SUMO-modification of the RNA-binding protein La enhances its binding to the translational start site of cyclin D1

Dissertation

Zur Erlangung der Würde des Doktors der Naturwissenschaften

des Fachbereiches Biologie, der Fakultät für Mathematik, Informatik und Naturwissenschaften, der Universität Hamburg

vorgelegt von Julia Kühnert

aus Berlin

Hamburg, 2012

Genehmigt vom Fachbereich Biologie

der Fakultät für Mathematik, Informatik und Naturwissenschaften

an der Universität Hamburg

auf Antrag von Professor Dr. T. DOBNER

Weiter Gutachter der Dissertation:

Priv.-Doz. Dr. D. WARNECKE

Tag der Disputation: 28.09.2012

Hamburg, den 14.09.2012
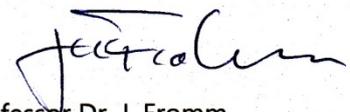
Professor Dr. J. Fromm

Vorsitzender des Promotionsausschusses
Biologie

Table of Contents

i

List of abbreviations

6-FAM	6-carboxyfluorescein
aa	amino acid
ADAR1	adenosine deaminase 1
APS	ammonium persulfate
ATP	adenosin-5'-triphosphate
ATP-γ-S	adenosine-5′-O-(3-thio)triphosphate
AUF1	AU-rich element RNA-binding protein 1
bp	base pairs
BSA	bovine serum albumin
CCND1	cyclin D1
cDNA	complimentary DNA; DNA copy without introns
CK2	creatin kinase II
cpm	counts per minute
CTD	C'-terminal domain
CTP	cytidine triphosphate
DNA	deoxyribonucleic acid
dNTP(s)	2'-deoxyribonucleotide-5'-triphosphate(s)
dsRNA	double stranded RNA
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
F	phenylalanine residue
FL	full length

GA	glycine alanine
GG	glycine glycine, digylcine motif
GST	glutathione S-transferase
FP	fluorescence polarization
HBV	hepatitis B virus
HCV	hepatitis C virus
HIV	human immunodeficiency virus
His	hexa histidine
hLa	human La
hnRNP	heterogenous nuclear protein
IPTG	isopropyl-1-thio-□-D-galactopyranoside
IRES	internal ribosomal entry site
ITAF	IRES *trans* acting factor
IVSA	*in vitro* sumoylation assay
K	lysine residue
K_D	dissociation constant
kDa	kilodalton
LAM	La motif
LB	lysogeny broth
Mdm2	murine double mutant 2
min	minutes
mRNA	messenger RNA
NCR	non coding region
Ni-NTA	nickel-nitriloacetic acid
NLS	nuclear localization signal
NMR	nuclear magnetic resonance
nt(s)	nucleotide(s)

NTD	N'-terminal domain
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PM	protein marker
PTM	posttranslational modification
R	arginine residue
RCD	RNA-chaperone domain
RNA	ribonucleic acid
RNP	ribonucleoprotein
RNP-1/-2	ribonucleoprotein consensus sequence -1/-2
SAE1/2	SUMO activating enzyme ½
Sam68	Src associated associated in mitosis, 68 kDa
SDS	sodium dodecyl sulfate
SENP	sentrin specific proteases
ssDNA	single stranded DNA
ssRNA	single stranded RNA
SUMO	small ubiquitin like modifier
T	threonine residue
TAE	tris base, acetic acid and EDTA
TBE	tris/borate/EDTA
TEMED	N,N,N',N'-Tetramethylethylenediamine
TOP	terminal oligopyrimidine
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
U	units
Ubc9	ubiquitin carrier protein 9

UTR	untranslated region
V	valine residue
v/v	volume per volume
WT	wild type
w/v	weight per volume
XIAP	X-linked inhibitor of apoptosis

Zusammenfassung

Das essentielle, humane La Protein (hLa) stimuliert die interne ribosomale Eintrittstelle (IRES, engl. internal ribosomal entry site)- Element vermittelte Translation viraler und zellulärer mRNAs. Die IRES-vermittelte Translation ermöglicht der kleinen Ribosomenuntereinheit eine direkte Bindung an den 5'-nichttranslatierten Bereich (5'-UTR, engl. untranslated region) bestimmter mRNAs. Die Aktivierung dieses Translationsmechanismusses findet vor allem unter Zellstress und in Krebszellen statt, wenn die 5'-Cap-abhängige Translation beeinträchtigt ist. IRES-transaktivierende Faktoren (ITAFs) werden oft für diese Translationsform benötigt. Humanes La wurde als erster zellulärer ITAF identifiziert und stimuliert die IRES-Translation des Zellzyklus regulierenden und kooperativen Onkogens cyclin D1 (CCND1) in HeLa-Zellen. Das hLa-Protein ist in verschiedenen Krebsarten überexprimiert und wird durch SUMO (small ubiquitin like modifier) posttranslationell modifiziert. Die genannten Erkenntnisse unterstützen die Hypothese, dass das hLa Protein mit dem CCND1-IRES-Element interagiert und diese Interaktion durch SUMO-Modifizierung reguliert wird. Zur Überprüfung dieser Hypothese wurden Experimente zur Identifizierung sowohl der Bindungsstelle des La Proteins in der 5'-UTR der CCND1 mRNA als auch der erforderlichen Bindungs-domänen. Zusätzlich wurden Untersuchungen zur Modulation der RNA-Bindungsaktivität des hLa-Proteins durch SUMO-Modifikation durchgeführt.

In dieser Studie wurde gezeigt, dass das hLa-Protein an oder in der Nähe der CCND1 Translationstartstelle bindet. RNA-Bindungsstudien ergaben, dass sowohl das RNA-Erkennungsmotiv 1 (RRM1) als auch RRM2 für die Vermittlung der CCND1 RNA-Bindung verantwortlich sind. Basische und aromatische Aminosäuren im C'-Terminus des La Proteins beeinflussen vermutlich die RNA-Bindungsaktivität negativ. Anhand eines etablierten *in vitro*-Sumoylierungsprotokolls für hLa wurden die Lysine 200 und 208 als Modifikationsstellen identifiziert und es konnte gezeigt werden, dass die RNA-Bindungsaktivität von hLa durch SUMO-Modifizierung gesteigert wird.

Die molekularen Erkenntnisse dieser Studie ermöglichen zukünftige Experimente zur Untersuchung der La-stimulierten IRES-vermittelten Translation. Des Weiteren kann die Modifizierung von hLa durch SUMO als potentieller neuer Mechanismus zur Regulation der aberranten CCND1-Expression in SUMO-La-überexprimierten Zellen untersucht werden.

Summary

The essential RNA-binding protein La is implicated in stimulating internal ribosome entry site (IRES)-mediated translation of cellular mRNAs. This type of translation enables the small ribosomal subunit (40S) to directly bind to the 5'-untranslated region (5'-UTR) of certain mRNAs. This mechanism is often activated in stressed and cancerous cells when cap-dependent translation is compromised. IRES *trans*-acting factors (ITAFs) have been shown to facilitate IRES-mediated translation. The human La (hLa) protein was the first cellular ITAF described and is known to stimulate the IRES-mediated translation of cell cycle progression regulator and cooperative oncogene cyclin D1 (CCND1) in HeLa cells. The hLa protein is overexpressed in various kinds of cancer, and is posttranslational modified by the small ubiquitin like modifier (SUMO). These findings support the hypothesis that the hLa protein interacts with the CCND1 IRES and may be regulated by SUMO-modification. To test this hypothesis experiments where designed to map the hLa binding site within the 5'-UTR of CCND1 mRNA, to identify minimal hLa domains required for binding, and to demonstrate that SUMO-modification modulates the RNA-binding activity of hLa.

Herein, it is shown that the hLa protein binds at or in close proximity to the translational start site of CCND1. The binding affinities for a number of La mutants were determined by electrophoretic mobility shift assays and fluorescence polarization, leading to the conclusion that both RRM1 and RRM2 of hLa are required for binding and that basic amino acids located in the C'-terminus have a negative effect on the RNA-binding activity. An *in vitro* sumoylation protocol for hLa was established, and utilized in identifying of lysine 200 and 208, which are located between RRM1 and RRM2, as SUMO-acceptor sites of hLa. The RNA-binding activity of hLa was enhanced upon SUMO-modification.

In conclusion, this study provides the molecular knowledge to aid future experiments into whether RRM1-RRM2-mediated binding of hLa to the CCND1 translational start site is critical for stimulating translational initiation. More importantly, it will aid in the investigation whether SUMO-modification of hLa is a novel mechanism by which the expression of the cooperative oncogene CCND1 is aberrantly regulated in cells overexpressing hLa.

1. Introduction

1.1 The La protein

The human La protein (hLa) was first described in the mid-1970's as an autoreactive antigen in serum samples from patients with systemic lupus erythematosus (SLE) and Sjogren's syndrome [2] [3], however, it was not cloned until 1985 [4]. Mammalian La is an essential protein, its knockout in mice results in embryogenic lethality [5]. The gene for the hLa autoantigen encodes at least two alternative mRNA isoforms, the 1.9 kb La1 and the 2.3 kb La1', which are both suggested to form functional La mRNAs [6] [7] [8]. The mRNA sequence contains 11 exons including a presumed G/C-rich (guanine/cytosine-rich) promoter upstream of the mRNA start site [9]. The only difference between those two mRNA isoforms are their 5'-UTRs, whereas the La1 5'-UTR contains 115 nucleotides the La1' mRNA contains 483 nucleotides, approximately 4-fold longer [8]. Further, the expression of those two isoforms are diverse, the predominant La1 transcripts are ubiquitously expressed in human tissues, however, La1' is expressed at a low level and tissue-specific, it is restricted to peripheral blood leukocytes, particularly to B, T, and natural killer cells [8]. Interestingly, the 5'-UTR of the La mRNA contains an internal ribosomal entry site (IRES) [8], which enables hLa expression under conditions when cap-dependent translation is compromised.

The La cDNA translates into a 408 amino acid 47 kDa RNA-binding protein, which is involved in many different aspects of the RNA metabolism and suggested to be aberrantly regulated in cancer [10] [11]. RNA-binding proteins are characterized by containing one or more RNA-binding domains such as the RRM, K homology (KH) domain, DEAD/DEAH box, Piwi/Argonaute/Zwille (PAZ) domain, or Zinc finger [12] [13] [14]. The RNA recognition motif is the most extensively studied RNA-binding domain [15], this consensus domain contains approximately 90 amino acids as determined by characterizing the mRNA polyadenylate binding protein (PABP) and heterogeneous nuclear ribonucleoprotein C (hnRNP C) [16]. As an RNA-binding protein human La contains three RNA-binding domains (figure 1.1): the N'-terminal La motif (LAM, aa 10-91) [17] [18], the canonical RNA recognition motif 1 (RRM1, aa 112-169) [17] [18], and the atypical RRM2 (RRM2, aa 231-327) [19] [17] [18] (reviewed in [20]). Each RRM comprises two ribonucleoprotein consensus sequences (RNP), RNP-1 and RNP-2 [15] [21]. The central RNP-1 is characterized by eight conserved mainly aromatic and positively charged amino acids [16] (reviewed in [22]). The RNP-2 was identified later as an N'-terminal consensus sequence consisting of six

less conserved amino acids [23] (reviewed in [22]). A spacer of approximately 30 amino acids separates both consensus motifs.
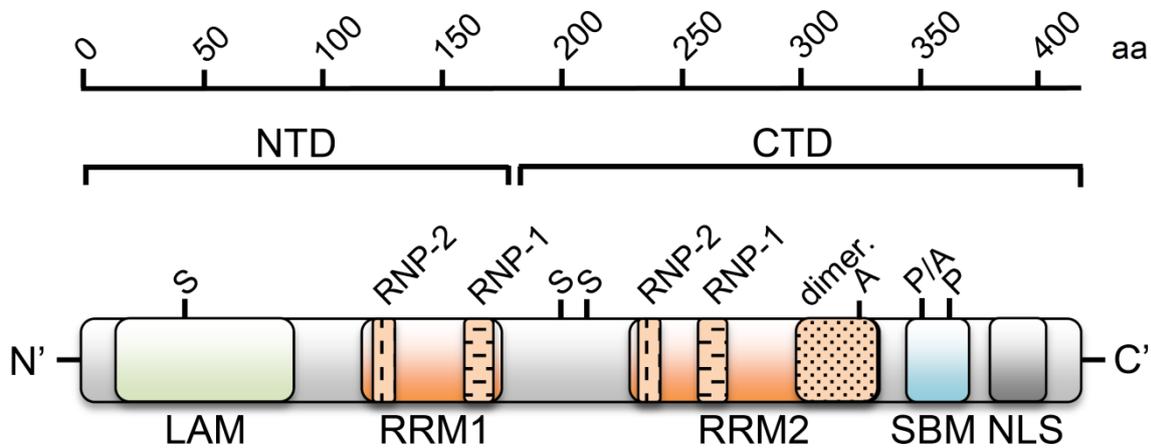
Figure 1.1: Domain organization of the human La protein. The cartoon figure displays selected domains of the hLa protein and posttranslational modification sites. NTD = N'-terminal domain, CTD = C'-terminal domain, LAM= La motif, RRM = RNA recognition motif, RNP = ribonucleoprotein, SBM = short basic motif, NLS= nuclear localization signal, N' = N'-terminus, C' = C'-terminus, S = sumoylation site, dimer = dimerization domain, A = acetylation site, P = phosphorylation site

The overall structure of the RRM folds into an α/β sandwich with a $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ topology [24] [25] [26] [22]. A large number of conserved residues are located in the hydrophobic core of the RRM [25]. The consensus sequences RNP-1 and RNP-2 are located coplanar in β_3 and β_1 respectively [27]. Structural analysis revealed that five conserved residues in the RNP-1 and RNP-2 are exposed to the solvent and directly mediate the RNA-binding [22] [28], whereas the positioning of the RNA at the RRM is mediated by aromatic residues outside of the β sheets [27] [22].

The 90-amino acid long La motif is highly conserved [18] [20], which adopts a winged helix-turn-helix confirmation [29]. A well-conserved aromatic patch on the LAM domain surface mediates RNA-binding and 3'- end recognition [30]. The La motif and the RRM1 are required for binding terminal uridylate RNAs. The co-crystal structure of RNA bound to the N'-terminal domain of La (NTD) reveals that neither the RRM1 nor the LAM are altered by RNA-binding, but residues of the linker polypeptide (aa 112-110) between those two domains form an α helix and the target RNA binds in this cleft [31] [32].

domain name	amino acid position	references
LAM	10-91	Maraia 2001/02
RRM1	112-169	Maraia 2001/02
RRM1, RNP-2	113-118	Maraia 2001/02
RRM1, RNP-1	151-158	Maraia 2001/02
RRM2	231-327	Jacks 2002,Maraia 2001/02
RRM2, RNP-2	235-240	Jacks 2002,Maraia 2001/02
RRM2, RNP-1	268-276	Jacks 2002,Maraia 2001/02
NRE	316-348	Intine 2002
NRE	165-337 and 266-269	Simons 1996
oligomerization	274-291	Craig 1997
oligomerization	298-348	Horke 2002
oligomerization	298-326	this study
NoLS	323-354	Horke 2004
short basic motif	348-368	Topfer 1993 Goodier 1997
WAM	333-339	Topfer 1993 Goodier 1997
NLS	383-400	Simons 1996
phosphorylation sites	S360, S366	Broekhuis 2000, Fan 1997, Choudhary
acetylation sites	K320, K360	Choudhary 2009
sumoylation sites *(in vivo)*	K41	van Niekerk 2007
sumoylation sites *(in vitro)*	K200, K208	this study

Table 1.1: Location of domains and posttranslational modification sites of the hLa protein. The domain names are indicated with their amino acid position and according reference, see text for citation number. LAM = La motif, RRM = RNA recognition motif, RNP = ribonucleoprotein, NRE = nuclear retention element, NoLS = nucleolar localization, WAM = Walker A motif, NLS = nuclear localization signal, S = serine, T = threonine, K = lysine

The C'-terminal domain (CTD) contains a nucleolar localization signal (NoLs, aa 323-354, [33]), a putative multimerization domain ([34] [35]); (aa 298-348 or aa 274-291), respectively, several phosphorylation sites (serine 360 and 366; [36] [37]), lysine acetylation sites (lysine residues 320 and 360) [38], and a short basic motif (aa 348-368) with a putative Walker A nucleotide binding motif (aa 333-339) [39] [40]. Murine La (mLa) is phosphorylated at threonine 309 by Akt [41], whereas the hLa is phosphorylated at threonine 389 by Akt [42]. Additional lysine acetylation sites are located in RRM1 at residues 116 and 128 [38]. An overview of the described domains and functional sites including their position in the human La protein can be found in table 1.1.

The La protein contains a bipartite nuclear localization signal (NLS) at the C'-terminus (aa 383 to 400) [43] which is bound by the nuclear import factor karyopherin, Kap108p [44] [45]. This interaction leads to the transport of La into the cell nucleus [44] [45]. Due to the nuclear retention signal (NRE, aa 316-348 [46] or aa 165-337 [43]) the hLa is mainly confined to the nucleus [47] [48] [49] [20]. The distribution of the La protein within the nucleus is regulated by the cell cycle, La localizes from the nucleoplasm to the nucleolus during late G_1- and early S-phase [50] [49] [51] [52], however, the signal which induces the localization to the nucleolus are not yet established.

Recently the nuclear-cytoplasmic shuttling of La has been documented [53], and certain cellular conditions are leading to an accumulation of La in the cytoplasm. During apoptosis La is enriched in the cytoplasm due to caspase 3-mediated proteolytic cleavage of the C'-terminus including the NLS [54] [55]. In addition, upon infection of cells with poliovirus the C'-terminus of the La protein is proteolytically removed by the poliovirus-specific protease 3C resulting in the localization of the protein to the cytoplasm [56]. A transient redistribution of the protein to the cytoplasm has also been described in herpes-simplex virus 1 infection [57]. The endogenous signals triggering the export of human La to the cytoplasm are still elusive. Phosphorylation and de-phosphorylation of hLa may be a suspected trigger. Human La can be phosphorylated at serine 366 by casein kinase II [37] [58]. Interestingly phosphorylated hLa is restricted to the nucleoplasm whereas the dephosphorylated protein is localized into the cytoplasm and nucleoplasm suggesting that de-phosphorylation triggers this re-distribution to the cytoplasm and into the nucleolus [59]. However, this concept was not confirmed by Broekhuis *et al.* [36], who showed that mutations of phosphorylation sites in La mutants did not affect the subcellular localization of La in *Xenopus laevis* oocytes and Hep2 cells. On the other hand, murine La is phosphorylated by serine-threonine protein kinase (Akt) at threonine 309 in murine glial cells [41]. Those experiments using green fluorescence protein-tagged La demonstrated that wild type La but not a La mutant, which could not be phosphorylated by Akt, shuttles to the cytoplasm [41]. Hence the signals and modifications inducing re-localization of La to the cytoplasm may be different in various cell types.

Phosphorylation is not the only posttranslational modification of the La protein; van Niekerk *et al.* [60] have shown the modification of human and murine La by the small ubiquitin like modifier (SUMO) in rat neuronal cells. The sumoylation at lysine 41 triggers the microtubule-based retrograde transport of the La protein to the nucleus in sensory axons [60]. Further, hLa is posttranslational modified by SUMO-2 in U2OS cells upon heat shock [61].

The oligomerization of the La protein is highly discussed with evidence for the role of higher-order complexes for the functionality of the protein in living cells. The dimerization domain has been mapped to amino acids 293-349 located in the C'-terminal domain (CTD) by Far-Western and the requirement of this domain for stimulating translation of mRNAs *in vitro* [34] [62]. Further, RNA-binding studies revealed that with increasing La concentrations additional complexes besides the primary complexes are formed [47] [63] [64] [40] (this dissertation). Those lower mobility complexes of La are suggested to be a result of homodimerization and not due to binding of more than one protein copy to less specific targets on the RNAs [64] as it has been shown for the poly(A)-binding protein [65]. The slower migrating complexes have been shown to have lower target specificity than the primary La complex by competition experiments *in vitro* [63]. On the other hand, Jacks *et al.* [66] could not confirm the homodimerization by either chemical shift analysis and ^{15}N backbone dynamics or by analytical ultracentrifugation *in vitro*. However, a trans-dominant negative La mutant, inhibited endogenous La functions in HCV IRES-mediated translation [67], which has been identified by overexpression experiments in HeLa cells. Hence, the oligomerization of the La protein still remains elusive and a potential functional role of oligomerization for hLa *in vivo* needs to be established.

Homologs of the human protein have been identified in various other eukaryotes [4] [9] [39] [68], such as *Saccharomyces cerevisie* and *Schizosaccharomyces pombe* [69], *Drosophila melanogaster* [69], and *Xenopus laevis* [70] [71]. Interestingly, in contrast to mammalian La, yeast La homologs are dispensable [72] [73] [74]. The N'-terminal domain including the La motif and the RNA recognition motif 1 (RRM1) is highly conserved, whereas the C'-terminus varies in sequence and length between species [20]. See the figure below (figure 1.2).

The human La protein, also referred to as LARP3, is a member of the La motif (LAM) family, which can be categorized into five families, the genuine La homologs and four La-related proteins (LARP) [28]. LARP proteins are characterized by the La motif followed by an RRM or RRM-like motif [28].
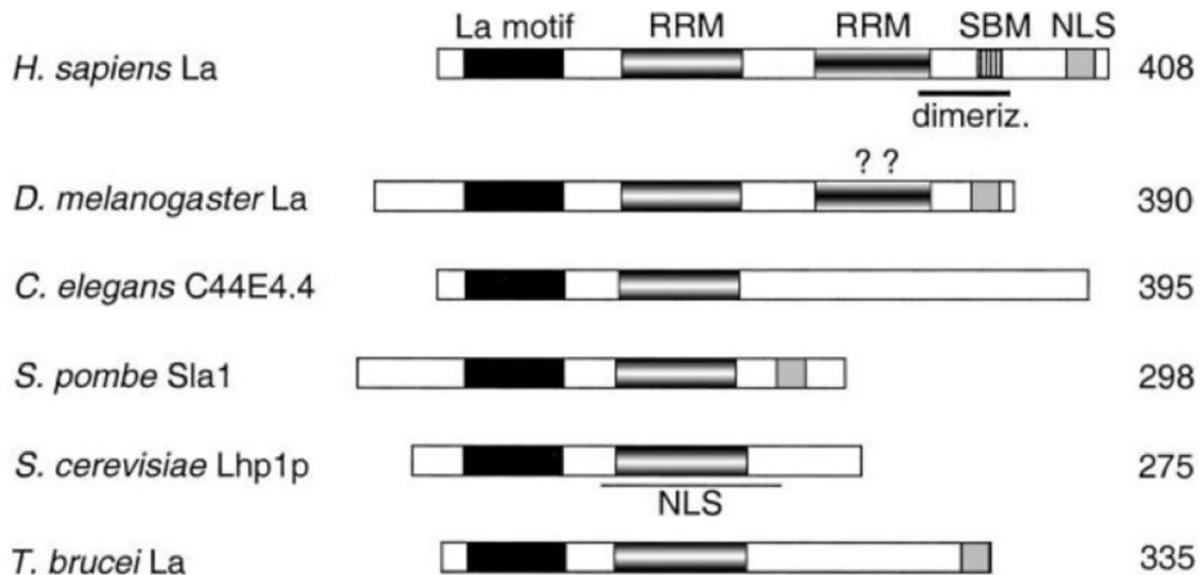
Figure 1.2: Domain organization of La proteins[20]. Alignment of human, *D. melanogaster, C. elegans, S. cervisiiae,* and *T. brucei* domain organization. The location of the short basic motif (SBM) and dimerization domain in the human La protein are indicted. The nuclear localization signals (NLS) are indicated if known. The protein length is indicated on the right. *H. sapiens = Homo sapiens*, *D. melanogaster = Drosophila melanogaster*, *C. elegans = Caenorhabditis elegans*, *S. = Saccharomyces, T. brucei = trypanosoma brucei*

1.1.1 The role of La in the RNA metabolism

The La protein functions in many different aspects of the RNA metabolism by interacting with a variety of different cellular and viral RNAs [18] [20]. The nuclear La protein displays a high affinity binding to the terminal 3'-UUU-OH motif, which is a hallmark of all RNA polymerase III transcripts [75] [20] [76]. Some RNA polymerase II transcripts in *S. cerevisiae* also contain a terminal 3'-UUU-OH motif allowing La binding [77] [78] [74]. The major role of La binding terminal polyuridine RNAs is the protection against 3'- exonuclease digestion [78] [77] [79] [80] [81]. The stabilization of nascent precursor RNAs is especially important for their maturation as shown for pre-tRNAs and U6 small nuclear RNA (snRNA) [82] [83] [20], not only cellular RNAs are stabilized and protected against endo-/ exonucleolytic degradation by the La protein but, it is suggested that viral RNAs such as hepatitis B virus (HBV) RNA [84] [85] and hepatitis C virus (HCV) RNA [86] are also stabilized and protected.

La has been reported to bind cellular [40] [87] [74] [88] [89] and viral mRNAs [86] [90], in addition, to its well-described function in terminal poly(U)-binding. This important protein

influences the translation efficiency of viral and cellular mRNAs, such as of the X-linked inhibitor of apoptosis (XIAP) [91], binding immunoglobulin protein (BiP) [92], 5'-terminal oligopyrimidine (5'-TOP) mRNAs [93], and murine double mutant 2 (Mdm2) [10], laminin B1 [94], and cyclin D1 (CCND1) [89].

Eukaryotic mRNAs are translated by two different mechanisms, the canonical cap-dependent and the internal translation, which is mediated by an internal ribosomal entry site (IRES). During the cap-dependent translation the 43S preinitiation complex is formed at the 5'-methylguanosine end of the mRNA, termed 5'-cap. In contrast, IRES-dependent translation relies on a *cis*-acting structural element comprised in the 5'-untranslated region (5'-UTR) of the mRNA enabling the recruitment of the 43S preinitiation complex independent of the 5'-cap, and often depends on additional factors called IRES *trans*-acting factors (ITAFs) [95]. This cap-independent mechanism has been described to be important under conditions that require the translation of specific mRNAs, such as factors involving apoptosis and stress-response [96]. The human La protein has been described as the first ITAF [97] [98] and has been shown to mainly stimulate IRES-dependent translation of viral mRNAs, e.g. hepatitis C virus (HCV) and poliovirus [99] [67] [100], however, hLa has been shown to inhibit the hepatitis A virus translation [101]. The internal translation via an IRES of cellular mRNAs, XIAP [91] and the cell cycle regulator cyclin D1 (CCND1, [89], has been demonstrated to be stimulated by the hLa protein. The mechanisms behind the stimulation of IRES-dependent translation by the La protein are not yet fully understood. Further, not only is the translation of cellular mRNAs with IRES-elements facilitated by hLa but also the cap-dependent translation of the murine double mutant 2 (Mdm2) [10] and HIV TAR RNA [102] [103], implying the involvement of La also in cap-dependent translation. It has yet to be firmly established whether the presumably cap-dependent translation of a subset of mRNAs containing terminal oligopyrimidine elements in their 5'-UTR, which are encoding ribosomal proteins and translation factors [104] [105] [93], are stimulated by hLa.

Whereas the La recognition motif for small RNAs has been identified as terminal polyuridylate, a consensus motif for La binding RNAs not ending in 3'-UUU-OH or binding sites within in the body of the mRNA has not been identified. However, long stem structures in RNAs lacking a terminal poly(U) motif are contributing to binding by vertebrate La [106] [107]. Ali and colleagues were able to identify the context of an initiator AUG of the HCV RNA as a La binding site [62]. Furthermore, the importance of the translational start site embedding Kozak sequence, a consensus sequence which plays a critical role in the translational initiation of all eukaryotic mRNAs [108] [109] [110], was studied for La-

binding *in vitro* [111]. The La protein was shown to only bind RNA oligoribonucleotides in which the initiation codon was embedded in a strong Kozak sequence context [111]. Binding sites for hLa were mapped in the structured 5'-UTR of XIAP [91] and Mdm2 mRNA [10], which are translated in an IRES-dependent and cap-dependent manner, respectively. In addition, Heise *et al.* [112] showed that La recognizes specifically a putative stem-loop structure in the HBV RNA. Mutational analyses using HCV oligoribonucleotides also suggested a structure driven La-binding [113]. Those findings point to, either sequence-driven, as shown for the HCV IRES [62] and Kozak sequence binding studies [111], or structure-driven binding by La as suggested for XIAP, Mdm2, HBV, and HCV oligoribonucleotides [91] [10] [112] [113].

Not only the main RNA-binding motifs (La motif, RRM1 and RRM2) but also the C'-terminal domain, which is not involved in binding to terminal polyuridylate RNA, of hLa may be important for internal mRNA binding, since it has been shown to recognize internal sequences in structured RNAs such as pre-tRNAs [103] and the IRES of hepatitis C virus [99], Horke *et al*. demonstrated the requirement of both RNP-2 motifs, located in RRM1 and RRM2, for hepatitis B RNA-binding using RNP-2 deletion mutants in native electrophoretic mobility shift assays (EMSAs) [35] suggesting a cooperative binding of RRM1 and RRM2 to HBV RNA.

One major concept to understand the many functions of the La protein is that it may act as an ATP-dependent helicase [114] and/or an RNA-chaperone [115] [116]. Proteins with RNA-chaperone activity either open up misfolded RNA or prevent RNAs from misfolding in an ATP-independent manner [117] [118]. Bachmann *et al.* [114] have reported that the La protein, isolated from cultured murine (NIH3T3) cells and calf thymus, is able to melt synthetic DNA-RNA hybrids, which mimics the previously suggested role of La in RNAPIII transcription termination factor [119] [120] [48]. They further suggested that the energy required for hybrid melting is located, as an ATPase-activity, within the La protein [114]. Other evidences for RNA-chaperone activity of the La protein was made by chemical and enzymatic footprinting experiments in yeast where the yeast La homolog Lhp1p promoted the correct folding of pre-tRNA stems [121]. Further RNA-restructuring activity of La, *cis*-splicing activity, was demonstrated by Belisova and coworkers [115] *in vitro* and in a bacterial assay *in vivo*. The α3 helix in RRM1 of recombinant hLa has been identified to be required for strand annealing *in vitro* as determined for FRET (Förster resonance energy transfer) [116]. However, only the N'-terminal and not the C'-terminal domain of La was analyzed. Intriguingly, the hLa C'-terminal tail is largely intrinsically disordered, i.e. the C'-

terminus lacks a well-defined tertiary structure. It has been shown for other proteins that those intrinsically disordered regions may adopt an ordered structure upon binding to their specific binding partner [122] [123] [124] [125]. More importantly, it has been proposed that intrinsic disordered regions contain RNA-chaperone activity [117] [126]. Nevertheless, the RNA-chaperone function of La has not been firmly proven.

In budding yeast, *Saccharomyces cerevisiae*, the La homolog Lhp1p has been shown to stabilize precursors to splicesomomal U small nucleolar RNAs (snRNAs) as well as the nucleolar U3 RNA [80] [77] [78]. Since Lhp1p also stabilizes the pre-U4 snRNA which is bound by the Smd1p protein, a small nucleolar ribonucleoprotein protein (snRNP) [78], it is proposed that La binding promotes the association of small RNAs into functional RNPs [80] [78] by assisting in RNA folding [78].

1.1.2 La in cancer

The human La protein was suggested to be aberrantly regulated in various types of cancer, e.g. expression of La is elevated in BCR/ABL transformed cells from patients suffering from chronic myeloid leukemia [10]. La stimulates expression of Mdm2 in those cells [10], which is a negative regulator of the tumor suppressor p53. In addition, using tissue microarrays La has been shown to be overexpressed in solid tumors, such as cervical cancer tissue [89] and oral squamous cell carcinomas (SCC) [11], furthermore, La protein levels are shown to be up-regulated in numerous cell lines [127]. The human La protein is with $2x10^7$ copies per cell in mammalian cells as highly abundant as ribosomal proteins [48]. Further, an estimated 50 nM of human La were quantified in human cervical cancer HeLa S100 cell extracts [47].

The Akt-dependent distribution of murine La to the cytoplasm was suggested to contribute to the oncogenic effects of deviant Akt activity in cancerous cells [41]. Other evidences suggest a tumorigenic role of hLa: La contributes to cell proliferation of different types of cell lines, such as cervical, prostate, and oral SCC cells [89] [11]. Further, La promotes migration and cell invasion of oral SCC cells *in vitro* [11], which further supports an important role of hLa in tumorigenesis. The translation of the cell cycle regulator protein cyclin D1 is decreased upon La depletion in different cell types [89] also suggesting a tumorigenic role for hLa. IRES-dependent cyclin D1 expression is stimulated when La is overexpressed and reduced when La levels are decreased [89]. The factors by which La is contributing to tumorigenesis are suggested to increased expression levels of the RNA-binding protein and the stimulation

of translation of specific mRNAs [10] [89] due to the re-localization of the protein to the cytoplasm in cancerous cells [41].

1.2 Cyclin D1

The product of the proto-oncogene cyclin D1 is a member of a highly conserved family of cyclins and displays pro-proliferative functions as a key regulator for the integration of extracellular mitogenic stimuli with cell proliferation [128] [129] [130]. Cyclin D1 (CCND1) activates cyclin dependent kinase (CDK) 4 or 6 to form a CDK4/6-CCND1 complex, this holoenzyme phosphorylates the tumor suppressor retinoblastoma protein (pRb) resulting in the inactivation of its cell-cycle inhibitory function. The pRb protein is critical for G_1 to S phase progression, it silences genes that are synthesized in S-phase and those genes are de-repressed and transcribed upon phosphorylation by CCND1 [131]. Recent studies revealed that CCND1 also displays multiple, CDK-independent functions [132] [133], which play a substantial role in cellular growth, metabolism, and cellular differentiation [133]. Cyclin D1 has been shown to directly bind more than 30 transcription factors or transcriptional co-regulators [134] [135] such as STAT3, C/EBPβ (CCAAT/enhancer binding protein), and the nuclear receptor superfamily, including the estrogen and androgen receptors [136] [137] for CDK-independent functions [138].

