATG (Krummheuer et al. 2007). Upon translation, the Env precursor polypeptide gp160 partially translocates into the lumen of the RER due to the N-terminal located signal-sequence (Sundquist and Kräusslich 2012). The complete translocation into the RER-lumen is, however, stopped by a hydrophobic sequence in the transmembrane domain (TMD), which serves as a membrane anchor. Thereby the future cytoplasmic tail is already located in the cytoplasm, while the future extracellular part is located in the lumen of the RER (Checkley, Luttge, and Freed 2011). During the translation process, the N-terminal signal sequence gets removed and degraded by cellular peptidases in the RER, and oligosaccharide chains are attached to the extracellular domain via N-glycosidic bonds. The subsequent trimerization of the individual gp160 precursor proteins causes the further transport to the Golgi complex, where the gp-160 precursor protein is cleaved into the surface glycoprotein gp120 and the transmembrane protein gp41 by cellular furin or furin-like proteases (cf. Figure 1.8) (Hallenberger et al. 1992; Checkley, Luttge, and Freed 2011).

The viral full-length RNA is either translated into the 160 kDa Gag-Pol polypeptide, or into the 55 kDa Gag precursor protein. The Gag precursor protein consists of the matrix (MA), capsid (CA), nucleocapsid (NC), and p6 domains, as well as the two spacer peptides, SP1 and SP2. Those Gag precursor proteins serve as the structural binding blocks of the later virions and are produced in large amounts. The viral enzymes, such as the viral protease, the reverse transcriptase, and integrase, are encoded within the Gag-Pol polypeptide. Even though these are equally important for the new virus particles, they are required in much smaller amounts. The genomic organization of HIV-1 genes on the full-length RNA allows the high expression of

Gag protein relative to the enzymes encoded downstream of Gag in the Pro and Pol genes (Freed 2015). To synthesize the 160 kDa Gag-Pol polypeptide from the full-length RNA, the stop codon of the gag gene must be overread, which results in the translation of the complete gene complex. This overreading is achieved by a so-called "slippery sequence", which provokes pausing of the translocating ribosome and results in a ribosomal frame shift event in 5-10% of all translation events. The slippery sequence consists of six uridines (UUUUUUA), followed by a hairpin structure (Girnary et al. 2007). If the ribosome translates this poly-U region correctly, the Gag translational stop codon, which is located approximately 200 nucleotides downstream, is recognized efficiently resulting in the 55 kDa Gag precursor protein (Pr55Gag). In the case of ribosome slipping across the slippery sequence, all six uridines will be recognized, but the adenine at the hairpin stem will be detected twice. This results in a changed open reading frame downstream of the poly-U region and translation elongation over the *pro* and *pol* genes, resulting in the translation of the 160 kDa Gag-Pro-Pol fusion protein. This mechanism ensures the production of approximately ten to twenty Gag molecules for each Gag-Pro-Pol polyprotein, maintaining a constant Gag to Gag-Pol ratio ensuring the right structural organization and the subsequent infectivity of the progeny virions (Jacks et al. 1988; Girnary et al. 2007).

1.3.1.6 Virus assembly, budding, and maturation

Upon translation of the Gag and Gag-Pol proteins, they translocate to the host plasma membrane, where they assemble into an immature hexameric network that packages the viral genomic RNA dimers and assembles with the membrane-bound heterotrimeric gp120/gp41 Env complexes (Rein 2019). In general, packaging of the viral plus-stranded RNA genome is driven by the NC domain of Gag which interacts with the viral full-length RNA, via the Ψ-packaging signal which is located in the 5' UTR (Lever et al. 1989; Chen et al. 2020). The NC furthermore contributes to the incorporation and placement on the tRNA[Lys] primer and reverse transcription (Sundquist and Kräusslich 2012). It has been proposed, that the 5' end of the viral RNA undergoes a conformational switch from a structure that favors translation to one that promotes packaging (Lu et al. 2011). One HIV-1 RNA dimer is the recognition unit for packaging, thereby ensuring that two copies of full-length viral RNA are packaged into each newly assembling virion. Additionally, the presence of two genomic RNAs per virion provides the opportunity for recombination during reverse transcription and allows reverse transcription to proceed if one of the RNA copies is damaged (Nikolaitchik et al. 2013).

Once all necessary components for virus assembly are located at the plasma membrane the p6 domain recruits the endosomal sorting complex required for transport (ESCRT), which catalyzes the membrane fission step by bringing the membrane sites into close contact to complete the budding process, which gets finally mediated by membrane fusion (Henne, Buchkovich, and Emr 2011; Sundquist and Kräusslich 2012). During the final buddying steps, no more Gag proteins can be incorporated into the virions, resulting in an unclosed lattice of Gag precursor proteins

in the buddying virions, hence a gap in the Gag layer at the point where the viruses are strangulated from the cell membrane (Carlson et al. 2008; Briggs and Kräusslich 2011). Upon buddying, the released virions are still immature particles, which must undergo maturation to become infectious viruses. Upon release, the surface of each new virion contains about ten trimers of the viral Env protein and about two-thirds of the inner viral membrane is covered with Gag (Carlson et al. 2008; Sundquist and Kräusslich 2012).

The viral maturation begins with self-cleavage of the viral protease, which subsequently cleaves the Gag-Pol polyproteins into the active enzymes RT and IN, and the Gag precursor into the mature Gag proteins MA, CA, NC, and p6. This cleavage of the Gag polyprotein results in the disassembly of the immature membrane-bound Gag lattice and to the assembly of the mature conical capsid (cf. Figure 1.8). Upon successful maturation, the two single-stranded viral RNA copies are located within the capsid, together with the viral proteins RT and IN (Pornillos, Ganser-Pornillos, and Yeager 2011; Freed 2015).

2 Material and Methods

2.1 Cloning of recombinant DNA vectors

Within this work different cell cultures were transfected with recombinant DNA expression vectors, to measure gene expression intensities, modulation of the NMD activity, as well as alternative splicing. Therefore, different recombinant DNA vectors had to be modified, a process enabled by different cloning strategies.

2.1.1 Cloning PCR

The primarily used method in this work to alter the nucleotide sequence of a given DNA plasmid was the polymerase chain reaction (PCR). Either one or two specific primers were designed for the respective PCR reaction, those primers contained the desired altered nucleotide sequence, as well as enough complementarity to the given plasmid sequence to allow primer annealing. Furthermore, those cloning primers contained restriction enzyme recognition sites at their 5' ends to ensure the compatibility of the new insertion site of the plasmid backbone of interest.

Cloning PCR reactions were carried out in 0.2µl PCR tubes (Starlab) in a final volume of 50µl. To set up the reaction, 2µl of the template plasmid (diluted 1:1000, approximately 1ng/µl) were mixed with 39.5µl ddH$_2$O, 5µl 10x Expand™ High Fidelity Buffer (with 15mM MgCl$_2$, Roche), 1µl dNTP mix (10mM dATP, dCTP, dGTP, and dTTP, Qiagen), 1µl undiluted forward primer (100pmol, Metabion), 1µl undiluted reverse primer (100pmol Metabion), and 0.5µl Expand™ High Fidelity DNA polymerase (3.5U/µl, Roche). After a short spin down (Sprout Plus Minicentrifuge, Biozym), the tubes were placed into the PCR cycler (Professional Biometra TRIO Thermocycler, Analytik Jena), the lid was tightly closed, and the following program was started.

	Temperature	Time [mm:ss]	
Initial denaturation, activation of the "hot-start" polymerase	94°C	03:00	
Denaturation	94°C	00:30	34x
Primer binding	53-57°C	01:00	
Elongation	72°C	01:00	
Final elongation	72°C	10:00	

2.1.1.1 Separation of DNA fragments on a 1% agarose gel

The success of the PCR amplification was tested by gel electrophoresis on a 1% agarose gel (0.5g LE Agarose, Biozym; 50ml 1x TBE (89mM Tris-borate (pH 8), 2mM EDTA, Sigma-Aldrich)). The gel was microwaved (800W, 30-90 seconds) until the agarose was fully dissolved. Two drops of ethidium bromide (EtBr solution 0.025% in dropper bottle, 250µg/ml, Roth) were added and the gel was poured into the prepared chamber. The chamber was filled with 1x TBE buffer until the gel and the electrodes were covered. In the case of the cloning PCR, 2µl of the total 50µl PCR

reaction were mixed with 1µl loading dye and 7µl ddH$_2$O. As a size marker, a DNA ladder (GeneRuler 1kb DNA-ladder, Thermo Fisher) was used. The gel electrophoresis was carried out at constant 70mA for about 10-30 minutes, followed by visualization of the PCR products by UV light (312nm, INTAS®UV-Systems). If the PCR bands had the expected band size the remaining 48µl of each PCR sample were either cleaned up or stored in the freezer at -20°C.

2.1.1.2 PCR clean-up

PCR products were cleaned up using the "Monarch® PCR & DNA Cleanup Kit (5µg)" from NEB. The remaining 48µl PCR product were mixed with 240µl DNA Cleanup Binding Buffer (Ratio 5:1 for dsDNA <2kb, NEB), loaded onto the provided column and centrifuged for 1 minute at 13,000rpm (Eppendorf Centrifuge 5430). The flow-through was discarded, the column re-inserted into the collection tube, followed by two washing steps using 200µl DNA Wash Buffer (NEB). To make sure the column was dry and had no contact with residential wash buffer, the column was centrifuged one more time, before it was transferred into a 1.5ml tube (SafeSeal reaction tube, Sarstedt). To elute the bound DNA, 17µl of ddH$_2$O were added to the center of the column and incubated for one minute, followed by 1-minute centrifugation at 13,000rpm. The elution volume was about 16µl and was directly used for the subsequent digestion with restriction enzymes.

2.1.2 Digestion

The restriction digestion of both, plasmid backbone and amplicon, was carried out in individual 1.5ml reaction tubes (SafeSeal, Sarstedt), each in a final volume of 20µl. The digestion was carried out in a heating block (Thermomixer comfort, Eppendorf) for 30-120 minutes. In general, all the used restriction enzymes were purchased from NEB (20U/µl) and the reactions were set up using the optimal buffers and temperatures, which were calculated with the help of the NEBcloner® website (https://nebcloner.neb.com/#!/redigest). The digestion of plasmid and insert were prepared as follows:

Digestion plasmid backbone		Digestion PCR insert	
DNA plasmid	1µg	Cleaned up PCR	16µl
NEB buffer	2µl	NEB buffer	2µl
NEB restriction enzyme I	1µl (20U/µl)	NEB restriction enzyme I	1µl (20U/µl)
NEB restriction enzyme II	1µl (20U/µl)	NEB restriction enzyme II	1µl (20U/µl)
ddH$_2$O	ad 20µl		20µl

After digestion, the DNA fragments were cleaned up before the backbone and insert were ligated. While the plasmid backbone had to be separated before from the cut-out insert, the digested PCR product could be cleaned up using the same method as before, with the "Monarch® PCR & DNA Cleanup Kit (5µg)" from NEB. This time the starting sample size was 20µl and was therefore mixed with 100µl DNA Cleanup Binding Buffer (Ratio 5:1 for dsDNA <2kb, NEB) by pipetting up and down. The other steps were executed as described above and the

elution of the DNA was again done by the addition of 17µl ddH$_2$O to the center of the column. To make sure that the PCR product was still present within the eluted sample, 2µl of the elution was mixed with 7µl ddH$_2$O and 1µl loading dye and was again used for gel electrophoresis on a 1% agarose gel.

2.1.2.1 Separation of digested DNA plasmids on a 0.8% low-melt agarose gel

The digested plasmid backbone was separated from the cut-out insert by gel electrophoresis using low-melt agarose (Sieve GP Agarose, Biozym). DNA-containing gel-fragments, which were cut out from this gel, were directly used for ligation without further purification, after melting of the gel fragment.

The low-melt gel (0.4g low-melt agarose (Sieve GP Agarose, Biozym), 50ml 1x TB$\frac{1}{10}$E buffer (89mM Tris borate (pH 8), 0.2mM EDTA)) was boiled in the microwave, like the agarose gel, but in contrast, it was poured at 4°C upon the addition of two drops of ethidium bromide (EtBr solution 0.025% in a dropper bottle, 250µg/ml, Roth). The gel electrophoresis was carried out at constant 35mA for about 30-60 minutes, depending on the size of the backbone and insert. Afterward, the 1x TB$\frac{1}{10}$E buffer was carefully removed, the gel was put onto a UV table and the DNA fragments were visualized by UV light (312nm, INTAS®UV-Systems). The backbone fragments of the expected sizes were cut out using a sharp, clean scalpel (Feather®) and transferred individually into a clean 1.5ml tube (SafeSeal reaction tube, Sarstedt). The excised plasmid backbone was directly used for ligation, after melting of the low-melt fragment for 10 minutes at 65°C (Thermomixer comfort, Eppendorf). Then, the gel fragments were stored in the freezer at -20°C.

2.1.3 Ligation

For efficient ligation, an ATP-dependent T4 DNA ligase (5U/µl, Thermo Fisher) was added to catalyze the ligation reaction. The setup of the ligation mixture was as follows: The low-melt gel fragment containing the plasmid backbone of interest was melted in a heating block at 65°C for 10 minutes (Thermomixer comfort, Eppendorf) and was mixed by flicking off the tube. For ligation 14µl ddH$_2$O were mixed with 1µl of the melted plasmid backbone, 2µl of the purified PCR product, 2µl 10x T4 DNA ligase buffer (100mM MgCl$_2$, 100mM DTT, 10mM ATP, 500mM Tris-HCl pH 7.5, Thermo Fisher) and 1µl T4 DNA ligase (5U, Thermo Fisher). The ligation setup was mixed by careful up and down pipetting and incubated for one hour at room temperature or overnight in a heating block at 16°C. Afterward, the ligation mix was directly used for the transformation of replication-competent *E. coli* cells or stored at -20°C.

2.2 Alternatives to the classical cloning procedures

2.2.1 Site-directed mutagenesis (SDM)

Using the Q5 polymerase a whole DNA plasmid can be amplified in a single PCR reaction, with the possibility of encoding substitutions, deletions, or insertions within the used primers. This method is especially useful when the region, which needs to be altered, lacks unique restriction sites, and therefore cannot be targeted by traditional cloning strategies.

The Q5-SDM was carried out in 0.2µl PCR tubes (Starlab) in a final volume of 50µl. The desired template plasmid was diluted 1:100 and 50ng of the plasmid were mixed with 1.5µl forward primer (1:10, 10µM, Metabion), 1.5µl reverse primer (1:10, 10µM, Metabion), 1µl dNTP mix (10mM, Qiagen), 10µl 5x Q5 reaction buffer (NEB), and 1µl Q5 high-fidelity DNA polymerase (2U/µl, NEB). After a short spin down (Sprout Plus Minicentrifuge, Biozym), the tubes were placed into the PCR cycler (Professional Biometra TRIO Thermocycler, Analytik Jena), and the following program was started.

	Temperature	Time [mm:ss]	
Initial denaturation	98°C	01:00	
Denaturation	98°C	00:30	
Primer binding	50-72°C	00:30	35x
Elongation	72°C	00:30/kb	
Final elongation	72°C	20:00	

The optimal primer annealing temperature was calculated using the NEB Tm calculator (www.neb.com/Tmcalculator), with a final primer concentration of 300nM. When the PCR was finished, the polymerase had amplified the whole plasmid, which remained linearized. To remove the initial plasmid DNA, which served as template during the PCR, the samples were mixed with DpnI, a restriction enzyme that only digests methylated DNA. Therefore, each PCR sample was mixed with 1µl DpnI (20U/µl, NEB) and incubated at 37°C for at least 30 minutes. Afterward, the PCR product was separated on a 1% agarose gel, the respective band was excised using a clean, sharp scalpel (Feather®) and purified using the "QIAquick Gel Extraction Kit" (Qiagen) and an elution volume of 30µl. Before the ends of the still linearized plasmid were ligated, they were phosphorylated by the addition of T4 polynucleotide kinase (T4 PNK). Therefore, 5µl of the purified PCR product were mixed with 12µl ddH$_2$O, 2µl of the ATP containing 10x T4 DNA ligase buffer (Thermo Fisher) and 1µl of the T4 PNK (10U/µl, Thermo Fisher), and incubated at 37°C for 30 minutes, followed by addition of 0.5µl ATP-dependent T4 DNA ligase (5U/µl Thermo Fisher), for subsequent ligation of the newly phosphorylated ends, and incubation for 1 hour at room temperature, or 16°C over-night. When the ligation was finished, the plasmid was ready for transformation into chemically competent *E. coli*.

44

2.2.2 Oligo annealing

If the desired nucleotide modification of a plasmid was located within a short fragment, which was surrounded by unique restriction sites, the respective complementary oligonucleotides were just annealed and inserted into the linearized plasmid backbone The complementary oligos were purchased from Metabion, 10µl of each undiluted primer (100µM each) were mixed in a 1.5ml tube (SafeSeal reaction tube, Sarstedt) and incubated for 3 minutes at 95°C (Thermomixer comfort, Eppendorf). Then, the heating block was switched off, and the primer mixture was left in the heating block for another 45 minutes, allowing the mixture to cool down slowly and the complementary sequences to anneal. The primer mixture was diluted 1:10 and 1:50 with water and those dilutions were used to set up the ligation reaction. 6µl of each primer dilution were mixed with 2µl of the digested backbone of interest, 1µl T4 DNA ligase (5U/µl, Thermo Fisher) and 11µl ddH$_2$O and incubated for 1 hour at room temperature.

2.3 Plasmid transformation into replication-competent *E. coli* cells

The ligated DNA plasmids were transformed into replication-competent *E. coli* cells, which amplified the plasmid and expressed the encoded ampicillin antibiotic resistance gene (ß-lactamase). The kanamycin-resistant *E. coli* strain DH5αF'IQ (Invitrogen) was used for plasmid transformation. The bacteria were stored in 20µl aliquots at -80°C in 1.5ml tubes (SafeSeal reaction tube, Sarstedt). For each transformation, one bacterial aliquot was thawed on ice, mixed with 6µl of the ligation mixture by gentle flicking against the tube, followed by incubation on ice for 15 minutes. To enhance the DNA uptake the bacteria were "heat-shocked" for 90 seconds at 42°C in a heating block (Thermomixer comfort, Eppendorf), followed by another incubation on ice for 5 minutes.

Each sample was mixed with 800µl of Luria Broth Base Medium (LB-Medium, Invitrogen) without antibiotics and placed into a spinning wheel (40 cycles per minute) at 37°C for 1.5-2 hours. Afterward, 400µl of the bacteria mixture were spread onto ampicillin-containing agar plates (100µg/ml, Roche) and incubated overnight at 37°C. Only bacteria that had taken up the plasmid and were therefore expressing the antibiotic-resistant gene were able to replicate and visible as white colonies.

2.4 Quantitative plasmid preparation (DNA-Mini-Prep)

Colonies were picked individually using a fresh pipette tip and transferred into a test tube containing 5ml of ampicillin-containing LB-medium (100µl/ml, Roche), followed by overnight incubation at 37°C in a spinning wheel (40 cycles per minute).

On the next day, 2ml of the 5ml liquid bacteria culture were transferred into individual 2ml tubes (SafeSeal reaction tube, Sarstedt) for the isolation of plasmid DNA, while the remaining 3ml were stored at 4°C. The bacteria were pelleted by centrifugation in a table centrifuge for 1 minute at

14,000rpm (Heraeus Megafuge 8R, Thermo Fisher). The supernatant was discarded, and the bacterial pellets were resuspended in 300µl pre-chilled buffer 1 (50mM Tris-HCl (pH 7.5), 10mM EDTA, 400µg/ml RNase A) by scratching of the closed tube over a metal grid. The bacteria were subsequently lysed by the addition of 300µl buffer 2 (0.2M NaOH, 1% SDS), multiple inversions of the tubes, followed by incubation for 5 minutes at room temperature. The alkaline lysis was stopped by the addition of 300µl pre-chilled buffer 3 (3M KAc (pH 5.5)), mixed again by inversion of the tubes. To pellet the cellular components of the bacteria, the samples were centrifuged for 15 minutes at 14,000rpm and 4°C, and the supernatant was transferred into fresh 1.5ml tubes (SafeSeal reaction tube, Sarstedt) containing 600µl isopropanol and vortexed. The plasmid DNA was precipitated by centrifugation for 30 minutes at 14,000rpm and 4°C, the supernatant was aspirated, and the pellet was washed two times with 200µl 70% ethanol. The DNA pellet was air-dried and resuspended in 50µl Millipore ddH$_2$O.

To analyze the obtained DNA plasmids, they were digested with specific restriction enzymes, which allowed a visible distinction between the original plasmid used as the backbone for the cloning reaction and the obtained new plasmid. The enzymes were chosen by reference to the *in-silico* build model of the plasmids using the Geneious software (Geneious Version R9.1, (https://www.geneious.com)). The software also provided the size of the expected band sizes, which were later compared to the observed bands after gel electrophoresis of the digested fragments.

For the analysis 4µl of the DNA mini preparation were digested with one, two, or up to three different restriction enzymes (20units/µl each, NEB) at once. The digestion was carried out in a 1.5ml reaction tube (SafeSeal reaction tube, Sarstedt) in an end volume of 20µl with the respective buffer for the used enzymes (used NEB 10x buffers: cut smart, buffer 2.1, or 3.1). For the digestion 4µl of the plasmid DNA were mixed with 0.3µl of each restriction enzyme, 2µl of the corresponding buffer, and Millipore ddH$_2$O to a final volume of 20µl. The digestion was carried out in a heating block (Thermomixer comfort, Eppendorf) for 1 hour at the temperature optimum of the used enzymes (mostly 37°C). Afterward, 10µl of the digestion mixture were mixed with 3µl of loading dye and transferred onto an ethidium bromide containing 1% agarose gel, where the DNA fragments were separated by size applying constant 70mA for about 30 minutes. To identify positive clones, a DNA size ladder (GeneRuler 1kb DNA-ladder, Thermo Fisher) was used to determine the size of the separated fragments by comparing them to the before-predicted fragment sizes. If the isolated plasmid DNA restriction pattern fitted the expectation, the remaining 3ml liquid bacteria culture of the corresponding sample were used to inoculate 100ml of Ampicillin-containing LB medium (100µl/ml, Roche) and incubated overnight at 37°C in a shaker.

2.5 Qualitative plasmid preparation (DNA-Midi-Prep)

To obtain a greater amount of purified plasmid DNA the DNA-Midi-Kit from Qiagen was used to isolate the plasmids from the 100ml overnight liquid bacterial culture. Before the bacteria were centrifuged and the supernatant was discarded, 700µl of each liquid culture were mixed with 300µl glycerin (100%) to generate glycerin stocks, which were stored at -80°C and used to inoculate a new over-night culture.

Each 100ml bacterial culture was equally split into two conical 50ml tubes (Greiner Bio-One™) and centrifuged at 4,000rpm and 4°C for 20 minutes (Eppendorf centrifuge 5810R). The supernatants were discarded, and the bacteria of both falcons reunited by resuspension with 4ml pre-chilled buffer 1 (50mM Tris-HCl (pH 7.5), 10mM EDTA (pH 8), 400µg/µl RNase A; Qiagen Midi-Kit). The bacteria were lysed by the addition of 4ml buffer 2 (0.2M NaOH, 1% SDS; Qiagen Midi-Kit), inversion of the tubes, and incubation for 5 minutes at room temperature. To stop the alkaline lysis 4ml of buffer 3 (3M KAc (pH 5.5); Qiagen Midi-Kit) were added to the mixture and again mixed by inverting the tube multiple times, followed by a 15-minute incubation at 4°C.

For the extraction of the DNA plasmids, the provided Qiagen-tip 100 (Qiagen Midi Kit) was used. After equilibration of the tip with 4ml QBT buffer (750mM NaCl, 50mM MOPS (pH 7), 15% Isopropanol (v/v), 0.15% Triton X-100 (v/v); Qiagen Midi-Kit) the lysed bacterial supernatants were put through a filter (595 1/2 Folded Filters, Whatman™ GE Healthcare), which was set on the equilibrated tip. When the supernatant had passed through filter and tip, the filter was removed, and the tip was washed two times with 10ml buffer QC (1M NaCl, 50mM MOPS (pH 7), 15% Isopropanol (v/v); Qiagen Midi-Kit), and the flow-through was discarded.

To elute the bound plasmid DNA the tip was placed on a round centrifugation plastic tube (Nalgene), which was filled with 3.5ml isopropanol, and 5ml buffer QF (1.25M NaCl, 50mM Tris-HCl (pH 8.5), 15% Isopropanol (v/v); Qiagen Midi-Kit) were added onto the tip.

The eluted and precipitated plasmid DNA was pelleted by centrifugation for 30 minutes at 10,000rpm and 4°C (Sigma laboratory centrifuge 3K30). The supernatant was discarded, and the pellet was washed two times with 2ml 70% ethanol followed by centrifugation for 10 minutes at 10,000rpm and 4°C each. Finally, the pellet was air-dried at room temperature, re-dissolved in 100µl TE buffer (pH 8), and transferred into a 1.5ml tube (SafeSeal reaction tube, Sarstedt). The concentration of the isolated plasmid DNA was measured at a wavelength of 260nm and 280nm using the NanoDrop 1000 spectral photometer (ND-1000 Version-3.7.0) and set to a final concentration of 0.5-1µg/µl by dilution with TE buffer (pH 8). The plasmids were stored at -20°C.

2.6 Plasmids and cloning strategies

The NMD reporter plasmids pCI-TPI-WT-4H and pCI-TPI-PTC40-4H were kindly provided by Dr. Nora Diehl (Diehl 2016), while the dual-luciferase plasmids SV40-RL-β-Globin-WT - SV40-FL$_{ref}$, as

well as SV40-RL-β-Globin-NS39 - SV40-FL$_{ref,}$ were cloned by Aljoscha Tersteegen during his master's thesis and kindly provided for this work (Tersteegen 2017).

The pNLA1 plasmid is a derivate of pNL4-3 and was kindly provided by Prof. Dr. Klaus Strebel (Strebel, Klimkait, and Martin 1988), and the pNLA1 env0 RRE variant from Prof Dr. Heiner Schaal (Schaal et al. 1995).

The *tat* expression plasmid SVctat was used as donor for the SV40 promoter sequence (Schaal et al. 1993; Krummheuer et al. 2001)

The respective cloning strategies can be found in the appendix in Table 6.1, while the corresponding cloning primers are also described in the appendix, in Table 6.2.

2.6.1 Gene strands

The used gene strands were ordered and synthesized from Eurofins genomics. They were delivered as dry dsDNA fragments and were ready for cloning upon resuspension in nuclease-free ddH$_2$O. The sequences of the ordered gene strands are given in the appendix, in Table 6.3.

2.7 Eukaryotic cell culture

Within this thesis different human cell lines were used for analysis. Those cell lines were either growing adherently (HeLa, HeLa T4$^+$, TZMbl, HEK293T, HEK293T CD4$^+$, Ghost CXCR4$^+$ CD4$^+$) in T75 cell culture flasks, in 12ml Dulbecco's Modified Eagle Medium (DMEM medium, Gibco) supplemented with 10% fetal calf serum (FBS Supreme, PAN™ Biotech) and 1% penicillin/streptomycin solution (Pen Strep, Gibco, 10,000U/ml Penicillin, 10,000µg/ml Streptomycin), or in suspension (Jurkat, PM1), using upstanding T75 cell culture flasks and 15ml RPMI medium 1640 (1x) + GlutaMAX™ (Gibco), as well supplemented with 10% FCS and 1% penicillin/streptomycin. All cell lines used in this work were kept in incubators at 37°C and 5% CO$_2$ (Heraeus BBD 6220, Thermo Fisher). To maintain the cells and preventing them from overgrowing the cell cultures were split twice a week and put into new flasks with fresh medium.

2.7.1 Adherent cell lines

To split the adherently growing cells lines, the medium was removed, and the cells were carefully washed two times with 10ml of room tempered phosphate-buffered saline (PBS, Gibco). To detach the cells from the bottom of the flask 1.5ml 0.05% Trypsin-EDTA (Gibco) was given into the flask, spread evenly over the cell lawn, and removed again. After incubation between 1-5 minutes at RT the cells were loosened from the bottom by multiple stokes against the flask and finally resuspended in 10ml fresh, pre-warmed DMEM medium (10% FCS, 1% Pen/Strep), which stopped the trypsin-induced digestion. The cell number was either determined with a Neubauer counting chamber and 0.5-2ml cell suspension (depending on the confluency of the cell-lawn before the trypsin incubation) were transferred into a new T75 cell flask, supplied with fresh

DMEM medium (10% FCS, 1% Pen/Strep), spread evenly by gentle movements of the flask and put into the incubator (37°C, 5% CO_2).

2.7.1.1 HeLa cells

The HeLa cell line is the oldest and most common immortal cell line used in scientific research. It was derived from the cervical epithelial carcinoma cells, which were unknowingly taken from 31-year-old patient Henrietta Lacks in 1951. HeLa cells helped with the development of the polio vaccine in 1953 (Scherer, Syverton, and Gey 1953). The genome of HeLa cells was completely sequenced and published in 2013 (Landry et al. 2013).

2.7.1.2 HeLa T4+ cells

HeLa T4+ cells were derived from HeLa cells and stably express the CD4 receptor on their surface, which was introduced by retrovirus-mediated gene transfer. Due to the CD4 receptor expression, HeLa T4+ cells can be infected by HIV-1 and exhibit syncytium formation upon infection. HeLa T4+ cell medium had to be supplemented with geneticin (G418, 500µg/ml) (Maddon et al. 1986; Mondor, Ugolini, and Sattentau 1998).

2.7.1.3 TZMbl cells

The TZMbl cell line is a HeLa cell line derived from the parental cell line JC.53, which expresses high amounts of CD4, CCR5, and CXCR4. TZMbl cells express separate integrated copies of the luciferase gene, as well as the β-galactosidase gene, both under control of the HIV-1 LTR promoter. TZMbl cells are highly sensitive to infections with diverse HIV-1 strains, both macrophage- (M-) and T-cell- (T-) tropic, and provide tools for infection control and titer determination (Platt et al. 1998).

2.7.1.4 HEK293T cells

Human embryonic kidney (HEK) 293T cells derived from kidney cells of an aborted healthy female fetus. The cells were transfected with adenovirus 5 DNA by Frank Graham in his 293rd experiment to produce a stable cell line (Graham et al. 1977). Another specialty is the expression of a stably transfected plasmid encoding a temperature-sensitive mutant of the SV40 large T-antigen (HEK293T), allowing replication of transfected plasmids which contain an SV40 origin of replication (ORI) (DuBridge et al. 1987). Within this work, HEK293T cells were used to produce retroviral vectors, as well as HIV-1 virus stocks.

2.7.1.5 HEK293T CD4+ cells

HEK293T CD4+ cells derived from HEK293T cells, which were made susceptible for HIV-1 infection by transfection with a CD4 expression plasmid (Reil et al. 1994)

2.7.1.6 Ghost CXCR4$^+$ CD4$^+$ cells

The Ghost cell line derived from human osteosarcoma cells, that stably expressed CD4. Ghost cells express GFP under the control of the HIV LTR promoter. A second transduction with a vector expressing the human CXCR4 gene was used to generate the Ghost CXCR4$^+$ cell line. Ghost cells can be used as HIV indicator cell line (Morner et al. 1999).

2.7.2 Suspension cell lines

Due to the upright storage of the suspension cells in the incubator, the cells gathered in the bottom corner of the flask. To split the suspension cell population, approximately 12ml of the 15ml old RPMI medium were removed and discarded. The cells were gently resuspended and 12ml of new, pre-warmed RPMI medium (10% FCS, 1% Pen/strep) was added. To wash the suspension cells and to remove dead cells from the flask, the complete cell culture was transferred from its flask to a conical 50ml tube (Greiner Bio-One™) and the cells were pelleted by a gentle centrifugation step for 3 minutes and 400g (Eppendorf centrifuge 5810 R), the old medium was removed and the cells were washed with 10ml room tempered PBS. After a second centrifugation, the PBS was discarded and the cell pellet resuspended in pre-warmed RPMI medium (10% FCS, 1% Pen/Strep), transferred to a new T75 flask, and stored upright in the incubator (37°C, 5% CO$_2$).

2.7.2.1 Jurkat cells

The Jurkat cell line is an immortalized T lymphocyte cell line with round cells, which grow in suspension. Jurkat cells were established from the peripheral blood of a 14-year-old boy with acute lymphoblastic leukemia (ALL) in 1976. The initial cell line was contaminated with mycoplasma and the cure of this infection resulted in the Jurkat E6-1 clone (Weiss, Wiskocil, and Stobo 1984).

2.7.2.2 PM1 cells

The suspension cell line PM1 was derived from the Hut78 T-cell line. The cells are CD4 positive and can efficiently be infected with a broad range of HIV-1 isolates, including primary isolates as well as macrophage- (M-) and T-cell- (T-) tropic isolates (Lusso et al. 1995).

2.7.3 Freezing and thawing of cells

To obtain a higher number of cells for freezing, the cells were kept in T175 cell culture flasks until they were confluent. The medium was removed, the cells were washed twice with PBS and incubated with 5ml trypsin + EDTA. The cells were resuspended in 13ml medium, transferred into conical 50ml tubes (Greiner Bio-One™), and centrifuged at 500g for 5 minutes. The supernatant was discarded, the cell pellet was resuspended in 11ml freezing medium with DMSO (Gibco) and transferred in 1ml aliquots into prepared 2ml cryovials (Nunc). The vials were put into an isopropanol-filled freezing container and stored at -80°C for 24 hours before they were transferred into liquid nitrogen.

Cells, which are stored at -196°C in liquid nitrogen, can easily be reactivated by thawing of the cell aliquot and resuspension of the cells and their freezing medium in 10ml pre-warmed medium. After transfer into a conical 50ml tube (Greiner Bio-One™) the cells were centrifuged for 5 minutes at 500g (Eppendorf centrifuge 5810 R), the supernatant was discarded and the cell pellet was resuspended in 5ml fresh medium supplied with 10% FCS and 1% Pen/Strep, transferred into T25 cell culture flasks and incubated at 37°C and 5% CO_2.

2.7.4 Determination of the cell number

The cell number of a given cell suspension was determined with a Neubauer counting chamber. Therefore, 15µl of the cell suspension was mixed with 15µl trypan blue stain 0.4% (Gibco) and loaded into the counting chamber (C-Chip, Neubauer Improved, NanoEnTek). The cells, which were located within the four big squares of the counting chamber were counted using a top-view microscope (Nikon Eclipse TS100) with a 40x magnifier lens and a click counter. For the determination of the cell number only the living cells were counted, as trypan blue stains dead cells. The cell number was calculated by the following equation:

$$\frac{total\ cell\ count}{number\ of\ counted\ sqares} * dilution\ factor * 10^4 = cells\ per\ ml$$

Adherent cells were seeded into 6-well plates (TPP®, Merck) for transfection experiments. Therefore, the required volume of cell suspension was mixed with fresh DMEM medium (10% FCS, 1% Pen/Strep) to an end volume of 500µl. This suspension was then added to the 6-well plate, in which each well was already supplied with 2ml DMEM (10%FCS, 1% Pen/Strep). The plates were gently waved back and forth to make sure that the cells are evenly spread throughout the well. The finished 6-well plates with an end-volume of 2.5ml were then placed in the incubator at 37°C and 5% CO_2.

To calculate the required volume of the cell suspension for a given number of cells the equation was modified as given below, here for a cell count of 250,000 cells per well:

$$\frac{1000µl}{cells\ per\ ml} * 250,000 = volume\ of\ cell\ suspension\ per\ well\ [µl]$$

If suspension cells were counted e.g. for infection experiments, those cells were not kept in 6-well plates but put into conical 15ml tubes (Greiner Bio-One™) in their corresponding RPMI medium for up to 6 hours and then transferred into T25 cell culture flasks.

2.7.5 Transfection of eukaryotic cell lines

Within this work human cell lines were transfected with different DNA plasmids, to either analyze the expression of the encoded genes, to detect differences in the splicing patterns, to generate lentiviral particles containing a gene of interest, or to produce infectious HI-viruses (HIV strain NL4-3). Furthermore, NMD-specific genes were downregulated by RNAi transfection.

For those complex experimental setup's different transfection agents and their corresponding protocols were used.

2.7.5.1 TransIT®-LT1 Transfection Reagent (Mirus)

DNA plasmids were transfected using the Mirus (TransIT®-LT1) transfection reagent. The transfection reagent was used in the ratio 2:1, meaning that for each µg of plasmid 2µl of Mirus were used. First, the calculated volume of plasmid, corresponding to the desired concentration (0.5µg-2µg), was pipetted into a fresh 1.5ml tube (SafeSeal reaction tube, Sarstedt) and mixed with 1µg of the transfection control plasmid (e.g. pXGH5). A second 1.5ml tube was prepared for each transfection, in which the transfection reagent (1µl-4µl) was mixed with serum-free DMEM medium under the sterile bench to a final volume of 100µl. This mixture was incubated for 5 minutes under the bench and then added to the plasmid-containing reaction tube. Mirus, medium, and plasmids were mixed by up and down pipetting and incubated for 15 minutes at room temperature. After the incubation time, the transfection mixture was carefully pipetted onto the prepared cells in 6-well-plates, gently distributed by spinning of the plate, and placed into the incubator at 37°C and 5% CO_2.

2.7.5.2 Polyethylenimine (PEI, Sigma-Aldrich)

Transfection experiments with polyethyleneimine (PEI) were performed for the generation of NL4-3 virus stocks, as well as to produce lentiviral particles. The PEI stock solution (100mg/ml) was diluted with PBS (Thermo Fisher) to an end concentration of 1mg/ml. This PEI dilution was then used to generate the transfection master mix, by mixing 675µl PEI (1mg/ml) with 14.3ml of serum-free DMEM medium. Depending on the scope of the experimental setup one 1.5ml tube (SafeSeal reaction tube, Sarstedt) was prepared for each transfection and filled with 9-15µg of plasmid DNA, depending on the transfection, and mixed with 500µl of the prepared PEI master mix. This mixture was incubated at room temperature for 30 minutes, while the medium of the cells was changed in the meantime and replaced with the minimal volume of fresh DMEM medium (10% FCS and 1% Pen/Strep). Enough medium to cover the cells, but to minimize the volume of medium on the cells during the transfection. When the incubation time was over the PEI/plasmid mix was pipetted into the T175 flask and spread evenly across the cell lawn by gentle movements of the flask. Then, the cells were put into the incubator at 37°C and 5% CO_2. The medium of the cells was changed on the next day and replaced with fresh, pre-warmed IMDM (Lonza) supplied with 10% FCS and 1% Pen/Strep.

2.7.5.3 Lipofectamine™ RNAiMAX Transfection Reagent (Thermo Fischer)

For transfection of small interfering RNAs (siRNAs) the Lipofectamine RNAiMAX transfection reagent was used. Under the sterile hood, two 1.5ml tubes (SafeSeal reaction tube, Sarstedt) were prepared for each transfection. In the first tube 248µl Opti-MEM (Gibco) were mixed with 1.5µl of the desired siRNA (either the UPF1 siRNA or the non-specific control (NSC) siRNA). In

the second tube, 248µl Opti-MEM were mixed with 2.5µl RNAiMAX transfection reagent. Both mixtures were combined within the first tube and incubated for 15 minutes at room temperature. Then the mixture was carefully pipetted onto the prepared cells in 6-well-plates and placed in the incubator at 37°C and 5% CO_2. The medium of the cells was changed on the next day and replaced with fresh, pre-warmed DMEM (10% FCS, 1% Pen/Strep).

2.7.6 siRNAs

The small interfering RNAs (siRNAs) were ordered from Eurofins and delivered in dry condition. They were resuspended in the provided 5x siMAX resuspension buffer (30mM HEPES, 100mM KCl, 1mM $MgCl_2$, pH 7.3) which was diluted with nuclease-free ddH_2O. The end concentration of the siRNAs was set to 20µM, they were stored in 10µl aliquots at -20°C. The sequence of the used siRNAs, which were ordered as 19-base-sequence-duplexes with "dTdT" overhangs was as following:

UPF1 - siRNA
sense: GAUGCAGUUCCGCUCCAUU - dTdT
antisense: dTdT - CUACGUCAAGGCGAGGUAA

Non-Specific Control 47% GC - siRNA
sense: AGGUAGUGUAAUCGCCUUG- dTdT
antisense: dTdT - UCCAUCACAUUAGCGGAAC

2.7.7 Inhibitors

2.7.7.1 Cycloheximide

Cycloheximide (CHX) was used as a translational inhibitor during this work. CHX was first reported in the 1950ties as a naturally occurring fungicide. When CHX enters a eukaryotic cell, it rapidly blocks translation elongation by binding to the E-site of the 60S large ribosomal subunit and interference with the residing deacetylated tRNA. This interference blocks the dissociation of the tRNA within the E-site and therefore the movement of the ribosomal subunit (Schneider-Poetsch et al. 2010). The used cycloheximide stock solution had a concentration of 50mg/ml and was stored at -20°C. It was kindly provided by Dr. Nora Diehl (Diehl 2016). Once one 1ml aliquot was thawed it was aliquoted in 20µl steps to avoid multiple thawing and freezing steps. For translational inhibition using cycloheximide, the stock solution was added to the cell culture medium (end concentration 50µg/ml) and incubated at 37°C and 5% CO_2 for six hours.

2.7.7.2 Ruxolitinib

Ruxolitinib is a direct inhibitor of the Janus-associated kinases (JAK) 1, 2, and 3, and thereby inhibits the Jak-STAT signal transduction, which would otherwise be activated upon interferon (IFN) stimulation. In this work ruxolitinib (Ruxo) was used as control during experiments with recombinant IFN and was added to the cell medium 10 minutes before the IFN treatment. Ruxo aliquots with a concentration of 20mM were kindly provided by Prof. Dr. Mirko Trilling and

stored at -20°C. The working condition was 4µM, which was the 1,200-fold the 50% inhibitory concentration [IC50] (Le-Trilling et al. 2018).

2.7.8 Recombinant Interferons

The treatment of cells with human recombinant interferons (IFNs) was done by the addition of the respective volume of IFN to the cell culture medium. The interferons had a stock volume of 100U/µl, were aliquoted a 100µl by Dr. Frank Hillebrand, and stored at -80°C. The used working concentration was 500U/ml so that 5µl recombinant IFN were added per ml cell culture medium.

Recombinant Human Interferon	Name	Catalog No	Lot No	Specific activity	Accession No	c_{Stock}
Alpha 2a	Hu-IFN-α-2a	11100-1	6146	$6.48 \cdot 10^8$ units/mg	V00549	100U/µl
Beta 1a	Hu-IFN-β-1a	11415-1	6391	$2.03 \cdot 10^8$ units/mg	V00543	100U/µl
Gamma	Hu-IFN-γ	11500-2	6350	$8.24 \cdot 10^6$ units/mg	V00543	100U/µl

2.8 Generation of HIV-1 (NL4-3) virus stocks

To generate infectious HIV-1 particles of the HIV-1 laboratory strain NL4-3, HEK293T cells were transfected with the proviral plasmid pNL4-3. With HIV-1 being classified as retrovirus of safety level 3 (S3**) by the ZKBS (central commission for the biological safety) the production of HI-viruses had to be done in our S3 laboratory.

On the first day, 6.5×10^6 HEK293T were seeded into T175 cell culture flasks in 25ml DMEM medium supplied with 10% FCS and 1% Pen/Strep and placed overnight in the incubator at 37°C and 5% CO_2. On the next day, the cells were transfected with the proviral plasmid pNL4-3 using the transfection reagent polyethyleneimine (PEI, Sigma-Aldrich). Therefore, the PEI stock solution (0.1g/ml) was diluted with PBS (Thermo Fisher) to an end concentration of 1mg/ml. This PEI dilution was then used to generate the transfection master mix, by mixing 675µl PEI (1mg/ml) with 14.3ml of serum-free DMEM. For each T175 cell culture flask or each transfection, 9µg of plasmid DNA were pipetted into individual 1.5ml tubes (SafeSeal reaction tube, Sarstedt). The plasmid DNA was then mixed with 500µl of the PEI master mix and incubated at room temperature for 30 minutes. In the meantime, the medium of the cells was fully removed and replaced with 15ml fresh DMEM medium (10% FCS and 1% Pen/Strep) to minimize the volume of medium on the cells during the transfection. When the incubation time was over the PEI/plasmid mix was pipetted into the T175 flask and spread evenly across the cell lawn by gentle movements of the flask. Then, the cells were again put into the incubator and left there overnight at 37°C and 5% CO_2.

On the third day, the DMEM medium was removed from the cells and carefully replaced with 15ml IMDM medium (Lonza), again supplied with 10% FCS and 1% Pen/Strep and put back into the incubator for another 24 hours.

On the fourth day, the viral particles in the supernatant were harvested. Therefore, the supernatant of each T175 cell culture flask was taken and pooled together in a fresh T175 flask. The supernatants were mixed and aliquoted into conical 50ml tubes (Greiner Bio-One™), which were centrifuged for 10 minutes at 500g (Hettich), to deplete all cellular residues from the HEK293T cells. The supernatant was carefully decanted into fresh 50ml tubes and then pipetted into pre-labeled tubes (HEK293T, NL4-3, IMDM, date, name) in 500µl aliquots. The virus stock was stored at -80°C.

2.8.1 Titer determination of HIV-1 (NL4-3) virus stocks

TZM-bl cells were used to determine the titer of the HIV-1 virus stocks. Those cells express the β-galactosidase gene (*lacZ*) of *E. coli* under the control of the HIV-1 LTR promotor. This promotor gets activated when a cell gets infected with HIV-1, by binding of the viral protein Tat to the tar region within the 5' LTR. Tat binding eventually results in the expression of β-galactosidase, which can be detected using a histochemical β-galactosidase assay, also called X-gal assay. The hydrolyzation of X-gal by β-galactosidase results in an intensive, insoluble, blue color which can be detected using a light microscope.

The assay was performed in 96-well plates. Therefore, 6,000 TZM-bl cells were seeded per well in 350µl DMEM medium, supplied with 10% FCS and 1% Pen/Strep, and kept in the incubator at 37°C and 5% CO_2. On the next day, the cells were infected with progeny viruses.

The infections were done in serial dilutions, eight replicas per plate, and twelve dilution steps. The virus-containing supernatants were prepared in a deep-well 96 well plate as follows: three virus aliquots (500µl each) were taken from the -80°C freezer, thawed, pooled, and mixed. As first dilution step, the virus stock was diluted with DMEM (10% FCS, 1% Pen/Strep) in a ratio of 1:10 or 1:20, before 350µl of the dilution were transferred into each of the 8 wells of the first column (A1-H1) of the deep-well mix plate. The other 88 wells of the deep-well plate (A2-H12) were filled with 220µl DMEM (10% FCS, 1% Pen/Strep) each. To start the dilution, 110µl of viral supernatant (A1-H1) were transferred into the second column (A2-H2) using a multichannel pipette. The tips were discarded and the virus-medium mixture in column 2 was mixed using fresh pipette tips, followed by transfer of 110µl to the next column. That process was repeated until the last column (A12-H12) was reached, and the pipette tips were changed one more time before mixing the medium in the last column. Now, that the preparation of the virus dilutions in the deep-well mix plate was finished, the medium of the TZM-bl cells was removed and replaced with 200µl of each corresponding well of the mix plate. Then the cells were incubated for up to 48 hours.

Before the cells were used for the X-gal assay, they were washed and fixated. Therefore, the medium was removed, and the cells were washed twice with 200µl pre-cooled PBS per well. To fixate the cell lawn, each well was covered with 100µl pre-cooled fixing solution (0.25%

55

glutaraldehyde, 0.8% formaldehyde in PBS) and incubated at 4°C for 10 minutes. The fixing solution was removed, and the cell lawn was again washed twice with 200µl PBS. Afterward, 100µl of staining solution (0.4mg/ml X-gal, 4mM $K_3[Fe(CN)_6]$, 4mM $K_4[Fe(CN)_6]$, 2mM $MgCl_2$ in PBS) was pipetted onto the cells and incubated at 37°C for a minimal duration of 4 hours. Then the cells were ready for the microscopic analysis, involving the detection and counting of wells that contained blue-colored cells.

2.8.1.1 Calculation of TCID$_{50}$

The tissue culture infection dose (TCID) represents the volume of viral supernatant, that is required to infect half of the cell number (TCID$_{50}$). The TCID$_{50}$ of the produced virus stock was calculated by the following formula:

$$TCID_{50}/ml = \frac{D_S^{\left(\frac{N}{R}+0.5\right)} \cdot D_0 \cdot 1000}{D_S \cdot V}$$

D_S = Dilution factor; N = Number of infected wells; R = Number of replicas of each dilution; D_0 = Dilution factor of the first dilution; V = Volume of virus dilution per well [µl]

The dilution factor of the consecutive dilutions (D_S) was three, while the dilution factor of the first dilution (D_0) was either 10 or 20. The total number of infected wells (N) was the result of the X-gal assay, with eight replications for each dilution (R). The volume of virus dilutions per well (V) was 200µl, which were transferred from the deep-well mixing plate onto the TZM-bl cells.

For the infection experiments in this work, a consistent multiplicity of infection (MOI) of 0.5 was used, to maintain the comparability between different experiments and virus stocks. The MOI indicates the number of viral particles per cell and is calculated by plaque-forming units per ml (PFU/ml). As HIV-1 does not produce plaques during an infection, the PFU/ml was estimated by multiplication of the acquired TCID$_{50}$/ml with 0.7, as the PFU/ml is approximately 0.7 times the TCID$_{50}$/ml. Therefore, the required volume of virus stock to infect a given number of cells can be calculated by the following formula:

$$V_{Virus\ stock}[ml] = \frac{MOI \cdot cell\ number}{TCID_{50}/ml \cdot 0.7}$$

2.9 Production of lentiviral vectors

Using pseudotyped lentiviruses, a broad range of cells can be analyzed, even cells that are hard to transfect. Another advantage is the stable integration of the lentiviral genome into its host DNA, resulting in the expression of the gene of interest (GOI) for a long period of time, as well as propagation to daughter cells. In general, the production of lentiviral vectors resembles the

experimental setup for HIV-1 (NL4-3) virus stocks, with the major difference being the transfected plasmid(s). Here, for the analysis of the nonsense-mediated mRNA decay, four different lentiviral stocks were produced. One pair encoded the WT version of the TPI (TPI-WT-4H), or the PTC containing version (TPI-PTC40-4H). The other pair encoded either the WT version of renilla luciferase fused to β-Globin (β-Globin-WT) or the PTC containing variant of renilla luciferase fused to β-Globin (β-Globin-NS39), here also both encoded the WT firefly luciferase, which was used as the internal reference.

On the first day, 3×10^6 HEK293T were seeded into each T75 cell culture flask, in 10ml DMEM medium supplied with 10% FCS and 1% Pen/Strep. After overnight incubation at 37°C and 5% CO_2, the cells were transfected with the mixture of the needed plasmids (5µg of each plasmid) using the transfection reagent polyethyleneimine (PEI, Sigma-Aldrich). The used plasmids were the following:

- pCD NL-BH, encoding for the viral RNAs *gag*, *pol*, *tat,* and *rev*

- pzVSV-G, encoding for the glycoprotein of VSV (vesicular stomatitis virus)

- puc2CL7, encoding for the GOI

    o puc2CL7EGwo TPI-WT-4H mod.

    o puc2CL7EGwo TPI-PTC40-4H mod.

    o puc2CL7EGwo RL-β-Globin-WT(-) FL_{ref}(-)

    o puc2CL7EGwo RL-β-Globin-NS39(-) FL_{ref}(-)

For the transfection, the PEI stock solution (100mg/ml) was diluted with PBS to an end concentration of 1mg/ml. This PEI dilution was then used to generate the transfection master mix, by mixing 30µl PEI (1mg/ml) with 470µl of serum-free DMEM. For each T75 cell culture flask or each transfection, 5µg of each plasmid DNA of the above mentioned, were pipetted into individual 1.5ml tubes (total of 15µg). The plasmid DNA was then mixed with the prepared 500µl PEI master mix and incubated at room temperature for 30 minutes. In the meantime, the medium of the cells was fully removed and replaced with 10ml fresh DMEM medium (10% FCS and 1% Pen/Strep). When the incubation time was over the PEI/plasmid mix was pipetted into the corresponding T75 flask and spread evenly across the cell lawn by gentle movements of the flask. Then, the cells were again put into the incubator and left there overnight at 37°C and 5% CO_2. On the third day, the DMEM medium was removed from the cells and carefully replaced with 10ml IMDM medium, again supplied with 10% FCS and 1% Pen/Strep and put back into the incubator for another 24 hours. On the fourth day, the lentiviral particles in the supernatant were harvested. Therefore, the supernatant of each T75 cell culture flask, from the same transfection setup, was taken and pooled together in a 50ml tube. The tubes were centrifuged

for 10 minutes at 500g (Eppendorf centrifuge 5810 R), to deplete all the cellular residues from the HEK293T cells. The supernatant was carefully decanted into fresh, labeled 50ml tubes and then pipetted into pre-labeled tubes in 500µl aliquots and stored at -80°C.

2.10 Total RNA isolation

To analyze the transcripts expressed from the transfected plasmids, the abundancies of the produced RNAs, or the splicing pattern of the genes, total RNA was isolated from eukaryotic cells. To remove all medium residues from the cells they were washed twice with PBS before they were lysed in 500µl solution D (SolD; 4M guanidinium thiocyanate, 25mM sodium citrate, 0.5% sarcosyl, 0.1M β-mercaptoethanol) per 6-well (approximately 10^6 cells per well). Adherent cells were additionally detached by gently scraping off the well with a cell scraper (cell scraper s, 240mm, TPP) and the mixture was transferred into a 1.5ml tube (SafeSeal reaction tube, Sarstedt). The RNA isolation was either interrupted at this point, by freezing the samples at -20°C or carried on directly by proceeding to the phenol/chloroform extraction. The following steps were carried out under a hood, the tubes and reagents were kept on ice and the centrifugation steps were carried out in a pre-cooled centrifuge at 4°C. Each tube was supplied with 7.2µl β-mercaptoethanol (Sigma Aldrich), 50µl 2M sodium acetate (pH 4), and 500µl phenol (Roti®-Aqua-Phenol, Roth), prepared as a master mix. Afterward, 100µl chloroform/IAA (24:1) was added to each sample and mixed by vortexing for 15 seconds, during which a white coloring could be observed. To separate the phases the samples were kept on ice for 15 minutes followed by 20-minute centrifugation at 10,000rpm and 4°C (Heraeus Megafuge 8R, Thermo Fisher). The upper, watery phase (2x 200µl) was then transferred into a fresh 1.5ml tube filled with 400µl isopropanol, vortexed, and stored at -20°C for 1 hour to overnight for RNA precipitation. To pellet the precipitated RNA the tubes were centrifuged in a pre-cooled centrifuge at 10,000rpm and 4°C for 30 minutes. The supernatant was carefully removed using a vacuum pump and the pellet was washed twice with 130µl 70% ethanol, followed by centrifugation for 10 minutes at 10,000rpm and 4°C. After removal of the ethanol, the tubes were kept on ice, with an open lid for 15 minutes to dry the RNA pellets. The RNA was resuspended in 10µl of DEPC-ddH$_2$O and the concentration of the isolated RNA was measured by photometric analysis using the NanoDrop 1000 spectral photometer (ND-1000 Version-3.7.0). Isolated RNA was stored at -80°C.

2.10.1 Northern blot analysis

With this technique, the total RNA of a given sample is separated by size, transferred, and cross-linked to a membrane. Using specific probes, which are complementary to the RNA of interest, it is possible to visualize those RNAs. This technique makes it possible to measure RNA abundancies without further amplification and offers great comparability between two samples.

For northern blot analysis, 3-5µg of total RNA were separated on a 100ml denaturing 1% agarose gel under the fume hood (1g agarose powder (Biozym LE Agarose), 85ml ddH$_2$O, 10ml 10x MEM

(200mM MOPS, 50mM Sodium acetate, 10mM EDTA, pH 7), 5.5ml formaldehyde (37%, Rotipuran®, Roth)). For the sample preparation the respective volume of each RNA was transferred into fresh 1.5ml tubes (SafeSeal reaction tube, Sarstedt) and mixed with 1µl recombinant DNase I (10U/µl, Roche) in a total volume of 7µl. The DNase digestion was carried out for 20 minutes at 37°C (Thermomixer comfort, Eppendorf), followed by 10 minutes at room temperature. Then, the samples were mixed with 7µl of 2x RNA loading dye (Thermo Fisher), incubated at 70°C for 10 minutes, and then put on ice. Each pocket of the prepared denaturing agarose gel was now loaded with 13.5µl of the RNA mixture and separated at 60V for 30 minutes up to 180 minutes.

Following separation, the ribosomal RNA bands were visualized on a UV table (312nm, INTAS®UV-Systems) followed by an overnight blotting step. The blot was set up in a plastic bowl from bottom to top in the following order: approx. 20 paper towels, which were halved and placed on another, three dry Whatman™ papers (3 MM CHR, GE Healthcare), one Whatman™ paper, pre-incubated with 20x SSC (3M NaCl, 300mM tri-sodium-citrate), the positively charged nylon membrane (Roche), pre-incubated with ddH$_2$O and then with 20x SSC. The denaturing RNA agarose gel, which was also washed with ddH$_2$O and then with 20x SSC, and three pre-incubated 20x SSC Whatman™ papers. Two plastic bowls filled with 20x SSC were put on either side of the build blotting setup and were connected by a Whatman™ paper stripe, reacting from one reservoir over the blot to the other reservoir. A plastic plate was put on top of the blotting setup and weight was put on the plate. The capillary blotting was done overnight.