Growth factors stimulate the expression of CCND1 through direct induction at its promoter [139]. The CCND1 promoter activity is modulated in a tissue specific manner by sequence specific transcription factors upon stimuli by growth factors and hormones [140] [141]. Once the mRNA is transcribed multiple other regulation levels exist for CCND1, including mRNA stability, mRNA translation, subcellular localization, and degradation [128] [129].

The mRNA of cyclin D1 has a 209 nucleotide long, GC-rich, and presumably structured 5'-UTR. These features in the 5'-UTR of some cellular mRNAs, particularly of cell proliferation and apoptosis key regulators, can impair efficient translation initiation by inhibiting ribosomal scanning [142]. The translation initiation for some of these mRNAs is accomplished by a cap-independent manner, referred to as internal ribosomal entry site (IRES)-mediated translation [143]. IRES-mediated translation appears to be of significant importance under conditions in which the cap-dependent translation pathway is inhibited [144] [145] [96]. Gera and his colleagues have shown that the CCND1 mRNA is translated by this alternative translation mechanism upon inhibition of cap-dependent translation [146] by the Akt signaling pathway inhibitor and tumor suppressor phosphatase and tensin homolog

(PTEN) [147]. Under these conditions they later identified the hnRNP A1 as IRES *trans*-acting factor (ITAF) regulating CCND1 IRES-dependent translation. The minimal sequence in the 5'-UTR enabling IRES-dependent translation spans nucleotides -209 to -44 upstream of the translational start site. The La protein has recently been shown to directly interact with the CCND1 mRNA in HeLa cells [11]. Respectively by overexpression or knockdown experiments it was shown that La stimulates or decreases IRES-dependent translation of CCND1 [11]. Moreover, different cancer cell lines were tested for the CCND1 protein expression and cell proliferation upon La depletion; in some tested cell lines the knockdown of La affected proliferation as well as CCND1 protein expression [11], thus the aberrant abundance of La is proposed to cause an overexpression of CCND1 in cancerous cells by stimulating IRES-dependent translation under certain cellular conditions, such as down regulation or inhibition of cap-dependent translation [11].

A hallmark of several tumor types is the overexpression and/or excessive activity of cyclin D1 [148] [131] [149]. An increased abundance of CCND1 occurs during the early stages of tumorigenesis [150] and coincides with poor prognosis for patients when levels of cyclin A are high [151]. Somatic mutations and rearrangements of the CCND1 gene are rare [152], but oncogenic signals are inducing CCND1 overexpression in most cancer types, such as lung, breast, and colon cancer [133]. Interestingly, CCND1 acts as a collaborative oncogene by enhancing transformation by other oncogenes, including Ras and Src, in tissue cultured cells [153] [154].

1.3 Protein modification by sumoylation

Chemical alterations of proteins following translation to regulate protein stability and protein function are called posttranslational modification (PTM). These usually reversible modifications are essential for cellular processes and represent another level of protein regulation. Well-studied examples for PTM are phosphorylation, methylation [155], and acetylation [156]. The posttranslational modification of proteins by small ubiquitin like modifier (SUMO), another type of PTM, regulates various cellular processes, such as transcription, DNA repair, and chromatin organization [157] [158] [159]. SUMO-modification of a target protein changes its functional properties; it may alter its protein-protein interaction, subcellular localization, transcription factor transactivation, or the activity of the target protein [160] [161] [162] [163] [164] [165].

Four SUMO homologues exist in mammals, designated SUMO-1 through SUMO-4. Despite their name as small ubiquitin like modifiers they only share 18% homology with ubiquitin and are 11 kDa large, 3 kDa larger than ubiquitin (8 kDa) [166]. SUMO-2 and SUMO-3 show 95% of sequence homology, differing from each other by only three N'-terminal residues and often referred to as SUMO-2/-3, both SUMO isoforms are capable of chain formation via internal lysine residues [167]. SUMO-1 only shares 50% homology with SUMO-2/-3 [158] and cannot form polychains [168], but acts as a SUMO-2/-3 chain terminator [169] [170]. SUMO-4 on the other hand has been proposed as SUMO paralogue by DNA sequence analysis, but due to the lack of introns in the SUMO-4 gene it is suggested to be a pseudogene [171].

The covalent conjugation of a member of the SUMO family to the target protein is a multistep process catalyzed by an enzymatic cascade, which is similar but distinct to ubiquitination [172] (figure 1.3), both pathways involve three enzymes, the E1 activating enzyme, the E2 conjugating enzyme, and the E3 ligase [173]. All eukaryotic SUMO proteins are translated as inactive precursors that have to undergo maturation. The mature form is generated by C'-terminal cleavage, catalyzed by a sentrin specific protease (SENP), resulting in the exposure of a di-glycine motif that is required for SUMO to be covalently conjugated to lysine residues in target proteins. In mammalian cells, this C'-terminal di-glycine motif gets adenylated by the SUMO E1 heterodimer SAE1/SAE2 (SUMO activating enzyme) and subsequently attacked by the active-site cysteine residue of SAE2 forming an E1-SUMO thioester [174] [163]. This step requires the hydrolysis of ATP as an energy donor so that the activated SUMO protein is then transferred to the active-site cysteine of the E2 conjugating enzyme Ubc9 forming another thioester linkage [175] [176] [177] [178]. The Ubc9 enzyme can then directly transfer SUMO to the acceptor lysine residue on the target protein, a process that is often facilitated by E3 ligases by acting as scaffolds [168]. The E3 ligase either brings the SUMO-Ubc9 into a complex with the target protein or stimulates the discharge of SUMO from Ubc9 to the substrate protein [170] [172]. A large number of SUMO E3 ligases have been described and assigned to three categories: the protein inhibitor of activated STAT (PIAS) family, the nuclear pore protein Ran binding protein 2 (RanBP2), and the polycomb group protein Pc2 [179] [180] [181]. The highly dynamic process of sumoylation results in rapid de-conjugation, where the isopeptide bond between SUMO and the target protein is cleaved by SENP, the same protease family catalyzing the maturation of pro-SUMO [182]. Seven SENP isopeptidase isoforms are known, SENP1-3 and SENP5-7, with varying cellular distribution and SUMO isoform specificity [161] [172].
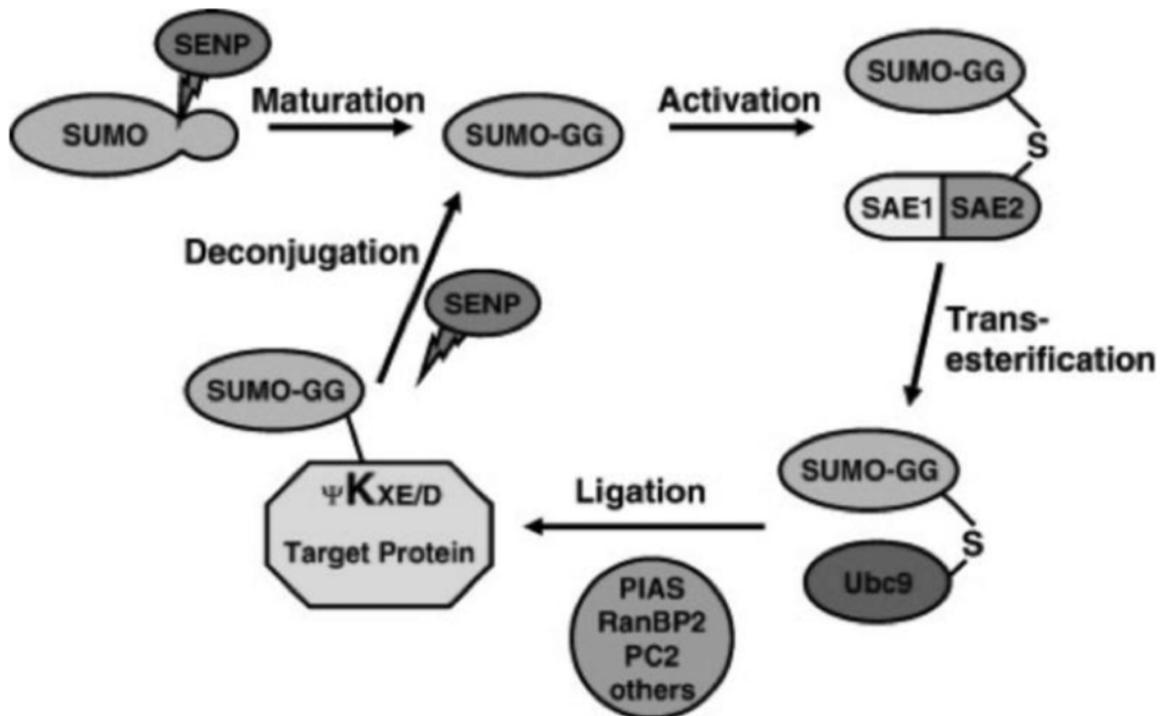
Figure 1.3: Schematic representation of the sumoylation cascade [321]. The maturation of the pre-cursor SUMO is catalyzed by the C'-terminal cleavage by SENP. The matured SUMO protein comprises a C'-terminal di-glycine motif, which gets activated by a heterodimer of the SUMO activating enzymes 1 and 2. The activated SUMO molecule forms a thioester bond with the conjugating enzyme Ubc9, which can directly ligate SUMO to a lysine residue in the target protein. This conjugation process is often facilitated by SUMO ligases, e.g. RanBP2 and PIAS. The rapid de-sumoylation is catalyzed by SENP to provide free SUMO that can undergo another sumoylation cycle. SENP = Sentrin specific proteases, SAE = SUMO activating enzyme, Ubc = ubiquitin conjugating complex

Many sumoylation target proteins contain a lysine residue in the consensus sequence $\psi Kx(D/E)$ with ψ being a hydrophobic amino acid, K a lysine residue, X corresponding to any amino acid, D as an aspartic acid, and E a glutamic acid residue [183] [177] [175] [184] [185]. The consensus sequence residues directly interact with the E2 conjugating enzyme Ubc9 [185] and most SUMO consensus sites are located in intrinsically disordered regions or extended loops [163]. Extended motifs containing additional elements have been identified, such as the phosphorylation-depenent SUMO motif [186], however, SUMO-modification does not only occur at specific lysine residues in consensus sequences, but non-consensus lysines have been described to be sumoylated [187] [188] [189]. Furthermore, some proteins do not even have the consensus motif in their sequence and yet they are modified by SUMO [190] [181].

Sumoylation is an essential process in mammals as knockout and knockdown studies of Ubc9 revealed [157] [191], thus SUMO-modification plays an important role in normal cell

functions [162]. The molecular consequences of target sumoylation results in three effects, which are not exclusive [159]: First, protein surfaces can be altered upon SUMO-modification influencing downstream interactions with other molecules, secondly, the covalent attachment of SUMO may mask or reveal binding sites within a protein that are required for downstream functions, and lastly, the conjugated SUMO can recruit other protein interacting factors to the target protein [172].

Many sumoylation target proteins are implicated in human diseases involving cancer and neurodegenerative diseases such as Alzheimer's, Parkinson's, and Huntigton's disease reviewded in [192]. The sumoylation pathway is implicated in tumorigenesis [193] by its involvement in pathways driving tumor formation such as cell growth, differentiation, senescence, apoptosis, and autophagy [194] [195]. The E2 conjugating enzyme Ubc9 has been shown to be overexpressed in several tumors such as melanoma, lung, and ovarian carcinomas, in addition its overexpression is linked to an increase in cancer cell growth [162] [196] [197]. Up-regulation of SUMO E1 activating enzyme is associated with a lower survival rate in hepatocellular carcinoma patients [198]. Importantly, the activities of the tumor suppressors such as p53, pRB, and Mdm2 are regulated by sumoylation [193]. Unsurprisingly, posttranslational modification by SUMO has been proposed as a target in cancer therapy [199] [197] [200].

Recently, SUMO-modification of a small number of RNA-binding proteins, such as hnRNP C1 (heterogenous nuclear protein C1) [201] and Sam68 (Src associated in mitosis, 68 kDa) [202], has been described. However, the effects of sumoylation on the RNA-binding activity are not well studied. It has been shown that the binding activity of hnRNP C, which mainly regulates pre-mRNA splicing, pre-mRNA polyadenylation, 3'-end processing [203] [204] [205], but also maintains single stranded DNA (ssDNA) extensions at telomeres [206] [207], to ssDNA is decreased upon sumoylation [201]. Sam68 (Src associated in mitosis, 68 kDa) is a member of the hnRNP K homology domain family of RNA-binding proteins and has been proposed to play a role in cell cycle progression [208] by repression of CCND1 and cyclin E on the transcriptional level [209], independent of its RNA-binding activity. Thus, Sam68 is implied as a potential tumor suppressor [202]. In contrast to hnRNP C sumoylation, Babic *et al.* have demonstrated that the RNA-binding activity of sumoylated Sam68 is not altered compared to native Sam68, but sumoylation was suggested to change its affinity for co-activators [202].

Desterro *et al.* [210] demonstrated that the RNA-editing activity of the double-stranded RNA-specific adenosine deaminase 1 (ADAR1) is affected upon sumoylation. Hence, the

sumoylation of an RNA-binding protein regulating translation and no positive effect of sumoylation on RNA-binding proteins have been reported in the literature. However, sumoylation of proteins involved in posttranscriptional gene expression processes have been shown to positively alter their function, e.g. sumoylation of the transcription factor eIF4E activates mRNA translation [211] and sumoylation of 3'-end mRNA processing factors symplekin and CPSF-73 enhances the activity of the 3'-mRNA processing complex [212].

1.4 Project aims

The multifunctional La protein has been shown to be involved in different aspects of the RNA metabolism in the cell and interacts with a variety of cellular and viral RNAs. Whereas the interaction of La with RNA polymerase III transcripts and other terminal polyuridylate RNAs has been extensively studied, the recognition and binding of internal RNA sequences by the La protein has not been well established. There are reports, however, pointing to the involvement of the atypical RRM2 as a binding domain for this specific set of RNAs, and recent studies propose that the cancer-associated La protein plays an important role in cancer pathology. This suggestion is based on the following facts: (a) La is overexpressed in solid tumors, (b) La knockdown results in a defect in cell proliferation in various cell lines, (c) cyclin D1 protein levels are decreased upon La depletion in some cultured cells, (d) elevated cyclin D1 levels were paralleled with increased La protein levels in tissue lysates, and (e) La is associated with the cyclin D1 mRNA in HeLa cells and stimulates the IRES-dependent translation of cyclin D1. However, the molecular mechanisms by which La stimulates CCND1 translation are not well established, therefore, this project aims to gain a better understanding of the interaction between the RNA-binding protein La with the 5'-UTR of cyclin D1, which contains an IRES element. Furthermore, the SUMO-modification of the hLa protein is to be analyzed *in vitro* to study whether it modulates the RNA-binding activity of the hLa protein.

Based on the current knowledge about La and the stimulation of IRES-mediated cyclin D1 translation it is hypothesized that sumoylation of the La protein alters its RNA-binding activity to bind the cyclin D1 mRNA at structural elements in its 5'-UTR. To test this hypothesis the following aims were developed: The first aim was to characterize the La interaction with the cyclin D1 mRNA and identify the La binding site, and the second aim was to analyze whether sumoylation of La modulates its RNA-binding activity.

Aim 1: The hLa protein has been shown to interact directly with the CCND1 mRNA and to stimulate the IRES-mediated translation; however, the molecular mechanisms behind this interaction are not clear.

In order to analyze the interaction of human La with the cyclin D1 5'-UTR, recombinant human La was prepared and its binding to *in vitro* transcribed cyclin D1 5'-UTR RNA was analyzed by native electrophoretic mobility shift assays (EMSAs) and fluorescence polarization. The aim was to identify the domains of the modular La protein that are required for cyclin D1 RNA-binding. Hence, mutants of the human La protein, including RNA-binding domain mutants, were purified and tested for cyclin D1 RNA-binding in native EMSAs. Fluorescence polarization assays were performed of the recombinant La proteins to determine the affinity of those proteins to the cyclin D1 RNA. In addition, a minimal La-binding site within the cyclin D1 mRNA was mapped by competitive EMSAs using recombinant wild type La protein and *in vitro* transcribed RNAs as well as a synthesized RNA oligoribonucleotide of cyclin D1.

Based on the current knowledge, the binding of La to a structural element in the cyclin D1 5'-UTR mediated by the canonical RRM1 and atypical RRM2 are expected.

Aim 2: The sumoylation of hLa has been described, however, the effects of SUMO-modification on the RNA-binding function of La has not been studied.

Therefore, a possible regulation of the RNA-binding activity through SUMO-modification was analyzed. An optimized *in vitro* sumoylation assay for recombinant human La had to be established to determine the effects of sumoylation on the RNA-binding activity of La. To study the impact of sumoylation on La function, the cyclin D1 RNA-binding of modified wild type La and non-modified La was analyzed by native EMSA. To validate the effects of sumoylation on RNA-binding, two fluorescent labeled synthesized RNA oligoribonucleotides representing a regulatory element in mRNAs, namely the TOP element in 5'-UTR of mRNAs encoding for ribosomal proteins and other translational factors, were also used as targets in native EMSAs with sumoylated and unmodified La. The SUMO acceptor sites had to be identified and mutated to further validate the modulation of La sumoylation on the RNA-binding activity.

Based on the affects of sumoylation on the RNA-binding protein hnRNP C1 [201] SUMO-modification of hLa is expected to negatively alter its RNA-binding activity.

Overall impact and outlook of this work will help in understanding the molecular mechanisms by which La is recognizing and binding internal mRNA elements, as exemplified by La recognition of the cooperative oncogene cyclin D1. Further, the observations made herein will help the understanding of how the RNA-binding activity of hLa is regulated, which will give an insight on the possibility of specifically targeting the functions of the cytoplasmic La protein in protein synthesis, but maintaining its essential functions in RNA polymerase III transcript processing in the nucleus, by novel therapeutic agents.

2. Materials

2.1 Kits

KinaseMax™ 5'End Labeling Kit	Ambion
MEGAclear Kit	Ambion
MEGAshortscript™ T7 High Yield Transcription Kit	Ambion
QIAGEN Plasmid Maxi Kit	Qiagen
QIAprep Spin MiniPrep Kit	Qiagen
QIAquick Gel Extraction Kit	Qiagen
QIAquick PCR Purification Kit	Qiagen
Rapid DNA Dephos & Ligation Kit	Roche

2.2 Equipment and instruments

Agarose Gel electrophoresis chamber B1A EasyCast Mini Gel System (7 x 8 cm)	Owl/ Thermo Fisher Scientific
Agarose Gel electrophoresis chamber B2 EasyCast Mini Gel System (12 x 14 cm)	Owl/ Thermo Fisher Scientific
Centrifuge 5810R, bench top centrifuge	Eppendorf
Centrifuge 5417R, bench top centrifuge	Eppendorf
Electrophoresis Power Supply EPS301	GE Healthcare
Electrophoresis Power Supply EPS601	GE Healthcare
GE (Amersham) Typhoon Imager	GE Healthcare
GE (Molecular Dynamics) Storm 865 Imager	GE Healthcare
Hoefer™ Slab Gel Dryer GD2000	Amersham Biosciences
ImageQuant RT ECL Imager	GE Healthcare

Labquake Shaker Rotisserie	Barnstaed
Mastercycler ep gradient	Eppendorf
MaxQ™ 6000 Refrigerated Stackable Shaker	Barnstaed
Microcentrifuge DW-41	Qualitron, Inc.
Microwave MW8117W professional series	Emerson
Mini Trans-Blot® Electrophoretic transfer cell	Biorad
Mini-Protean® Tetra Cell	Biorad
MLS-3750 Top-Loading Autoclave	Sanyo
Multi-Purpose Rotator	Barnstead
NanoDrop ND-1000 Spectrophotometer	Thermo Scientific
Pipetman® P10/P20/P200/P1000	Gilson, Inc.
PB 153-S / FACT precision balance	Mettler Toledo
PB 1502-S / FACT precision balance	Mettler Toledo
POLARstar Omega	BMG Labtech
Powerpette® Turbo	Jencons Scientific, LLC
Scintillation Counter LS6500	Beckman Coulter
SevenGo™ pH-Meter SG2	Mettler Toledo
Sonic Dismembrator Model 100	Fisher Scientific
Thermomixer® R	Eppendorf
Thermomixer®	Eppendorf
UV-Transilluminator Universal Hood	Biorad
UV/White Light Transluminator	UVP, LLC.
Vortex Genie 2	Scientific Industries, Inc.
VWR® Ceramic hot plate	VWR International
VWR® Labmax™ Bottle-Top Dispensers	VWR International

VWR® Stir plate VS-C7	VWR International

2.3 Consumables

2V qualitative filter circles, Ø 185 mm	Whatman
384-well plate, black	Greiner Bio-One, Inc.
Amicon® Utra Centrifugal Filter Units Ultracel®10K membrane	EMD Millipore
Autoradiography Film B Plus, full blue	Labscientific, Inc.
Costar assay plate, 96 well no lid	Corning
Culture tubes, 13 ml	Sarstedt
NitroPure Pure Nictrocellulose, 0.45 micron	GE Water & Processing Technologies
Petri dish 100/15 with vents	Greiner Bio-One, Inc.
Safe-Lock tubes, 1.5 ml and 2 ml	Sarstedt
Screw Cap tube conical, 15 ml and 50 ml	Sarstedt
Serological pipette (5 ml, 10 ml, 25 ml)	Sarstedt
TipOne 1-200 µl Natural Beveled Tips	USA Scientific, Inc.
TipOne 101-1000 µl Blue Graduated Pipet Tips	USA Scientific, Inc.
TipOne Beveled Filter Tips 1-20 µl	USA Scientific, Inc.
TipOne Graduated Filter Tips 1-200 µl	USA Scientific, Inc.
Spectra/Por® membrane dialysis tubing MWCO: 6-8,000	Spectrum Labs
VWR® Centrifuge Filters modified PES 10K	VWR International
Whatman® filter paper 3	Whatman

2.4 Chemicals

All non-listed chemicals used were EMD/OmniPur molecular biology grade.

cOmplete protease inhibitor	Roche
Ethyl alcohol 190 proof ACS/USP grade	Pharmco-Aaper
Ethyl alcohol 200 proof absolute, anhydrous	Pharmco-Aaper
Formamide, ≥99.5% (GC), liquid	Sigma-Aldrich Corp.
Heparin sodium	Acros
Imidazole, 99+%, crystalline	Acros
Nonidet P-40	Fluka
Protein Assay, dye reagent concentrate	Biorad
Sodium azide	Fluka
Sodium phosphate, monobasic, anhydrous	Sigma
TEMED	National diagnostics
Thiourea	Sigma
Tween® 20, molecular biology grade	Promega
UltraPure ProtoGel 30% (w/v) acrylamide: 0.8% (w/v) Bis-Acrylamide stock solution (37.5:1)	National diagnostics
Xylene cyanole FF	Acros

2.5 Buffers, solutions, and media

annealing buffer	10 mM Tris/HCl pH 7.4, 40 mM NaCl, 0.2 mM EDTA
APS, 10%	10% (w/v) ammonium persulfate in water
Adenosine 5'-[γ-thio]triphosphate tetralithium salt (ATP-γ-S)	Sigma
ATP-Mix, 10x	100 mM creatine phosphate, 50 mM $MgCl_2$, 20 mM rATP
BSA, 2%	2% (w/v) BSA Fraction V in 1x wash buffer (IB)
BSA, 10 mg/ml	New England Biolabs
Coomassie destaining solution	45% (v/v) methanol, 10% (v/v) acetic acid

Coomassie staining solution	45% (v/v) methanol, 10% (v/v) acetic acid, 0.1% (w/v) Coomassie brilliant blue R-250, 0.1% (w/v) Coomassie brilliant blue G-250
dialysis buffer, allowing nuclease treatment	50 mM Tris/HCl pH 7.6, 75 mM NaCl, 5 mM $CaCl_2$
dialysis buffer	10 mM Tris/HCl pH7.4, 150 mM NaCl, 3 mM $MgCl_2$, 5% (v/v) glycerol
DNA-loading buffer, 6x	0.25% (w/v) bromphenol blue, 0.25% (w/v) xylene cyanole FF, 30% (v/v) glycerol
elution buffer	50 mM NaH_2PO_4, 300 mM imidazole, 300 mM NaCl, 1% (w/v) cOmplete protease inhibitor
FP assay buffer	20 mM Tris/HCl pH 8.0, 150 mM NaCl, 1.5 mM $MgCl_2$, 0.5 mM EDTA, 0.5% (v/v) Nonidet P-40, 2 mg/ml (w/v) bovine gamma globulin
LB Agar (Miller) selection plates	37 g granulated medium in 1 L water, autoclaved and 100 μg/ml ampicillin or 30 μg/ml kanamycin added at temperature < 50°C and poured into petri dishes; stored at 4°C
LB Broth (Miller) medium	25 g granulated medium in 1 L water, autoclaved
LB Broth (Miller) ampicillin selection medium	25 g granulated medium in 1 L water, autoclaved, 100 μg/ml ampicillin added freshly before use
LB Broth (Miller) kanamycin selection medium	25 g granulated medium in 1 L water, autoclaved, 30 μg/ml kanamycin added freshly before use
lysis buffer	50 mM NaH_2PO_4, 10 mM imidazole, 300 mM NaCl, 1 mg/ml lysozyme, 1% (w/v) cOmplete protease inhibitor
NEBuffer 1, restriction endonuclease buffer 1	New England Biolabs
NEBuffer 2, restriction endonuclease buffer 2	New England Biolabs
NEBuffer 3, restriction endonuclease buffer 3	New England Biolabs

primary antibody solutions	antibody in 1x IB wash buffer, 1:1,000 (v/v) sodium azide
RNA-binding buffer, 1x	20 mM Tris/HCl pH 8.0, 150 mM NaCl, 1.5 mM $MgCl_2$, 0.5 mM EDTA, 0.5% (v/v) Nonidet P-40
RNA-binding buffer, 2x	40 mM Tris/HCl pH 8.0, 300 mM NaCl, 3 mM $MgCl_2$, 1 mM EDTA, 1% (v/v) Nonidet P-40
RNA-loading buffer	5% glycerol (v/v), 0.25% (w/v) bromphenol blue in 1x TBE buffer
RNA-loading buffer, denaturing	80% (v/v) formamide, 1 mM EDTA, 50 mM 1x TBE buffer, 0.05% (w/v) bromphenol blue, 0.05% (w/v) xylene cyanole FF
rATP, 100 mM	Promega
SDS-loading buffer, 3x	60 mM Tris/HCl pH 6.8, 10% (v/v) glycerol, 6% (v/v) SDS, 0.01% (w/v) bromphenol blue, 5% (v/v) β-mercaptoethanol (added freshly)
SDS-running buffer	250 mM Tris, 19.2 mM glycine, 0.35 mM SDS
siRNA-buffer, 5x	Dharmacon, working stock dilution 1x in nuclease free water
sumoylation buffer	50 mM Tris/HCl pH 8.0, 100 mM NaCl, 0.1% (v/v) Nonidet P-40, 10% (v/v) glycerol, 1 mM Dithiothreitol
TAE-buffer	50 mM Tris/HCl pH 7.4, 20 mM sodium acetate, 2 mM EDTA
TBE-buffer, 40x	1.8 M Tris/HCl pH 8.5, 1.8 M boric acid, 40 mM EDTA
TE-buffer, 1x	10 mM Tris/HCl pH 8.0, 1 mM EDTA
Transfer buffer	25 mM Tris/HCl pH 8.3, 192 mM glycine, 20% (v/v) ethanol 190 proof
wash buffer	50 mM NaH_2PO_4, 42.5 mM imidazole, 1 M NaCl, 0.11% (v/v) Triton X-100, 1% (w/v) cOmplete protease inhibitor
wash buffer, immunoblotting (IB)	10 mM Tris/HCl pH 7.5, 100 mM NaCl, 0.1% (v/v) Tween 20
water, nuclease-free	Fisher Scientific

2.6 Enzymes and antibodies

GC-RICH PCR System	Roche
Micrococcal nuclease	New England Biolabs
Platinum® *Taq* DNA Polymerase High Fidelity	Invitrogen
Platinum® *Taq* DNA Polymerase High Fidelity, 10x High Fidelity PCR buffer	Invitrogen
Platinum® *Taq* DNA Polymerase High Fidelity, 50 mM magnesium sulfate	Invitrogen
Platinum® *Pfx* DNA polymerase	Invitrogen
Platinum® *Pfx* DNA polymerase, 10x *Pfx* amplification buffer	Invitrogen
Platinum® *Pfx* DNA polymerase, 50 mM magnesium sulfate	Invitrogen
Restriction endonuclease BamH I	New England Biolabs
Restriction endonuclease Dpn I	New England Biolabs
Restriction endonuclease EcoR I-HF	New England Biolabs
Restriction endonuclease Hind III	New England Biolabs
Restriction endonuclease Pst I	New England Biolabs
rRNasin® ribonuclease inhibitor	Promega
Peroxidase– AffiniPure Goat Anti-Mouse IgG(H+L)	Jackson ImmunoResearch
Peroxidase– AffiniPure Goat Anti-Rabbit IgG(H+L)	Jackson ImmunoResearch
IgG2α,κ, mouse isotype control, clone eBM2a	eBioscience
La 3B9, mouse monoclonal antibody	kind gift of Bachmann
La SW5, mouse monoclonal antibody	kind gift of Bachmann
SUMO1 (C'-terminal), rabbit polyclonal antibody	Enzo Life Sciene
SUMO2/3 (N'-terminal), rabbit polyclonal antibody	Enzo Life Science

2.7 Standards

PageRuler™ Prestained Protein Ladder	Fermentas
Prestained Protein Marker, Broad Range	New England Biolabs
Quickload® 1 kb DNA ladder	New England Biolabs
Quickload® 100 bp DNA ladder	New England Biolabs

2.8 Radioactive isotypes

ATP, [γ-^{32}P]- 3000Ci/mmol 10mCi/ml	Perkin Elmer
CTP, [α-^{32}P]- 800Ci/mmol 10mCi/ml Lead	Perkin Elmer

2.9 Prokaryotic cells

Subcloning Efficiency™ *Escherichia coli* DH5α™ Competent Cells	Invitrogen
BL21 Star™ (DE3)pLysS One Shot® Chemically Competent *Escherichia coli*	Invitrogen
Rosetta™(DE3)pLysS Competent Cells	Novagen

2.10 Plasmids

pEGFPc1-LaWT	previously cloned by lab members [35]
pet11d-SUMO-1/-2/-3WT	kind gift of C.-M. Chiang, University of Texas, Dallas, TX, USA
pet11d-SUMO-1/-2/-3GA	kind gift of C.-M. Chiang, University of Texas, Dallas, TX, USA
pet15b-Ubc9	kind gift of R. Hay, University of Dundee, Dundee, UK
pet28b(+) vector	Novagen
pet28b(+)-LaWT	kind gift of E. Chan, Scripps Research Institute, San Diego, CA, USA
pet28b(+)-LaΔ1	previously cloned by lab members [35]
pet28b(+)-LaΔ2	previously cloned by lab members [35]

pet28b(+)-LaΔ4	previously cloned by lab members [35]
pet28b(+)-LaΔ7	previously cloned by lab members [35]
pet28b(+)-LaSM3	kind gift of Jeff Twiss, Drexel University, Philadelphia, PA, USA
pet28b(+)-LaSM23	kind gift of Jeff Twiss, Drexel University, Philadelphia, PA, USA
pGEX2T-BP2ΔFG	kind gift of F. Melchior, DKFZ-ZMBH Alliance, Heidelberg, Germany
pGEX3T-IR1+M	kind gift of F. Melchior, DKFZ-ZMBH Alliance, Heidelberg, Germany
pGEX2T-SAE1	kind gift of R. Hay, University of Dundee, Dundee, UK
pGEX2T-SAE2	kind gift of R. Hay, University of Dundee, Dundee, UK
pGEX2T-SUMO-1	kind gift of
pT7HTb vector	kind gift of D. Fedarovich, Medical University of South Carolina, Charleston, SC, USA
pRCD1F	kind gift of J. Gera, University of California, Los Angeles, CA, USA

2.11 Oligoribonucleotides

All oligoribonucleotides were synthesized (and labeled) by Integrated DNA Technologies, Inc.