The next day, the blot setup was dismantled, the membrane was cut to the size of the gel and for orientation, one corner of the blot was removed. The RNA was UV cross-linked to the membrane (CL-1000 Ultraviolet Crosslinker, UVP, Energy: 1200x100µJ/cm^2), and the large and small rRNAs bands were marked with a pencil, to serve as size standard. The human 18S ribosomal subunit ran at a size of 1.9kb, and the human 28S ribosomal subunit at 5.0kb (Thermo Fisher, Ribosomal RNA Sizes). The membrane was transferred into a hybridization bottle (Thermo Fisher) and washed twice with ddH$_2$O. After 2h of pre-hybridization with 10ml 1x DIG Easy Hyb hybridization solution (Roche) at 55°C (hybridization oven, Biometra OV 5, 5rpm), the membrane was hybridized with specific digoxigenin (DIG)-labeled PCR probes (DIG-11-dUTP alkali-labile; Roche) overnight at 55°C.

On the next day, the membrane was washed twice with ddH$_2$O and two times with stringent wash buffer I (2x SSC, 0.1% SDS) at room temperature, followed by two 20-minute washing steps in stringent wash buffer II (0.2x SSC, 0.1% SDS) at 68°C in the UVP hybridization oven. After two additional washing steps with ddH$_2$O, the membrane was taken from the hybridization bottle, transferred into a plastic box, washed with maleic acid buffer (0.1M maleic acid, 150mM NaCl, pH 7.5), and blocked with 1x northern blot blocking reagent (Roche) dissolved in maleic acid

buffer for 45 minutes. Anti-digoxigenin-AP, Fab fragments (sheep, Roche) were diluted 1:20,000 in 1x blocking solution and incubated for 1h at RT. The membrane was washed three times with maleic acid buffer (10 minutes each), and the RNA bands were visualized by using CDP star (Roche) for chemiluminescent reactions (1:100 in AP buffer [0.1M Tris HCl, 0.1M NaCl, pH 9.5]; Roche). The blots were visualized using the Lumi-Imager™ F1 (INTAS).

2.10.1.1 DIG-labeled PCR probes

To produce specific digoxigenin (DIG)-labeled PCR probes, which can be used for the detection of complementary RNAs in a northern blot, two PCR reactions were necessary. First, a regular PCR with unlabeled dNTPs (Qiagen), using plasmid DNA with the desired sequence as template, followed by a second PCR with specific digoxigenin (DIG)-labeled PCR probes (DIG-11-dUTP alkali-labile; Roche) using the purified PCR product of the first round as template.

The first PCR reaction was carried out in 0.2µl PCR tubes (Starlab) in a final volume of 50µl. Therefore, 2µl of the template plasmid (diluted 1:1000, approximately 1ng/µl) were mixed with 39.5µl ddH$_2$O, 5µl 10x Expand™ High Fidelity Buffer (with 15mM MgCl$_2$, Roche), 1µl dNTPs (10mM, Qiagen), 1µl forward primer (1:10, 10pmol, Metabion), 1µl reverse primer (1:10, 10pmol, Metabion), and 0.5µl Expand™ High Fidelity DNA polymerase (3.5U/µl, Roche). After a short spin down (Sprout Plus Minicentrifuge, Biozym), the tubes were placed into the PCR cycler (Professional Biometra TRIO Thermocycler, Analytik Jena), and the following program was started.

	Temperature	Time [mm:ss]	
Initial denaturation, activation of the "hot-start" polymerase	95°C	03:00	
Denaturation	95°C	00:30	
Primer binding	57°C	01:00	35x
Elongation	72°C	01:00	
Final elongation	72°C	10:00	

To verify the success of the first PCR, and to verify the size of the PCR product, 2µl of the PCR reaction were separated on a 1% agarose gel. The rest of the PCR was cleaned up using the "Monarch® PCR & DNA Cleanup Kit (5µg)" from NEB, eluted with 20µl ddH$_2$O, and used as the template for the DIG-labeled PCR.

The DIG-labeled dUTPs were mixed with unlabeled dNTPs prior to the DIG-PCR reaction. The DIG-dNTPs were prepared as a 100µl stock solution and stored in 10µl aliquots at -20°C. 35µl of the DIG-labeled dUTPs (1mM DIG-11-dUTP alkali-labile; Roche) were mixed with 1µl dATP (10mM), 1µl dGTP (10mM), 1µl dCTP (10mM), 0.65µl dTTP (10mM), and 61.35µl DEPC-ddH$_2$O.

The following DIG-PCR was carried out in 0.2µl PCR tubes (Starlab) in a final volume of 40µl, using the same primer pair as for the first PCR step, as well as the same PCR program. For the PCR setup, 4µl of the prepared DIG-labeled dNTPs were mixed with 2µl of the prepared, purified PCR probe, 4µl 10x Expand™ High Fidelity Buffer (with 15mM $MgCl_2$, Roche), 2µl forward primer (1:10, 20pmol, Metabion), 2µl reverse primer (1:10, 20pmol, Metabion), 25.5µl ddH_2O, and 0.5µl Expand™ High Fidelity DNA polymerase (3.5U/µl, Roche). The DIG-PCR product was monitored by gel electrophoresis, where 1µl of the newly obtained DIG-labeled probe was separated on a 1% agarose gel, next to 1µl of the PCR template. The DIG-labeled PCR product migrated more slowly through the gel. The DIG-labeled probe was stored at -20°C.

2.10.1.1.1 DIG-labeled Probes and primer pairs

The DIG-labeled probes which were used for the northern blot analysis are given in Table 2.1 and the corresponding primer numbers and sequences in Table 2.2.

Table 2.1: The DIG-labeled PCR probes with their corresponding name, the used template plasmid DNA, the corresponding primer pairs, and obtained product sizes.

Probe name	Template DNA	Used primer pair	Product size
HIV-1 Exon 7	pNLA1	#3387/#3388	153bp
HIV-1 3' UTR	pNLA1	#5854/#5855	186bp
4x HBB	TPI-WT-4H	#5651/#5599	465bp

Table 2.2: Primer numbers and corresponding sequences used for the generation of DIG-labeled PCR probes.

Primer number	Primer sequence 5' → 3'
#3387	TTGCTCAATGCCACAGCCAT
#3388	TTTGACCACTTGCCACCCAT
#5854	GACTTTCCAGGGAGGCGTG
#5855	TGAAGCACTCAAGGCAAGCT
#5651	CAGGGACTCGAGCACC
#5599	ACCCTCACTAAAGGGAAGCG

2.11 Analysis of isolated RNAs by PCR amplification

To allow transcript-specific amplification by polymerase chain reaction (PCR) the single-stranded cellular RNAs had to be converted into complementary DNA (cDNA).

2.11.1 cDNA synthesis

To eliminate all possible DNA contaminations of the isolated RNA before cDNA synthesis 3µg of RNA were incubated with 1µl of recombinant DNase I (10U/µl, Roche) in a total volume of 10µl using 0.5ml reaction tubes (Eppendorf). The DNA digestion was carried out in a heating block (Thermomixer comfort, Eppendorf) at 37°C for 20 minutes, followed by a 10-minute incubation at room temperature. After heat inactivation of the DNase by a 5-minute incubation at 70°C, the samples were put on ice.

4.5µl of the DNase I treated RNAs were transferred into a fresh 0.5ml reaction tube and mixed with 6.5µl ddH$_2$O, 1µl dNTP mix (10mM dATP, dCTP, dGTP, and dTTP, Qiagen), and 1µl oligo(dT) primers (diluted 1:20, Roche). To denature RNA secondary structure and allow binding of the oligo(dT) primers, the samples were incubated at 65°C for 5 minutes and then placed on ice.

For cDNA amplification, the RNA, dNTP, primer-mixtures were transferred into 0.2µl PCR reaction tubes (Starlab) und mixed with 4µl 5x first strand buffer (FSB, Invitrogen), 1µl DTT (0.1M, Invitrogen), 1µl RNasin® Ribonuclease Inhibitor (40U/µl, Promega), and 1µl SuperScript™ III Reverse Transcriptase (200U/µL, Invitrogen). The synthesis of the cDNA was carried out in a PCR cycler (Professional Biometra TRIO Thermocycler, Analytik Jena) for 60 minutes at 50°C, followed by 15 minutes at 72°C. Afterward, the cDNA was either used for PCR analysis or stored at -20°C.

2.11.2 Qualitative RT-q-PCR (real-time PCR)

Real-time PCRs or qPCRs can monitor the amplification of PCR products in a quantitative way, which is measured by the intensity of fluorescence during each step of the reaction. Real-time PCRs can be done with the usage of specific probes, which bind to the DNA product of interest, or with the addition of SYBR-green to the PCR reaction mixture. SYBR green is a fluorescent dye, which intercalates into double-stranded DNA. Therefore, the ratio of fluorescence intensifies with every cycle of the PCR amplification, as the amount of double-stranded DNA doubles in each step. The number of cycles needed to pass a predefined threshold (C_T) allows the calculation of the initial amount of cDNA when the PCR result of the primer pair of interest is normalized to a reference primer pair amplifying a housekeeping gene (e.g. GAPDH or β-actin), resulting in the ΔC_T, which can be compared between different samples of the same experiment.

The real-time PCR reactions were performed in 20µl LightCycler® glass capillaries (Roche) in a final volume of 20µl. The capillaries were filled with 18µl of master mix, which was prepared individually for each primer pair (10µl of "PrecisionPLUS qPCR Master Mix" premixed with SYBR green (Primerdesign Ltd), 1µl forward primer (1:10, 10pmol, Metabion), 1µl reverse primer (1:10, 10pmol, Metabion), and 6µl ddH$_2$O). The cDNA was diluted 1:10 and 2µl of diluted cDNA was added to the capillaries. As negative control, a water sample was included in each run for each primer pair. The capillaries were closed and the qPCR reaction mix was spun down into the tip of the capillary by short centrifugation for 10 seconds at 300rpm (Eppendorf Centrifuge 5430). Afterward, the capillaries were transferred into the LightCycler®1.5 Sample Carousel (20µl, Roche), put into the LightCycler®1.5 Instrument (Roche) and the following program was started:

	Temperature	Time [mm:ss]	
Denaturation	95°C	02:00	
Amplification	95°C	00:10	42x
	60°C	01:00	
Melting curve	95°C	00:00	
	63°C	00:30	
	95°C	00:00	
Cooling	40°C	00:30	

2.11.2.1 Calculation of the relative gene expression

To calculate the relative gene expression from the obtained C_T values, the delta C_T (ΔC_T) value between the GOI and the reference gene was calculated first. That was done by subtraction of the C_T(Ref) - C_T(GOI). Then, the expression of each GOI was determined by calculating $e^{\Delta CT}$. To bring the different conditions into reference, the ratio between the untreated sample and the treated sample was determined by the division of the calculated expression of the treated sample through the untreated sample ($e^{\Delta CT}$(treated)/ $e^{\Delta CT}$(untreated)).

2.11.3 Quantitative RT-PCR (endpoint PCR)

These PCR reactions were carried out in 0.2µl PCR tubes (Starlab) in a final volume of 50µl. Therefore, 4µl of the amplified cDNA were mixed with 37.75µl ddH$_2$O, 5µl 10X PCR buffer I (Thermo Fisher), 1µl dNTP mix (10mM, Qiagen), 1µl forward primer (1:10, 10pmol, Metabion), 1µl reverse primer (1:10, 10pmol, Metabion), and 0.25µl AmpliTaq® DNA polymerase (5U/µl, Thermo Fisher). After a short spin down (Sprout Plus Minicentrifuge, Biozym), the tubes were placed into the PCR cycler (Professional Biometra TRIO Thermocycler, Analytik Jena), and the following program was started:

	Temperature	Time [mm:ss]	
Initial denaturation	95°C	03:00	
Denaturation	95°C	00:30	
Primer binding	53-57°C	01:00	26x
Elongation	72°C	01:00	
Final elongation	72°C	10:00	

2.11.4 Primer pairs for (q)PCR amplification

The primer pairs which were used for PCR and qPCR amplifications are described below in Table 2.3 with the name of the amplified gene and the primers which were used for the amplification are given with the respective primer numbers and the corresponding sequences.

Table 2.3: Primer pairs used for PCR amplification in RT-(q)-PCR reactions. Shown are the names of the amplified gene fragments, the corresponding primer numbers, as well as the primer sequences.

Amplified GOI	Primer pair	Primer sequences 5' → 3'
Actin	#6056/#6057	CATCGAGCACGGCATCGTCA / TAGCACAGCCTGGATAGCAAC
Erk2	#4646/#4647	CAACGTGTGGCCACATATTCT / CTGCTTATGATAATGTCAACAAAGTTCG
HGH	#1224/#1225	TCTTCCAGCCTCCCATCAGCGTTTGG / CAACAGAAATCCAACCTAGAGCTGCT
HIV-1 exon 1-4	#1544/#3632	CTTGAAAGCGAAAGTAAAGC / TTGATGCTTCCAGGGCTC
HIV-1 Exon 7	#3387/#3388	TTGCTCAATGCCACAGCCAT / TTTGACCACTTGCCACCCAT
SRSF2 (SC35)	#6165/#6166	CGGTGTCCTCTTAAGAAAATGATGTA / CTGCTACACAACTGCGCCTTTT
SRSF3	#4004/#4005	GCCGCATTTTTTAACCCT / CTCTCTCTTCTCCTATCTCT
SRSF7	#5784/#5785	CGCTGGCAAAGGAGAGTTAGA / ATCGCCTTCCTCTGGATCGA
SRSF7 WT	#6078/#6076	CCAGACCTAGATCTTCTGAG / CGAAGAAGAAGCAGGTCAC
SRSF7 NMD	#6078/#6077	CCAGACCTAGATCTTCTGAG / CTGCAAACTTGCAAGGTCAC
hnRNPDL WT	#6086/#6084	GTTACTTTAACATCTCCTCCTGTT / AGGGCCAAAACTGGAACCAA
hnRNPDL NMD	#6087/#6084	GCAATGAAGTTCCCGCTGTT / AGGGCCAAAACTGGAACCAA
hnRNPL	#5778/#5779	GAGCGTGAACAGTGTGCTTC / ACTGGTGGACCCATCCTTCT
RPS24	#5776/#5777	AAAGCAACGAAAGGAACGCA / ACAGTGGCCACAGCTAACAT
UPF1	#5427/#5428	CTTCCCATCCAACATCTTCTACGA / ACACAGGACAGGATGATGAAGTC
ISG15	#6054/#6055	GAGAGGCAGCGAACTCATCT / AGGGACACCTGGAATTCGTT
IRF-1	#6058/#6059	TTTGTATCGGCCTGTGTGAATG / AAGCATGGCTGGGACATCA
Plasmid TPI	#4324/#5665	TAATACGACTCACTATAGG / CTCCTTGCAGGTTGCC
Endogenous TPI	#5664/#5665	CCAGCACCATGGCG / CTCCTTGCAGGTTGCC
Lentiviral TPI	#6128/#5673	GACAGACTGAGTCGCCCG / CCCAAAGACATGCCTTCTCTCTG
Lentiviral TPI	#6128/#5665	GACAGACTGAGTCGCCCG / CTCCTTGCAGGTTGCC
Renilla	#6167/#6168	GCGTTGATCAAATCTGAAGAAGG / TTGGACGACGAACTTCACCT
Firefly	#3279/#3280	GTCCGTTCGGTTGGCAGAAGCTATG / GTCGTTCGCGGGCGCAACTG

2.11.5 Separation of DNA fragments on a 10% polyacrylamide (PAA) gel

To ensure sufficient separation of PCR amplification fragments, polyacrylamide gels were used for the analysis of quantitative RT-PCRs, which were poured horizontally between two glass plates. The PAA gel mixture was prepared in conical 50ml tubes (Greiner Bio-One™) by mixing 13.8ml ddH$_2$O, 10ml PAA-stock-solution (ROTIPHORESE®Gel 30, Roth), 6ml 5x TBE (Sigma Aldrich), and 420µl APS (Ammoniumperoxydisulfate 10%, Merck). The radical acrylamide-polymerization was started by the addition of 21µl TEMED (Tetramethylethylendiamin, Roth). The gel was mixed by vortexing and bubbles were removed by gentle swaying of the tube. Then the gel was poured, the comb was added, and the gel was left to polymerize for about 15 minutes. The gel was put vertically into the electrophoresis chamber and fixated with clips at both sides. The points of contact between the upper buffer chamber and the glass plate were sealed with agarose gel (1%, Biozym LE Agarose). After the agarose had dried, the upper and

lower chamber was filled with 1x TBE (Sigma Aldrich), the comb was removed, and the pockets were washed with 1x TBE using a syringe with a needle.

Following the PCR reaction, 10µl PCR product were mixed with 3µl of blue loading dye loaded onto the gel. As a size reference, a DNA ladder (GeneRuler 1kb DNA-ladder, Thermo Fisher) was used, and the electrophoresis was carried out at 35mA for 30-90 minutes. Then, the glass plates were separated, the gel was transferred onto a plastic bag and incubated for 5 minutes with 25ml 1x TBE and two drops of ethidium bromide (EtBr solution 0.025% in a dropper bottle, 250µg/ml, Roth). The DNA bands were visualized with the gel documentation system Gel iX Imager (312nm, INTAS®UV-Systems, Germany) and the obtained pictures were processed using the ImageJ software.

2.11.6 Isolation, reamplification, and sequencing of DNA bands from PAA gels

If the observed PCR bands had to be verified by sequencing, they were carefully cut out of the PAA gel using a clean, sharp scalpel (Feather®) and transferred into a 1.5ml tube (SafeSeal reaction tube, Sarstedt). The gel slice was incubated with 100µl diffusion buffer (0.5M NH_4Ac, 10mM MgAc, 1mM EDTA, 0.1% SDS) for 30 minutes in a heating block at 50°C (Thermomixer comfort, Eppendorf). After short centrifugation for 1 minute at 13,000rpm (Eppendorf Centrifuge 5430), the supernatant was carefully transferred into a fresh 1.5ml tube and mixed with 100µl Isopropanol. From this step onward the buffers and columns of the "QIAquick Gel Extraction Kit" from Qiagen were used. The supernatant was mixed with 300µl of buffer QG and transferred onto a QIAquick spin column in a 2ml collection tube followed by 1-minute centrifugation at 13,000rpm. The flow-through was discarded and the column was washed with 750µl buffer PE by another centrifugation step of 1 minute. The flow-through was again discarded and the spin column centrifuged one more time to remove potential buffer PE residues from the column. Afterward, the column was transferred into a fresh 1.5ml microcentrifugation tube and 30µl of buffer EB or ddH_2O were added to the center of the column, incubated for 1 minute, and then centrifuged for another minute at 13,000rpm to elute the DNA. The eluted DNA had to be reamplified by PCR using the same primer pair as for the original PCR, which resulted in the designated DNA band in the PAA gel. This time the PCR was set up like described above in chapter 2.11.3, but this time using 4µl of the purified PAA band, instead of 4µl cDNA. Additionally, the number of PCR cycles was altered from 26 to 34 repetitions.

The PCR amplification of excised PAA gel bands often resulted in multiple PCR products, especially when the excised band was not the smallest PCR product. Therefore, the whole 50µl PCR reaction volume was mixed with 5µl of loading dye and transferred into the pocket of an 1% agarose gel (0.5g LE Agarose (Biozym), 50ml 1x TBE (Sigma-Aldrich). A current of 75mA was applied and the PCR products were separated by size for about 10-30 minutes. Using a UV table,

the PCR product of the expected band size was cut out using a clean, sharp scalpel (Feather®) and transferred into a 1.5ml reaction tube. The weight of the gel fragment was determined, and the DNA was isolated from the gel by the usage of the "QIAquick Gel Extraction Kit" from Qiagen, according to the manufacturer's protocol. The DNA was eluted from the provided columns by the addition of 30µl ddH$_2$O into clean 1.5ml reaction tubes. The DNA concentration was subsequently measured by nanodrop analysis (ND-1000 Version-3.7.0) and the samples were prepared for sequencing. For PCR products smaller than 300bp, 1ng/µl and products larger than 300bp, 5ng/µl were sent to Eurofins in an end volume of 15µl (Eurofins Genomics Germany GmbH) for sequencing.

2.12 Protein analysis

2.12.1 Protein isolation from cells

To isolate proteins from human cell lines, the supernatant was removed except for 1ml, in which the cells were resuspended (suspension cell lines) or detached from the bottom of their well using a cell scraper (TPP, adherent cell lines). The cells were transferred into prepared 1.5ml tubes (SafeSeal reaction tube, Sarstedt) and pelleted by a short centrifugation at full speed (14,000rpm) for 15 seconds. The supernatant was discarded, except if proteins were to be isolated from the supernatant, then 350µl of the supernatant were transferred into a new 1.5ml tube and only the rest was discarded. The cell pellet was washed twice with 900µl PBS, each time followed by resuspension, centrifugation, and removal of the supernatant. If both, RNA and protein were to be isolated from one cell population, the resuspended cells were split equally into two 1.5ml tubes before the last centrifugation step. To isolate the proteins from the cells, the cell pellet was resuspended and lysed in 30µl RIPA-buffer (25mM Tris-HCl (pH 7.6), 150mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS, protease inhibitor cocktail (Roche)) and put into the -80°C freezer for at least 10 minutes. The frozen samples were then put into the centrifuge and centrifuged at full speed for 10 minutes, to pellet the dispensable cell residues. The supernatant, containing the cellular proteins was transferred into a fresh 1.5ml tube, mixed with 20µl 5x sample buffer (60mM Tris-HCl (pH 6.8), 24% glycerol, 2% SDS, 14.4mM β-Mercaptoethanol, 1% bromophenol blue) and incubated for 10 minutes at 95°C in a heating block (Thermomixer comfort, Eppendorf). Protein samples were stored at -20°C.

2.12.2 Protein isolation from viral supernatant

To isolate proteins from the cellular supernatant, which contained the virus in case of HIV-1 infection, 350µl of supernatant was carefully underlayed with 200µl 20% sucrose (1ml 1M Tris (pH 7.5), 3.3ml 3M NaCl, 200µl 0.5M EDTA, 20g sucrose, ad. 100ml ddH$_2$O) and centrifuged for 45 minutes at 50,000g and 4°C (Sigma laboratory centrifuge 3K30) to pellet the viruses in the supernatant. The supernatant was removed, and the virus pellet was resuspended in 20µl 2.5x sample buffer, vortexed, and incubated for 10 minutes at 95°C.

2.12.3 SDS-PAGE

For the molecular weight-dependent separation of proteins, the samples were incubated in sodium dodecyl sulfate (SDS) containing sample buffer and loaded onto polyacrylamide gels for electrophoresis (SDS-PAGE). The components of the "Mini-PROTEAN® Tetra Vertical Electrophoresis Cell" from Bio-Rad were used for pouring the hand-casted gels and the subsequent electrophoresis. The gel consisted of a 12% separation gel and a 5% stacking gel with pockets for the probes.

	Separation gel (12%)	Stacking gel (5%)
dH2O	3.38ml	3.15ml
PAA-30	3.6ml	1.13ml
0.5M Tris (pH 6.8)	-	0.9ml
2M Tris (pH 8.8)	1.9ml	-
20% SDS	45µl	34µl
60% Sucrose	-	1.58ml
APS	108µl	90µl
TEMED	18µl	9µl

Both gels were prepared without the addition of APS and TEMED, which start and catalyze the radical polymerization. The lower gel, the separation gel, had to be poured first. Therefore, APS and TEMED were added to the prepared gel and mixed by vortexing. The gel was poured with the help of a 10ml pipette until the liquid level reached about a two-finger space to the top edge of the glass plates. The separation gel was overlaid with isopropanol until the gel had polymerized. The isopropanol was removed, APS and TEMED were added to the prepared stacking gel, again the mixture was vortexed and poured on top of the separation gel using a 5ml pipette. The comb was added, and the gel was left to polymerize.

When the gel was solid, the comb was carefully removed, and the pockets were washed with 1x laemmli buffer (25mM Tris (pH 8.8), 192mM glycine, 0.1% SDS) using a syringe with a needle. Then, the gel was put into the "Mini-PROTEAN® Tetra Electrode Assembly" (Bio-Rad) and set into the tank, which was filled with 1x laemmli buffer. Before loading, the protein samples were incubated in a heating block (Thermomixer comfort, Eppendorf) at 95°C for 10 minutes. The "PageRuler Plus Prest Protein Maker V" (PeqLab) was used as molecular weight. The SDS-PAGE was carried out for about 1 hour using a current of constant 25mA for each gel (Power Pac 3000, Bio-Rad), followed by separation of the glass plates, removal of the stacking gel, and blotting of the separation gel.

2.12.4 Western blot

If the proteins on the SDS-PAGE were to be analyzed with specific antibodies they had to be transferred to a membrane (Amersham™ Protran™ 0.45µm NC, Nitrocellulose Blotting Membrane, GE Healthcare), a process called western blotting. For this step, the "Mini Trans-

Blot® Cell" from Bio-Rad was used. All used materials for blotting were pre-incubated with tank-blot buffer (laemmli buffer + 20% methanol). Gel and membrane were assembled and fixated within the "Mini Gel Holder Cassette" in the following order: black plastic grid, one foam pad, two Whatman™ papers (3 MM CH, GE Healthcare), separation gel, membrane, two Whatman™ papers, again another foam pad, and the white plastic grid. To remove all bubbles in between the gel and the membrane, a glass pipette was gently rolled over the Whatman™ papers on top of the membrane. To start the blotting process, the cassette was put into the "Mini Trans-Blot Central Core", according to the color code, and then put into the tank. The tank was filled with tank-blot buffer and a current of constant 150mA, 100V for 15 minutes, followed by constant 300mA, 100V for 20 minutes was applied (Power Pac 3000, Bio-Rad).

The success of the blotting was checked by the transference of the used protein molecular weight marker. The membrane was once washed with 1x TBST (100mM Tris-HCl (pH 8), 1.5M NaCl, 0.5% Tween20) and then incubated with 5% milk powder in TBST for 60 minutes on a shaker. The primary antibody was added (in 5% milk powder TBST) and incubated overnight on a shaker at 4°C. On the next day, the primary antibody was removed and stored in conical 50ml tubes (Greiner Bio-One™) at -20°C (the primary antibodies were used up to three times). The membrane was washed three times for 10 minutes with TBST, on a shaker at room temperature, followed by the addition of the secondary antibody (in 5% milk powder in TBST) and incubation for one hour at room temperature on the shaking plate. Afterward, the blot had to be washed again three times for 10 minutes with TBST, to remove the unbound secondary antibodies. Then the membrane was washed once with ddH$_2$O and dried with Whatman™ paper. The bands of the molecular weight marker were marked with the WesternSure® Pen (Li-Cor) and the membrane was transferred to a new plastic tray. The blot was developed by a 3-minute incubation with ECL solution (mixed 1:2, Amersham™ ECL™ Western Blotting Detection Reagents, GE Healthcare). The blots were visualized using the Lumi-Imager™ F1 (INTAS).

If necessary, the membrane was incubated with another primary antibody, therefore the membrane was washed once with TBST and then incubated with re-blot solution (Merck, mixed 1:10 with ddH$_2$O) for 10-15 minutes. After another washing step with TBST, the membrane was again blocked twice with 5% milk powder in TBST for 15 minutes each at room temperature, followed by the addition of the new primary antibody in milk powder + TBST at 4°C on a shaker over-night.

2.12.4.1 Primary antibodies

Antibody	Dilution	Company	Cat. No.
Mouse anti-actin	1:5000	Sigma Aldrich	A5316-2ML
Rabbit anti-UPF1	1:10000	Abcam	ab109363

2.12.4.2 Secondary antibodies

Antibody	Dilution	Company	Cat. No.
Rabbit anti-mouse	1:2000	Invitrogen	A27025
Goat anti-rabbit	1:2000	Thermo Fisher	A27036

2.13 Luciferase activity

Cells that had been transfected with plasmids encoding for renilla luciferase (*Renilla Reniformis*) and firefly luciferase (*Photinus Pyralis*) could be analyzed based on their luciferase activity. Therefore, the cells were harvested in specific cell lysis juice and incubated with the respective luciferase substrate. The subsequent light emission was measured and the relative light units (RLU) from the renilla luciferase were normalized to the corresponding firefly luciferase activity from the same cell population.

2.13.1 Sample preparation

The cells medium was removed, and the cells were washed twice with PBS before the lysis juice was added (500µl/Well in 6-well plates). The 2x lysis-juice (PJK GmbH) was stored at 4°C and had to be diluted 1:2 with ddH$_2$O before use. After the addition of the lysis juice, the cells were either detached from the bottom of the well using a cell scraper (cell scraper s, 240mm, TPP), or by up and down pipetting. The lysis was improved by freezing and thawing the cells at -80°C. The samples were stored in 1.5ml tubes (SafeSeal reaction tube, Sarstedt) at -80°C.

2.13.2 Measurement of luciferase activity

Before the measurement of the luciferase activities, the substrates had to be prepared. The beetle-juice reaction buffer (PJK) was mixed with the provided D-luciferin (PJK) and ATP (PJK), and stored in 2ml aliquots at -80°C. The volume of firefly substrate, which was needed for the respective experiment was taken from the freezer, thawed, and pooled in a conical 50ml tube (Greiner Bio-One™), shielded from light with aluminum foil. In contrast to the firefly substrate, the renilla substrate had to be prepared freshly for every experiment. Therefore, the calculated volume of renilla-juice reaction buffer (PJK, stored at 4°C) was mixed in a ratio of 50:1 with the substrate mixture (coelenterazine in reconstitution buffer, PJK, stored at -80°C) in a conical 50ml tube (Greiner Bio-One™), again shielded from light with aluminum foil.

The measurement of the luciferase activity was performed in flat white bottom 96-well plates (Nunc™, Thermo Fisher). Each sample was measured in four technical replicas, therefore, 20µl of each sample was pipetted into four neighboring wells, one plate for the measurement of firefly activity, and one plate for renilla activity. The measurement was performed using the Tecan Infinite® 200 machine and the i-control 1.12 software. The injection volume was 100µl, and the measurement of the luminescence had an integration time of 5000ms. In the first run,

the relative light units (RLUs) of the firefly luciferase were measured, followed by a second run in which the same samples were measured for their renilla luciferase activity.

2.13.3 Calculation of the relative luciferase activity

To calculate the relative luciferase activity from the measured relative light units (RLUs) the RLUs between the Renilla and Firefly measurements for each sample were divided ($RLU_{Renilla}/RLU_{Firefy}$) to obtain a ratio ($\Delta RLU$) for each sample. This ratio was then used to calculate the differences in the luciferase activity between the cells transfected with the WT- vs. the NS39-variant for each treatment condition, by division of $\Delta RLU_{WT}/\Delta RLU_{NS39}$. The different treatments were then normalized and set to 100% by dividing (treated/untreated) and multiplying that ratio with 100.

3 Results

3.1 Cycloheximide-dependent accumulation of HIV-1 2kb *tat* mRNAs

This work aimed to investigate the impact of the cellular nonsense-mediated decay (NMD) pathway regarding its potential antiviral capacities, and the interplay with HIV-1 upon infection. This work is based on a preliminary observation by Dr. Nora Diehl, who discovered that treatment of HIV-1 infected cells with cycloheximide (CHX), a translational inhibitor, increased the viral 2kb mRNA class in northern blot analysis (Diehl 2016).

Since it was described that the NMD-mediated degradation of target mRNAs is dependent on translation (Belgrader, Cheng, and Maquat 1993; Carter et al. 1995; Thermann et al. 1998), the most obvious explanation was that the HIV-1 2kb mRNA class might be an NMD target.

First, to test whether the HIV-1 2kb mRNA class is an NMD target and if so whether it is specific for certain cell types, peripheral blood mononuclear cells (PBMCs) from two healthy donors, as well as PM1 and Ghost cell culture cells were infected with the HIV-1 lab strain NL4-3 for 48 hours (MOI 0.5) and incubated with 50 µg/ml CHX for six hours. The total RNA was harvested and used for northern blot analysis, using a DIG-labeled HIV-1 exon 7 probe, or for RT-PCR analysis, using primers specifically amplifying the HIV-1 *tat* mRNAs (Figure 3.1).
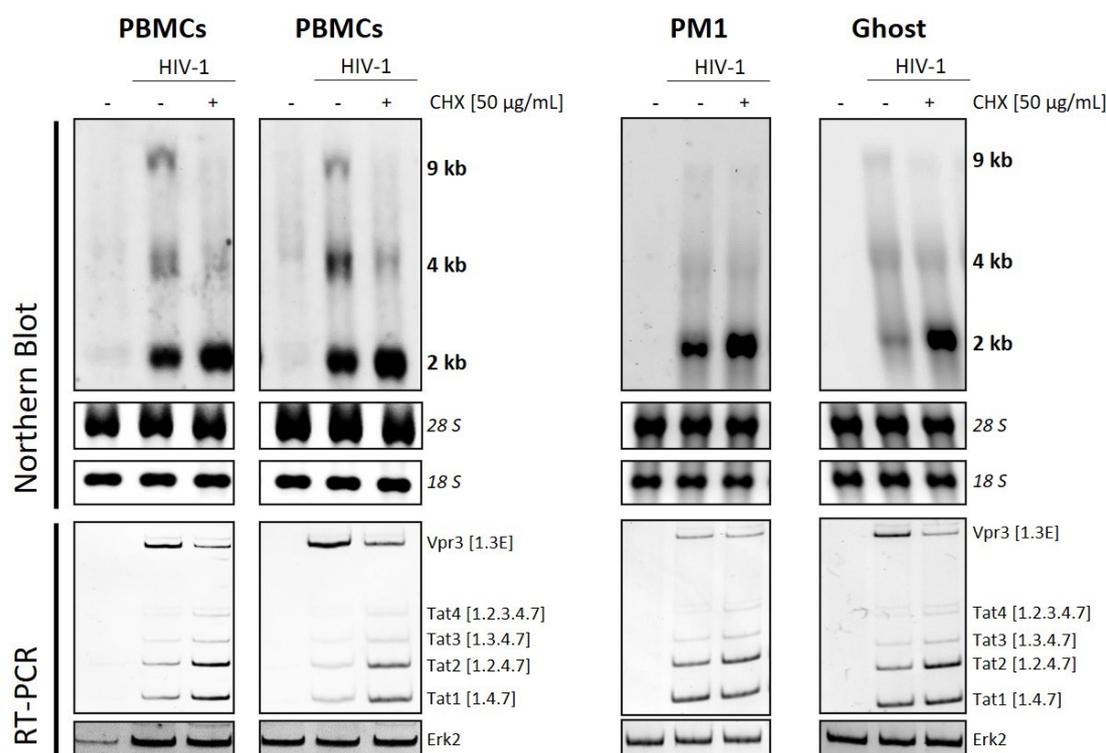

Figure 3.1: CHX-dependent accumulation of the HIV-1 2kb class and *tat*-mRNAs. Northern blot analysis of the viral mRNA classes and RT-PCR analysis of the viral splicing pattern. **(Left)** PBMCs from two healthy donors were pre-stimulated with IL-2 and infected with HIV-1. **(Right)** PM1 and Ghost cell culture cells were infected with HIV-1 (MOI 0.5) for 48 hours. Six hours before the RNA was harvested, the cells were treated with 50 µg/ml CHX. **(Top)**

Northern blot analysis. 3 µg of total RNA were separated on a formaldehyde-containing 1% agarose gel and capillary blotted overnight onto a nitrocellulose membrane. For the detection of the viral mRNA classes, a specific DIG-labeled HIV-1 exon 7 probe was used (#3387/#3388). The ribosomal RNAs (18S 1.9kb; 28S 5.0kb) served as size standard and loading control. **(Bottom)** RT-PCR analysis of the viral exon 1-4 splicing pattern. RNA from the same preparation as used for the northern blot was reverse transcribed and amplified with specific primer pairs (#1544/#3632), separated on a 10% PAA gel, and stained with EtBr. The amplification of Erk2 (#4646/#4647) served as loading control. The previously performed PBMC experiments (left) were done by Dr. Frank Hillebrand and the results were generously provided for this work.

The results in Figure 3.1 confirmed, that the observed CHX-dependent accumulation of the HIV-1 2kb mRNA class was independent of the analyzed cell type. Furthermore, the accumulation was specific for the 2kb class, as the abundance of the HIV-1 4kb and 9kb mRNA classes remained on the same level or decreased (Figure 3.1, top). The subsequent RT-PCR analysis showed that specifically the HIV-1 *tat* mRNAs accumulated upon CHX treatment, while other HIV-1 mRNAs, like *vpr3*, decreased (Figure 3.1, bottom). The *vpr* mRNAs belong to the 4kb class of HIV-1 mRNAs, hence the CHX-dependent increase in the amount of the 2kb class could be confirmed by RT-PCR.

To test whether the observed CHX-induced accumulation of the HIV-1 2kb class was restricted to infected cells, or intrinsic to the mRNA classes, HeLa cells were transfected with the proviral DNA (pNL4-3), instead of being infected with NL4-3 viruses. The total RNA was isolated 24 hours after transfection and a six-hour cycloheximide treatment and used for northern blot analysis (Figure 3.2 B, lanes 1 and 2). The CHX-induced accumulation of the HIV-1 2kb class upon transfection strongly resembled the experimental outcome of the prior infection experiments, even though the decrease of 4kb mRNAs, which was observed in the infection experiments, was not detectable. However, the consistent accumulation of the 2kb class upon CHX treatment allowed the investigation of potential NMD-triggering features within the 2kb class of HIV-1 mRNAs, which was tested by modification of the transfected plasmid DNA.

Due to its genome structure, all sequence features of the alternatively spliced 2kb class mRNAs are also present in the mRNAs of the 4kb and the 9kb class, unless an RNA secondary structure would be created via an exon-exon junction, which would serve as a signal. However, it is unlikely to find the potential element causing the CHX-dependent accumulation within the 2kb class. This potential sequence element, that protects against the NMD-induced RNA degradation, is, therefore, more likely to be found within the 4kb and 9kb classes.

What structurally differentiates the mRNAs of the 2kb class from the 4kb, and 9kb class among other things, is the overall number of splicing events, hence the subsequent number of putative deposited exon-junction-complexes on the mRNA. Therefore, to address the question of whether the possible number of deposited exon-junction-complexes affects the NMD-induced RNA degradation, the plasmid-encoded HIV-1 introns were stepwise removed to generate a fully spliced *tat* transcript, that was assumed to not be loaded with any exon-junction-complexes (this transcript consisted of the HIV-1 exons 1, 2, 4, and 7). The generated plasmids are shown in

Figure 3.2 A and the results of the transfection experiments, measured by the mRNA levels are shown in Figure 3.2 B.
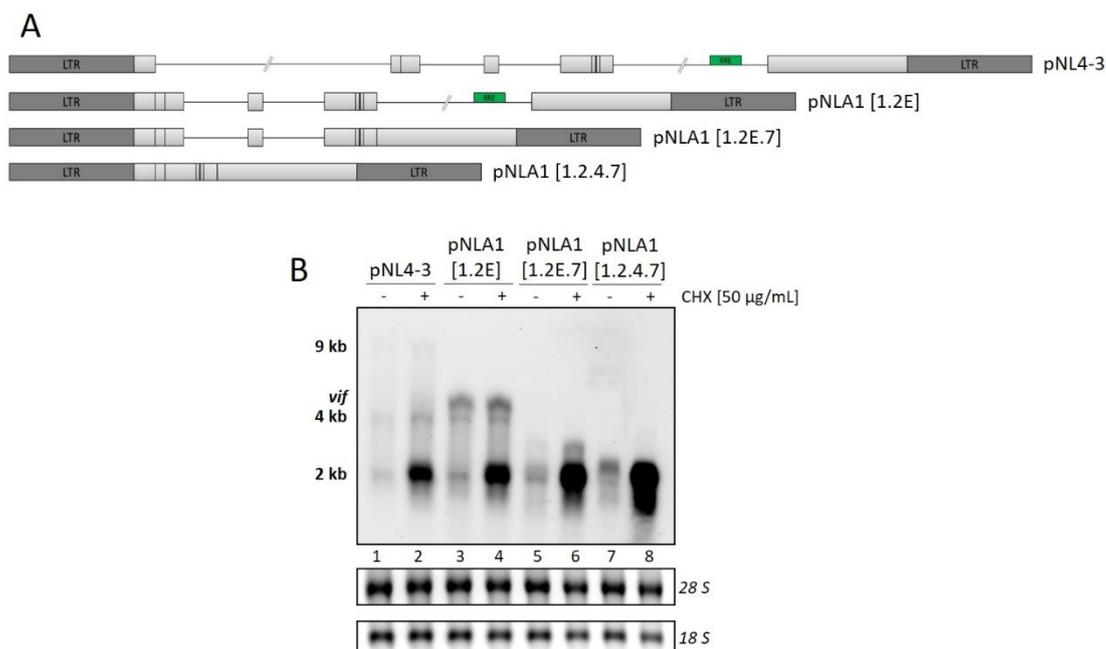

Figure 3.2: The CHX-dependent accumulation of the 2kb class is splicing-independent. (A) Schematic drawing of the used reporter constructs. Grey boxes indicate exons, black lines introns. The dark grey boxes show the viral LTRs and the green box indicates the location of the RRE within intron 4. The plasmid on the top represents the full-length HIV-1 genome, encoded on the pNL4-3 plasmid. The pNLA1 construct lacks the first intron, the exon-exon junction is indicated by the [1.2]. Removal of intron 4 generates pNLA1 [1.2E.7] and the additional removal of the remaining introns encodes a fully spliced *tat* mRNA with the HIV-1 exons 1, 2, 4, and 7. **(B)** Northern blot analysis of the constructs depicted in (A). HeLa cells were transfected with 1 µg of the indicated plasmid for 24 hours. 50 µg/ml CHX was added to the cell culture media 6 hours before harvesting. After RNA isolation 3.5 µg of total RNA was digested with DNase I and separated on a 1% formaldehyde containing agarose gel and transferred to a nylon membrane by overnight capillary blotting. For hybridization, a probe specifically detecting HIV-1 exon 7 was used (#3387/#3388). The ribosomal bands served as size standards and loading control.

As depicted in Figure 3.2 (lanes 1 and 2) the CHX-dependent accumulation of the HIV-1 2kb class was visible not only for infection experiments but was also observed when cells were transfected with the proviral plasmid DNA pNL4-3 (cf. Figure 3.1 and Figure 3.2). However, transfection of cells with the pNL4-3 plasmid resulted in the production of virus particles by the transfected cells. To investigate the impact of the first major splicing event, and to abolish virus production, the non-infectious derivate pNLA1 (Strebel, Klimkait, and Martin 1988) was used. This construct lacks the first major intron (between SD1 and SA1) of HIV-1. Due to the loss of the *gag* and *pol* messages no 9kb class could be detected by northern blot analysis (Figure 3.2, lanes 3 and 4). The ratio of the 4kb class was not altered by the addition of CHX to the media, however, after intron 1 removal, consequently, the amount of *vif* mRNA was much higher (cf. lanes 3 and 4 with lanes 1 and 2), while the CHX-dependent accumulation of the 2kb class remained similar.

To further decrease the number of splicing events, also the second major intron of HIV-1 was removed, joining SD5 and SA7. Transfection of this plasmid (pNLA1 [1.2E.7], Figure 3.2, lanes 5 and 6) consequently resulted in no detectable 4kb class mRNAs, while the CHX-dependent accumulation of the 2kb class remained. The removed intron 4 also contained the rev-responsive-element (RRE), which allowed Rev-mediated nuclear export of the 4kb class mRNAs (Malim et al. 1989). Thus, all transcripts produced from pNLA1 [1.2E.7] are classified as 2kb class mRNAs and exported to the cytoplasm independent of the interaction with Rev. Additional to the prominent accumulation of the intronless mRNAs, there was a slightly larger band, potentially mRNAs which still contained intron 2, 3, or both, which were also sensitive to CHX treatment.

Next, to exclude any splicing process, a fully spliced *tat*-transcript was generated, only consisting of HIV-1 exons 1, 2, 4, and 7 (Figure 3.2, lanes 7 and 8). Those exons were chosen based on the most prominent CHX-effect on this *tat*-message in RT-PCR experiments (cf. Figure 3.1, bottom). This pNLA1 [1.2.4.7] *tat* expression plasmid-encoded an intronless *tat* cDNA. Still, the accumulation through CHX treatment was visible, maybe even more prominent (Figure 3.2, lanes 7 and 8), indicating that the observed CHX-dependent accumulation of the HIV-1 2kb class was splicing-independent. However, additionally to the accumulation of the 2kb class upon CHX treatment, which was observed for all constructs, the 2kb mRNAs of pNLA1 [1.2E.7] and [1.2.4.7] (Figure 3.2, lanes 5-8) showed a smear upon CHX treatment, indicating some degradation of these transcripts.

Based on the prominent effect of CHX treatment on the splicing-independent pNLA1 [1.2.4.7] *tat* transcript, the potential sequence region responsible for the observed NMD-induced RNA degradation was further narrow down, by modification of the plasmid-encoded transcript in three different ways (see Figure 3.3, top). The most obvious difference between the RNAs of the 2kb class to the RNAs of the 4kb and 9kb class was the presence of HIV-1 intron 4, which contained the RRE. Since the RRE-containing 4kb and 9kb class RNAs did not show accumulation upon CHX treatment (cf. Figure 3.1 and Figure 3.2) the influence of the RRE was investigated. Therefore, the RRE sequence was re-inserted into the *tat* mRNA, within the sequence of the *tat* 3' UTR, just downstream of the *nef* ATG. The RRE generates a prominent secondary structure, that can be bound by the Rev protein (Cochrane, Chen, and Rosen 1990) and is essential for the nuclear export of intron-containing HIV-1 mRNAs. In this experimental setup the RRE was investigated without the interaction with the Rev protein, but as a structural element, since it has been shown that RNA secondary structures can stabilize RNAs, as for the example of the Rous Sarcoma Virus RNA stability element (RSE) (Weil and Beemon 2006). The comparison of CHX treatment of pNLA1 [1.2.4.7] and pNLA1 [1.2.4.7] RRE (Figure 3.3 B, lanes 5-8) showed that the addition of the RRE sequence, however, had no stabilizing impact on the *tat* mRNA.
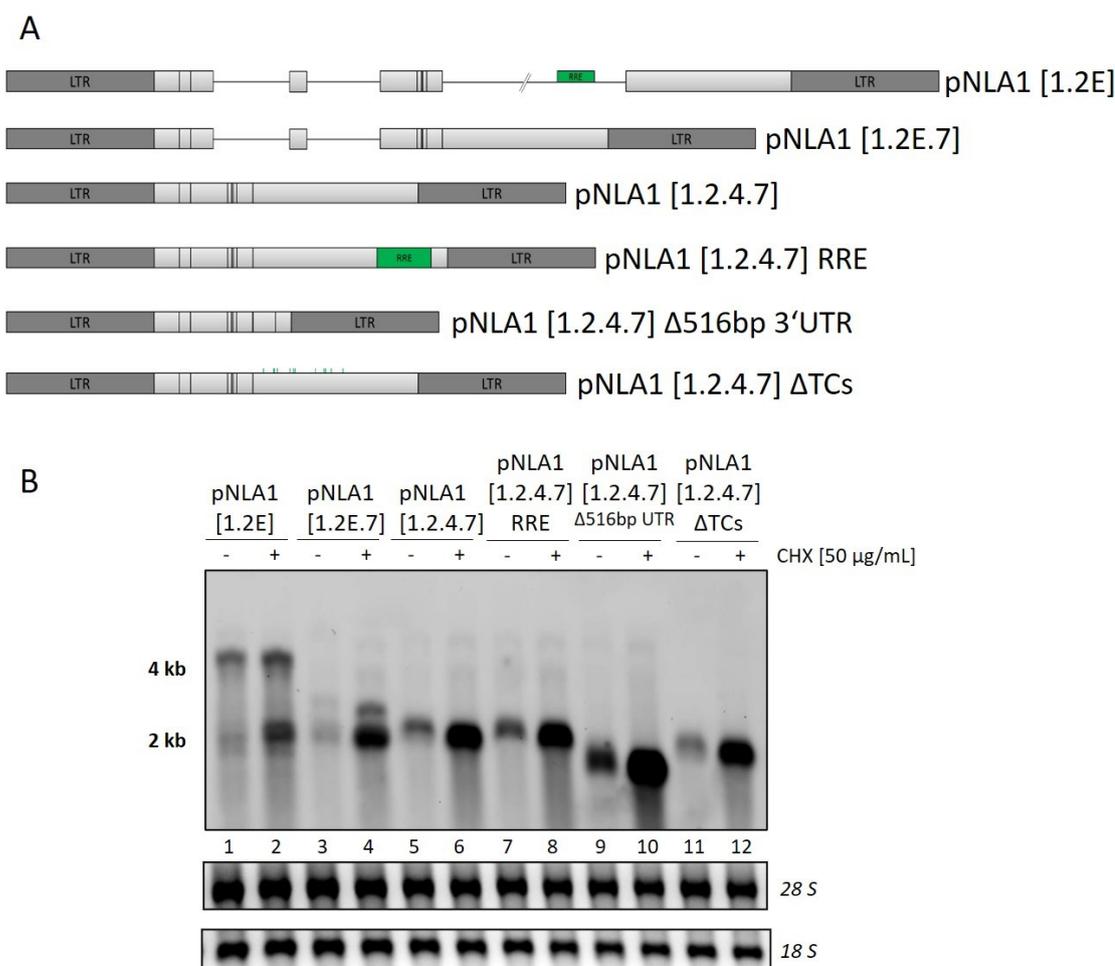
Figure 3.3: Modifications of the *tat* mRNA still resulted in CHX-dependent accumulation. (A) Schematic drawing of the reporter constructs used in this experiment. Light grey boxes indicate exons, while black lines indicate introns. The dark grey boxes show the viral LTRs and the green box indicates the location of the RRE. The top 3 plasmids were already introduced in Figure 3.2 and served as reference, the other 3 plasmids were modified: The pNLA1 [1.2.4.7] RRE construct encodes the rev-responsive element within the 3' UTR. The size of the pNLA1 [1.2.4.7] Δ516bp UTR construct was decreased by deletion of 516bp in the 3' UTR of *tat*. The *tat* ORF was artificially extended in the pNLA1 [1.2.4.7] ΔTCs reporter construct, by mutation of all in-frame termination codons (TCs) until the *nef* TC, each mutated TC is indicated by a green vertical line. **(B)** Northern blot analysis of the constructs depicted in (A). HeLa cells were transfected with 1 µg of each of the indicated plasmids for 24 hours. 6 hours before harvesting 50 µg/ml CHX was added to the cell culture media. After RNA isolation 3.5 µg of total RNA was digested with DNase I, separated on a 1% formaldehyde containing agarose gel, and transferred to a nylon membrane by overnight capillary blotting. For hybridization, a probe specifically detecting the 3' UTR of HIV-1 was used (#5854/#5855). The ribosomal bands served as size standards and loading control.

Another critical feature of RNA stability is its overall 3' UTR length and structure (Bühler et al. 2006). To study the impact of the untranslated region on NMD-mediated RNA degradation, two different approaches were used. Firstly, the length of the UTR was shortened by the deletion of 516 nucleotides downstream of the *tat* translational stop codon (TC) (Figure 3.3, lane 9 and 10). Secondly, all in-frame termination codons between the *tat* TC and the downstream *nef* start codon were mutated, to extend the encoded ORF from the *tat* ATG up to the *nef* TC (Figure 3.3, lanes 11 and 12). However, the northern blot results indicated that neither shortening the UTR

nor extending the ORF had an impact on the CHX-induced accumulation in this experimental setup, indicating that those features were not sufficient to counteract the NMD-mediated RNA degradation. However, as the 2kb class is independent of the interaction of Rev and the RRE for nuclear export, the HIV-1 4kb and 9kb classes require that interaction. Hence, the potential dependency of the RRE on the presence of Rev in the context of the NMD-induced RNA degradation was further investigated. Therefore, HeLa cells were transfected with the pNLA1 [1.2.4.7] constructs (with and without RRE) in the presence and absence of Rev. The RNA abundancies were again visualized via northern blot analysis (Figure 3.4).
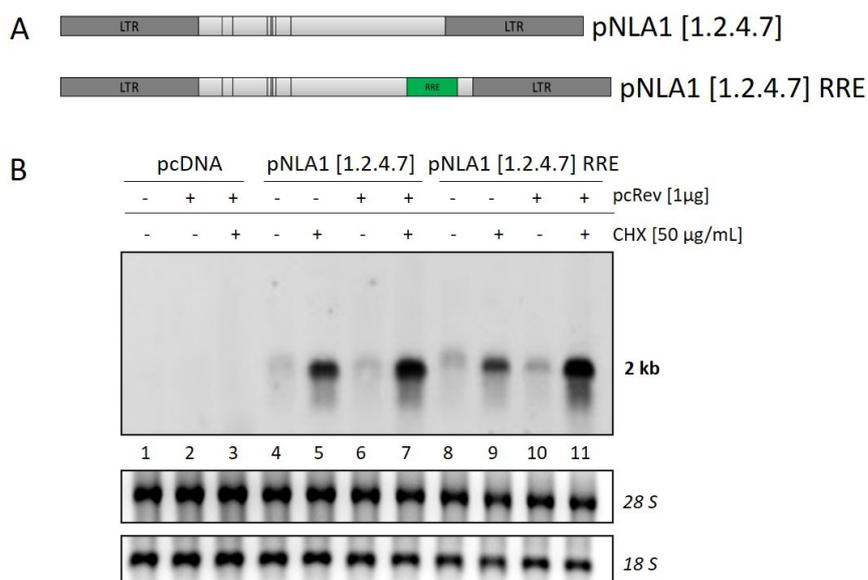

Figure 3.4: Co-transfection of *rev* further increased the CHX-dependent accumulation of *tat* mRNAs independent from the presence of the RRE. (A) Schematic drawing of the *tat* reporter constructs, which were already introduced in Figure 3.3. The light grey boxes indicate joined exons, the dark grey boxes show the viral LTRs and the green box indicates the location of the RRE within the *nef* ORF. **(B)** Northern blot analysis of the constructs depicted in (A). HeLa cells were co-transfected with 1 µg of the indicated plasmids, and either 1 µg of the empty vector pcDNA3.1(+), or 1 µg of pcRev, encoding for the Rev protein for 24 hours. 50 µg/ml CHX was added to the cell culture media 6 hours before harvesting. After RNA isolation 3.5 µg of total RNA was digested with DNase I, separated on a 1% formaldehyde containing agarose gel, and transferred to a nylon membrane by overnight capillary blotting. For hybridization, a probe specifically detecting HIV-1 exon 7 was used (#3387/#3388). The ribosomal bands served as loading control.

The co-transfection of pcRev with the pNLA1 [1.2.4.7] plasmids resulted in an advanced accumulation of the *tat* RNA levels upon CHX treatment, independent of the presence of the RRE (cf. Figure 3.4, lanes 4-7 and lanes 8-11). That indicated an increase of the NMD-induced RNA degradation by the presence of Rev protein in the cell, which was, however, independent of the presence of the RRE sequence within the transcript, in this experimental setup.

However, while NMD target verification is still often performed indirectly by translation inhibition through cycloheximide, NMD can be inhibited more specifically by the usage of siRNAs, targeting core NMD factors, like UPF1.