D1-ATG:
5'-GCUGCCCAGGAAGAGCCCCAGCCAUGGAACACCAGCUCCUGUGCUGC-3'

6-FAM D1-ATG:
5'-GCUGCCCAGGAAGAGCCCCAGCCAUGGAACACCAGCUCCUGUGCUGC//6-FAM-3'

mu2: 5'-CAGGAAGAGCCCCAGCCAGGGAACACCAGCUCCUG-3'

mu3: 5'-CAGGAAGAGCCCCACAGAUGCACCACCAGCUCCUG-3'

Cy3- L22 (GenBank Accession number NM_031104):
5'-CUCCUCUGCCGCCAUGGCUCCUGUGAAAAAGCUUGUG//Cy3-3'

Cy3- L37 (GenBank Accession number NM_031106):
5'-UUCCGGUCUCUUUGGCCUCGCCGGUAGAAGCAAGAUG//Cy3-3'

`5'- CUGAUCGCUGUAGACCUUACAGUUGCUGCUAACUGCC//Cy3-3'`

2.12 Oligonucleotides

2.12.1 Oligonucleotides for *in vitro* transcription templates

template D1-FL:

D1-1S-T7 sense primer (T7 promoter sequence is underlined):
`5'-`<u>`ACTCCGGGATCCTAATACGACTCACTATAGGG`</u>`CACACGGACTACAGGGGAG-3'`

D1-4AS antisense primer: `5'-CATGGCTGGGGCTCTTCC-3'`

template part A:

D1-1S-T7 sense primer: vide supra

D1-3AS antisense primer: `5'- GGAGCGTGCGGACTCTGC-3'`

template part B:

D1-2S-T7 sense primer (T7 promoter sequence is underlined):
`5'-`<u>`ACTCCGGGATCCTAATACGACTCACTATAGGG`</u>`CGAGGGGCAGAAGAGCGCGAGG`
`GAGCGC-3'`

D1-4AS antisense primer: vide supra

template ΔB1:
a) 5'-terminus upstream of deletion (sequence for overhang is in bold):
D1-1S-T7 sense primer: vide supra
D1-30AS: `5'-`**`GCAGCTGGGGAGGGCTGTGG`**`CGGAGCGTGCGGACTCT-3'`

b) 3'-terminus downstream of deletion (sequence for overhang is in bold):
D1-29S: `5'-`**`GCAGAGTCCGCACGCTCCG`**`CCACAGCCCTCCCCAGCT-3'`
D1-4AS antisense primer: vide supra

c) fusion PCR
D1-1S-T7 sense primer: vide supra
D1-4AS antisense primer: vide supra

template ΔB2:
D1-1S-T7 sense primer: vide supra
D1-13AS antisense primer: `5'-CTGTGGGTCCTGGCTGGGTCC-3'`

template ΔB2.1:
D1-1S-T7 sense primer: vide supra

D1-35AS antisense primer: 5'-TGGGCAGCTGGGGAGGG-3'

<u>template ΔB2.2:</u>
D1-1S-T7 sense primer: vide supra
D1-34AS antisense primer: 5'- CTCTTCCTGGGCAGCTGGGG-3'

2.12.2 Oligonucleotides for DNA:RNA hybrid formation

1AS: 5'-TCTGGAGGCTCCAGGACTTTGCAACTTCAACAAAACTCCCCTGTAGTCCG
      TGTGCTATAGTGAGTCGTATTAGGATCCCT-3'

2AS: 5'-CTCGCCGGAGCGTGCGGACTCTGCTGCTCGCTGCTACTGCGCCGACAGCC
      CCTATAGTGAGTCGTATTAGGATCCCT -3'

3AS: 5'-GCGCTCGGCTCTCGCTTCTGCTGCCCCGCGCTCCCTCGCGCTCTTCTGCC
      CCTATAGTGAGTCGTATTAGGATCCCT-3'

4AS: 5'-GGCTGGGGCTCTTCCTGGGCAGCTGGGGAGGGCTGTGGGTCCTGGCTGGG
      TCCCTATAGTGAGTCGTATTAGGATCCCT-3'

5AS: 5'-TGCTCGCTGCTACTGCGCCGACAGCCCTCTGGAGGCTCCAGGACTTTGCA
      ACCTATAGTGAGTCGTATTAGGATCCCT-3'

6AS: 5'-GCGCTCCCTCGCGCTCTTCTGCCCCTCGCCGGAGCGTGCGGACTCTGCCT
      ATAGTGAGTCGTATTAGGATCCCT-3'

7AS: 5'-TGGGGAGGGCTGTGGGTCCTGGCTGGGTCCGCGCTCGGCTCTCGCTTCTG
      CTGCCCCCTATAGTGAGTCGTATTAGGATCCCT-3'

2.12.3 Oligonucleotides for mutation of the human La protein

<u>hLa mutant: LaSM123 (introducing a lysine to arginine substitution at K41)</u>
LaK41R-S sense primer (spanning nucleotides 94-137, mutation in bold):
5'-GGACAAGTTTCTAAAGGAACAGATAA**G**ACTGGATGAAGGCTGG-3'

LaK41R-AS antisense primer (spanning nucleotides 94-137, mutation in bold):
5'-CCAGCCTTCATCCAGT**C**TTATTGTTCCTTTAGAAACTTGTCC-3'

<u>hLa mutant: RRM1+2 (terminal deletion of NTD and CTD, aa 1-113 and aa 335- 408)</u>
69S sense primer (spanning nts 319 to 340, start codon is in bold, *EcoR* I restriction site underlined):
5'-GTCCC<u>GAATTC</u>TACC**ATG**AAAAATGATGTAAAAAACAGA-3'

70AS antisense primer (spanning nts 988 to 1005, stop codon in bold *Hind* III restriction site underlined):
5'-CGCCCA<u>AAGCTT</u>**CTA**ACGACCTTTTGACTTCCATTT-3'

<u>hLa mutant: RRM1+2 K208R (terminal deletion of NTD and CTD, aa 1-113 and aa 335- 408)</u>
69S sense primer: vide supra

70AS antisense primer: vide supra

hLa mutant: RRM1+2 K200RK208R (introducing a lysine to arginine substitution at K200 and terminal deletion of NTD and CTD, aa 1-104 and aa 335- 408)

a) Generation of 5'-terminus and substitution of lysine 200 to arginine and introduction of restriction sites for cloning into pT7HTb vector

pT7HTb sense primer (spanning nts 108 to 144 (NM_003142-3), *BamH* I restriction site underlined):
```
5'-GCACTGGGATCCATGGCTGAAAATGGTGATAATGAAAAGATGGCTGCC-3'
```

LaK200R-AS antisense primer (spanning nucleotides 586-619, mutation in bold):
```
5'-TTTTGTTTTATCTCACCTTCGATTTAATTCTCC-3'
```

b) Generation of 3'-terminus and substitution of lysine 200 to arginine and introduction of restriction sites for cloning into pT7HTb vector

LaK200R-S sense primer (spanning nucleotides 591-614, mutation in bold):
```
5'-CAAAATAGAGTGGAAGCTAAATTA-3'
```

pT7HTb antisense primer (spanning nts 1307 to 1334 (NM_003142-3), *Pst* I (underlined) and *Hind* III restriction site double-underlined):
```
5'-ACGTCAGCTGCAGCCAAGCTTCTACTACTGGTCTCCAGCACCATTTTCTG-3'
```

c) fusion PCR

pT7HTb sense primer: vide supra
pT7HTb antisense primer: vide supra

d) terminal deletion of NTD (aa 1-113) and CTD (aa 335- 408)

69S sense primer: vide supra
70AS antisense primer: vide supra

2.13 Software/ Web servers

DISOPRED2 Prediction of Protein Disorder	http://bioinf.cs.ucl.ac.uk/disopred
Image Lab™ Gel documentation	Biorad
ImageQuant TL	GE Healthcare
MacVector sequence analysis	MacVector, Inc.
Mfold/ UNAfold	M. Zuker, State University of NY, USA
MARS, data analysis software for POLARstar Omega plate reader	BMG Labtech GmbH
Prism 4 and Prism 5	GraphPad, Inc.
SUMOsp2.0, version 2.0.4	CUCKOO Workgroup

SUMOplot^{TM} Analysis Program   Abgent

2.14 List of suppliers

Abgent, San Diego, CA, USA

Acros/ Thermo Fisher Scientific, Asheville, NC, USA

Ambion/ Life Technologies, Grand Island, NY, USA

Amersham Biosciences/ GE Healthcare, Pittsburgh, PA, USA

Barnstead/ Thermo Fisher Scientific, Asheville, NC, USA

Beckman Coulter, Brea, CA, USA

Biorad, Hercules, CA, USA

BMG Labtech GmbH, Ortenberg, Germany

Corning, Tewksbury, MA, USA

eBioscience, Inc., San Diego, CA, USA

EMD Millipore/ Merck KGaA, Darmstadt, Germany

Emerson, St. Louis, MO, USA

Enzo Life Science, Farmingdale, NY, USA

Eppendorf, Hamburg, Germany

Fermentas/ Thermo Fisher Scientific, Asheville, NC, USA

Fisher Scientific/ Thermo Fisher Scientific, Asheville, NC, USA

Fluka/ Sigma-Aldrich Corp., St. Louis, MO, USA

GE Healthcare, Pittsburgh, PA, USA

GE Water and Processing Technologies/ GE Healthcare, Pittsburgh, PA, USA

Gilson, Inc, Middleton, MI, USA

GraphPad, Inc., La Jolla, CA, USA

Greiner Bio-One North America, Inc. Monroe, NC, USA

Invitrogen/ Life Technologies, Grand Island, NY, USA

Integrated DNA Technologies, Inc. Coralville, IA, USA

Jackson ImmunoResearch Laboratories, Inc. West Grove, PA, USA

Jencons Scientific, LLC., Bridgeville, PA, USA

MacVector, Inc. Cary, NC, USA

Mettler Toledo, Columbus, OH, USA

National Diagnostics, Atlanta, GA, USA

New England Biolabs, Ipswich, MA, USA

Novagen/ EMD Millipore/ Merck KGaA, Darmstadt, Germany

Perkin Elmer, Waltham, MA, USA

Pharmco-Aaper, Charlotte, NC, USA

Promega, Madison, WI, USA

Qiagen, Valencia, CA, USA

Qualitron, Inc./ Krackeler Scientific, Inc., Albany, NY, USA
Roche, Indianapolis, IN, USA

Sarstedt, Newton, NC, USA

Scientific Industries, Inc., Bohemia, NY, USA

Sigma-Aldrich Corp., St. Louis, MO, USA

Spectrrum Labs, Rancho Dominguez, CA, USA

Thermo Fisher Scientific, Asheville, NC, USA

USA Scientific, Inc., Ocala, FL, USA

UVP, LLC, Upland, CA, USA

VWR International, Radnor, PA, USA

Whatman/ GE Healthcare, Pittsburgh, PA, USA

3. Methods

3.1 Molecular methods of DNA technology – cloning, extraction, and analysis of DNA

3.1.1 Introduction of mutations in the cDNA sequence of human La by PCR

This dissertation analyzes the binding of recombinant human La (hLa) to the 5'-UTR of cyclin D1 mRNA. In order to determine the domain required for cyclin D1 RNA-binding and to investigate whether SUMO-modification of La affects its RNA binding activity several mutations had to be introduced in hLa as detailed below.

Point mutations (lysine to arginine substitutions) were introduced to create the LaSM123 mutant and the RRM1+2 K200RK208R mutant (aa 114-326). The expression vector was mutated by oligo-directed mutagenesis of the specific sites. Both primers, the sense and the antisense primer, are mutagenic by containing the desired mutation used to introduce point mutations. Mutagenic primers are longer than usual primers to form stable double strands and contain the substitution rather in the middle than at the edges of the oligonucleotide (oligo). The following already generated La mutants were used for the production of new mutants:

- LaSM23 (vector pet28b(+), mutation: lysines 185 and 208 to arginine)
- LaSM3 (vector pet28b(+), mutation: lysine 208 to arginine)

The LaSM123 mutant:

The LaSM123 mutant was created by introducing a lysine to arginine substitution at lysine 41 in the pet28b(+) LaSM23 vector, which contains lysine to arginine substitutions at lysines 185 and 208, based on the whole plasmid mutagenesis after Braman [320] and the QuickChange site directed mutagenesis protocol from Stratagene. The primer LaK41R-S was used as a sense primer (2.12.3) and LaK41R-AS as antisense primer (2.12.3) in a PCR reaction using the high fidelity and non-strand displacing Platinum *Pfx* DNA polymerase to amplify both strands of the entire plasmid:

ad 50.0 µl   molecular grade water

    5.0 µl   10x Pfx amplification buffer

    1.5 µl   10 mM dNTP mixture (final concentration 0.3 mM)

    1.0 µl   50 mM magnesium sulfate (final concentration 2 mM)

    1.0 µl   10 µM LaK41R-S sense primer (final concentration 0.2 µM)

    1.0 µl   10 µM LaK41R-AS antisense primer (final concentration 0.2 µM)

  10.0 ng   pet28b(+) LaSM12 template DNA

  10.0 U   Platinum *Pfx* DNA polymerase

thermal profile

initial denaturation	95°C	30 sec	
denaturation	95°C	30 sec	
annealing	55°C	60 sec	18 cycles
elongation	68°C	6 min	

The PCR product contains the parental template plasmid as well as a nicked mutated plasmid. The parental plasmid is eliminated by endonucleolytic digestion with the restriction enzyme *Dpn* I, which specifically recognizes methylated or hemimethylated DNA. The parental plasmid template was prepared from *E. coli*, which methylated the template DNA. The PCR product was digested with 10 units *Dpn* I for 1 hour at 37°C. 10 units of *EcoR* I, as a unique cutter for this sequence, was used to linearize 200 ng of the template DNA and a 5 µl aliquot of PCR product in 1x NEB1 buffer for 2 hours at 37°C. The linearized DNAs were separated on a 1% (w/v) agarose gel to assess the correct size of the amplified product. Once the correct size was confirmed, 1 µl of the *Dpn* I digested PCR product was used for transformation in *E. coli* DH5α cells (3.1.3).

The La K200RK208R mutant:

The template for the substitution of lysine 200 to arginine in the minimal RNA-binding competent RRM1+2 context was the pet28b(+) LaSM3 plasmid, this allows the exclusion of C'-terminal putative sumoylation sites,. The pet28b(+) LaSM3 plasmid harbors the lysine to

arginine substitution at lysine 208. This cloning was a multistep process, the substitution had to be introduced by oligo-directed mutagenesis, and terminal deletions had to be introduced to create the minimal La protein RRM1+2 context. The point mutation was introduced by two mutagenic primers, LaK200R-S sense and LaK208R-AS antisense primers (see 2.12.3). The goal was to clone the cDNA of the double substitution mutant K200R and K208R into the pT7HTb expression vector. Hence, another approach called PCR site-directed mutagenesis was taken for this mutagenesis. The 5'-terminal end of the La cDNA was amplified using the antisense mutagenic primer as well as the pT7HTb sense primer, which contains a restriction site for the restriction enzyme *BamH* I (product A; indicated in 2.12.3). The mutagenic sense primer LaK200R-S and the pT7HTb antisense primer, which introduces a *Pst* I restriction site, were used for the amplification of the 3'-terminal end of the La cDNA (product B). Both products were designed so that they have sequence overlap of 23 nucleotides (nts). This overlap allowed to use product A and B as templates for the amplification of the entire La K200RK208R cDNA (product C, 1224 nts) by fusion PCR. All amplifications were performed using a Platinum *Taq* DNA Polymerase High Fidelity, 10 ng of the 5'-terminus (product A) and the 3'-terminus (product B) were combined to serve as a template for the amplification of the entire cDNA:

reaction mix for fusion PCR

ad 50.0 μl   molecular grade water

5.0 μl   10x High Fidelity PCR buffer

5.0 μl   2 mM dNTP mixture (final concentration 0.2 mM)

2.0 μl   50 mM magnesium sulfate (final concentration 2 mM)

1.5 μl   10 μM sense primer (final concentration 0.3 μM)

1.5 μl   10 μM antisense primer (final concentration 0.3 μM)

10.0 ng   template product A

10.0 ng   template product B

2.0 U   Platinum *Taq* DNA Polymerase High Fidelity

thermal profile

initial denaturation	94°C	60 sec
denaturation	94°C	30 sec
annealing	55°C	25 sec
elongation	68°C	1 min

35 cycles

Each PCR-product was purified using the QIAquick PCR-purification and QIAquick Gel Extraction Kit from Qiagen. In order to clone the PCR product La K200RK208R (product C) into the pT7HTb vector the PCR products were digested with *BamH* I and *Pst* I, then gel purified using the QIAquick Gel Extraction Kit from Qiagen, and ligated as described in section 3.1.2, and finally transformed into *E. coli* DH5α cells (3.1.3).

The RRM1+2 protein and its mutants:

The minimal RNA-binding competent La proteins RRM1+2, RRM1+2 K208R, and RRM1+2 K200RK208R were created amplifying nucleotides 340 to 1005 (aa 114 to 326) with primer 69S (nts 319 to 340, *EcoR* I site; 5'-end of RRM1) and primer 70AS (988 to 1005, *Hind* III *site,* 3'-end of RRM2.) For primer sequence and template information refer to 2.12.3. The DNA was amplified by the Platinum *Taq* DNA Polymerase High Fidelity as described above using the 69S and 70AS as sense and antisense primer, respectively. The PCR products were purified using the QIAquick PCR-purification and QIAquick Gel Extraction Kit from Qiagen, analyzed for correct size and purification success on a 1.5 % agarose gel, and ligated into the pet28b(+) vector (refer to 3.1.2).

Plasmid DNAs were prepared as described in section 3.1.4 and analyzed for correct mutagenesis by DNA-sequencing, which was performed either on-campus by the MUSC Nucleic Acid Analysis Facility or off-campus by GENEWIZ.

3.1.2 Ligation of PCR products in expression vectors

The La K200RK208R PCR product C had to be cloned into the pT7HTb. The RRM1+2, RRM1+2 K208R, and RRM1+2 K200RK208R PCR products had to be cloned into the pet28b(+) prokaryotic expression vector to be eventually expressed in *E. coli* BL21 cells for protein purification.

For cloning pet28b(+) La K200RK208R the La K200RK208R PCR product C and the pet28b(+) vector were digested with *BamH* I and *Pst* I overnight at 37°C as detailed:

> *digestion mix (BamH I and Pst I)*
> ad 50.0 μl molecular grade water
>   30.0 μl purified La K200RK208R PCR product
>   4.0 μl 10 μg/ml BSA
>   4.0 μl 10x NEB buffer 3
>   20 U   *BamH* I
>   20 U   *Pst* I

For cloning RRM1+2, RRM1+2 K208R, and RRM1+2 K200RK208R the PCR products and the vector pet28b(+) were digested with *EcoR* I-HF and *Hind* III overnight at 37°C as detailed:

> *digestion mix (EcoR I and Hind III)*
> ad 50.0 μl molecular grade water
>   30.0 μl purified PCR product
>   4.0 μl 10x NEB buffer 2
>   20 U   *EcoR* I-HF
>   20 U   *HindIII*

The digestion products were gel purified using the QIAquick Gel Extraction Kit and eluted in 30 μl elution buffer. The empty vectors pT7HTb and pet28b(+) were previously prepared in the lab. In order to prevent the DNA strands from self-ligating, the 5'-end of the plasmid DNA was dephosphorylated using the Rapid DNA Dephos & Ligation Kit. 3 μl of the plasmid DNA were dephosphorylated by 1 U rAPid Alkaline Phosphatase in 1x dephosphorylation buffer in a total volume of 20 μl for 10 at 37°C followed by enzyme inactivation for three min at 75°C. An aliquot of the purified, digested insert and an aliquot of the plasmid DNA were separated on a 1.2% agarose gel to determine the appropriate ratio of dephosphorylated plasmid DNA and insert. The vector DNA:insert DNA ratio of 1:3 was chosen for the ligation using the Rapid DNA Dephos & Ligation Kit according to the manufacturer. 5 μl of the ligation product was subsequently used for the transformation into competent *E. coli* DH5α cells.

3.1.3 Plasmid-DNA transformation

The transformation of a plasmid into competent bacteria is a common technique to amplify the plasmid DNA, usually containing a gene of interest, in large quantities. Expression vectors do not only carry the cDNA of the gene of interest cloned into the multiple cloning site, but are also engineered to contain, inter alia, a selectable marker and an origin of replication. The selectable marker is usually a gene that encodes for antibiotic resistance, e.g. the β-lactamase gene encodes for ampicillin resistance. By adding the appropriate antibiotic to the growth medium it allows for the selection of bacteria transformed with the resistance gene carrying expression vector. For DNA-amplification competent *E. coli* DH5α cells and *E. coli* BL21 cells or *E. coli* Rosetta cells were transformed for the expression of proteins. To introduce the plasmid into the bacterial cell the cells were heat shocked, which opens the pores of the competent allowing the plasmid intake. The plasmid is retained within the cell after the pores are closed again which is achieved by a cold shock.

50 μl of chemical competent bacterial cells were thawed on ice and incubated with plasmid DNA for 30 min on ice. The heat shock was performed by placing the cells for 90 sec at 42°C in a thermomixer. The cells were immediately returned to ice for two min for the cold shock. The cells were allowed to recover and grow for 45 min at 37°C and 300 rpm in 700 μl sterile LB Agar broth without the addition of any antibiotics. Bacteria transformed with ligation products were pelleted for 60 sec at maximum speed in a bench top centrifuge, and 650 μl of the supernatant was discarded. The pellet was resuspended in the remaining supernatant and plated on selective LB agar Miller plates. 25 μl of the bacteria culture transformed with DNA other than ligation products were directly plated on selective LB agar Miller plates. Transformed bacteria were grown in an incubator overnight at 37°C.

3.1.4 Preparation of plasmid DNA

The isolation of plasmid DNA was performed using either the QIAprep Spin MiniPrep Kit for a small scale preparation or the QIAGEN Plasmid Maxi Kit for larger scales. The DNA concentration was determined using the NanoDrop ND-1000 Spectrophotometer after the manufacturer's instructions. The quality of the DNA preparation was analyzed by linearizing the plasmid DNA with either restriction enzyme (3.1.2) and separating the digestion product on agarose gels (3.1.5).

3.1.5 Agarose gel electrophoresis of DNA and ethidium bromide staining

DNA was analyzed for its size and quality in analytic DNA agarose gel electrophoresis and preparative agarose gel electrophoresis were used to purify DNA samples. By applying an electric current the different-sized molecules are separated according to their size through the porous agarose matrix, the overall negatively charged DNA moves towards the positive pole with smaller DNA fragments migrating faster and further through the matrix than larger fragments. The fluorescent dye ethidium bromide is intercalating into the double helix, therefore it can be visualized by exciting the gel with UV-light.

The matrix concentration was chosen accordingly to the size of fragments that had to be separated, higher percentage gels of 2% or more agarose (w/v) in 1x TAE (50 mM Tris/HCl pH 7.4, 20 mM sodium acetate, 2 mM EDTA) were prepared for the separation of small *in vitro* transcription templates. Larger DNA fragments were separated on a 1% (w/v) or 1.5% (w/v) agarose gel in 1x TAE buffer. The agarose in 1x TAE buffer was dissolved by heating the mixture to its melting point in a microwave and allowed to cool down to approximately 50°C before the gel solution was poured into the electrophoresis chamber. To create pockets for loading the samples later, a plastic comb was immediately inserted after the gel solution was poured. The agarose gel was polymerized for at least one hour at room temperature before samples were loaded with one-sixth volume of 6x DNA-loading buffer (0.25% (w/v) bromphenol blue, 0.25% (w/v) xylene cyanole FF, 30% (v/v) glycerol). The DNA separation was achieved by applying a current between 70V (small agarose gels) and 100V (larger gels) in 1x TAE running buffer. A DNA standard of defined DNA fragments sizes was also separated and used as a reference. To visualize the DNA, the agarose gels were carefully removed and incubated in a 0.5 μg/ml (w/v) ethidium bromide solution for 10 min. Gels were destained in water for 10 min and subsequently excited by UV-light using either a UV light transluminator or a UV-Transilluminator Universal Hood that allowed photographic documentation.

Since ethidium bromide is highly mutagenic, a reference lane loaded with an aliquot of the DNA sample was separated in parallel, cut out and stained with ethidium bromide. After identification and marking the DNA fragment in the reference lane, the reference gel was aligned with the rest of the gel and the DNA was cut according to the marked reference lane. A successful DNA-excision was checked by ethidium bromide staining the remaining gel.

3.2 Protein biochemistry

3.2.1 Recombinant protein expression in prokaryotic cells

In order to synthesize and purify recombinant N'-terminal hexa-histidine-tagged proteins (overview in table 3.1) their expression vectors were expressed in bacterial *E. coli* BL21 or *E. coli* Rosetta cultures. The human La protein and the mutants LaΔ1, LaΔ2, LaΔ4, LaΔ7, including RRM1+2, RRM1+2 K208R, RRM1+2 K200RK208R, LaK41R, and LaSM123 were cloned into the pet28b(+) vector (vector map in appendix). The cDNA of the La mutants RCD1 and RCD2 were cloned into the pT7Htb vector (vector information in appendix). The SUMO-1/-2/-3 wild type (SUMO-1/-2/-3-GG), their mutants (SUMO-1/-2/-3-GA) and the Ubc9 cDNA were cloned into the pet11d vector. Those four vectors, pet28b(+), pT7HTb, and pet11d carry an Isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible T7 promoter, transcription from this promoter only occurs in the presence of the inducer IPTG (Jacob 1961) (Hansen 1998). The open reading frames (ORF) for the La proteins as well as for the SUMO isoforms are under control of the T7 promoter.

The expression vectors were transformed into *E. coli* BL21 competitive cells (3.1.3). The transformed cells were plated on LB agar plates containing 0.1 mg/ml Ampicillin (pet11d expression vector) or 30 μg/ml Kanamycin (pet28b(+) expression vector) and after incubating them overnight, a single colony was picked and then used to inoculate 3 ml of LB medium with the appropriate antibiotic. Bacterial cultures were allowed to grow overnight at 37°C and 300 rpm and used to inoculate 100 ml LB with antibiotic (0.1 mg/ml Ampicillin or 30 μg/ml Kanamycin) for small scale purification or 2x 400 ml for large scale purification. The bacterial growth at 37°C and 300 rpm was monitored by measuring the optical densities of the cultures. Protein expression was induced at an $OD_{605nm}= 0.4$ by adding 100 mM IPTG (His-SUMO-1/-2/-3 wild types and its respective mutants; His-La wild type) or 50 mM IPTG (His-La mutants) directly to the cultures. The temperature was reduced to 20°C and bacterial growth was allowed for an additional 12 hours. Cells were harvested by centrifugation at 4,000 x g for 15 min at 4°C and the supernatant was discarded. In order to purify the expressed proteins, cells were lysed depending on the cell pellet size in 4-6 ml lysis buffer, incubated for at least four hours under rotation and sonicated using a sonic dismembrator every hour for 3 x 15 sec during the incubation. Lysozyme, which is contained in the lysis buffer, is a glycoside hydrolase that damages prokaryotic cell walls by hydrolyzing the linkages in peptidoglycans. By applying ultrasound to the lysates another way of cell

disruption was chosen to increase the efficiency of cell lysis. The lysates were centrifuged at 20,000 x g for 20 min and 4°C to settle cell debris. The supernatant containing the soluble, recombinant His-tagged protein was then subjected to protein purification using Nickel-NTA spin columns as described in section 3.2.2.

The cDNA of the proteins SAE1, SAE2, BP2ΔFG, IR1M+1, and SUMO-1/-2/-3 fused with the glutathione S-transferase (GST)- tag were cloned in the pGEX-2T vector in other laboratories. The pGEX-2T expression vectors were transformed into *E. coli* BL21 or Rosetta 2 (DE3) cells and positively transformed cells were selected on agar plates in the presence of ampicillin. The pGEX-2T vector contains the β-lactamase gene for ampicillin resistance and a T7-promoter. The expression in either cell lines was performed as described above for the His-tagged proteins.

protein name	vector	fusion tag	protein purification by
La	pet28b(+)	His	J. Kühnert& PPL (MUSC)
LaΔ1	pet28b(+)	His	previously by lab members
LaΔ2	pet28b(+)	His	PPL (MUSC)
LaΔ4	pet28b(+)	His	PPL (MUSC)
LaΔ7	pet28b(+)	His	PPL (MUSC)
RCD1	pT7HTb	His	J. Kühnert
RCD2	pT7HTb	His	J. Kühnert
RRM1+2	pet28b(+)	His	J. Kühnert
LaK41R	pet28b(+)	His	J. Kühnert
LaSM123	pet28b(+)	His	J. Kühnert
RRM1+2 K208R	pet28b(+)	His	J. Kühnert
RRM1+2 K200RK208R	pet28b(+)	His	J. Kühnert
SAE1	pGEX-2T	GST	PPL (MUSC)
SAE2	pGEX-2T	GST	PPL (MUSC)
Ubc9	pet11d	His	PPL (MUSC)
SUMO-1	pGEX-2T	GST	PPL (MUSC)
SUMO-2	pGEX-2T	GST	PPL (MUSC)
SUMO-3	pGEX-2T	GST	PPL (MUSC)
SUMO-1WT	pet11d	His	PPL (MUSC)
SUMO-2WT	pet11d	His	PPL (MUSC)
SUMO-3WT	pet11d	His	PPL (MUSC)
SUMO-1GA	pet11d	His	PPL (MUSC)
SUMO-2GA	pet11d	His	PPL (MUSC)
SUMO-3GA	pet11d	His	PPL (MUSC)

Table 3.1: Overview of recombinant proteins. Displayed are the protein names and vectors that cDNA was cloned into, either a hexahistidine (His) tag or glutathione S-transferase (GST)- tag have been fused to the N'-terminus of the cDNA to allow protein purification. The protein purificiation was either performed in small scale by Julia Kühnert or large scale using either HiTrap HP Ni-NTA (His-tagged proteins) or a GSTrap FF column (GST-tagged proteins) by the Protein Production Lab at the Medical University of South Carolina, USA. His = hexa histidine tag, GST = glutathione S-transferase tag, PPL = protein production lab, MUSC = Medical University of South Carolina, Charleston, SC, USA

3.2.2 Microspin technology for affinity purification of His-tagged recombinant proteins using Nickel-NTA spin columns

The recombinant protein purification is based on the small affinity tag consisting of six consecutive histidine residues, the 6x His-tag, which has been introduced in the cDNA of the proteins of interest. The Ni-NTA affinity purification system is based on the high affinity of the Ni-NTA resins for hexa-histidine-tagged proteins. Whereas nickel ions are a highly specific histidine residue ligand, those nickel ions are bound with a high specificity by the chelating group of nitrilotriacetic acid (NTA). Therefore, the system allows an efficient purification of His-tagged recombinant proteins. Imidazole is a competitor for histidine when in surplus it displaces the histidine-tagged proteins from the Ni-NTA matrix. Thus, the elution of the recombinant proteins is carried out by using a high concentration of imidazole in the elution buffer.

For the purification of the His-tagged recombinant proteins wild type La (LaWT, RCD1, RCD2, RRM1+2, RRM1+2 K200R, and RRM1+2 K200RK208R were expressed in *E. coli* BL21 (3.2.1), then purified using Ni-NTA spin columns following the manufacturer's instructions (Qiagen Ni-NTA Spin Handbook, 2nd edition, 2008). Protein solutions as well as buffers were either kept on ice, or at 4°C, at all times and 1x cOmplete protease inhibitor were added freshly to all buffers.

Spin columns were equilibrated with 600 µl lysis buffer (50 mM NaH$_2$PO$_4$, 10 mM imidazole, 300 mM NaCl, 1 mg/ml lysozyme, 1% (w/v) cOmplete protease inhibitor) by centrifugation at 890 x g for 2 min prior to binding of His-tagged proteins to the Ni-NTA resin. For binding 600 µl of the cleared lysate containing the protein of interest were loaded on a spin column and centrifuged at 270 x g for 5 min. This step was repeated until all of the lysate was applied to the spin column. The flow through was collected and re-applied twice in order to ensure high efficient protein binding by saturating the Ni-NTA resin with His-tagged protein. The columns were subsequently washed six times with 600 µl wash buffer (50 mM NaH$_2$PO$_4$, 42.5 mM imidazole, 1 M NaCl, 0.11% (v/v) Triton X-100, 1% (w/v) cOmplete protease inhibitor) for removal of contaminants, which were unspecifically adsorbed by the resin, by centrifugation for 2 min at 700 x g. Proteins were eluted three times in 150 µl elution buffer (50 mM NaH$_2$PO$_4$, 300 mM imidazole, 300 mM NaCl, 1% (w/v) cOmplete protease inhibitor) and centrifuged for 2 min at 700 x g.

Aliquots for analysis by SDS-PAGE and subsequent Coomassie staining were collected after binding, from wash fractions 1, 3, 5, and of the eluate, which contains the purified recombinant protein.

3.2.3 Dialysis of recombinant proteins

The purification of proteins requires buffer conditions, such as high salt and imidazole concentrations, which can have negative effects on downstream applications. Dialysis is a common method allowing the adjustment of salt concentrations and eliminatation of interfering buffer components. During dialysis molecules in solution are separated by diffusion through a semi-permeable dialysis membrane. The protein in solution is filled into the dialysis tubing, a membrane with a defined molecular weight cut-off (MWCO), then placed into a container with the buffer of interest. Molecules that are below the MWCO will freely diffuse into the solution whereas proteins and molecules larger than the MWCO will be retained in the tubing. This allows the exchange of the purification buffer to the buffer of interest.

The Ni-NTA purified recombinant proteins were placed into membrane dialysis tubing with a MWCO of 6-8,000 Da, which was then sealed on both sides. The buffer was exchanged twice after one hour against 500 ml dialysis buffer (10 mM Tris/HCl pH7.4, 150 mM NaCl, 3 mM MgCl$_2$, 5% (v/v) glycerol) and overnight against 1,000 ml buffer. The buffer was pre-chilled and dialysis was carried out at 4°C, the buffer was continuously stirred to ensure an even distribution of the buffer and salt concentrations. Dialyzed proteins were carefully removed from the tubing and potential precipitated proteins were removed by centrifuging the protein solution for 5 min at 18,000 x g.

3.2.4 Protein concentration

In some cases the protein concentration was inapplicable for downstream applications and had to be concentrated. A low-binding, anisotropic and hydrophilic regenerated cellulose membrane with a MWCO of 10,000 Da from Amicon was used for concentrating purified recombinant proteins.

The protein solution was added to the membrane, centrifuged at 14,000 x g and 4°C to the desired volume and subsequently recovered by placing the sample reservoir upside down and centrifuging for 3 min at 1,000 x g to transfer the concentrate to a clean tube.