3.2 Efficient identification of endogenous NMD targets by translation inhibition through Cycloheximide

The observed CHX-dependent accumulation of the 2kb class, hence the CHX-dependent decrease in NMD-induced RNA degradation, was still observed upon various plasmid and transcript modifications. Since neither shortening of the *tat* 3' UTR, nor the additional presence of the RRE sequence within the *tat* mRNAs, with and without interaction of the RRE with Rev protein had an impact on the CHX-dependent RNA accumulation, the indirect NMD inhibition by cycloheximide was questioned. To address this question, three different cell culture cell lines (HeLa, Ghost, and HEK293T cells) were incubated with 50 µg/ml CHX for six hours. Total RNA was reverse transcribed and analyzed via semi-quantitative RT-PCR for the expression of endogenous genes, which encoded alternative splice isoforms that harbored PTCs. Those naturally occurring PTCs resulted in pre-mature translation termination and the subsequent degradation of those messenger RNAs by NMD. Those transcripts should therefore not be present in the RT-PCR analysis without translation inhibition. Five endogenous genes were tested, and the results are depicted in Figure 3.5.
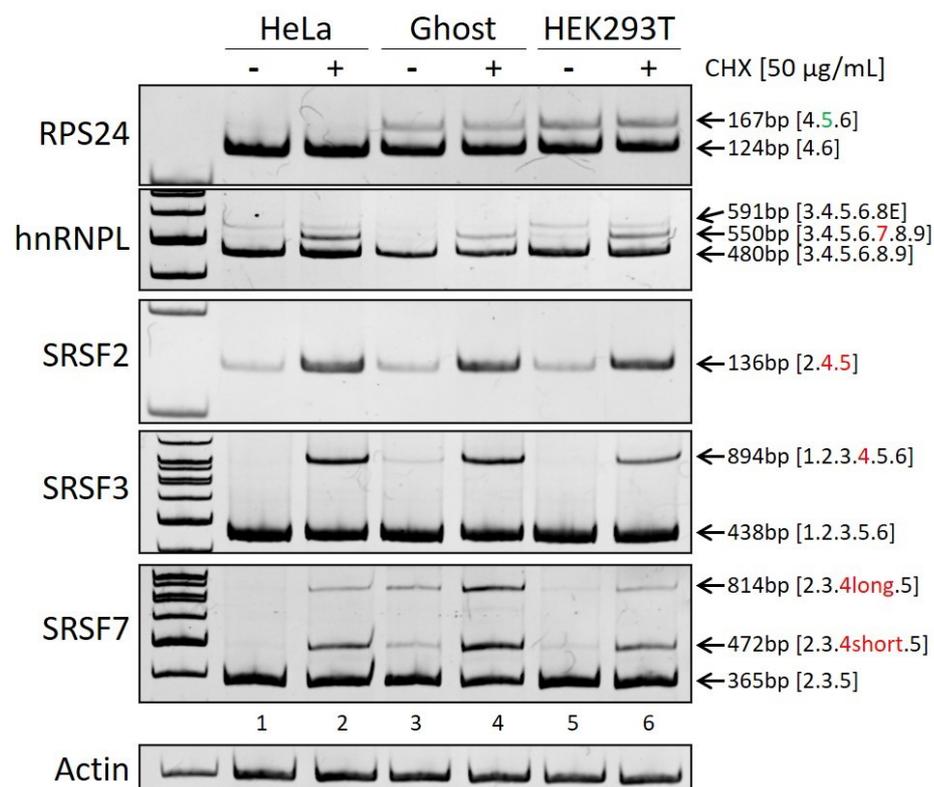
Figure 3.5: The CHX-dependent accumulation is specific for splice isoforms that are targeted by NMD. RT-PCR analysis of HeLa, Ghost, and HEK293T cell culture cells. 5×10^5 cells of each cell type were seeded in 6-well plates for 24 hours. 6 hours before harvesting of total RNA, 50 µg/ml CHX was added to one well of each cell type. The RNA was reverse transcribed and amplified with specific primer pairs. RPS24 (#5776/#5777, 26 cycles), hnRNPL (#5778/#5779, 26 cycles), SRSF2 (#6165/#6166, 26 cycles (Ramage et al. 2015)), SRSF3 (#4004/#4005, 35 cycles (Diehl 2016)), SRSF7

(#5784/#5785, 26 cycles), Actin (#6056/#6057, 26 cycles). The PCR fragments were separated on a 10% PAA gel and visualized by EtBr staining and UV-light exposure. The distinct bands, indicating the alternative splice isoforms, were excised from the gel, re-amplified, and confirmed by sequencing analysis. Numbers in brackets indicate the exon numbers found within the PCR product. Red numbers indicate exons with PTCs resulting in NMD-sensitive transcripts. Green numbers indicate exons with PTCs not resulting in NMD-sensitive transcripts.

Regarding endogenous genes, which encoded alternative splice isoforms that harbored PTCs (Figure 3.5) it was observed that some splice isoforms of the tested genes were only visible upon CHX treatment, while other isoforms increased in abundance, indicating different NMD-sensitivities of these splice isoforms. That ratio was, at least to some extent, cell type-specific, indicating a further layer of complexity regarding NMD-sensitivity of certain endogenous splice isoforms in certain cell types. For example, the protein-coding splice isoform of the ribosomal protein S24 (RPS24) excluded exon 5 from the mature mRNA, while the inclusion of exon 5 resulted in a pre-mature termination codon (PTC), 31nt upstream of the normally used TC, but only 22nt upstream of the final exon-exon junction and was therefore not recognized as NMD target. PTCs are only efficiently recognized when they are located at least 50-55 nucleotides upstream of an exon-exon junction (Cheng et al. 1994). Thus, the results in Figure 3.5 support this distance-dependent effect, as the amplification of transcripts which included RPS24 exon 5 did not accumulate upon CHX treatment. The only observable differences were the expression levels of the RPS24 [4.5.6] between the different cell types.

The transcript abundance of the other four genes which were tested in this setup (hnRNPL, SRSF2, SRSF3, and SRSF7) showed noticeable differences between the three tested cell types regarding the splice isoforms in the untreated control sample, as well as in the abundance of NMD target mRNAs without CHX treatment (Figure 3.5, compare lanes 1, 3 and 5), suggesting that the NMD-dependent degradation process had different efficiencies in the different cell types. However, upon CHX treatment all these four genes showed NMD-sensitive splice isoforms, which included PTC-containing exons, and accumulated upon CHX treatment.

Regarding the effect of translation inhibition by the addition of cycloheximide to the cell culture, the observed outcome for the five analyzed genes was rather specific. The CHX-dependent accumulation of transcripts was restricted to predicted NMD targets and did not alter the abundance of not-NMD-transcripts (cf. Figure 3.5, RPS24 and actin). These results suggested a rather specific NMD inhibition through CHX.

However, as the treatment with CHX not only interfered with the NMD pathway, but with the global translation apparatus, the next step was to interfere more precisely with the components of the NMD machinery, to reduce potential side effects. Therefore, small interfering RNAs (siRNAs) were tested in comparison with the so far used CHX control. Those siRNAs either targeted a non-specific control sequence (siNSC), or the major NMD factor UPF1 (siUPF1). First, the concentration of siRNAs, sufficient for NMD inhibition, had to be determined. Therefore,

HeLa cells were transfected with three different concentrations of siRNAs for 48 hours and analyzed for their UPF1 levels, cell viability, and NMD target degradation via western blot and RT-PCR (Figure 3.6).
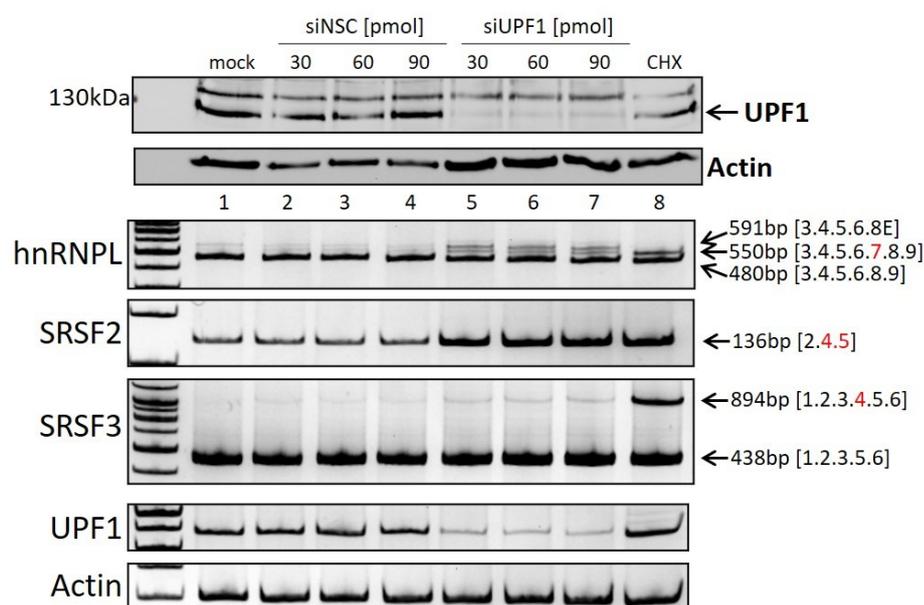

Figure 3.6: NMD inhibition through UPF1 siRNA transfection works efficiently at low concentrations. 2×10^5 HeLa cells were seeded into 6-well plates and transfected with the depicted concentration of siRNAs on the next day. The medium was changed 24h post-transfection and the cells were incubated for another 24h before RNA and proteins were harvested from each well. 6h before harvesting one of the not-transfected wells was treated with 50 µg/ml CHX as control. **(Top)** Western blot analysis of the isolated proteins using specific antibodies against UPF1 or actin, respectively. **(Bottom)** RT-PCR analysis of the total RNA isolated from the same cells. The RNA was reverse transcribed and amplified with specific primer pairs. hnRNPL (#5778/#5779, 26 cycles), SRSF2 (#6165/#6166, 26 cycles (Ramage et al. 2015)), SRSF3 (#4004/#4005, 35 cycles (Diehl 2016)), UPF1 (#5427/#5428, 26 cycles), actin (#6056/#6057, 26 cycles). The PCR fragments were separated on a 10% PAA gel and visualized by EtBr staining and UV-light exposure.

The results in Figure 3.6 showed that the siRNA-mediated knockdown of UPF1 worked even with the smallest tested concentration of siRNAs. Even 30 pmol of transfected UPF1 siRNAs were sufficient to reduce the levels of UPF1 RNAs, as well as UPF1 protein, as shown by RT-PCR analysis and western blot. In addition, none of the tested UPF1 siRNA concentrations resulted in a reduction of cell viability, as indicated by the consistent actin levels in the western blot (Figure 3.6, top, lanes 5-7).

Next, to monitor the efficiency of NMD inhibition by UPF1 siRNAs, the splicing pattern of three different endogenous genes (hnRNPL, SRSF2, and SRSF3) were analyzed by RT-PCR and compared to the CHX control (Figure 3.6, bottom). Regarding SRSF2, the accumulation of the NMD-specific isoform upon UPF1 knockdown strongly resembled the CHX control, while the splicing patterns for SRSF3 and hnRNPL differed to some extent from the observed CHX-dependent accumulation. For hnRNPDL, the siRNA-mediated UPF1 knockdown resulted in increased detection of the exon 7 containing splice variant, but furthermore, also resulted in the

detection of an alternative splice isoform, which had not been detected so far. Within this new splice isoform, exon 7 of hnRNPL was excluded, but the intron between exon 8 and 9 was retained. This result could indicate that UPF1 not only functions as the core component of the NMD machinery but might also function as splicing regulatory protein, facilitating the usage of hnRNPL splice donor 8, as SD8 is not used upon UPF1 knockdown. The third analyzed gene was SRSF3, where UPF1 knockdown increased the NMD-specific, exon 3 containing isoform, although not as strong as upon CHX treatment.

In conclusion, the UPF1 knock-down-induced accumulation of annotated NMD-sensitive splice isoforms strongly resembled the effect of CHX-induced translational inhibition. Thus, it can be concluded that CHX-mediated translation inhibition mimics the more specific NMD inhibition by siRNA knockdown of UPF1.

Now, to compare the observed indirect effect of CHX treatment on the fully spliced 2kb *tat* mRNAs (cf. chapter 3.1) to the direct siRNA-mediated UPF1 knockdown, and to investigate the afore mentioned potential stabilizing impact of the Rous Sarcoma Virus RNA stability element (RSE) on pNLA1 [1.2.4.7] transcripts, northern blots were used. For the experimental setup, HeLa cells were either treated with a non-specific siRNA (siNSC) or siRNAs against UPF1 (siUPF1), respectively. On the next day, the cells were transfected with the corresponding pNLA1 [1.2.4.7] *tat* plasmid for another 24 hours. As control, one well of each experimental condition was incubated with 50 µg/ml CHX six hours before the RNA and protein harvest. The results of the protein analysis, hence the UPF1 protein levels, are shown by western blot in Figure 3.7 (WB, top), while the RNA levels of the HIV-1 2kb *tat* message are shown by northern blot in Figure 3.7 (NB, bottom).
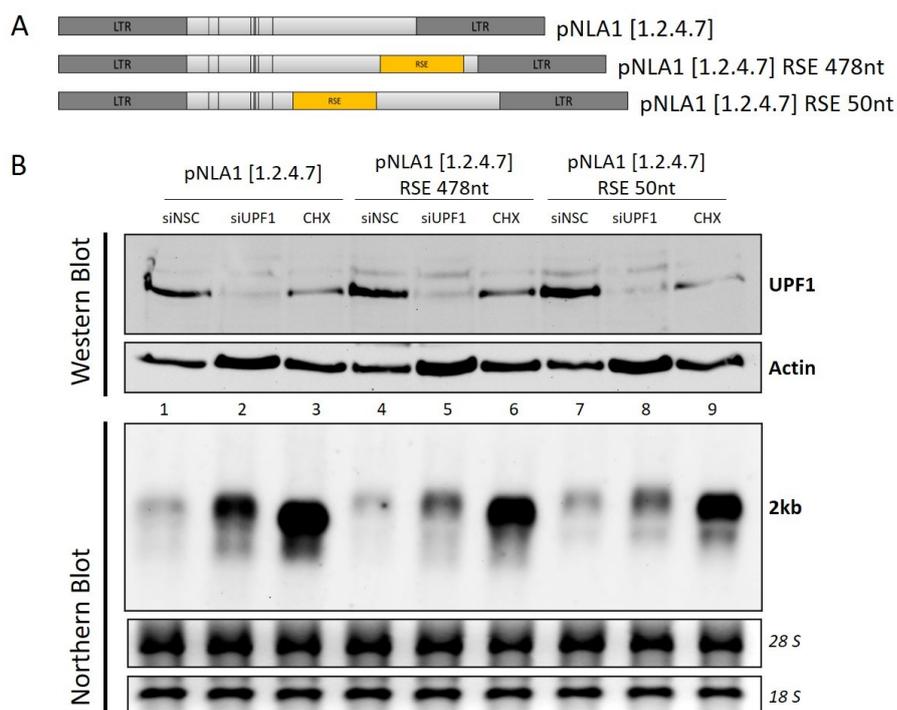
Figure 3.7: The NMD-induced RNA degradation of HIV-1 *tat* mRNAs can be decreased by the addition of the RSE.
(A) Schematic drawing of the pNLA1 [1.2.4.7] *tat* reporter constructs, originally introduced in Figure 3.3. The light grey boxes indicate joined exons, the dark grey boxes show the viral LTRs. The yellow boxes indicate the location of the inserted RSE sequence either 478 nt downstream or 50 nt downstream of the *tat* TC. **(B)** $2x10^5$ HeLa cells were seeded into 6-well plates and transfected with 30 pmol of the indicated siRNAs. The medium was changed 24h later and the cells were transfected with 1 µg of the indicated plasmids for another 24h. For the CHX-control wells, 50 µg/ml CHX was added to the cell culture medium 6h before harvesting the total RNA and protein from each well. **(B, Top)** Western blot analysis of the isolated proteins with specific antibodies detecting UPF1 or actin, respectively. **(B, Bottom)** Northern blot analysis of the *tat* mRNAs. 3 µg of total RNA were digested with DNase I, separated on a 1% formaldehyde containing agarose gel, and transferred to a nylon membrane by overnight capillary blotting. For hybridization, a probe specifically detecting HIV-1 exon 7 was used (#3387/#3388). The ribosomal bands served as loading control.

The siRNA-mediated UPF1 knockdown was successful in all samples which were treated with the siRNA targeting UPF1, as shown by the western blot for the UPF1 protein levels (Figure 3.7 B, top). However, the UPF1 protein levels not only decreased upon siRNA targeting but also by the six-hour CHX treatment, indicating a short half-life of UPF1 proteins.

Regarding the *tat* RNA levels, which were produced by transcription of the transfected plasmids (Figure 3.7 B, bottom), the pNLA1 [1.2.4.7] *tat* RNA levels increased upon UPF1 knockdown almost as prominently as upon CHX treatment (cf. lanes 1-3), indicating a similar efficiency of the UPF1 siRNA knock-down-mediated NMD inhibition as observed prior for the translation inhibition based NMD interference. Again, as observed before (cf. Figure 3.3), the size of the *tat* mRNAs upon CHX treatment, but not upon UPF1 knockdown, appeared minimally smaller and as a smear, which could indicate degradation of these transcripts upon translation termination.

Since the CHX-induced accumulation, hence the CHX-induced decrease in NMD-dependent RNA degradation, of the pNLA1 [1.2.4.7] *tat* transcripts was further enhanced by the insertion of the Rev responsive element (RRE) (cf. Figure 3.3 and Figure 3.4), a different structural element, which was described as stability element was tested. Therefore, the sequence of the Rous Sarcoma Virus RNA stability element (RSE) was inserted into the 3' UTR of the *tat* ORF in the context of the pNLA1 [1.2.4.7] plasmid. As the RSE was shown to act in a position-dependent manner, it was inserted in two different positions (Ge et al. 2016). Hence, the RSE was located either 478nt or 50nt downstream of the *tat* termination codon, and the results of the northern blot analysis are shown in Figure 3.7 (lanes 4-9). The addition of the RSE downstream of the *tat* TC indeed stabilized the RNAs, however, independent of the inserted position. Furthermore, the NMD-induced RNA degradation was reduced by the insertion of the RSE in the context of siRNA-mediated NMD inhibition, while the CHX-dependent accumulation was still clearly detectable, even though it decreased as well with the addition of the RSE to the transcript.

Based on the observed differences of the NMD-mediated RNA degradation of RSE containing *tat* transcripts between the direct NMD inhibition by siRNA-mediated UPF1 knockdown and the indirect NMD inhibition by cycloheximide treatment, the CHX-dependent effect was unlikely to be only mediated by the NMD-sensitivity of a given transcript. Since CHX was used as a translational inhibitor, it was assumed that it only influenced transcripts that are actively translated and would be degraded co-translationally based upon their NMD triggering features.

To test the hypothesis of whether CHX treatment only influences transcripts that are actively translated, the expression of transcripts encoded by the plasmid pNLA1 to the plasmid pNLA1 tat⁻rev⁻ was compared. Within the pNLA1 tat⁻rev⁻ plasmid, the translational start-codons of *tat* and *rev* mRNAs were mutated, resulting in the absence of Tat and Rev protein. Since the Tat protein is essential for the enhanced activation of the viral LTR promotor, and the Rev protein is essential for the nuclear export of intron-containing RNAs, those proteins had to be provided in *trans*. The obtained RNA levels were analyzed by northern blot (Figure 3.8).
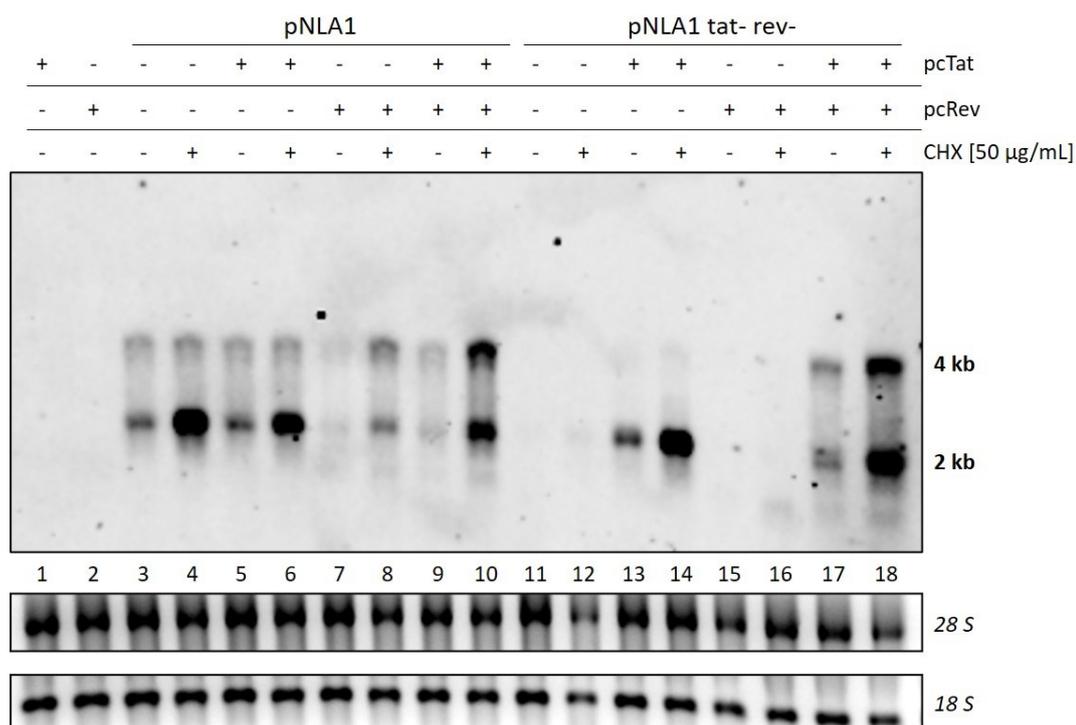
Figure 3.8: The accumulation of the HIV-1 2kb class is not dependent on a functional *tat* ORF and the HIV-1 4kb class accumulates upon Rev co-expression. Northern blot analysis of 3 µg of total RNA isolated from HeLa cells. 2.5×10^5 HeLa cells were transfected with 1 µg of each of the indicated plasmids, or pcDNA3.1(+) for 24 hours. 50 µg/ml CHX was added to the cell culture media 6 hours before harvesting. After RNA isolation 3 µg of total RNA were digested with DNase I, separated on a 1% formaldehyde containing agarose gel, and transferred to a nylon membrane by overnight capillary blotting. A specific probe, detecting HIV-1 exon 7 (#3387/#3388), was used for hybridization. The ribosomal RNA bands served as size markers and loading control.

Only co-transfection of the pNLA1 tat⁻rev⁻ plasmid together with the expression plasmids for Tat and Rev (pcTat and pcRev) resulted in high enough RNA levels for detection by northern blot analysis (Figure 3.8, cf. lanes 11, 13, and 17). The *tat* and *rev* mRNAs, which were produced from the provided expression plasmids, were not visible in the northern blot as the used DIG-labeled probe specifically detected HIV-1 exon 7. This sequence was absent in the expression plasmids and the corresponding mRNAs (Figure 3.8, lanes 1 and 2).

The results in Figure 3.8 mimic the prior observed effect of the CHX-dependent accumulation of the 2kb class, even though the translational start codons of the *tat* and *rev* mRNAs were mutated in the pNLA1tat⁻rev⁻ construct. However, what was not observed before, the co-transfection of the pNLA1 plasmids together with the Rev expression plasmid resulted in the CHX-dependent accumulation of both, not only the 2kb class RNAs, but also the 4kb class. Generally, the transfection of the pNLA1 plasmid alone resulted, as expected, in detectable levels of HIV-1 2kb and 4kb class mRNAs without the need of Tat or Rev co-transfection (lanes 3 and 4). The additional transfection of pcTat did not affect the RNA levels demonstrating no limiting amounts of Tat expressed from pNLA1 (lanes 5 and 6), but the co-transfection with pcRev, however, resulted in an unexpected decrease of 2kb and 4kb RNAs, indicating a destabilization of the viral

RNAs by high levels of Rev (lanes 7 and 9). The transfection of the pNLA1 tat⁻rev⁻ plasmid alone resulted, as expected, in almost no detectable HIV-1 mRNAs (lanes 11 and 12), and the co-transfection with pcTat (without pcRev) enhanced the activity of the viral LTR promotor, resulting in high levels of 2kb class RNAs, which accumulated upon CHX treatment (lanes 13 and 14). The co-transfection with pcRev (without pcTat) resulted in no evaluable result, as the mRNA levels were under the detection level (lanes 15 and 16). However, the co-transfection with both, pcTat and pcRev resulted in detectable amounts of 2kb and 4kb mRNAs, both of which accumulated upon CHX treatment (lanes 17 and 18).

Concluding from Figure 3.8, the expression of additional Rev protein reduced the overall RNA abundance in the context of pNLA1, which indicated an increase of the NMD-induced RNA degradation (cf. lane 3 to lanes 7 and 9), as it was observed priorly for the pNLA1 [1.2.4.7] transcripts upon *rev* co-transfection (cf. Figure 3.4). Furthermore, upon co-transfection with pcRev, now not only the HIV-1 2kb class mRNAs but also the 4kb class of HIV-1 mRNAs accumulated upon CHX treatment, suggesting an influence of cellular Rev protein levels on the HIV-1 4kb class stability. The higher amount of Rev protein should enhance the nuclear export of 4kb mRNAs, but instead, a decrease in the amount of 4kb class RNAs was detected in the northern blot analysis, which furthermore showed an accumulation upon CHX treatment (cf. lane 3 and 4 to lanes 7-9). This CHX-dependent accumulation of the 4kb class mRNAs has not been observed so far, which might be due to the normally much lower intracellular Rev levels, and the absence of the 4kb class in prior experiments with pcRev co-transfection (cf. Figure 3.4).

It was described that high levels of Rev inhibit the NFX1-mediated nuclear export, which might explain the reduced detection of 2kb class mRNAs upon pcRev co-transfection (Taniguchi, Mabuchi, and Ohno 2014). Furthermore, if high levels of Rev interfere with the NFX1-dependent nuclear export, that interference might reduce the overall cellular mRNA levels, thereby increasing the chance of recognizing the 4kb mRNAs as NMD targets.

However, this experiment aimed to investigate whether the deletion of the *tat* and *rev* translational start codons influenced the NMD-sensitivity of these messages and if the removal of the ATGs would decrease the CHX-dependent accumulation of the 2kb class. However, as shown by the northern blot in Figure 3.8, that accumulation was still prominent. Even though the 2kb *nef* mRNA still possessed its ATG, the comparison of the 2kb class accumulation between the pNLA1 and the pNLA1tat⁻rev⁻ messages (cf. Figure 3.8 lanes 3, 4 with lanes 13, 14 and 17, 18) is less likely to be only mediated by *nef* mRNAs. Therefore, it was questioned whether the degradation of target mRNAs by NMD is at all dependent on translation and if it might be influenced by the intrinsic strength of the translational start codon, as also HIV-1 transcripts show differences in the intrinsic strength of their translational start codons.

3.3 The efficiency of NMD target recognition depends on the intrinsic strength of the translational start codon

To investigate the translation dependency on NMD target recognition, a dual-reporter system was used, which encoded two independent open reading frames on a single plasmid. This plasmid was cloned by Aljoscha Tersteegen, during his Master Thesis and kindly provided for this work. Within this vector system, the firefly luciferase (FL) served as reference and was expressed under the control of an SV40 promotor. The gene on interest (GOI) in this reporter system was the renilla luciferase (RL), fused to exon 1, 2, and 3 of the human β-globin sequences, with corresponding introns. The RL open reading frame (ORF) was thereby extended from the RL start codon to the β-globin translational stop codon, which was located within exon 3 of β-globin. The expression of RL-β-globin was regulated by its own SV40 promotor sequence. In addition to the RL-β-globin WT plasmid, the introduction of a pre-mature termination codon (PTC) within exon 2 of β-globin resulted in the early translation termination at amino acid position 39 (NS39) of β-globin. This PTC was located more than 55 nucleotides upstream of the last exon-exon junction and therefore made the corresponding mRNAs targets for degradation by NMD. A schematic drawing of the dual-luciferase reporter is depicted in Figure 3.9 A. Upon normalization of the renilla levels to the corresponding firefly levels, this reporter system not only allowed a PCR-based readout of the mRNA abundancies but was also used to analyze the relative luciferase activity of the produced luciferase proteins, which allowed a functional readout of the translational activity.

The transfection of HeLa cells with these dual-luciferase reporters resulted, as expected, in high FL mRNA levels, as well as a high firefly luciferase activity. The measurable reference firefly-parameters were comparable between the WT and the NS39 plasmid, as expected by the same sequence of the FL ORF in both plasmids (cf. Figure 3.9 D, firefly RT-PCR, lane 1 (RL-ATG β-globin-WT) and lane 2 (RL-ATG β-globin-NS39)). The measured RL activity, as well as the renilla RNA levels, however, differed strikingly between cells transfected with the RL-ATG β-globin-WT and the RL-ATG β-globin-NS39 variant (cf. Figure 3.9 B-D). The decreased renilla luciferase mRNA levels of the RL-ATG β-globin-NS39 plasmid in the RT-(q)-PCRs showed the NMD-dependent degradation of the PTC-containing mRNA. Even though the RL-ATG β-globin-NS39 variant contained a PTC within the β-globin encoded sequence, the upstream RL encoded sequence was intact. But based on the PTC-dependent degradation of the mRNA the subsequent translation of the functional hybrid RL-β-globin protein was abolished, which therefore resulted in minimal measurable renilla luciferase activity (Figure 3.9 C, D).

To further investigate the importance of translation and the influence of the intrinsic strength of the translational start codon on the NMD-dependent degradation of PTC containing mRNAs, the translational start codon was altered. The plasmids described above contained a strong

translational start codon (ATG), which almost perfectly matched the well-known Kozak consensus sequence (GCCRCC ATG G, R = Purine (Kozak 1986)) and was efficiently recognized by the small ribosomal subunit. Only the position +1 consisted of an adenine (A) instead of a guanine (G), which was necessary to maintain the amino acid sequence of the renilla luciferase. The nucleotide sequences of the used translational start sites are shown beneath the schematic overview of the used dual-luciferase plasmids in Figure 3.9 A.
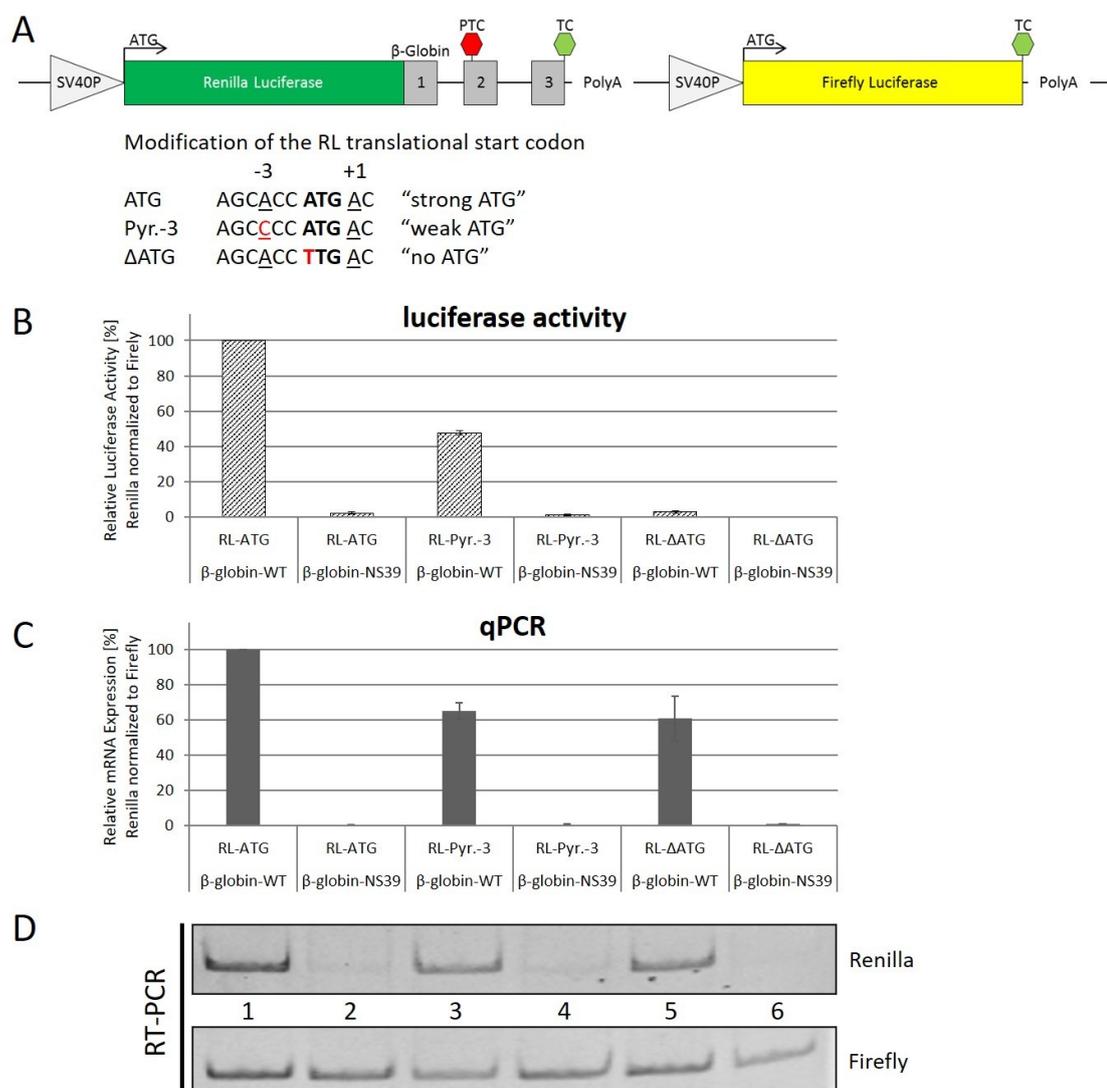

Figure 3.9: The sequence of the translational start codon influences the overall RNA stability, but not the NMD-dependent degradation of target mRNAs. (A) Schematic drawing of the used dual-luciferase reporter constructs. The ORF of the renilla luciferase (RL) was fused to exon 1-3 of β-globin with corresponding introns. The ORF of the WT-RL-β-globin starts at the drawn arrow labeled with ATG and ends at the TC within exon 3 of β-globin. The NS39-RL-β-globin variant contained a PTC within exon 2 of β-globin (PTC). The firefly luciferase (FL) served as an internal reference and the FL ORF is indicated by the translational start ATG and the green stop TC. Both ORFs are expressed independently from the same plasmid using their own SV40 promotor sequence. The RL start codon was modified to either create a translational suboptimal Kozak surrounding with a pyrimidine at position -3 (Pyr.-3), or the ATG was removed by mutation to TTG (ΔATG). **(B-D)** 5×10^5 HeLa cells were transfected with the depicted reporter constructs. 24h post-transfection, one-half of the cells was used for the measurement of the luciferase activity, the

other half was used for total RNA isolation. **(B)** Results of the measured luciferase activity. The activity of the renilla luciferase was normalized to the corresponding firefly luciferase. **(C)** RT-q-PCR results of the total RNA after DNase I digestion and reverse transcription using specific primer pairs (RL: #6167/#6168; FL: #3279/#3280). **(D)** RT-PCR analysis of the same cDNA and the same primer pairs as in (C). After PCR amplification with 26 cycles, the PCR products were separated on a 10% PAA gel and visualized by staining with EtBr and UV-light exposure.

To investigate the translation dependency of RNA target degradation through NMD, the translational start codon of the RL ORF was modified in two ways. First, the purine (adenine, A) in position -3 was replaced by a pyrimidine (cytosine, C) resulting in a translational "weak" ATG, with suboptimal Kozak surrounding (Pyr.-3). Second, the ATG itself was removed by substitution of the "ATG" with a "TTG", without further modification of the surrounding sequence (ΔATG; cf. Figure 3.9 A). The subsequent transfection of HeLa cells with these constructs and the measurement of the relative luciferase activities showed that the relative luciferase activity of RL-Pyr.-3 β-globin-WT decreased to 50% upon substitution of the purine in position -3 for a pyrimidine (A>C) (cf. Figure 3.9 B). However, the insertion of the PTC within exon 2 of β-globin (RL-Pyr.-3 β-globin-NS39) still abolished the renilla luciferase activity, as already observed for the RL-ATG β-globin-NS39 containing construct with optimal Kozak surrounding (Figure 3.9 B). The mutation of the RL translational start codon (ATG>TTG, RL-ΔATG β-globin) resulted, as expected, in no measurable renilla luciferase activity, regardless of whether the ORF contained a PTC or not.

The so-far convincing results, however, were contradicted by the analysis of the corresponding RNA levels, which were done by real-time PCR and semi-quantitative PCR (Figure 3.9 C and D). While the renilla mRNA levels of the WT constructs decreased to a rate of about 60% by weakening (Pyr.-3), or removal of the ATG (ΔATG), the insertion of a PTC resulted in degradation in both cases. These results suggested either that the target degradation through NMD was indeed not dependent on translation, or that instead of the first translational start codon, another ATG further downstream was used. The latter could explain the loss of protein activity of the RL-ΔATG β-globin-WT construct, because of an N-terminal shortened protein.

To address this question the full sequence of the RL-β-globin gene was scanned for potential start codons in all three different potential open reading frames, reaching from the original renilla ATG to the polyadenylation signal downstream of β-globin exon 3. This search identified 46 potential translational start sites within all three ORFs, which were subsequently mutated to abolish translation initiation from these sites (Figure 3.10 A, red lines). Due to the location of some of the identified ATGs within the original open reading frame, it was not always possible to substitute bases as silent mutation, which would maintain the amino acid sequence of the renilla luciferase. These necessary substitutions could therefore have an impact on the overall renilla protein function and thereby influence the measurable renilla luciferase protein activity. However, the chosen mutations were selected with respect to the HEXplorer profile and without the generation of new potential splice donor or acceptor sites (the respective HEXplorer plot

can be found in the Appendix, Figure 6.1). Furthermore, using the HEXplorer profile as a prediction tool for potential binding sites of splicing regulatory proteins, the generation of new potential splicing regulatory elements was avoided by minimalizing the mutational changes of the profile. HeLa cells were transfected with the modified plasmids, and the results of the RT-(q)-PCRs are shown in Figure 3.10.
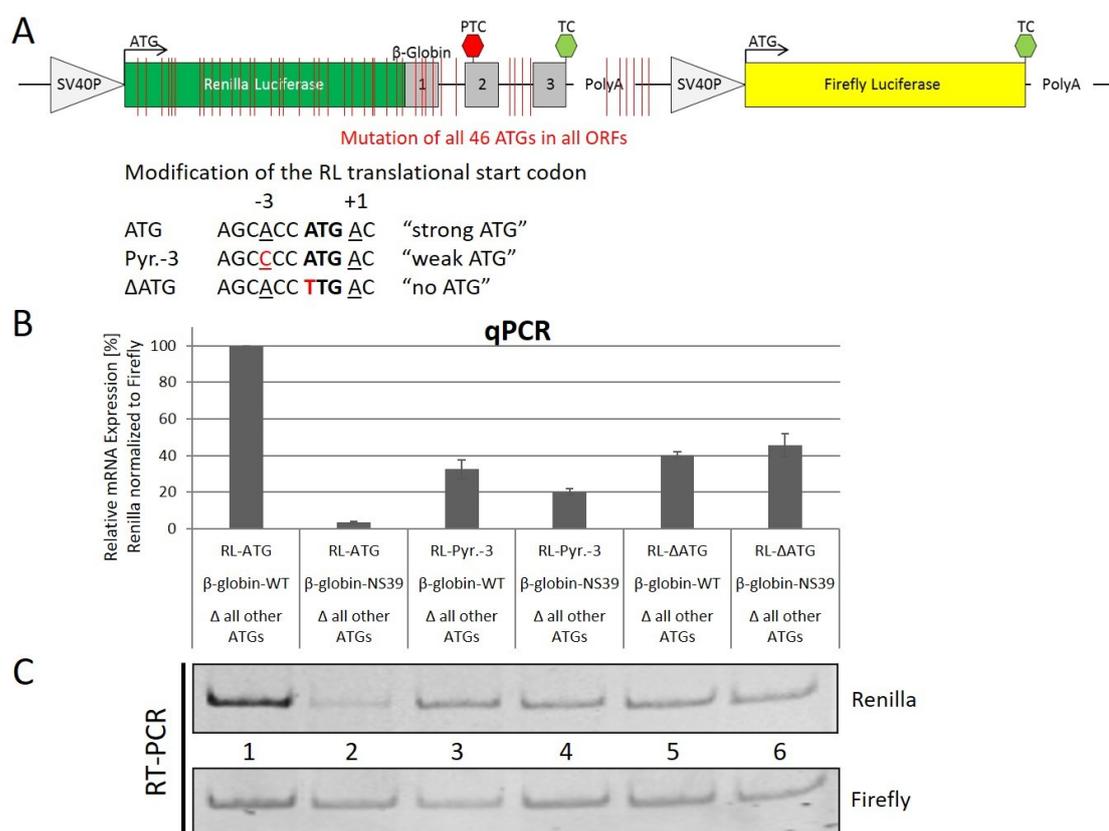

Figure 3.10: Removal of all in-frame translational start codons abolishes NMD target recognition in the context of a suboptimal, or no ATG. (A) schematic drawing of the used dual-luciferase reporter constructs, as already described in Figure 3.9. The red lines within RL-β-globin indicate the 46 ATGs in all three potential open reading frames, that were mutated for this analysis. **(B-C)** $5x10^5$ HeLa cells were transfected with the depicted reporter constructs. 24h later, one-half of the cells was used for the measurement of the luciferase activity, the other half was used for total RNA isolation. **(B)** RT-q-PCR results of the total RNA after DNase I digestion and reverse transcription using specific primer pairs (RL: #6167/#6168; FL: #3279/#3280). **(C)** RT-PCR analysis of the same cDNA and the same primer pairs as in (B). After PCR amplification with 26 cycles, the PCR products were separated on a 10% PAA gel and visualized by staining with EtBr and UV-light exposure.

Transfection of HeLa cells with the modified constructs depicted in Figure 3.10 A resulted in no measurable renilla luciferase activity, while the activity of the firefly luciferase remained comparable to the prior experiments with the dual-luciferase reporter (data not shown). As mentioned before, it was not possible to substitute all bases within the renilla ORF as silent mutations, hence the amino acid sequence of the renilla luciferase was slightly altered. These nucleotide alterations were sufficient to change the corresponding amino acids to an extend that had an impact on the protein function.

However, due to the design of the reporter system, it was still possible to analyze the impact of the mutations on the mRNA levels. The two transcripts which contained the original renilla ATG (RL-ATG β-globin-WT Δ all other ATGs and RL-ATG β-globin-NS39 Δ all other ATGs) showed similar PCR intensities as their corresponding plasmid pair without the 46 mutations (cf. Figure 3.10 B and C with Figure 3.9 C and D). The highest renilla mRNA transcript levels were observed for the RL-ATG β-globin-WT Δ all other ATGs variant, while much fewer renilla transcripts were found for the RL-ATG β-globin-NS39 Δ all other ATGs variant, which indicated an efficient degradation of these mRNAs by NMD (cf. Figure 3.10 C and D, lanes 1 and 2). Surprisingly, the renilla RNA levels of the WT constructs harboring a weak ATG (RL-Pyr.-3 β-globin-WT Δ all other ATGs), or no ATG (RL-ΔATG β-globin-WT Δ all other ATGs) were only around 30-40% of the ATG WT (RL-ATG β-globin-WT Δ all other ATGs) levels. Hence, that ratio was decreased in comparison to the 60% ratio, which was observed for the WT RNA levels in Figure 3.9.

The additional insertion of a pre-mature termination codon (NS39) into the Renilla-β-globin ORF, in the context of the RL, construct with the suboptimal start codon (RL-Pyr.-3 β-globin-NS39 Δ all other ATGs), reduced the RNA levels by only 10% (cf. Figure 3.10 B, C, lanes 3 and 4). That indicated an increase in the RNA stability, based on the impaired recognition of these target mRNAs by the NMD pathway due to decreased translation initiation rates.

Finally, the addition of a PTC into the construct without any translational start codon (RL-ΔATG β-globin-NS39 Δ all other ATGs) did not decrease the RNA levels at all, supporting the hypothesis that RNA degradation through NMD is indeed dependent on translation (cf. Figure 3.10 B, C, lanes 5 and 6). However, due to the impaired degradation of the NS39 Pyr-3 transcript, NMD is not only dependent on translation at all, but also on the recognition efficiency of the translational start codon. Therefore, mRNAs which harbor a semi-optimal Kozak surrounding might escape degradation through NMD, even if they encode NMD triggering features. By that mechanism, HIV-1 mRNAs with weak translational start sites, like *vpu* could be expressed within the host cell, even though the overlapping *vpu-env* message is a potential NMD triggering feature. An overview of the translational start sites of the different HIV-1 mRNAs is shown in Table 3.1. The nucleotides in positions -3 and +1 were used to score the strength of the translational start sites according to the Kozak sequence. Translational suboptimal start codon surroundings were found for *rev*, *vpu*, and *vpr*, which might mediate the escape of these mRNAs from the nonsense-mediated mRNA decay. However, the observed NMD-sensitivity of the HIV-1 *tat* mRNAs cannot be explained by a suboptimal Kozak surrounding, as *tat* has a strong translational start codon.

Table 3.1: The translational start codons of the different HIV-1 mRNAs in comparison to the optimal Kozak sequence surrounding GCC RCC **ATG** G, R = Purine (Kozak 1986). The HIV-1 translational start codon sequences were taken from the sequence of the pNL4-3 plasmid. Besides the ATG, the nucleotides in positions -3 and +1 were shown to be most important. Nucleotides that match the consensus sequence are highlighted in green (guanine (G) or adenine (A) in position -3; and guanine (G) in position +1), mismatches in these positions are highlighted in red. The intrinsic strength of the HIV-1 ATGs was defined by the match to the Kozak sequence in the positions -3 and +1

HIV-1 RNA class	HIV-1 mRNA	Translational start codon	Intrinsic strength
2kb	*tat*	CAA GAA **ATG** GA	strong
	nef	TAT AAG **ATG** GG	strong
	rev	TCT CCT **ATG** GC	weak
4kb	*vpu*	CAT GTA **ATG** CA	weak
	env	GTG GCA **ATG** AG	strong
	vif	GGG ATT **ATG** GA	strong
	vpr	GGA CAG **ATG** GA	weak
9kb	*gag*	AGA GAG **ATG** GG	strong

The initial observations of HIV-1 transcripts being potential NMD targets were found by the accumulation of the HIV-1 2kb class mRNAs upon translation inhibition through cycloheximide. Furthermore, this work is partially based on previous research results of Dr. Nora Diehl, who used a northern blot-based NMD reporter system to characterize the impact of HIV-1 infection on the cellular NMD machinery (Diehl 2016). Based on the observed complexity of the dual-luciferase NMD reporter system, this northern blot-based reporter system was investigated for its reliability of NMD target modulations.

3.4 Commonly used NMD reporter systems generate alternative splice isoforms with NMD-triggering features

A broad spectrum of what is known today about the nonsense-mediated mRNA decay pathway was discovered with the use of NMD reporter systems, which mostly consist of multiple exons, with a WT and a PTC containing variant. Not only the three exon β-globin reporter is commonly used, but another also well-known reporter system is the seven-exon triose-phosphate-isomerase (TPI) reporter (Böhm et al. 2016). In theory, NMD reporters offer a simplified evaluation system and are independent of the potentially different NMD-sensitivities of individual endogenous transcripts. Regarding the priorly observed difficulty of the NMD reporter situation, hence the number of modifications that had to be made to the RL-β-globin reporters to answer the initial question, the TPI NMD reporter system was analyzed next. The TPI reporter system is of great interest, as it offers an easy detection and quantification of the RNA levels by northern blot analysis. The pCI-TPI-WT-4H reporter, as well as its corresponding PTC containing variant pCI-TPI-PTC40-4H were originally designed by Prof. Dr. Niels Gehring's lab (Böhm et al.

2016) and kindly provided to Dr. Nora Diehl, who already used those plasmids during her Ph.D. thesis (Diehl 2016).

One great advantage of this reporter system was the probe binding site, downstream of the translation termination codon. This binding site consisted of four 100 nt repeats of the human β-globin (HBB) sequence, which were used as the binding site for the northern blot DIG-labeled probe, which resulted in efficient detection of the TPI RNA levels by northern blot analysis. Furthermore, another advantage of this reporter system was the overall number of encoded TPI exons, hence the corresponding number of exon-junction-complexes (EJCs) on the mRNA post splicing. Due to the position of the PTC at amino acid position 40 (PTC40) within TPI exon 2, the five downstream exon junctions exhibited an efficient NMD-dependent degradation of the PTC-containing isoform. The following figure represents the result from Dr. Nora Diehl, which gave the initial hint of an HIV-1 infection-mediated modulation of the cellular NMD pathway. For this experiment (Figure 3.11, adapted from (Diehl 2016)) HeLa T4+ cells were transiently transfected with the TPI reporter constructs (either WT or PTC-containing). On the next day, the same cells were either transfected with the proviral plasmid DNA pNL4-3 or infected with HIV-1 (MOI 0.5). The total RNA was harvested 48 hours later and used for northern blot analysis.
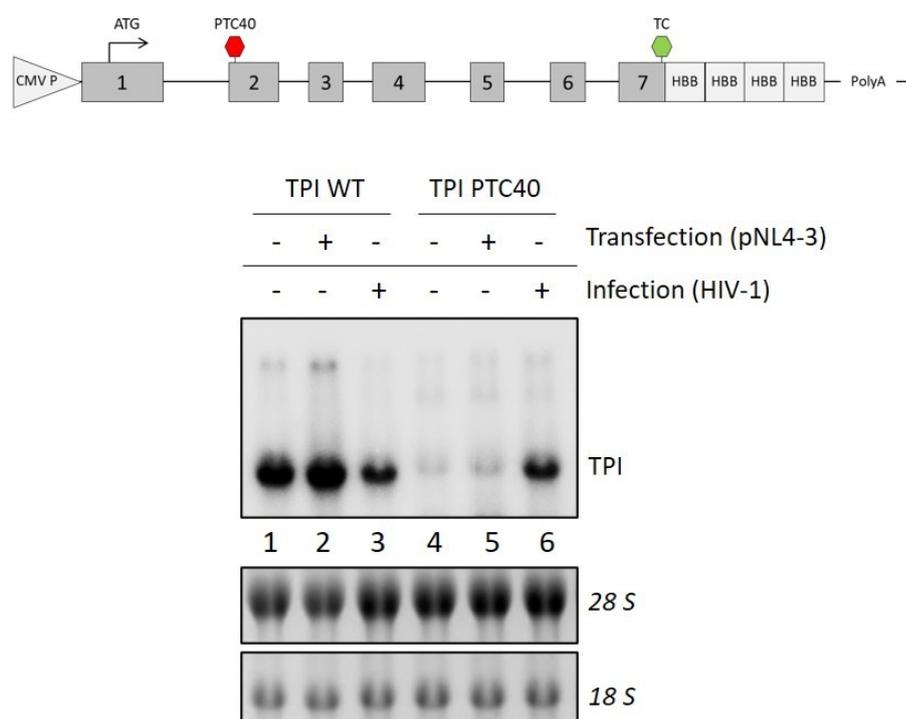
Figure 3.11: Infection with HIV-1, but not transfection with pNL4-3, reduced the cellular NMD activity. (Top) Schematic drawing of the transiently transfected triosephosphate isomerase (TPI) reporter constructs. The PTC-containing variant had a pre-mature stop codon (PTC) at position 40. The HBB repeats served as the binding sequence for the northern blot probe. **(Bottom)** HeLa T4+ cells were transiently transfected with the corresponding reporters. 24 hours later the cells were either transfected with pNL4-3 or infected with HIV-1 (MOI 0.5). 48 hours later, the total RNA was isolated, separated on formaldehyde containing 1% agarose gel, and transferred to a nylon membrane. Detection was performed with a specific probe. The detection of the ribosomal RNAs served as the loading control.

The Northern Blot was performed in cooperation with Dr. Volker Böhm, Institute for Genetics in Cologne. The figure was adapted from Dr. Nora Diehl Ph.D. thesis (Diehl 2016).

Based on the result in Figure 3.11 it appeared as the infection with HIV-1, but not the transfection of proviral DNA resulted in a down regulation of the cellular NMD machinery, visualized by the accumulation of the PTC containing transcript upon infection. However, as promising as this result looked, the attempt to reproduce this experiment led to controversial outcomes. Not only the PTC-containing variant was affected by an HIV-1 infection or transfection of proviral DNA, but also the WT mRNA levels (data not shown). Therefore, the basic structure of the TPI WT and NMD reporters as well as their corresponding NMD-sensitivities were investigated. As cycloheximide treatment priorly resulted in specific accumulation of NMD-sensitive splice isoforms (cf. Figure 3.5) HeLa cells, which were transiently transfected with the TPI reporter constructs, were treated with cycloheximide. Additional to the co-translational NMD inhibition with CHX two other potential NMD inhibitors were tested, those were puromycin and caffeine. Puromycin inhibits the translation elongation, by its structural resemblance to the 3' end of an amino acylated tRNA (aa-tRNA). Hence, when puromycin enters the A-site of the translocating ribosome it gets incorporated into the C-terminus of the growing peptide chain, thereby preventing further extension, resulting in pre-mature translation termination (reviewed in (Aviner 2020)). Caffeine, on the other hand, was described to reduce the NMD activity by interference with the UPF1 phosphorylation cycle, by inhibition of SMG1 (Keeling et al. 2013). To test the potential NMD inhibitors on the before described pCI-TPI reporters, HeLa cells were transiently transfected with either TPI WT or TPI PTC40 for 24 hours. Before harvesting the total RNA, the cells were treated with either CHX (50 µg/ml for 6 hours), puromycin (100 µg/ml for 4 hours), or 10 mM caffeine for 4 hours.
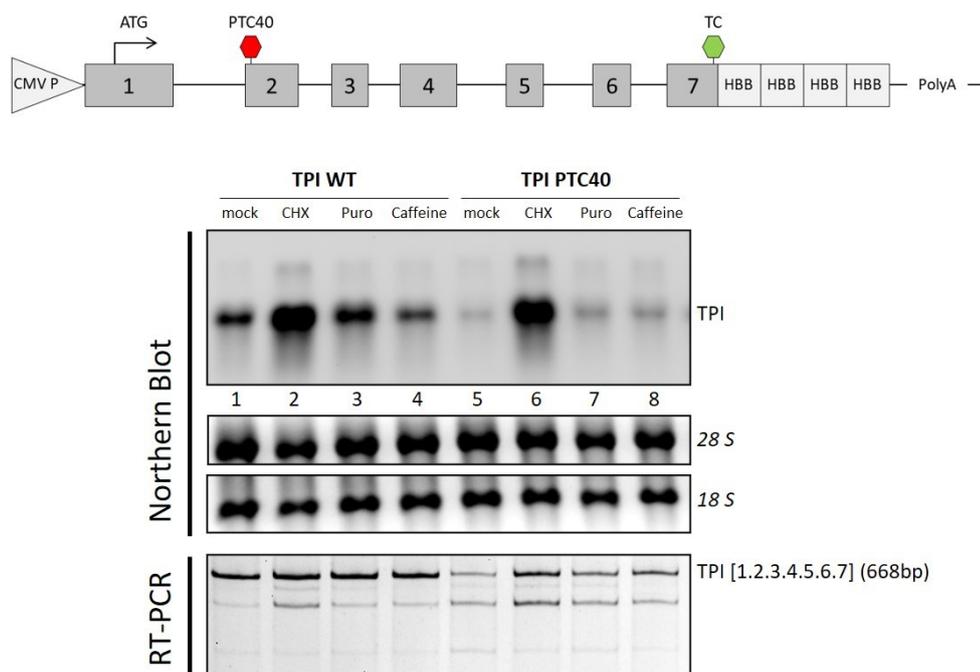
Figure 3.12: Transiently transfected TPI reporters produce unexpected alternative splice isoforms. Test of the TPI WT and TPI PTC40 reporters upon CHX treatment in comparison with potential other NMD inhibitors. The transiently transfected reporters produced alternative splice isoforms, resulting in NMD-sensitive mRNAs even in the WT. **(Top)** schematic drawing of the used TPI reporter constructs. **(Bottom)** HeLa cells were transiently transfected with 1 µg of the respective TPI reporter construct for 24 hours and treated with either 50 µg/ml CHX (6h), 100 µg/ml Puromycin (4h), or 10 mM Caffeine (4h). The northern blot of the isolated RNA. 5 µg of total RNA was treated with DNase I and separated on a 1% formaldehyde-containing agarose gel, capillary blotted to a nylon membrane and stained with a specific 4x HBB DIG-probe (#5651/#5599). The ribosomal bands served as loading control. 3 µg of the same RNAs were reverse transcribed and amplified with specific primer pairs (#5664/#5665, 20 cycles). The PCR products were separated on a 10% PAA gel, stained with EtBr, and visualized by UV light. The expected band size for TPI with this primer pair was 668bp.

The northern blot analysis of HeLa cells, which were transiently transfected with the original pCI-TPI reporters, as they were previously used by Dr. Nora Diehl, is shown in Figure 3.12. HeLa cells, which were transfected with the PTC-containing variant (TPI PTC40) showed lower TPI mRNA levels than the cells transfected with the TPI WT plasmid, as expected as these PTC-containing mRNAs were degraded by NMR (cf. Figure 3.12 top, mock samples, lanes 1 and 5). The PTC-containing TPI mRNA levels varied, depending on the treatment with one of the three chemicals (lanes 5-8). The most prominent effect of the TPI PTC40 mRNA levels was obtained by cycloheximide treatment (cf. lanes 5 and 6), while puromycin and caffeine treatment only resulted in a minimal increase of the mRNA levels (cf. lane 5 to lanes 7 and 8). Surprisingly, not only the PTC-containing TPI mRNA levels were affected by the chemical treatment, but also the WT-TPI mRNA levels (lanes 1-4). The CHX treatment resulted in a strong, unexpected accumulation of the WT RNA (cf. lanes 1 and 2). This accumulation was surprisingly almost as prominent as the accumulation of the PTC-containing RNAs upon CHX treatment. Upon puromycin or caffeine treatment, however the RNA levels of both, TPI WT, as well as their PTC-

containing counterparts only marginally increased (cf. Figure 3.12, lane 1 to 3 and 4, as well as lane 5 to 7 and 8).

Since the highest specificity of the CHX-dependent accumulation of NMD-targets was priorly observed in RT-PCR experiments (cf. Figure 3.5), RNA from the same experiment was used for RT-PCR analysis (Figure 3.12, bottom). The TPI messages were amplified with specific primers, that bind within exon 1 and 7 of triosephosphate isomerase. However, the PAA-gel revealed, that not only the expected PCR product of 668bp was amplified, but the gel showed multiple bands which indicated alternative splice isoforms of TPI. These alternative splice isoforms were not only unexpected, but they also increased upon CHX treatment, indicating NMD-sensitivity. Since the used primer pair not only amplified the plasmid-derived TPI RNAs but also the endogenous TPI RNAs, a different forward primer was designed, which only amplified the plasmid-encoded transcripts. To directly compare the transcript isoforms of cellular and plasmid-encoded TPI RNAs another experiment was performed, in which the TPI expression of HeLa cells was compared between cells that have been transfected with the TPI WT plasmid and not-transfected cells. The results are shown in Figure 3.13. All TPI splice variants were identified by isolation of the different PCR bands, reamplification, and sequencing. The amplification of HGH served as transfection control, while SRSF3 was used as CHX-specific control regarding NMD inhibition.
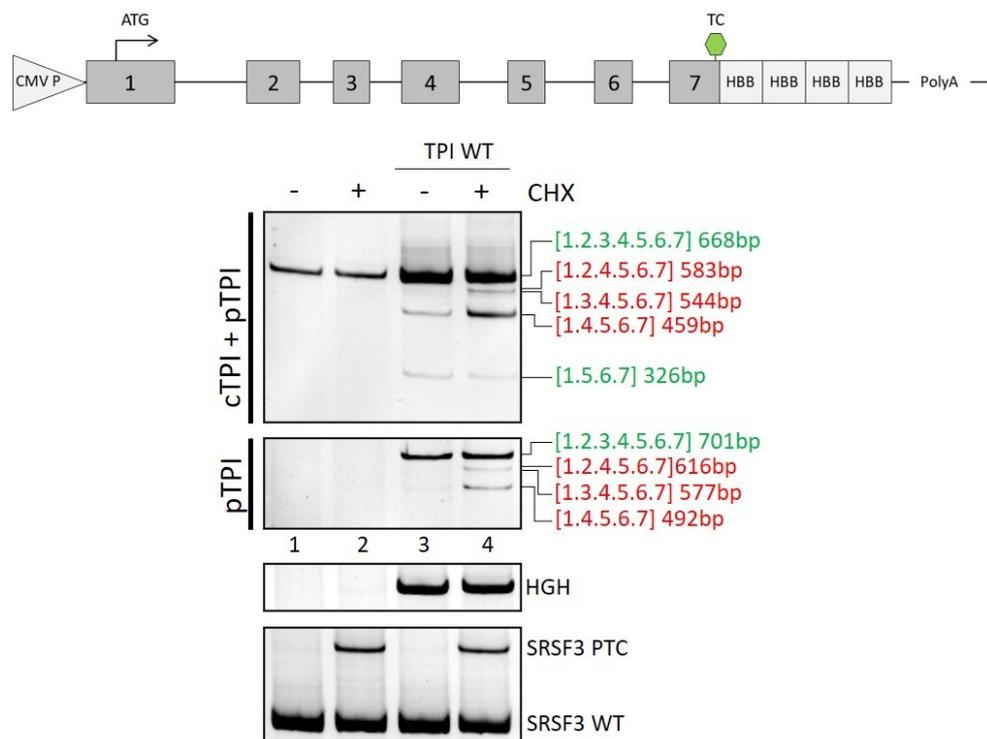

Figure 3.13: The multiple TPI splice isoforms derived from the TPI plasmid. (Top) Schematic drawing of the used TPI WT reporter plasmid. **(Bottom)** RT-PCR results of untreated HeLa cells (lanes 1 and 2) and HeLa cells transfected with the pCI-TPI-WT plasmid (lanes 3 and 4). The genes of interest were amplified with specific primer pairs (cellular TPI

(cTPI) + plasmid TPI (pTPI): #5664/#5665 26 cycles, pTPI: #4324/#5665 18 cycles, HGH transfection control #1224/#1225 26 cycles, NMD control SRSF3 #4003/#4004 (Diehl 2016) 35 cycles), separated on a 10% PAA gel and visualized by EtBr staining and UV light exposure. The TPI bands were excised, reamplified, and sequenced. Numbers in brackets represent the exon numbers that were included in the splice isoform. Predicted NMD targets with frameshift mutations are highlighted in red letters, while not-NMD targets with intact ORFs are represented by green letters.

The results in Figure 3.13 confirmed that the endogenous cellular TPI only produced a single transcript, which consisted of the TPI exons [1.2.3.4.5.6.7] and was not sensitive to CHX treatment (cf. Figure 3.13 cTPI + pTPI, lanes 1 and 2). Hence, all other observed bands derived from the transfected plasmid TPI and were amplified by both used TPI primer pairs (cf. Figure 3.13 cTPI + pTPI (#5664/#5665) with pTPI (#4324/5665)). Furthermore, some of the amplified splice isoforms accumulated upon CHX treatment, even though only the WT plasmid was transfected. The subsequent excision, amplification, and sequencing of these bands allowed the assignment of the included TPI exons to each splice isoform. While the expected WT TPI transcripts consisted of the exons [1.2.3.4.5.6.7] and encoded for a 250 amino acid protein (ORF of 750 nucleotides), the other observed isoforms all excluded one or more TPI exons. The skipping of some TPI exons resulted in the disruption of the ORF, hence the generation of frameshift mutations and pre-mature translation termination. For example, the exclusion of TPI exon 3 resulted in a frameshift mutation, as exon 3 had a length of 85bp ([1.2.4.5.6.7]). The same observation was made for the exclusion of TPI exon 2, with a length of 124bp ([1.3.4.5.6.7]), or the exclusion of both, TPI exon 2 and 3 (together 209bp, [1.4.5.6.7]). The only exception for the observed alternative splicing-dependent NMD-sensitivity of the obtained alternative splice isoform was found for the simultaneous exclusion of TPI exon 2, 3, and 4 (together 342bp, [1.5.6.7]). The exclusion of all three exons resulted in no frameshift mutation and therefore generated no NMD susceptibility. As expected, from the prior CHX experiments, the TPI splice isoforms with predicted frameshift mutations accumulated upon CHX treatment. Therefore, not only the TPI PTC plasmid encoded for NMD-sensitive splice isoforms but also transfection of cells with the TPI WT plasmid resulted in NMD-sensitive splice isoforms.

Since the cellular encoded TPI, in difference to the plasmid-encoded TPI, only produced a single transcript, the sequences were aligned and scanned for differences. This comparison of the exonic and intronic sequences is shown in Figure 3.14 and revealed four sequence differences between the endogenous TPI and the plasmid-encoded TPI, which might explain the differences in splice site recognition.
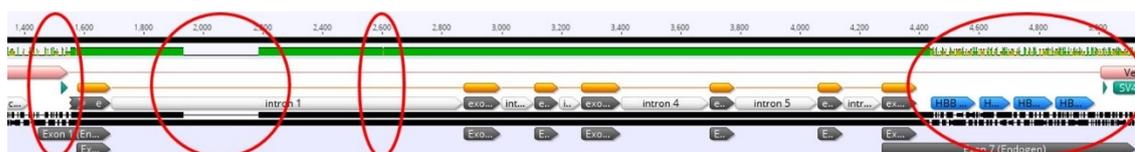

Figure 3.14: Sequence differences between the cellular encoded TPI and plasmid-encoded TPI. Four differences were found: **1)** The plasmid TPI missed the first 109 nucleotides of exon 1, counted from cellular TPI-ATG. Furthermore, the Kozak sequence of the plasmid-encoded translational start codon was modified by a G>A mutation

nucleotides in position -3 upstream of the ATG (AGC GCC <u>ATG</u> G>AGC ACC <u>ATG</u> G). **2)** Intron 1 of the plasmid-encoded TPI missed 250 nucleotides, which were present in the cellular TPI. **3)** Furthermore, one single nucleotide exchange (C>G) was found within intron 1 of the plasmid-encoded TPI, 271 nucleotides upstream of exon 2. **4)** The endogenous and the plasmid-encoded TPI share the same TC, but the downstream sequence of the plasmid-encoded TPI exon 7 was replaced by four HBB binding site repeats. The sequence alignment was done and illustrated using the Geneious software.

The first difference between the sequence of the cellular encoded TPI versus the plasmid-encoded TPI was found directly within exon 1. Here, the plasmid-encoded TPI used a different endogenous ATG codon, which was located 109 nucleotides downstream of the original TPI start codon. Furthermore, these 109 upstream bases were not included in the plasmid TPI sequence and the originally internal ATG, which was now used as translational start codon in the plasmid version of TPI was modified in position -3 by a G to A mutation. The second major difference between the cellular and plasmid-encoded TPI was found within the first intron. Here, intron 1 of the plasmid-encoded TPI lacked 250 nucleotides, which were most likely excluded from the plasmid for size reasons. The third observed difference was minor, a single nucleotide exchange from C to G within intron 1 of the plasmid TPI, 271 nucleotides upstream of exon 2. The last difference was found within exon 7, directly downstream of the endogenous translational stop codon. While the endogenous 3' UTR consisted of the TPI exon 7 sequence, the untranslated region of the plasmid-encoded TPI was replaced with a probe binding site. That probe binding site consisted of four repeats of a 100-nucleotide sequence of the human β-globin gene (HBB). While the single nucleotide exchange within intron 1 and the alterations of the 3' UTR most likely had no impact on the splicing process, the impact of the first two alterations from the endogenous to the plasmid TPI were investigated regarding the splicing outcome. Therefore, the missing sequences of either exon 1 or intron 1 were amplified from isolated HeLa DNA and inserted into the TPI plasmid sequence. The splicing outcome was again monitored by transfection of HeLa cells with the generated modified TPI WT plasmids, and RT-PCR analysis using specific primers. The results are shown in Figure 3.15.
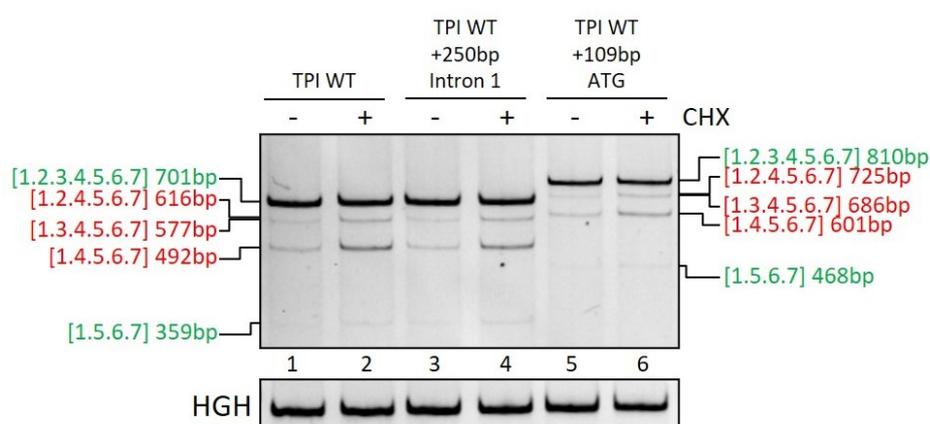
Figure 3.15: The plasmid-encoded TPI splice isoforms are independent of the performed sequence alterations. RT-PCR results of HeLa cells transfected with the depicted pCI-TPI-WT plasmids. The plasmid-encoded TPI messages were amplified with specific primer pairs (#4324/#5665, 26 cycles), HGH served as transfection control (#1224/#1225, 26 cycles). The PCR products were separated on a 10% PAA gel and visualized by EtBr staining and UV light exposure.

As a reference for the splicing pattern, HeLa cells were transfected with the original pCI-TPI-WT-4H plasmid and the obtained alternative splice isoforms are shown in Figure 3.15 lanes 1 and 2, with the labeling on the left-hand site. The insertion of the missing 250 nucleotide sequence into intron 1 had no impact on the splicing outcome (cf. lanes 3 and 4 to lanes 1 and 2). Likewise, the insertion of the 109-nucleotide sequence upstream of the plasmid-encoded ATG increased the PCR product sizes (labeled on the right-hand site) but had no further impact on the ratio of splice site usage (cf. lanes 5 and 6 to lanes 1 and 2). Hence, those sequence modifications, if at all, had only a minor impact on splice site usage of the plasmid-encoded TPI.

Since the observed alternative TPI splice isoforms from the plasmid-encoded TPI were not found in the endogenous TPI and were not abolished by adaption of the plasmid TPI structure to the endogenous sequence another approach was used. The idea was to increase the exon recognition of the plasmid TPI by manipulating the intrinsic splice donor strength. For that the splice donor of TPI exon 3 (HBond score 16.9) was chosen, by the introduction of a single point mutation (gAGGTtAGTAg>gAGGTAAGTAg, HBS 16.9 > 20.8) (Figure 3.16 A).
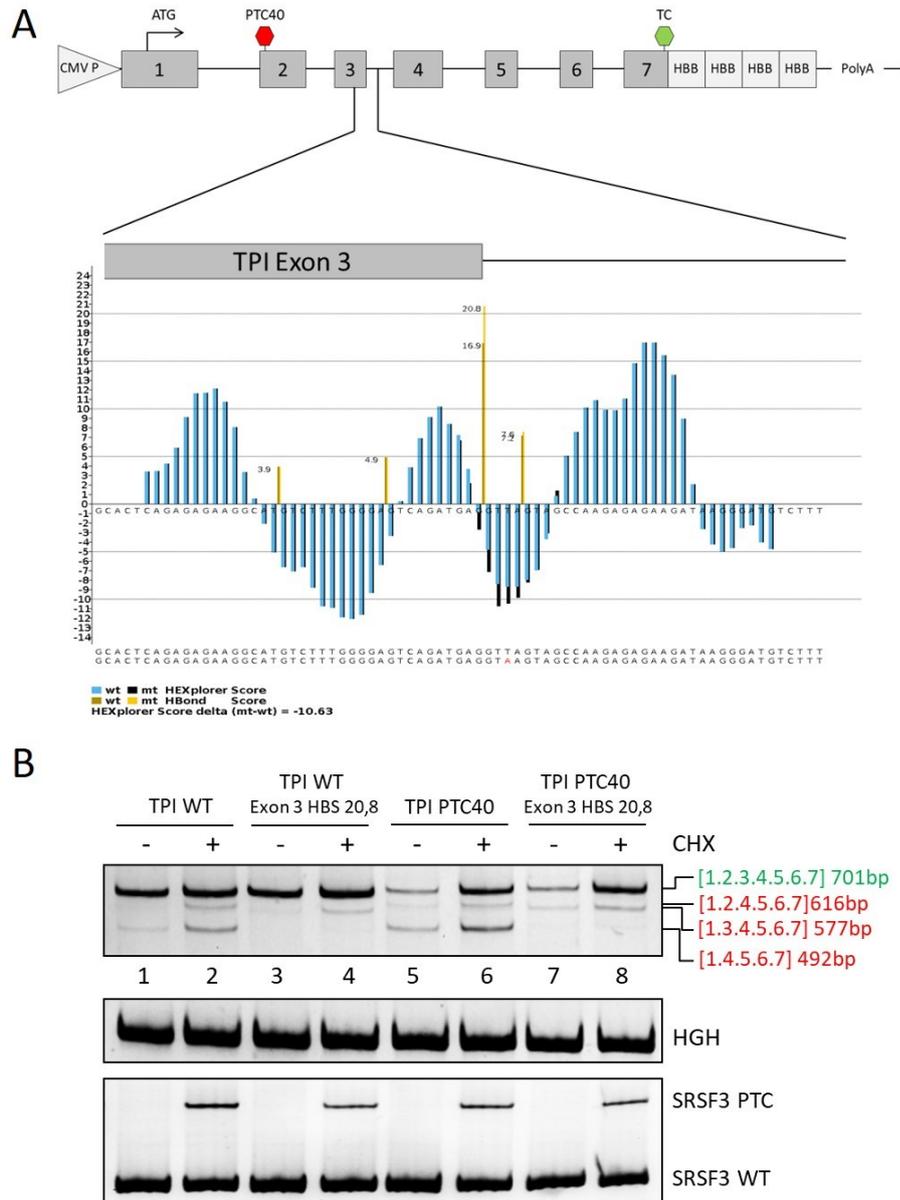
Figure 3.16: Enhanced inclusion of plasmid-encoded TPI exon 3 upon increased splice donor complementarity. (A) schematic drawing of the used TPI reporter construct. The 3' end of exon 3 and the 5' end of intron 3 were enlarged to show the introduced mutation. The SD was modified by a single point mutation (gAGGTtAGTAg>gAGGTAAGTAg, HBS 16.9 > 20.8). **(B)** RT-PCR results were obtained from the transfection of HeLa cells with the modified plasmids. HeLa cells were transiently transfected with the depicted reporter constructs for 24 hours. Six hours before the total RNA was harvested the cells were treated with 50 µg/ml CHX. The RNA was reverse transcribed and amplified with specific primer pairs (#4324/#5665 pTPI, 18 cycles). HGH (#1224/1225, 26 cycles) served as transfection control and SRSF3 (#4003/#4004, 35 cycles) as control for CHX treatment. The PCR products were separated on a 10% PAA gel, stained with ethidium bromide, and visualized via UV light.

The comparison of HeLa cells which were transfected with the SD3 modified TPI plasmids showed fewer pTPI splice isoforms than HeLa cells which were transfected with the original pCI-TPI constructs. The results in Figure 3.16 B showed that the single nucleotide substitution within spice donor 3 removed the TPI [1.2.4.5.6.7] and the TPI [1.4.5.6.7] splice isoforms from

the pattern, both of which lacked exon 3. The one alternative splice isoform which was still observable upon SD3 modification was consisting of the TPI exons [1.3.4.5.6.7]. That splice isoform lacked exon 2 and was even more abundant than before. The exclusion of the 124nt-long exon 2 resulted in a frameshift mutation, that made the resulting mRNAs sensitive to degradation by the NMD machinery, which was nicely illustrated by the accumulation upon CHX treatment (cf. Figure 3.16 B lanes 3-4 and 7-8). In conclusion, the single-base modification of the plasmid-encoded TPI SD3 resulted in increased recognition and inclusion of exon 3. However, this modification was not sufficient, as the remaining exclusion of TPI exon 2 from some splice isoforms of the TPI WT construct still resulted in NMD-sensitive transcripts. Therefore, those alternative splice isoforms had additionally to be removed before this reporter system would mimic the endogenous TPI splicing pattern.

As the intrinsic strength of TPI splice donor 2 (CAGGTGAGatc, HBond score 17.4) was higher than TPI splice donor 3, another approach was used to improve splice site recognition. Instead of modifying the splice donor site, the increased inclusion of TPI exon 2 was mediated by modification of its surrounding nucleotides increasing its Splice Site HEXplorer Weight (SSHW) (Erkelenz et al. 2014). The aim of this modification was the modulation of the HEXplorer profile to generate HEXplorer positive areas upstream of TPI SD2 and HEXplorer negative areas downstream of TPI SD2 predicted to bind SR proteins upstream and hnRNP proteins downstream of the splice site supportive for splice site recognition (Erkelenz et al. 2014). That profile modification was accomplished by the mutation of four nucleotides, two up and two downstream of TPI SD2 (cf. Figure 3.17 A). The substitutions of the exon 2 nucleotides upstream of SD2 were chosen with respect to the coding sequence, as that region belonged to TPI exon 2. Therefore, only the third base of each codon was modified, which resulted in silent mutations not altering the amino acid code (cf. Figure 3.17 A, ACT-GGG-GAG-AT<u>C</u>-<u>AG</u> > AC<u>C</u>-GG<u>A</u>-GAG-AT<u>C</u>-<u>AG</u>). Downstream of SD2 it was not necessary to maintain the amino acid code, as this sequence was located within TPI intron 2. Therefore, the mutations were chosen based on their impact on the HEXplorer profile, concerning the 11-nucleotide splice donor sequence (<u>GTGAGATC</u>GAGGTGG > <u>GTGAGATC</u>TAGG<u>C</u>GG). The overall modifications of the HEXplorer profile were monitored by the HEXplorer score delta (mt-wt), also known as the ΔHZ_{EI}, with a score of 35.3. The newly generated reporter plasmids, with modifications of either TPI exon 2, exon 3, or both, were transfected into HeLa cells and the results of the RT-PCRs are shown in Figure 3.17 B.
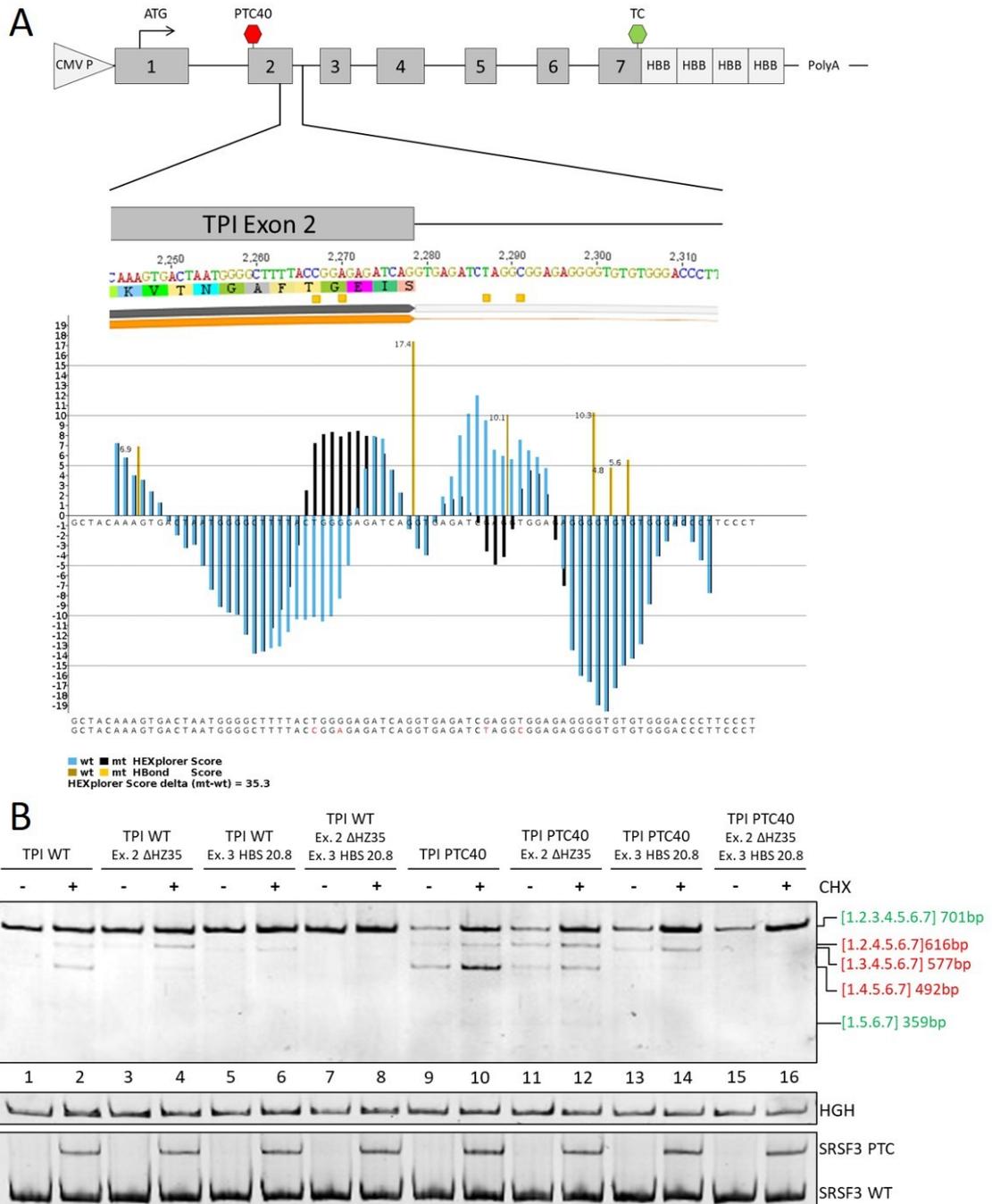
Figure 3.17: HEXplorer-based modifications increased the TPI exon 2 inclusion. The combination with increased TPI splice donor 3 complementarities finally resulted in a single TPI transcript. **(A)** schematic drawing of the used TPI reporter construct. The 3' end of exon 2 and the 3' end of intron 2 were enlarged to illustrate the inserted mutations. The HEXplorer profile was modified by two silent point mutations within exon 2, which maintained the amino acid sequence (ACT-GGG-GAG-ATC-AG > ACC-GGA-GAG-ATC-AG), and two point-mutations within intron 2, which were chosen concerning the 11-nucleotide splice donor site (CAG/GTGAGATCGAGGTGG > CAG/GTGAGATCTAGGCGG). **(B)** HeLa cells were transiently transfected with one of the depicted reporter constructs for 24 hours. Six hours before the total RNA was harvested the cells were treated with 50 µg/ml CHX. The RNA was reverse transcribed and amplified with specific primer pairs (#4324/#5665 pTPI, 20 cycles). HGH (#1224/1225, 26 cycles) served as transfection control and SRSF3 (#4003/#4004, 35 cycles) as control for CHX treatment. The PCR products were separated on a 10% PAA gel, stained with ethidium bromide, and visualized via UV light.

As mentioned before, the modification of TPI exon 3 was sufficient to abolish the TPI [1.2.4.5.6.7], as well as the TPI [1.4.5.6.7] splice isoforms, which both lacked TPI exon 3 (cf. Figure 3.16 B and Figure 3.17 B). The modification of TPI exon 2 alone, however, showed no difference in the splicing pattern in comparison to the pCI-TPI original plasmid (cf. Figure 3.17 B, WT: lanes 1-4 and PTC40: lanes 9-12). However, the combination of both, the increased HBond score of the TPI SD3 together with the TPI exon 2 HEXplorer score modification finally resulted in a single splice isoform of plasmid-encoded TPI. This splice isoform included all TPI exons ([1.2.3.4.5.6.7]) and only the PTC-containing transcript showed an accumulation to CHX treatment (cf. Figure 3.17 B, lanes 7 and 8 (WT) and lanes 15 and 16 (PTC40)).

3.5 Transiently transfected NMD reporter plasmids reveal no modulation of the NMD activity by HIV-1 infection

Finally, the newly generated TPI reporter plasmids with a single splice variant were used to repeat the initial experiment (cf. Figure 3.11) to analyze the impact of HIV-1 infection in comparison to the transfection of proviral DNA on the activity of the nonsense-mediated RNA decay in HeLa T4$^+$ cells. Therefore, the cells were transiently transfected with the TPI NMD reporter system 24 hours before the infection, or pro-viral transfection. To include the infection process and the ongoing viral replication into the analysis, three different time points were chosen for the sample collection. Therefore, the total RNA was harvested 6 hours, 24 hours, or 48 hours post-infection, or pNL4-3 transfection, and used for northern blot analysis. For RNA detection, two different northern blot probes were used. The first probe targeted the 4x HBB repeat of the TPI reporter constructs to monitor the NMD activity. The other probe targeted the HIV-1 exon 7 sequence, visualized the viral mRNA classes, and thereby monitored the infection progression. The results are shown in Figure 3.18.
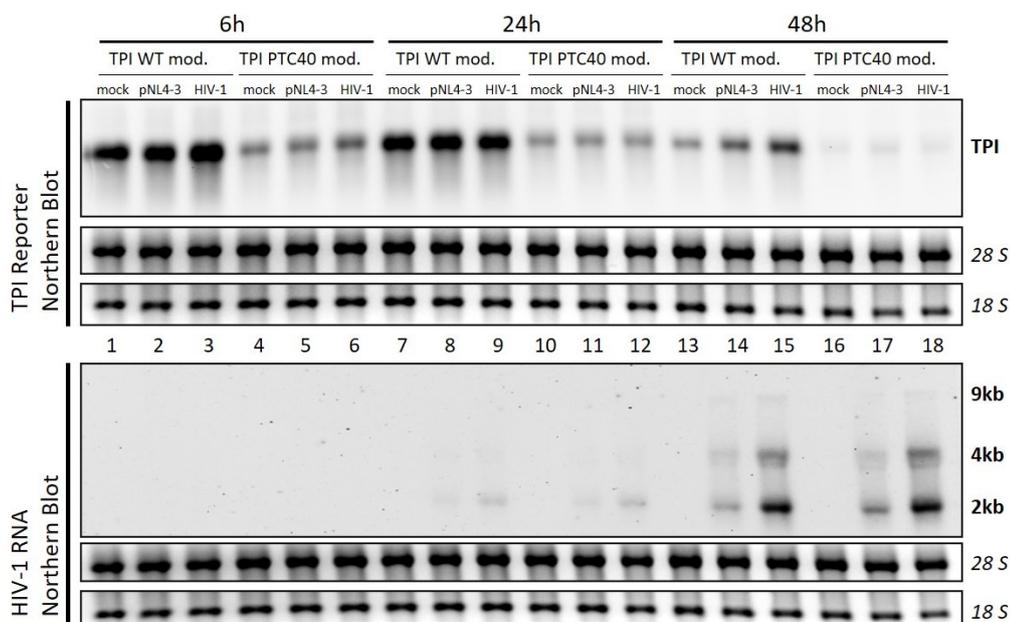
Figure 3.18: Neither the HIV-1 infection nor the pNL4-3 transfection modified the abundance of the NMD-sensitive TPI Reporter in HeLa T4+ cells. HeLa T4+ were transiently transfected with 1 µg of the modified versions of the depicted TPI reporter constructs (TPI Exon 2 ΔHZ35 and Exon 3 HBS 20.8, abbreviated in the Figure with "TPI mod."). Twenty-four hours post reporter transfection the cells were either transfected with 1 µg of the proviral DNA (pNL4-3) or infected with HIV-1 (MOI 0.5). The total RNA was harvested 6 hours, 24 hours, or 48 hours later and used for northern blot analysis. For each blot 4 µg of total RNA were DNase I treated, separated on a 1% formaldehyde-containing agarose gel and capillary blotted overnight onto a nitrocellulose membrane. The TPI reporter RNAs were detected using the 4x HBB probe (#5651/#5599, top), the HIV RNAs were detected using the HIV-1 exon 7 probe (#3387/#3388, bottom). The ribosomal RNAs served as loading control.

The results in Figure 3.18 showed that the viral RNA levels increased over time, as expected from the different time points post HIV-1 infection or pNL4-3 transfection (Figure 3.18 B, bottom, HIV-1 RNA). Upon the first analyzed time point, which was 6 hours post-infection or transfection, no viral mRNA classes were detected in the northern blot analysis (Figure 3.18 B, bottom, lanes 1-6). That observation differed 18 hours later, as the 24h time point already showed detectable levels of HIV-1 2kb and 4kb RNAs (Figure 3.18 B, bottom, lanes 7-12). Another 24 hours later, on the 48-hour time point, even the HIV-1 9kb class was visible in the northern blot (Figure 3.18 B, bottom, lanes 13-18). In general, the HIV-1 specific RNA levels produced by HeLa T4+ cells were higher upon infection (HIV-1 MOI 0.5) than upon transfection with proviral DNA (1 µg pNL4-3). Even though the infection, as well as the transfection with proviral DNA, was productive, shown by the increased detection of the viral RNAs over time (cf. Figure 3.18 B, bottom), the detection of the transfected plasmid TPI reporters was difficult at later time points. At the 6-hour time point, which corresponded to 30 hours post TPI reporter transfection, the levels of TPI WT RNA were, as expected, much higher than the mRNA levels of the corresponding TPI PTC40 reporter. Neither the HIV-1 infection nor the transfection with pNL4-3 influenced the RNA levels of the PTC-containing TPI RNAs (cf. Figure 3.18 B, top, lanes 1-6). The next analyzed time point (24 hours, lanes 7-12) equaled 48h post TPI reporter transfection. Here, as the viral RNAs became visible in the northern blot, again neither the TPI WT, nor the TPI PTC40 RNA levels were

influenced by infection with HIV-1, or pNL4-3 transfection (Figure 3.18 B, top, lanes 7-12). As mentioned before, the last chosen time point, 48 hours post-HIV-1 infection or pNL4-3 transfection, nicely showed the HIV-1 mRNA classes (Figure 3.18 B, bottom, lanes 13-18). However, these samples were taken 72h post TPI reporter transfection and the plasmid-derived TPI RNAs were almost not detectable (cf. Figure 3.18 B, top, lanes 13-18). That either indicated the degradation of the transiently transfected plasmids or the loss of the plasmids due to cell division without plasmid propagation. However, independent of the reason for the plasmid loss, the potential virus-induced modification of the NMD pathway, monitored by the TPI reporter plasmids, could not be determined for this time point. In conclusion, transiently transfected NMD reporter plasmids show different layers of complexity. Not only did the TPI WT reporter produce NMD-sensitive isoforms by alternative splicing, but these reporter systems are furthermore not suited for time-course experiments, as they were expressed unreliably at time points ≥ 72h.

However, one more shot was taken with the transiently transfected reporter plasmids using a different cell line, which were the HEK293T CD4$^+$ cells. These cells are not only sensitive for HIV-1 infection, as they express the CD4-receptor on their surface, but these cells also expressed the SV40 large T antigen. This expression should circumvent the problem of plasmid loss over time, as the large T antigen allows the amplification of plasmids that contained an SV40 origin of replication (ori). Therefore, the HEK293T CD4$^+$ cells should maintain the transiently transfected plasmid for longer periods than the HeLa T4$^+$ cells and the experimental time frame was extended up to 72h post-HIV-1 infection, or pNL4-3 transfection. The experimental setup, however, was kept similar as for the HeLa T4$^+$ cells (Figure 3.18), and the HEK293T CD4$^+$ cells were transiently transfected with the modified TPI WT or PTC40 reporter constructs for 24 hours followed by HIV-1 infection or pNL4-3 transfection. The total RNA was harvested 24h, 48h, and 72h later, which resembled 96 hours post reporter transfection, and used for northern blot analysis. The results are shown in Figure 3.19.
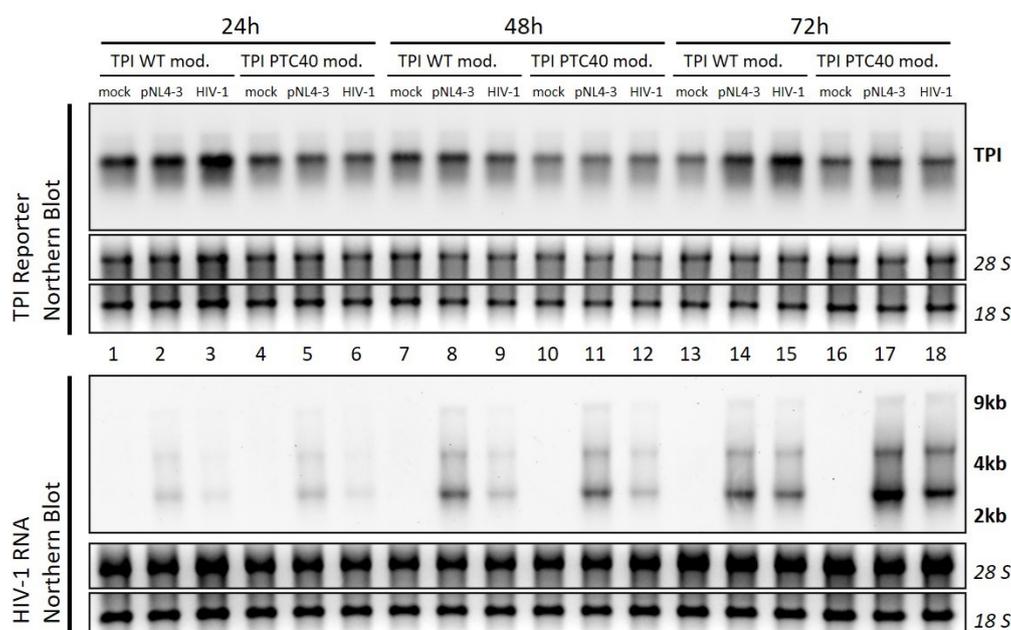
Figure 3.19: Neither the efficient HIV-1 infection nor the transfection with proviral DNA modulated the TPI NMD reporter RNA abundance in transient transfection experiments with HEK293T CD4+ cells. HEK293T CD4+ cells were transiently transfected with 1 µg of the modified versions of the depicted TPI reporter constructs (TPI Exon 2 ΔHZ35 and Exon 3 HBS 20.8, abbreviated with "TPI mod."). 24 hours later they were either transfected with 1 µg of the proviral DNA (pNL4-3) or infected with HIV-1 (MOI 0.5). The total RNA was harvested either 24 hours, 48 hours, or 72 hours later and used for northern blot analysis. For each blot, 4 µg of total RNA were DNase I treated and separated on a 1% formaldehyde-containing agarose gel and capillary blotted overnight onto a nitrocellulose membrane. The TPI reporter RNAs were detected using the 4x HBB probe (#5651/#5599, top), the HIV RNAs were detected using the HIV-1 exon 7 probe (#3387/#3388, bottom). The ribosomal RNAs served as loading control.

As the HEK293T CD4+ cells should amplify the transiently transfected plasmids, the sample time was extended in comparison to the HeLa CD4+ experimental setup. The first HEK293T CD4+ RNA samples were taken 24 hours post-HIV-1 infection, or pNL4-3 transfection, so 48 hours post TPI reporter transfection. Already in this first analyzed time point, the HIV-1 mRNA classes were detectable, even though the viral RNAs had higher intensities upon pNL4-3 transfection than upon HIV-1 infection (cf. Figure 3.19, bottom, lanes 1-6). The amount of viral RNA increased further over the time course of the experiment, indicating an efficient viral replication (cf. Figure 3.19, bottom). Regarding the RNA levels of the transiently transfected TPI reporter plasmids the abundance of the TPI PTC40 mRNAs in HEK293T CD4+ cells was much higher than observed before in HeLa T4+ cells (cf. Figure 3.18). That suggested a decreased NMD-dependent degradation of PTC-containing mRNAs in HEK293T CD4+ cells in comparison to HeLa T4+ cells. That result mirrored the priorly observed cell type-dependent abundancies of NMD targets (cf. Figure 3.5). However, regardless of the infection with HIV-1, or the transfection with pro-viral DNA, the TPI PTC40 mRNA levels in the HEK293T CD4+ cells remained at a constant level, suggesting no impact on the general activity of the NMD pathway in this transfection-based experimental setup (cf. Figure 3.19, top).

However, since we observed a cell type-dependent difference in the efficiency of NMD target recognition (cf. Figure 3.5, Figure 3.18, and Figure 3.19) and it was furthermore reported that transient transfections can impair the NMD efficiency of different cell types (Gerbracht, Boehm, and Gehring 2017), another approach was tested to elucidate the impact of HIV-1 infection on the abundance of NMD reporter systems. Therefore, a lentiviral vector system was designed encoding for the NMD reporter system of interest. The subsequent transduction of cell culture cells with these lentiviral particles would therefore result in the stable integration of the reporter construct into the cellular DNA.

3.6 The establishment of a lentiviral vector system for NMD reporter transduction causes unexpected difficulties

As mentioned before, transfection experiments using the modified TPI reporter system were not suited for time-course experiments. That was either due to the loss of the transiently transfected plasmids (cf. Figure 3.18) or based on the impairment of the cellular NMD pathway upon transfection (cf. Figure 3.19). Therefore, it was decided to generate a lentiviral vector system, which contained either the WT or PTC-containing NMD reporter as gene of interest (GOI). The use of lentiviral vectors offered two major advantages. First, the target cells do not need to be transfected with reporter plasmids, which might already impair the NMD activity. Second, the GOI gets stably integrated into the target cell's genome, ensuring the transmission of the GOI to daughter cells upon cell division.

The lentiviral vector backbone, used for this work was described in detail by Dr. Katharina Röllecke (Röllecke 2016) and a structural overview of the vector structure is shown in Figure 3.20 A. The HIV-1 long terminal repeats (LTRs) were modified to obtain a self-inactivating vector, which remained inactive upon integration into the host cells genome. The lentiviral vector was expressed under the control of a CMV promotor, which made it independent of *tat* levels. The expression of the GOI was mediated by the internal SFFV U3 promoter. As the promoter was already included within this vector system, the modified TPI WT or TPI PTC40 cassettes were cloned in between the *KpnI* and *NotI* restriction sites (cf. Figure 3.20 A and B). Next, to generate infectious lentiviruses, HEK293T cells were transfected with the depicted third-generation lentiviral vector, containing the TPI cassette of interest, as well as the structural elements which are required for RNA packaging (Figure 3.20 B), together with two helper plasmids. The first helper plasmid encoded for gag and pol, as well as other accessory HIV-1 genes (pCD/NL-BH), while the second helper plasmid encoded for the glycoprotein G of vesicular stomatitis virus (pczVSV-G).
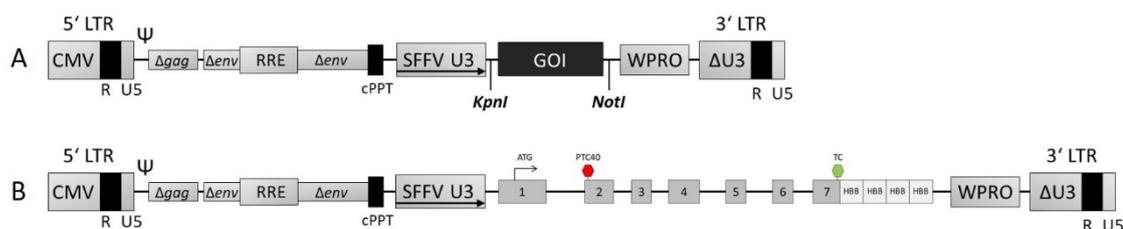
Figure 3.20: Schematic drawing of the third-generation lentiviral vector systems. (A) Structure of the used lentiviral vector before insertion of the GOI. The modified HIV-1 5' LTR contained the CMV promotor, which made it independent of *tat* levels. The promotor was followed by the R, U5, and HIV-1 leader sequence, containing the packaging signal and packaging relevant sequences within the gag and env genes. The expression of the GOI was controlled by the endogenous SFFV U3 promoter and the optimized woodchuck hepatitis virus posttranscriptional regulatory element (WPRO). To ensure the self-inactivating (SIN) capacities of the lentiviral vector system, the U3 promoter, and enhancer region were deleted which inactivated the newly generated 5' LTR upon reverse transcription. **(B)** The TPI reporter cassette was inserted between the *KpnI* and *NotI* restriction sites. Either the TPI WT or the TPI PTC40 mod. reporter cassette with its corresponding introns (TPI Exon 2 ΔHZ35 and Exon 3 HBS 20.8, abbreviated with "TPI mod."). The schematic drawing was adapted from Dr. Katharina Röllecke (Röllecke 2016). The full plasmid card is shown in the Appendix, Figure 6.2.

First, to determine the infectivity of the produced lentiviruses, the integration success, and the subsequent expression of the GOI, HeLa T4$^+$ cells were transduced with different volumes of viral supernatant, either TPI WT or TPI PTC40, and incubated for 48 hours. The total RNA was harvested and used for northern blot experiments and RT-PCRs and the results are shown in Figure 3.21.
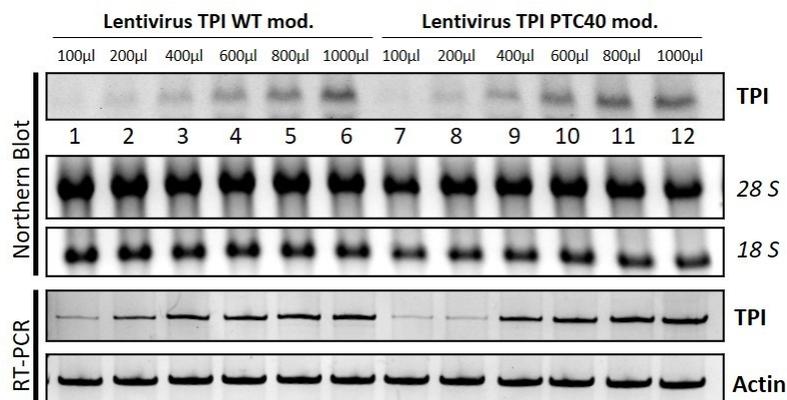
Figure 3.21: Test of the generated TPI lentiviruses showed similar RNA levels for either TPI WT or TPI PTC40 derived transcripts. HeLa T4$^+$ cells were transduced with the depicted volume of lentiviral supernatant, encoding either the TPI WT mod. or the TPI PTC40 mod. Sequences (TPI Exon 2 ΔHZ35 and Exon 3 HBS 20.8, abbreviated with "TPI mod."). The total RNA was harvested 48h later. **(Top)** Northern blot analysis: 5 µg of the total RNA were separated on a 1% formaldehyde containing agarose gel, capillary blotted overnight onto a positively charged nylon membrane, and used for northern blot analysis. The TPI reporter RNAs were detected using the 4x HBB probe (#5651/#5599). The ribosomal bands served as loading control. **(Bottom)** RT-PCR analysis: 3 µg of the total RNA were reverse transcribed into cDNA and amplified with specific primer pairs (TPI (#6128/#5565, 26 cycles), actin (#6056/#6057, 26 cycles)). The PCR products were separated on a 10% PAA gel and visualized by EtBr staining and UV-light exposure.

To monitor the total TPI RNA levels, which were produced from the lentiviral TPI transgene, the isolated HeLa T4$^+$ RNA was used for northern blot analysis and stained using the DIG-labeled 4x HBB probe. As expected from the increasing volumes of used viral supernatant, the TPI RNA

106

levels increased in a stepwise manner (cf. northern blot in Figure 3.21, top). However, that increase was expected for the TPI WT samples (lanes 1-6), but not for the TPI PTC40 samples (lanes 7-12), as those RNAs should be co-translationally degraded due to the pre-mature termination codon within exon 2. To further analyze those TPI messages, the RNAs were reverse transcribed and amplified with specific primer pairs, which detected only the lentiviral encoded TPI. Here, the observed PCR pattern resembled the results of the northern blot analysis (cf. Figure 3.21, bottom). The TPI transcript levels increased hand in hand with the used volume of lentiviral supernatant, independent of the TPI version with or without a pre-mature termination codon.

One possibility for the loss of NMD-sensitivity might be the exclusion of the TPI introns via splicing before packaging of the lentiviral genome. That might happen when the HEK293T producer cells express the transfected lentiviral genome plasmid by the usage of the CMV promoter, and the cellular spliceosome recognized and excised the TPI introns. Hence, the lentiviral RNA genome, which finally got packaged contained an already spliced TPI message. That spliced TPI then got stably integrated into the HeLa T4$^+$ genome, was expressed under the control of the SFFV U3 promoter, and left the nucleus without further need of splicing. Hence, the obtained transcripts were not loaded with EJCs and the pre-mature termination codons (PTC40) were not detected by the NMD machinery.

To investigate whether the integrated TPI gene contained introns, HeLa T4$^+$ cells were again transduced with the lentiviral supernatant and analyzed 48 hours later, this time not only for their TPI RNA levels but also on DNA levels. The results are shown in Figure 3.22.
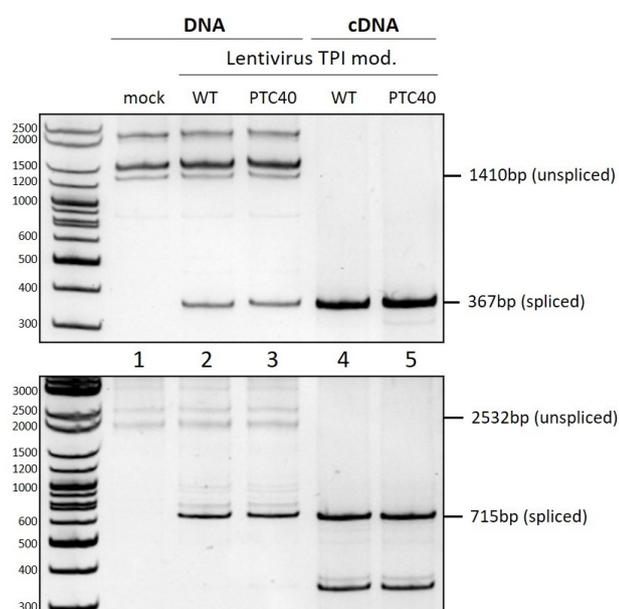

Figure 3.22: DNA analysis upon lentiviral transduction confirmed the integration of the spliced TPI isoform. HeLa T4$^+$ cells were transduced with 1000 μl of lentiviral containing supernatant of either TPI WT mod. or TPI PTC40 mod. for

48 hours. The cells were harvested, and one-half of the cells was used for DNA isolation, while the other half was used for total RNA isolation. Upon reverse transcription of the RNA, the obtained DNA and cDNA were used for PCR amplification using specific primers. Location of the primer binding sites: #6128 3' end of the SFFV U3 promoter, #5673 within TPI exon 3, #5665 within TPI exon 7. **(Top)** Primer pair #6128/#5673, expected band size with introns: 1410bp, without introns: 367bp. **(Bottom)** Primer pair #6128/#5665, expected band size with introns: 2532bp, without introns: 715bp.

The isolated HeLa T4$^+$ DNA, as well as the reverse transcribed RNA, were amplified with two different primer pairs. The first primer pair, which was used for the analysis of the lentiviral transduced TPI, was #6128/#5673 (Figure 3.22, top). Using this primer pair, the sequence from the SFFV U3 promoter to TPI exon 3 was amplified. For the isolated DNA, the expected band size with all TPI introns was 1410bp, and for the cDNA without TPI introns 367bp. The spliced TPI isoform was detected in both, the HeLa T4$^+$ DNA, as well as the cDNA, in both lentiviral vectors (TPI WT mod. and TPI PTC40 mod.). The intron-containing TPI DNA was expected to produce a PCR product of 1410bp. However, the observed band was not only found in the transduced samples but also in the mock sample and was therefore not specific for the TPI messages. To verify the observed results a second primer pair was used, which amplified the whole TPI gene from the SFFV U3 promoter to TPI exon 7 (#6128/#5665), resulting in the expected band sizes of 2532bp for the unspliced variant and 715bp for the spliced variant. Again, no clear PCR product was observed for the unspliced TPI DNA. Furthermore, the spliced 715bp product was found within the DNA samples, again indicating the TPI intron excision before the assembly of the lentiviruses within the producer HEK293T cell line. Within the cDNA samples, not only the expected 715bp product was detected, but also a smaller product of around 350bp, which could represent a new splice isoform from the already spliced TPI message, as the joined sequence of TPI exon 1 and 2 generated a new potential splice donor site (gAGGTggtTtg, HBS 10.1). However, since these messages showed no difference between the TPI WT and PTC40 sample (cf. Figure 3.22, bottom, lanes 4 and 5) this splicing event did not restore the NMD-sensitivity of the PTC containing variant and was therefore not further investigated.

As the insertion of an intron-containing NMD reporter into the lentiviral vector structure resulted in splicing of the desired reporter before the assembly of the lentiviruses, and no further splicing was needed for the nuclear export of the NMD reporter upon integration into the transduced cells genome, the NMD reporter turned out not to be suitable for this approach. Hence, a strategy was needed that would allow the retention of the NMD reporter-introns until the GOI was stably integrated into the host DNA. Therefore, the coding sequence of the NMD reporter cassette was inserted in reverse orientation into the lentiviral vector. Since the SFFV U3 promoter, which usually drives the transcription of the transgene was located on the (+) strand it had to be replaced with a promoter sequence on the (-) strand. Additionally, the dual-luciferase NMD reporter system – introduced in chapter 3.3 – was inserted as this reporter system contains the firefly luciferase as an internal reference. Both, the firefly luciferase, as well as the renilla luciferase were expressed under the control of their own SV40 promoter sequence.

The renilla luciferase was again fused to exon 1, 2, and 3 of β-globin with the corresponding introns, containing either no PTC (WT) or a PTC within β-globin exon 2 (NS39) resulting in NMD-sensitivity of the transcribed mRNA. A schematic drawing of the generated lentiviral vector structure is shown in Figure 3.23, the full plasmid card is shown in the Appendix, Figure 6.3.
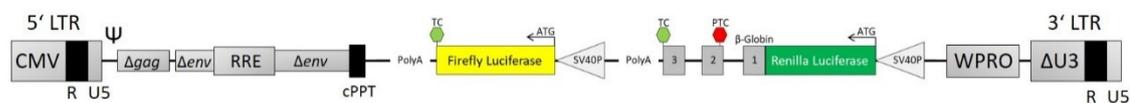

Figure 3.23: Schematic drawing of the newly designed lentiviral vector structure encoding for the dual-luciferase NMD reporter system with its coding sequence on the minus strand. The sequence between the cPPT and WPRO was deleted and replaced with the dual-luciferase NMD reporter system, either WT or NS39, each with its own SV40 promoter and polyadenylation site (cf. Figure 3.9). To prevent unwanted splicing events within the producer cell line the two GOIs were inserted in 3' to 5' direction. The other structural features were kept as in the TPI lentiviral vector (Figure 3.20), as adapted from (Röllecke 2016). The full plasmid card is shown in the supplemental Figure 6.3.

For the production of lentiviruses, HEK293T cells were transiently transfected with the designed lentiviral vectors, together with the two helper plasmids pCD/NL-BH and pczVSV-G. As the viral titer was still not measurable, HeLa T4$^+$ cells were test-wise incubated with different volumes of lentiviral supernatant, either containing the RL-WT, or RL-NS39 isoform, for 48 hours and analyzed for their luciferase activity. As control served HeLa T4$^+$ cells which were transiently transfected with the WT or NS39 dual-luciferase reporter plasmid. The measured luciferase expression in relative light units [RLU] is shown in Figure 3.24.
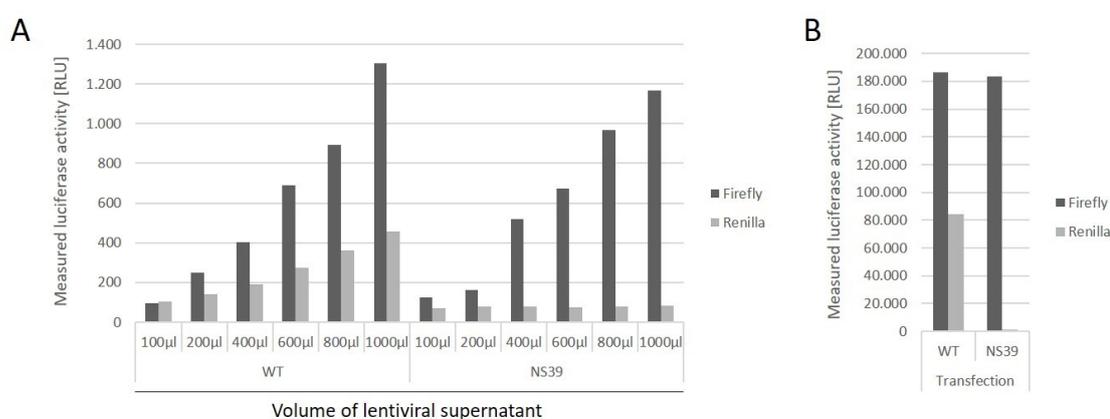
Figure 3.24: The measured luciferase activity of transduced cells resembled the transfection control. (A) HeLa T4$^+$ cells were incubated with the depicted volume of lentiviral supernatant (either WT or NS39) for 48 hours. **(B)** As a control, HeLa T4$^+$ cells were either transfected with 0.5 µg of the WT or NS39 dual-luciferase plasmid. The cells were harvested and lysed with one-step lysis juice and analyzed for the activity of the firefly luciferase, as well for the renilla luciferase activity. Based on the massive differences in the RLU levels, the transfection experiments are shown in a different diagram (B).

The RLUs for both firefly and renilla luciferase upon transduction are shown in Figure 3.24 A, individually for each volume of the two used lentiviral stocks (WT or NS39). As expected, the levels of firefly luciferase activity (dark grey) increased with the used volume of lentivirus for both, the WT and the NS39 virions. However, the same increase could be observed for the renilla

luciferase activity (shown in light grey), but only for cells that were transduced with the WT virus. The renilla luciferase activity of HeLa T4+ cells which were transduced with either volume of the NS39 lentivirus showed no increase. That minimal but stable detection of renilla luciferase activity in the NS39 samples indicated the degradation of the PTC containing renilla-β-globin mRNAs via the NMD pathway, resulting in the absence of protein translation and hence, only minimal measurable luciferase activity. However, the comparison of the luciferase activity levels upon lentiviral transduction (Figure 3.24 A) to the transfection experiment (Figure 3.24 B) showed clearly that the overall luciferase activity of both, firefly and renilla luciferase was a lot less upon lentiviral gene transfer.