3.2.5 Protein quantification

All purified recombinant proteins were quantified using the Bradford method [317], which is a rapid and sensitive method for protein quantification. This method is based on the observation that the absorbance maximum of Coomassie Brilliant Blue G-250 shifts from 495 nm to 595 nm when a protein binds to the dye in acetic solution occurs. The visible color change, monitored with a photometer, is caused by the stabilization of the anionic form of the dye. In order to quantify the protein sample, the color change is compared to the binding of known amounts of bovine serum albumin (BSA) as a standard curve.

All protein quantification assays were performed in a 96-well plate, the absorption at 595 nm was determined using a multiplate reader. A standard curve was prepared by carefully pipetting BSA in concentrations between 0.1 and 10 mg/ml, adding an aliquot of diluted dye reagent concentrate (1:5 (v/v) in water) and measuring the absorbance after a 5 min incubation. The optical densities (ODs) were corrected versus the blank, the diluted dye reagent without protein, by subtracting the blank value from the data. The blank corrected raw data were graphically plotted against the known protein concentrations.

Protein samples were measured in triplicates, according to the expected protein concentration 1 - 10 μl of protein was placed in a 96-well assay plate and the diluted dye reagent was added ad 180 μl. After the assay was thoroughly mixed, it was incubated for 5 min and the OD was measured at 595 nm. The average of the optical densities was determined after blank correcting the raw data. In order to obtain the protein concentration of the test samples the average OD_{595nm} was compared with the standard curve by reading off the protein concentration.

Since some of the recombinant purified proteins were contaminated and/or partially degraded, the true concentration of those proteins could not be determined with a calorimetric method as described above. Therefore, proteins were immunoblotted against La (as described in 3.2.9) and the chemiluminescent signal of the protein samples were quantified and compared with recombinant La of a known concentration that was blotted in parallel. The advantage of this method was only La specific bands were quantified and not any unspecific protein.

3.2.6 Protein modification - optimized *in vitro* sumoylation assay

In order to study the role of SUMO-modification on La RNA-binding activity recombinant human La protein and La protein mutants were modified by the small ubiquitin modifier SUMO *in vitro*.

The optimized *in vitro* assay to modify 425 nM recombinant La proteins was performed by using 89 nM GST-SAE1, 80 nM GST-SAE2, 0.4 μM His-Ubc9 and 7.3 μM His-SUMO-1/-2/-3WT or GST-SUMO-1 and parallel reactions with the conjugation deficient His-SUMO-1/-2/-3GA mutants. The reaction mixture was prepared in a total volume of 25 μl in sumoylation buffer (50 mM Tris/HCl pH 8.0, 100 mM NaCl, 0.1% (v/v) Nonidet P-40, 10% (v/v) glycerol, 1 mM Dithiothreitol) containing 1 x ATP mix (10 mM creatine phosphate, 5 mM $MgCl_2$, 2 mM rATP), 1 U creatin kinase, and 1x cOmplete protease inhibitor mix. Reactions were incubated for 2 hours at 30°C and either stored at 4°C for functional EMSA assays (3.3.6.2) or acetone precipitated with 250 μl ice-cold 100% acetone over night at -20°C for immunoblot analysis. The precipitated samples were pelleted for 20 min at 18,000 x g at 4°C in a bench top centrifuge and were washed twice in 250 μl 80% (v/v) ice cold acetone, then the pellet was air-dried. The pellet was prepared and separated by SDS-PAGE and subjected to immunoblotting as described in 3.2.7 and 3.2.9.

3.2.7 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

A common method for separating proteins is electrophoresis, which is the separation of macromolecules according to size and charge in an electric field, using a polyacrylamide gel as matrix. Sodium dodecylsulfate (SDS) is an anionic detergent that denatures hydrophobic molecules such as polypeptides. The number of amino acids bound by SDS is proportional to the length of a polypeptide. Further, intrinsic charges of polypeptides become ignorable due to the negative charges contributed by SDS. A polyacrylamide gel allows smaller molecules to migrate faster through the matrix than larger molecules. The separation of proteins in a SDS-polyacrylamide gel electrophoresis (SDS-PAGE) is independent of charges since all polypeptides are masked by anionic SDS, but dependent on differences by their molecular mass. To ensure that proteins are truly in their monomeric form, sulfhydryl-reducing agents, e.g. β-mercaptoethanol, are often used to reduce disulfide bonds between polypeptide chains. SDS-PAGEs were performed following the method described by Laemmli [318] using the Mini-PROTEAN-3 and Tetra Cell system. The concentration of the resolving polyacrylamide

gel was either 7.5% or 12.5% depending on the molecular weight of the protein of interest; the analysis of *in vitro* sumoylation of wild type La required a lower gel percentage whereas 12.5% resolving gels were used for all other analyses. The stacking gel concentration was 5% for all assays. All reagents for the resolving gel were combined before adding TEMED and APS, which are initiators of gel polymerization. During polymerization, one hour for 7.5% and 45 min for 12.5% polyacrylamide gels, the gel was covered with a layer of isopropanol to ensure an even surface without air bubbles. After discarding this layer and drying the gel with a tissue, the stacking gel was poured and a 1 mm 10-well comb was inserted immediately after pouring followed by allowing the gel to polymerize for at least 30 minutes. Protein samples were prepared by adding 3x SDS-loading buffer (60 mM Tris/HCl pH 6.8, 10% (v/v) glycerol, 6% (v/v) SDS, 0.01% (w/v) bromphenol blue, and freshly added 5% (v/v) β-mercaptoethanol) and boiling them at 95°C for 10 min at 1,400 rpm in a thermomixer to denature the protein samples and coat them with SDS. Denatured samples centrifuged at 14,000 x g for 1 min and loaded on the SDS-polyacrylamide gel. The proteins were separated by electrophoresis at 200 V for 55 min in SDS-running buffer (250 mM Tris, 19.2 mM glycine, 0.35 mM SDS).

3.2.8 Protein staining in polyacrylamide gels with Coomassie-Brilliant-Blue

Proteins can be visualized in a polyacrylamide gel after electrophoresis by staining proteins with Coomassie-Brilliant-Blue R-250 according to Meyer and Lambert [319]. Proteins in the gel must be fixed prior or simultaneously to staining in order to prevent diffusion of proteins and remove SDS, which may interfere with the Coomassie staining. Therefore, gels were stained under slow agitation for 20-30 min in Coomassie staining solution (45% (v/v) methanol, 10% (v/v) acetic acid, 0.1% (w/v) Coomassie brilliant blue R-250, 0.1% (w/v) Coomassie brilliant blue G-250). After the gel was washed with water, it was destained in Coomassie destaining solution (45% (v/v) methanol, 10% (v/v) acetic acid). Gels were stored in double distilled water until drying or scanning.

The Coomassie staining as well as the destaining solutions are re-usable for a limited time, the destaining solution containing Coomassie dye was regenerated by incubating it with activated charcoal. The activated carbon binds the dye and the destaining solution can be reused after filtration.

3.2.9 Immunological detection of proteins by immunblotting

The verification of *in vitro* sumoylation as well as the quantification of purified recombinant La proteins was achieved by immunoblotting, which is a sensitive immunological method for detecting specific proteins.

Proteins were transferred to a nitrocellulose membrane after electrophoretic separation according to their molecular size by SDS-PAGE (3.2.7). The SDS-PAGE gels, as well as the nitrocellulose membrane, were equilibrated in pre-cooled (4°C) transfer buffer (25 mM Tris/HCl pH 8.3, 192 mM glycine, 20% (v/v) ethanol 190 proof) for 30 min under slight horizontal shaking. The Mini Trans-Blot Electrophoretic Transfer Cell was used for immunoblotting. A so called transfer sandwich was set up in a cassette by aligning the gel with the membrane in between soaked pieces of Whatman filter paper and fiber pads. The cassette with a cooling unit (pre-cooled at -20°C) was placed in the gel transfer cell and filled with ice-cold transfer buffer. The transfer was carried out at 100 V and 350 mA for 3.5 hours in a cold room for the detection of sumoylated La, and for 1.5 hours at room temperature for the quantification of recombinant La. After electrophoretic protein transfer to the nitrocellulose membrane the transfer sandwich was carefully dissembled, the protein marker highlighted with a pen, and the membrane put in a 50 ml conical screw cap tube. In order to block unspecific antibody binding to the membrane, the membrane was incubated for 30-45 min at room temperature in blocking solution, 2% BSA (w/v) in wash buffer (10 mM Tris/HCl pH 7.5, 100 mM NaCl, 0.1% (v/v) Tween 20), under continuous rocking on a Labquake Rotisserie Shaker. The blocking solution was discarded and the membrane was incubated with the primary antibody over night at 4°C. As primary antibodies for hLa immunodetection served the monoclonal mouse anti-La antibodies 3B9 and SW5. For the detection of SUMO signals, the polyclonal rabbit-SUMO-1 and rabbit-SUMO-2/-3 antibodies were used as primary antibodies. The primary La antibodies 3B9 and SW5 were diluted 1:500 and the SUMO antibodies 1:4,000 in 10 ml 2% BSA in wash buffer (w/v) containing 1:1,000 sodium azide (v/v). The primary antibodies were recollected and stored at 4°C. The membrane was washed three times for 5-7 min in wash buffer. The secondary antibody specifically recognizes and binds an epitope on the primary antibody. The secondary antibody horseradish peroxidase conjugated anti-mouse or anti-rabbit antibodies were diluted 1:20,000 in 10 ml 2% BSA (w/v) in wash buffer and applied for 1 hour at room temperature or 3 hours at 4°C. The membrane was washed again as described above. The SuperSignal West Pico Chemiluminescent Substrate was applied to the membrane. This substrate then

gets cleaved by the horse-radish peroxidase, which is linked to the secondary antibody, resulting in a luminescent signal that was either detected with an autoradiography film or the ImageQuant RT ECL imager.

3.3 RNA synthesis and RNA-protein-interactions

3.3.1 Synthesis of DNA templates for *in vitro* transcription

The pRCD1F plasmid served as a DNA template for the generation of DNA-templates for *in vitro* transcription (described in 3.3.2). The cDNA of the cyclin D1 5'-UTR (209 nts) and the AUG start codon (3 nts) was cloned between two luciferase genes in the bicistronic pRCD1F plasmid. The oligonucleotides serving as PCR primers for the amplification of DNA templates are listed in 2.12.1. The T7 RNA polymerase was used for *in vitro* transcription of the DNA templates to RNA. During PCR a 5'-terminal T7 promoter, which is required for T7 RNA polymerase-mediated transcription, was introduced into all of the DNA templates by the sense PCR primers. RNAs with terminal deletions, D1-FL, part A, part B, ΔB2, ΔB2.1, and ΔB2.2 were *in vitro* transcribed from DNA templates with terminal 5'- or 3'- terminal deletions amplified by a single PCR using the primers listen in section 2.12.1. The GC-RICH PCR System was used to ensure efficient template amplification. Because the 5'-UTR of cyclin D1 is known to be GC-rich, thus amplification of such a GC-rich template can be a difficult task. For each transcript three parallel PCR reactions were prepared using independent master mixes after the manufacturer's instructions:

    reaction mix 1

ad 35.0 µl   PCR grade water

    5.0 µl   2 mM dNTPs (final concentration 200 nM of each dNTP)

    1.0 µl   10 µM sense primer (final concentration 200 µM)

    1.0 µl   10 µM antisense primer (final concentration 200 µM)

  10.0 ng   pRCD1F template DNA

    2.0 µl   5 M GC-RICH resolution solution (final concentration 2 M)

reaction mix 2

  4.0 µl PCR Grade water

 10.0 µl 5x GC-RICH PCR reaction buffer with DMSO

  1.0 µl GC-RICH PC system enzyme 1 (final concentration 2 U)

Both reaction mixes were combined on ice and gently vortexed before the PCR was performed in the Mastercycler PCR cycler (Eppendorf) after the following thermal profile:

initial denaturation	95°C	3 min	
denaturation	95°C	30 sec	
annealing	55°C	30 sec	10 cycles
elongation	68°C	30 sec	
denaturation	95°C	30 sec	
annealing	55°C	30 sec	25 cycles
elongation	68°C	45 sec	
final elongation	68°C	7 min	

The PCR products were purified in two steps to ensure a high purity DNA template. First, the PCR products were removed from salts and oligonucleotide primers using the QIAquick PCR Purification Kit from Qiagen after the manufacturer's instructions. The purified samples were subsequently separated on an analytical 2.3% agarose gel, the template DNA bands were excised and purified using the QIAquick Gel Extraction Kit from Qiagen. The purified products were assessed for their integrity by 2.3% agarose gel electrophoresis and subsequent ethidium bromide staining.

The DNA template for the ΔB1 transcript (-209 to 110 and -37 to +3) has a 73 nucleotide internal deletion between nucleotides 99 and 172. The internal deletion was achieved by a fusion PCR strategy (as displayed in figure 3.1). In separate PCR reactions the 5'-end and 3'-end of the cyclin D1 5'-UTR were amplified using the D1-1S-T7 as sense primer for the 5'-terminus and D1-4AS as antisense primer for the 3'-end. The antisense primer CD1-30AS required for amplification of the 5'-terminus contains an additional 20 nucleotides. Those additional nucleotides are complimentary to the sequence downstream of the deletion. A 19 nucleotide long overhang upstream of the deleted sequence was created by the D1-29S sense

primer, which was used in combination with the D1-4AS primer to amplify the 3'-end of the cyclin D1 5'-UTR. The D1-30AS antisense and D1-29S sense primer are annealing in a way so both amplification products lack the nucleotides downstream of nucleotide 99 and upstream of nucleotide 172, respectively. In a fusion PCR using the sense D1-1S-T7 primer and the antisense D1-4AS primer both PCR products were fused due to the complimentary overhangs created in the first PCR. All three GC-rich PCRs were performed using the GC-RICH PCR system from Roche as described above, but using 7.5 ng of each PCR product fusion PCR. The PCR products were separated on an analytical 2.3% agarose gel and purified after each PCR amplification using the QIAquick Gel Extraction kit from Qiagen.
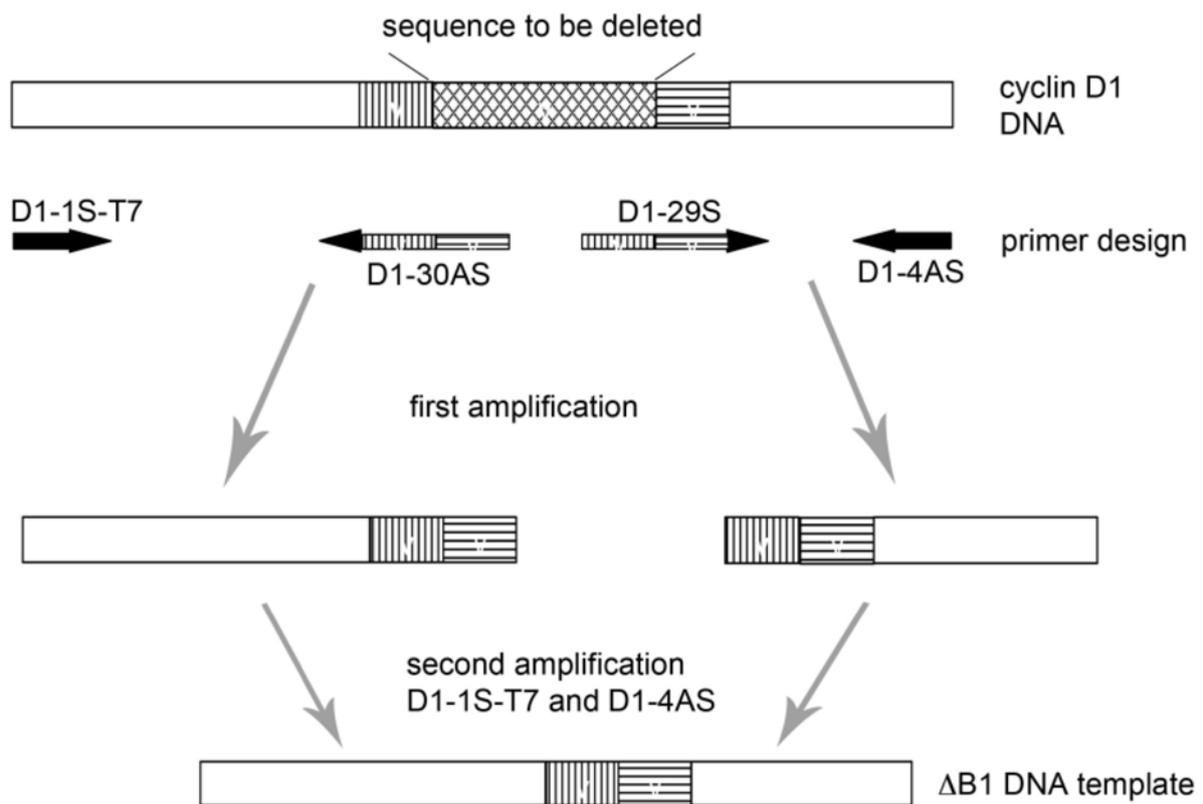
Figure 3.1: Cartoon of internal deletion strategy to create the ΔB1 DNA template. The overhangs up- and downstream of the deleted sequence created by D1-30AS and D1-29S primers are indicated by vertical and horizontal lines, respectively.

3.3.2 Synthesis and purification of RNA by *in vitro* transcription

RNA probes, non-radioactive as well as radioactive [^{32}P]-CTP labeled, were synthesized using the MEGAshortscript High Yield Transcription Kit from Ambion according to the manufacturer's instructions.

For the transcription of internally labeled RNAs reactions were assembled in the following order 2 µl T7 10x reaction buffer, 8 µl of 75 mM T7 ATP-GTP-UTP Mix (18.75 mM each), 1.5 mM cold CTP, 40 µCi [α-^{32}P]-CTP, 100 nM of DNA template containing a T7 promoter, 2 µl T7 enzyme mix, and nuclease-free water ad final volume of 20 µl. For DNA templates smaller than 75 nucleotides, 150 nM of template DNA was used for the *in vitro* transcription. After gently mixing, the reactions were incubated for 3.5 hours at 37°C. In order to remove the DNA template, the samples were incubated for an additional 15 min at 37°C with 1 µl TURBO DNase. Non-radioactive RNA transcripts were synthesized by using a similar reaction, but using a mix of 100 mM T7 NTP Mix (ATP, GTP, UTP, CTP: 25 mM each) instead of unlabeled and [^{32}P]-labeled CTP. An overview of the RNA transcripts is displayed in table 3.2.

The *in vitro* transcribed RNA was in solution with a high content of free nucleotides, salts and enzymes, thus it was purified using a glass-filter based MEGAclear Kit from Ambion according to the manufacturer's instructions.

The freshly synthesized RNA was prepared as follow before being bound to the glass-filter: The RNA was brought to a volume of 100 µl by adding 80 µl of Elution Solution, and subsequently 350 µl of Binding Solution Concentrate followed by 250 µl 200 proof ethanol were added to the samples. The reactions were gently mixed after each step by gently pipetting. To bind the RNA, the RNA mixture was pipetted onto the membrane and centrifuged for 1 min at 13,000 x g. After discarding the flow-through, the membrane was washed twice by applying 500 µl Wash Solution Concentrate and centrifuging for 1 min at 13,000 x g after each wash step. An additional centrifugation step for 30 sec at 13,000 x g ensured the complete removal of wash solution. The RNA was eluted in two steps: 50 µl of pre-heated 95°C Elution Solution were applied to the center of the filter, which was placed into a clean Collection/Elution tube, and centrifuged for one min at 13,000 x g. This step was repeated to maximize recovery.

The RNA yield was determined in a 1:10 dilution by diluting 3 µl of RNA to 27 µl 1x TE buffer (10 mM Tris/HCL pH 8.0, 1 mM EDTA). The absorbance at 260 nm was spectrophotometrically determined using a NanoDrop. The RNA concentration was

calculated based on A_{260} of 1 equivalent to 40 ng RNA/µl: RNA concentration = $A_{260}*10*40$ ng/µl, with 10 being the dilution factor.

transcript name	mutation type	nts position	length
D1-FL	none	-209 to +3	212
part A	5'-terminal deletion	-209 to -110	99
part B	3'-terminal deletion	-110 to +3	113
ΔB1	internal deletion	-209 to -110 and -37 to +3	139
ΔB2	3'-terminal deletion	-209 to -37	172
ΔB2.1	3'-terminal deletion	-209 to -15	194
ΔB2.2	3'-terminal deletion	-209 to -8	201

Table 3.2: Overview of RNA transcripts used for electrophoretic mobility shift assays. Type of mutation introduced into the D1-FL RNA and the nucleotide length of the transcripts. The nucleotide positions of the transcripts in reference to the D1-FL RNA are indicated. nts = nucleotides

3.3.3 5'-end labeling and purification of synthesized RNA oligoribonucleotides

To test whether La has RNA-binding activity to an RNA oligoribonucleotide containing the cyclin D1 translational start side in its original context, the D1-ATG RNA oligoribonucleotide was synthesized by Integrated DNA technologies, Inc. (sequence information in section 2.12). The lyophilized RNA was resuspended in ribonuclease-free 1x siRNA buffer to a stock solution concentration of 100 nM. The RNA was 5'-end labeled with [γ-^{32}P]-ATP using the KinaseMax Kit from Ambion, which uses the enzyme T4 polynucleotide kinase to catalyze the gamma phosphate transfer from ATP to the 5'-hydroxyl group of the RNA. The manufacturer's instructions were followed. Since the D1-ATG RNA is a synthesized oligoribonucleotide it has a 5'-OH group and does not have to be de-phosphorylated in contrast to *in vitro* transcribed RNA, which has a 5'-phosphate. The kinase reaction was performed in a total volume of 20 µl. A set of three parallel reactions were prepared when large amounts of radiolabeled D1-ATG RNA were anticipated. 20 pmol D1-ATG RNA oligoribonucleotide were phosphorylated using 3 µl 3,000 Ci/mmol, 10 mCi/ml [γ-^{32}P]-ATP, 2 µl 10x Kinase Buffer, 10 U T4 polynucleotide nuclease ad 20 µl nuclease-free water for 1 hour at 37°C. The reaction was quenched by adding EDTA to a final concentration of 1 mM and heating the samples to 95°C for 2 min.

In order to remove free, unincorporated nucleotides and salts the kinase product was purified by spin-column chromatography using NucAway Spin columns after manufacturer's instructions. Major advantages of using spin columns compared to gel purification followed

by phenol:chloroform extraction include time savings and the elimination of hazardous phenols. During the kinase reaction the spin column were prepared by settling the dry gel in the bottom of the column and adding 650 µl RNase-free water, tapping out air bubbles and hydrating the dry gel for 10 min at room temperature. The rehydrated column was spun for 2 min at 750 x g to achieve a proper column packing. The kinase max reaction was loaded on the center of the column, reactions where pooled and applied to one column when parallel kinase reactions were prepared, and the purified RNA was then eluted by spinning the column at 750 x g for 2 min. The integrity of the 5'-end labeled RNA was assessed by denaturing TBE-urea PAGE (section 3.3.4).

3.3.4 Denaturing TBE-urea polyacrylamide gel electrophoresis for the analysis of radioactive labeled RNAs

In order to verify the correct size as well as the homogeneity and quality of *in vitro* transcribed [^{32}P]-labeled RNAs they were subjected to denaturing urea PAGE. RNA can form very stable and compact secondary structures, which are falsifying the relation between mobility and molecular weight. The bases of RNA may form hydrogen bonds, which are difficult to denature, thus, a denaturing agent such as urea must be present in a polyacrylamide gel to disrupt the hydrogen bonds and separate RNA under denaturing conditions.

A 5% TBE-urea polyacrylamide mini gel for the Mini-PROTEAN Tetra Cell was prepared for the analysis of larger radiolabeled transcripts (D1-FL, part A, part B) and unlabeled transcripts (D1-FL, part A, part B, ΔB1, ΔB2, ΔB2.1, ΔB2.2) with 8 M urea, 0.6 ml 40x TBE buffer (1.8 M Tris/HCl pH 8.5, 1.8 M boric acid, 40 mM EDTA), 1.25 ml 30% acrylamide/bisacrylamide (29:1, v/v) ad 7.5 ml deionized water, then polymerization was initiated by adding 0.01% (w/v) APS and 0.001% (v/v) TEMED. For the smaller D1-ATG oligoribonucleotides and related RNAs (labeled D1-ATG, unlabeled D1-ATG, mu2, and mu3) a 10% denaturing gel was prepared accordingly. The comb was inserted and the gel was allowed to polymerize for one hour before it was pre-run at 240 V for 35 min, or until current was invariant, in 1x TBE buffer (45 mM Tris/HCl pH 8.5, 45 mM boric acid, 1 mM EDTA) to remove all traces of APS and to ensure that buffer and gel are at the same temperature. 10 µl of the *in vitro* transcribed RNA was denatured in 10 µl denaturing RNA-loading buffer (80% (v/v) formamide, 1 mM EDTA, 50 mM 1x TBE buffer, 0.05% (w/v) bromphenol blue, 0.05% (w/v) xylene cyanole FF) at 80°C for 10 min and cooled down to room temperature

prior to loading. Electrophoreses were carried out at 240 V for 20-35 min depending on the RNA size. The RNA-loading buffer contains bromphenol blue as well as xylene cyanol, which migrate at 35 and 140 base pairs respectively in a 5% denaturing polyacrylamide gel and at around 12 and 55 base pairs in a 10% denaturing gel. Therefore, electrophoresis times were estimated according to the dyes' migration and the expected RNA sizes. After the electrophoresis the gel system was dissembled, the gel was placed on a Whatman filter and wrapped in plastic wrap. The RNA quality was determined by exposing the gel to a storage phosphor screen and visualized using a Storm Phosphorimager scanner.

3.3.5 Evaluation of RNA quality by denaturing urea polyacrylamide gel electrophoresis of non-radioactive labeled RNA

To assess the quality as well as correct size of *in vitro* transcribed non-radioactive labeled RNA the detection method as described above could not be performed. Nucleic acids can be detected by using ethidium bromide (EtBr) stain. EtBr intercalates in the spaces between base pairs in the double helix of nucleic acids. Single stranded RNA can also be detected using EtBr, but it is less efficient. The denaturing PAGE was performed as indicated above, but the TBE polyacrylamide gel contained 8 M urea. 400 ng RNA were loaded in 5 μl denaturing RNA loading buffer (80% (v/v) formamide, 1 mM EDTA, 50 mM 1x TBE buffer, 0.05% (w/v) bromphenol blue, 0.05% (w/v) xylene cyanole FF) after being denatured at 90°C for 10 minutes. An electric field was applied for 25 min at 240 V. Upon dissembling the gel system the gel was placed in a solution of 0.5 μg/ml EtBr for 5 min and subsequently washed for 2 min in distilled water. The EtBr intercalated into the RNA was visualized under UV light using the UV-Transilluminator Universal Hood and Image Lab software.

3.3.6 Electrophoretic mobility shift assay using radioactive labeled RNA

Electrophoretic mobility shift assay (EMSA) is a sensitive quantitatively as well as qualitatively technique allowing the analysis of protein-nucleic acid interactions. Protein and nucleic acid solutions are combined and the mixture is separated by native polyacrylamide gel electrophoresis (PAGE). The species of nucleic acid is usually determined by autoradiography of the [32P]-labeled nucleic acid. Commonly, the free nucleic acid has a higher mobility compared to nucleic acids shifted into a complex with the protein and migrating slower through the polyacrylamide gel. In this study the La protein was combined

with [32P]-labeled RNAs and the protein-RNA solution subjected to native TBE polyacrylamide gel electrophoresis. The [32P]-labeled nucleic acid species were detected by phosphor screen imaging (3.3.6.7).

3.3.6.1 Standard EMSA

Several recombinant La mutants were tested for their ability to form complexes with the 5'-UTR of cyclin D1. The binding affinity of recombinant human La to the 5'-UTR was determined by titrating La against the RNA. A 6% native TBE polyacrylamide gel was prepared in a Mini-PROTEAN Tetra Handcast System, after polymerization it was pre-run for 35 min at 240 V during sample preparation. The RNA was re-annealed for 10 min at 85°C in 1x RNA-binding buffer (20 mM Tris/HCl pH 8.0, 150 mM NaCl, 1.5 mM MgCl$_2$, 0.5 mM EDTA, 0.5% (v/v) Nonidet P-40) and slowly cooled down to 37°C in a thermomixer. The binding reaction was prepared in a total volume of 20 μl in 1x RNA-binding buffer. Increasing concentrations of recombinant wild type La protein: 40, 80, 160, 320, 640, and 960 nM were placed in a 96-well microtiter assay plate and 10 nM of re-annealed [32P]-labeled cyclin D1-FL RNA was added. A reaction with only the RNA transcript and no protein served as negative control. The samples were equilibrated for 10 min at room temperature before 2.5 μl RNA loading buffer (5% glycerol (v/v), 0.25% (w/v) bromphenol blue in 1x TBE buffer) was added and then samples were subjected to electrophoresis for 10 min at 240 V followed by 60 min at 100 V. The gel was carefully removed from the glass plates, placed on Whatman filter paper and dried on a gel dryer for 45 min at 80°C before placing it in a cassette and finally exposing it to a storage phosphor screen. The phosphor screen was scanned using a Storm PhosphorImager. The La:RNA complex formation was quantified using the ImageQuant TL software and plotted as a function of the La concentration in Prism 5. The dissociation constant was determined using the one-site hyperbolic binding fit in Prism 5.

The binding affinity of wild type recombinant human La protein to the translational start site context of the cyclin D1 RNA, represented by the D1-ATG RNA oligoribonucleotide, was determined by native EMSA. A 10% native TBE polyacrylamide gel was prepared and allowed to polymerize for one hour before it was pre-run for 40 min at 260 V in 1x TBE buffer (45 mM Tris/HCl pH 8.5, 45 mM boric acid, 1 mM EDTA). In the meantime the [32P]-D1-ATG RNA oligoribonucleotide in 1x RNA-binding buffer was re-annealed for 10 min at

85°C and subsequently cooled down slowly to 37°C in a thermomixer. The binding reaction was prepared in a total volume of 20 μl in 1x RNA-binding buffer. Increasing concentrations of human recombinant His-tagged La protein of 10 nM, 60 nM, 200 nM, 400 nM, 600 nM, 800 nM, 1 M, 1.6 μM, and 3 μM were placed into a 96-well microtiter plate. 10 nM of the re-annealed [^{32}P]-labeled D1-ATG RNA was added to the La protein. The protein was allowed to bind the RNA for 10 min at room temperature before 2.5 μl RNA-loading buffer was added to each sample and loaded on the pre-run native TBE polyacrylamide gel. The electrophoresis was carried out for 10 min at 260V followed by 50 min at 140V. The gel was carefully removed, placed on Whatman filter paper, and dried for 60 min at 80°C under vacuum. The dried gels were exposed to a storage phosphor screen overnight and visualized using a Storm PhosphorImager scanner. The ribonucleoprotein complex formation was quantified using the ImageQuant TL software. As described above, the dissociation constant was determined in Prism 5 after plotting the La:D1-ATG RNA complex formation as a function of the La protein concentration.

In order to determine the La domains required for D1-ATG RNA-binding, recombinant La protein mutants were titrated against 10 nM 5'-end labeled [^{32}P]-D1-ATG RNA oligoribonucleotide and separated by electrophoresis on a 10% native TBE polyacrylamide gel. The native EMSA was carried out as described above, but only four protein concentrations for each La protein mutant were chosen: 40 nM, 80 nM, 160 nM, and 320 nM. The oligomerization of RRM1+2 was analyzed by using the highly negatively charged polyanion heparin as an unspecific competitor. The EMSA was performed as described above, but 99, 198, 298, and 596 nM of RRM1+2 were incubated with [^{32}P]-D1-FL RNA in the presence or absence of 0.5 μg heparin.

The La binding site within the 5'-UTR of cyclin D1 was attempted to be mapped using different [^{32}P]-labeled cyclin D1 RNA transcripts containing deletions. The *in vitro* transcribed and internally radiolabeled RNAs, D1-FL, part A and part B (refer to 3.3.2 and 3.3.3), were used as targetes for RNA-binding studies with 320 nM recombinant La protein. The EMSA was carried out as described before for D1-FL RNA binding studies using 10 nM of each transcript, thus allowing it to be bound by 320 nM recombinant His-tagged La protein.

3.3.6.2 EMSA studying SUMO-modified La binding activity

In order to study the effects of SUMO-modification of the La protein on RNA-binding recombinant human La was SUMO-modified and analyzed by native EMSA. Initially, optimal salt conditions for the electrophoretic mobility shift assay had to be created by exchanging the sumoylation buffer (50 mM Tris/HCl pH 8.0, 100 mM NaCl, 0.1% (v/v) Nonidet P-40, 10% (v/v) glycerol, 1 mM Dithiothreitol) to 2x RNA-binding buffer (40 mM Tris/HCl pH 8.0, 300 mM NaCl, 3 mM MgCl$_2$, 1 mM EDTA, 1% (v/v) Nonidet P-40) by utilizing centrifugal filter tubes from Amicon with a MWCO of 10 kDa. The IVSA products were applied to the filter tubes and spun for 10 min at 14,000 x g. A three-fold volume of 2x RNA-binding buffer was added after the concentration of the IVSA product and centrifuged at 10 min for 14,000 x g. This step was repeated 3 times. The end volume was brought to the same volume as the IVSA product with 2x RNA-binding buffer. In order to study the RNA-binding activity of sumoylated wild type La vs. native La 10 nM [^{32}P]-labeled D1-FL RNA was incubated with 27, 80, and 106 nM of sumoylated and native La, which was incubated with the conjugation deficient SUMO-GA mutant, and separated by native PAGE as described above. 80 nM of sumoylated and native RRM1+2 and its respective mutants, RRM1+2 K208R and RRM 1+2 K200RK208R, were allowed to form RNP complexes with 10 nM [^{32}P]-labeled D1-ATG RNA and finally analyzed by native EMSA as described above.