Even though the measured luciferase activity upon lentiviral gene transfer was decreased in comparison to the transfection-based luciferase activity, the dual-luciferase lentiviruses were used to investigate the influence of HIV-1 infection on the cellular NMD activity. Therefore, HeLa T4+ cells were transduced with either the WT or the NS39 lentiviral supernatant for 48 hours, before they were infected with HIV-1 (MOI 0.5). The cells were harvested 6 hours, 24 hours, 48 hours, and 72 hours later and analyzed for their luciferase activity. The results are shown in Figure 3.25.
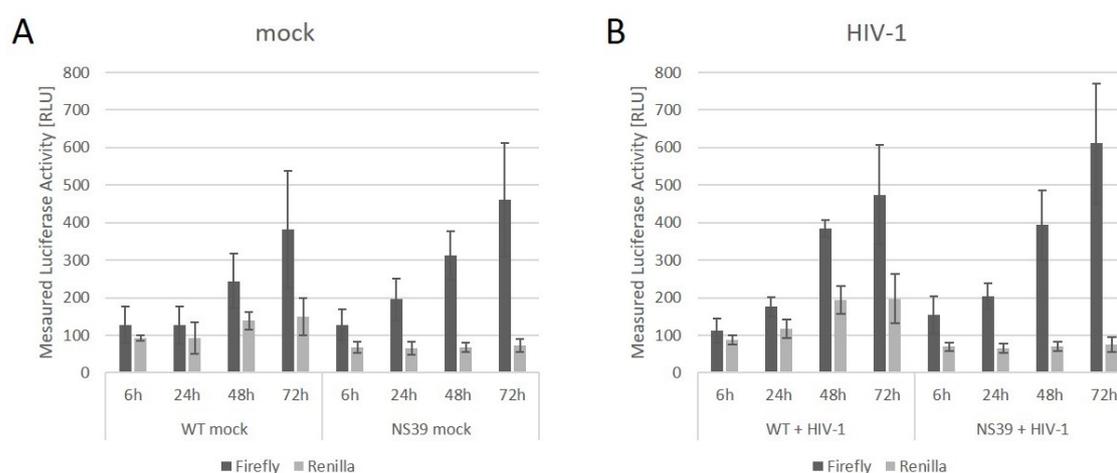
Figure 3.25: The measured luciferase activity between untreated and HIV-1 infected cells show only minor differences. HeLa T4+ cells were transduced with 1 ml of the lentiviral supernatant containing either the WT or the NS39 reporter variant for 48 hours. **(A)** The cells were not infected (mock), or **(B)** the cells were infected with HIV-1 (MOI 0.5) and incubated further for either 6h, 24h, 48h, or 72h. As all samples were collected, they were analyzed for their firefly and renilla luciferase activity.

The measured luciferase activities are shown in Figure 3.25 A regarding the not infected (mock) cells, and in Figure 3.25 B upon infection with HIV-1 (MOI 0.5). The overall luciferase activity pattern did not change strikingly upon HIV-1 infection. Noticeable was the increase in the overall luciferase activity upon HIV-1 infection, regarding both, the firefly and renilla luciferase activity in the WT samples (cf. Figure 3.25 A and B, WT). For the PTC-containing NS39 samples, that

increase of firefly activity was observed again, the renilla luciferase activity, however, was not altered, but remained constant. As the firefly luciferase activity increased, and the renilla luciferase activity remained on the same level, that could indicate an infection-based increase of NMD activity upon HIV-1 infection.

To test whether these steady-state levels of renilla luciferase activity were indeed due to efficient degradation of the PTC-containing NS39 β-globin mRNA, or if these measurements were an unspecific side effect produced by the firefly luciferase reacting with the renilla substrate the co-reactivity of the two luciferases with the opposite substrate was tested. Therefore, the dual-luciferase reporter plasmids were modified to encode only for either one of the two luciferases. First, the firefly luciferase was removed, generating a plasmid which was solely encoding for the renilla luciferase, fused to β-globin exon 1, 2, and 3 (ΔFL). Second, another plasmid was generated by the removal of the renilla luciferase ORF and β-globin exon 1, 2, and 3, which then only encoded for the firefly luciferase (ΔRL). A schematic sketch of these new single-luciferase plasmids is shown in Figure 3.26 A. HeLa T4$^+$ cells were transiently transfected with one of the single-luciferase plasmids for 24 hours and analyzed for their luciferase activity regarding both substrates, the firefly luciferase substrate, as well as the renilla luciferase substrate. The measured luciferase activity in relative light units is shown on a logarithmic scale in Figure 3.26 B.
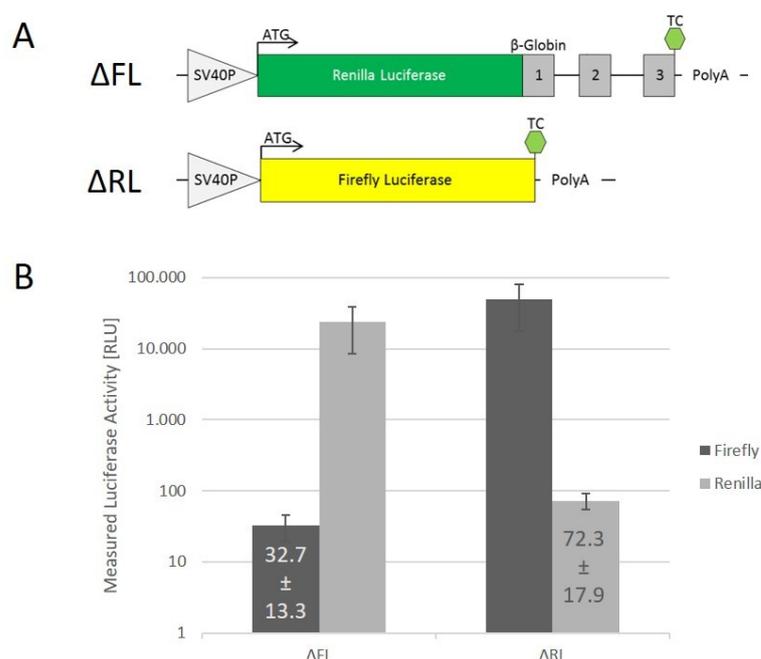
Figure 3.26: Determination of the minimum detection threshold of Firefly luciferase and Renilla luciferase. (A) The dual-luciferase plasmids were modified and either one of the two luciferase ORFs was removed, generating two single-luciferase plasmids. Removal of the firefly luciferase ORF (ΔFL) generated a plasmid which only encoded for renilla luciferase, fused to β-globin exon 1, 2, and 3 with corresponding introns. Removal of the renilla-β-globin ORF (ΔRL) generated a plasmid which only encoded for the firefly luciferase. **(B)** HeLa T4$^+$ cells were transiently transfected

with the constructs depicted in (A) for 24 hours. The cells were harvested and lysed with one-step lysis juice and analyzed for their luciferase activity.

The measured luciferase activities of the transfected luciferase reporter plasmids are shown on a logarithmic scale, as the encoded luciferase with its corresponding substrate showed a very high luciferase activity (cf. Figure 3.26 B ΔFL, Renilla, and ΔRL, Firefly) and the not-corresponding substrate showed much lower, but still measurable luciferase activities (cf. Figure 3.26 B ΔFL, Firefly, and ΔRL, Renilla). From this experiment, the minimal luciferase threshold was calculated based on the transfected luciferase and the measured luciferase activity with the not-corresponding substrate. That threshold was calculated from the mean value of measured luciferase activity plus the obtained standard deviation. For firefly luciferase, that threshold was 46 RLU and for renilla luciferase 90 RLU.

Now, coming back to the experiments with the lentiviral vectors which encoded for the dual-luciferase reporters (cf. Figure 3.25) the threshold for the firefly luciferase activity was exceeded in all samples and time points. The minimal measured firefly luciferase activity was 112 RLU and found in the 6h WT + HIV-1 sample (Figure 3.25 B, WT + HIV-1 6h). Hence, the obtained firefly values were above the threshold levels.

That was different regarding the measured renilla luciferase activity. Here, the calculated threshold was 90 RLU (cf. Figure 3.26) but only the measured renilla activities of the renilla luciferase β-globin WT exceeded this threshold value. The measured renilla activities of the renilla luciferase β-globin NS39 variant were all below the threshold value, with the highest measured renilla activity of 75.06 RLU 72h post-HIV-1 infection (Figure 3.25 B, NS39 + HIV-1 72h). Hence, the measured luciferase activities of the renilla luciferase were all under the detection threshold level and could therefore not be used for the calculation of the NMD activity. As promising as this lentiviral vector system looked, upon calculation of the threshold value, even the measured renilla luciferase activity of the NS39 variant of the test samples which were treated with 1000µl of lentiviral supernatant (cf. Figure 3.24 A) only reached 82 RLU and therefore remained below the determined threshold value. Hence, the lentiviral vector system was not suited for the determination of the infection-based modulation of the NMD machinery, and the transfection-based readout had to be used further for the reporter-based analysis.

3.7 The cellular NMD activity can be modulated by recombinant Interferons

So far, this work has shown that the HIV-1 2kb RNAs, especially the *tat* mRNAs are susceptible to degradation through the nonsense-mediated decay pathway. Even though the infection-based modulation of the cellular NMD machinery, which was described by Dr. Nora Diehl (Diehl 2016), could not be reproduced, the potential role of infection-based innate immunity activation and the interplay with the NMD activity was to be investigated.

To investigate the innate immune response on the activity of the nonsense-mediated decay pathway, HeLa cells were stimulated with recombinant interferon (α2, β, or γ) at a working concentration of 500 U/ml and analyzed for their NMD activity at different time points. As readout for the NMD activity the HeLa cells were transiently transfected with the dual-luciferase reporter system, already introduced in chapter 3.3, before IFN treatment. As control that the observed effects were indeed IFN-induced, the Jak/Stat inhibitor Ruxolitinib (Ruxo) was added to the cell media before the IFN addition. The Ruxolitinib stock solution was kindly provided by Prof. Dr. Mirko Trilling and used in the recommended working concentration of 4 µM (Le-Trilling et al. 2018). The measured luciferase activities of the transiently transfected reporters were used as readout for the NMD activity. First, the measured activity of the renilla luciferase was normalized to the measured firefly luciferase activity of the same sample. Second, the obtained ratio of the PTC containing reporter was normalized to the WT reporter ratio for each treatment condition. The thereby calculated relative NMD activity was plotted against the different time points and the results are shown in Figure 3.27.
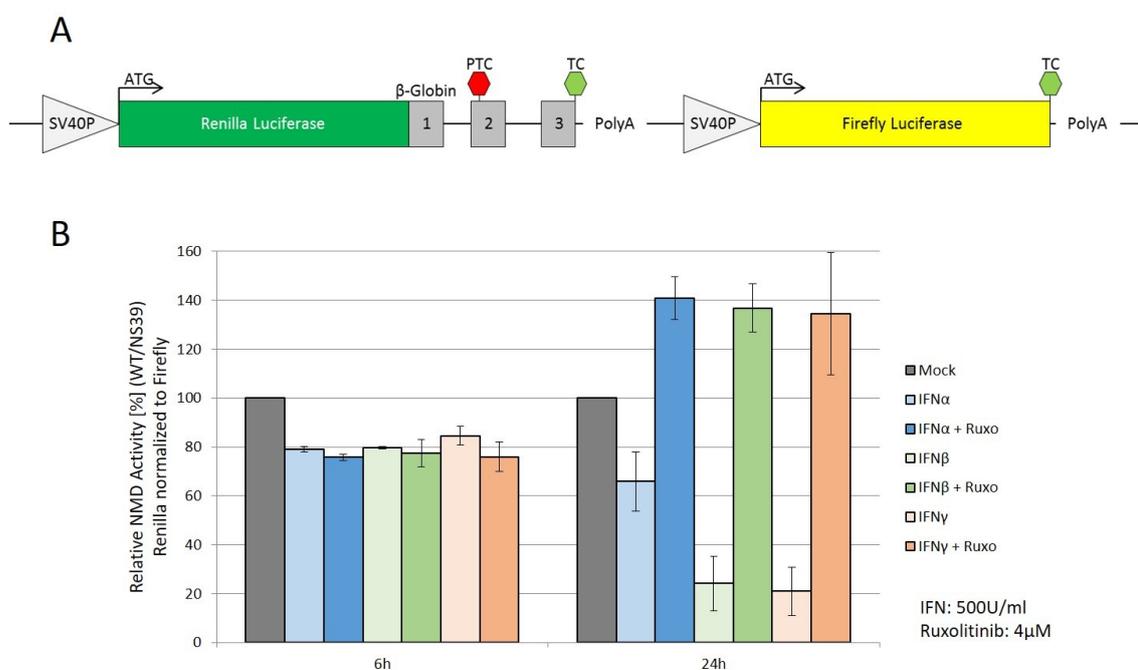
Figure 3.27: The cellular NMD activity is reduced by the addition of recombinant interferons. (A) schematic drawing of the used dual-luciferase reporter constructs, as described in Figure 3.9. (B) 3x10^5 HeLa cells were transfected with the depicted reporter constructs. After 24h the media was changed, and the designated interferon was added to its corresponding well. Ruxolitinib treatment served as control and was added to the media 10 minutes before the IFNs. The cells were harvested either 6h or 24h post-IFN treatment and the luciferase activity was measured. The activity of the renilla luciferase was normalized to the corresponding firefly luciferase. The relative NMD activity was calculated by the division of WT through NS39.

The results in Figure 3.27 showed that each of the three used interferons suppressed the cellular NMD activity, as shown by the decreased WT/NS39 ratios. The treatment with recombinant IFNα2, however, had the smallest impact on the cellular NMD activity, in comparison with IFNβ

and IFNγ. The effects of IFN treatment seemed to be specific, as the prior incubation with Ruxolitinib inhibited that effect. Hence, this experiment indicated an influence of IFN treatment and the subsequent signaling cascade on the regulation of the nonsense-mediated decay.

To gain a better understanding of our recombinant interferons and the subsequent modulations of gene expression, the IFNs were tested based on their functionality in cooperation with Dr. Marek Widera. For the experimental setup, our recombinant IFNs (α2, β or γ; 500 U/ml each) were compared with two interferon α subtypes (IFNα2 and IFNα14, working conditions 10,000 pg/ml) from Dr. Wideras lab. To analyze the IFN-based modulation of gene expression, Jurkat and PM1 T-cell lines were incubated with the depicted interferons for either 6 hours, or 24 hours, with or without prior Ruxolitinib treatment. The RNA was harvested, reverse transcribed, and analyzed via quantitative PCR for the expression of the interferon-stimulated genes (ISGs) ISG15 and IRF-1. Since this experimental setup was independent of transfections, the two T-cell lines Jurkat and PM1 were used. These immortalized T cell lines resemble the natural host cells of HIV-1 more than the previously used HeLa cells but are unfortunately extremely difficult to transfect. Hence, as this experiment did not require transfection, the influence of our recombinant interferons on these suspension T-cell lines was tested.
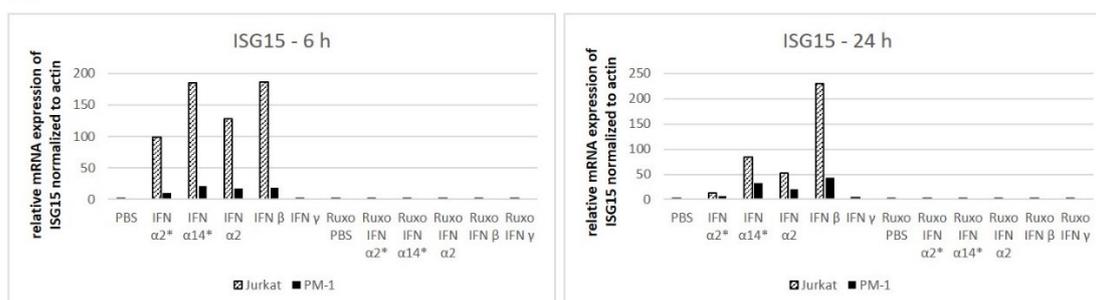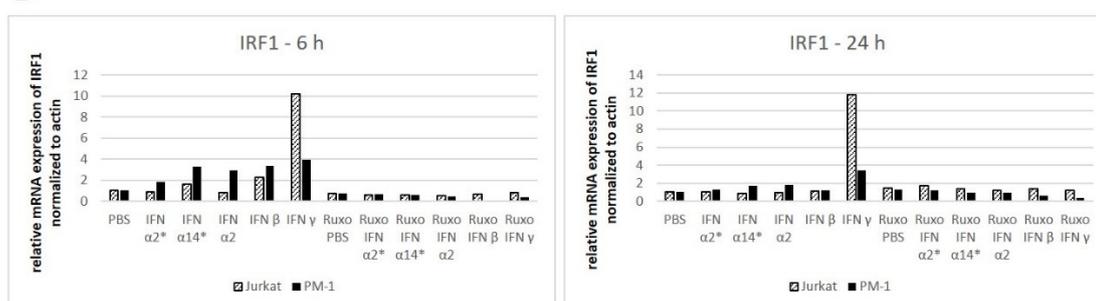
Figure 3.28: Specific induction of ISG15 by type I and IRF-1 by type II interferons. 10,000 cells (Jurkat or PM1) were seeded into 48-well plates and incubated with interferons (IFNα2*, IFNα14* 10,000 pg/ml, AG Widera; IFNα2, IFNβ, IFNγ 500 U/ml, AG Schaal) either with or without prior treatment with 4 µM Ruxolitinib. The total RNA was harvested either 6h or 24h post-IFN treatment, incubated with DNase I, reverse transcribed into cDNA, and used for quantitative PCR analysis. **(A)** ISG15 qPCRs (#6054/#6055), **(B)** IRF-1 qPCRs (#6058/#6059). The relative mRNA expression was calculated by normalizing to actin (#6056/#6057).

The analyzed ISGs were chosen based on their inducibility by either type I (ISG15), or type II (IRF-1) interferons. The results of IFN treatment on the expression of the two tested ISGs showed, that the intensity of ISG induction was cell type-dependent. The overall interferon-induced gene expression of Jurkat cells was much higher than for PM1 cells (cf. Figure 3.28). Furthermore, depending on the used IFN, the induction of the analyzed ISG showed differences in the temporal expression. While the ISG15 induction after IFNα treatment was reduced after 24h, in comparison with the 6-hour time point, the ISG15 expression increased even further by IFNβ stimulation from 6h to 24h (cf. Figure 3.28 A). The interferon-induced expression of IRF-1 in Jurkat cells, however, was only enhanced upon IFNγ treatment. The IRF-1-induction after 6h of IFNγ treatment further increased until 24h post-treatment (cf. Figure 3.28 B). In general, the comparison of the two cell types showed that the PM1 cells had overall weaker inducibility by IFNγ, as already observed for the type I interferons (cf. Figure 3.28 A). Furthermore, the specificity of IRF-1 expression in PM1 cells was reduced, shown by the enhanced gene expression of IRF-1 by stimulation with IFNα2, α14, or IFNβ. The addition of Ruxolitinib showed an efficient inhibition of the interferon signaling pathway in both cell types and time points, as the IFN stimulation of cells that were pre-incubated with Ruxolitinib, resulted in almost no detectable ISG expression.

Since the effect of recombinant interferon on transiently transfected HeLa cells had such a great impact on the cellular NMD activity (compare Figure 3.27), and the presence of interferons can be monitored by the expression of ISGs (compare Figure 3.28), the next step was to analyze the role of endogenous interferons upon HIV-1 infection. To stick as close to the natural target cells of HIV-1 and keeping the cell type-dependent differences in mind, the two T-cell lines (Jurkat and PM1) were used again, now for HIV-1 infection experiments. As interferon-induction is a very fast cellular response, the first RNA samples were collected already six hours post-infection, to monitor potential endogenous IFN production and the subsequent expression of ISGs.

Jurkat and PM1 cells were infected with HIV-1 (MOI 0.5) and the total RNA was harvested either 6, 24, 48, or 72 hours post-infection. The RNAs were isolated, reverse transcribed, and analyzed by RT-(q)-PCR for their potential ISG expression (ISG15 and IRF-1), as well as for the expression of endogenous splice isoforms which encode NMD targets (SRSF3 and hnRNPL).
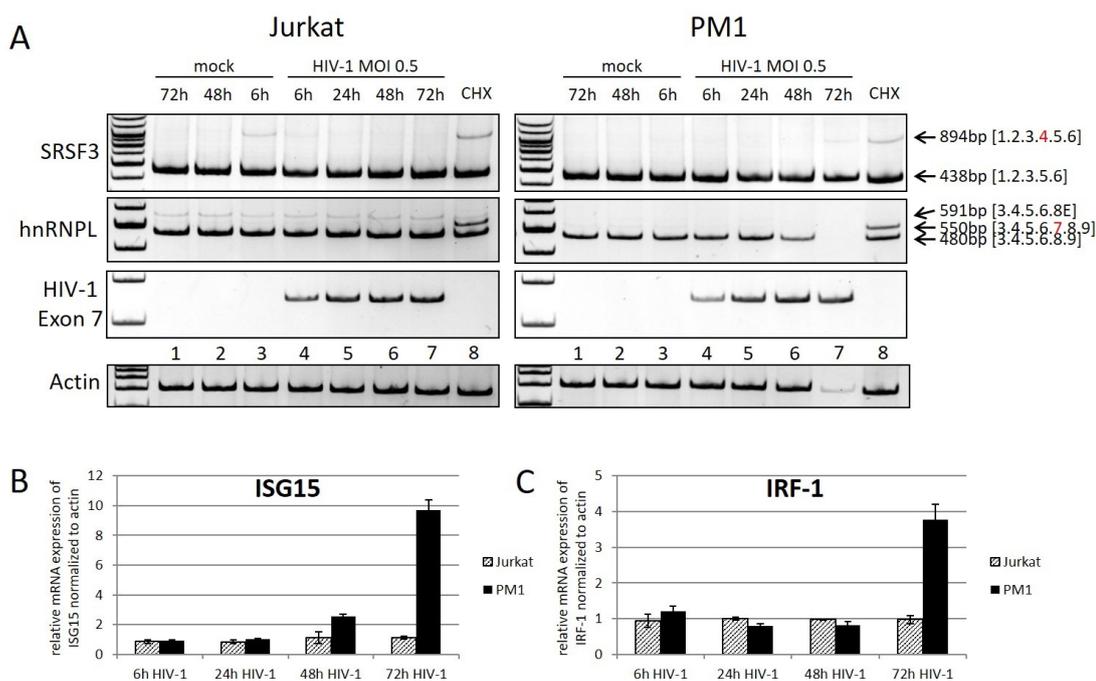
Figure 3.29: Infection with HIV-1 results neither in NMD inhibition nor IFN response in Jurkat and PM1 cells. 5×10^5 Jurkat or PM1 cells were infected with HIV-1 (MOI 0.5). The total RNA was harvested 6h, 24h, 48h, or 72h post-infection. For the cycloheximide control, 50 µg/ml CHX were added to one not-infected (mock) flask of each cell type, 6 hours before harvesting. **(A)** RT-PCR analysis of Jurkat and PM1 cells. The RNA was reverse transcribed and amplified with specific primer pairs. SRSF3 (#4004/#4005, 35 cycles (Diehl 2016)), hnRNPL (#5778/#5779, 26 cycles), HIV-1 exon 7 (#3387/#3388, 26 cycles), actin (#6056/#6057, 26 cycles). The PCR fragments were separated on a 10% PAA gel and visualized by EtBr staining and UV-light exposure. **(B, C)** RT-q-PCR results of the same cDNA as used for (A), using specific primer pairs: ISG15 (#6054/#6055), IRF-1 (#6058/#6059), normalized to actin (#6056/#6057).

Based on the observation that recombinant interferons can suppress the cellular NMD pathway (cf. Figure 3.27) and that IFN treatment can be monitored by the induction of ISGs (cf. Figure 3.28) the HIV-1 infection-based IFN response, as well as the potential influence on NMD, was investigated in this experimental setup.

As described before (cf. Figure 3.5 and Figure 3.6), the splice isoforms of the two endogenous genes SRSF3 and hnRNPL transcripts were checked via RT-PCR to monitor the potential effect of the HIV-1 infection on the NMD-sensitive isoforms. A six-hour CHX treatment of uninfected cells again served as a control to monitor the PCR amplification of the NMD specific isoform (Figure 3.29 A, lane 8). The amplification of HIV-1 exon 7 served as infection marker (cf. Figure 3.29 A, lanes 4-7) and suggested a productive infection, as the amounts of viral RNA increased over time. The abundance of the amplified NMD-sensitive splice isoforms, however, was not altered upon HIV-1 infection in both cell lines and both tested genes (cf. Jurkat and PM1, SRSF3 and hnRNPL). The NMD-sensitive splice isoforms were only clearly detectable within the CHX-control lane.

Based on the observation of IFN-induced NMD inhibition, the missing upregulation of the NMD-sensitive splice isoforms of SRSF3 and hnRNPL upon HIV-1 infection led to the assumption that there was no enhanced expression of IFNs and hence, ISGs. Indeed, there was no remarkable induction of the interferon-stimulated genes ISG15 and IRF-1 (cf. Figure 3.29 B and C). The only noteworthy upregulation of both ISGs was observed in PM1 cells 72h post-infection, but that was most likely an artifact, based on the decreased actin levels of these cells, suggesting apoptosis due to the infection. Additionally, based on the prior IFN-induced ISG-induction (cf. Figure 3.28), the induction of ISGs upon HIV-1 infection was expected earlier than 72h post-infection. Furthermore, based on the observed cell-type dependency, a stronger ISG induction of Jurkat cells was expected in respect to the PM1 cells.

Inferentially it is to mention, that recombinant interferons at a concentration of 500 U/ml efficiently induced ISGs within our tested cell culture cells, and interfered with the activity of the cellular nonsense-mediated decay. In a more natural setup, however, we could not detect induction of ISG expression or influence on the NMD activity upon HIV-1 infection in Jurkat or PM1 T-cell lines.

3.8 Quantitative PCR analysis indicate an influence of high viral replication on some endogenous NMD targets

One last approach was used to quantitatively investigate the influence of HIV-1 infection on the cellular NMD machinery regarding endogenous genes, without the need for transfection or transduction of reporter genes. The initial working hypothesis suggested an infection-based inhibition of the NMD machinery. Within this work, however, the NMD susceptibility of the HIV-1 2kb class mRNAs was described and investigated, but a clear influence of the HIV-1 infection on the global NMD machinery could not be shown. So far, the analysis of NMD reporters, as well as endogenous NMD targets were performed using RT-PCRs, which focused on the potential up-regulation of NMD targets, and the subsequent UV-visualization of PCR products. To gain a higher specificity, a quantitative PCR-based system was established to analyze the impact of HIV-1 infection on endogenous genes. This approach allowed the normalization of each transcript isoform, not only to a house-keeping gene, but also the normalization of the NMD-target splice isoform to its not-NMD-target counterpart in relation to the HIV-1 transcript levels.

Two genes were selected for this experimental validation. The selection of these genes was based on the data of the PM1 deep sequencing experiments, which were generated by Dr. Nora Diehl (Diehl 2016), and re-analyzed by Johannes Ptok (dissertation in preparation). Based on that analysis, the selected candidate genes were the two splicing regulatory proteins SRSF7 and hnRNPDL, as they both encoded for at least one NMD-sensitive isoform and were both highly expressed in PM1 cells. So far, the detection of endogenous NMD splice isoforms was performed

using one primer pair which detected both, the NMD-sensitive, as well as the not-NMD-sensitive isoform. Now, for the quantitative detection of NMD-specific, as well as not-NMD-specific splice isoforms, specific primer pairs were needed which amplified only one isoform, either NMD-specific or not-NMD-specific. Therefore, specific exon junction primers were designed, which allowed to individually detect the abundance of each splice isoform. Schematic sketches of the exon-intron structures of SRSF7 and hnRNPDL, as well as the designed primer pairs, and primer binding sites, are shown in Figure 3.30. For both selected genes, one exon had to be excluded from the final transcript to produce the protein-coding isoform. For SRSF7, exon 4 had to be excluded and for hnRNPDL, exon 8 had to be excluded. The inclusion of these so-called "poison-cassette exons" into the final mRNA of both test genes, resulted in pre-mature translation termination more than 55 nucleotides upstream of the last exon-exon junction, hence in NMD-sensitivity and degradation. Each exon junction primer was designed to detect either, the NMD-sensitive, or the not-NMD-sensitive splice isoform. The primers were tested using two different cell types, HeLa and HEK293T cells with and without cycloheximide treatment. The results of the primer test are shown in Figure 3.30 beneath the exon-intron structure of the respective gene.
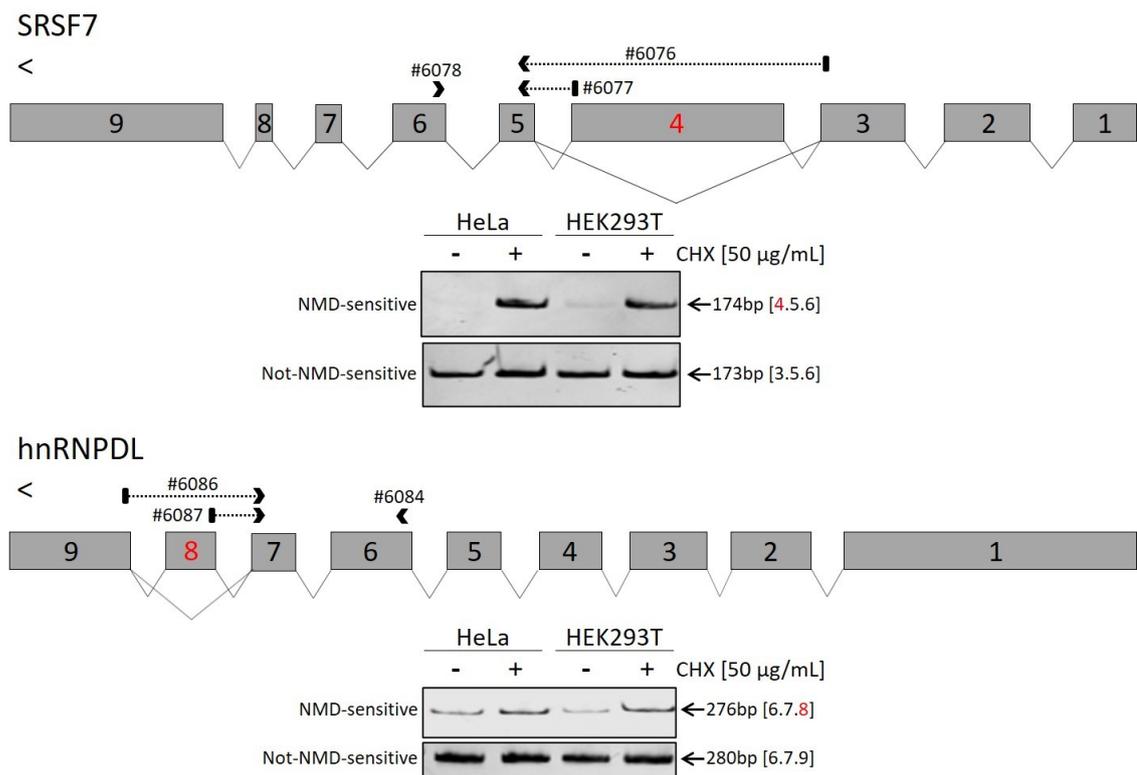
Figure 3.30: The cycloheximide-based primer test for the selected candidate genes SRSF7 and hnRNPDL amplified the predicted transcript isoforms. Schematic drawing of the exon-intron structures of SRSF7 and hnRNPDL with positions of the primer binding sites and the corresponding primer numbers. Exon junction primers and their binding sites are indicated by arrowheads, dots, and tails. The poison-cassette exons are marked by red letters. The corresponding RT-PCR results are shown beneath the sketches. The total RNA from HeLa and HEK293T cells was

harvested and reverse transcribed. As NMD control, one sample was incubated with 50 µg/ml CHX for 6 hours before harvesting. The cDNA was diluted 1:10 and used for RT-PCRs (28 cycles). The PCR products were separated on a 10% PAA gel, stained with ethidium bromide, and visualized by UV light. The experiments were performed by Lucienne Bogun during her master thesis.

The RT-PCR test of the designed primer pairs for the two tested genes SRSF7 and hnRNPDL showed reliable results regarding the primer efficiencies in HeLa and HEK293T cells with and without cycloheximide treatment (cf. Figure 3.30). The not-NMD-sensitive isoforms showed no accumulation upon CHX treatment in either cell type, while the NMD-sensitive splice isoforms accumulated upon CHX treatment, which was consistent with the co-translational degradation of NMD targets. The basal levels of the NMD-sensitive isoforms, however, differed not only between the cell types (compare Figure 3.30 lanes 1 and 3) but also between the tested genes. The basal level of the NMD-sensitive isoform of SRSF7 was much lower than the one of hnRNPDL. Furthermore, the observed accumulation upon CHX treatment was much higher for SRSF7, than for hnRNPDL, indicating a different sensitivity for NMD target recognition, which is not only dependent on the cell type but the tested gene. That difference in the NMD-sensitivity might be due to the number of splicing events, hence the number of deposited EJCs downstream of the PTC, as the NMD-sensitive splice isoform of SRSF7, had five downstream splicing events and hnRNPDL only one (cf. Figure 3.30).

Now, following the design and test of the primers, PM1 cells were infected with HIV-1 (MOI 0.5) and the total RNA was harvested 0h, 24h, 48h, or 72h post-infection. For each time point, one not-infected PM1 cell sample was harvested as mock and reference. Due to observed variations in the productivity of the HIV-1 infection between the biological replicates, the fold change ratios of NMD-sensitive transcripts to not-NMD-sensitive transcripts were plotted against the fold change of HIV-1 exon 7, rather than the time points post-infection (cf. Figure 3.31).
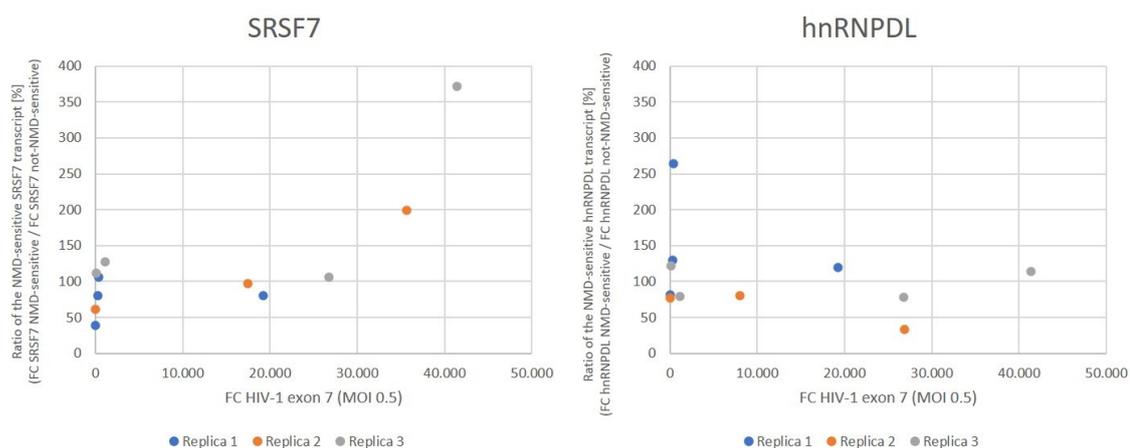
Figure 3.31: The ratio of NMD-sensitive SRSF7 transcripts, but not hnRNPDL transcripts increased upon productive HIV-1 infection in PM1 cells. The fold change (FC) ratios of NMD-sensitive transcripts to the not-NMD-sensitive transcript are plotted against the fold change of HIV-1 exon 7 for SRSF7 **(left)**, and hnRNPDL **(right)**. PM1 cells were infected with HIV-1 (MOI 0.5) and the total RNA was harvested at different time points (0h, 24h, 48h, 72h p.i.). The

fold change of HIV exon 7 (#3387/#3388) within each sample was used to detect the productivity of the infection. The primer pairs used for the detection of SRSF7 were #6076/#6078 (not-NMD-sensitive) and #6077/#6078 (NMD-sensitive). The primer pairs used for the detection of hnRNPDL were #6084/#6086 (not-NMD-sensitive) and #6084/#6087 (NMD-sensitive). The qPCRs were performed by Lucienne Bogun during her master thesis.

The results of the qPCR analysis of SRSF7 showed an increasing ratio of NMD-sensitive transcripts to not-NMD-sensitive transcripts with increasing amounts of HIV-1 transcripts (cf. Figure 3.31, left). That increase of NMD-sensitive SRSF7 transcripts suggested a down-regulation of the NMD-mediated target degradation during HIV-1 infection. The obtained qPCR results for hnRNPDL, however (cf. Figure 3.31, right) showed no impact of the HIV-1 infection regarding the ratio of NMD-sensitive transcripts to not-NMD-sensitive transcripts.

From the set of experiments, that were necessary to show an effect of HIV-1 infection on the regulation of NMD targets (cf. Figure 3.18, Figure 3.19, Figure 3.29, and Figure 3.31) this work concludes that HIV-1 infection, in general, has no direct impact on the global cellular nonsense-mediated decay pathway, hence the virus does not seem to directly interfere with the NMD machinery. Even though the viral 2kb class, especially the *tat* mRNAs showed sensitivity to NMD (cf. Figure 3.1, Figure 3.2, Figure 3.3, Figure 3.4, and Figure 3.7) that sensitivity was restrictive enough that interference of the virus with the NMD machinery was necessary. However, as the HIV-1 2kb class mRNAs were always detectable within the performed experiments, the level of transcribed RNAs is higher than the amounts of degraded transcripts. That effect resembled the NMD-sensitive isoforms of hnRNPDL, which was detectable even without CHX treatment (cf. Figure 3.30). Furthermore, some HIV-1 RNAs encode suboptimal translational start codons, which might also protect these RNAs from degradation by NMD (cf. Figure 3.10).

4 Discussion

This work aimed to characterize the sensitivity of HIV-1 transcripts to co-translational degradation through the nonsense-mediated decay pathway, as well as the infection-based modulations of the cellular NMD machinery. The results of this work showed that specifically HIV-1 transcripts of the 2kb class were targeted by NMD. In addition, this work illustrated the complexity of even simple-looking splice-dependent reporter systems, due to the splice-, translation and RNA stability-dependent influence of the read-out.

4.1 HIV-1 transcripts are targeted by NMD

The initial experiments indicated that the 2kb class of HIV-1 transcripts are targeted for degradation by the nonsense-mediated decay pathway, as they showed cycloheximide-induced accumulation in infection-, as well as in transfection experiments (cf. Figure 3.1 and Figure 3.2). That CHX-dependent accumulation was independent of the number of splicing events, hence the number of deposited EJCs (cf. Figure 3.2). Subsequent modifications of the 3' UTR of the fully spliced *tat* message, such as shortening of the 3' UTR by deletion of 516 nucleotides, or extension of the *tat* ORF to the *nef* TC, did not stabilize the viral transcript (cf. Figure 3.3). Furthermore, the insertion of the RRE sequence, as a structural element, into the 3' UTR of the *tat* message had no impact on the viral RNA stability. The subsequent co-transfection with viral *rev*, which allowed RRE-Rev interactions, further destabilized the transcript. Interestingly, that destabilization, which was shown by an increase in the CHX-induced accumulation, was, however, independent of the presence of the RRE within the *tat* transcript, and thus RRE-Rev interaction (cf. Figure 3.4).

Like all retroviruses, HIV-1 stably integrates its reverse-transcribed genome into the DNA of its host cell and uses the cellular transcription, splicing, and translation apparatus to express its proteins and assemble new virions. Due to their compact genome structure, viral mRNAs encode features like very long 3' UTRs or overlapping ORFs, which distinguishes them from cellular mRNAs and potentially marks them for degradation by NMD. Within this work, however, it was shown that not the HIV-1 mRNAs with the longest 3' UTRs (*gag*, *gag/pol*, *vif*, and *vpr*), nor the bicistronic *gag/pol* transcripts were targeted by NMD, but only transcripts of the shortest class, the viral 2kb class, especially the *tat* mRNAs showed a CHX-induced accumulation.

Retroviruses use alternative splicing to generate multiple mRNA species from a single pre-mRNA, and the order of co-transcriptional intron removal from HIV-1 genomic RNA is tightly regulated to ensure that no viral transcript harbors a 3' UTR EJC, which would trigger NMD (Bohne, Wodrich, and Kräusslich 2005). Therefore, it was postulated that retroviruses like HIV-1 indirectly escape NMD by controlled splicing (Popp, Cho, and Maquat 2020).

However, if the regulated splicing process protects the viral mRNAs from EJC-dependent NMD the viral transcripts still encode very long 3' UTRs, which might mark them as NMD targets and result in degradation. Within this work, the CHX-based translation inhibition showed that only the fully spliced 2kb class of HIV-1 mRNAs accumulated and that this NMD-sensitivity was not reduced by modifications of the 3' UTR, like shortening of the 3' UTR of the *tat* transcript. That observation, however, is in line with the observations of Balistreri and colleagues, who shortened the 3' UTR of Semliki Forest virus (SFV) to 62 nucleotides and still observed NMD-sensitivity of the SFV genomic RNA (Balistreri et al. 2014). In general, the sensitivity of HIV-1 transcripts to degradation through NMD was only slightly investigated so far. In pNL4-3 transfection experiments, Andrew Mouland's group reported that the depletion of UPF1 resulted in a massive reduction of HIV-1 RNAs, which could not be rescued by UPF1 overexpression (Ajamian et al. 2008). They interpreted that UPF1 has a major impact on HIV-1 RNA stability and enhanced the viral RNA expression, a role that was independent of the function of UPF1 in NMD. However, when they expressed a Rev⁻ variant of the used provirus, they observed a decrease of 9kb RNAs and a slight increase of 2kb RNAs upon UPF1 knockdown. Even though they interpreted that result as the decay of the 9kb class upon UPF1 knockdown, it could also suggest the NMD-sensitivity of the 2kb class of HIV-1 mRNAs, as they were completely spliced and exported to the cytoplasm for translation and the 9kb RNAs were kept in the nucleus. In general, the observation of a decreased 9kb and an increase in 2kb RNAs could also be interpreted differently, as NMD-sensitivity of the 2kb class upon UPF1 knockdown, like observed in this work for the translation inhibition through cycloheximide.

Already in 2008 Ajamian and colleagues reported, that UPF1 particles were found within HIV-1 virus particles and the subsequent research of Heinrich Göttlinger's group showed, that the uptake of UPF1 particles into HIV-1 virions was dependent on the NC domain of HIV-1 Gag and that this UPF1 uptake is critical for the infectivity of the HIV-1 progeny viruses (Serquina et al. 2013). Even though they did not investigate the levels of viral RNA in infected cells upon UPF1 knock-down, they used HI-viruses, which were produced in UPF1 depleted cells. They observed a decreased infectivity of these virions, even though the UPF1 knockdown did not affect the overall virus production. Furthermore, the observations that UPF1 co-purifies with a model HIV-1 3' UTR (Hogg and Goff 2010) and that the beneficial effect of UPF1 on HIV-1 progeny virus infectivity was independent of the interaction of UPF1 with UPF2 (Ajamian et al. 2015) strongly suggest a role of UPF1 on HIV-1 replication, which is independent of the role of UPF1 in NMD.

The results of this work, however, suggest that UPF1 and its role in NMD also have an impact on HIV-1 transcripts, even though only on the 2kb class mRNAs. As the HIV-1 2kb class mRNAs, in comparison with the 4kb and 9kb class transcripts, have the shortest 3' UTRs, the described beneficial UPF1 HIV-1 interactions might only function for the intron-containing 4kb class and the full-length 9kb class. Differences between the mRNAs of the 2kb class to the mRNAs of the

4kb and 9kb class are not only the overall number of splicing events, which was shown to not have an impact on the CHX-based accumulation (cf. Figure 3.2) but also the way of nuclear export to the cytoplasm. While the fully spliced 2kb class mRNAs are exported via the canonical NXF1/Tap-mediated cellular mRNA export way, the intron-containing mRNAs of the 4kb and 9kb class are bound by the Rev protein at the Rev responsive element (RRE), which mediates the export of these transcripts via the CRM1 export pathway (Malim et al. 1989; Malim and Cullen 1993; Fischer et al. 1995). It was shown that UPF1 interaction helps the nuclear export of viral RNAs, via the CRM1 and not the NXF1 export pathway, by the formation of a UPF1-Rev-CRM1-DDX3 complex (Ajamian et al. 2015). This complex formation abolished the interaction of UPF1 with UPF2 and might therefore mediate functions of UPF1 which are isolated from the role of UPF1 in NMD target recognition. That difference between the export mechanisms of the HIV-1 mRNA classes and the NMD-independent interactions of the 4kb and 9kb transcripts with UPF1 might explain the targeting of the viral 2kb class mRNAs via the NMD pathway, even though the intron-containing 4kb and 9kb transcripts would be the more favorable target.

4.2 Chemical NMD inhibition through CHX provides comprehensible results regarding endogenous targets

The reliability of the NMD target detection by transcript accumulation upon translation inhibition through cycloheximide was analyzed. The amplification of five different endogenous genes which encoded alternative splice isoforms showed specific accumulations of NMD target transcripts upon CHX treatment in all three tested cell types (cf. Figure 3.5). The induced effects of NMD target accumulation by inhibition of the translation apparatus were confirmed in the next step using specific inhibition of the NMD machinery by siRNA-mediated knockdown of UPF1, which mimicked the CHX-dependent accumulation effect of NMD target transcripts (cf. Figure 3.6). Furthermore, the UPF1 knock-down not only resembled the CHX-induced transcript accumulation in RT-PCR experiments, but it also demonstrated the accumulation of the HIV-1 2kb *tat* message in northern blot experiments, confirming that these transcripts are indeed targeted by NMD for degradation (cf. Figure 3.7). The approach to stabilize the HIV-1 *tat* message by the addition of the well-studied RSE element was partially successful, shown by a decreased transcript accumulation. This effect was independent of the inserted position of the RSE sequence (cf. Figure 3.7). Based on the reliable results between CHX-mediated translation inhibition and the more specific NMD inhibition by UPF1 knockdown, the investigation of translation-dependency of NMD targets, using a pNLA1 derivate with mutated internal *tat* and *rev* translational start codons, surprisingly showed a prominent accumulation of the HIV-1 2kb class, even though only the *nef* message should be translated (cf. Figure 3.8). Furthermore, upon Rev co-transfection, the overall HIV-1 RNA levels decreased but here, the CHX-dependent accumulation was not only visible for the 2kb RNAs, but also for the 4kb class mRNAs.

Even though we detected a smear in the accumulating bands upon CHX treatment, indicating degradation of these transcripts, the cycloheximide-dependent NMD target detection was very efficient. The observed smear could result from degradation of the transcripts poly-A tails, as it was shown that cycloheximide treatment results in the accumulation of deadenylated mRNAs (Beelman and Parker 1994). Nonetheless, diverse research groups successfully used the translational inhibitor CHX to analyze the NMD-sensitivity of diverse transcripts and reporter systems. Almost 30 years ago, it was reported that incubation with CHX dramatically induced mRNAs from PTC-containing constructs, while it had little or no effect on the mRNA levels of the WT in-frame construct (Carter et al. 1995). Already a few years later, Ishigaki and colleagues analyzed the levels of nonsense-containing CBP80-bound β-globin mRNA upon CHX treatment and reported an increase in the RNA abundance from 15% to 42% (Ishigaki et al. 2001). Linde and colleagues investigated the role of NMD in genetic diseases by examining the levels of cystic fibrosis transmembrane conductance regulator (CFTR) nonsense transcripts by treatment with CHX and observed that the inhibition of NMD by CHX is specific to transcripts that harbor a PTC (Linde et al. 2007). However, they not only analyzed the impact of cycloheximide treatment but also used siRNAs to specifically reduce the cellular UPF1 levels. They reported a higher fold increase of PTC transcripts upon CHX treatment, then in comparison to UPF1 downregulation, which they claimed to be probably due to the incomplete NMD inhibition by siRNA-mediated UPF1 knockdown. The results of Linde et al., however, fit the results of this work, where the siRNA-mediated UPF1 knockdown mimicked the CHX-dependent transcript accumulation, albeit to a smaller extent. That might be due to basal activity of the NMD pathway, possible through UPF1 knockdown and not UPF1 knockout, or by CHX-based artifacts, as CHX not only inhibits NMD but the global translation apparatus. A complete knockout of UPF1 is not possible, as UPF1 is essential for cell maintenance and survival in animals. For example, in cultured Drosophila SL2 cells, deletion of UPF1 resulted in cell cycle arrest (Rehwinkel et al. 2005), while zebrafishes and mice showed early embryonic lethality upon UPF1 knock out (Wittkopp et al. 2009; Medghalchi et al. 2001). For cultured human cell lines, it was reported, that UPF1 is required for S-phase progression (Azzalin and Lingner 2006).

Even though the NMD target determination through translation inhibition is not without side effects, the siRNA-mediated knockdown of UPF1 used in this work, allowed a comparable detection of NMD targets as observed by CHX treatment, even though to a smaller extent. However, another difference between the UPF1 knockdown and CHX treatment was observed regarding the splice site usage of the analyzed test gene hnRNPL (cf. Figure 3.6), suggesting a potential influence of UPF1 on alternative splicing, as the usage of hnRNPL SD8 was suppressed. That UPF1 might play a role in splicing and therefore might influence alternative splicing on a genome-wide scale was recently suggested by research in Arabidopsis (Raxwal et al. 2020). Using UPF1-null plants Raxwal and colleagues showed that the abundancies of different splice variants

switched upon UPF1 inactivation, which they explained with the deregulation of splicing factors and proteins involved in RNA processing. But they also proposed a more direct influence of UPF1 on splicing, as studies in human and Drosophila cells showed a co-transcriptional association with RNA in the nucleus and might therefore interfere with splicing factors on the transcribed RNA (Raxwal et al. 2020).

Based on the usage of translational inhibitors, like cycloheximide, and the introduction of stem-loops in the 5' UTR, the nonsense-mediated decay was proposed as a translation-dependent degradation pathway (Carter et al. 1995; Thermann et al. 1998). As these results were in line with the observed NMD target accumulations upon CHX treatment in this work, it was tested whether the observed accumulation of the HIV-1 2kb class mRNAs can be reduced by the deletion of the translational start codons of *tat* and *rev* in the context of the HIV-1 derivate pNLA1 (cf. Figure 3.8). Interestingly, the accumulation of the 2kb class was not affected by the deletion of the *tat* and *rev* ATGs, which suggested on ongoing degradation of these transcripts, possibly due to an ongoing translation from internal translational start codons, which then, of course, did not lead to the authentic Tat and Rev proteins. However, the co-transfection of pcRev resulted in decreased RNA levels in the pNLA1 context without CHX treatment, and in a surprising accumulation of both, the HIV-1 2kb and 4kb class upon translation inhibition. That newly observed 4kb accumulation might suggest a previously undetected NMD-sensitivity of this viral RNA class, or an artifact which resulted from the decreased RNA levels upon rev co-transfection, which were counteracted by translation inhibition through CHX. Already some years ago an enhancement of translation by Rev was proposed by different groups (Lawrence et al. 1991; D'Agostino et al. 1992; Perales, Carrasco, and Gonzalez 2005). Groom and colleagues specifically analyzed the influence of Rev on the translation of HIV-1 mRNAs in a cell-free assay, which allowed the separation of the nuclear transport function of Rev to its role in translation (Groom et al. 2009). For their analysis, they used RNAs that contained the HIV-1 5' UTR, which were reported to encode an additional highly conserved Rev binding site (Greatorex et al. 2002). They observed a significant inhibition of translation at the highest used Rev concentration, which was however reversed by lower Rev concentrations, which stimulated translation but were independent of the presence of the RRE (Groom et al. 2009). They concluded that low translation levels can destabilize RNAs and decrease the RNA levels, which might explain the lower 2kb and 4kb RNA levels upon pcRev co-transfection (cf. Figure 3.8). The reduced levels of 2kb class mRNAs could also be explained differently, as a decrease of the fully spliced HIV-1 RNAs was reported in the presence of Rev, even though they observed a rather positive effect of Rev on the RNA stability of unspliced viral RNAs (Felber et al. 1989). However, other studies do not report an effect of Rev on the 2kb transcript expression (Malim and Cullen 1993). Furthermore, Lata and colleagues reported that Rev causes the specific degradation of Tat, resulting in the

reduction of HIV-1 gene expression (Lata et al. 2015), which might explain the reduced mRNA levels upon Rev co-transfection but seems unlikely upon additional Tat co-transfection.

The probably most important function of Rev is the coordinated interaction with the RRE to mediate the export of intron-containing and unspliced viral transcripts from the nucleus to the cytoplasm via the Crm1 nuclear export receptor (Behrens et al. 2017). Taniguchi et al. showed that Rev reduces the levels of cellular and viral RNAs which are exported by Crm1-independent pathways by inhibition of the NXF1/TAP-specific export of RNAs (Taniguchi, Mabuchi, and Ohno 2014), which might also be a reason for the decrease in 2kb RNA levels upon Rev co-transfection.

However, off-target effects cannot be excluded within this experimental setup, as Rev normally is a low-abundant HIV-1 protein. Therefore, in experimental conditions, Rev is typically overexpressed what might lead to artifacts (Truman et al. 2020). For example, in non-dividing cells, it was shown that the accumulation of Rev resulted in cell death suggesting that high concentrations of Rev are toxic as they saturate the Crm1 receptor (Levin et al. 2010), therefore an off-target effect, resulting in the observed accumulation of the HIV-1 4kb class upon Rev co-transfection and CHX treatment cannot be excluded.

4.3 The translational link to RNA stability

As the CHX-dependent accumulation of the HIV-1 2kb class was still observed with mutated *tat* and *rev* translational start sites (cf. Figure 3.8), the translation dependency of NMD target recognition and degradation was questioned. Therefore, a dual-luciferase NMD reporter system was established, which allowed the readout of translational activity via luciferase activity measurements, and the determination of RNA levels via RT-(q)PCR. In both cases, firefly luciferase was used as an internal reference and renilla luciferase fused to β-globin exon 1-3 as GOI. With this reporter system, it was shown that a single nucleotide substitution within the Kozak sequence was sufficient to decrease or abolish the protein activity of the renilla luciferase and influence the RNA stability of the renilla WT constructs, shown by decreased RNA levels (cf. Figure 3.9). However, as the PTC-containing renilla RNAs were still degraded, there was no effect of the first ATG alterations on the NMD-sensitivity of the PTC-containing RNAs. That result suggested the usage of another translational start codon further downstream, which abolished the protein activity, but not the NMD-sensitivity (cf. Figure 3.9). Indeed, the subsequent mutation of all downstream located 46 potential translational start codons confirmed that another downstream ATG was used for translation initiation within the renilla β-globin RNA, as the translation did not result in a functional protein but was sufficient to induce degradation of the PTC containing RNA (cf. Figure 3.10). As the inserted mutations were not always silent, the renilla protein activity was not maintained and could not be measured. However, the RT-(q)PCR readout confirmed two things. First, the RNA levels of not-PTC-containing renilla β-globin constructs with weakened ATG (Pyr.-3) or without ATG (ΔATG) were decreased, suggesting an

impaired RNA stability. Second, the NMD-dependent target degradation was indeed dependent on translation, as the construct with a single but strong ATG showed efficient depletion of the PTC-containing transcript. Furthermore, it was shown that the NMD-dependent degradation was not only dependent on translation, but also on the level of translation initiation, and thereby possibly the number of translation events, as the PTC-containing transcripts with a weakened ATG were not degraded efficiently (cf. Figure 3.10). As these results indicate that the intrinsic strength of the translational start codon might influence the NMD-sensitivity of transcripts, the intrinsic strength of HIV-1 mRNAs was analyzed, and some viral transcripts, like *rev*, *vpu,* and *vpr* might indeed escape degradation through NMD by translation initiation at suboptimal Kozak sequences (cf. Table 3.1).

The translation dependency of NMD targets was initially shown by the introduction of a stable hairpin loop into the 5' UTR of a PTC-containing TPI reporter, which blocked the scanning of the 40S ribosomal subunit and drastically reduced the degradation of these target mRNAs (Belgrader, Cheng, and Maquat 1993). That block of translation was further enhanced by the usage of suppressor tRNAs in combination with the 5' UTR hairpin, as suppressor tRNAs insert an amino acid in competition with the translational termination codon and release factor association, they thereby inhibit the pre-mature translation termination and production of NMD targets (Belgrader, Cheng, and Maquat 1993). These observations were supported by the usage of multiple protein synthesis inhibitors which, even though they acted by different mechanisms, all efficiently stabilized PTC-harboring TCR-β-transcripts what resulted in the conclusion that the *de novo* protein synthesis is required for PTC-mediated degradation (Carter et al. 1995). Further on, instead of blocking translation initiation completely, Thermann and colleagues used an iron-response element (IRE) which they inserted into the 5' UTR of a human β-globin reporter. By interaction of the IRE with iron-regulatory proteins (IRPs), they were able to specifically regulate the translation of mRNAs with IREs within their 5' UTRs. They observed stable levels of β-globin WT transcripts independent of the availability of iron, while the PTC-containing variant got degraded co-translationally and was stabilized upon iron-depletion (Thermann et al. 1998). They concluded that ribosome association and cytoplasmic translation are required for NMD.

Even though these results strongly suggest NMD to be translation-dependent, the experiments above abolished translation, either by the introduction of a 5' hairpin structure, the global translation inhibition, or the iron-dependent IRE, which all abolished the translocation of ribosomes over the mRNAs. Therefore, the influence of translation rates was not investigated within the described studies. The analysis of the translational start codon in chapter 3.3 not only suggested a dependency of NMD target recognition by translation but also on the levels of translation (cf. Figure 3.10). The main investigations regarding the intrinsic strength of a translational start codon and translation initiation were performed by Marilyn Kozak back in the 1980ties, which resulted in the discovery of the Kozak sequence for optimal translation

initiation. Already in 1981, she claimed that the favored sequence for eukaryotic translation initiation was (A/G)XXATGG, with a highly conserved purine in position -3 and in position +4, which was based on the analysis of 153 messages (Kozak 1981). For the recognition of the ATG by the eukaryotic ribosomes, she showed that at least one of the two flanking nucleotides was required to be a purine for translation initiation to occur, which she measured by *in vitro* binding assays of synthetic oligonucleotides to wheat germ ribosomes (Kozak 1981). A few years later the consensus sequence for eukaryotic translation initiation was extended to CC(A/G)CCATGG by adding the translational initiation sequences of 166 published cellular mRNAs to the analysis (Kozak 1984a). Furthermore, within the same year, Marilyn Kozak performed an analysis of single nucleotide exchanges near the translational start site of a cloned preproinsulin gene, to determine the influence of the sequence contest on the ATG recognition (Kozak 1984b). She found that sequence changes around the initiator codon drastically affected the yield of proinsulin and was thereby the first person to show that the sequence context influenced the recognition of a given ATG triplet by eukaryotic ribosomes. She showed that the highest impact on protein levels was induced by a change in the nucleotide position -3, where the substitution of C to A enhanced the translation of rat proinsulin more than 15-fold (Kozak 1984b). Even though she analyzed protein levels, she also discussed that the production and stability of mature mRNA might be lower in transcripts with suboptimal initiation sites in comparison to transcripts that harbor the consensus sequence (Kozak 1984b), which fits the results of this work, where the levels of mRNAs without a PTC were reduced upon modification of the translational start codon (cf. Figure 3.10). However, when she extended her experiments on the preproinsulin gene a few years later, she reported no influence of the single base substitutions around the ATG initiator codon on the cytoplasmic RNA levels even though these sequence alterations modulated the yields of proinsulin over a 20-fold range (Kozak 1986).

For high expression genes in yeast, it was shown that the overall ribosome density is higher on transcripts with reduced 5' UTR secondary structure (Ingolia et al. 2009; Robbins-Pianka, Rice, and Weir 2010). Translation rates, however, are not only regulated by the intrinsic strength of a translational start codon and its surrounding nucleotides, but also by the elongation speed, which is dependent on the availability of the desired tRNAs. Even though the genetic code was shown to be degenerated, meaning that a single amino acid can be encoded by various nucleotide codons, and these amino acids are considered synonymous, the cellular amounts of the corresponding tRNAs can be different, thereby regulating the speed by which they can be recognized by the ribosome (Varenne et al. 1984). Therefore, the codon-mediated effects on gene expression are not only dependent on translation but on the number of ribosomes on a given mRNA (Wu et al. 2019), which fits the observed translation-initiation experiments in Figure 3.10. Most research was, however, performed, not on translational start codons, but on the so-called 'codon optimality' which refers to the ability of a given codon to affect mRNA stability in

a translation-dependent manner (Presnyak et al. 2015). Apparently, mRNAs which are enriched in 'optimal' codons showed higher stabilities and abundancies, while 'non-optimal' codons result in unstable RNAs with decreased transcript levels (Bazzini et al. 2016; Mishima and Tomari 2016; Radhakrishnan et al. 2016; Webster et al. 2018; Hanson et al. 2018; Wu et al. 2019). However, as the sequences of the renilla-β-globin transcripts and therefore the encoded amino acid codons were similar within the analyzed constructs, except for the sequence of the translation initiation codon (cf. Figure 3.10), the influence of the codon optimality seems less likely to explain the differences in the transcript levels. Therefore, the performed experiment sheds light on the impact of the Kozak sequence not only on the levels of translation but furthermore on the RNA stability, especially regarding the NMD target recognition by degradation through RNA surveillance mechanisms, such as the nonsense-mediated decay pathway. The observation that PTC-containing transcripts can escape recognition through NMD when the translation initiates at an ATG with a suboptimal Kozak consensus sequence which contains a pyrimidine in position -3 (Pyr.-3) might explain the escape of some transcripts, like HIV-1 transcripts from NMD degradation, based on their low translational initiation levels.

4.4 Alternative splicing and its link to mRNA decay is not restricted to endogenous transcripts

As some HIV-1 transcripts, especially the transcripts of the 2kb class were targeted for degradation by the NMD machinery, an infection-mediated interference of HIV-1 with the cellular NMD activity seemed plausible. The preliminary work of Dr. Nora Diehl using a triosephosphate isomerase NMD reporter showed an accumulation of the NMD-sensitive transcripts upon HIV-1 infection, but not transfection (cf. Figure 3.11). However, the reproduction of this experiment revealed unexpected difficulties, as not only the PTC-containing NMD transcript but also the TPI WT transcript showed NMD-sensitivity by accumulation upon translation inhibition (cf. Figure 3.12). These accumulations were due to alternative splice isoforms of TPI, which generated frameshift mutations resulting in NMD-sensitivity of the generated transcript. These alternative splice isoforms solely derived from the plasmid TPI and arose from skipping of one more multiple TPI exons, not by the usage of cryptic splice sites (cf. Figure 3.12 and Figure 3.13). The sequence analysis confirmed differences between the endogenous encoded TPI and the plasmid-encoded TPI (cf. Figure 3.14), but the adjustment of the plasmid TPI to the sequence of the endogenous TPI did not restore the endogenous splicing pattern (cf. Figure 3.15). Finally, the desired single splice isoform was achieved by increasing the HBond score, hence the complementarity of TPI splice donor 3 to the U1 snRNP by a single base substitution (cf. Figure 3.16) and combining that increased exon 3 inclusion with increased exon 2 recognition, which was achieved by four HEXplorer-based modifications shaping the potential binding sites of SPRs in a way that favored exon 2 inclusion (cf. Figure 3.17). These

modifications finally resulted in a splicing-dependent NMD reporter that exclusively generated a single splice isoform, which was either NMD-sensitive (PTC40) or not (WT).

While it was reported that cryptic splice sites can get activated once the major splice site is deleted (Brillen, Schöneweis, et al. 2017) and that the correct deposition of EJCs upstream of splice sites prevents the usage of cryptic splice sites (Böhm et al. 2018), no reports were found regarding the alternative splicing of an endogenous gene which was transferred into an NMD-splicing reporter. Interestingly, the observed alternative splicing pattern was solely generated by the skipping of one or more exons, but not by the usage of new splice sites which are endogenously not used and would therefore be referred to as cryptic.

The human triosephosphate isomerase gene encodes for a housekeeping enzyme of the glycolytic pathway, which acts as a dimer and catalyzes the inter-conversion of dihydroxyacetone phosphate (DHAP) and D-glyceraldehyde-3-phosphate (D-GAP) (Wierenga, Kapetaniou, and Venkatesan 2010). Over the years, different types of TPI deficiency have been described, some of them encoding nonsense mutations which result in pre-mature translation termination, a truncated protein, and decreased mRNA stability (Daar and Maquat 1988; Cheng and Maquat 1993; Belgrader et al. 1994; Schneider and Cohen-Solal 1996). The seven exon TPI reporter cassettes which were used in this work included four copies of a short β-globin 3' UTR sequence (4H) in the 3' UTR of TPI and were designed by Boehm and colleagues (Böhm et al. 2014). Within their experimental setup, they used the TPI NMD reporters exclusively for northern blot-based analysis, in which they described no observed alternative splicing of the TPI constructs. However, without the RT-PCR-based analysis, one would not expect that the TPI reporter construct generates alternative splice isoforms, as the size difference were not drastic enough to be visible as distinct bands in the northern blot experiments. Therefore, when the WT TPI levels are compared to different PTC-harboring isoforms without the global inhibition of NMD, which would also result in an accumulation of the WT transcript, these alternative splice isoforms might not fall into account. That could be a reason why these splice isoforms have not been reported so far. Regarding the performed RT-PCR experiments within this work, the observed accumulation of NMD target mRNAs in comparison with WT transcripts regarding the CHX-dependent transcript accumulation was rather specific (cf. Figure 3.5). The analysis of the splicing pattern of the used TPI reporter confirmed, that the TPI WT reporter itself generated transcript isoforms that were sensitive to NMD (cf. Figure 3.13) what questioned the usability of this reporter for quantitative NMD experiments. Furthermore, the comparison of the sequences of the endogenous TPI and the plasmid-derived TPI showed differences, like the 5' shortened exon 1 which consists of 246 nucleotides in the Ensembl encoded TPI version and the only 117 nucleotides in the plasmid-encoded TPI exon 1 version (cf. Figure 3.14). As these alterations suggested the usage of another translational start codon in the TPI plasmid version, which was located further downstream and therefore encodes a potentially N-terminally shortened

protein, the missing bases were re-inserted into the plasmid TPI and were shown to have no impact on the splice site usage (cf. Figure 3.15). Furthermore, it was reported that the used plasmid-encoded TPI translational start codon was the same as for the endogenous TPI, therefore the translation of the plasmid-encoded TPI should not result in a truncated protein (Ralser et al. 2006).

It is not uncommon that a variety of genes are regulated by the coordinated action between alternative splicing and NMD. It was recently estimated that more than 80,000 protein-coding transcripts are produced from the 20,000 human genes and that alternative splicing coupled to NMD is widely used to fine-tune the expression of many transcripts and thereby shapes the transcriptome (Garcia-Moreno and Romao 2020). Interestingly, it was estimated that approximately one-third of all alternative splicing events result in the inclusion of an in-frame nonsense codon (Lewis, Green, and Brenner 2003). As the selection of splice sites is not only dependent on the intrinsic strength of a given splice site, but also on the binding of SRPs to SREs within the pre-mRNA, it is remarkable that many RNA binding proteins are regulated by autoregulatory negative feedback loops. Depending on their intracellular protein levels, these RBPs bind their own pre-mRNA and regulate the splice site usage by to fine tune the protein levels, as for example reported for SR and hnRNP protein family members (Lareau et al. 2007; Ni et al. 2007). Furthermore, in addition to the influence of SRPs on splice site recognition, there are reports which claim an influence of the promoter choice on the splice site regulation (Nogues et al. 2002; Xin, Hu, and Kong 2008), which might explain the observed differences in the splice site recognition between the endogenous and plasmid encoded TPI, as the plasmid derived TPI transcripts were expressed under the control of the CMV promoter, which allows high transcription levels in a broad range of cells but obviously differs from the endogenously used TPI promoter (Boshart et al. 1985; Ho et al. 2015).

4.5 Virus-induced modulations of the cellular NMD activity

Upon the successful modifications of the TPI NMD reporter transcriptional unit, which after all only generated the desired single transcript isoform, this system was used to investigate the potential modulations of the cellular NMD machinery by HIV-1 infection or transfection with proviral DNA. While HeLa T4[+] cells nicely demonstrated high RNA levels of the TPI WT transcript and reduced RNA levels of TPI PTC40, neither an infection nor transfection had an impact on the cellular NMD machinery, even though the HIV-1 infection, as well as the pNL4-3 transfection, resulted in increasing levels of viral RNAs over time (cf. Figure 3.18). Furthermore, over the time course of the experiment, the loss of the transiently transfected TPI reporter plasmids was observed, which made it difficult to evaluate the later time points in terms of NMD activity in HeLa T4[+] cells (cf. Figure 3.18). The usage of the HEK293T CD4[+] cell line helped to circumvent the problem of plasmid loss over time, as these cells can amplify transiently transfected plasmids

by the expression of their large T antigen. However, even though the HIV-1 infection and the pNL4-3 transfection again caused increasing viral RNA levels over the analyzed time points, still no effect on the NMD activity was observed (cf. Figure 3.19). Furthermore, the HEK293T CD4+ cells generally showed an impaired degradation of the PTC-containing TPI variant, even in the absence of HIV-1 infection or pNL4-3 transfection, which may suggest a generally reduced NMD activity of these cells after reporter transfection (cf. Figure 3.19).

It has been described that the NMD activity not only varies between different cell strains, as shown for HeLa cells (Viegas et al. 2007) and epithelial cells (Linde et al. 2007), it also differs between different tissues (Zetoune et al. 2008). However, additionally to these cell-type-specific differences, it was reported that the transient transfection of cells can also impact their NMD activity (Gerbracht, Boehm, and Gehring 2017). Within their work, Gerbracht and colleagues observed that three out of six tested cell lines did not show efficient degradation of NMD reporter-derived mRNAs in transient transfection experiments. They concluded that not all human cell lines degrade NMD targets upon transient transfection and that this efficiency of degradation depends on the manner of transfection, not on the amounts of transfected plasmid DNA or the used transfection reagent. Furthermore, they showed that the reduced NMD activity was not due to cellular stress and concluded that mRNAs derived from the expression of transiently transfected DNA might be more resistant to mRNA turnover than endogenous mRNAs (Gerbracht, Boehm, and Gehring 2017). Within their work, they reported an impaired NMD activity in HEK293 cells post-transfection, which fits the results of this work, regarding the enhanced levels of TPI PTC40 RNA levels in comparison to the observed RNA levels in HeLa cells. In general, however, HEK cells showed a slightly decreased NMD activity even without prior transfection, in comparison to HeLa cells, which might be further decreased by the transient transfection (cf. Figure 3.5). With the transiently transfected reporter plasmids the question of whether an HIV-1 infection or the transfection with proviral DNA influences the activity of the cellular NMD pathway could not be answered. Therefore, the major question regarding the potential escape of the HI-virus from NMD and the thereby achieved protection of the viral RNAs from degradation remains to be answered. So far different strategies of viral evasion from NMD have been described, which either protect the viral transcripts themselves from being recognized by the degradation machinery, hence in *cis*, or by direct interference of viral proteins with components of the NMD machinery, thereby blocking NMD in *trans*. One of the best described examples for NMD-interference in *cis* is given by the positive-sense Rous Sarcoma retrovirus (RSV). The RSV gag transcript encodes the RSE, a stability element with a minimal length of 250bp, which is located downstream of the gag TC. The RSE recruits the hosts' polypyrimidine tract-binding protein 1 (PTBP1), thereby preventing binding of UPF1 to the viral RNA and subsequently the degradation by NMD (Barker and Beemon 1994; Withers and Beemon 2011; Ge et al. 2016). Another way of NMD antagonization is the expression of viral proteins

that directly bind to NMD factors and thereby inhibit NMD function. One favorite example of NMD inhibition in *trans* is given by the human T-cell lymphotropic virus type 1 (HTLV-1), another retrovirus. Here, the viral proteins Tax and Rex were shown to inhibit NMD in addition to their roles in the viral replication cycle (Nakano et al. 2013; Fiorini et al. 2018).

It was shown that not only retroviruses, which replicate in the nucleus of its host cell and make use of the splicing machinery, can interfere with the host's NMD machinery. For example, the causative agent of the COVID-19 pandemic – the enveloped, positive-sense, single-stranded RNA virus of the Coronaviridae family – severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) replicates in the cytoplasm and does not need the splicing machinery for its gene expression (Kim et al. 2020). Using mouse hepatitis virus (MHV) as a model for SARS-CoV, the group of Shinji Makino showed that the presence of unusually long 3' UTRs of the viral RNAs made these transcripts targets for degradation through NMD, which the virus counteracted by expression of the viral N-protein, which was shown to inhibit NMD in *trans* (Wada et al. 2018). The N-protein of SARS-CoV-2 was shown to copurify with UPF1 and MOV10 what suggests that the interference of SARS-CoV and SARS-CoV-2 with the NMD machinery might be conserved (Gordon et al. 2020).

As simple as it may sound, the co-evolution of viruses and their host cells not only resulted in a variety of strategies regarding antiviral defense mechanisms, as one of which NMD might also function but also in diverse strategies of viral escape, for example from the degradation of viral transcripts through NMD. This situation, however, seems more complicated regarding HIV-1, as there is still no clear picture of the NMD-based interference with HIV-1 transcripts. If an HIV-1 infection, or even the presence of the proviral DNA, would indeed interfere with the cellular NMD machinery that effect was expected to be visible through the accumulation of the TPI encoded NMD reporter (cf. Figure 3.18 and Figure 3.19). Almost twenty years ago, it was postulated that HIV-1 controls NMD by the tight control over splicing and EJC deposition (Bohne, Wodrich, and Kräusslich 2005), which fits the results of this work where only the 2kb class of HIV-1 mRNAs was targeted by NMD (cf. chapter 3.1). However, while one group reported an NMD-mediated restriction of HIV-1 replication, as the depletion of UPF2 and SMG6 enhanced viral RNA expression in primary monocyte-derived macrophages (Rao et al. 2019), other groups reported UPF1 to be a positive regulator of HIV-1 gene expression, which gets hijacked by the virus to assist in reverse transcription as well as export of the viral RNA (Ajamian et al. 2008; Serquina et al. 2013). Therefore, the impact of HIV-1 infection on the cellular NMD activity remains to be deciphered. In general, it was proposed that viruses rather shape the NMD activity of the cell, than inhibiting it completely, as the elimination of NMD might have toxic side effects (Popp, Cho, and Maquat 2020).

4.6   Lentiviral vector systems for NMD reporter delivery

Based on the observed sensitivity of the cellular NMD activity upon transient transfection experiments, a new approach was used to circumvent the transfection process but maintain the desired readout couple of a WT and NMD-sensitive splicing reporter. Therefore, a third-generation lentiviral vector system was established, which encoded the TPI exon 1-7 cassette with corresponding introns either with or without PTC (cf. Figure 3.20). The initial analysis of cells, which were transduced with either one of these lentiviral vectors, however, showed no differences in the TPI RNA levels regarding the WT or PTC-containing isoform (cf. Figure 3.21). Subsequent analysis confirmed that this was due to the previous excision of the TPI introns by the splicing apparatus of the lentivirus-producing cells, even before packaging of the transgene into the assembling lentiviral particles (cf. Figure 3.22). These reporter-encoded introns, however, were critical for the desired NMD-sensitivity of the transduced reporter constructs, as the splicing process and subsequent EJC deposition would mark the encoded PTC as pre-mature, due to its position upstream of the last exon-exon junction. To circumvent splicing of the producer cell line, a new approach was used to establish the lentiviral vector system. Therefore, the dual-luciferase NMD reporter system (cf. Figure 3.9), which contained the firefly luciferase as internal control, was inserted in reverse orientation to abolish splice site recognition by the producer cell line (cf. Figure 3.23). The generated lentiviral particles produced promising results, as the transduction experiments resembled the transfection control experiments (cf. Figure 3.24). Using this system, the question of whether HIV-1 infection modulates the cellular NMD machinery was investigated using HeLa T4$^+$ cells, which were stably transduced with either the WT or the NMD version of the generated lentiviral particles. However, based on the observed lack of difference between the luciferase activities of the NMD reporter constructs upon HIV-1 infection in comparison with the mock samples, no significant effect of HIV-1 infection on the activity of the cellular NMD machinery was observed (cf. Figure 3.25). Furthermore, the determination of the minimum threshold for Firefly and Renilla luciferase revealed that the Renilla luciferase activity in the PTC-containing variant was below the detection limit (cf. Figure 3.26). Therefore, the possible fine-tuning of cellular NMD activity that might be mediated by HIV-1 infection could not be measured in this setup. Based on these results, however, a complete silencing of NMD activity by HIV-1 infection seems unlikely, as this should enhance Renilla luciferase above the threshold.

Lentiviral vectors are commonly used to generate stable cell lines, as the evolution of lentiviral vectors and the establishment of third-generation lentiviral vector systems dramatically increased the safety and reliability of these constructs. As the first vectors of the split-component system still contained a significant portion of the HIV genome, most of these viral sequences were deleted over the lentiviral vector generations (Milone and O'Doherty 2018). While the second generation already lacked the accessory genes *vif*, *vpr*, *vpu*, and *nef*, the third

generation of lentiviral vectors allowed the removal of *tat*, as a constitutively active promoter replaced the viral LTR promoter. Furthermore, deletions in the viral 3' LTR resulted in self-inactivating (SIN) lentiviral vectors which greatly enhanced the safety of this system (Dull et al. 1998).

For cells that were transduced with lentiviral vectors, the general expression of the encoded transgene is not only dependent on the used promotor, but also on the chromatin density at the integration site. Apparently, the place of integration is not random, as many retroviruses were shown to have a characteristic preference, which exemplary for HIV-1 is the preferential integration into active transcription sites (Mitchell et al. 2004). Furthermore, not only the site of integration and the promoter activity are important for the success of a lentiviral transduction system and transgene expression, it was shown that also the length of the transgene has an impact on the generation of lentiviruses, as the viral titers decreased with the size of the insert (Kumar et al. 2001).

Here, the analysis of HeLa T4$^+$ cells, which were transduced with the TPI-encoding lentiviruses, showed not only the loss of TPI introns, which were removed before packaging but also alternative splice isoforms of TPI which were not expected (cf. Figure 3.22). These isoforms might arise from the integration process of the lentiviral vector, as this was shown to induce alternative splicing and can promote the generation of aberrant transcripts (Moiani et al. 2012). Within their work, Moiani and colleagues observed alternative splicing and chimeric transcripts upon lentiviral transduction and showed that these transcripts arose from the usage of constitutive and cryptic splice sites not only within the HIV-1 5' LTR and *gag* sequence but also within their inserted transgenes (Moiani et al. 2012).

However, as the proper function of the used NMD reporters required the excision of the encoded introns by splicing and the deposition of EJCs on the final transcript, it was necessary to include the reporter encoded introns within the lentiviral vector setup. As mentioned before, this was difficult, as the HEK293T producer cells already removed the introns before RNA packaging (cf. Figure 3.22). This pre-mature splicing of transgenic RNA in the nucleus of the producing cell line is a major problem in the production of lentiviral vectors that are supposed to contain genes with intron sequences since these are already removed by the splicing machinery of the producer cell before they are packaged into lentiviruses. One approach to circumvent intron excision was described by Li and Garoff, who used the Semliki Forest virus (SFV) expression system to synthesize the viral RNA in the cytoplasm of the producer cell, not in the nucleus (Li and Garoff 1998). Using this SFV system, they observed efficient packaging of the intron-containing RNAs into retroviral vectors. Generally, the inclusion of an antisense β-globin gene with its corresponding introns and the genomic promoter into a murine leukemia virus vector was already described more than 30 years ago (Karlsson et al. 1987). However, the

presence of a constitutively active antisense promoter might produce antisense transcripts, that can basepair with the vectors RNA genome, generating dsRNAs in the packaging cell line that might activate the innate antiviral immune response. Therefore another method to include an intron into a lentiviral vector and even enhance the transduction-induced protein levels was recently described by Bryan Cullen's group, who generated lentiviral vectors in which the transgene was expressed under the control of an antisense-orientated, inducible promoter which drove the expression of an antisense expression cassette which contained a functional intron (Poling et al. 2017). However, even though several methods have been described to generate lentiviral particles which contain a gene cassette with introns the reverse orientation of the target gene appears to be the most feasible method to trick the cellular splicing machinery from recognizing the intron-exon structures and preventing intron excision prior to RNA packaging.

4.7 The absence of IFN response and modulation of the cellular NMD activity upon HIV-1 infection

To investigate whether the innate immune system might be linked to the activity of the cellular NMD machinery, the impact of recombinant interferons on the NMD activity was tested. Using the transiently transfected dual-luciferase NMD reporter system as readout, a clear impact of interferons on the cellular NMD activity was observed (cf. Figure 3.27). All tested interferons were able to strongly reduce the NMD activity, a process that was inhibited by the addition of Ruxolitinib before IFN treatment. Furthermore, the used recombinant IFNs also induced the expression of interferon-stimulated genes (ISGs) in cell culture T-cell lines (cf. Figure 3.28), ensuring that the recombinant IFNs were functional. In HIV-1 infection experiments, however, there was neither an enhanced expression of ISGs over the measured time course nor an impact on the expression of cellular genes with NMD isoforms (SRSF3 and hnRNPL), hence the cellular NMD activity (cf. Figure 3.29).

Interferons are major players of the innate immune response and can rapidly induce the expression of a large network of ISGs, which alter the gene expression of the cell to exhibit diverse functions in infection control and viral defense (Uehata and Takeuchi 2020). Also, in HIV-1 infection, innate immunity and IFNs serve as the first line of defense, based on their antiviral activity (Poli et al. 1989; Levy, Scott, and Mackewicz 2003). In general, the production of IFNs is not limited to immune cells and can arise from many cell types, including endothelial cells (Haque et al. 2018). Not only the expression of ISGs is controlled by IFNs, but also post-transcriptional regulations like pre-mRNA splicing are important regulators of gene expression, as the disruption of splicing can result in the production of potentially harmful proteins and disease (Chabot and Shkreta 2016). As alternative splicing is often coupled with the degradation of target mRNAs by NMD, the potential link between transcript degradation and antiviral

immune response should be investigated. In plants, for example, two different groups reported that the NMD efficiency was reduced upon infection with bacteria or viruses, which was shown by the accumulation of endogenous NMD targets (Gloggnitzer et al. 2014; Garcia, Garcia, and Voinnet 2014). The described reduction of the cellular NMD activity fits to the results of these work, in which the treatment of cells with recombinant IFNs had a similar effect (cf. Figure 3.27). Even though NMD factors are constitutively expressed and are not classified as ISGs, there were shown to modulate the cellular levels of immune receptors in plants (Schoggins et al. 2011; Rigby and Rehwinkel 2015; Gloggnitzer et al. 2014). Within this work, the dual-luciferase NMD reporter system was used to measure the impact of recombinant interferons on the activity of the cellular NMD machinery (cf. Figure 3.27). This readout system allowed the normalization of the GOI renilla luciferase to firefly as internal control with comparable efficiency, as both luciferases were encoded on the same plasmid. This was important to exclude transfection-based differences as it was described before that IFN treatment can have an impact on the expression of luciferase reporters (Ghazawi et al. 2005).

Even though cells can switch into an antiviral state upon IFN induction, an IFN expression upon HIV-1 infection is not measurable in all cell types, like it was shown for primary monocyte-derived macrophages (Rasaiyaah et al. 2013). Here, type I IFN expression was only induced upon HIV-1 capsid destabilization and subsequent leakage of reverse transcripts into the cytoplasm, which were sensed by the cyclic GMP-AMP (cGAMP) synthase (cGAS) (Rasaiyaah et al. 2013). Within immortalized human T cell lines, however, the expression of cGAS and therefore the cGAS-dependent cellular responses differed drastically between the cell lines (Elsner et al. 2020). Even though some T cell lines, like PM1, showed high levels of cGAS, while Jurkat cell lines were shown to express only low levels of cGAS RNA and protein, they could not detect cGAS-dependent cellular responses upon productive HIV-1 infection, indicating efficient escape of cGAS recognition by HIV-1 (Elsner et al. 2020). Furthermore, the reverse transcribed HIV-1 cDNA can also be sensed intracellularly by the IFN inducible protein 16 (IFI16) which, like cGAS, can activate the stimulator of interferon gene (STING) to produce IFNs (Jakobsen et al. 2013; Soper et al. 2017). However, Berg and colleagues reported that, even though T cells can detect the intracellular viral DNA via IFI16, they fail to induce IFN response and ISG induction (Berg et al. 2014), which fits the observed absence of ISG induction upon productive HIV-1 infection (cf. Figure 3.29).

4.8 Splicing regulatory proteins, negative feedback loops, and NMD - the many facets of HIV-1 infection

As transient transfection might impair the cellular NMD activity, and transduction of cells with lentiviral encoded NMD reporters did not result in sufficient reporter levels, the last step of this work focused on the expression of endogenous genes during HIV-1 infection. The selected genes

showed high expression levels and encoded alternative splice variants which were targeted by NMD. Using quantitative PCR as readout, the potential HIV-1 infection-mediated influence on the cellular NMD activity, which could not be detected by the usage of NMD reporter plasmids, should be investigated. The two selected endogenous genes were chosen based on their Ensembl annotated NMD target splice isoforms and their high expression levels in our PM1 cells deep sequencing data (Diehl 2016). The amplification of these two genes, which were SRSF7 and hnRNPDL, was performed with specific exon junction primers. The amplification of either the NMD-sensitive or the not-NMD-sensitive isoform revealed two things. First, the two tested genes differed in their sensitivity for degradation through NMD, as the NMD-sensitive isoform of hnRNPDL was more abundant than the NMD-sensitive splice isoform of SRSF7 in untreated cells (cf. Figure 3.30). Second, as observed before (cf. Figure 3.5), the NMD efficiency differed between the two tested cell lines (cf. Figure 3.30). However, as the designed exon-junction primers allowed the specific detection of either the NMD-sensitive or the not-NMD-sensitive splice isoform, those primers were used in quantitative PCR reactions to investigate the abundance of these splice isoforms in PM1 cells over the course of HIV-1 infection (cf. Figure 3.31). By dividing the fold change of NMD-sensitive splice isoforms to the fold change of the not-NMD-sensitive splice isoforms, the relative abundance of the NMD-sensitive splice isoform was plotted against the fold change of HIV-1 exon 7 as marker for the infection progression, an accumulation of the SRSF7 NMD-sensitive transcripts but not hnRNPDL NMD-sensitive transcripts was observed (cf. Figure 3.31). Therefore, throughout HIV-1 infection there was an impaired degradation of NMD-sensitive SRSF7 splice isoforms, which could suggest an impaired NMD activity during HIV-1 infection. However, as this accumulation was not observed for the hnRNPDL NMD-sensitive transcripts, HIV-1 does not globally inhibit the cellular NMD machinery.

Throughout the literature, the investigation of the NMD activity of a given cell line was often performed by the amplification of 'physiologic NMD substrates', using a single primer pair to measure the abundancies of NMD targets, like *ASNS* and *CARS* (Mendell et al. 2004; Linde et al. 2007). Here, the usage of specific exon-junction primers allowed the direct distinction between NMD-sensitive and not-NMD-sensitive transcript isoforms of SRSF7 and hnRNPDL, as this allowed the amplification of each transcript isoform, containing either the NMD-specific or the not-NMD-specific exon-exon junction. Therefore, this system allowed a much more specific readout of transcript abundancies (cf. Figure 3.30).

In general, the NMD-sensitivity of most NMD reporter constructs, but also of the two selected genes with NMD-sensitive splice isoforms, not only relies on the presence of a PTC more than 50-55 nucleotides upstream of the last exon-exon junction, but also on the presence of EJCs downstream of the PTC for efficient NMD target recognition and degradation. Using single-molecule imaging Hoek and colleagues showed that not only the distance of a PTC to the next EJC defines the NMD-sensitivity, but that the number of introns that are located downstream of

the PTC, hence the number of deposited EJCs on the mRNA correlates with the efficiency of target degradation through NMD (Hoek et al. 2019). These results are in line with the observed abundancies of the SRSF7 and hnRNPDL NMD-sensitive splice isoforms in untreated HeLa and HEK293T cells in RT-PCR experiments (cf. Figure 3.30). Here, the NMD-sensitive splice isoform of SRSF7, with five exon-exon junctions downstream of its PTC, was much less abundant than the NMD-sensitive splice isoform of hnRNPDL with only one exon-exon junction downstream of the PTC (cf. Figure 3.30).

The splicing regulatory protein SRSF7, which is also known as 9G8, belongs to the SR protein family and is the only SR protein that contains a zinc knuckle in addition to its RNA recognition motif (RRM) (Cavaloc et al. 1999). HnRNPDL, however, is a paralog of hnRNPD, with whom it shares high sequence similarity of its RNA binding domains and was therefore suggested to recognize similar RNA target sequences (Doi et al. 1998). The transcript levels of both genes, SRSF7, and hnRNPDL, are regulated through negative feedback loops. SRSF7 was shown to regulate its RNA levels by the inclusion of an ultraconserved alternative exon, called poison cassette exon (PCE) which results in the inclusion of a PTC and rapid degradation of the transcript by NMD (Lareau et al. 2007; Pervouchine et al. 2019). The autoregulation of this feedback loop was recently described by Michaela Müller-McNicolls group, who showed that SRSF7, at low intrinsic protein levels, binds to the flanking exons of the PCE within its own pre-mRNA and thereby mediates PCE skipping and production of functional SRSF7 protein (Königs et al. 2020). Upon overexpression, however, SRSF7 binds to the splice sites of the PCE and promotes its inclusion to generate the NMD-sensitive SRSF7 transcript isoform which gets degraded by NMD and thereby decreases the SRSF7 protein levels ensuring protein homeostasis (Königs et al. 2020). Furthermore, SRSF7 levels are not only regulated by the intrinsic SRSF7 levels but PCE inclusion can also be promoted via cross-regulation through SRSF3 binding (Änkö et al. 2012). Therefore, this regulatory mechanism might also represent an antiviral defense mechanism, as some viruses, like HIV-1, hijack SR proteins for the processing of their genome (Königs et al. 2020). However, the increasing amounts of NMD-sensitive SRSF7 splice isoform in PM1 cells upon HIV-1 infection (cf. Figure 3.31) might therefore be not a direct effect of infection-mediated NMD inhibition, but rather an indirect effect which gets mediated by SRSF7 hijacking through HIV-1.

Even though NMD-sensitive isoforms of hnRNPD were described already 15 years ago, the auto- and cross-regulation of hnRNPD and hnRNPDL transcripts was only recently described (Banihashemi et al. 2006; Kemmerer, Fischer, and Weigand 2018). Exon 8, which encodes the PCE of hnRNPDL, is located within the ultraconserved element, within the 3' UTR of hnRNPDL. Kemmerer and colleagues observed increased detection of hnRNPDL exon 8 containing transcripts upon puromycin treatment and UPF1 knock-down, suggesting that this splice isoform is targeted by NMD like observed with CHX treatment in this work (cf. Figure 3.30) (Kemmerer,

Fischer, and Weigand 2018). Furthermore, by fusing the hnRNPDL 3' UTR to a luciferase reporter gene and overexpressing hnRNPDL, they showed that the hnRNPDL protein levels themselves had an impact on PCE inclusion, confirming the expected autoregulatory feedback loop for hnRNPDL (Kemmerer, Fischer, and Weigand 2018).

As both genes, SRSF7 and hnRNPDL, are regulated through negative feedback loops, high expression of each protein should result in binding to its own pre-mRNA, thereby promoting PCE inclusion, resulting in enhanced levels of NMD-sensitive mRNA isoforms and reduced levels of WT mRNAs. Upon HIV-1 infection, the levels of the NMD-sensitive SRSF7 splice isoform increased in relation to the not-NMD-sensitive isoform, while this increase was not observed for the NMD-sensitive isoform of hnRNPDL (cf. Figure 3.31). Therefore, an outstanding question is whether HIV-1 infection selectively interferes with NMD, or if the infection enhances the levels of SRSF7 protein, but not hnRNPDL protein.

Even though there are no reports about the expression of hnRNPDL during HIV-1 infection, and only little is known about the role of SRSF7 in HIV-1 splicing, there are some data on SR protein regulations and overexpression during HIV-1 infection. For example, the expression of SRSF2 in H9 T-lymphocytic cells was shown to increase 2-3-fold two days post HIV-1 infection (Maldarelli et al. 1998). While the characterization of differently expressed genes in HIV-1 infected human T-cells showed a down-regulation of SRSF7 mRNAs 60 hours post-infection (Ryo et al. 2000). Furthermore, the group of Marek Widera very recently analyzed the expression levels of SRSF mRNAs in chronically HIV-1 infected patients, with or without antiretroviral (ART) therapy and reported significantly lower mRNA levels of SRSF1, SRSF3, SRSF7, and SRSF10 in chronically infected patients in comparison to healthy controls (Sertznig et al. 2021). In general, the effect of distinct SR proteins on HIV-1 alternative splicing was mostly investigated by overexpression studies, even though SRSF7 was reported to be down-regulated in HIV-1 infected cells. However, the over-expression of either tested SR protein reduced the levels of genomic HIV-1 RNA, structural proteins, and virion production in 293T cells, while each SR protein modified the HIV-1 splicing pattern in a distinct way, where the overexpression of SRSF1 increased the levels of *vpr* mRNA, while SRSF2 and SRSF7 over-expression resulted in enhanced levels of *tat* mRNA (Jacquenet et al. 2005). Furthermore, SR proteins can not only influence the *tat* mRNA levels, they can furthermore impair Tat activity as it was shown by Massimo Caputi's group, who described that HIV-1 transcription is regulated mostly by SRSF1, but also other members of the SR protein family, like SRSF3 and SRSF7, by binding to the viral TAR element and thereby competing with Tat and downregulating Tat activity (Paz, Krainer, and Caputi 2014).

In conclusion, the described reports suggest an HIV-1 infection-mediated reduction of SRSF7 mRNA and protein levels, which would normally result in similarly decreased levels of NMD-sensitive SRSF7 splice isoforms, as this isoform is regulated through a negative feedback loop

and should be induced upon high SRSF7 protein levels. Therefore, the observed accumulation of NMD-sensitive SRSF7 transcript might be due to a reduced cellular NMD activity, resulting in higher detection of the normally degraded transcript (cf. Figure 3.31). However, this remains to be deciphered, as the SRSF7 protein levels were not measured in this experimental setup and the NMD-sensitive isoform of hnRNPDL showed no accumulation upon HIV-1 infection. Alternatively, the observed difference in NMD-target accumulation between SRSF7 and hnRNPDL might also result from the different NMD-sensitivities of the two genes (cf. Figure 3.30). If the potential HIV-1-induced NMD downregulation does not fall into account regarding the already abundant hnRNPDL transcript isoforms, that might explain the difference to SRSF7, where the NMD-sensitive splice isoform was only hardly detectable in the mock situation and without translational inhibitors. Additionally, the observed accumulation of SRSF7 NMD-sensitive splice isoforms was only visible upon very high viral exon 7 levels (cf. Figure 3.31). These high exon 7 levels might also suggest apoptosis of the host cell which might also lead to NMD-target accumulation. Another possibility might be the infection-based UPF1-depletion in the cytoplasm, as UPF1 proteins get hijacked by HIV-1 for replication and virus production as reported by Serquina, Ajamian, and colleagues (Serquina et al. 2013; Ajamian et al. 2015).

Overall, the results in this work suggest that, even though some HIV-1 transcripts, like the viral 2kb class are targeted by degradation through NMD, there seems to be no active inhibition of the NMD machinery by HIV-1. The observed dysregulation of the NMD activity might result from IFN production, as this can interfere with the NMD activity (cf. Fig 3.27), or very efficient viral replication due to UPF1 hijacking by HIV-1 which results in the accumulation of designated target transcripts like the NMD-sensitive isoform of SRSF7 (cf. Fig 3.31).

5   Sources

Achsel, T., H. Brahms, B. Kastner, A. Bachi, M. Wilm, and R. Lührmann. 1999. 'A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro', *EMBO J*, 18: 5789-802.

Agafonov, D. E., B. Kastner, O. Dybkov, R. V. Hofele, W. T. Liu, H. Urlaub, R. Lührmann, and H. Stark. 2016. 'Molecular architecture of the human U4/U6.U5 tri-snRNP', *Science*, 351: 1416-20.

Ajamian, L., K. Abel, S. Rao, K. Vyboh, F. Garcia-de-Gracia, R. Soto-Rifo, A. E. Kulozik, N. H. Gehring, and A. J. Mouland. 2015. 'HIV-1 Recruits UPF1 but Excludes UPF2 to Promote Nucleocytoplasmic Export of the Genomic RNA', *Biomolecules*, 5: 2808-39.

Ajamian, L., L. Abrahamyan, M. Milev, P. V. Ivanov, A. E. Kulozik, N. H. Gehring, and A. J. Mouland. 2008. 'Unexpected roles for UPF1 in HIV-1 RNA metabolism and translation', *RNA*, 14: 914-27.

Akari, H., M. Fujita, S. Kao, M. A. Khan, M. Shehu-Xhilaga, A. Adachi, and K. Strebel. 2004. 'High level expression of human immunodeficiency virus type-1 Vif inhibits viral infectivity by modulating proteolytic processing of the Gag precursor at the p2/nucleocapsid processing site', *J Biol Chem*, 279: 12355-62.

Anderson, J. L., A. T. Johnson, J. L. Howard, and D. F. Purcell. 2007. 'Both linear and discontinuous ribosome scanning are used for translation initiation from bicistronic human immunodeficiency virus type 1 env mRNAs', *J Virol*, 81: 4664-76.

Änkö, M. L., M. Müller-McNicoll, H. Brandl, T. Curk, C. Gorup, I. Henry, J. Ule, and K. M. Neugebauer. 2012. 'The RNA-binding landscapes of two SR proteins reveal unique functions and binding to diverse RNA classes', *Genome Biol*, 13: R17.

Arhel, N. 2010. 'Revisiting HIV-1 uncoating', *Retrovirology*, 7: 96.

Arias-Palomo, E., A. Yamashita, I. S. Fernandez, R. Nunez-Ramirez, Y. Bamba, N. Izumi, S. Ohno, and O. Llorca. 2011. 'The nonsense-mediated mRNA decay SMG-1 kinase is regulated by large-scale conformational changes controlled by SMG-8', *Genes Dev*, 25: 153-64.

Asang, C., I. Hauber, and H. Schaal. 2008. 'Insights into the selective activation of alternatively used splice acceptors by the human immunodeficiency virus type-1 bidirectional splicing enhancer', *Nucleic Acids Res*, 36: 1450-63.

Aviner, R. 2020. 'The science of puromycin: From studies of ribosome function to applications in biotechnology', *Comput Struct Biotechnol J*, 18: 1074-83.

Aznarez, I., T. T. Nomakuchi, J. Tetenbaum-Novatt, M. A. Rahman, O. Fregoso, H. Rees, and A. R. Krainer. 2018. 'Mechanism of Nonsense-Mediated mRNA Decay Stimulation by Splicing Factor SRSF1', *Cell Rep*, 23: 2186-98.

Azzalin, C. M., and J. Lingner. 2006. 'The human RNA surveillance factor UPF1 is required for S phase progression and genome stability', *Curr Biol*, 16: 433-9.

Balistreri, G., C. Bognanni, and O. Mühlemann. 2017. 'Virus Escape and Manipulation of Cellular Nonsense-Mediated mRNA Decay', *Viruses*, 9.

Balistreri, G., P. Horvath, C. Schweingruber, D. Zund, G. McInerney, A. Merits, O. Mühlemann, C. Azzalin, and A. Helenius. 2014. 'The host nonsense-mediated mRNA decay pathway restricts Mammalian RNA virus replication', *Cell Host Microbe*, 16: 403-11.

Banerji, J., S. Rusconi, and W. Schaffner. 1981. 'Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences', *Cell*, 27: 299-308.

Banihashemi, L., G. M. Wilson, N. Das, and G. Brewer. 2006. 'Upf1/Upf2 regulation of 3' untranslated region splice variants of AUF1 links nonsense-mediated and A+U-rich element-mediated mRNA decay', *Mol Cell Biol*, 26: 8743-54.

Barash, Y., J. A. Calarco, W. Gao, Q. Pan, X. Wang, O. Shai, B. J. Blencowe, and B. J. Frey. 2010. 'Deciphering the splicing code', *Nature*, 465: 53-9.

Barberan-Soler, S., N. J. Lambert, and A. M. Zahler. 2009. 'Global analysis of alternative splicing uncovers developmental regulation of nonsense-mediated decay in C. elegans', *RNA*, 15: 1652-60.

Barker, G. F., and K. Beemon. 1994. 'Rous sarcoma virus RNA stability requires an open reading frame in the gag gene and sequences downstream of the gag-pol junction', *Mol Cell Biol*, 14: 1986-96.

Barre-Sinoussi, F., J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum, and L. Montagnier. 1983. 'Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS)', *Science*, 220: 868-71.

Bazzini, A. A., F. Del Viso, M. A. Moreno-Mateos, T. G. Johnstone, C. E. Vejnar, Y. Qin, J. Yao, M. K. Khokha, and A. J. Giraldez. 2016. 'Codon identity regulates mRNA stability and translation efficiency during the maternal-to-zygotic transition', *EMBO J*, 35: 2087-103.

Beelman, C. A., and R. Parker. 1994. 'Differential effects of translational inhibition in cis and in trans on the decay of the unstable yeast MFA2 mRNA', *J Biol Chem*, 269: 9687-92.

Behrens, R. T., M. Aligeti, G. M. Pocock, C. A. Higgins, and N. M. Sherer. 2017. 'Nuclear Export Signal Masking Regulates HIV-1 Rev Trafficking and Viral RNA Nuclear Export', *J Virol*, 91.

Belgrader, P., J. Cheng, and L. E. Maquat. 1993. 'Evidence to implicate translation by ribosomes in the mechanism by which nonsense codons reduce the nuclear level of human triosephosphate isomerase mRNA', *Proc Natl Acad Sci U S A*, 90: 482-6.

Belgrader, P., J. Cheng, X. Zhou, L. S. Stephenson, and L. E. Maquat. 1994. 'Mammalian nonsense codons can be cis effectors of nuclear mRNA half-life', *Mol Cell Biol*, 14: 8219-28.

Berg, R. K., S. H. Rahbek, E. Kofod-Olsen, C. K. Holm, J. Melchjorsen, D. G. Jensen, A. L. Hansen, L. B. Jorgensen, L. Ostergaard, M. Tolstrup, C. S. Larsen, S. R. Paludan, M. R. Jakobsen, and T. H. Mogensen. 2014. 'T cells detect intracellular DNA but fail to induce type I IFN responses: implications for restriction of HIV replication', *PLoS One*, 9: e84513.

Bhattacharya, A., K. Czaplinski, P. Trifillis, F. He, A. Jacobson, and S. W. Peltz. 2000. 'Characterization of the biochemical properties of the human Upf1 gene product that is involved in nonsense-mediated mRNA decay', *RNA*, 6: 1226-35.

Black, D. L. 2003. 'Mechanisms of alternative pre-messenger RNA splicing', *Annu Rev Biochem*, 72: 291-336.

Böhm, V., T. Britto-Borges, A. L. Steckelberg, K. K. Singh, J. V. Gerbracht, E. Gueney, L. Blazquez, J. Altmüller, C. Dieterich, and N. H. Gehring. 2018. 'Exon Junction Complexes Suppress Spurious Splice Sites to Safeguard Transcriptome Integrity', *Mol Cell*, 72: 482-95 e7.

Böhm, V., J. V. Gerbracht, M. C. Marx, and N. H. Gehring. 2016. 'Interrogating the degradation pathways of unstable mRNAs with XRN1-resistant sequences', *Nat Commun*, 7: 13691.

Böhm, V., N. Haberman, F. Ottens, J. Ule, and N. H. Gehring. 2014. '3' UTR length and messenger ribonucleoprotein composition determine endocleavage efficiencies at termination codons', *Cell Rep*, 9: 555-68.

Bohne, J., H. Wodrich, and H. G. Kräusslich. 2005. 'Splicing of human immunodeficiency virus RNA is position-dependent suggesting sequential removal of introns from the 5' end', *Nucleic Acids Res*, 33: 825-37.

Bono, F., J. Ebert, E. Lorentzen, and E. Conti. 2006. 'The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA', *Cell*, 126: 713-25.

Boshart, M., F. Weber, G. Jahn, K. Dorsch-Hasler, B. Fleckenstein, and W. Schaffner. 1985. 'A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus', *Cell*, 41: 521-30.

Briggs, J. A., and H. G. Kräusslich. 2011. 'The molecular architecture of HIV', *J Mol Biol*, 410: 491-500.

Brillen, A. L., K. Schöneweis, L. Walotka, L. Hartmann, L. Müller, J. Ptok, W. Kaisers, G. Poschmann, K. Stuhler, E. Buratti, S. Theiss, and H. Schaal. 2017. 'Succession of splicing regulatory elements determines cryptic 5ss functionality', *Nucleic Acids Res*, 45: 4202-16.

Brillen, A. L., L. Walotka, F. Hillebrand, L. Müller, M. Widera, S. Theiss, and H. Schaal. 2017. 'Analysis of Competing HIV-1 Splice Donor Sites Uncovers a Tight Cluster of Splicing Regulatory Elements within Exon 2/2b', *J Virol*, 91.

Bringmann, P., B. Appel, J. Rinke, R. Reuter, H. Theissen, and R. Lührmann. 1984. 'Evidence for the existence of snRNAs U4 and U6 in a single ribonucleoprotein complex and for their association by intermolecular base pairing', *EMBO J*, 3: 1357-63.

Bringmann, P., and R. Lührmann. 1986. 'Purification of the individual snRNPs U1, U2, U5 and U4/U6 from HeLa cells and characterization of their protein constituents', *EMBO J*, 5: 3509-16.

Bühler, M., S. Steiner, F. Mohn, A. Paillusson, and O. Mühlemann. 2006. 'EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length', *Nat Struct Mol Biol*, 13: 462-4.

Burdick, R. C., C. Li, M. Munshi, J. M. O. Rawson, K. Nagashima, W. S. Hu, and V. K. Pathak. 2020. 'HIV-1 uncoats in the nucleus near sites of integration', *Proc Natl Acad Sci U S A*, 117: 5486-93.

Burley, S. K., and R. G. Roeder. 1996. 'Biochemistry and structural biology of transcription factor IID (TFIID)', *Annu Rev Biochem*, 65: 769-99.

Campbell, E. M., and T. J. Hope. 2015. 'HIV-1 capsid: the multifaceted key player in HIV-1 infection', *Nat Rev Microbiol*, 13: 471-83.

Carlson, L. A., J. A. Briggs, B. Glass, J. D. Riches, M. N. Simon, M. C. Johnson, B. Müller, K. Grünewald, and H. G. Kräusslich. 2008. 'Three-dimensional analysis of budding sites and released virus suggests a revised model for HIV-1 morphogenesis', *Cell Host Microbe*, 4: 592-9.

Carninci, P., A. Sandelin, B. Lenhard, S. Katayama, K. Shimokawa, J. Ponjavic, C. A. Semple, M. S. Taylor, P. G. Engstrom, M. C. Frith, A. R. Forrest, W. B. Alkema, S. L. Tan, C. Plessy, R. Kodzius, T. Ravasi, T. Kasukawa, S. Fukuda, M. Kanamori-Katayama, Y. Kitazume, H. Kawaji, C. Kai, M. Nakamura, H. Konno, K. Nakano, S. Mottagui-Tabar, P. Arner, A. Chesi, S. Gustincich, F. Persichetti, H. Suzuki, S. M. Grimmond, C. A. Wells, V. Orlando, C. Wahlestedt, E. T. Liu, M. Harbers, J. Kawai, V. B. Bajic, D. A. Hume, and Y. Hayashizaki. 2006. 'Genome-wide analysis of mammalian promoter architecture and evolution', *Nat Genet*, 38: 626-35.

Cartegni, L., S. L. Chew, and A. R. Krainer. 2002. 'Listening to silence and understanding nonsense: exonic mutations that affect splicing', *Nat Rev Genet*, 3: 285-98.

Carter, M. S., J. Doskow, P. Morris, S. Li, R. P. Nhim, S. Sandstedt, and M. F. Wilkinson. 1995. 'A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts in vivo is reversed by protein synthesis inhibitors in vitro', *J Biol Chem*, 270: 28995-9003.

Carter, M. S., S. Li, and M. F. Wilkinson. 1996. 'A splicing-dependent regulatory mechanism that detects translation signals', *EMBO J*, 15: 5965-75.

Cavaloc, Y., C. F. Bourgeois, L. Kister, and J. Stevenin. 1999. 'The splicing factors 9G8 and SRp20 transactivate splicing through different and specific enhancers', *RNA*, 5: 468-83.

Chabot, B., and L. Shkreta. 2016. 'Defective control of pre-messenger RNA splicing in human disease', *J Cell Biol*, 212: 13-27.

Chakrabarti, S., U. Jayachandran, F. Bonneau, F. Fiorini, C. Basquin, S. Domcke, H. Le Hir, and E. Conti. 2011. 'Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2', *Mol Cell*, 41: 693-703.

Chamieh, H., L. Ballut, F. Bonneau, and H. Le Hir. 2008. 'NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity', *Nat Struct Mol Biol*, 15: 85-93.

Chan, W. K., L. Huang, J. P. Gudikote, Y. F. Chang, J. S. Imam, J. A. MacLean, 2nd, and M. F. Wilkinson. 2007. 'An alternative branch of the nonsense-mediated decay pathway', *EMBO J*, 26: 1820-30.

Chang, J. C., and Y. W. Kan. 1979. 'beta 0 thalassemia, a nonsense mutation in man', *Proc Natl Acad Sci U S A*, 76: 2886-9.

Chang, S. T., P. Sova, X. Peng, J. Weiss, G. L. Law, R. E. Palermo, and M. G. Katze. 2011. 'Next-generation sequencing reveals HIV-1-mediated suppression of T cell activation and RNA processing and regulation of noncoding RNA expression in a CD4+ T cell line', *MBio*, 2.

Charenton, C., M. E. Wilkinson, and K. Nagai. 2019. 'Mechanism of 5' splice site transfer for human spliceosome activation', *Science*, 364: 362-67.

Charneau, P., M. Alizon, and F. Clavel. 1992. 'A second origin of DNA plus-strand synthesis is required for optimal human immunodeficiency virus replication', *J Virol*, 66: 2814-20.

Checkley, M. A., B. G. Luttge, and E. O. Freed. 2011. 'HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation', *J Mol Biol*, 410: 582-608.

Chen, J., C. Umunnakwe, D. Q. Sun, O. A. Nikolaitchik, V. K. Pathak, B. Berkhout, A. T. Das, and W. S. Hu. 2020. 'Impact of Nuclear Export Pathway on Cytoplasmic HIV-1 RNA Transport Mechanism and Distribution', *MBio*, 11.

Cheng, H., K. Dufu, C. S. Lee, J. L. Hsu, A. Dias, and R. Reed. 2006. 'Human mRNA export machinery recruited to the 5' end of mRNA', *Cell*, 127: 1389-400.

Cheng, J., P. Belgrader, X. Zhou, and L. E. Maquat. 1994. 'Introns are cis effectors of the nonsense-codon-mediated reduction in nuclear mRNA abundance', *Mol Cell Biol*, 14: 6317-25.

Cheng, J., and L. E. Maquat. 1993. 'Nonsense codons can reduce the abundance of nuclear mRNA without affecting the abundance of pre-mRNA or the half-life of cytoplasmic mRNA', *Mol Cell Biol*, 13: 1892-902.

Cho, H., S. Han, J. Choe, S. G. Park, S. S. Choi, and Y. K. Kim. 2013. 'SMG5-PNRC2 is functionally dominant compared with SMG5-SMG7 in mammalian nonsense-mediated mRNA decay', *Nucleic Acids Res*, 41: 1319-28.

Cho, H., K. M. Kim, and Y. K. Kim. 2009. 'Human proline-rich nuclear receptor coregulatory protein 2 mediates an interaction between mRNA surveillance machinery and decapping complex', *Mol Cell*, 33: 75-86.

Chun, T. W., S. Moir, and A. S. Fauci. 2015. 'HIV reservoirs as obstacles and opportunities for an HIV cure', *Nat Immunol*, 16: 584-9.

Ciuffi, A. 2016. 'The benefits of integration', *Clin Microbiol Infect*, 22: 324-32.

Clark, F., and T. A. Thanaraj. 2002. 'Categorization and characterization of transcript-confirmed constitutively and alternatively spliced introns and exons from human', *Hum Mol Genet*, 11: 451-64.

Cochrane, A. W., C. H. Chen, and C. A. Rosen. 1990. 'Specific interaction of the human immunodeficiency virus Rev protein with a structured region in the env mRNA', *Proc Natl Acad Sci U S A*, 87: 1198-202.

Colombo, M., E. D. Karousis, J. Bourquin, R. Bruggmann, and O. Mühlemann. 2017. 'Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive

redundancy between SMG6- and SMG7-mediated degradation pathways', *RNA*, 23: 189-201.

Company, M., J. Arenas, and J. Abelson. 1991. 'Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes', *Nature*, 349: 487-93.

Contu, L., S. Steiner, V. Thiel, and O. Mühlemann. 2019. 'The Role of Stress Granules and the Nonsense-mediated mRNA Decay Pathway in Antiviral Defence', *Chimia (Aarau)*, 73: 374-79.

Czaplinski, K., Y. Weng, K. W. Hagan, and S. W. Peltz. 1995. 'Purification and characterization of the Upf1 protein: a factor involved in translation and mRNA degradation', *RNA*, 1: 610-23.

D'Agostino, D. M., B. K. Felber, J. E. Harrison, and G. N. Pavlakis. 1992. 'The Rev protein of human immunodeficiency virus type 1 promotes polysomal association and translation of gag/pol and vpu/env mRNAs', *Mol Cell Biol*, 12: 1375-86.

Daar, I. O., and L. E. Maquat. 1988. 'Premature translation termination mediates triosephosphate isomerase mRNA degradation', *Mol Cell Biol*, 8: 802-13.

Daugherty, M. D., I. D'Orso, and A. D. Frankel. 2008. 'A solution to limited genomic capacity: using adaptable binding surfaces to assemble the functional HIV Rev oligomer on RNA', *Mol Cell*, 31: 824-34.

De Conti, L., M. Baralle, and E. Buratti. 2013. 'Exon and intron definition in pre-mRNA splicing', *Wiley Interdiscip Rev RNA*, 4: 49-60.

Deeks, S. G., J. Overbaugh, A. Phillips, and S. Buchbinder. 2015. 'HIV infection', *Nat Rev Dis Primers*, 1: 15035.