3.3.6.3 DNA:RNA hybrid EMSAs

An alternative approach to the classical EMSA was performed by using antisense DNA oligoribonucleotides complimentary to the cyclin D1 RNA to map the binding site of La within the 5'-UTR of cyclin D1 *in vitro*.

Antisense oligoribonucleotides of about 50 nts complimentary to the D1-FL RNA were synthesized by Integrated DNA Technologies Inc., their sequence information can be found in 2.12.2. The lyophilized oligoribonucleotides were resuspended in nuclease-free water to a concentration of 100 nM. The DNA concentration was confirmed by spectrophoretic analysis using a NanoDrop. A DNA:RNA hybrid was created by annealing 10 nM of [^{32}P]-D1-FL RNA with 20 nM antisense DNA oligoribonucleotide at 85°C for 10 min in annealing buffer (10 mM Tris/HCl pH 7.4, 40 mM NaCl, 0.2 mM EDTA) and subsequently cooling it down to room temperature. The EMSA was performed as described in 3.3.6.1 using 320 nM recombinant human La protein. The samples were prepared in the presence of 0.5 μg heparin

and separated for 8 min at 240V followed by 27 min at 160V a non-denaturing 4% TBE polyacrylamide gel.

3.3.6.4 Competitive EMSA

A variation of the gel retardation assay, a competitive EMSA, allows analyzing the specificity of protein-RNA interactions. In this study, cold competitor RNAs were used to map the La binding site in the cyclin D1 mRNA. Only in the presence of an excess amount of unlabeled competitor RNA including the hLa binding site out-competes the interaction of hLa with the labeled RNA resulting in a reduction or complete loss of an EMSA shift.

Electrophoretic mobility shift assays were conducted as described above (3.3.6.1), but increasing amounts of cold competitor RNA was added and equilibrated for 10 min with 160 nM recombinant La protein before 10 nM radioactive labeled D1-FL or D1-ATG RNA was added. Cold competitors were added in excessive amounts of 100 nM 10x), 250 nM (25x), and 500 nM (50x) in order to challenge the binding. The specificity of the assay was tested by utilizing the cold transcripts of the radiolabeled D1-FL RNA or the unlabeled D1-ATG RNA oligoribonucleotide as competitor RNAs.

In addition to the RNA transcripts listed in table 3.2, the following synthetic RNA oligoribonucleotides were used as competitors:

- D1-ATG RNA (nts -23 to +24)
- mu2 RNA (nts -17 to +18)
- mu3 RNA (nts -17 to +18)
-

3.3.6.5 Supershift assay

An EMSA displays protein:RNA interactions, but it is unable to identify the protein mediating this interactions. A modification of the classical EMSA, supershift assay, allows for the specific immuno-identification of proteins actually binding the RNA. Incubating the protein of interest with a highly specific antibody leads to a much larger complex when bound to the RNA. The lower mobility of the antibody:protein:RNA complex results in a so called "super shift" of the RNA into a lower mobility complex in an EMSA. 300 ng of human recombinant La protein was incubated for one hour at 4°C with 5 μl of the mouse monoclonal La SW5 antibody in 10 mM Tris/HCl pH 7.4 and 20 units RNasin in a total reaction volume

of 10µl. A parallel reaction with 5 µl of the isotype control mouse IgG2α,κ antibody was performed. The antibody pre-incubated La protein was subsequently used for EMSA as described in 3.3.6.1.

3.3.6.6 Non-radioactive electrophoretic mobility shift assay

To examine the RNA-binding activity of the La protein to TOP elements, fluorescence-based native EMSAs were performed. The use of fluorescent labeled RNA probes compared to radioactive labeled RNA has advantages such as longer probe shelf life and the handling of less hazardous materials [220].

The TOP RNA-oligoribonucleotides L22 and L37, and the oligonribonucleotide of the unrelated GAP43 mRNA were synthesized and labeled with Cy3 cyanine at their 3'-end by Integrated DNA Technologies Inc. The lyophilized RNAs were resuspended in 1x siRNA buffer (Dharmacon) to a stock solution concentration of 100 nM.

A native 10% TBE polyacrylamide gel was prepared and allowed to polymerize for 1.5 hours. The RNA-binding reaction was prepared during the pre-run of the polyacrylamide gel at 240 V for 2.5 hours in 1x TBE buffer (45 mM Tris/HCl pH 8.5, 45 mM boric acid, 1 mM EDTA). The RNA from stock solutions was diluted 1:1,000 in annealing buffer (10 mM Tris/HCl pH 7.4, 40 mM NaCl, 0.2 mM EDTA) and re-annealed at 80°C for 5 min and then immediately placed on ice.

For EMSAs to determine the dissociation constant, 22 nM of re-annealed L22, L37, or GAP43 RNA were added to increasing recombinant La protein concentrations ranging from 6.7 nM to 638 nM in 1x RNA-binding buffer (20 mM Tris/HCl pH 8.0, 150 mM NaCl, 1.5 mM MgCl$_2$, 0.5 mM EDTA, 0.5% (v/v) Nonidet P-40) in a total reaction volume of 40 µl. The RNA:protein mixture was incubated for 10 min at room temperature, subsequently loaded on the pre-run gel in the absence of a RNA-loading buffer and separated by native PAGE at 240V for 20 min and 140V at 3 hours. After electrophoresis, the gel was removed carefully from the plates and wrapped in plastic wrap. The fluorescent signals of La:Cy3-RNA complexes were visualized with a Typhoon Imager and quantified using ImageQuant TL software. The quantified RNP complex formation was plotted as a function of the La protein concentration.

For the analysis, RNA-binding activity of SUMO-modified La protein the Cy3-labeled L37, L22, and GAP43 RNAs were re-annealed and the native TBE polyacrylamide gel was prepared as described. In order to create the salt conditions optimal for the electrophoretic

mobility shift assay the salt conditions had to be adjusted. The sumoylation buffer (50 mM Tris/HCl pH 8.0, 100 mM NaCl, 0.1% (v/v) Nonidet P-40, 10% (v/v) glycerol, 1 mM Dithiothreitol) was exchanged to 2x RNA-binding buffer (40 mM Tris/HCl pH 8.0, 300 mM NaCl, 3 mM MgCl$_2$, 1 mM EDTA, 1% (v/v) Nonidet P-40) by using centrifugal filter tubes with a modified polyethersulfone membrane and a MWCO at 10 kDa (VWR) as described in section 3.3.6.2. For the binding reaction 22 nM of L22 or L37 RNA were added to 27, 80, and 106 nM La protein from the IVSA reaction (S-1:La and con La) in 1x RNA-binding buffer in a total reaction volume of 40 μl. Binding was allowed for 10 min at room temperature, samples were loaded without RNA-loading buffer onto the pre-run native 10% TBE polyacrylamide gel and separated by electrophoresis for 30 min at 240 V followed by 2 hours at 160 V. The gels were carefully removed after gel electrophoresis and wrapped in plastic wrap to transport them to the Typhoon imager, which allowed the visualization of the fluorescent-labeled La-RNP complexes. The fluorescent Cy3-RNP complexes were quantified using ImageQuant TL software. The value of unmodified La at 27 nM was set to 1000 arbitrary units and all other data were normalized against unmodified La. The significance (*p*-value) of S-1:La- and control La-binding to L37 and L22 RNA-oligoribonucleotides were calculated using a paired two-tailed Sutdent's t-test in Prism 4.

3.3.6.7 Quantitative EMSA analysis

For radioactive EMSAs dried native PAGE gels were exposed to a phosphor screen for 1.5 hours or overnight depending on the signal strength. The phosphor screen was scanned on a Storm PhosphorImager, which exposes the storage phosphor screen to ionizing radiation and induces latent image formation. Upon laser scanning the crystals in the phosphor screen emit blue light that is collected by the instrument and formed into a quantitative representation of the sample.

The electric signal of the scanned gels were digitalized and analyzed by the ImageQuant TL software, areas were assigned to all RNP complexes to determine the signal strength, which is expressed in pixel/mm^2. These data were corrected by subtracting the signal of an equal sized area of the background. The RNP complex formation was plotted as a function of the La concentration in the Prism 5 software. The signal strength of the La-RNP formation at the lowest control La concentration was set as 100 % or as 1,000 arbitrary units for EMSAs with SUMO-modified La proteins. The signal strengths of all other samples were quantified and put in relation to the 100% or 1,000 arbitrary units.

For fluorescence EMSAs, plastic wrapped gels were carefully transported and scanned using a Typhoon PhosphorImager. The Cy3 fluorescent dye, which is 3'-end labeled to the RNA, was excited at a wavelength of 532 nm and the emitted light of 580 nm was measured by the Typhoon laser using the 580 DF 30 filter. The visualization and quantification of the RNP complex formation by the ImageQuant software was carried out as described above.

3.3.7 Fluorescence polarization assay

An alternative quantitative method to monitor RNA-protein-interactions is based on the alteration in tumbling properties of a fluorescent ligand upon binding to a larger molecule [220]. Jean Perrin described in 1926 that fluorescent molecules in solution emit light into a fixed plane if the molecules remain stationary during the excitation with plane-polarized light. Molecules are not stationary in solution, but rotate and tumble, smaller molecules tumble especially more than larger molecules. If the fluorophore attached to a molecule tumbles out of plane during excitation the emitted light is in a different plane, i.e. the emitted light is depolarized, measured in parallel and perpendicular planes to the excitation plane. A larger molecule rotates and tumbles less and upon excitation the emitted light remains polarized (figure 3.2). Polarization is expressed in units of millipolarization (*mP*) and is related to the fluorescence intensity (*I*) in the parallel (*para*) and perpendicular (*perp*) planes. Fluorescence polarization is defined as [220]:

$$mp = 1000 \; x \; [\frac{Ipara-Iperp}{Ipara+Iperp}]$$

Fluorescence polarization can be used to study RNA-protein interactions by labeling the RNA with a fluorophore and measuring the polarization upon the addition of varying protein concentrations. The polarization is influenced by the relative concentration of each fluorescent species in the reaction [220], thus the polarization is directly proportional to the protein-complexed RNA at each protein concentration [220].

The RNA of interest, D1-ATG, was synthesized and 3'-end labeled with the fluorophore 6-carboxyfluorescin (6-FAM) by Integrated DNA Technologies, Inc (2.11.4). The lyophilized RNA was resuspended in 1x siRNA (Dharmacon) buffer to a stock concentration of 100 μM. The 3'-FAM-labed D1-ATG RNA stock solution was diluted to 10 μM in FP assay buffer (20

mM Tris/HCl pH 8.0, 150 mM NaCl, 1.5 mM MgCl$_2$, 0.5 mM EDTA, 0.5% (v/v) Nonidet P-40, 2 mg/ml (w/v) bovine gamma globulin) and re-annealed for 10 min at 80° and slowly cooled down to 37°C. A 0.1 µM La protein stock solution in FP assay buffer was prepared at increasing concentrations (1 nM- 300 nM) of the La protein and its mutants were placed in black 384-well plates. 10 nM of re-annealed RNA was added to the proteins and FP assay buffer was added to a total volume of 100 µl per reaction. 100 µl of FP assay buffer was used as blank sample. To determine the background polarization 10 nM free 3'-FAM-labed D1-ATG RNA in FP assay buffer was used as control. The POLARstar Omega plate reader was utilized to measure fluorescence polarization after 5 min incubation at room temperature. The assay was done in quadruplicates for each protein concentration, to reduce the standard deviation a 5x RNA-protein master mix of 500 µl was prepared and 100 µl was added per well. The background polarization was subtracted from the blank corrected data. The change in millipolarization (ΔmP) was plotted as a function of the La concentration in Prism 5. The dissociation constant K$_D$ was determined by the one-site (hyperbola) non-linear fit algorithm in Prism 5.
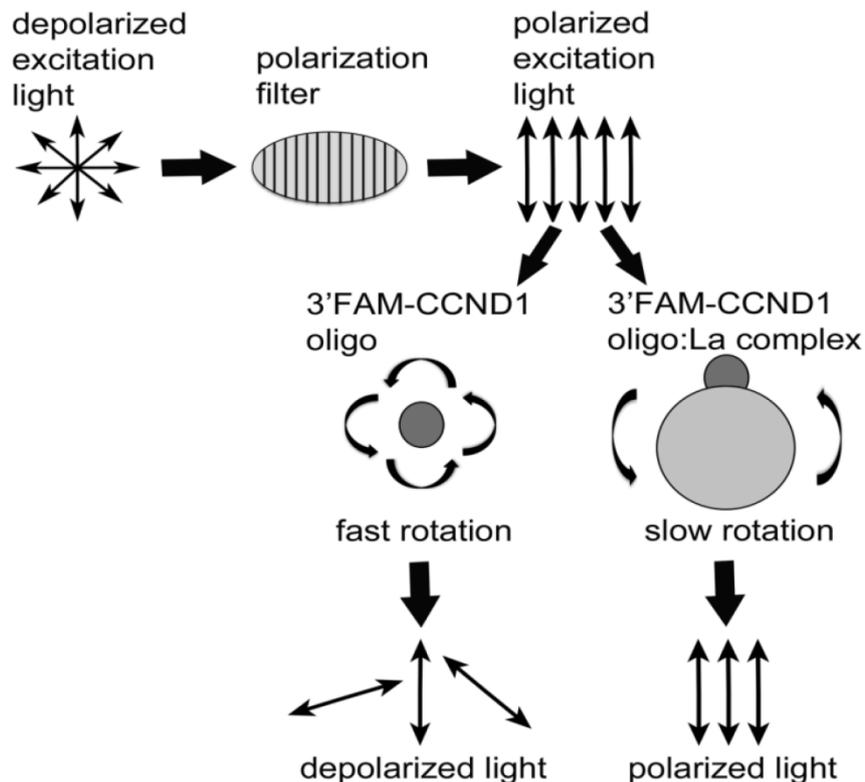
Figure 3.2: Schematic representation of a fluorescence polarization assay. A polarization filter, was used to excite the FAM fluorophore 3'-end labeled to the D1-ATG RNA (3'FAM-CCND1 oligo). The small RNA oligoribonucleotide rotates fast in solution and thus emits depolarized light. Upon binding of the 3'-FAM D1-ATG RNA to the La protein or La protein mutant (3'FAM-CCND1 oligo:La complex) the emitted light is more polarized because of the slower rotation of the larger molecule complex.

4. Results

4.1 Interaction of La with the 5'-UTR of cyclin D1

The human RNA-binding protein La (hLa) has recently been shown to directly interact with the cyclin D1 (CCND1) mRNA in HeLa cells [11]. More interestingly, siRNA mediated knockdown of La in HeLa cells resulted in decreased IRES-dependent translation of cyclin D1, and stimulatory effects of hLa on cyclin D1 IRES-dependent translation were observed when La was overexpressed, [11] suggesting that La facilitates D1-IRES-mediated translation. However, the precise molecular mechanism by which the hLa protein stimulates CCND1 IRES-dependent translation is still elusive.

As described in the introduction (refer to 1.1), there are implications that the La protein may bind in close proximity to the translational start site of mRNAs [62] [111] and supports 48S preinitiation complex formation by recruitment of the 40S subunit to the mRNA, [67] suggesting that hLa plays a role in translational initiation. The notion that La may support the formation of the 48S complex was also suggested by Pudi *et al.* [100]. Because no consensus binding site for the La protein has been described, the binding site of this protein within the CCND1 5'-UTR has to be experimentally mapped as described below.

4.1.1 Protein purification of recombinant human His-tagged La protein

In order to study the interaction of the human La protein with the cyclin D1 RNA *in vitro* the human La protein, containing a hexa histidine tag (His-tag), was expressed in *E. coli* and subsequently purified using Ni-NTA columns as described (3.2.1 and 3.2.2). The prokaryotic expression vector pet28b(+) containing the cDNA sequence of human La has been prepared previously by other lab members.

During the purification of the hLa protein aliquots were taken in order to monitor the purification process, those aliquots were then subjected to SDS-PAGE and Coomassie staining (figure 4.1.1). As seen in figure 4.1.1 the cell lysate (IN lane) contains a variety of proteins including the La protein (indicated by an arrow). Most of those proteins including the hLa protein are found in the flow through (FT lane, figure 4.1.1). The first wash step (W1) removes the majority of unspecific bound proteins; however, subsequent wash steps do not have such a dramatic effect (compare W1 with W3 and W5). As a result, the recombinant

La protein is the prominent band in the purified elution fraction (as indicated by an arrow, EL), which also included two less prominent bands of higher molecular weight. The prominent band was identified as human La by immunoblotting for La (data not shown).

Human La as an RNA-binding protein is able to bind RNA from the prokaryotic cells, thus the contamination of the purified La protein with nucleic acids was analyzed by determining the ratio of spectrophotometric absorbance at 260 nm and 280 nm using a NanoDrop. This ratio is an indication of purity whereas pure proteins without nucleic acid contaminations are considered to have an extinction coefficient ratio of $A_{260/280}=0.57$. The higher the extinction coefficient the more nucleic acids are in the protein sample. The A_{260}/A_{280} ratio of the purified recombinant human La protein was determined as 1.52 using a NanoDrop spectrophotometer. Digestion of the extracts with micrococcal nuclease resulted in a calculated ratio of 1.38, reflecting a slightly less nucleic acid contamination of the hLa protein.
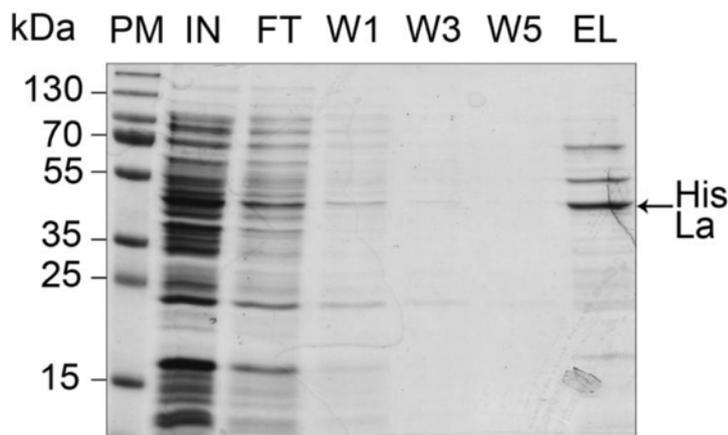
Figure 4.1.1: Protein purification of recombinant human His-tagged La using Ni-NTA columns. During protein purification aliquots were taken and equal volumes were separated on a 12.5% SDS-PAGE and subsequent Coomassie staining. The protein marker, left lane, and the referring molecular weights in kDa are indicated. The main band in the elution fraction below 55 kDa, indicated by an arrow, was identified as human La protein by immunoblotting (data not shown). PM = protein marker, IN = input (lysate), FT = flow-through, W = wash fraction, EL = elution fraction

4.1.2 Characterization of La binding to the cyclin D1 5'-UTR

As mentioned previously the La protein has been shown to directly interact with the cyclin D1 (CCND1) mRNA in HeLa cells and that IRES-mediated cyclin D1 translation is stimulated by the La protein [89]. The minimal functional IRES element that can be stimulated by the IRES-transacting factor (ITAF) hnRNPA1 was mapped to nucleotide -44 to -209 upstream of the initiator start codon of cyclin D1 [213]. Although La has been shown to interact directly with the RNA [89] the binding site has not been determined. Since the La protein is known to bind within the 5'-UTR of the hepatitis C virus (HCV), the polio virus,

and XIAP IRES [62] [67] [91] it is expected that La would bind the 5'-UTR of the cyclin D1 mRNA. Therefore, the binding of recombinant human La (hLa) to the cyclin D1 5'-UTR was analyzed.
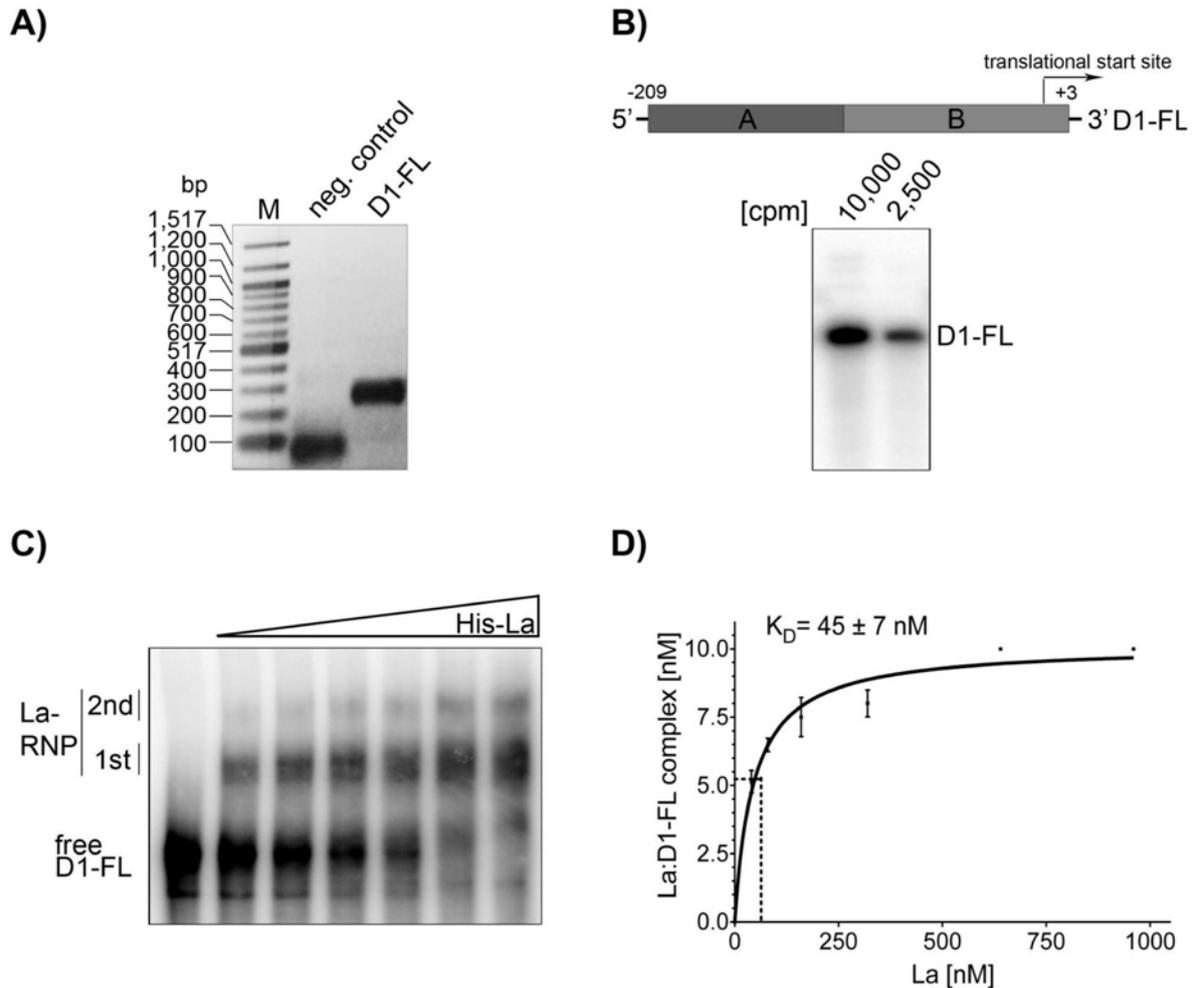
Figure 4.1.2: Binding of recombinant human La to the 5'-UTR of cyclin D1. A) The template for *in vitro* transcription was amplified from the pRCD1F vector by GC-rich PCR and analyzed on a 2 % agarose gel. The DNA ladder and the corresponding DNA sizes are displayed on the very left. The PCR product is of expected size, 244 basepairs. A PCR without the template vector was run in parallel as negative control. The band smaller than 100 bp is the product of primer dimerization. **B)** Cartoon figure of the D1-FL transcript comprised of 209 nts from the 5'-UTR (-209 to -1) and the AUG start codon (+1 to +3). The *in vitro* transcribed [^{32}P]-labeled D1-FL RNA was purified and radioisotope labeling assessed by scintillation counting. To assess the RNA integrity 2,500 and 10,000 cpm RNA were separated by denaturing PAGE and visualized using the Storm and ImageQuant TL phosphorimager system. The gel shows the high purity of the RNA represented by one RNA band for both concentrations. **C)** Recombinant human La was analyzed for binding the CCND1 FL RNA by EMSA. The RNA concentration was held at 10 nM and 40, 80, 160, 320, 640, and 960 nM recombinant La were titrated for the EMSA assay. The RNA is shifted into a La-RNP complex in a La concentration dependent manner. RNA without the addition of protein served as a negative control. The positions of the free RNA and RNA protein complex are indicated on the left. **D)** EMSA quantification of the bound RNA and plotted against the La concentration allowed the determination of the K_D of 45 ± 7 nM, n = 4. bp = basepair, M = marker, CCND1 = cyclin D1, FL = full length, cpm = counts per minute, LaWT = wild type La, RNP = ribonucleoprotein, K_D = dissociation constant

In order to test whether hLa binds the CCND1 5'-UTR an *in vitro* transcript of the cyclin D1 5'-UTR (D1-FL, 212 nts, -209 to +3) was synthesized (refer to 3.3.2). The CCND1 5'-UTR consists of 209 nucleotides, traditionally referred to as -209 to -1 in the 5' to 3' direction. In addition to the 5'-UTR the D1-FL RNA also contained the translational start site codon AUG, where the first and last nucleotides are referred to as +1 and +3 respectively. In order to transcribe the RNA *in vitro*, a DNA template had to be amplified by PCR. The PCR product was analyzed on a 2% agarose gel in order to check the quality and size of the PCR product. Figure 4.1.2.A shows that the PCR product has the correct size of 244 base pairs (212 bp CCND1 5'-UTR+ATG and 32 bp T7 promoter). More importantly, no secondary PCR product was amplified. The pure amplified DNA was used as a template for *in vitro* transcription and internal radiolabeling with $[^{32}P]$-CTP. The quantity of radioactive material was first detected using a scintillation counter, and then the RNA quality was subsequently analyzed by subjecting 2,500 cpm and 10,000 cpm of RNA to denaturing polyacrylamide gel electrophoresis. The gel was visualized by exposing it to a storage phosphor screen. Figure 4.1.2B shows the high quality of *in vitro* transcribed D1-FL RNA, no degradation product or other RNA species were detected.

Utilizing the recombinant La and the *in vitro* transcribed D1-FL RNA, affinity for the La and D1-FL RNA interaction was determined. Since La is an RNA-binding protein, the binding of the La protein to the D1-FL RNA is expected to obey the rules of saturation kinetics displaying a high affinity, whereas a low affine protein to RNA interaction only results in saturation at very high protein concentrations. The dissociation constant K_D for the La:D1-FL interaction is the binding constant and reflects the La concentration when half the protein binding sites are associated with the D1-FL RNA at equilibrium. A prominent method used to study RNA-protein interactions is electrophoretic mobility shift assays (EMSA), this method enables the visualization of the protein-nucleic acid interaction and allows for the determination of the dissociation constant. Hence, quantitative EMSAs were used to determine dissociation constant (K_D) of the interaction between La and D1-FL RNA.

Human recombinant La was titrated between 40 and 960 nM in presence of 10 nM radioactive labeled D1-FL RNA (3.3.6.1). The D1-FL RNA bound by La was quantitatively determined using the Storm phosphorimager and ImageQuant software (3.3.6.7). Figure 4.1.2C shows that a ribonucleoprotein (RNP) was formed by the recombinant La protein shifting the D1-FL RNA into a complex. The RNP formation occurs in a La concentration-dependent manner (figure 4.1.2C). Figure 4.1.2C further shows that at low La concentrations

merely one primary La-RNP (1st La-RNP) is formed, but the D1-FL RNA is shifted into an additional secondary La-RNP (2nd RNP) in the presence of higher La concentrations (160, 320, 640, and 960 nM). This additional secondary La-RNP is of lower mobility. The formation of both La-RNPs by hLa was considered as one La:D1-FL RNA complex for the EMSA quantification, which is displayed in figure 4.1.2D. A dissociation constant of $K_D \approx 45$ nM was calculated. The dissociation constant of 45 nM is in the lower nanomolar range, suggesting an affine binding of La to the cyclin D1 5'-UTR.

4.1.3 La binds the 5'-UTR of cyclin D1 at the 3'-terminus

After having established that human recombinant La binds the cyclin D1 5'-UTR (refer to 4.1.2), the next step was to map the La binding site within the CCND1 5'-UTR. Therefore, two *in vitro* transcripts (figure 4.1.3A) spanning the 5'-half (part A, 99 nts, nts -209 to -110) and the 3'-half (part B, 113 nts, nts -110 to +3) of D1-FL were synthesized and controlled as described in (3.3.1 and 3.3.2). All transcripts display high purity as displayed in figure 4.1.3C. In order to determine the binding of La to D1-FL, part A and part B RNA 320 nM, the hLa protein was incubated with 10 nM RNA and subjected to native EMSA (3.3.6.1). As shown earlier recombinant hLa protein forms a complex with the D1-FL RNA (La:D1-FL RNP, figure 4.1.3B). Importantly, hLa forms a complex with part B RNA and only a weak complex with part A as indicated by an asterisk in figure 4.1.3B. The faster mobility of the hLa:part B and partA RNPs is due to the smaller RNA molecules. Several slower mobility part A and part B RNA species are present in the absence of the La protein, most likely reflecting the presence of RNA molecules folded into different conformations.

These findings suggest that La binds within the 3'- terminal 113 nucleotides of the cyclin D1 5'-UTR. An additional contact site for hLa in the 5'-terminal half is possible since part A is shifted weakly into a complex with La.

To further map the binding site of La within the cyclin D1 5'-UTR utilization of synthesized antisense DNA oligoribonucleotides complimentary to the CCND1 5'-UTR to form a DNA:[32P]-D1-FL hybrid was used, this is described in 3.3.6.3. A cartoon of the DNA oligoribonucleotides and their complementary region in the cyclin D1 5'-UTR is displayed in figure 4.1.4A, the sequences of those oligonucleotides are listed in 2.12.2.

After hybridization of the radioactive labeled D1-FL RNA with the DNA oligoribonucleotides, RNA-protein binding reactions were carried out using 10 nM

DNA:RNA hybrid and 320 nM of recombinant La and analyzed by native PAGE figure 4.1.4B). The hybridization of DNA oligoribonucleotides to the RNA results in different RNA folding represented by the aberrant running behavior of free DNA:D1-FL RNA hybrids (figure 4.1.4B: 2 AS, 3 AS, 5 AS, 6 AS) in the native gel. Further, some hybrids, e.g. 3AS:D1-FL and 6AS:D1-FL, display more than one hybrid species due to alternative folding events.
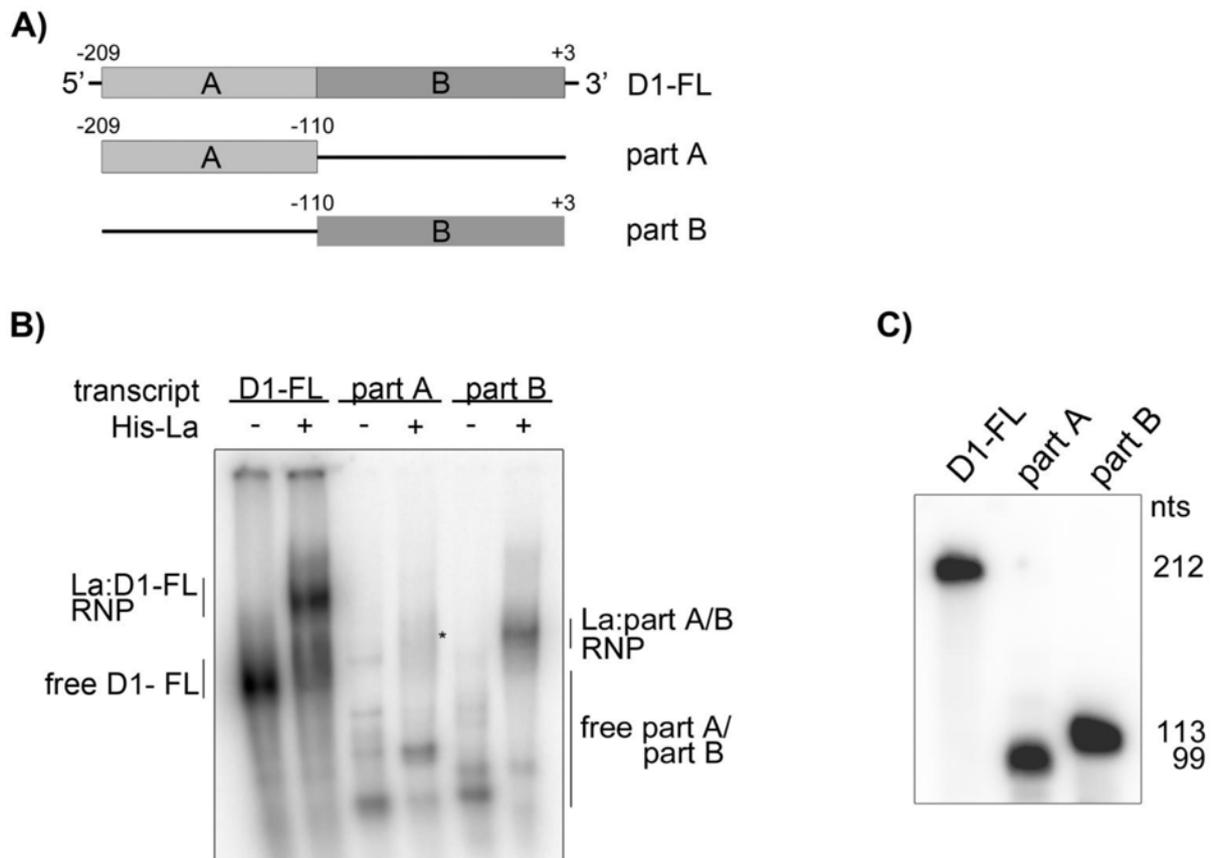
Figure 4.1.3: La binds the 3'-terminus of the cyclin D1 5' UTR. To map the binding site of La the protein's ability to bind terminal deletion mutants of the cyclin D1 RNA was studied by EMSA. **A)** Cartoon of the RNA deletion mutants created, part A consists of the 99 nucleotides of the 5' terminus of the cyclin D1 5'-UTR and lack the 3'- terminal. Part B lacks the 5'- terminal part, but contains the 110 nucleotides of the 3'- terminus. **B)** 320 nM of recombinant human La was incubated with 10 nM [^{32}P]-labeled cyclin D1 transcripts and separated by native EMSA. The free full length RNA and protein FL complex are indicated on the left. Free part A and part B RNA positions are indicated on the right, the part B and protein complex is indicated on the right as well. The hyphen represents reactions without protein. **C)** Denaturing PAGE of [^{32}P]-labeled *in vitro* transcribed RNAs. All transcripts were of the expected size and of high purity. FL = cyclin D1 5' UTR full length, A = part A, B = part B, nts = nucleotides

The complex formation of the hLa with DNA:D1-FL hybrid (figure 4.1.4B) was not as efficient as the EMSA described above and shown for the D1-FL RNA. Because of the different mobilities of the DNA:D1-FL hybrids it is difficult to identify La:DNA:D1-FL hybrid complexes. An asterisk in figure 4.1.4B indicates those potential complexes, which are formed with 2AS:D1-FL, 5AS:D1-FL, and 6AS:D1-FL RNA hybrids. As described before the annealing of the 3AS oligoribonucleotide to the D1-FL RNA resulted in an additional low mobility RNA species in the absence of hLa. Upon La addition the distinct band representing this slower mobility hybrid RNA species is more diffused and of weaker signal intensity, which does not allow for a conclusion regarding La-binding to this region (nts -103 to -53) in the D1-FL RNA. Hybridizing the oligoribonucleotides 4AS or 7AS to the radiolabeled D1-FL RNA resulted in one DNA:D1-FL species. In addition, those two oligoribonucleotides were able to disrupt the La:D1-FL complex formation, suggesting nucleotide -79 to +3 are critical for La binding to the cyclin D1 5'-UTR.
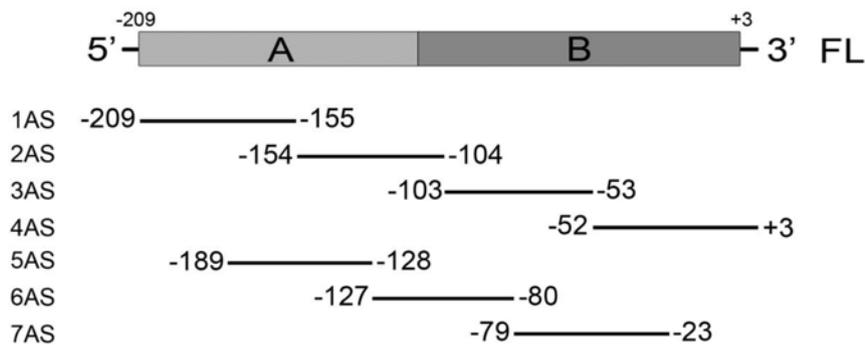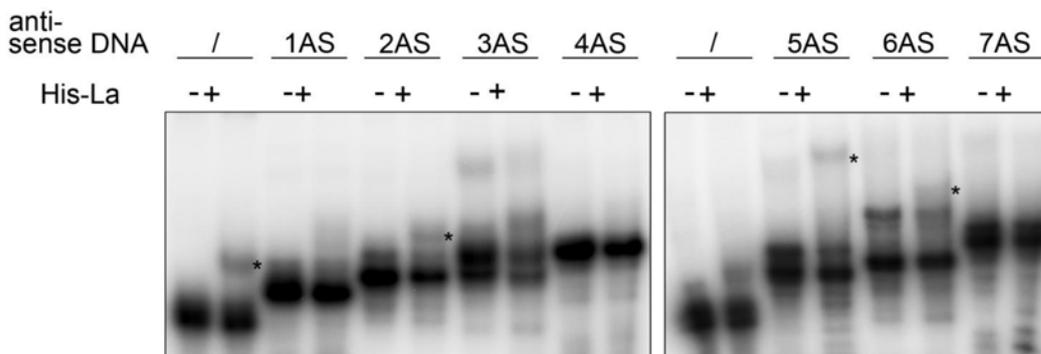
Figure 4.1.4: La probably binds between nts -79 to +3 of the CCND1 5'-UTR. A) Cartoon of the location of the DNA oligoribonucleotides complimentary regions to the cyclin D1 5'-UTR RNA. The DNA oligoribonucleotide names are indicated on the left and the nucleotide positions are left and right of the oligoribonucleotides. **B)** For binding studies 320 nM recombinant hLa protein was incubated with the RNA alone and with RNA:DNA hybrids and separated by native EMSAs. Asterisks indicate possible complexes of La and the RNA:DNA hybrid. Reactions without the protein served as negative control (indicated by a hyphen). FL = full length, AS = antisense

Competitive EMSAs were performed to further map the binding site of La in the cyclin D1 5'-UTR more precisely. The RNP formation of La and the D1-FL RNA can be challenged in the presence of competitors, thus an unlabeled (cold) competitor RNA containing a binding site for the La protein can out-compete the La:[^{32}P]-D1-FL complex formation when added in excess amounts. This out-competition is then represented by a reduction or loss of the EMSA shift. Unlabeled RNAs (D1-FL, part A and B) were synthesized (3.3.1 and 3.3.2) and used as competitors.
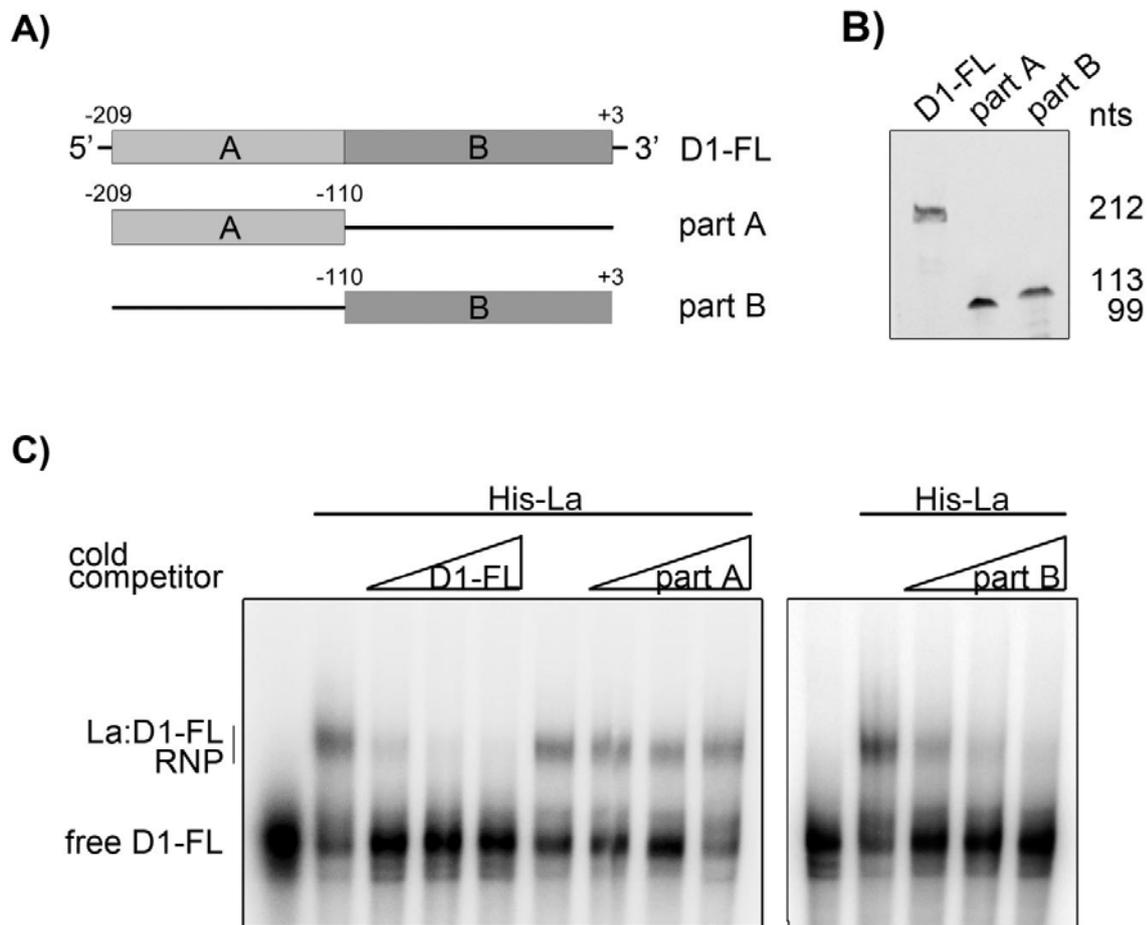
Figure 4.1.5: Recombinant hLa binds to the 3'-terminus of the cyclin D1 5'-UTR. Cold, unlabeled RNA competitors were used to map the binding site of La in the CCND1 5'-UTR. **A)** Cartoon of the cold competitors RNAs, D1-FL, part A, and part B with the location of their deletion indicated. **B)** Ethidium bromide stained denaturing PAGE gel of unlabeled RNA transcripts. All transcripts were of expected size and high purity. **C)** 160 nM recombinant human La were pre-incubated with increasing concentrations of cold transcripts (10-, 25-, and 50-fold excess to labeled probe) before 10 nM [^{32}P]-labeled cyclin D1 full length transcript was added and separated by native EMSA. The free full RNA and La:D1-FL RNP complexes are indicated on the left. Cold FL and part B probes are competing for La binding with the radioactive labeled FL transcript. Cold part A transcript competes weakly when the highest concentration is used.
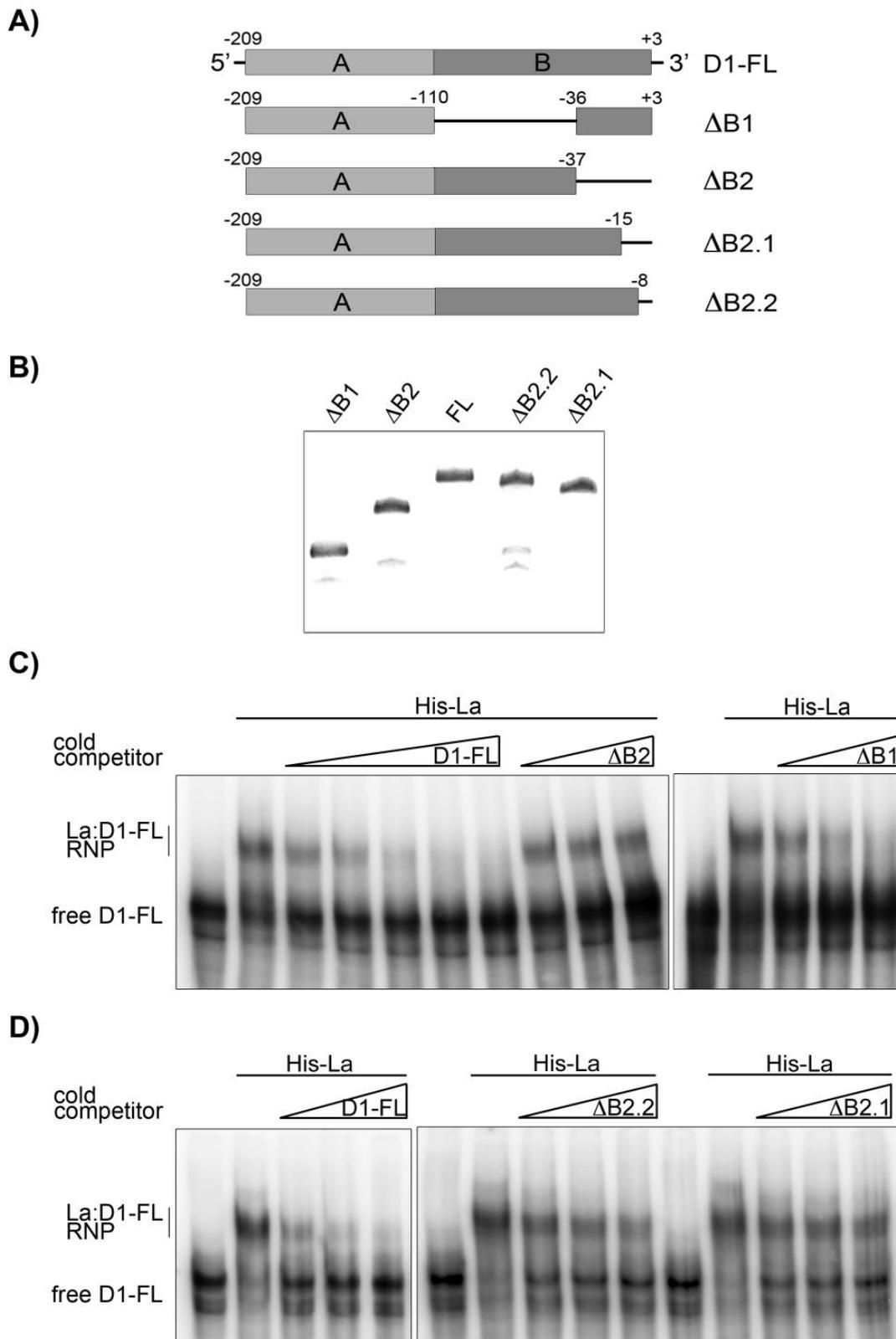
Figure 4.1.6: The La binding site is in close proximity to the translational start site. **A)** Cartoon of analyzed cold competitor RNAs D1-FL and part B deletion transcripts ΔB1, ΔB2, ΔB2.1, and ΔB2.1. Nucleotide deletions are indicated above each transcript. **B)** 400 ng of each unlabeled transcript was separated by denaturing PAGE and stained with ethidium bromide. All transcripts were of expected size and high purity. **C)** and **D)** Cold RNA probes in excess amounts of 10-, 25-, and 50-fold, and additional 2.5- and 5-fold for D1-FL **C)** pre-incubated with 160 nM recombinant La protein before binding to D1-FL.

The cold transcripts were allowed to interact with 160 nM recombinant human La for 10 min before 10 nM of $[^{32}P]$-labeled D1-FL was added (refer to 3.3.6.4 for details). According to the binding curve in figure 4.1.2D the La binding site was not fully saturated at a concentration of 160 nM La, thus this La concentration allows binding of more RNA as either competitor RNA or labeled D1-FL RNA. As positive binding control served the interaction of La with labeled D1-FL RNA in absence of competitors, the unlabeled cold D1-FL served as a positive competition control in the experiments. The La:$[^{32}P]$-D1-FL interaction was analyzed on native PAGE and displayed in figure 4.1.5C. La is shown to form a complex with the labeled D1-FL in the absence of competitor RNAs (figure 4.1.5C). As expected, the interaction of hLa with the labeled D1-FL probe is lost upon the addition of increasing amounts (100 nM, 250 nM, and 500 nM; 10-, 25-, and 50-fold) of cold D1-FL RNA (figure 4.1.5C). Similarly, adding increasing amounts of part B competitor RNA reduces the complex formation of La with the labeled RNA dramatically (figure 4.1.5C). However, the part A competitor RNA is only slightly reducing the complex La:D1-FL RNP formation. Note that only one La:D1-FL complex is formed and no other slower mobility complexes are detected. These EMSA findings and observations utilizing DNA:RNA hybrids as targets for the hLa protein (figure 4.1.4B) strongly suggest that the main La binding site is located in the 3'-terminal half of the cyclin D1 5'-UTR.

With this information as a starting point, cold competitors with deletions in part B were transcribed to again further map the La binding site more precisely by competitive EMSA. Figure 4.1.6A shows the deletions within part B utilized for additional mapping experiments. The primer sequences for the amplification of different DNA templates from the pRCD1F vector are listed in 2.12.1. The internal deletion in ΔB1 (nts -209 to -110 and -37 to +3) was achieved by a fusion PCR strategy as described in 3.3.1. As described above and in 3.3.2, the DNA templates were used for *in vitro* transcription and the RNA quality was assessed on a denaturing PAGE with subsequent ethidium bromide staining (figure 4.1.6B). The competitive EMSA was carried out as described above, as positive binding control served a La:D1-FL reaction without cold competitors. The radiolabeled D1-FL RNA was shifted into one complex with recombinant La in the absence of competitor RNAs and was successfully competed with increasing amounts of cold D1-FL RNA (25, 50, 100, 250, and 500 nM) or ΔB1 (100, 250, and 500 nM) (figure 4.1.6C). In contrast increasing concentrations (100 nM, 250 nM, and 500 nM) of competitor ΔB2, ΔB2.1, or ΔB2.2 were not able to efficiently compete for binding (figure 4.1.6C). These findings suggest that hLa binds to nts -8 and +3 of the cyclin D1 5'-UTR containing the start codon AUG. As mentioned, the translational start

site is part of the D1-FL RNA and was deleted in all terminal B-deletion mutants (ΔB2, ΔB2.1, and ΔB2.2), thus it cannot be excluded that the La protein directly binds the AUG translational start site.

4.1.4 Characterization of La binding to the cyclin D1 translational start site context embedded in a strong Kozak sequence

Mapping the binding site of La to nucleotides in close proximity of the start codon of cyclin D1 coincides with La binding to the HCV translational start codon [62]. Earlier, McBratney and Sarnow [111] have shown the interaction of La with the AUG start codon using synthetic oligoribonucleotides. Further, they have demonstrated the positive influence of a strong Kozak consensus sequence on oligoribonucleotide binding by the La protein.

To test whether La binds directly to the cyclin D1 translational start codon and to demonstrate whether the Kozak sequence is of importance for hLa binding a synthetic RNA oligoribonucleotide, D1-ATG (47 nts; nts -23 to +24), was synthesized. Figure 4.1.7A indicates this synthetic RNA, which consists of the 23 nucleotides upstream of the AUG, the start codon, and 21 additional nucleotides of the cyclin D1 open reading frame (ORF). Utilizing this tool, the binding of La to the synthetic D1-ATG RNA was analyzed by EMSA and the binding affinity was determined.

Quantitative EMSAs were carried out in order to study the interactions of human recombinant La with the synthesized $[^{32}P]$-D1-ATG RNA (3.3.6.1). The D1-ATG RNA oligoribonucleotide was synthesized by Integrated DNA Technologies, Inc. and 5'-end-labeled with $[^{32}P]$-ATP using the KinaseMax Kit as described in 3.3.3. Ni-NTA purified hLa protein was titrated between 1 nM and 30 μM, whereas the RNA concentration was kept unaltered at 10 nM. The La:D1-ATG complex formation was visualized and quantified using the Storm phosphorimager and ImageQuant software. Figure 4.1.7B shows that in the presence of recombinant La several La:D1-ATG RNPs are formed, similarly to RNP formation of hLa with the $[^{32}P]$-D1-FL RNA (figure 4.1.2C), which results in a primary and weak secondary La-RNP complex. As displayed in figure 4.1.7B at the lowest La concentration only one La:D1-ATG RNP complex (1st La-RNP) is formed. Increasing the La concentration results in secondary and tertiary La:D1-ATG complexes (2nd and 3rd La-RNP) leading to a total of three La:D1-ATG RNPs as shown in figure 4.1.7B. The 2nd and 3rd La-RNPs are preferentially formed in the presence of 600 nM or more of the La protein

compared to the 1st La-RNP. Further, those two complexes are formed at the expense of the primary complex, which is represented by a fading signal intensity of the 1st La-RNP. Note the additional higher mobility band, marked by a pound sign in figure 4.1.7B, which may represent by hLa refolded free RNA rather than a La-RNP. These multiple complex formation is discussed in detail in section 5.5.2. The dissociation constant for the formation of all three complexes is $K_D \approx 80$ nM, thus the affinity to the synthesized RNA is less than to the D1-FL RNA ($K_D \approx 45$ nM).

A)

-23   -3   +1   +4   +24

ATG: GCUGCCCAGGAAGAGCCCCAGCCAUGGAACACCAGCUCCUGUGCUGC

Figure 4.1.7: La binds to the translational start site context of cyclin D1 mRNA. A) Sequence of the D1-ATG, the AUG start codon is underlined and the most important positions -3 and +4 of the Kozak sequence are indicated. **B)** 10 nM of labeled D1-ATG RNA was incubated with increasing amounts of recombinant human La protein, 10, 60, 200, 400, 800, 1000, 1600, and 3000 nM and separated by a native EMSA. The free RNA and the La-RNP complexes are indicated on the left. # = free restructured D1-ATG RNA. **C)** Quantitative EMSA analysis. The La-RNP complexes are plotted against the molar La protein concentration. The dissociation constant was determined by non-linear regression as 80.3 nM.

In order to demonstrate that the complex formation in the EMSAs is mediated specifically by human recombinant La and not by contaminants from the protein purification, an antibody supershift assay was performed. This was done by utilizing 300 ng of recombinant La protein pre-incubated with 5 μl of the specific monoclonal La antibody SW5 or the isotype control mouse IgG2α,κ antibody as described in methods 3.3.6.4. Both La antibody mixtures were subsequently used for binding studies using D1-ATG RNA. Due to the higher mass of the La:SW5 complex a very low mobility La:SW5:D1-ATG complex was formed (figure 4.1.8). The La:control antibody mixture had no effect on the mobility of the La:D1-ATG RNP. Hence, this experiment demonstrates that the recombinant hLa protein binds the D1-ATG RNA and not any contaminating protein from the protein purification process. The signal

intensity and the mobility 1st La-RNP is slightly affected in the presence of either antibody, SW5 or IgG2α,κ, which is likely due to different buffer conditions.

Figure 4.1.8: Supershift assay identifies human La as D1-ATG RNA-binding partner. Monoclonal anti-La antibody SW5 and control antibody IgG2α,κ were incubated with recombinant Ni-NTA purified La protein before subjected to a standard EMSA with radiolabeled D1-ATG RNA as substrate. A control reaction without protein and one without antibody was included in the EMSA study (left two lanes). The free RNA, La-RNP complexes, and the supershift are indicated on the left. The D1-ATG RNA is shifted specifically into a complex with hLa.

In order to study the importance of the cyclin D1 AUG start codon and the authentic Kozak sequence for La binding mutations were introduced in the D1-ATG sequence. In the mu2 RNA mutant the AUG was changed to AGG in order to test the importance of the uridine [62] for La recognition, whereas in the mu3 RNA mutant the Kozak sequence was changed from the strong authentic D1 sequence (5'-GCC_AUG_GAA-3') to a weak Kozak sequence (5'-CAG_AUG_CAC-3') [214] [215] [216].

The RNA oligoribonucleotides, mu2 (34 nts, -17 to +18) and mu3 (34 nts; -17 to +18), were synthesized by Integrated DNA Technologies Inc., for sequence information refer to figure 4.1.9A and 2.11. These RNAs contain the original cyclin D1 sequence from the 5'-UTR as well as from the ORF. The changes in the sequence are indicated in bold and the start codon is underlined (figure 4.1.9A).

The effects of these mutations on La binding to D1-ATG were studied by competitive EMSAs using mu2 and mu3 as cold competitor RNAs and labeled D1-ATG (figure. 4.1.9B). The RNA oligoribonucleotide quality was assessed on a 10% denaturing PAGE with subsequent ethidium bromide staining (figure 4.1.9C), the RNAs were pure and of similar concentrations. The faster mobility of the mu2 and mu3 RNAs was caused by the smaller size of those oligoribonucleotides, which were 13 nucleotides smaller than the D1-ATG RNA. The unlabeled D1-ATG RNA oligoribonucleotide served as a positive competition control, a control reaction for La binding to D1-ATG in the absence of competitor was included as well. The D1-ATG RNA is shifted into three complexes with the La protein, the fastest mobility complex (1st La-RNP) is the major RNP complex, and the intensities of the other two

complexes, 2nd and 3rd La-RNP are weaker. As expected unlabeled D1-ATG RNA was a strong competitor for the La:[^{32}P]-D1-ATG complex formation (figure 4.1.9B). Increasing concentrations (100 nM, 250 nM, and 500 nM) of the start codon RNA mutant mu2 out-competes the binding of La to the [^{32}P]-D1-ATG very efficiently (figure 4.1.9B). This suggests that the AUG to AGG substitution had no effect on La binding. In contrast the mu3 competitor RNA was not able to compete for La: [^{32}P]-D1-ATG RNA-binding when added in excess amounts (100 nM, 250 nM, and 500 nM) suggesting that the cyclin D1 Kozak sequence surrounding the authentic CCND1 translational start site is critical for La binding. In summary, these studies identified for the first time the La binding site within the cyclin D1 5'-UTR. This binding site is located between nts -8 and + 3 and efficient binding depends mainly on a strong Kozak consensus sequence context.
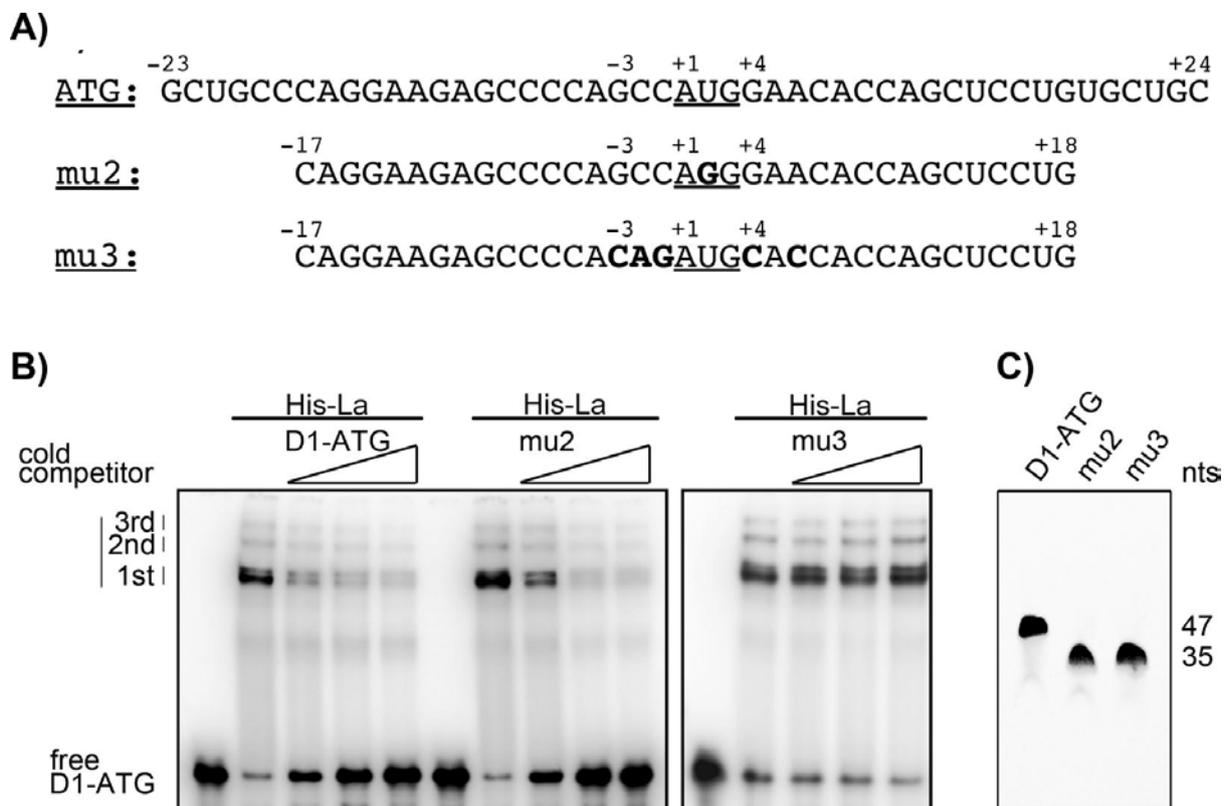
Figure 4.1.9: The binding of the hLa protein to D1-ATG RNA depends on a strong Kozak consensus sequence. Competitive EMSAs were performed to identify the role of the translational start site codon and its context. **A)** Sequences of the synthesized RNA oligoribonucleotides. The start codon is underlined; changes in the nucleotide sequence are in bold. The most important positions of the consensus Kozak consensus sequence are indicated. **B)** Competitive EMSA were performed using 10-, 50-, 100-fold excess amounts of unlabeled RNA and 10 nM [^{32}P]-labeled cyclin D1 ATG RNA. As negative binding and competition control served a reaction without hLa and competitor RNA, respectively. Cold D1-ATG RNA and mu2 RNA, but not mu3 RNA, are out-competing binding of La to radiolabeled D1-ATG RNA. **C)** Ethidium bromide staining of 400 ng cold D1-ATG, mu2, and mu3 RNAs separated by denaturing PAGE. The purity and concentrations of all RNAs are similar.

4.2 The RNA binding domains of La required for mRNA binding

In the nucleus, La binds the 3'-UTR of nascent RNAPIII transcripts resulting in stabilized, properly folded and mature transcripts (for references see introduction 1.1.). The binding of nascent RNAP III transcripts studied at the structural level revealed that hLa cooperatively binds RNA mediated by two N'-terminal domains, the La motif and the RNA recognition motif 1 (RRM1), which forms an RNAP III transcript-binding surface that adopts a wing-helix fold [217] [218] [31]. However, the interaction of the La protein with mRNAs in the cytoplasm is not fully understood. In general it is believed that the RRM1 and RRM2 is involved in binding to internal RNA elements [112] [35], mRNAs with 5'-terminal oligopyrimidine (TOP) elements [66] [59] and mRNAs with IRES elements in their 5'-UTR [91] [67] [89]. The following work elucidates the domains that are required for binding the cyclin D1 RNA.

4.2.1 Quantitative monitoring of La:RNA interaction by fluorescence polarization

The identification of the La domains required for binding the cyclin D1 5'-UTR can be analyzed quantitatively by several methods [219]. Alternatives to commonly used radioisotopes are probes labeled with commercially available fluorescent dyes. Those fluorescent probes can be used for EMSA, but also for other assays, such as fluorescence polarization (FP), and also allow for the determination of the dissociation constant. The FP assay is described in detail in 3.3.7. In short, a small molecule (e.g. RNA oligoribonucleotide) labeled with a fluorescent dye rotates fast in solution. Upon excitation with polarized light the fluorophore-labeled RNA molecule emits depolarized light when the molecule is rotating rapidly in solution. The depolarization is reduced and polarization increased when the labeled probe is bound to a protein and forms an RNA:protein complex with higher mass which rotates slower in solution.

Although the D1-ATG RNA is larger than recommended by Pagano [220] for an FP assay, it was 5'-end labeled with 6-carboxyfluorescin (6-FAM) by Integrated DNA Technologies, Inc. and used for fluorescence polarization assays to test whether this method is an alternative method to study La:D1-ATG RNA-binding. Equilibration reactions using the fixed amount of 100 nM 6-FAM-RNA and increasing amounts of La were carried out in quadruplicates in a 96-well plate. After 5 min incubation at room temperature, to allow for equilibration, the
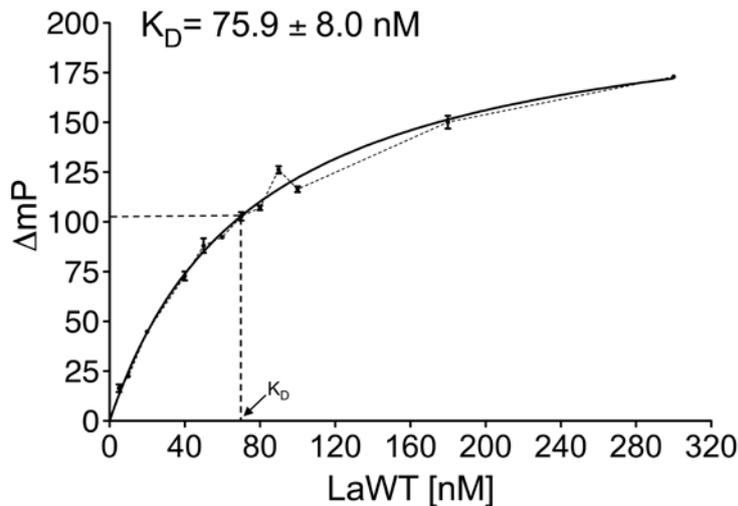
Figure 4.2.1: Binding affinity for the hLa:D1-ATG RNA interaction as determined by fluorescence polarization assay. Human La was used in binding studies with 6-FAM labeled D1-ATG RNA. The difference in polarization was plotted against the protein concentration. Using the one-site (hyperbola) non-linear regression fit the K_D was determined as 76 nM, as indicated by an arrow, in Prism 5 from quadruplicates of two independent experiments (n=2). mP = milli-polarization, K_D = dissociation constant

fluorescence plate reader POLARstar OPTIMA was used for detection of the emitted light, a control using the 6-FAM labeled D1-ATG RNA was used to determine the background polarization, and the reaction buffer itself was used as a blank. The effective polarization is expressed in millipolarization (mP), which is related to the fluorescence intensity in the parallel and perpendicular planes, the equation can be found in 3.3.7. The blank corrected data was analyzed by subtracting the mP values of the control reaction from the mP values of the La:D1-ATG RNA reactions. These corrected data were plotted as a function of the La protein concentration and the dissociation constant was calculated in Prism 5 using the one-site binding (hyperbola) non-linear fit. As a result, figure 4.2.1 shows the binding curve of two independent experiments each done in quadruplicates. The dissociation constant $K_D \approx 76$ nM $\pm$ 8 nM resembles the dissociation constant determined by EMSA ($K_D \approx 80$ nM; 4.1.4). Therefore, the fluorescence polarization assay was determined to be a suitable alternative for radioactive EMSAs. As a result, both assays were used in the following experiments to determine which La domains are required for D1-ATG RNA-binding.