Deniaud, A., M. Karuppasamy, T. Bock, S. Masiulis, K. Huard, F. Garzoni, K. Kerschgens, M. W. Hentze, A. E. Kulozik, M. Beck, G. Neu-Yilik, and C. Schaffitzel. 2015. 'A network of SMG-8, SMG-9 and SMG-1 C-terminal insertion domain regulates UPF1 substrate recruitment and phosphorylation', *Nucleic Acids Res*, 43: 7600-11.

Dever, T. E., and R. Green. 2012. 'The elongation, termination, and recycling phases of translation in eukaryotes', *Cold Spring Harb Perspect Biol*, 4: a013706.

Diehl, N. 2016. 'Der Einfluss von HIV-1 induzierten Veränderungen des PI3K/Akt Signalwegs auf das alternative Spleißen in T-Zellen', Dissertation, Heinrich-Heine-Universität.

Doi, A., T. Shiosaka, Y. Takaoka, K. Yanagisawa, and S. Fujita. 1998. 'Molecular cloning of the cDNA encoding A + U-rich element RNA binding factor', *Biochim Biophys Acta*, 1396: 51-6.

Doms, R. W., and J. P. Moore. 2000. 'HIV-1 membrane fusion: targets of opportunity', *J Cell Biol*, 151: F9-14.

Dostie, J., and G. Dreyfuss. 2002. 'Translation is required to remove Y14 from mRNAs in the cytoplasm', *Curr Biol*, 12: 1060-7.

DuBridge, R. B., P. Tang, H. C. Hsia, P. M. Leong, J. H. Miller, and M. P. Calos. 1987. 'Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system', *Mol Cell Biol*, 7: 379-87.

Dull, T., R. Zufferey, M. Kelly, R. J. Mandel, M. Nguyen, D. Trono, and L. Naldini. 1998. 'A third-generation lentivirus vector with a conditional packaging system', *J Virol*, 72: 8463-71.

Durand, S., T. M. Franks, and J. Lykke-Andersen. 2016. 'Hyperphosphorylation amplifies UPF1 activity to resolve stalls in nonsense-mediated mRNA decay', *Nat Commun*, 7: 12434.

Dyhr-Mikkelsen, H., and J. Kjems. 1995. 'Inefficient spliceosome assembly and abnormal branch site selection in splicing of an HIV-1 transcript in vitro', *J Biol Chem*, 270: 24060-6.

Eberle, A. B., S. Lykke-Andersen, O. Mühlemann, and T. H. Jensen. 2009. 'SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells', *Nat Struct Mol Biol*, 16: 49-55.

Eberle, A. B., L. Stalder, H. Mathys, R. Z. Orozco, and O. Mühlemann. 2008. 'Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region', *PLoS Biol*, 6: e92.

Elsner, C., A. Ponnurangam, J. Kazmierski, T. Zillinger, J. Jansen, D. Todt, K. Dohner, S. Xu, A. Ducroux, N. Kriedemann, A. Malassa, P. K. Larsen, G. Hartmann, W. Barchet, E. Steinmann, U. Kalinke, B. Sodeik, and C. Goffinet. 2020. 'Absence of cGAS-mediated type I IFN responses in HIV-1-infected T cells', *Proc Natl Acad Sci U S A*, 117: 19475-86.

Erkelenz, S., W. F. Mueller, M. S. Evans, A. Busch, K. Schöneweis, K. J. Hertel, and H. Schaal. 2013. 'Position-dependent splicing activation and repression by SR and hnRNP proteins rely on common mechanisms', *RNA*, 19: 96-102.

Erkelenz, S., G. Poschmann, S. Theiss, A. Stefanski, F. Hillebrand, M. Otte, K. Stuhler, and H. Schaal. 2013. 'Tra2-mediated recognition of HIV-1 5' splice site D3 as a key factor in the processing of vpr mRNA', *J Virol*, 87: 2721-34.

Erkelenz, S., S. Theiss, M. Otte, M. Widera, J. O. Peter, and H. Schaal. 2014. 'Genomic HEXploring allows landscaping of novel potential splicing regulatory elements', *Nucleic Acids Res*, 42: 10681-97.

Exline, C. M., Z. Feng, and C. M. Stoltzfus. 2008. 'Negative and positive mRNA splicing elements act competitively to regulate human immunodeficiency virus type 1 vif gene expression', *J Virol*, 82: 3921-31.

Fairbrother, W. G., R. F. Yeh, P. A. Sharp, and C. B. Burge. 2002. 'Predictive identification of exonic splicing enhancers in human genes', *Science*, 297: 1007-13.

Fassati, A., and S. P. Goff. 2001. 'Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1', *J Virol*, 75: 3626-35.

Felber, B. K., M. Hadzopoulou-Cladaras, C. Cladaras, T. Copeland, and G. N. Pavlakis. 1989. 'rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA', *Proc Natl Acad Sci U S A*, 86: 1495-9.

Fica, S. M., and K. Nagai. 2017. 'Cryo-electron microscopy snapshots of the spliceosome: structural insights into a dynamic ribonucleoprotein machine', *Nat Struct Mol Biol*, 24: 791-99.

Fica, S. M., C. Oubridge, W. P. Galej, M. E. Wilkinson, X. C. Bai, A. J. Newman, and K. Nagai. 2017. 'Structure of a spliceosome remodelled for exon ligation', *Nature*, 542: 377-80.

Fica, S. M., N. Tuttle, T. Novak, N. S. Li, J. Lu, P. Koodathingal, Q. Dai, J. P. Staley, and J. A. Piccirilli. 2013. 'RNA catalyses nuclear pre-mRNA splicing', *Nature*, 503: 229-34.

Fiorini, F., M. Boudvillain, and H. Le Hir. 2013. 'Tight intramolecular regulation of the human Upf1 helicase by its N- and C-terminal domains', *Nucleic Acids Res*, 41: 2404-15.

Fiorini, F., J. P. Robin, J. Kanaan, M. Borowiak, V. Croquette, H. Le Hir, P. Jalinot, and V. Mocquet. 2018. 'HTLV-1 Tax plugs and freezes UPF1 helicase leading to nonsense-mediated mRNA decay inhibition', *Nat Commun*, 9: 431.

Fischer, U., J. Huber, W. C. Boelens, I. W. Mattaj, and R. Lührmann. 1995. 'The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs', *Cell*, 82: 475-83.

Fischer, U., V. W. Pollard, R. Lührmann, M. Teufel, M. W. Michael, G. Dreyfuss, and M. H. Malim. 1999. 'Rev-mediated nuclear export of RNA is dominant over nuclear retention and is coupled to the Ran-GTPase cycle', *Nucleic Acids Res*, 27: 4128-34.

Forsythe, S. S., W. McGreevey, A. Whiteside, M. Shah, J. Cohen, R. Hecht, L. A. Bollinger, and A. Kinghorn. 2019. 'Twenty Years Of Antiretroviral Therapy For People Living With HIV: Global Costs, Health Achievements, Economic Benefits', *Health Aff (Millwood)*, 38: 1163-72.

Fox-Walsh, K. L., and K. J. Hertel. 2009. 'Splice-site pairing is an intrinsically high fidelity process', *Proc Natl Acad Sci U S A*, 106: 1766-71.

Freed, E. O. 2015. 'HIV-1 assembly, release and maturation', *Nat Rev Microbiol*, 13: 484-96.

Freund, M., C. Asang, S. Kammler, C. Konermann, J. Krummheuer, M. Hipp, I. Meyer, W. Gierling, S. Theiss, T. Preuss, D. Schindler, J. Kjems, and H. Schaal. 2003. 'A novel approach to describe a U1 snRNA binding site', *Nucleic Acids Res*, 31: 6963-75.

Frolova, L., X. Le Goff, H. H. Rasmussen, S. Cheperegin, G. Drugeon, M. Kress, I. Arman, A. L. Haenni, J. E. Celis, M. Philippe, and et al. 1994. 'A highly conserved eukaryotic protein family possessing properties of polypeptide chain release factor', *Nature*, 372: 701-3.

Galej, W. P., M. E. Wilkinson, S. M. Fica, C. Oubridge, A. J. Newman, and K. Nagai. 2016. 'Cryo-EM structure of the spliceosome immediately after branching', *Nature*, 537: 197-201.

Gallo, R. C., P. S. Sarin, E. P. Gelmann, M. Robert-Guroff, E. Richardson, V. S. Kalyanaraman, D. Mann, G. D. Sidhu, R. E. Stahl, S. Zolla-Pazner, J. Leibowitch, and M. Popovic. 1983. 'Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS)', *Science*, 220: 865-7.

Garcia-Moreno, J. F., and L. Romao. 2020. 'Perspective in Alternative Splicing Coupled to Nonsense-Mediated mRNA Decay', *Int J Mol Sci*, 21.

Garcia, D., S. Garcia, and O. Voinnet. 2014. 'Nonsense-mediated decay serves as a general viral restriction mechanism in plants', *Cell Host Microbe*, 16: 391-402.

Ge, Z., B. L. Quek, K. L. Beemon, and J. R. Hogg. 2016. 'Polypyrimidine tract binding protein 1 protects mRNAs from recognition by the nonsense-mediated mRNA decay pathway', *Elife*, 5.

Gehring, N. H., J. B. Kunz, G. Neu-Yilik, S. Breit, M. H. Viegas, M. W. Hentze, and A. E. Kulozik. 2005. 'Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements', *Mol Cell*, 20: 65-75.

Gerbracht, J. V., V. Boehm, and N. H. Gehring. 2017. 'Plasmid transfection influences the readout of nonsense-mediated mRNA decay reporter assays in human cells', *Sci Rep*, 7: 10616.

Geuens, T., D. Bouhy, and V. Timmerman. 2016. 'The hnRNP family: insights into their role in health and disease', *Hum Genet*, 135: 851-67.

Ghazawi, I., S. J. Cutler, P. Low, A. S. Mellick, and S. J. Ralph. 2005. 'Inhibitory effects associated with use of modified Photinus pyralis and Renilla reniformis luciferase vectors in dual reporter assays and implications for analysis of ISGs', *J Interferon Cytokine Res*, 25: 92-102.

Girnary, R., L. King, L. Robinson, R. Elston, and I. Brierley. 2007. 'Structure-function analysis of the ribosomal frameshifting signal of two human immunodeficiency virus type 1 isolates with increased resistance to viral protease inhibitors', *J Gen Virol*, 88: 226-35.

Gloggnitzer, J., S. Akimcheva, A. Srinivasan, B. Kusenda, N. Riehs, H. Stampfl, J. Bautor, B. Dekrout, C. Jonak, J. M. Jimenez-Gomez, J. E. Parker, and K. Riha. 2014. 'Nonsense-mediated mRNA decay modulates immune receptor levels to regulate plant antibacterial defense', *Cell Host Microbe*, 16: 376-90.

Gordon, D. E., G. M. Jang, M. Bouhaddou, J. Xu, K. Obernier, K. M. White, M. J. O'Meara, V. V. Rezelj, J. Z. Guo, D. L. Swaney, T. A. Tummino, R. Huttenhain, R. M. Kaake, A. L. Richards, B. Tutuncuoglu, H. Foussard, J. Batra, K. Haas, M. Modak, M. Kim, P. Haas, B. J. Polacco, H. Braberg, J. M. Fabius, M. Eckhardt, M. Soucheray, M. J. Bennett, M. Cakir, M. J. McGregor, Q. Li, B. Meyer, F. Roesch, T. Vallet, A. Mac Kain, L. Miorin, E. Moreno, Z. Z. C. Naing, Y. Zhou, S. Peng, Y. Shi, Z. Zhang, W. Shen, I. T. Kirby, J. E. Melnyk, J. S. Chorba, K. Lou, S. A. Dai, I. Barrio-Hernandez, D. Memon, C. Hernandez-Armenta, J. Lyu, C. J. P. Mathy, T. Perica, K. B. Pilla, S. J. Ganesan, D. J. Saltzberg, R. Rakesh, X. Liu, S. B. Rosenthal, L. Calviello, S. Venkataramanan, J. Liboy-Lugo, Y. Lin, X. P. Huang, Y. Liu, S. A. Wankowicz,

M. Bohn, M. Safari, F. S. Ugur, C. Koh, N. S. Savar, Q. D. Tran, D. Shengjuler, S. J. Fletcher, M. C. O'Neal, Y. Cai, J. C. J. Chang, D. J. Broadhurst, S. Klippsten, P. P. Sharp, N. A. Wenzell, D. Kuzuoglu-Ozturk, H. Y. Wang, R. Trenker, J. M. Young, D. A. Cavero, J. Hiatt, T. L. Roth, U. Rathore, A. Subramanian, J. Noack, M. Hubert, R. M. Stroud, A. D. Frankel, O. S. Rosenberg, K. A. Verba, D. A. Agard, M. Ott, M. Emerman, N. Jura, M. von Zastrow, E. Verdin, A. Ashworth, O. Schwartz, C. d'Enfert, S. Mukherjee, M. Jacobson, H. S. Malik, D. G. Fujimori, T. Ideker, C. S. Craik, S. N. Floor, J. S. Fraser, J. D. Gross, A. Sali, B. L. Roth, D. Ruggero, J. Taunton, T. Kortemme, P. Beltrao, M. Vignuzzi, A. Garcia-Sastre, K. M. Shokat, B. K. Shoichet, and N. J. Krogan. 2020. 'A SARS-CoV-2 protein interaction map reveals targets for drug repurposing', *Nature*, 583: 459-68.

Görlich, D., and U. Kutay. 1999. 'Transport between the cell nucleus and the cytoplasm', *Annu Rev Cell Dev Biol*, 15: 607-60.

Gottlieb, M. S., R. Schroff, H. M. Schanker, J. D. Weisman, P. T. Fan, R. A. Wolf, and A. Saxon. 1981. 'Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency', *N Engl J Med*, 305: 1425-31.

Grabowski, P. J., R. A. Padgett, and P. A. Sharp. 1984. 'Messenger RNA splicing in vitro: an excised intervening sequence and a potential intermediate', *Cell*, 37: 415-27.

Graham, F. L., J. Smiley, W. C. Russell, and R. Nairn. 1977. 'Characteristics of a human cell line transformed by DNA from human adenovirus type 5', *J Gen Virol*, 36: 59-74.

Greatorex, J., J. Gallego, G. Varani, and A. Lever. 2002. 'Structure and stability of wild-type and mutant RNA internal loops from the SL-1 domain of the HIV-1 packaging signal', *J Mol Biol*, 322: 543-57.

Groom, H. C. T., E. C. Anderson, J. A. Dangerfield, and A. M. L. Lever. 2009. 'Rev regulates translation of human immunodeficiency virus type 1 RNAs', *J Gen Virol*, 90: 1141-47.

Haberle, V., and A. Stark. 2018. 'Eukaryotic core promoters and the functional basis of transcription initiation', *Nat Rev Mol Cell Biol*, 19: 621-37.

Hallenberger, S., V. Bosch, H. Angliker, E. Shaw, H. D. Klenk, and W. Garten. 1992. 'Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160', *Nature*, 360: 358-61.

Hampsey, M. 1998. 'Molecular genetics of the RNA polymerase II general transcriptional machinery', *Microbiol Mol Biol Rev*, 62: 465-503.

Hang, J., R. Wan, C. Yan, and Y. Shi. 2015. 'Structural basis of pre-mRNA splicing', *Science*, 349: 1191-8.

Hanson, G., N. Alhusaini, N. Morris, T. Sweet, and J. Coller. 2018. 'Translation elongation and mRNA stability are coupled through the ribosomal A-site', *RNA*, 24: 1377-89.

Haque, N., R. Ouda, C. Chen, K. Ozato, and J. R. Hogg. 2018. 'ZFR coordinates crosstalk between RNA decay and transcription in innate immunity', *Nat Commun*, 9: 1145.

Haselbach, D., I. Komarov, D. E. Agafonov, K. Hartmuth, B. Graf, O. Dybkov, H. Urlaub, B. Kastner, R. Lührmann, and H. Stark. 2018. 'Structure and Conformational Dynamics of the Human Spliceosomal B(act) Complex', *Cell*, 172: 454-64 e11.

Hashimoto, C., and J. A. Steitz. 1984. 'U4 and U6 RNAs coexist in a single small nuclear ribonucleoprotein particle', *Nucleic Acids Res*, 12: 3283-93.

Henderson, B. R., and P. Percipalle. 1997. 'Interactions between HIV Rev and nuclear import and export factors: the Rev nuclear localisation signal mediates specific binding to human importin-beta', *J Mol Biol*, 274: 693-707.

Henne, W. M., N. J. Buchkovich, and S. D. Emr. 2011. 'The ESCRT pathway', *Dev Cell*, 21: 77-91.

Herz, H. M., D. Hu, and A. Shilatifard. 2014. 'Enhancer malfunction in cancer', *Mol Cell*, 53: 859-66.

Ho, S. C., Mariati, J. H. Yeo, S. G. Fang, and Y. Yang. 2015. 'Impact of using different promoters and matrix attachment regions on recombinant protein expression level and stability in stably transfected CHO cells', *Mol Biotechnol*, 57: 138-44.

Hoek, T. A., D. Khuperkar, R. G. H. Lindeboom, S. Sonneveld, B. M. P. Verhagen, S. Boersma, M. Vermeulen, and M. E. Tanenbaum. 2019. 'Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay', *Mol Cell*, 75: 324-39 e11.

Hogg, J. R., and S. P. Goff. 2010. 'Upf1 senses 3'UTR length to potentiate mRNA decay', *Cell*, 143: 379-89.

Hong, X., D. G. Scofield, and M. Lynch. 2006. 'Intron size, abundance, and distribution within untranslated regions of genes', *Mol Biol Evol*, 23: 2392-404.

Hoshino, S., M. Imai, T. Kobayashi, N. Uchida, and T. Katada. 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', *J Biol Chem*, 274: 16677-80.

Hosoda, N., Y. K. Kim, F. Lejeune, and L. E. Maquat. 2005. 'CBP80 promotes interaction of Upf1 with Upf2 during nonsense-mediated mRNA decay in mammalian cells', *Nat Struct Mol Biol*, 12: 893-901.

Hu, W. S., and S. H. Hughes. 2012. 'HIV-1 reverse transcription', *Cold Spring Harb Perspect Med*, 2.

Hug, N., and J. F. Cáceres. 2014. 'The RNA helicase DHX34 activates NMD by promoting a transition from the surveillance to the decay-inducing complex', *Cell Rep*, 8: 1845-56.

Huntzinger, E., I. Kashima, M. Fauser, J. Sauliere, and E. Izaurralde. 2008. 'SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan', *RNA*, 14: 2609-17.

Hurt, J. A., A. D. Robertson, and C. B. Burge. 2013. 'Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay', *Genome Res*, 23: 1636-50.

Hwang, J., H. Sato, Y. Tang, D. Matsuda, and L. E. Maquat. 2010. 'UPF1 association with the cap-binding protein, CBP80, promotes nonsense-mediated mRNA decay at two distinct steps', *Mol Cell*, 39: 396-409.

Imamachi, N., K. A. Salam, Y. Suzuki, and N. Akimitsu. 2017. 'A GC-rich sequence feature in the 3' UTR directs UPF1-dependent mRNA decay in mammalian cells', *Genome Res*, 27: 407-18.

Ingolia, N. T., S. Ghaemmaghami, J. R. Newman, and J. S. Weissman. 2009. 'Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling', *Science*, 324: 218-23.

International Human Genome Sequencing, Consortium. 2004. 'Finishing the euchromatic sequence of the human genome', *Nature*, 431: 931-45.

Isel, C., C. Ehresmann, G. Keith, B. Ehresmann, and R. Marquet. 1995. 'Initiation of reverse transcription of HIV-1: secondary structure of the HIV-1 RNA/tRNA(3Lys) (template/primer)', *J Mol Biol*, 247: 236-50.

Isel, C., C. Ehresmann, and R. Marquet. 2010. 'Initiation of HIV Reverse Transcription', *Viruses*, 2: 213-43.

Ishigaki, Y., X. Li, G. Serin, and L. E. Maquat. 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*, 106: 607-17.

Isken, O., Y. K. Kim, N. Hosoda, G. L. Mayeur, J. W. Hershey, and L. E. Maquat. 2008. 'Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay', *Cell*, 133: 314-27.

Ivanov, A., T. Mikhailova, B. Eliseev, L. Yeramala, E. Sokolova, D. Susorov, A. Shuvalov, C. Schaffitzel, and E. Alkalaeva. 2016. 'PABP enhances release factor recruitment and stop codon recognition during translation termination', *Nucleic Acids Res*, 44: 7766-76.

Ivanov, P. V., N. H. Gehring, J. B. Kunz, M. W. Hentze, and A. E. Kulozik. 2008. 'Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways', *EMBO J*, 27: 736-47.

Jacks, T., M. D. Power, F. R. Masiarz, P. A. Luciw, P. J. Barr, and H. E. Varmus. 1988. 'Characterization of ribosomal frameshifting in HIV-1 gag-pol expression', *Nature*, 331: 280-3.

Jacquenet, S., D. Decimo, D. Muriaux, and J. L. Darlix. 2005. 'Dual effect of the SR proteins ASF/SF2, SC35 and 9G8 on HIV-1 RNA splicing and virion production', *Retrovirology*, 2: 33.

Jäger, S., P. Cimermancic, N. Gulbahce, J. R. Johnson, K. E. McGovern, S. C. Clarke, M. Shales, G. Mercenne, L. Pache, K. Li, H. Hernandez, G. M. Jang, S. L. Roth, E. Akiva, J. Marlett, M. Stephens, I. D'Orso, J. Fernandes, M. Fahey, C. Mahon, A. J. O'Donoghue, A. Todorovic, J. H. Morris, D. A. Maltby, T. Alber, G. Cagney, F. D. Bushman, J. A. Young, S. K. Chanda, W. I. Sundquist, T. Kortemme, R. D. Hernandez, C. S. Craik, A. Burlingame, A. Sali, A. D. Frankel, and N. J. Krogan. 2011. 'Global landscape of HIV-human protein complexes', *Nature*, 481: 365-70.

Jakobsen, M. R., R. O. Bak, A. Andersen, R. K. Berg, S. B. Jensen, J. Tengchuan, A. Laustsen, K. Hansen, L. Ostergaard, K. A. Fitzgerald, T. S. Xiao, J. G. Mikkelsen, T. H. Mogensen, and S. R. Paludan. 2013. 'IFI16 senses DNA forms of the lentiviral replication cycle and controls HIV-1 replication', *Proc Natl Acad Sci U S A*, 110: E4571-80.

Kadonaga, J. T. 2012. 'Perspectives on the RNA polymerase II core promoter', *Wiley Interdiscip Rev Dev Biol*, 1: 40-51.

Kammler, S., C. Leurs, M. Freund, J. Krummheuer, K. Seidel, T. O. Tange, M. K. Lund, J. Kjems, A. Scheid, and H. Schaal. 2001. 'The sequence complementarity between HIV-1 5' splice site SD4 and U1 snRNA determines the steady-state level of an unstable env pre-mRNA', *RNA*, 7: 421-34.

Kammler, S., M. Otte, I. Hauber, J. Kjems, J. Hauber, and H. Schaal. 2006. 'The strength of the HIV-1 3' splice sites affects Rev function', *Retrovirology*, 3: 89.

Kao, S. Y., A. F. Calman, P. A. Luciw, and B. M. Peterlin. 1987. 'Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product', *Nature*, 330: 489-93.

Karlsson, S., T. Papayannopoulou, S. G. Schweiger, G. Stamatoyannopoulos, and A. W. Nienhuis. 1987. 'Retroviral-mediated transfer of genomic globin genes leads to regulated production of RNA and protein', *Proc Natl Acad Sci U S A*, 84: 2411-5.

Karn, J., and C. M. Stoltzfus. 2012. 'Transcriptional and posttranscriptional regulation of HIV-1 gene expression', *Cold Spring Harb Perspect Med*, 2: a006916.

Karousis, E. D., S. Nasif, and O. Mühlemann. 2016. 'Nonsense-mediated mRNA decay: novel mechanistic insights and biological impact', *Wiley Interdiscip Rev RNA*.

Kashima, I., A. Yamashita, N. Izumi, N. Kataoka, R. Morishita, S. Hoshino, M. Ohno, G. Dreyfuss, and S. Ohno. 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 Dev*, 20: 355-67.

Katahira, J. 2012. 'mRNA export and the TREX complex', *Biochim Biophys Acta*, 1819: 507-13.

Katahira, J., K. Strässer, A. Podtelejnikov, M. Mann, J. U. Jung, and E. Hurt. 1999. 'The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human', *EMBO J*, 18: 2593-609.

Keeling, K. M., D. Wang, Y. Dai, S. Murugesan, B. Chenna, J. Clark, V. Belakhov, J. Kandasamy, S. E. Velu, T. Baasov, and D. M. Bedwell. 2013. 'Attenuation of nonsense-mediated mRNA decay enhances in vivo nonsense suppression', *PLoS One*, 8: e60478.

Kemmerer, K., S. Fischer, and J. E. Weigand. 2018. 'Auto- and cross-regulation of the hnRNPs D and DL', *RNA*, 24: 324-31.

Kervestin, S., C. Li, R. Buckingham, and A. Jacobson. 2012. 'Testing the faux-UTR model for NMD: analysis of Upf1p and Pab1p competition for binding to eRF3/Sup35p', *Biochimie*, 94: 1560-71.

Khiytani, D. K., and N. J. Dimmock. 2002. 'Characterization of a human immunodeficiency virus type 1 pre-integration complex in which the majority of the cDNA is resistant to DNase I digestion', *J Gen Virol*, 83: 2523-32.

Kim, D., J. Y. Lee, J. S. Yang, J. W. Kim, V. N. Kim, and H. Chang. 2020. 'The Architecture of SARS-CoV-2 Transcriptome', *Cell*, 181: 914-21 e10.

Kim, S. Y., R. Byrn, J. Groopman, and D. Baltimore. 1989. 'Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression', *J Virol*, 63: 3708-13.

Kishor, A., Z. Ge, and J. R. Hogg. 2019. 'hnRNP L-dependent protection of normal mRNAs from NMD subverts quality control in B cell lymphoma', *EMBO J*, 38.

Kistler, A. L., and C. Guthrie. 2001. 'Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for sub2, an essential spliceosomal ATPase', *Genes Dev*, 15: 42-9.

Klaver, B., and B. Berkhout. 1994. 'Comparison of 5' and 3' long terminal repeat promoter function in human immunodeficiency virus', *J Virol*, 68: 3830-40.

Köhler, A., and E. Hurt. 2007. 'Exporting RNA from the nucleus to the cytoplasm', *Nat Rev Mol Cell Biol*, 8: 761-73.

Kolitz, S. E., J. E. Takacs, and J. R. Lorsch. 2009. 'Kinetic and thermodynamic analysis of the role of start codon/anticodon base pairing during eukaryotic translation initiation', *RNA*, 15: 138-52.

Königs, V., C. de Oliveira Freitas Machado, B. Arnold, N. Blümel, A. Solovyeva, S. Löbbert, M. Schafranek, I. Ruiz De Los Mozos, I. Wittig, F. McNicoll, M. H. Schulz, and M. Müller-McNicoll. 2020. 'SRSF7 maintains its homeostasis through the expression of Split-ORFs and nuclear body assembly', *Nat Struct Mol Biol*, 27: 260-73.

Kozak, M. 1979. 'Inability of circular mRNA to attach to eukaryotic ribosomes', *Nature*, 280: 82-5.

———. 1981. 'Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes', *Nucleic Acids Res*, 9: 5233-52.

———. 1984a. 'Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs', *Nucleic Acids Res*, 12: 857-72.

———. 1984b. 'Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin in vivo', *Nature*, 308: 241-6.

———. 1986. 'Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes', *Cell*, 44: 283-92.

Krummheuer, J., A. T. Johnson, I. Hauber, S. Kammler, J. L. Anderson, J. Hauber, D. F. Purcell, and H. Schaal. 2007. 'A minimal uORF within the HIV-1 vpu leader allows efficient translation initiation at the downstream env AUG', *Virology*, 363: 261-71.

Krummheuer, J., C. Lenz, S. Kammler, A. Scheid, and H. Schaal. 2001. 'Influence of the small leader exons 2 and 3 on human immunodeficiency virus type 1 gene expression', *Virology*, 286: 276-89.

Kugler, W., J. Enssle, M. W. Hentze, and A. E. Kulozik. 1995. 'Nuclear degradation of nonsense mutated beta-globin mRNA: a post-transcriptional mechanism to protect heterozygotes from severe clinical manifestations of beta-thalassemia?', *Nucleic Acids Res*, 23: 413-8.

Kumar, M., B. Keller, N. Makalou, and R. E. Sutton. 2001. 'Systematic determination of the packaging limit of lentiviral vectors', *Hum Gene Ther*, 12: 1893-905.

Kurosaki, T., W. Li, M. Hoque, M. W. Popp, D. N. Ermolenko, B. Tian, and L. E. Maquat. 2014. 'A post-translational regulatory switch on UPF1 controls targeted mRNA degradation', *Genes Dev*, 28: 1900-16.

Kurosaki, T., and L. E. Maquat. 2013. 'Rules that govern UPF1 binding to mRNA 3' UTRs', *Proc Natl Acad Sci U S A*, 110: 3357-62.

Kurosaki, T., K. Miyoshi, J. R. Myers, and L. E. Maquat. 2018. 'NMD-degradome sequencing reveals ribosome-bound intermediates with 3'-end non-templated nucleotides', *Nat Struct Mol Biol*, 25: 940-50.

Kurosaki, T., M. W. Popp, and L. E. Maquat. 2019. 'Quality and quantity control of gene expression by nonsense-mediated mRNA decay', *Nat Rev Mol Cell Biol*.

Landry, J. J., P. T. Pyl, T. Rausch, T. Zichner, M. M. Tekkedil, A. M. Stutz, A. Jauch, R. S. Aiyar, G. Pau, N. Delhomme, J. Gagneur, J. O. Korbel, W. Huber, and L. M. Steinmetz. 2013. 'The genomic and transcriptomic landscape of a HeLa cell line', *G3 (Bethesda)*, 3: 1213-24.

Lareau, L. F., M. Inada, R. E. Green, J. C. Wengrod, and S. E. Brenner. 2007. 'Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements', *Nature*, 446: 926-9.

Lata, S., A. Ali, V. Sood, R. Raja, and A. C. Banerjea. 2015. 'HIV-1 Rev downregulates Tat expression and viral replication via modulation of NAD(P)H:quinine oxidoreductase 1 (NQO1)', *Nat Commun*, 6: 7244.

Lawrence, J. B., A. W. Cochrane, C. V. Johnson, A. Perkins, and C. A. Rosen. 1991. 'The HIV-1 Rev protein: a model system for coupled RNA transport and translation', *New Biol*, 3: 1220-32.

Le-Trilling, V. T. K., K. Wohlgemuth, M. U. Ruckborn, A. Jagnjic, F. Maassen, L. Timmer, B. Katschinski, and M. Trilling. 2018. 'STAT2-Dependent Immune Responses Ensure Host Survival despite the Presence of a Potent Viral Antagonist', *J Virol*, 92.

Le Hir, H., D. Gatfield, E. Izaurralde, and M. J. Moore. 2001. 'The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay', *EMBO J*, 20: 4987-97.

Le Hir, H., E. Izaurralde, L. E. Maquat, and M. J. Moore. 2000. 'The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions', *EMBO J*, 19: 6860-9.

LeBlanc, J. J., and K. L. Beemon. 2004. 'Unspliced Rous sarcoma virus genomic RNAs are translated and subjected to nonsense-mediated mRNA decay before packaging', *J Virol*, 78: 5139-46.

Lee, S. R., G. A. Pratt, F. J. Martinez, G. W. Yeo, and J. Lykke-Andersen. 2015. 'Target Discrimination in Nonsense-Mediated mRNA Decay Requires Upf1 ATPase Activity', *Mol Cell*, 59: 413-25.

Leeds, P., S. W. Peltz, A. Jacobson, and M. R. Culbertson. 1991. 'The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon', *Genes Dev*, 5: 2303-14.

Lejeune, F., X. Li, and L. E. Maquat. 2003. 'Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities', *Mol Cell*, 12: 675-87.

Lenhard, B., A. Sandelin, and P. Carninci. 2012. 'Metazoan promoters: emerging characteristics and insights into transcriptional regulation', *Nat Rev Genet*, 13: 233-45.

Lerner, M. R., and J. A. Steitz. 1979. 'Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus', *Proc Natl Acad Sci U S A*, 76: 5495-9.

Leung, A. K., K. Nagai, and J. Li. 2011. 'Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis', *Nature*, 473: 536-9.

Lever, A., H. Gottlinger, W. Haseltine, and J. Sodroski. 1989. 'Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions', *J Virol*, 63: 4085-7.

Levin, A., Z. Hayouka, A. Friedler, and A. Loyter. 2010. 'Over-expression of the HIV-1 Rev promotes death of nondividing eukaryotic cells', *Virus Genes*, 40: 341-6.

Levine, M., and R. Tjian. 2003. 'Transcription regulation and animal diversity', *Nature*, 424: 147-51.

Levy, J. A., I. Scott, and C. Mackewicz. 2003. 'Protection from HIV/AIDS: the importance of innate immunity', *Clin Immunol*, 108: 167-74.

Lewis, B. P., R. E. Green, and S. E. Brenner. 2003. 'Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans', *Proc Natl Acad Sci U S A*, 100: 189-92.

Li, K. J., and H. Garoff. 1998. 'Packaging of intron-containing genes into retrovirus vectors by alphavirus vectors', *Proc Natl Acad Sci U S A*, 95: 3650-4.

Li, M., M. Mizuuchi, T. R. Burke, Jr., and R. Craigie. 2006. 'Retroviral DNA integration: reaction pathway and critical intermediates', *EMBO J*, 25: 1295-304.

Liang, W. W., and S. C. Cheng. 2015. 'A novel mechanism for Prp5 function in prespliceosome formation and proofreading the branch site sequence', *Genes Dev*, 29: 81-93.

Linde, L., S. Boelz, M. Nissim-Rafinia, Y. S. Oren, M. Wilschanski, Y. Yaacov, D. Virgilis, G. Neu-Yilik, A. E. Kulozik, E. Kerem, and B. Kerem. 2007. 'Nonsense-mediated mRNA decay affects nonsense transcript levels and governs response of cystic fibrosis patients to gentamicin', *J Clin Invest*, 117: 683-92.

Loh, B., S. Jonas, and E. Izaurralde. 2013. 'The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2', *Genes Dev*, 27: 2125-38.

Lopez, P. J., and B. Seraphin. 1999. 'Genomic-scale quantitative analysis of yeast pre-mRNA splicing: implications for splice-site recognition', *RNA*, 5: 1135-7.

Losson, R., and F. Lacroute. 1979. 'Interference of nonsense mutations with eukaryotic messenger RNA stability', *Proc Natl Acad Sci U S A*, 76: 5134-7.

Louder, R. K., Y. He, J. R. Lopez-Blanco, J. Fang, P. Chacon, and E. Nogales. 2016. 'Structure of promoter-bound TFIID and model of human pre-initiation complex assembly', *Nature*, 531: 604-9.

Lu, K., X. Heng, L. Garyu, S. Monti, E. L. Garcia, S. Kharytonchyk, B. Dorjsuren, G. Kulandaivel, S. Jones, A. Hiremath, S. S. Divakaruni, C. LaCotti, S. Barton, D. Tummillo, A. Hosic, K. Edme, S. Albrecht, A. Telesnitsky, and M. F. Summers. 2011. 'NMR detection of structures in the HIV-1 5'-leader RNA that regulate genome packaging', *Science*, 334: 242-5.

Lusso, P., F. Cocchi, C. Balotta, P. D. Markham, A. Louie, P. Farci, R. Pal, R. C. Gallo, and M. S. Reitz, Jr. 1995. 'Growth of macrophage-tropic and primary human immunodeficiency virus type 1 (HIV-1) isolates in a unique CD4+ T-cell clone (PM1): failure to downregulate CD4 and to interfere with cell-line-tropic HIV-1', *J Virol*, 69: 3712-20.

Lykke-Andersen, J. 2002. 'Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay', *Mol Cell Biol*, 22: 8114-21.

Lykke-Andersen, J., M. D. Shu, and J. A. Steitz. 2000. 'Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon', *Cell*, 103: 1121-31.

Lykke-Andersen, S., Y. Chen, B. R. Ardal, B. Lilje, J. Waage, A. Sandelin, and T. H. Jensen. 2014. 'Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes', *Genes Dev*, 28: 2498-517.

Lykke-Andersen, S., and T. H. Jensen. 2015. 'Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes', *Nat Rev Mol Cell Biol*, 16: 665-77.

Mabin, J. W., L. A. Woodward, R. D. Patton, Z. Yi, M. Jia, V. H. Wysocki, R. Bundschuh, and G. Singh. 2018. 'The Exon Junction Complex Undergoes a Compositional Switch that Alters mRNP Structure and Nonsense-Mediated mRNA Decay Activity', *Cell Rep*, 25: 2431-46 e7.

Maddon, P. J., A. G. Dalgleish, J. S. McDougal, P. R. Clapham, R. A. Weiss, and R. Axel. 1986. 'The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain', *Cell*, 47: 333-48.

Maderazo, A. B., F. He, D. A. Mangus, and A. Jacobson. 2000. 'Upf1p control of nonsense mRNA translation is regulated by Nmd2p and Upf3p', *Mol Cell Biol*, 20: 4591-603.

Maldarelli, F., C. Xiang, G. Chamoun, and S. L. Zeichner. 1998. 'The expression of the essential nuclear splicing factor SC35 is altered by human immunodeficiency virus infection', *Virus Res*, 53: 39-51.

Malecki, M., S. C. Viegas, T. Carneiro, P. Golik, C. Dressaire, M. G. Ferreira, and C. M. Arraiano. 2013. 'The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway', *EMBO J*, 32: 1842-54.

Malim, M. H., and B. R. Cullen. 1993. 'Rev and the fate of pre-mRNA in the nucleus: implications for the regulation of RNA processing in eukaryotes', *Mol Cell Biol*, 13: 6180-9.

Malim, M. H., J. Hauber, S. Y. Le, J. V. Maizel, and B. R. Cullen. 1989. 'The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA', *Nature*, 338: 254-7.

Mandal, D., C. M. Exline, Z. Feng, and C. M. Stoltzfus. 2009. 'Regulation of Vif mRNA splicing by human immunodeficiency virus type 1 requires 5' splice site D2 and an exonic splicing enhancer to counteract cellular restriction factor APOBEC3G', *J Virol*, 83: 6067-78.

Manley, J. L., and A. R. Krainer. 2010. 'A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins)', *Genes Dev*, 24: 1073-4.

Marini, B., A. Kertesz-Farkas, H. Ali, B. Lucic, K. Lisek, L. Manganaro, S. Pongor, R. Luzzati, A. Recchia, F. Mavilio, M. Giacca, and M. Lusic. 2015. 'Nuclear architecture dictates HIV-1 integration site selection', *Nature*, 521: 227-31.

Masuda, S., R. Das, H. Cheng, E. Hurt, N. Dorman, and R. Reed. 2005. 'Recruitment of the human TREX complex to mRNA during splicing', *Genes Dev*, 19: 1512-7.

Matera, A. G., and Z. Wang. 2014. 'A day in the life of the spliceosome', *Nat Rev Mol Cell Biol*, 15: 108-21.

McDonald, D., M. A. Vodicka, G. Lucero, T. M. Svitkina, G. G. Borisy, M. Emerman, and T. J. Hope. 2002. 'Visualization of the intracellular behavior of HIV in living cells', *J Cell Biol*, 159: 441-52.

Medghalchi, S. M., P. A. Frischmeyer, J. T. Mendell, A. G. Kelly, A. M. Lawler, and H. C. Dietz. 2001. 'Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability', *Hum Mol Genet*, 10: 99-105.

Meinel, D. M., C. Burkert-Kautzsch, A. Kieser, E. O'Duibhir, M. Siebert, A. Mayer, P. Cramer, J. Soding, F. C. Holstege, and K. Strässer. 2013. 'Recruitment of TREX to the transcription

machinery by its direct binding to the phospho-CTD of RNA polymerase II', *PLoS Genet*, 9: e1003914.

Mendell, J. T., and H. C. Dietz. 2001. 'When the message goes awry: disease-producing mutations that influence mRNA content and performance', *Cell*, 107: 411-4.

Mendell, J. T., N. A. Sharifi, J. L. Meyers, F. Martinez-Murillo, and H. C. Dietz. 2004. 'Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise', *Nat Genet*, 36: 1073-8.

Miller, M. D., C. M. Farnet, and F. D. Bushman. 1997. 'Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition', *J Virol*, 71: 5382-90.

Milone, M. C., and U. O'Doherty. 2018. 'Clinical use of lentiviral vectors', *Leukemia*, 32: 1529-41.

Mishima, Y., and Y. Tomari. 2016. 'Codon Usage and 3' UTR Length Determine Maternal mRNA Stability in Zebrafish', *Mol Cell*, 61: 874-85.

Mitchell, R. S., B. F. Beitzel, A. R. Schroder, P. Shinn, H. Chen, C. C. Berry, J. R. Ecker, and F. D. Bushman. 2004. 'Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences', *PLoS Biol*, 2: E234.

Mocquet, V., J. Neusiedler, F. Rende, D. Cluet, J. P. Robin, J. M. Terme, M. Duc Dodon, J. Wittmann, C. Morris, H. Le Hir, V. Ciminale, and P. Jalinot. 2012. 'The human T-lymphotropic virus type 1 tax protein inhibits nonsense-mediated mRNA decay by interacting with INT6/EIF3E and UPF1', *J Virol*, 86: 7530-43.

Moiani, A., Y. Paleari, D. Sartori, R. Mezzadra, A. Miccio, C. Cattoglio, F. Cocchiarella, M. R. Lidonnici, G. Ferrari, and F. Mavilio. 2012. 'Lentiviral vector integration in the human genome induces alternative splicing and generates aberrant transcripts', *J Clin Invest*, 122: 1653-66.

Mondor, I., S. Ugolini, and Q. J. Sattentau. 1998. 'Human immunodeficiency virus type 1 attachment to HeLa CD4 cells is CD4 independent and gp120 dependent and requires cell surface heparans', *J Virol*, 72: 3623-34.

Moriarty, P. M., C. C. Reddy, and L. E. Maquat. 1998. 'Selenium deficiency reduces the abundance of mRNA for Se-dependent glutathione peroxidase 1 by a UGA-dependent mechanism likely to be nonsense codon-mediated decay of cytoplasmic mRNA', *Mol Cell Biol*, 18: 2932-9.

Morner, A., A. Bjorndal, J. Albert, V. N. Kewalramani, D. R. Littman, R. Inoue, R. Thorstensson, E. M. Fenyo, and E. Bjorling. 1999. 'Primary human immunodeficiency virus type 2 (HIV-2) isolates, like HIV-1 isolates, frequently use CCR5 but show promiscuity in coreceptor usage', *J Virol*, 73: 2343-9.

Mort, M., D. Ivanov, D. N. Cooper, and N. A. Chuzhanova. 2008. 'A meta-analysis of nonsense mutations causing human genetic disease', *Hum Mutat*, 29: 1037-47.

Nakano, K., T. Ando, M. Yamagishi, K. Yokoyama, T. Ishida, T. Ohsugi, Y. Tanaka, D. W. Brighty, and T. Watanabe. 2013. 'Viral interference with host mRNA surveillance, the nonsense-mediated mRNA decay (NMD) pathway, through a new function of HTLV-1 Rex: implications for retroviral replication', *Microbes Infect*, 15: 491-505.

Newman, A. J., and C. Norman. 1992. 'U5 snRNA interacts with exon sequences at 5' and 3' splice sites', *Cell*, 68: 743-54.

Nguyen, T. H., W. P. Galej, X. C. Bai, C. G. Savva, A. J. Newman, S. H. Scheres, and K. Nagai. 2015. 'The architecture of the spliceosomal U4/U6.U5 tri-snRNP', *Nature*, 523: 47-52.

Ni, J. Z., L. Grate, J. P. Donohue, C. Preston, N. Nobida, G. O'Brien, L. Shiue, T. A. Clark, J. E. Blume, and M. Ares, Jr. 2007. 'Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay', *Genes Dev*, 21: 708-18.

Nicholson, P., A. Gkratsou, C. Josi, M. Colombo, and O. Mühlemann. 2018. 'Dissecting the functions of SMG5, SMG7, and PNRC2 in nonsense-mediated mRNA decay of human cells', *RNA*, 24: 557-73.

Nickless, A., J. M. Bailis, and Z. You. 2017. 'Control of gene expression through the nonsense-mediated RNA decay pathway', *Cell Biosci*, 7: 26.

Nikolaitchik, O. A., K. A. Dilley, W. Fu, R. J. Gorelick, S. H. Tai, F. Soheilian, R. G. Ptak, K. Nagashima, V. K. Pathak, and W. S. Hu. 2013. 'Dimeric RNA recognition regulates HIV-1 genome packaging', *PLoS Pathog*, 9: e1003249.

Nogues, G., S. Kadener, P. Cramer, D. Bentley, and A. R. Kornblihtt. 2002. 'Transcriptional activators differ in their abilities to control alternative splicing', *J Biol Chem*, 277: 43110-4.

O'Reilly, M. M., M. T. McNally, and K. L. Beemon. 1995. 'Two strong 5' splice sites and competing, suboptimal 3' splice sites involved in alternative splicing of human immunodeficiency virus type 1 RNA', *Virology*, 213: 373-85.

Ocwieja, K. E., S. Sherrill-Mix, R. Mukherjee, R. Custers-Allen, P. David, M. Brown, S. Wang, D. R. Link, J. Olson, K. Travers, E. Schadt, and F. D. Bushman. 2012. 'Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule enrichment and long-read sequencing', *Nucleic Acids Res*, 40: 10345-55.

Ohnishi, T., A. Yamashita, I. Kashima, T. Schell, K. R. Anders, A. Grimson, T. Hachiya, M. W. Hentze, P. Anderson, and S. Ohno. 2003. 'Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7', *Mol Cell*, 12: 1187-200.

Okada-Katsuhata, Y., A. Yamashita, K. Kutsuzawa, N. Izumi, F. Hirahara, and S. Ohno. 2012. 'N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD', *Nucleic Acids Res*, 40: 1251-66.

Ottens, F., and N. H. Gehring. 2016. 'Physiological and pathophysiological role of nonsense-mediated mRNA decay', *Pflugers Arch*, 468: 1013-28.

Padgett, R. A., M. M. Konarska, P. J. Grabowski, S. F. Hardy, and P. A. Sharp. 1984. 'Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors', *Science*, 225: 898-903.

Pan, Q., A. L. Saltzman, Y. K. Kim, C. Misquitta, O. Shai, L. E. Maquat, B. J. Frey, and B. J. Blencowe. 2006. 'Quantitative microarray profiling provides evidence against widespread coupling of alternative splicing with nonsense-mediated mRNA decay to control gene expression', *Genes Dev*, 20: 153-8.

Pan, Q., O. Shai, L. J. Lee, B. J. Frey, and B. J. Blencowe. 2008. 'Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing', *Nat Genet*, 40: 1413-5.

Parker, R., P. G. Siliciano, and C. Guthrie. 1987. 'Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA', *Cell*, 49: 229-39.

Patikoglou, G. A., J. L. Kim, L. Sun, S. H. Yang, T. Kodadek, and S. K. Burley. 1999. 'TATA element recognition by the TATA box-binding protein has been conserved throughout evolution', *Genes Dev*, 13: 3217-30.

Paz, S., A. R. Krainer, and M. Caputi. 2014. 'HIV-1 transcription is regulated by splicing factor SRSF1', *Nucleic Acids Res*, 42: 13812-23.

Pemberton, L. F., G. Blobel, and J. S. Rosenblum. 1998. 'Transport routes through the nuclear pore complex', *Curr Opin Cell Biol*, 10: 392-9.

Pennings, P. S. 2013. 'HIV Drug Resistance: Problems and Perspectives', *Infect Dis Rep*, 5: e5.

Perales, C., L. Carrasco, and M. E. Gonzalez. 2005. 'Regulation of HIV-1 env mRNA translation by Rev protein', *Biochim Biophys Acta*, 1743: 169-75.

Perriman, R., and M. Ares, Jr. 2010. 'Invariant U2 snRNA nucleotides form a stem loop to recognize the intron early in splicing', *Mol Cell*, 38: 416-27.

Pervouchine, D., Y. Popov, A. Berry, B. Borsari, A. Frankish, and R. Guigo. 2019. 'Integrative transcriptomic analysis suggests new autoregulatory splicing events coupled with nonsense-mediated mRNA decay', *Nucleic Acids Res*, 47: 5293-306.

Pestova, T. V., I. B. Lomakin, J. H. Lee, S. K. Choi, T. E. Dever, and C. U. Hellen. 2000. 'The joining of ribosomal subunits in eukaryotes requires eIF5B', *Nature*, 403: 332-5.

Pisarev, A. V., M. A. Skabkin, V. P. Pisareva, O. V. Skabkina, A. M. Rakotondrafara, M. W. Hentze, C. U. Hellen, and T. V. Pestova. 2010. 'The role of ABCE1 in eukaryotic posttermination ribosomal recycling', *Mol Cell*, 37: 196-210.

Plaschka, C., P. C. Lin, and K. Nagai. 2017. 'Structure of a pre-catalytic spliceosome', *Nature*, 546: 617-21.

Platt, E. J., K. Wehrly, S. E. Kuhmann, B. Chesebro, and D. Kabat. 1998. 'Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1', *J Virol*, 72: 2855-64.

Poli, G., J. M. Orenstein, A. Kinter, T. M. Folks, and A. S. Fauci. 1989. 'Interferon-alpha but not AZT suppresses HIV expression in chronically infected cell lines', *Science*, 244: 575-7.

Poling, B. C., K. Tsai, D. Kang, L. Ren, E. M. Kennedy, and B. R. Cullen. 2017. 'A lentiviral vector bearing a reverse intron demonstrates superior expression of both proteins and microRNAs', *RNA Biol*, 14: 1570-79.

Ponjavic, J., B. Lenhard, C. Kai, J. Kawai, P. Carninci, Y. Hayashizaki, and A. Sandelin. 2006. 'Transcriptional and structural impact of TATA-initiation site spacing in mammalian core promoters', *Genome Biol*, 7: R78.

Pope, M., M. G. Betjes, N. Romani, H. Hirmand, P. U. Cameron, L. Hoffman, S. Gezelter, G. Schuler, and R. M. Steinman. 1994. 'Conjugates of dendritic cells and memory T lymphocytes from skin facilitate productive infection with HIV-1', *Cell*, 78: 389-98.

Popp, M. W., H. Cho, and L. E. Maquat. 2020. 'Viral subversion of nonsense-mediated mRNA decay', *RNA*, 26: 1509-18.

Pornillos, O., B. K. Ganser-Pornillos, and M. Yeager. 2011. 'Atomic-level modelling of the HIV capsid', *Nature*, 469: 424-7.

Presnyak, V., N. Alhusaini, Y. H. Chen, S. Martin, N. Morris, N. Kline, S. Olson, D. Weinberg, K. E. Baker, B. R. Graveley, and J. Coller. 2015. 'Codon optimality is a major determinant of mRNA stability', *Cell*, 160: 1111-24.

Ptok, J., L. Muller, S. Theiss, and H. Schaal. 2019. 'Context matters: Regulation of splice donor usage', *Biochim Biophys Acta Gene Regul Mech*, 1862: 194391.

Purcell, D. F., and M. A. Martin. 1993. 'Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity', *J Virol*, 67: 6365-78.

Radhakrishnan, A., Y. H. Chen, S. Martin, N. Alhusaini, R. Green, and J. Coller. 2016. 'The DEAD-Box Protein Dhh1p Couples mRNA Decay and Translation by Monitoring Codon Optimality', *Cell*, 167: 122-32 e9.

Raghunathan, P. L., and C. Guthrie. 1998. 'RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2', *Curr Biol*, 8: 847-55.

Ralser, M., G. Heeren, M. Breitenbach, H. Lehrach, and S. Krobitsch. 2006. 'Triose phosphate isomerase deficiency is caused by altered dimerization--not catalytic inactivity--of the mutant enzymes', *PLoS One*, 1: e30.

Ramage, H. R., G. R. Kumar, E. Verschueren, J. R. Johnson, J. Von Dollen, T. Johnson, B. Newton, P. Shah, J. Horner, N. J. Krogan, and M. Ott. 2015. 'A combined proteomics/genomics

approach links hepatitis C virus infection with nonsense-mediated mRNA decay', *Mol Cell*, 57: 329-40.

Ramdas, P., A. K. Sahu, T. Mishra, V. Bhardwaj, and A. Chande. 2020. 'From Entry to Egress: Strategic Exploitation of the Cellular Processes by HIV-1', *Front Microbiol*, 11: 559792.

Rao, S., R. Amorim, M. Niu, Y. Breton, M. J. Tremblay, and A. J. Mouland. 2019. 'Host mRNA decay proteins influence HIV-1 replication and viral gene expression in primary monocyte-derived macrophages', *Retrovirology*, 16: 3.

Rasaiyaah, J., C. P. Tan, A. J. Fletcher, A. J. Price, C. Blondeau, L. Hilditch, D. A. Jacques, D. L. Selwood, L. C. James, M. Noursadeghi, and G. J. Towers. 2013. 'HIV-1 evades innate immune recognition through specific cofactor recruitment', *Nature*, 503: 402-05.

Raxwal, V. K., C. G. Simpson, J. Gloggnitzer, J. C. Entinze, W. Guo, R. Zhang, J. W. S. Brown, and K. Riha. 2020. 'Nonsense-Mediated RNA Decay Factor UPF1 Is Critical for Posttranscriptional and Translational Gene Regulation in Arabidopsis', *Plant Cell*, 32: 2725-41.

Rehwinkel, J., I. Letunic, J. Raes, P. Bork, and E. Izaurralde. 2005. 'Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets', *RNA*, 11: 1530-44.

Reil, H., M. Hoxter, D. Moosmayer, G. Pauli, and H. Hauser. 1994. 'CD4 expressing human 293 cells as a tool for studies in HIV-1 replication: the efficiency of translational frameshifting is not altered by HIV-1 infection', *Virology*, 205: 371-5.

Rein, A. 2019. 'RNA Packaging in HIV', *Trends Microbiol*, 27: 715-23.

Rigby, R. E., and J. Rehwinkel. 2015. 'RNA degradation in antiviral immunity and autoimmunity', *Trends Immunol*, 36: 179-88.

RKI, Robert Koch Institut. 2020. "HIV/AIDS in Deutschland." In, Epidemiologische Kurzinformation des Robert Koch-Instituts zu den HIV/AIDS-Eckdaten (Stand: Ende 2019).

Robbins-Pianka, A., M. D. Rice, and M. P. Weir. 2010. 'The mRNA landscape at yeast translation initiation sites', *Bioinformatics*, 26: 2651-5.

Röllecke, K. 2016. 'A modified P450 Cytochrome as Safety Mechanism in Adoptive T-cell Therapy', Dissertation, Heinrich-Heine University Düsseldorf.

Ruskin, B., A. R. Krainer, T. Maniatis, and M. R. Green. 1984. 'Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro', *Cell*, 38: 317-31.

Ryo, A., Y. Suzuki, M. Arai, N. Kondoh, T. Wakatsuki, A. Hada, M. Shuda, K. Tanaka, C. Sato, M. Yamamoto, and N. Yamamoto. 2000. 'Identification and characterization of differentially expressed mRNAs in HIV type 1-infected human T cells', *AIDS Res Hum Retroviruses*, 16: 995-1005.

Schaal, H., M. Klein, P. Gehrmann, O. Adams, and A. Scheid. 1995. 'Requirement of N-terminal amino acid residues of gp41 for human immunodeficiency virus type 1-mediated cell fusion', *J Virol*, 69: 3308-14.

Schaal, H., P. Pfeiffer, M. Klein, P. Gehrmann, and A. Scheid. 1993. 'Use of DNA end joining activity of a Xenopus laevis egg extract for construction of deletions and expression vectors for HIV-1 Tat and Rev proteins', *Gene*, 124: 275-80.

Scherer, W. F., J. T. Syverton, and G. O. Gey. 1953. 'Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix', *J Exp Med*, 97: 695-710.

Schneider-Poetsch, T., J. Ju, D. E. Eyler, Y. Dang, S. Bhat, W. C. Merrick, R. Green, B. Shen, and J. O. Liu. 2010. 'Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin', *Nat Chem Biol*, 6: 209-17.

Schneider, A., and M. Cohen-Solal. 1996. 'Hematologically important mutations: triosephosphate isomerase', *Blood Cells Mol Dis*, 22: 82-4.

Schoggins, J. W., S. J. Wilson, M. Panis, M. Y. Murphy, C. T. Jones, P. Bieniasz, and C. M. Rice. 2011. 'A diverse range of gene products are effectors of the type I interferon antiviral response', *Nature*, 472: 481-5.

Schuller, A. P., and R. Green. 2018. 'Roadblocks and resolutions in eukaryotic translation', *Nat Rev Mol Cell Biol*, 19: 526-41.

Schwartz, S., B. K. Felber, D. M. Benko, E. M. Fenyo, and G. N. Pavlakis. 1990. 'Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type 1', *J Virol*, 64: 2519-29.

Schwer, B. 2008. 'A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release', *Mol Cell*, 30: 743-54.