4.2.2 Mapping the La binding domains for cyclin D1 RNA

The RNA binding protein La contains three RNA binding domains, the La motif, RNA recognition motif 1 (RRM1), and the RNA recognition motif 2 (RRM2). The La motif and RRM1, often referred to as La module, cooperatively bind RNAP III transcripts as described in 1.1. However, the RNP-2 consensus sequences in RRM1 and RRM2 have been demonstrated to be essential for HBV.B RNA-binding [35] [113] confirming the involvement of the non-canonical RRM2 in HCV IRES binding. In order to identify the La domains

required for D1-ATG RNA binding, mutations of/or within the RNA binding domains were required. Prior to this work, the cDNA of human recombinant La mutants with several mutations were already cloned into the prokaryotic expression vector pet28b(+), namely LaΔ1 (deletion of aa 11-99), LaΔ2 (deletion of aa 113-118), and LaΔ4 (deletion of aa 235-242) (refer to table 4.2.2) [35].
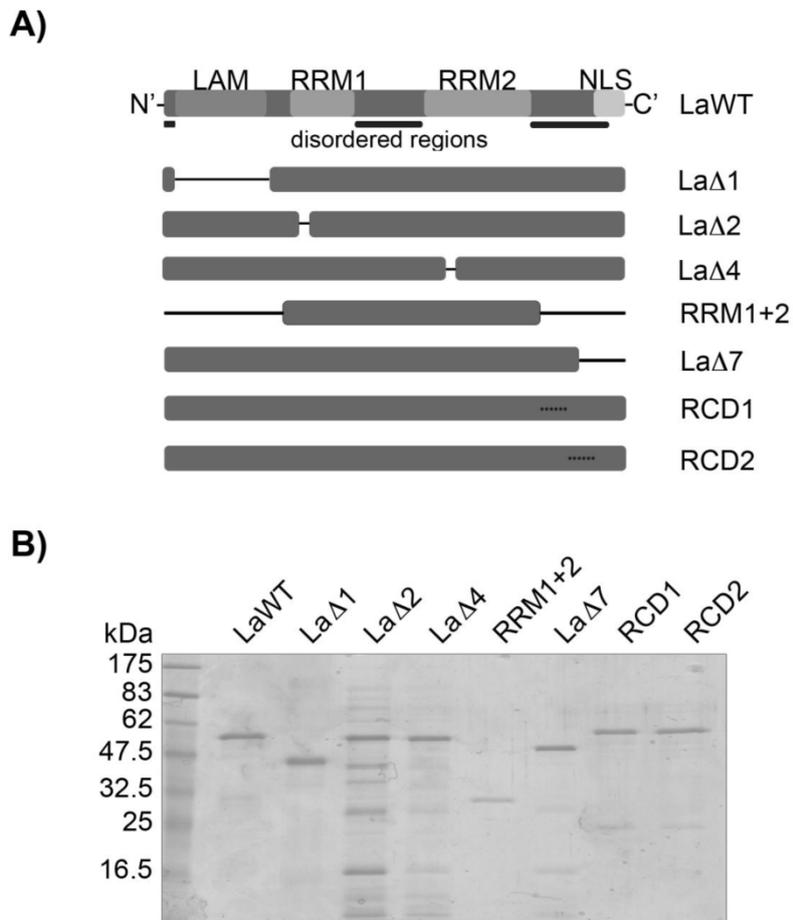
Figure 4.2.2: La protein mutants analyzed in RNA-binding studies. A) Cartoon of human La wild type and its respective mutants. The location of the amino acid deletion is indicated by the black line. The La motif was deleted in LaΔ1, the RNP-2 consensus sequence was deleted in RRM1 and RRM2 for LaΔ2 and LaΔ4, respectively. The NTD and CTD upstream and downstream of RRM1, the interdomain linker, and RRM2 are deleted in RRM1+2. The LaΔ7 has a CTD deletion, including the NLS. Basic and aromatic stretches in the disordered CTD were substituted by neutral amino acids (indicated by asterisks). **B)** Equal molarities of the recombinant proteins, LaWT (46.8 kDa), LaΔ1 (36.3 kDa), LaΔ2 (46.1 kDa), LaΔ4 (46.1 kDa), RRM1+2 (25 kDa), LaΔ7 (42.3 kDa), RCD1 (46.8 kDa), and RCD2 (46.8 kDa), were separated by SDS-PAGE and subsequently stained with Coomassie. The protein marker and the corresponding molecular weights in kDa are on the left. LAM = La motif, RRM = RNA recognition motif, kDa = kilodalton

Human recombinant His-tagged wild type La (hLaWT; 46.8 kDa), LaΔ2 (46.1 kDa), and LaΔ4 (46.1 kDa), were purified by D. Fedarovich from the protein production lab at the Medical University of South Carolina, USA. The His-tagged LaΔ1 (36.3 kDa) was expressed and purified by previous lab members. The integrity of the recombinant proteins was analyzed by separation of equal molarities by SDS-PAGE with subsequent Coomassie staining (figure 4.2.2). The wild type hLa and LaΔ4 did not display major degradation products or contaminations; however, purified LaΔ2 displayed a fragmentation pattern as shown in (figure 4.2.2B). Horke *et al*. [35] identified the 16.5 kDa, ~ 25 kDa, and ~ 40 kDa fragments in the LaΔ2 preparation as N'-terminal fragments by immunoblotting for La with the mouse monoclonal specific La 3B9 antibody. The LaΔ4 mutant shows a similar fragmentation pattern, which is by far not as extensive compared to LaΔ2.

Initially, the importance of the La motif for cyclin D1 RNA binding was studied by using fixed amounts of [^{32}P]-D1-ATG RNA for EMSA and 6-FAM D1-ATG RNA for the FP assay and increasing amounts (40, 80, 160, and 320 nM) of recombinant LaΔ1. Evidently, shown in figure 4.2.3B the loss of the La motif does not affect RNA binding activity for the oligoribonucleotide in native EMSA. As noted, hLaWT and LaΔ1 have the ability to shift the D1-ATG RNA into La-RNP complexes as shown in figure 4.2.3B. The hLaWT protein forms three complexes with the D1-ATG RNA, however, the fastest mobility RNP (# in figure 4.2.3B) may represent an RNA molecule with different structure and mobility. As figure 4.2.3B shows three La-RNPs (1st, 2nd, and 3rd La-RNP) are formed in a protein concentration dependent manner. Binding of the LaΔ1 protein mutant with D1-ATG RNA results in the formation of two La-RNP complexes (1st La-RNP, 2nd La-RNP; figure 4.2.3B), the signal intensity of the 1st LaΔ1:D1-ATG RNP is decreasing with increasing La concentrations, whereas the 2nd LaΔ1:D1-ATG RNP is saturating. The running behavior of the mutant LaΔ1:D1-ATG complexes differ from the hLaWT:D1-ATG RNP by showing a higher mobility. In a native PAGE the protein mobility is based on primarily the protein's electric charge rather than its molecular mass. The hLaWT protein has a net charge of -2.7 at pH 8, whereas the net charge of the LaΔ1 mutant is slightly higher at -2.4 at the same pH potentially explaining the different mobilities of those proteins. The dissociation constant for the LaΔ1:D1-ATG binding reaction was determined by fluorescence polarization as $K_D \approx 68$ nM. Hence, the LaΔ1 was shown to be very similar to the wild type hLa dissociation constant of $K_D \approx 76$ nM $\pm$ 8 (figure 4.2.1C). These data suggest that the La motif does not contribute to binding of La to D1-ATG RNA.
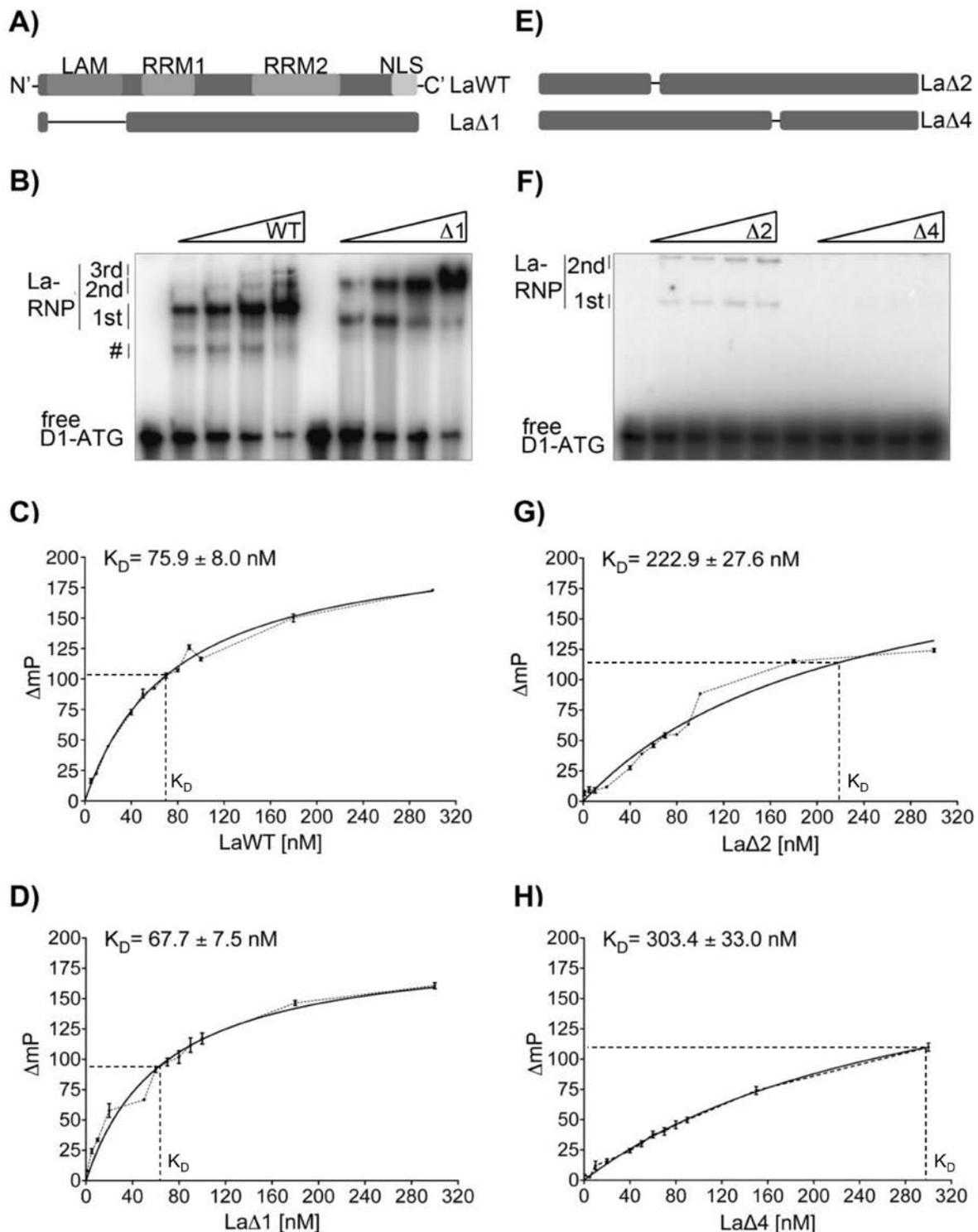
Figure 4.2.3: The RNP-2 consensus sequences in RRM1 and RRM2 are required for cyclin D1 RNA-binding. The hLa RNA-binding domains were identified by EMSA and fluorescence polarization assays **A)** and **E)** Cartoon of LaWT and RNA-binding mutants LaΔ1, LaΔ2, and LaΔ4. The deletions are indicated by a black line. **B)** and **F)** EMSAs were carried out with 40, 80, 160, and 320 nM recombinant La proteins as described in 3.3.6.1. Free D1-ATG RNA and La-RNP complexes are indicated on the left. # = free restructured RNA **C), D), G), H)** Binding curves of LaWT and RNA-binding domain mutants (indicated) as determined by FP assay using 6-FAM labeled D1-ATG RNA. Blank and background corrected data are plotted as a function of La protein concentration. The dissociation constants were calculated using the one-site (hyperbola) non-linear regression fit in the Prism 5 software from quadruplicates.

Additionally, the requirement of the RNA recognition motifs 1 and 2 were tested for D1-ATG RNA binding by EMSA and FP. The RNA recognition motifs contain consensus sequence stretches, called ribonucleoprotein domain-1, RNP-1, and ribonucleoprotein domain-2, RNP-2, which are essential for mediating RNA-binding [15]. These RNP sequences have been shown to be required for binding HBV.B RNA [35]. Therefore, in order to study the role of the RNP-2 sequence in RNA-binding recombinant La mutants with RNP-2 sequence deletions of either RNA recognition motif 1 or 2 as described by Horke *et al.* [35], were analyzed for their D1-ATG RNA-binding activity. As mentioned earlier, the RNP-2 in RRM1 was deleted in LaΔ2 (deletion aa113-118), whereas the RNP-2 of RRM2 is deleted in LaΔ4 (deletion aa235-242) (figure 4.2.3E and table 4.2.1)

The synthetic D1-ATG RNA was 5'-end labeled with [^{32}P]-ATP for EMSA studies and with 6-FAM for fluorescence polarization assays. For native EMSAs increasing concentrations of LaΔ2 and LaΔ4 proteins (40, 80, 160, and 320 nM) were allowed to form complexes with 10 nM radiolabeled D1-ATG RNA. The D1-ATG RNA was shown not to be shifted into a complex with the LaΔ4 mutant as shown in the four lanes on the right in figure 4.2.3F. Thus, the deletion of the RNP-2 sequence (5'-KFSGDL-3') in RRM2 in LaΔ4 is essential for binding according to the gel retardation assay. The polarization assay was carried out as described in (3.3.7), and the RNA-binding affinity $K_D \approx 303$ nM $\pm$ 33 nM was determined by FP assay which is almost four fold higher compared to the wild type protein $K_D \approx 76$ nM $\pm$ 8 nM (compare figure 4.2.1 with 4.2.3H). In contrast to the EMSA the FP result suggests that the binding of hLa to the D1-ATG oligoribonucleotide is not completely lost, however, the deletion of the RNP-2 in RRM2 is changing the binding kinetics of the LaΔ4 mutant dramatically so that no La-RNP complex formation is observed in EMSAs.

The LaΔ2 mutant forms two La-RNP complexes, 1st La-RNP and 2nd La-RNP, in the presence of D1-ATG RNA shown by the four lanes on the left in figure 4.2.3F. Those complexes are of very weak intensity and only the 1st La-RNP is formed in a hLa-concentration dependent manner. Hence, the deletion of the RNP-2 sequence (5'-VYIKGF-3') of the RRM1 in the LaΔ2 results in a near total loss of RNA-binding activity. The affinity of LaΔ2 to the D1-ATG RNA was determined by fluorescence polarization as $K_D \approx 223$ nM $\pm$ 28 nM, hence showing a three-fold increase in the dissociation constant compared to the wild type hLa protein $K_D \approx 76$ nM (compare figure 4.2.3G with figure 4.2.1). Interestingly, the slightly better RNA-binding activity of LaΔ2 in the EMSA is also reflected in the lower dissociation constant of LaΔ2 $K_D \approx 223$ nM compared to LaΔ4 $K_D \approx 303$ nM (compare 4.2.4C with 4.2.4D).

Since RRM1 and RRM2 are required for cyclin D1 RNA oligoribonucleotide binding, a minimal La protein RRM1+2 (aa 114-326, 25 kDa) containing RRM1, the inter-RRM peptide linker region, and the RRM2 were analyzed for their ability to bind D1-ATG RNA in EMSA and FP assays and also to determine whether other domains in the La protein are contributing to RNA-binding.

name	mutation type	location of mutation
LaΔ1	internal deletion	La motif (aa 11-99)
LaΔ2	internal deletion	RNP-2 RRM1 (aa 113-118)
LaΔ4	internal deletion	RNP-2 RRM2 (aa 235-242)
LaΔ7	internal deletion	NLS (aa 353-393)
RCD1	substitution	K332G, R334G, R335G, F336V, K337G, K339G
RCD2	substitution	F357V, K360G, K361G, K363G, F364V
RRM1+2	terminal deletions	La motif and C'-terminus (aa 1-112, 327-408)

Table 4.2.1: Overview of recombinant human La protein mutants. The mutation types as well as their positions are listed for the La mutants used in electrophoretic mobility shift or fluorescence polarization assays. aa = amino acid, RNP = ribonucleoprotein consensus sequence, RRM = RNA recognition motif, NLS = nuclear localization signal, K = lysine residue, R = arginine residue, F = phenylalanine residue

The minimal La protein RRM1+2 had to be cloned, the primer sequences needed are listed in section 2.12.3 and the cloning strategy is discussed in detail in sections 3.1.1 and 3.1.2. The protein was expressed (refer to 3.2.1) and purified using the Ni-NTA technology as described in section 3.2.2. The integrity of the recombinant RRM1+2 protein (25 kDa) was analyzed by SDS-PAGE with subsequent Coomassie staining (figure 4.2.2) revealing high protein purity. Hence, increasing RRM1+2 concentrations of 40, 80, 160, and 320 nM were titrated against 10 nM [^{32}P]-labeled D1-ATG RNA for the EMSA separated by native PAGE (3.3.6.1). As shown in figure 4.2.4B the minimal La protein RRM1+2 was able to shift the D1-ATG RNA into three La-RNPs, (1st, 2nd, 3rd La-RNP). The 1st La-RNP is not formed in a protein concentration-dependent manner, since the intensity was unaltered upon increasing RRM1+2 concentrations. The secondary La-RNP, 2nd La-RNP, is a very weakly formed La:D1-ATG complex, however, the tertiary RNA complex, 3rd La-RNP, is only formed in the presence of the highest RRM1+2 protein concentrations (160 nM and 320 nM, figure 4.2.4B). The signal intensity of the tertiary La-RNP complex is similar to the primary complex in the presence of 320 nM protein. The RNA-binding affinity for the RRM1+2 mutant to the D1-ATG RNA was determined by fluorescence polarization assay as $K_D \approx 111$ nM $\pm$ 6 nM (figure 4.2.4C), note that value compared to the wild type $K_D \approx 76$ nM (figure 4.2.1). As a result, RRM1 and RRM2 are the most important RNA-binding motifs for binding of hLa to D1-ATG RNA. The N'-terminal domain (NTD) as well as the C'-terminal domain (CTD) are not essential for D1-

ATG RNA-binding activity but are clearly contributing to binding as indicated by an increase in the $K_D \approx 111$ nM of RRM1+2.
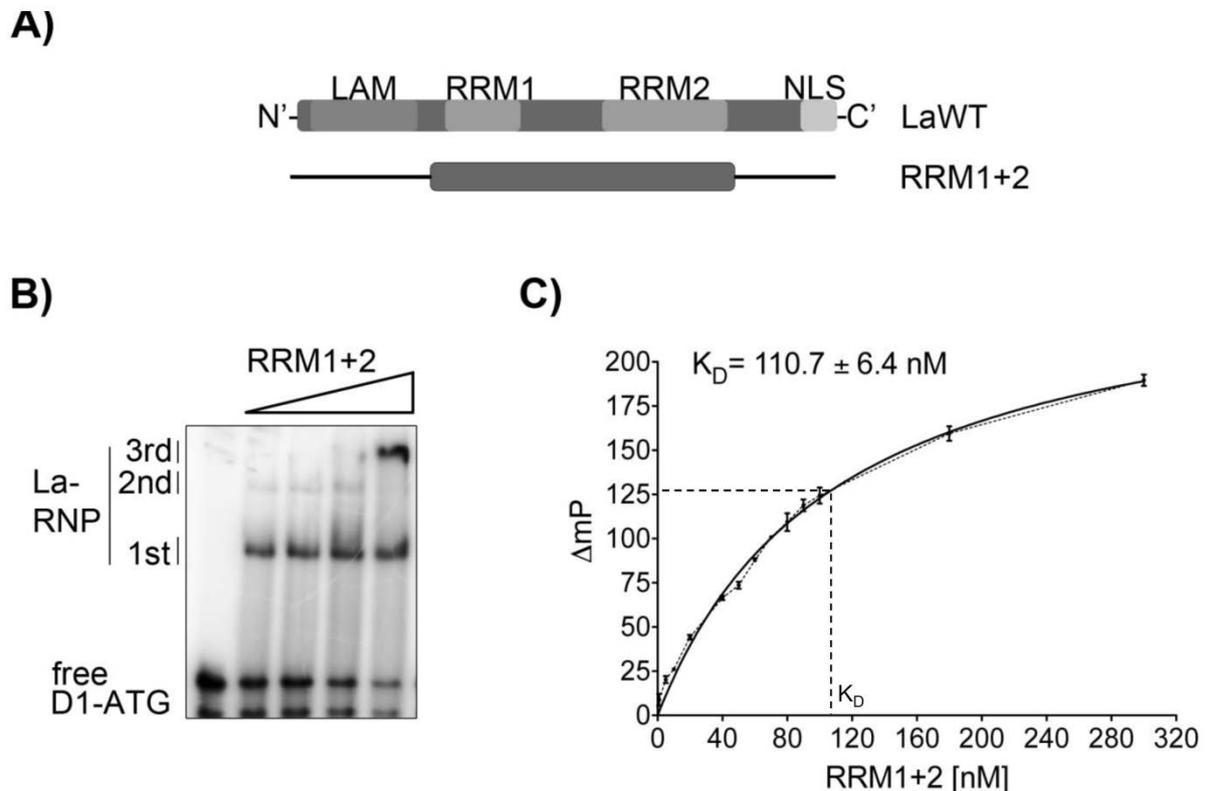
Figure 4.2.4: The hLa N'-terminal and C'-terminal domains are not required for D1-ATG RNA-binding. The RNA-binding activity of the recombinant minimal La protein RRM1+2 was analyzed by EMSA and fluorescence polarization. **A)** Cartoon of La wild type and the minimal RRM1+2 protein lacking the N'-terminus as well as the C'-terminus. The terminal deletions are represented by a black line. **B)** EMSA with 40, 80, 160, 320 nM of recombinant RRM1+2 protein binding [^{32}P]-labeled D1-ATG RNA. Free RNA and La-RNP complexes are indicated on the left. **C)** Binding curve of RRM1+2 with 6-FAM labeled D1-ATG as determined by FP assay. The dissociation constant was determined by one-site (hyperbola) nonlinear regression fit using the Prism 5 software from quadruplicates.

The oligomerization of the La protein is heavily discussed in the literature (refer to the introduction 1.1). The proposed multimerization domain is located between either amino acid 274-291 referenced by Craig [34] or 298-348 referenced by Horke [35]. It is important to note that the dimerization of the La protein is a functional requirement to stimulate translation [34]. Note that secondary and tertiary La-RNP complexes have been observed in the presence of high RRM1+2 concentrations (figures 4.2.4) and that this minimal La protein, RRM1+2, constitutes amino acids 113-326. In order to test for oligomerization located between amino acids 293-348 as proposed by Horke *et al.* [35], the minimal RNA-binding competent La protein was added in high concentrations (99 nM, 198 nM, 298 nM, and 596 nM) to radiolabeled D1-FL RNA. The formation of multimeric La-RNP complexes are suggested to

result from unspecific binding events [65] [64] [63], thus EMSAs were performed in the presence or absence of 0.5 μg of the highly negatively charged polyanion heparin. Since heparin mimics the RNA phosphate backbone it is often used to reduce non-specific and background protein binding to RNA [221] [222] [223]. In the presence of excessive heparin amounts, unspecific RNA-binding to the D1-FL RNA is expected to be abrogated and the formations of specific monomeric La-RNPs are also expected.

The RNA-binding activity of the RRM1+2 at high concentrations was analyzed by native EMSA in the absence of heparin. The minimal La protein RRM1+2 was able to form more than one RRM1+2:D1-FL RNP complex in a protein concentration-dependent manner as shown in figure 4.2.5B. Note that only one complex of high mobility (monomeric La:D1-FL) was formed at the lowest RRM1+2 protein concentration (99 nM), however, upon increasing RRM1+2 concentrations (198 , 298, and 596 nM) the D1-FL RNA was shifted into a secondary and tertiary La-RNP also noted in figure 4.2.5B as multimeric La:RNP, this was done at the expense of the primary complex. Note that the lowest RRM1+2 protein concentration of 99 nM shifts all of the D1-FL RNA into a La-RNP, as a result there is no free RNA left in the samples.

On the other hand, in the presence of 0.5 μg of the unspecific RNA competitor heparin, RRM1+2 forms merely one complex with the D1-FL RNA in a protein concentration-independent manner (figure 4.2.5B, four lanes on the right). This complex has the same mobility as the primary RNP also noted in figure 4.2.5B as monomeric La:RNP in the absence of heparin.

Taking these findings into consideration it can be concluded that the recombinant RRM1+2 protein forms multimeric RRM1+2:D1-FL RNPs in the absence of an unspecific competitor, and an unspecific multimeric La:D1-FL RNP formation is abrogated in the presence of excessive amounts of the unspecific competitor heparin. The question remains whether the multimerization of La in absence of heparin is caused by an unspecific aggregation or by a specific dimerization domain driven association. Further, it is unclear whether multimers are formed by binding of up to three La molecules to the same RNA or by hLa:hLa interactions. In conclusion, *in vitro* analyses for mapping the dimerization domain of hLa are unconsolidated.
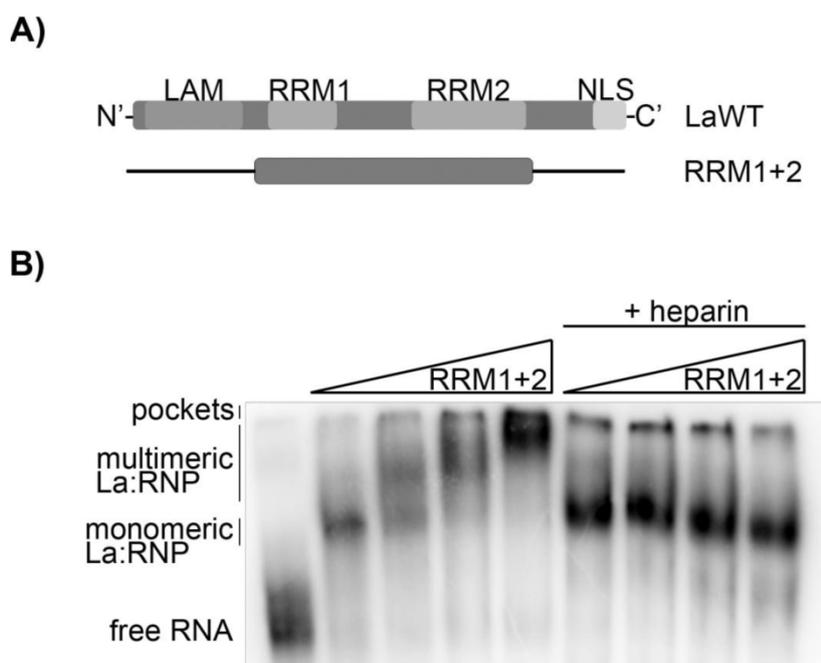
Figure 4.2.5: La-RNP complex formation in the absence and presence of heparin. A) Cartoon of RRM1+2 mutant La protein in comparison to human LaWT. **B)** A standard EMSA performed with 99, 198, 298, and 596 nM RRM1+2 and D1-FL RNA in the presence or absence of 0.5 μg heparin as indicated. In the absence of heparin multimeric La:RNP complexes are formed. Contrary, one La:RNA complex is formed in the presence of heparin.

Since the RNA-binding affinity of the minimal RRM1+2 protein, which lacks the NTD and CTD, is lower compared to the wild type hLa it can be inferred that the NTD and/or CTD may contribute to D1-ATG RNA-binding. It is known that the C'-terminal amino acids 353-393 are contributing to HBV.B RNA-binding [35]. Contrary, the basic residues 329-363 have been reported to be functionally important for transcription factor activity [40] and the basic amino acid residues between amino acids 328-344 of hLa contribute to HCV IRES-binding [99]. Furthermore, the CTD of hLa is intrinsically disordered [224] and RNA-chaperone activities are proposed to be located in those kinds of unstructured regions [117] [126]. Taking all this into consideration, mutants with C'-terminal deletions or amino acid substitutions were analyzed by EMSA and FP assays to study whether they contribute the D1-ATG RNA-binding.

His-tagged recombinant C'-terminal deletion mutant LaΔ7 (deletion of aa 353-393, 42.3 kDa), described by Horke *et al.* [35], and basic and aromatic substitution mutants, RCD1 (K332G, R334G, R335G, F336V, K337G, K339G; 46.8 kDa) and RCD2 (F357V, K360G, K361G, K363G, F364V; 46.8 kDa) were expressed and purified as described in sections 3.2.1 and 3.2.2. Equal molarities of the recombinant mutant proteins were subjected to SDS-PAGE with subsequent Coomassie staining to confirm the correct size and analyze the integrity of

the recombinant proteins (figure 4.2.2). In addition, native EMSA analyses were carried out with 10 nM D1-ATG RNA and increasing concentrations (40, 80, 160, and 320 nM) of recombinant LaΔ7 protein (figure 4.2.6B). The D1-ATG RNA is shifted into two RNPs, 1st RNP and 2nd RNP, with the LaΔ7. The 2nd RNP complex is LaΔ7-concentration dependent; however, the 1st RNP is slightly decreasing. The RNA-binding affinity was determined by FP assay as described in 3.3.7 resulting in a $K_D \approx 51$ nM ± 3 nM (figure 4.2.6C). This dissociation constant is slightly lower than the wild type La:D1-ATG RNA-affinity of $K_D \approx$ 76 nM (figure 4.2.1). Interestingly, the LaΔ7 mutant behaves similar to the deletion of the La motif mutant LaΔ1, the band pattern is similar as well as the slightly lower dissociation constants (compare 4.2.6B and C with 4.2.3B and C). In summary, the deletion of aa 353 to 393 in the C'-terminus of La Δ7 does not dramatically affect D1-ATG RNA-binding in EMSA and fluorescence polarization as shown in figure 4.2.6B and C. However the higher RNA-binding activity of the mutant hLa protein suggests that aa 353-393 negativly affect LaWT binding to D1-ATG RNA.
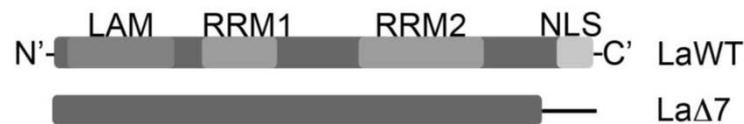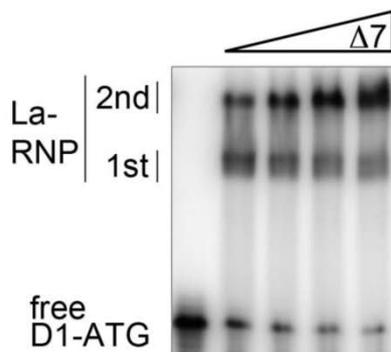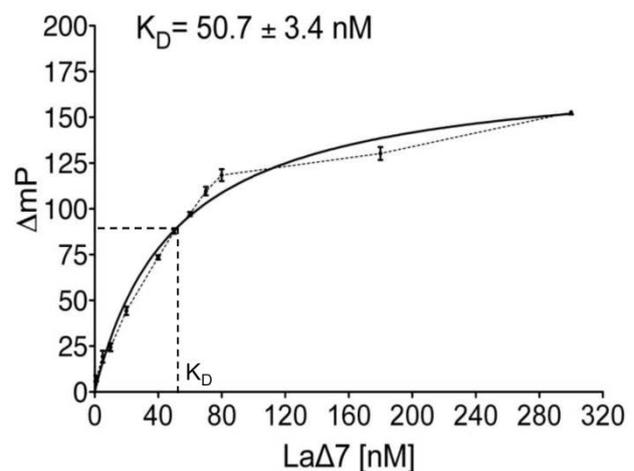
Figure 4.2.6: The extreme C'-terminus of La is not required for cyclin D1 RNA-binding. EMSA and FP are used to determine the RNA-binding activity of LaΔ7. **A)** Cartoon of human LaWT and C'-terminal deletion mutant LaΔ7, which lacks 40 amino acids including the nuclear localization signal in the C'-terminus of the protein. The amino acid deletions are represented by a black line. **B)** Standard EMSAs were performed as described before with recombinant LaΔ7 protein and [32P]-labeled D1-ATG RNA. **C)** Binding curve of FP assay using recombinant LaΔ7 and of 6-FAM labeled D1-ATG RNA.