Selyutina, A., M. Persaud, K. Lee, V. KewalRamani, and F. Diaz-Griffero. 2020. 'Nuclear Import of the HIV-1 Core Precedes Reverse Transcription and Uncoating', *Cell Rep*, 32: 108201.

Senapathy, P., M. B. Shapiro, and N. L. Harris. 1990. 'Splice junctions, branch point sites, and exons: sequence statistics, identification, and applications to genome project', *Methods Enzymol*, 183: 252-78.

Seraphin, B. 1995. 'Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs', *EMBO J*, 14: 2089-98.

Serquina, A. K., S. R. Das, E. Popova, O. A. Ojelabi, C. K. Roy, and H. G. Gottlinger. 2013. 'UPF1 is crucial for the infectivity of human immunodeficiency virus type 1 progeny virions', *J Virol*, 87: 8853-61.

Sertznig, H., F. Hillebrand, S. Erkelenz, H. Schaal, and M. Widera. 2018. 'Behind the scenes of HIV-1 replication: Alternative splicing as the dependency factor on the quiet', *Virology*, 516: 176-88.

Sertznig, H., F. Roesmann, B. Bleekmann, C. Elsner, M. Santiago, J. Schuhenn, Y. Benatzy, R. Snodgrass, S. Esser, K. Sutter, U. Dittmer, and M. Widera. 2021. 'Interferon-α subtype treatment induces the repression of SRSF1 in HIV-1 target cells and affects HIV-1 post integration steps (preprint)', *bioRxiv*.

Sherman, F., G. McKnight, and J. W. Stewart. 1980. 'AUG is the only initiation codon in eukaryotes', *Biochim Biophys Acta*, 609: 343-6.

Sheth, N., X. Roca, M. L. Hastings, T. Roeder, A. R. Krainer, and R. Sachidanandam. 2006. 'Comprehensive splice-site analysis using comparative genomics', *Nucleic Acids Res*, 34: 3955-67.

Shlyueva, D., G. Stampfel, and A. Stark. 2014. 'Transcriptional enhancers: from properties to genome-wide predictions', *Nat Rev Genet*, 15: 272-86.

Silva, A. L., P. Ribeiro, A. Inacio, S. A. Liebhaber, and L. Romao. 2008. 'Proximity of the poly(A)-binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay', *RNA*, 14: 563-76.

Singh, G., A. Kucukural, C. Cenik, J. D. Leszyk, S. A. Shaffer, Z. Weng, and M. J. Moore. 2012. 'The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus', *Cell*, 151: 750-64.

Singh, G., I. Rebbapragada, and J. Lykke-Andersen. 2008. 'A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay', *PLoS Biol*, 6: e111.

Smith, C. W., T. T. Chu, and B. Nadal-Ginard. 1993. 'Scanning and competition between AGs are involved in 3' splice site selection in mammalian introns', *Mol Cell Biol*, 13: 4939-52.

Sokabe, M., C. S. Fraser, and J. W. Hershey. 2012. 'The human translation initiation multi-factor complex promotes methionyl-tRNAi binding to the 40S ribosomal subunit', *Nucleic Acids Res*, 40: 905-13.

Sontheimer, E. J., and J. A. Steitz. 1993. 'The U5 and U6 small nuclear RNAs as active site components of the spliceosome', *Science*, 262: 1989-96.

Soper, A., I. Kimura, S. Nagaoka, Y. Konno, K. Yamamoto, Y. Koyanagi, and K. Sato. 2017. 'Type I Interferon Responses by HIV-1 Infection: Association with Disease Progression and Control', *Front Immunol*, 8: 1823.

Staley, J. P., and C. Guthrie. 1999. 'An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p', *Mol Cell*, 3: 55-64.

Steckelberg, A. L., V. Boehm, A. M. Gromadzka, and N. H. Gehring. 2012. 'CWC22 connects pre-mRNA splicing and exon junction complex assembly', *Cell Rep*, 2: 454-61.

Stoltzfus, C. M. 2009. 'Chapter 1. Regulation of HIV-1 alternative RNA splicing and its role in virus replication', *Adv Virus Res*, 74: 1-40.

Strässer, K., J. Bassler, and E. Hurt. 2000. 'Binding of the Mex67p/Mtr2p heterodimer to FXFG, GLFG, and FG repeat nucleoporins is essential for nuclear mRNA export', *J Cell Biol*, 150: 695-706.

Strebel, K., T. Klimkait, and M. A. Martin. 1988. 'A novel gene of HIV-1, vpu, and its 16-kilodalton product', *Science*, 241: 1221-3.

Sundquist, W. I., and H. G. Kräusslich. 2012. 'HIV-1 assembly, budding, and maturation', *Cold Spring Harb Perspect Med*, 2: a006924.

Tani, H., R. Mizutani, K. A. Salam, K. Tano, K. Ijiri, A. Wakamatsu, T. Isogai, Y. Suzuki, and N. Akimitsu. 2012. 'Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals', *Genome Res*, 22: 947-56.

Taniguchi, I., N. Mabuchi, and M. Ohno. 2014. 'HIV-1 Rev protein specifies the viral RNA export pathway by suppressing TAP/NXF1 recruitment', *Nucleic Acids Res*, 42: 6645-58.

Terry, L. J., and S. R. Wente. 2009. 'Flexible gates: dynamic topologies and functions for FG nucleoporins in nucleocytoplasmic transport', *Eukaryot Cell*, 8: 1814-27.

Tersteegen, A. 2017. 'Generation of a reporter system to detect the influence of cytokines on the degradation of aberrant mRNAs', Master Thesis, Heinrich-Heine-University Düsseldorf.

Thermann, R., G. Neu-Yilik, A. Deters, U. Frede, K. Wehr, C. Hagemeier, M. W. Hentze, and A. E. Kulozik. 1998. 'Binary specification of nonsense codons by splicing and cytoplasmic translation', *EMBO J*, 17: 3484-94.

Toma, K. G., I. Rebbapragada, S. Durand, and J. Lykke-Andersen. 2015. 'Identification of elements in human long 3' UTRs that inhibit nonsense-mediated decay', *RNA*, 21: 887-97.

Trcek, T., H. Sato, R. H. Singer, and L. E. Maquat. 2013. 'Temporal and spatial characterization of nonsense-mediated mRNA decay', *Genes Dev*, 27: 541-51.

Trono, D., C. Van Lint, C. Rouzioux, E. Verdin, F. Barre-Sinoussi, T. W. Chun, and N. Chomont. 2010. 'HIV persistence and the prospect of long-term drug-free remissions for HIV-infected individuals', *Science*, 329: 174-80.

Truant, R., and B. R. Cullen. 1999. 'The arginine-rich domains present in human immunodeficiency virus type 1 Tat and Rev function as direct importin beta-dependent nuclear localization signals', *Mol Cell Biol*, 19: 1210-7.

Truman, C. T., A. Jarvelin, I. Davis, and A. Castello. 2020. 'HIV Rev-isited', *Open Biol*, 10: 200320.

Uehata, T., and O. Takeuchi. 2020. 'RNA Recognition and Immunity-Innate Immune Sensing and Its Posttranscriptional Regulation Mechanisms', *Cells*, 9.

Ule, J., and B. J. Blencowe. 2019. 'Alternative Splicing Regulatory Networks: Functions, Mechanisms, and Evolution', *Mol Cell*, 76: 329-45.

Unterholzner, L., and E. Izaurralde. 2004. 'SMG7 acts as a molecular link between mRNA surveillance and mRNA decay', *Mol Cell*, 16: 587-96.

Varenne, S., J. Buc, R. Lloubes, and C. Lazdunski. 1984. 'Translation is a non-uniform process. Effect of tRNA availability on the rate of elongation of nascent polypeptide chains', *J Mol Biol*, 180: 549-76.

Viegas, M. H., N. H. Gehring, S. Breit, M. W. Hentze, and A. E. Kulozik. 2007. 'The abundance of RNPS1, a protein component of the exon junction complex, can determine the variability in efficiency of the Nonsense Mediated Decay pathway', *Nucleic Acids Res*, 35: 4542-51.

Wada, M., K. G. Lokugamage, K. Nakagawa, K. Narayanan, and S. Makino. 2018. 'Interplay between coronavirus, a cytoplasmic RNA virus, and nonsense-mediated mRNA decay pathway', *Proc Natl Acad Sci U S A*, 115: E10157-E66.

Wahl, M. C., C. L. Will, and R. Lührmann. 2009. 'The spliceosome: design principles of a dynamic RNP machine', *Cell*, 136: 701-18.

Wang, E. T., R. Sandberg, S. Luo, I. Khrebtukova, L. Zhang, C. Mayr, S. F. Kingsmore, G. P. Schroth, and C. B. Burge. 2008. 'Alternative isoform regulation in human tissue transcriptomes', *Nature*, 456: 470-6.

Wang, Z., and C. B. Burge. 2008. 'Splicing regulation: from a parts list of regulatory elements to an integrated splicing code', *RNA*, 14: 802-13.

Webster, M. W., Y. H. Chen, J. A. W. Stowell, N. Alhusaini, T. Sweet, B. R. Graveley, J. Coller, and L. A. Passmore. 2018. 'mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases', *Mol Cell*, 70: 1089-100 e8.

Weil, J. E., and K. L. Beemon. 2006. 'A 3' UTR sequence stabilizes termination codons in the unspliced RNA of Rous sarcoma virus', *RNA*, 12: 102-10.

Weischenfeldt, J., J. Waage, G. Tian, J. Zhao, I. Damgaard, J. S. Jakobsen, K. Kristiansen, A. Krogh, J. Wang, and B. T. Porse. 2012. 'Mammalian tissues defective in nonsense-mediated mRNA decay display highly aberrant splicing patterns', *Genome Biol*, 13: R35.

Weiss, A., R. L. Wiskocil, and J. D. Stobo. 1984. 'The role of T3 surface molecules in the activation of human T cells: a two-stimulus requirement for IL 2 production reflects events occurring at a pre-translational level', *J Immunol*, 133: 123-8.

Wente, S. R., and M. P. Rout. 2010. 'The nuclear pore complex and nuclear transport', *Cold Spring Harb Perspect Biol*, 2: a000562.

WHO, World Health Organization. 2021. 'HIV/AIDS Key Facts', Accessed July 2021. https://www.who.int/news-room/fact-sheets/detail/hiv-aids.

Widera, M., S. Erkelenz, F. Hillebrand, A. Krikoni, D. Widera, W. Kaisers, R. Deenen, M. Gombert, R. Dellen, T. Pfeiffer, B. Kaltschmidt, C. Munk, V. Bosch, K. Kohrer, and H. Schaal. 2013. 'An intronic G run within HIV-1 intron 2 is critical for splicing regulation of vif mRNA', *J Virol*, 87: 2707-20.

Widera, M., F. Hillebrand, S. Erkelenz, A. A. Vasudevan, C. Munk, and H. Schaal. 2014. 'A functional conserved intronic G run in HIV-1 intron 3 is critical to counteract APOBEC3G-mediated host restriction', *Retrovirology*, 11: 72.

Wierenga, R. K., E. G. Kapetaniou, and R. Venkatesan. 2010. 'Triosephosphate isomerase: a highly evolved biocatalyst', *Cell Mol Life Sci*, 67: 3961-82.

Wilen, C. B., J. C. Tilton, and R. W. Doms. 2012. 'HIV: cell binding and entry', *Cold Spring Harb Perspect Med*, 2.

Will, C. L., and R. Lührmann. 2011. 'Spliceosome structure and function', *Cold Spring Harb Perspect Biol*, 3.

Withers, J. B., and K. L. Beemon. 2011. 'The structure and function of the rous sarcoma virus RNA stability element', *J Cell Biochem*, 112: 3085-92.

Wittkopp, N., E. Huntzinger, C. Weiler, J. Sauliere, S. Schmidt, M. Sonawane, and E. Izaurralde. 2009. 'Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic development and survival', *Mol Cell Biol*, 29: 3517-28.

Wong, M. S., J. B. Kinney, and A. R. Krainer. 2018. 'Quantitative Activity Profile and Context Dependence of All Human 5' Splice Sites', *Mol Cell*, 71: 1012-26 e3.

Wu, J., and J. L. Manley. 1989. 'Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing', *Genes Dev*, 3: 1553-61.

Wu, Q., S. G. Medina, G. Kushawah, M. L. DeVore, L. A. Castellano, J. M. Hand, M. Wright, and A. A. Bazzini. 2019. 'Translation affects mRNA stability in a codon-dependent manner in human cells', *Elife*, 8.

Wu, X., Y. Li, B. Crise, S. M. Burgess, and D. J. Munroe. 2005. 'Weak palindromic consensus sequences are a common feature found at the integration target sites of many retroviruses', *J Virol*, 79: 5211-4.

Xin, D., L. Hu, and X. Kong. 2008. 'Alternative promoters influence alternative splicing at the genomic level', *PLoS One*, 3: e2377.

Yamashita, A., T. C. Chang, Y. Yamashita, W. Zhu, Z. Zhong, C. Y. Chen, and A. B. Shyu. 2005. 'Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover', *Nat Struct Mol Biol*, 12: 1054-63.

Yamashita, A., N. Izumi, I. Kashima, T. Ohnishi, B. Saari, Y. Katsuhata, R. Muramatsu, T. Morita, A. Iwamatsu, T. Hachiya, R. Kurata, H. Hirano, P. Anderson, and S. Ohno. 2009. 'SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay', *Genes Dev*, 23: 1091-105.

Yamashita, A., T. Ohnishi, I. Kashima, Y. Taya, and S. Ohno. 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 Dev*, 15: 2215-28.

Yedavalli, V. S., and K. T. Jeang. 2010. 'Trimethylguanosine capping selectively promotes expression of Rev-dependent HIV-1 RNAs', *Proc Natl Acad Sci U S A*, 107: 14787-92.

Yeo, G., and C. B. Burge. 2004. 'Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals', *J Comput Biol*, 11: 377-94.

Yeo, G., D. Holste, G. Kreiman, and C. B. Burge. 2004. 'Variation in alternative splicing across human tissues', *Genome Biol*, 5: R74.

Yepiskoposyan, H., F. Aeschimann, D. Nilsson, M. Okoniewski, and O. Mühlemann. 2011. 'Autoregulation of the nonsense-mediated mRNA decay pathway in human cells', *RNA*, 17: 2108-18.

Zetoune, A. B., S. Fontaniere, D. Magnin, O. Anczukow, M. Buisson, C. X. Zhang, and S. Mazoyer. 2008. 'Comparison of nonsense-mediated mRNA decay efficiency in various murine tissues', *BMC Genet*, 9: 83.

Zhan, X., C. Yan, X. Zhang, J. Lei, and Y. Shi. 2018. 'Structures of the human pre-catalytic spliceosome and its precursor spliceosome', *Cell Res*, 28: 1129-40.

Zhu, T., B. T. Korber, A. J. Nahmias, E. Hooper, P. M. Sharp, and D. D. Ho. 1998. 'An African HIV-1 sequence from 1959 and implications for the origin of the epidemic', *Nature*, 391: 594-7.

Zhuang, Y. A., A. M. Goldstein, and A. M. Weiner. 1989. 'UACUAAC is the preferred branch site for mammalian mRNA splicing', *Proc Natl Acad Sci U S A*, 86: 2752-6.

Zhuang, Y., and A. M. Weiner. 1986. 'A compensatory base change in U1 snRNA suppresses a 5' splice site mutation', *Cell*, 46: 827-35.

Zila, V., E. Margiotta, B. Turonova, T. G. Müller, C. E. Zimmerli, S. Mattei, M. Allegretti, K. Börner, J. Rada, B. Müller, M. Lusic, H. G. Kräusslich, and M. Beck. 2021. 'Cone-shaped HIV-1 capsids are transported through intact nuclear pores', *Cell*, 184: 1032-46 e18.

Zünd, D., A. R. Gruber, M. Zavolan, and O. Mühlemann. 2013. 'Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs', *Nat Struct Mol Biol*, 20: 936-43.

6 Appendix
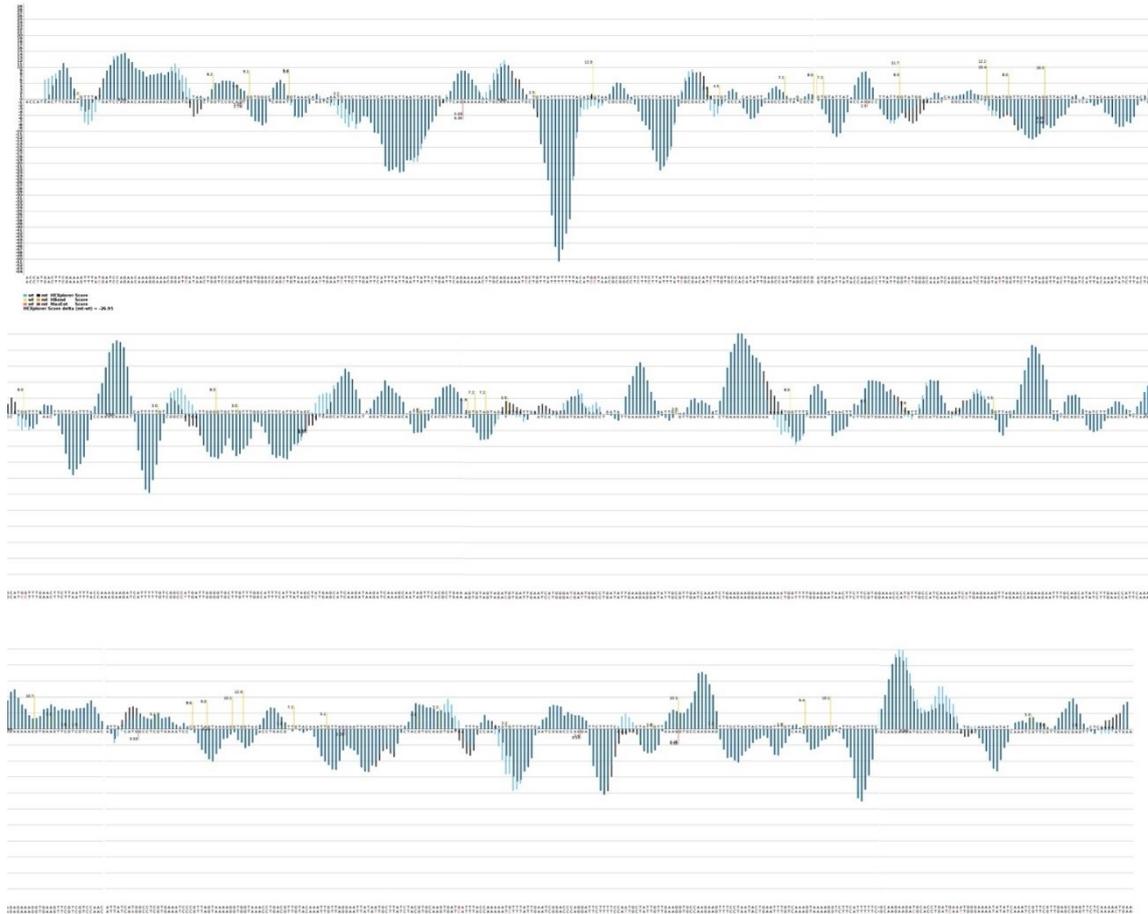
Figure 6.1: HEXplorer profile of the RL-β-Globin gene. All 46 potential translational start codons in all three potential open reading frames were mutated. Starting at the original renilla ATG until the polyadenylation signal downstream of β-Globin exon 3 (compare Figure 3.10A, red lines). For ATGs which were located within the original open reading frame, it was not always possible to substitute bases in form of a silent mutation, but the mutations were selected with respect to the HEXplorer profile, to keep the changes to the profile as minimal as possible.
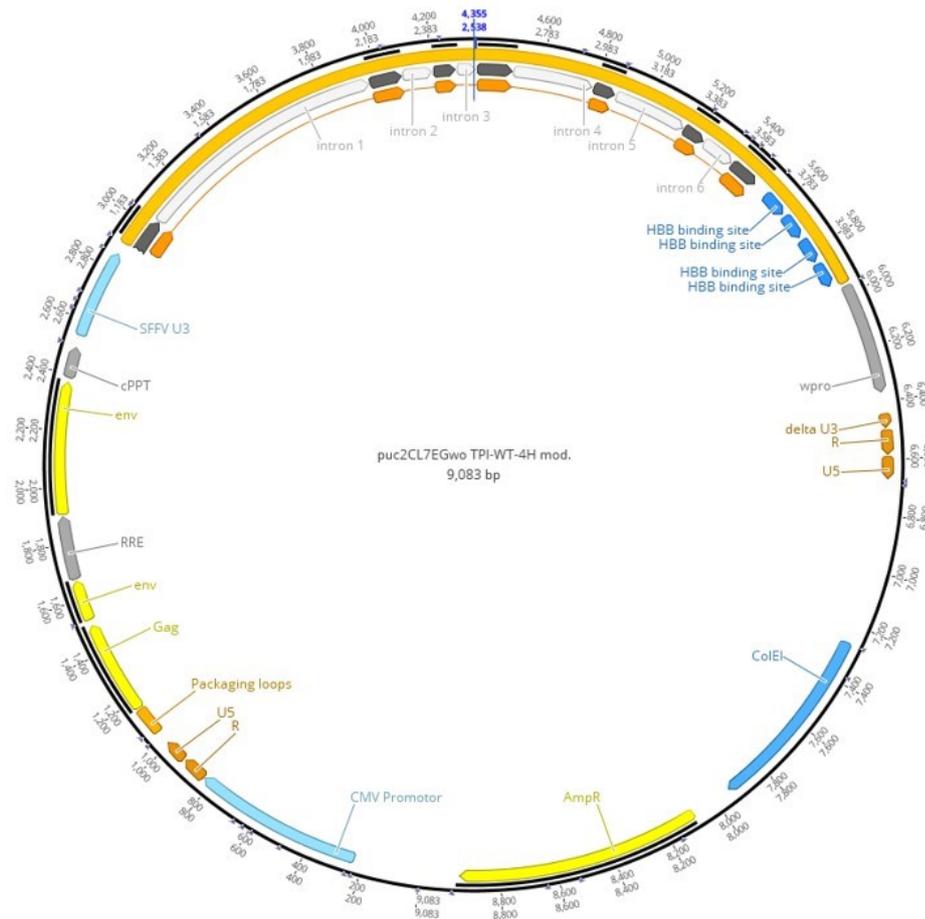
Figure 6.2: Plasmid card of the used TPI lentiviral vector structure. The plasmid had a size of 9,083 nucleotides and encoded the ampicillin resistance gene, as well as the ColEI ORI for amplification in *E. coli*. The modified HIV-1 5'LTR contained the CMV promotor followed by the R and U5 sequence. The CMV promoter made the plasmid expression independent of tat levels. The LTR was followed by the HIV-1 leader sequence, which contained the packaging signal, as well as packaging relevant sequences of the gag and env genes and the central polypurine tract (cPPT). The expression of the TPI reporter cassette was controlled by the endogenous SFFV U3 promoter and the optimized woodchuck hepatitis virus posttranscriptional regulatory element (WPRO). To ensure the self-inactivating (SIN) capacities of the lentiviral vector system, the U3 promoter and enhancer region was deleted which inactivated the newly generated 5'LTR upon reverse transcription.
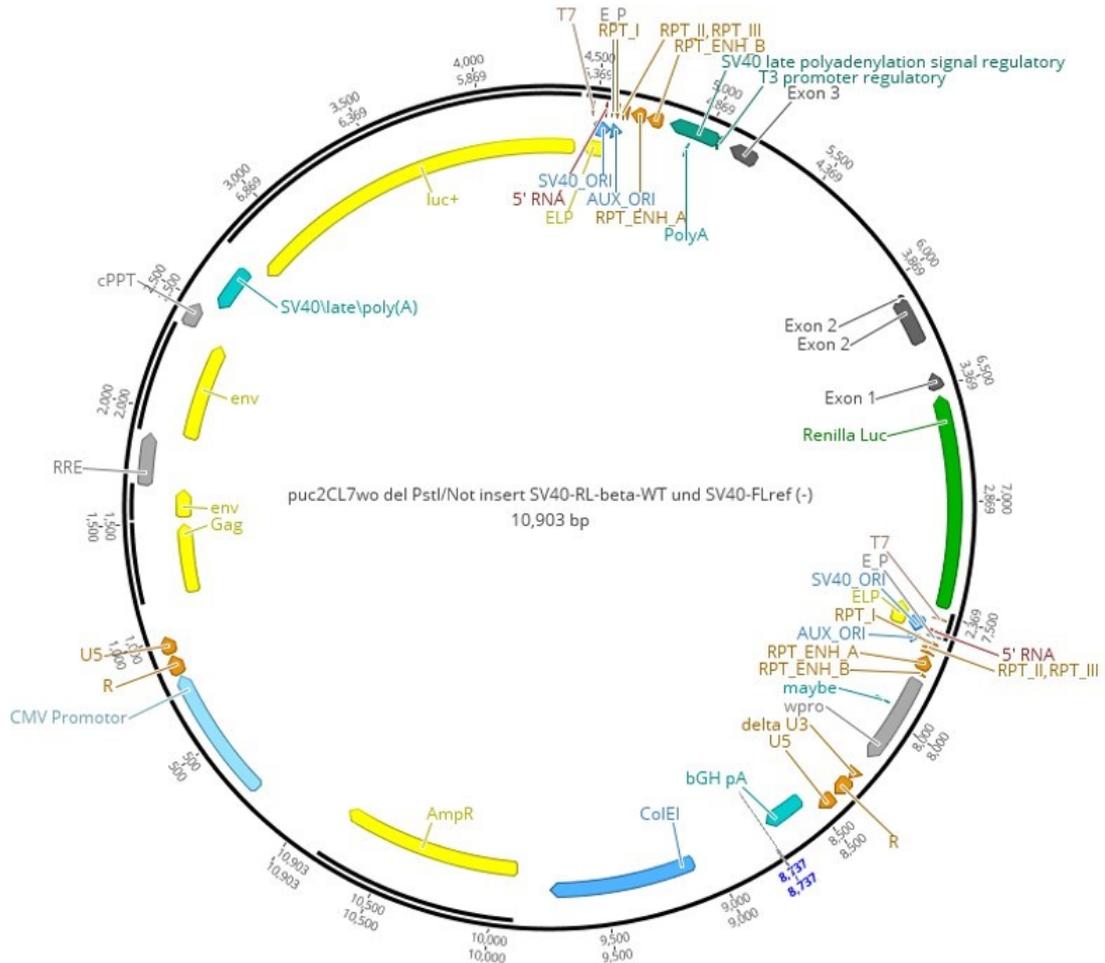
Figure 6.3: Plasmid card of the designed lentiviral vector structure encoding for the dual-luciferase reporter system located on the minus strand. The plasmid had a size of 10,903 nucleotides and encoded the ampicillin resistance gene, as well as the ColEI ORI for amplification in *E. coli*. The modified HIV-1 5'LTR contained the CMV promotor followed by the R and U5 sequence. The CMV promoter made the plasmid expression independent of tat levels. The LTR was followed by the HIV-1 leader sequence, which contained the packaging signal, as well as packaging relevant sequences of the gag and env genes and the central polypurine tract (cPPT). The sequence between the cPPT and WPRO were removed and replaced by the dual-luciferase NMD reporter system, either WT or NS39, each with its own SV40 promoter and polyadenylation site. To prevent unwanted splicing events within the producer cell line the transgenes were inserted in 3' to 5' direction. To ensure the self-inactivating (SIN) capacities of the lentiviral vector system, the U3 promotor sand enhancer region was deleted which inactivated the newly generated 5'LTR upon reverse transcription.

167

Table 6.1: Cloning strategies for the newly generated plasmids. IC = intermediate clone

Plasmid name	Plasmid backbone	Insert	Template	Restriction enzymes
pCI-TPI-WT-4H Exon 3 HBS 20.8	-	PCR product Q5 SDM #5672/#5673	pCI-TPI-WT-4H	-
pCI-TPI-PTC40-4H Exon 3 HBS 20.8	-	PCR product Q5 SDM #5672/#5673	pCI-TPI-PTC40-4H	-
pCI-TPI-PTC40-4H Exon 2 ΔHZ 35.3	pCI-TPI-PTC40-4H	PCR product #5607/#5665	PCR products #5607/#5683 and #5682/#5665 pCI-TPI-PTC40-4H	XcmI/ BstEII-HF
pCI-TPI-WT-4H Exon 2 ΔHZ 35.3	pCI-TPI-WT-4H	Digested fragment	pCI-TPI-PTC40-4H Exon 2 ΔHZ 35.3	XcmI/ PmlI
pCI-TPI-PTC40-4H Exon 2 ΔHZ 35.3 + Exon 3 HBS 20.8	pCI-TPI-PTC40-4H Exon 2 ΔHZ 35.3	Digested fragment	pCI-TPI-PTC40-4H Exon 3 HBS 20.8	PmlI/ BstEII-HF
pCI-TPI-WT-4H Exon 2 ΔHZ 35.3 + Exon 3 HBS 20.8	pCI-TPI-WT-4H Exon 2 ΔHZ 35.3	Digested fragment	pCI-TPI-WT-4H Exon 3 HBS 20.8	PmlI/ BstEII-HF
pCI-TPI-WT-4H +250bp Intron 1	pCI-TPI-WT-4H	PCR product #5743/#5673	Isolated HeLa DNA	AgeI-HF/ SacII
pCI-TPI-WT-4H cellular ATG	pCI-TPI-WT-4H	PCR product #5744/#5673	Isolated HeLa DNA	NheI-HF/ AgeI-HF
pNLA1 [1.2.4.7]	pNLA1	PCR product #1544/#3392	Isolated tat [1.2.4.7] PAA-PCR band from HeLa cells transfected with pNLA1	BssHII/ BamHI-HF
pNLA1 [1.2E.7]	pNLA1	Digested fragment	pNLA1 [1.2.4.7]	SalI/ BamHI
pNLA1 [1.2.4.7] Δ516bp UTR	pNLA1 [1.2.4.7]	Oligo Annealing #5838/#5839	-	BamHI-HF/ KpnI-HF
pNLA1 [1.2.4.7] Δ all in-frame TCs	pNLA1 [1.2.4.7]	Gene Strand I	-	MfeI-HF/ XhoI
pNLA1 [1.2.4.7] RRE	pNLA1 [1.2.4.7]	Digested fragment	pNLA1 env0 RRE	XhoI/ KpnI-HF
pNLA1 [1.2.4.7] RSE (478nt distance to TC)	pNLA1 [1.2.4.7]	Gene strand II	-	XhoI/ KpnI-HF
SV40-RL-β-Globin-WT-ΔATG - SV40-FL$_{ref}$	SV40-RL-β-Globin-WT - SV40-FL$_{ref}$	PCR product #5925/#5814	SV40-RL-β-Globin-WT - SV40-FL$_{ref}$	NheI-HF/ XhoI
SV40-RL-β-Globin-WT-Pyr.-3 - SV40-FL$_{ref}$	SV40-RL-β-Globin-WT - SV40-FL$_{ref}$	PCR product #5926/#5814	SV40-RL-β-Globin-WT - SV40-FL$_{ref}$	NheI-HF/ XhoI
SV40-RL-β-Globin-NS39-ΔATG - SV40-FL$_{ref}$	SV40-RL-β-Globin-NS39 - SV40-FL$_{ref}$	PCR product #5925/#5814	SV40-RL-β-Globin-WT - SV40-FL$_{ref}$	NheI-HF/ XhoI
SV40-RL-β-Globin-NS39-Pyr.-3 - SV40-FL$_{ref}$	SV40-RL-β-Globin-NS39 - SV40-FL$_{ref}$	PCR product #5926/#5814	SV40-RL-β-Globin-WT - SV40-FL$_{ref}$	NheI-HF/ XhoI
pNLA1 [1.2.4.7] RSE (50nt distance to TC)	pNLA1 [1.2.4.7]	PCR product #5931/#5932	pNLA1 [1.2.4.7] RSE (478nt distance to TC)	BamHI-HF/ BclI
IC-1 SV40-ΔELP-RL-β-Globin-WT-ΔATG - SV40-FL$_{ref}$	SV40-RL-β-Globin-WT-ΔATG - SV40-FL$_{ref}$	PCR product #6096/#6097	SV40-RL-β-Globin-WT-ΔATG - SV40-FL$_{ref}$	KpnI-HF/ NheI-HF

IC-1 SV40-ΔELP-RL-β-Globin-NS39-ΔATG - SV40-FL$_{ref}$	SV40-RL-β-Globin-NS39-ΔATG - SV40-FL$_{ref}$	PCR product #6096/#6097	SV40-RL-β-Globin-WT-ΔATG - SV40-FL$_{ref}$	KpnI-HF/ NheI-HF
IC-1 SV40-ΔELP-RL-β-Globin-WT-ATG - SV40-FL$_{ref}$	SV40-RL-β-Globin-WT - SV40-FL$_{ref}$	PCR product #6096/#6097	SV40-RL-β-Globin-WT-ΔATG - SV40-FL$_{ref}$	KpnI-HF/ NheI-HF
IC-1 SV40-ΔELP-RL-β-Globin-NS39-ATG - SV40-FL$_{ref}$	SV40-RL-β-Globin-NS39 - SV40-FL$_{ref}$	PCR product #6096/#6097	SV40-RL-β-Globin-WT-ΔATG - SV40-FL$_{ref}$	KpnI-HF/ NheI-HF
IC-1 SV40-ΔELP-RL-β-Globin-WT-Pyr.-3 - SV40-FL$_{ref}$	SV40-RL-β-Globin-WT-Pyr.-3 - SV40-FL$_{ref}$	PCR product #6096/#6097	SV40-RL-β-Globin-WT-ΔATG - SV40-FL$_{ref}$	KpnI-HF/ NheI-HF
IC-1 SV40-ΔELP-RL-β-Globin-NS39-Pyr.-3 - SV40-FL$_{ref}$	SV40-RL-β-Globin-NS39-Pyr.-3 - SV40-FL$_{ref}$	PCR product #6096/#6097	SV40-RL-β-Globin-WT-ΔATG - SV40-FL$_{ref}$	KpnI-HF/ NheI-HF
IC-2 SV40-ΔELP-RL-β-Globin-WT-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	IC-1 SV40-ΔELP-RL-β-Globin-WT-ΔATG - SV40-FL$_{ref}$	Gene Strand III	-	NheI-HF/ XhoI
IC-2 SV40-ΔELP-RL-β-Globin-NS39-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	IC-1 SV40-ΔELP-RL-β-Globin-NS39-ΔATG - SV40-FL$_{ref}$	Gene Strand III	-	NheI-HF/ XhoI
IC-2 SV40-ΔELP-RL-β-Globin-WT-ATG ΔRL-ATGs - SV40-FL$_{ref}$	IC-2 SV40-ΔELP-RL-β-Globin-WT-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	PCR product #6101/#5632	IC-2 SV40-ΔELP-RL-β-Globin-WT-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	NheI-HF/ XhoI
IC-2 SV40-ΔELP-RL-β-Globin-NS39-ATG ΔRL-ATGs - SV40-FL$_{ref}$	IC-2 SV40-ΔELP-RL-β-Globin-NS39-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	PCR product #6101/#5632	IC-2 SV40-ΔELP-RL-β-Globin-WT-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	NheI-HF/ XhoI
IC-2 SV40-ΔELP-RL-β-Globin-WT-Pyr.-3 ΔRL-ATGs - SV40-FL$_{ref}$	IC-2 SV40-ΔELP-RL-β-Globin-WT-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	PCR product #6129/#5632	IC-2 SV40-ΔELP-RL-β-Globin-WT-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	NheI-HF/ XhoI
IC-2 SV40-ΔELP-RL-β-Globin-NS39-Pyr.-3 ΔRL-ATGs - SV40-FL$_{ref}$	IC-2 SV40-ΔELP-RL-β-Globin-NS39-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	PCR product #6129/#5632	IC-2 SV40-ΔELP-RL-β-Globin-WT-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	NheI-HF/ XhoI
IC-3 SV40-ΔELP-RL-β-Globin-WT-ΔATG ΔRL-ATGs Δβ-Globin ATGs exon 1&2 - SV40-FL$_{ref}$	IC-2 SV40-ΔELP-RL-β-Globin-WT-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	PCR product #6071/#6072	PCR products #6068/#6069 and #6070/#6072 IC-2 SV40-ΔELP-RL-β-Globin-WT-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	XhoI/ EcoRI-HF
IC-3 SV40-ΔELP-RL-β-Globin-NS39-ΔATG ΔRL-ATGs Δβ-Globin ATGs exon 1&2 - SV40-FL$_{ref}$	IC-2 SV40-ΔELP-RL-β-Globin-NS39-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	PCR product #6071/#6072	PCR products #6068/#6069 and #6070/#6072 IC-2 SV40-ΔELP-RL-β-Globin-NS39-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	XhoI/ EcoRI-HF
IC-3 SV40-ΔELP-RL-β-Globin-WT-Pyr.-3 ΔRL-ATGs Δβ-Globin ATGs	IC-2 SV40-ΔELP-RL-β-Globin-WT-Pyr.-3 ΔRL-ATGs - SV40-FL$_{ref}$	PCR product #6071/#6072	PCR products #6068/#6069 and #6070/#6072	XhoI/ EcoRI-HF

exon 1&2 - SV40-FL$_{ref}$			IC-2 SV40-ΔELP-RL-β-Globin-WT-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	
IC-3 SV40-ΔELP-RL-β-Globin-NS39-Pyr.-3 ΔRL-ATGs Δβ-Globin ATGs exon 1&2 - SV40-FL$_{ref}$	IC-2 SV40-ΔELP-RL-β-Globin-WT-Pyr.-3 ΔRL-ATGs - SV40-FL$_{ref}$	PCR product #6071/#6072	PCR products #6068/#6069 and #6070/#6072 IC-2 SV40-ΔELP-RL-β-Globin-NS39-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	XhoI/ EcoRI-HF
IC-3 SV40-ΔELP-RL-β-Globin-WT-ATG ΔRL-ATGs Δβ-Globin ATGs exon 1&2 - SV40-FL$_{ref}$	IC-2 SV40-ΔELP-RL-β-Globin-WT-ATG ΔRL-ATGs - SV40-FL$_{ref}$	PCR product #6071/#6072	PCR products #6068/#6069 and #6070/#6072 IC-2 SV40-ΔELP-RL-β-Globin-WT-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	XhoI/ EcoRI-HF
IC-3 SV40-ΔELP-RL-β-Globin-NS39-ATG ΔRL-ATGs Δβ-Globin ATGs exon 1&2 - SV40-FL$_{ref}$	IC-2 SV40-ΔELP-RL-β-Globin-WT-ATG ΔRL-ATGs - SV40-FL$_{ref}$	PCR product #6071/#6072	PCR products #6068/#6069 and #6070/#6072 IC-2 SV40-ΔELP-RL-β-Globin-NS39-ΔATG ΔRL-ATGs - SV40-FL$_{ref}$	XhoI/ EcoRI-HF
SV40-ΔELP-RL-β-Globin-WT-ΔallATGs - SV40-FL$_{ref}$	IC-3 SV40-ΔELP-RL-β-Globin-WT-ΔATG ΔRL-ATGs Δβ-Globin ATGs exon 1&2 - SV40-FL$_{ref}$	Gene Strand IV	-	EcoRI/ BglII
SV40-ΔELP-RL-β-Globin-NS39-ΔallATGs - SV40-FL$_{ref}$	IC-3 SV40-ΔELP-RL-β-Globin-NS39-ΔATG ΔRL-ATGs Δβ-Globin ATGs exon 1&2 - SV40-FL$_{ref}$	Gene Strand IV	-	EcoRI/ BglII
SV40-ΔELP-RL-β-Globin-WT-ATG ΔallotherATGs - SV40-FL$_{ref}$	IC-3 SV40-ΔELP-RL-β-Globin-WT-ATG ΔRL-ATGs Δβ-Globin ATGs exon 1&2 - SV40-FL$_{ref}$	Gene Strand IV	-	EcoRI/ BglII
SV40-ΔELP-RL-β-Globin-NS39-ATG ΔallotherATGs - SV40-FL$_{ref}$	IC-3 SV40-ΔELP-RL-β-Globin-NS39-ATG ΔRL-ATGs Δβ-Globin ATGs exon 1&2 - SV40-FL$_{ref}$	Gene Strand IV	-	EcoRI/ BglII
SV40-ΔELP-RL-β-Globin-WT-Pyr.-3 ΔallotherATGs - SV40-FL$_{ref}$	IC-3 SV40-ΔELP-RL-β-Globin-WT-Pyr.-3 ΔRL-ATGs Δβ-Globin ATGs exon 1&2 - SV40-FL$_{ref}$	Gene Strand IV	-	EcoRI/ BglII
SV40-ΔELP-RL-β-Globin-NS39-Pyr.-3 ΔallotherATGs - SV40-FL$_{ref}$	IC-3 SV40-ΔELP-RL-β-Globin-NS39-Pyr.-3 ΔRL-ATGs Δβ-Globin ATGs exon 1&2 - SV40-FL$_{ref}$	Gene Strand IV	-	EcoRI/ BglII
SV40-RL-β-Globin-WT-ΔallATGs - SV40-FL$_{ref}$	SV40-ΔELP-RL-β-Globin-WT-ΔallATGs - SV40-FL$_{ref}$	Digested fragment	SV40-RL-β-Globin-WT - SV40-FL$_{ref}$	KpnI-HF/ NheI-HF
SV40-RL-β-Globin-NS39-ΔallATGs - SV40-FL$_{ref}$	SV40-ΔELP-RL-β-Globin-NS39-ΔallATGs - SV40-FL$_{ref}$	Digested fragment	SV40-RL-β-Globin-WT - SV40-FL$_{ref}$	KpnI-HF/ NheI-HF
SV40-RL-β-Globin-WT-ATG ΔallotherATGs - SV40-FL$_{ref}$	SV40-ΔELP-RL-β-Globin-WT-ATG ΔallotherATGs - SV40-FL$_{ref}$	Digested fragment	SV40-RL-β-Globin-WT - SV40-FL$_{ref}$	KpnI-HF/ NheI-HF
SV40-RL-β-Globin-NS39-ATG	SV40-ΔELP-RL-β-Globin-NS39-ATG	Digested fragment	SV40-RL-β-Globin-WT - SV40-FL$_{ref}$	KpnI-HF/ NheI-HF

ΔallotherATGs - SV40-FL$_{ref}$	ΔallotherATGs - SV40-FL$_{ref}$			
SV40-RL-β-Globin-WT-Pyr.-3 ΔallotherATGs - SV40-FL$_{ref}$	SV40-ΔELP-RL-β-Globin-WT-Pyr.-3 ΔallotherATGs - SV40-FL$_{ref}$	Digested fragment	SV40-RL-β-Globin-WT - SV40-FL$_{ref}$	KpnI-HF/ NheI-HF
SV40-RL-β-Globin-NS39-Pyr.-3 ΔallotherATGs - SV40-FL$_{ref}$	SV40-ΔELP-RL-β-Globin-NS39-Pyr.-3 ΔallotherATGs - SV40-FL$_{ref}$	Digested fragment	SV40-RL-β-Globin-WT - SV40-FL$_{ref}$	KpnI-HF/ NheI-HF
IC puc2CL7wo new MCS ΔPsiI/NotI	Puc2CL7EGwo	Oligo Annealing #6172/#6173	-	PstI/ NotI-HF
Puc2CL7wo SV40-RL-β-Globin-WT(-) SV40-FL$_{ref}$(-)	IC puc2CL7wo new MCS ΔPsiI/NotI	Digested fragment	SV40-RL-β-Globin-WT SV40-FL$_{ref}$	KpnI-HF/ ClaI
Puc2CL7wo SV40-RL-β-Globin-NS39(-) SV40-FL$_{ref}$(-)	IC puc2CL7wo new MCS ΔPsiI/NotI	Digested fragment	SV40-RL-β-Globin-NS39 SV40-FL$_{ref}$	KpnI-HF/ ClaI
IC puc2CL7EGwo with NheI site	Puc2CL7EGwo	Oligo Annealing #6073/#6074	-	KpnI-HF/ AgeI-HF
Puc2CL7EGwo TPI-WT-4H mod.	IC puc2CL7EGwo with NheI site	Digested fragment	pCI-TPI-WT-4H Exon 2 ΔHZ 35.3 + Exon 3 HBS 20.8	NheI-HF/ NotI-HF
Puc2CL7EGwo TPI-PTC40-4H mod.	IC puc2CL7EGwo with NheI site	Digested fragment	pCI-TPI-PTC40-4H Exon 2 ΔHZ 35.3 + Exon 3 HBS 20.8	NheI-HF/ NotI-HF

Table 6.2: Numbers and sequences of the used cloning primers. The combination of each primer pair that was used for cloning is documented in Table 6.1.

Primer	Sequence 5' → 3'
#1544	CTTGAAAGCGAAAGTAAAGC
#3392	CGTCCCAGATAAGTGCTAAGG
#5607	GGTTGCTCCCTGGAGAATG
#5632	GTAGGGAGATTATGAATATGC
#5665	CTCCTTGCAGGTTGCC
#5672	GAGTCAGATGAGGTAAGTAGCCAAG
#5673	CCCAAAGACATGCCTTCTCTCTG
#5682	GGAGAGATCAGGTGAGATCTAGGCGGAGAG
#5683	TCTCCGCCTAGATCTCACCTGATCTCTCCAGTAAA
#5743	GAGCTCATCGGCACTCT
#5744	CTATAGGCTAGCGCCTCGGCTCCAGCACCATGGCGGAGGACGGC
#5814	GTGTCTCGAGGTTGTTCATTTTTG
#5838	GATCCTTATATCTCGAGACCTCAGGTAC
#5839	CTGAGGTCTCGAGATATAAG
#5925	GCTTGCTAGCACCTTGACTTCGAAAGTTTATGATCCAGA
#5926	GCTTGCTAGCCCCATGACTTCGAAAGTTTATGATC
#5931	AACGGATCCGACAAATTTATAGGGAGGGCC
#5932	TCATGATCATCTAGAGTAAATGCAAAAGCTTCGCGATC
#6068	CAACCTCGAGACACCATAGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGCTGAAGTTGGTGGTGAGGC
#6069	GAGCCTTCACCTTAGGGTTGCCCAGAACAGTATCAGGAGTGGACAGATCC
#6070	AACCCTAAGGTGAAGGCTCTTGGCAAGAAAGTGCTCGGTGCCTTTAGTGATCGCCTGGCTCACCTGG
#6071	CAACCTCGAGACACCATAGT
#6072	CCAGCCACCACTTTCTGA
#6073	CGCGGGCCCGCTAGCGGATCCA
#6074	CCGGTGGATCCGCTAGCGGGCCCGCGGTAC
#6096	GATAGGTACCGCATCTCAATTAG
#6097	AGGTGCTAGCAAGCTGGGGCGGAGAATGG
#6101	GCTTGCTAGCACCATGACTTCGAAAGTTTACGATCCAGAA
#6129	GCTTGCTAGCCCCATGACTTCGAAAGTTTACGATCCAGAA
#6172	GAGGATCGATCGCGCTAGCAGCGGTACCTACGC
#6173	GGCCGCGTAGGTACCGCTGCTAGCGCGATCGATCCTCTGCA

Table 6.3: Sequences of the Gene Strands which were used for cloning. The restriction sites, which were used for cloning are marked by red letters.

Name	Sequence (5' -> 3')
Gene Strand I: pNLA1 mutation of all in-frame TCs (677bp)	TACCAATTGCTATTGTAAAAAGTGTTGCTTTCATTGCCAAGTTTGTTTCATGACAAAAGCCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGAGCTCATCAGAACAGTCAGACTCATCAAGCTTCTCTATCAAAGCAACCCACCTCCCAATCCCGAGGGGACCCGACAGGCCCGAAGGAATACAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTATTGGACGGATCCTTACCACTTATCTGGGACGATCTGCGGAGCCTGTGCCTCTTCAGCTACCACCGCTTGCGAGACTTACTCTTGCTTGTACCGAGGATTGTGGAACTTCTGGGACGCAGGGGGTGGGAAGCCCTCAAATATTGGTGGAATCTCCTACAGTATTGGAGTCAGGAACTACAGAATAGTGCTGTTAACTTGCTCAATGCCACAGCCATATCAGTATCTGAGGGGACAGATAGGGTTATACAAGTATTACAAGCAGCTTATAGAGCTATTCGCCACATACCTAGAAGAATACGACAGGGCTTGGAAAGGATTTTGCTATAAGATGGGTGGCAAGTGGTCAAAAAGTAGTGTGATTGGATGGCCTGCTGTAAGGGAAAGAATGAGACGAGCTGAGCCAGCAGCAGATGGGGTGGGAGCAGTATCTCGAGACC
Gene Strand II: Rous sarcoma stability element (RSE) (433bp)	GTATCTCGAGGACAAATTTATAGGGAGGGCCACTGTTCTCACTGTTGCGCTACATCTGGCTATTCCGCTCAAATGGAAGCCAGACCACACGCCTGTGTGGATTGACCAGTGGCCCCTCCCTGAAGGTAAACTTGTAGCGCTAACGCAATTAGTGGAAAAAGAATTACAGTTAGGACATATAGAACCTTCACTTAGTTGTTGGAACACACCTGTCTTCGTGATCCGGAAGGCTTCCGGGTCTTACCGCTTACTGCATGATTTGCGCGCTGTTAACGCCAAGCTTGTTCCTTTTGGGGCCGTCCAACAGGGGGCGCCAGTTCTCTCCGCGCTCCCGCGTGGCTGGCCCCTGATGGTCTTAGACCTCAAGGATTGCTTCTTTTCTATCCCTCTTGCGGAACAAGATCGCGAAGCTTTTGCATTTACGGTACCTTTA
Gene Strand III: Mutation of all Renilla (RL) ATGs (957bp)	GCTTGCTAGCACCTTGACTTCGAAAGTTTACGATCCAGAACAAAGGAAACGGATCATAACTGGTCCGCAGTGGTGGGCCAGCTGTAAACAATTGAATCTTCTTGATTCATTTATTAATTATTCTGATTCAGAAAAACTTGCAGAAAATCCTGTTATTTTTTTACATCCTAACGCGGCCTCTTCTTATTTATCGCGACATCTTGTGCCACATATTGAGCCAGTAGCGCGGTGTATTATACCAGACCTTATTGGTCTGGGCAAATCAGGCAAATCTGGTATTGGTTCTTATAGGTTACTTGATCATTACAAATATCTTACTGCATCCTTTGAACTTCTTAATTTACCAAAGAAGATCATTTTTGTCGGCCTTGATTGGGGTGCTTGTTTGGCATTTCATTATAGCTCTGAGCATCAAGATAAGATCAAAGCAATAGTTCACGCTGAAAGTGTAGTAGACGTGATTGAATCCTGGGACGATTGGCCTGATATTGAAGAGGATATTGCGTTGATCAAATCTGAAGAAGGAGAAAAACTGATTTTGGAGAATAACTTCTTCGTGGAAACCATCTTGCCATCAAAAATCCTGAGAAAGTTAGAACCAGAAGAATTTGCAGCATATCTTGAACCATTCAAAGAGAAAGGTGAAGTTCGTCGTCCAACATTATCAAGGCCTCGTGAAATCCCGTTAGTAAAAGGTGGTAAACCTGACGTTGTACAAATTGTTAGGAATTATATTGCTTATCTACGTGCAAGTGATCATTTACCAAAAATCTTTATTGAATCGGACCCAGGATTCTTTTCCATTGCTATTGTTGAAGGTGCCAAGAAGTTTCCTAATACTGAATTTGTCAAAGTAAAAGGTCTTCATTTTTCGCAAGAAGACGCACCTGATCAATTGGGAAAATATATCAAATCGTTCGTTGAGCGAGTTCTCAAAACTGAACAACCTCGAGACAC
Gene Strand IV: Mutation of all ATGs downstream of β-Globin Exon 2 (370bp)	CAAAGAATTCACCCCACCAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTGGCTATTGCCCTGGCCCACAAGTATCACTAATCTAGAGTCGACCCGGGCGGCCGCTTCCCTTTAGTGAGGGTTAATCCTTCGAGCAGACTTGATAAGATACATTGATCAGTTTGGACAAACCACAACTAGACTGCAGTGAAAAAAATCCTTTATTTGTGAAATTTGTGCTGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGATGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAAGATCTGCTG

7 Publications

- Xu, H. C., R. Wang, P. V. Shinde, **L. Walotka**, A. Huang, G. Poschmann, J. Huang, W. Liu, K. Stuhler, H. Schaal, A. Bergthaler, A. A. Pandyra, C. Hardt, K. S. Lang, and P. A. Lang. 2021. 'Slow viral propagation during initial phase of infection leads to viral persistence in mice', *Commun Biol*, 4: 508.

- Müller, L., M. Andree, W. Moskorz, I. Drexler, **L. Walotka**, R. Grothmann, J. Ptok, J. Hillebrandt, A. Ritchie, D. Rabl, P. N. Ostermann, R. Robitzsch, S. Hauka, A. Walker, C. Menne, R. Grutza, J. Timm, O. Adams, and H. Schaal. 2021. 'Age-dependent immune response to the Biontech/Pfizer BNT162b2 COVID-19 vaccination', *Clin Infect Dis*, ciab381.

- Erkelenz, S., S. Theiss, W. Kaisers, J. Ptok, **L. Walotka**, L. Muller, F. Hillebrand, A. L. Brillen, M. Sladek, and H. Schaal. 2018. 'Ranking noncanonical 5' splice site usage by genome-wide RNA-seq analysis and splicing reporter assays', *Genome Res*, 28: 1826-40.

- Brillen, A. L., **L. Walotka**, F. Hillebrand, L. Müller, M. Widera, S. Theiss, and H. Schaal. 2017. 'Analysis of Competing HIV-1 Splice Donor Sites Uncovers a Tight Cluster of Splicing Regulatory Elements within Exon 2/2b', *J Virol*, 91.

- Brillen, A. L., K. Schöneweis, **L. Walotka**, L. Hartmann, L. Müller, J. Ptok, W. Kaisers, G. Poschmann, K. Stuhler, E. Buratti, S. Theiss, and H. Schaal. 2017. 'Succession of splicing regulatory elements determines cryptic 5ss functionality', *Nucleic Acids Res*, 45: 4202-16.

8 Erklärung

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf" erstellt worden ist. Diese Arbeit wurde bei keiner anderen Fakultät vorgelegt. Weiterhin habe ich bisher keine erfolglosen oder erfolgreichen Promotionsversuche unternommen.

Lara Andrea Walotka

Düsseldorf, den

9 Danksagung

An dieser Stelle möchte ich allen danken, die mir ermöglicht haben, diese Forschungsarbeit durchzuführen und durch ihre Unterstützung dafür gesorgt haben, dass diese Doktorarbeit nun endlich abgeschlossen, geschrieben und gedruckt werden konnte.

An aller erster Stelle möchte ich mich bei Prof. Dr. Heiner Schaal bedanken. Dafür, dass ich meine Doktorarbeit in seiner Arbeitsgruppe durchführen konnte und er immer mit Freude und Interesse bei der Sache war, immer ein offenes Ohr für Fragen hatte und nie um Ideen und Anregungen verlegen war. Danke, dass du immer für mich da warst – auch wenn es mal schwierig wurde.

Außerdem möchte ich mich herzlich bei der Jürgen Manchot Stiftung bedanken, deren Stipendium es mir ermöglich hat, in meiner Wunscharbeitsgruppe zu dem von mir gewähltem Thema zu promovieren. Danke auch dafür, dass ich als assoziiertes Mitglied bei der dritten Runde der Jürgen Manchot Graduiertenschule „molecules of infection" teilnehmen durfte.

Weiterhin möchte ich mich noch herzlich bei Prof. Dr. Michael Feldbrügge dafür bedanken, dass er meine Promotion als Mentor begleitet hat. Danke auch an Prof. Dr. Jörg Timm, dass ich meine Forschung am Institut für Virologie durchführen konnte.

Auch gilt mein großer Dank meinen Arbeitskollegen und Freunden, ohne die die Zeit im Labor und Institut bei weitem nicht so schön gewesen wäre. Danke für die vielen schönen Stunden im Großraumlabor, wo wir zusammen laut WDR4 mitgesungen haben. Danke für die vielen hilfreichen Diskussionen und Gespräche zu jeder Tageszeit, sowie die vielen Gelegenheiten des sozialen Miteinanders in den Büros und im Sozialraum – unter Genuss des ein oder anderen alkoholischen Getränks. Vielen Dank an all meine Lieblingskollegen, die mich immer motiviert haben, morgens voller Freude zur Arbeit zu gehen, egal wie lang die Tage auch geworden sind.

Danke auch an meine Familie und meine Freunde, die mir immer zugehört haben und für mich da waren, auch wenn sie nie wussten, wovon ich rede, sobald es um mein Forschungsthema ging. Danke, dass ihr trotzdem zugehört habt. Danke, dass ihr meine schlechten Launen ertragen und mich aufgemuntert habt, wenn die Frustrationen des Forscherlebens mal wieder zu groß wurden. Danke, dass ihr euch mit mir gefreut habt, wenn es gut lief. Danke, dass ihr mich immer bestärkt und an mich geglaubt habt, wenn ich es selbst nicht mehr konnte.