The D1-ATG RNA-binding activity of the RCD1 and RCD2 La mutants were next analyzed by EMSA and FP. In the RCD1 protein five basic (lysines 332, 337, and 339, and arginines 334 and 335) and one aromatic residue (phenylalanine 336) were substituted by the neutral amino acid glycine and valine, respectively (table 4.2.1). Substitutions in the RCD2 protein comprise three basic (lysines 360, 361, and 363) and two aromatic amino acids (phenylalanine residues 357 and 364) (table 4.2.1). In order to test whether these basic and aromatic amino acids are required or contribute to the binding of the D1-ATG RNA native EMSAs and FP assays were performed.
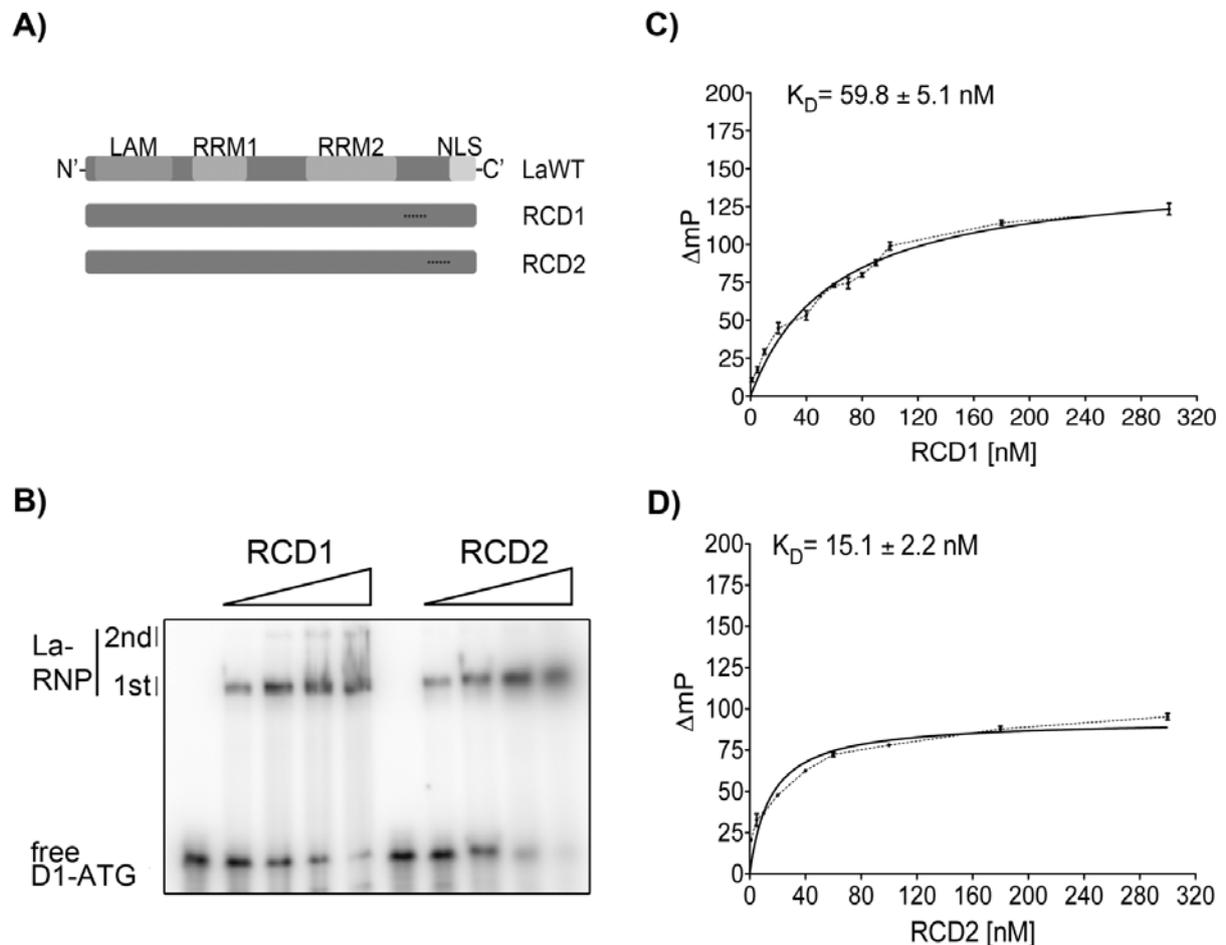
Figure 4.2.7: D1-ATG RNA-binding is not mediated by C'-terminal basic and aromatic amino acid stretches. The RNA-binding activity of the recombinant La mutants RCD1 and RCD2 was determined by native EMSA and FP. **A)** Cartoon of the recombinant La mutants in comparison to human LaWT. A stretch of basic and aromatic residues were substituted with glycine and valine, respectively (the mutations are indicated by asterisks in the cartoon). **B)** Standard EMSAs were performed with purified RCD1 and RCD2 and D1-ATG RNA. **C)** and **D)** Binding curve obtained from 6-FAM D1-ATG RNA and RCD1 (C) or RCD2 (D) by FP assay.

In the gel retardation assays, recombinant RCD1 or RCD2 were titrated (40, 80, 160, and 320 nM) in the presence of 10 nM labeled D1-ATG RNA and separated on a native PAGE. The hLa-RCD1 and RCD2 formed mainly a single complex with D1-ATG (figure 4.3.7B). A minor secondary La-RCD1:D1-ATG RNA La-RNP (figure 4.3.7B, 2nd La-RNP), is formed which was not detectable with RCD2. Interestingly, the RNA-binding affinities for RCD1 and RCD2 differed dramatically, the K_D for RCD1:D1-ATG RNA interaction was $K_D \approx 60$ nM $\pm$ 5 nM, thus similar to the dissociation constant of the wild type hLa, LaΔ1, and LaΔ7, where the RCD2:D1-ATG RNA interaction was dramatically lower at $K_D \approx 15$ nM $\pm$ 2 nM. In conclusion, the basic and aromatic aa 332-339 substituted in RCD1 do not contribute to the D1-ATG RNA-binding affinity. However, the substitutions of basic and aromatic residues between amino acids 357-364 in RCD2 suggest that those residues negatively influence hLa binding to D1-ATG RNA.

La protein	EMSA binding	K_D [nM]
hLaWT	+++	76 ± 8
LaΔ1	+++	68 ± 7
LaΔ2	(-)	223 ± 28
LaΔ4	-	303 ± 33
LaΔ7	+++	51 ± 3
RCD1	+++	60 ± 5
RCD2	++++	15 ± 2
RRM1+2	++	111 ± 6

Table 4.2.2: Summary of D1-ATG RNA-binding studies using La protein mutants. The results of the RNA-binding studies by electrophoretic mobility shift assay and fluorescence polarization (FP) show a high degree of similarity. The affinities determined by FP of LaΔ1, LaΔ7, RCD1, and RRM1+2 are similar to the K_D of the wild type La protein. This is also represented by the EMSA studies. The RNA-binding activity of the RCD2 mutant is higher compared to the other binding mutants, which is reflected in a low K_D. EMSA = electrophoretic mobility shift assay, K_D = dissociation constant

In summary, the RNP-2 sequences in both RNA recognition motifs, RRM1 and RRM2, were identified as important requirements for D1-ATG RNA-binding. However, the minimal RRM1+2 protein, lacking the NTD and CTD, displayed a lower RNA-binding affinity compared to the dissociation constant determined for hLaWT suggesting additional domains contributing to La binding D1-ATG RNA. The NTD (including the La motif), as well as amino acids 353-393, and the basic and aromatic residues located between aa 332 to 339 of the intrinsically disordered CTD are not essential for D1-ATG RNA-binding. However, the

basic and aromatic residues spanning the C'-terminal aa 357-364, are negatively contributing to D1-ATG RNA-binding by hLa.

The fluorescence polarization experiments represents a time saving RNA-binding assay that accurately displays RNA:protein interactions, however, the formation of monomeric or multimeric complexes cannot be monitored by FP assay. Fluorescence polarization is an attractive method to rapidly screen proteins for RNA-binding affinities.

4.3 SUMO-modification of human recombinant La protein

Sumoylation is the posttranslational modification of a target protein by covalent and reversible conjugation of a peptide called small ubiquitin-like modifier (SUMO), this modification has been found to play an essential regulatory role in many biological processes by maintaining protein stability, transcriptional regulation, and modifying specific transcription factors [225] [192]. A small number of RNA-binding proteins, e.g. hnRNP C1, eIF4E [211], PSF [226], and Sam68, have been identified to be modified by SUMO [201] [202]. It has been shown that murine La is sumoylated in cultured dorsal root ganglions, that GFP-tagged human La is modified by SUMO when transfected into rat sensory axons, and that this modification is required for retrograde transport of hLa [60]. However, at this point it is not clear what the role and function of sumoylation of murine and human La may be biologically. The most studied function of La is the interaction with a broad variety of RNA molecules. Therefore it is reasonable to speculate that sumoylation may modulate the RNA-binding activity of hLa. To test this idea a robust *in vitro* sumoylation assay (IVSA) for hLa was established and the RNA binding activity of SUMO-La was tested in EMSAs using D1-FL RNA and RNA oligoribonucleotides, representing substrates with terminal sequences of 5'-UTR corresponding to different TOP mRNA.

4.3.1 Purification of SUMO-cascade components

Conjugation of SUMO to its target protein is a multi-step process involving maturation of the SUMO precursor; activation, conjugation, ligation, and de-sumoylation as described in detail in the introduction (refer to 1.3). In short, all SUMO isoforms are expressed as precursors, which undergo maturation mediated by SENP (sentrin specific proteases). The mature SUMO

has to be activated by a heterodimer of the SUMO activating enzymes, SAE1 and SAE2, which requires ATP- hydrolysis [158]. The activated SUMO peptide is transferred to the E2 conjugating enzyme Ubc9, which then transfers SUMO to its target protein. An efficient conjugation requires the co-catalyzation by a specific E3 ligase *in vivo* but this is not required *in vitro* [227]. SENPs not only mediate the maturation of SUMO precursors but also are responsible for de-sumoylation of SUMO-modified proteins. The *in vitro* SUMO-modification using recombinant proteins does not require SENPs for maturation of SUMO nor de-sumoylation, therefore, the only essential proteins for the IVSA are SUMO, SUMO-activating enzymes SAE1 and SAE2, Ubc9, and hLa all of which needed to be expressed as recombinant proteins.

The 3'-terminal di-glycine motif of SUMO proteins (referred to as SUMO-GG) is essential for the ligation of SUMO to its target. Substitution of the terminal glycine to alanine (SUMO-GA) inhibits the conjugation of the small peptide to its target and serves as a negative sumoylation control [228] [229] [230]. In addition, RanBP2 is an E3 ligase, two fragments of RanBP2 (BP2ΔFG and IR1+M) were described as active E3 ligases [180] [231]. Those two truncated RanBP2 fragments where used in the IVSA to test if their E3 ligase activity facilitates sumoylation of hLa.

Expression vectors encoding His-Ubc9, GST-SAE1, GST-SAE2, GST-IR1M+1, GST-BP2ΔFG, GST-SUMO1, GST-SUMO-2, GST-SUMO3, His-SUMO1-GG, His-SUMO-1GA, His-SUMO-2WT, His-SUMO-2GA, His-SUMO-3WT, and His-SUMO3-GA were provided as a courtesy by Ron T. Hay, F. Melchior, and C. M.Chiang (2.12). All constructs, except His-SUMO-1WT/GA, His-SUMO-2WT/GA, and His-SUMO-3WT/GA, were purified by the protein production lab at the Medical University of South Carolina, USA using either a HiTrap HP Ni-NTA or a GSTrap FF column. His-tagged SUMO proteins were expressed and purified in a smaller scale using Ni-NTA columns (refer to 3.2.1 and 3.2.2). All purified proteins were quantified by Bradford and 1 μg aliquots were analyzed by SDS-PAGE and Coomassie staining. The protein isolates migrated at their expected molecular weight (figure 4.3.1), note also the molecular weight shifts of the proteins due to the approximately 26 kDa glutathione S-transferase tag: His-La (47 kDa), SAE1 (38 kDa), SAE2 (71 kDa), His-Ubc9 (18 kDa), SUMO-1/-2/-3 (11 kDa, aberrant running behavior between 11 and 20 kDa), BP2ΔFG (33 kDa), and IR1+M (7 kDa). His-tagged proteins are overall of higher purity than GST-tagged SAE1 and SAE2 proteins. Since GST-SAE1 and GST-SAE2 preparations are either partially degraded or contaminated with other proteins, the protein concentration of the

full length proteins determined by Bradford assay may be not correct. To judge the amount of full length GST-SAE1 (64 kDa including GST-tag) and SAE2 (97 kDa including GST-tag) (as indicated by an asterisk in figure 4.3.1) were used to estimate the protein concentrations compared to the intensity of the Coomassie stain with 1 μg of Ubc9, which displayed the highest purity. Furthermore, commercially available BSA (10 mg/ml) was diluted (0.25 mg/ml to 1.5 mg/ml) and loaded next to 1 μg GST-SAE1 and 1 mg/ml GST-SAE2 proteins on a SDS-PAGE in order to judge the protein concentrations of GST-SAE1 and GST-SAE2 by comparing their band intensities after Coomassie staining (data not shown). The GST-tagged SUMO-1 protein as well as GST-BP2ΔFG, and GST-IR1+M indicated that they are of high purity (figure 4.3.1C).
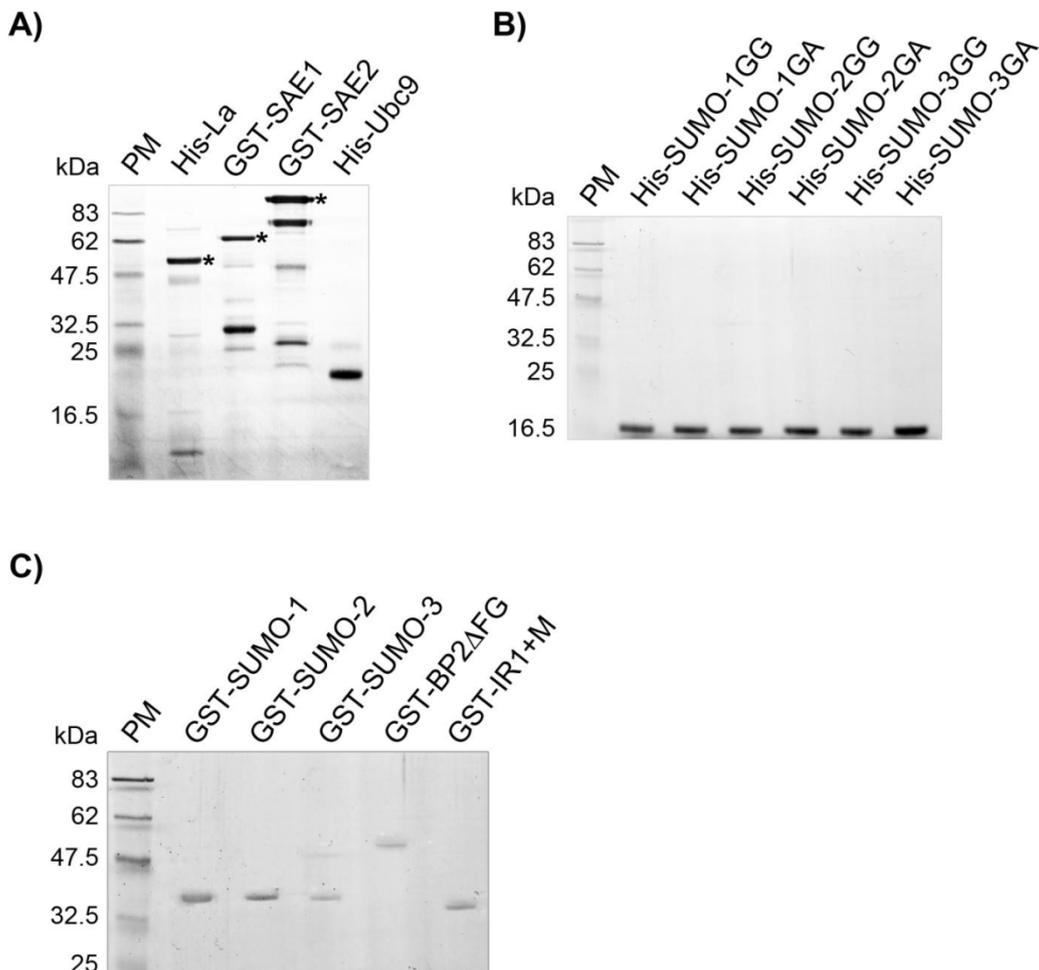
Figure 4.3.1: Coomassie staining of human recombinant proteins of the sumoylation cascade. Proteins were purified by the Protein Production Lab at MUSC. **A)-C)** Equal amounts of recombinant proteins, His-La (46.8 kDa), GST-SAE1 (64 kDa), GAST-SAE2 (97 kDa), His-Ubc9 (18 kDa), His-SUMO-1/-2/-3WT(GG)/-GA (11 kDa), GST-SUMO-1/-2/-3 (37 kDa), GST-BP2ΔFG (59 kDa), and GST-IR1+M (32 kDa), were separated by SDS-PAGE and visualized by Cooomassie staining. Asterisks indicate the correct protein size. The protein marker is on the left lane and the corresponding molecular weights are indicated left of the gel. PM = protein marker, kDa = kilodalton, His = hexahistidine tag, GST = glutathione S-transferase tag

4.3.2 Establishing an *in vitro* sumoylation assay (IVSA) for recombinant human La

In order to investigate whether sumoylation of hLa has an effect on its RNA-binding activity, an *in vitro* sumoylation assay was developed to produce SUMO-modified hLa. To establish IVSA, several previously described *in vitro* sumoylation methods were directly compared. These methods varied widely in the amount of target proteins used, in the Ubc9:SUMO ratios, and in the incubation parameters, such as the temperature and incubation time.
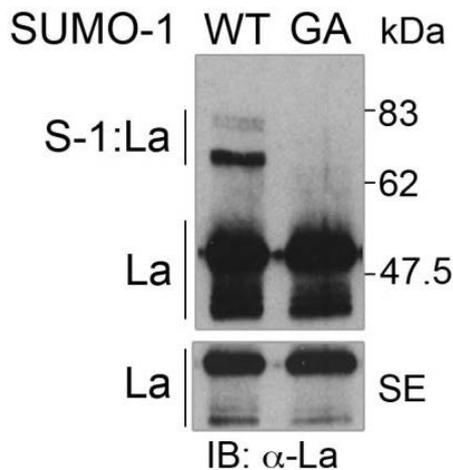
Figure 4.3.2: Modification of recombinant human La protein by recombinant SUMO-1 *in vitro*. In order to establish a robust control for the IVSA, a conjugation deficient mutant of SUMO-1 (His-SUMO-1GA) was purified and used as negative control. An optimized *in vitro* sumoylation was carried out for 2 hours at 30°C, acetone precipitated overnight, separated by SDS-PAGE and immunblotted for La. The immunoblot shows La-specific high molecular weight bands (S-1:La) only when SUMO-1WT was used for the IVSA and not when the conjugation deficient SUMO-1GA mutant was used. The molecular weights are indicated in kDa on the right site. SE = short exposure, IB = immunoblot

With this information as a starting point, a series of experiments were designed to establish conditions eventually leading to an optimal sumoylation of recombinant human La as described in section 3.2.6. Immunoblot analysis of IVSA reactions using the La-specific mouse monoclonal antibody La3B9 revealed two high molecular La-specific bands (SUMO-1:La, referred to as S-1:La) running at ~ 67.5 kDa and ~ 77 kDa (figure 4.3.2) as calculated by SDS-PAGE. These bands are La and SUMO-1WT specific, since using the conjugation-deficient SUMO-1GA protein as a control did not result in higher molecular weight bands (figure 4.3.2). As observed, only a small fraction of recombinant hLa was modified by SUMO-1 and was quantified at five percent (figure 4.3.6). Clearly visible in the short exposure of the immunoblot (figure 4.3.2, SE), the hLa antibody detects two bands for the unmodified recombinant hLa at approximately 50 and 43 kDa, the higher and more intense band represents full-length La while the lower band at 43 kDa is a fragment of hLa. The S-1:La double bands are likely the product of two SUMO peptides being covalently attached to the La protein. SUMO-1 is unable to form poly-SUMO chains, thus the S-1:La band at 77 kDa cannot be explained by SUMO-SUMO protein interaction. Any additional high molecular La-specific bands were rarely detected; suggesting that recombinant hLa is mainly

modified by a single SUMO-1 peptide *in vitro*, to a small extent by two, and rarely by three SUMO-1 peptides. The molecular weight of SUMO is approximately 11 kDa, however, with an aberrant running behavior on SDS-PAGEs sumoylation causes an approximately 20 kDa shift to the target protein. Nonetheless, the main S-1:La band at 67.5 kDa may also be a result of the conjugation of two SUMO-1 peptides and the S-1:La band at 77 kDa may represent the modification of hLa with three SUMO molecules.
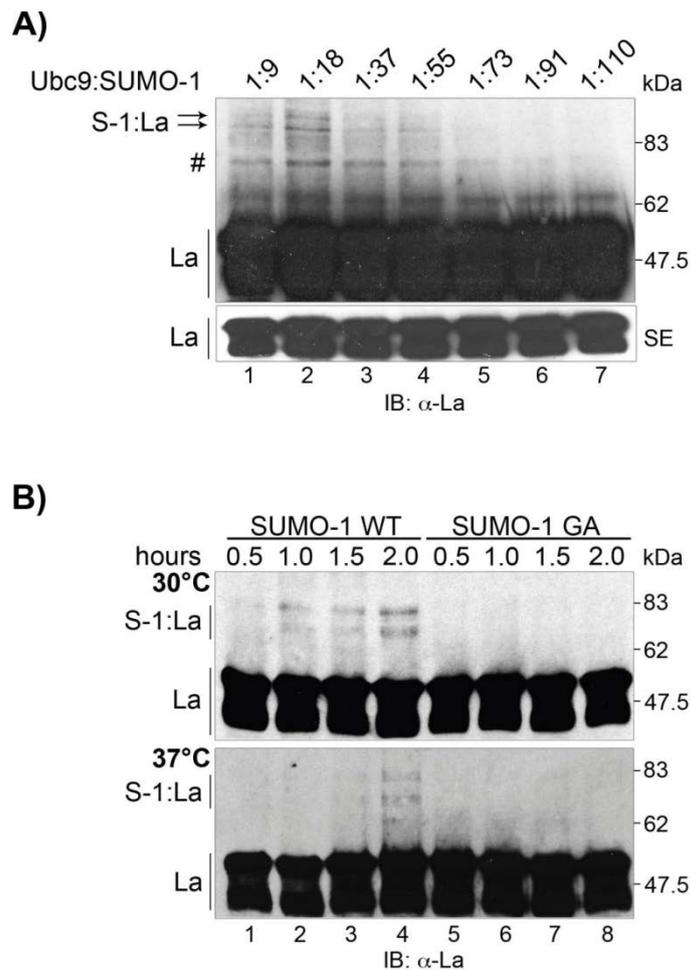
Figure 4.3.3: Optimization of *in vitro* sumoylation of recombinant human La. A) The *in vitro* sumoylation assay using varying Ubc9:SUMO-1 ratios was performed with GST-tagged SUMO-1 for 2 hours at 37°C. The reaction was quenched with 4-fold addition of 3x SDS-loading buffer; the samples were boiled for 10 min at 95°C and separated on a 7.5% SDS-PAGE. The IVSA products were visualized by immunoblotting for La. The native La and sumoylated La protein fraction (S-1:La) are indicated on the left. Ratios of Ubc:SUMO-1 beyond 1:55 result in a low level La sumoylation. The molecular weights are indicated the right. # = possible sumoylation of the La fragment **B)** Incubation parameters were analyzed for sumoylation of recombinant La protein. IVSA was carried out in parallel at 30°C and 37°C, aliquots were taken at 0.5, 1, 1.5, and 2 hours and the reaction was quenched by adding 3x SDS-loading buffer (1:4). In parallel reactions of His-SUMO-1GA were performed and served as negative control. The incubation at 30°C resulted in S-1:La after an hour (upper blot), whereas the incubation at 37°C results in an overall weaker sumoylation of La after 1.5 hours (lower blot). The native La and S-1:La species are indicated on the left, the molecular weights in kDa on the right. SE = short exposure, IB = immunoblot, kDa = kilodalton, S-1:La = SUMO-1 modified La

As mentioned earlier the ratio of the conjugating enzyme Ubc9 to SUMO-1 differed in the previously described IVSAs with SUMO-1 concentrations remaining in excess amounts to Ubc9. [177] [180] [232]. To test whether the Ubc9:SUMO ratio impacts the IVSA efficiency a wide range of Ubc9:SUMO-1 molar ratios were analyzed (figure 4.3.3A) in the IVSA (3.2.6). Standard IVSA reactions with varying Ubc9:GST-SUMO-1 ratios (1:9, 1:18, 1:37, 1:55 1:73, 1:91, and 1:110) were performed and analyzed as described (3.2.6, 3.2.7, and 3.2.9), note that the glutathione S-transferase (GST) tag adds approximately 26 kDa to the SUMO-1 protein. In consequence, the sumoylated hLa protein has a higher molecular mass of approximately 83 kDa or more as indicated by arrows in figure 4.3.3A, resulting from the aberrant running behavior of the GST-SUMO peptide. In figure 4.3.3A the intensity of the higher molecular S-1:La bands above 83 kDa depends on the Ubc9:SUMO-1 ratio. The ratio of 1:18 (0.4 μM His-Ubc9 to 7.3 μM GST-SUMO-1) is the most efficient for La SUMO-modification, whereas ratios beyond 1:55 (lanes 4-7, figure 4.3.3A) do not result in sumoylation of the La protein. However, note that the ~ 80 kDa band (# sign in figure 4.3.3A) is also Ubc9:SUMO-1 ratio-dependent and is suggested to represent either SUMO-modification of the La fragment or represents mono-sumoylated hLa, whereas the larger band above 83 kDa is poly-sumoylated hLa. For further studies, an optimal 18-fold excess of SUMO to the conjugation enzyme Ubc9 was used for all subsequent IVSA analyses.

The ratio of Ubc9 and SUMO not only differed in various publications for IVSA protocols, but also the reaction temperatures and incubation times varied. The effect of temperature and incubation time on sumoylation efficiency was studied by incubating reactions at either 30°C or 37°C, the physiological relevant temperature, for 0.5, 1, 1.5, and 2 hours. These reactions were stopped with SDS-loading buffer and separated on a 7.5% SDS-PAGE. The detection was carried out by immunoblotting for La using an α-La specific 3B9 antibody (figure 4.3.3B). The conjugation-deficient His-SUMO-1GA mutant was used as a negative control for the IVSA, displayed in lanes 5-8 of figure 4.3.3B. Comparing the upper and lower blot in figure 4.3.3B indicates that the incubation time of 2 hours resulted in the most efficient SUMO-modification. Sumoylation products shown in the upper blot of figure 4.3.3B occur as shortly after 30 minutes and increase over time. The lower blot from figure 4.3.3B indicates very weak signals for modified La and was detected only after 1.5 hours of incubation. Since an incubation of two hours was the maximum incubation time in this assay and resulted in the most efficient sumoylation of La, an additional assay was performed extending the incubation time to 4 hours in 30 min increments. However, extending the incubation time beyond 2

hours did not have any positive effect, because no increase in sumoylation efficiency of the La protein was detected (data not shown). The incubation temperature of 30°C resulted in an overall increased sumoylation efficiency as seen by more intense S-1:La bands, this intensity is about twice as much as that from the S-1:La bands incubated at 37°C, therefore, following *in vitro* sumoylation assays were carried out for 2 hours at 30°C as the ideal experimental conditions.
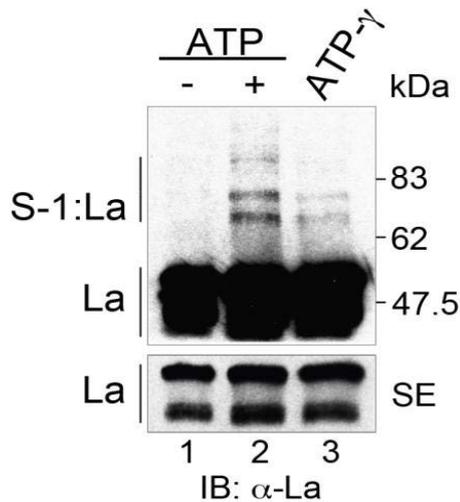
The activation of SUMO by the E1 enzymes SAE1 and SAE2 is an ATP-dependent process [177]. In the absence of ATP sumoylation should be inhibited because no energy for the activation of SUMO by ATP-hydrolysis can be supplied. To test whether the removal of an energy supply is a suitable negative control for the IVSA, IVSA reactions without the addition of ATP or with the non-hydrolysable ATP analog, phosphorothioate ATP-γ-S, were performed. After incubation for 2 hours at 30°C, samples were stopped by SDS-loading buffer, separated by SDS-PAGE and the La protein species were detected by immunoblotting. In the presence of ATP the formation of SUMO-1:La was observed (figure 4.3.4) and, as hypothesized, no SUMO-modification occurred in the absence of ATP. Replacing ATP with ATP-γ-S results in an inefficient SUMO-modification of hLa (figure 4.3.4, compare lane 2 and 3). Observe in figure 4.3.4 the high molecular band above 83 kDa in the presence of ATP in addition to the previously described double band, this 83 kDa band may represent hLa modified by three or more SUMO-peptides (figure 4.3.4, lane 2). These findings confirm that sumoylation is an energy-dependent process. In conclusion, the use of ATP-γ-S as a non-hydrolysable ATP analog is not a suitable negative control, because

sumoylation still occurred, thus the conjugation-deficient SUMO-GA mutant was used as a negative control for all following IVSA protocols.

Although the SAE1-SAE2 heterodimer and the E2 enzyme Ubc9 are sufficient for SUMO-modification, sumoylation of a target protein may be specifically enhanced by SUMO ligases *in vitro* [59, 62]. Pichler *et al.* have shown that the SUMO transfer efficiency for target proteins is strongly enhanced by the SUMO E3 ligase RanBP2. Furthermore, they identified two fragments of the 358 kDa RanBP2 ligase, namely BP2ΔFG and IR1M+1 (internal repeat 1 and middle domain), to capably act as SUMO E3 ligase.

To test whether the sumoylation of La is improved in the presence of RanBP2 ligase activity, the two functional protein fragments were added to IVSA reactions. IVSAs were performed in the presence or absence of either 13 nM BP2ΔFG or 15 nM IR1+M and immunoblotted for hLa. The conjugation deficient His-SUMO-1GA was included as negative sumoylation control. Figure 4.3.5 shows that hLa was SUMO-1-modified in the presence and absence of BP2ΔFG (lanes 1 and 2) and IR1+M (lanes 5 and 6) as represented by the two high molecular bands at 67.5 kDa and 77 kDa, both BP2ΔFG and IR1+M were not able to stimulate sumoylation of hLa, suggesting that RanBP2 is probably not a E3 ligase facilitating the sumoylation of hLa *in vitro.* However, BP2ΔFG had a small negative effect whereas in the presence of IR1M+1 sumoylation of hLa was reduced by approximately 50 percent. In conclusion, the E2 ligase RanBP2 has rather an inhibitory effect on SUMO-modification of the La protein *in vitro* rather than enhancing the IVSA efficiency.

Previously, only partial amounts of the IVSA product were analyzed by SDS-PAGE and immunoblotting (one fourth), which may inhibit the detection of weak signals of poly-SUMO-modified, more than two SUMO sites, hLa. In order to enhance the detections sensitivity and to allow for the quantification of the S:1-La species versus native hLa the IVSA products were subjected to overnight acetone precipitation, analysis of the whole reaction product was achieved by concentrating the samples after the overnight acetone precipitation, which also removed the samples from salts (3.2.6). An aliquot of the reaction product prior to precipitation was taken as a control to compare signal quality and exposure times (figure 4.3.6A). The acetone-precipitated samples showed a much stronger signal intensity for S-1:La bands than the control. As known from the literature, only a small fraction of targets are sumoylated at a given time in cells [232] [233].
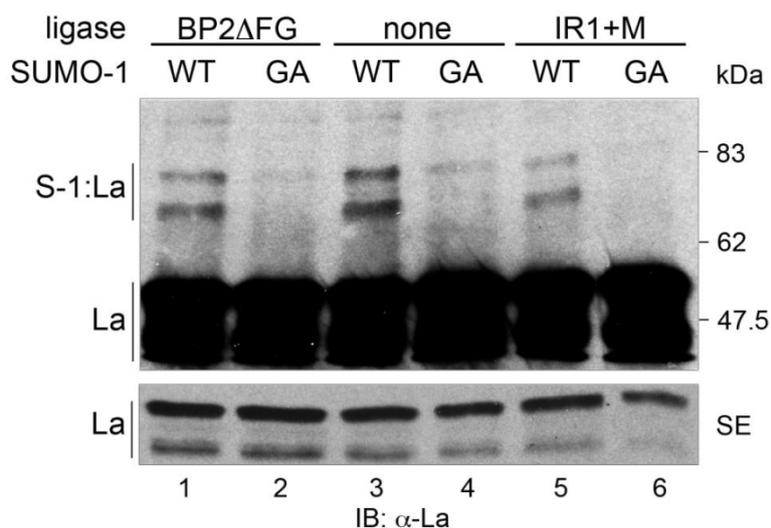
Figure 4.3.5: RanBP2 ligase activity does not facilitate SUMO-1 modification of hLa. In order to determine if RanBP2 ligase activity can increase the efficiency *in vitro* sumoylation of recombinant human La, two minimal functional recombinant RanBP2 proteins were purified and used in the IVSA. The sumoylation assay was carried out for two hours at 30°C and samples were analyzed as described. SUMO-1 modified La is indicated on the left. Modification of La is observed only when SUMO-1WT is used and not in the presence of the conjugation deficient SUMO-1GA mutant. The short exposure shows similar La levels (lower blot). The molecular weight is indicated on the right. SE = short exposure.

To determine the percentage of S1-La in IVSA reactions four sumoylation reactions were acetone precipitated, thus allowing for the determination of the modified vs. native hLa stoichiometric ratios (S-1:La to hLa). The quantification reveals that 5 ± 2.3 % of the hLa were modified by SUMO-1 (Fig. 4.3.6B), therefore, approximately 25 ng SUMO-modified hLa is produced in one IVSA. Because of the improvement seen with the precipitated IVSAs, all following IVSA reaction products were acetone precipitated prior to SDS-PAGE and immunoblotting.

In summary, the IVSA was optimized by determining the appropriate Ubc9:SUMO ratio, finding appropriate incubation parameters, and omitting the ligase activity of RanBP2. The use of the non-hydrolysable ATP analog ATP-γ-S does not serve as an adequate negative sumoylation control, but the conjugation deficient SUMO-GA mutant does. The detection sensitivity of the IVSA was increased by acetone precipitation of IVSA products overnight prior to SDS-PAGE and immunoblotting. The established robust IVSA enabled downstream functional analyses of SUMO-modulated La functions.
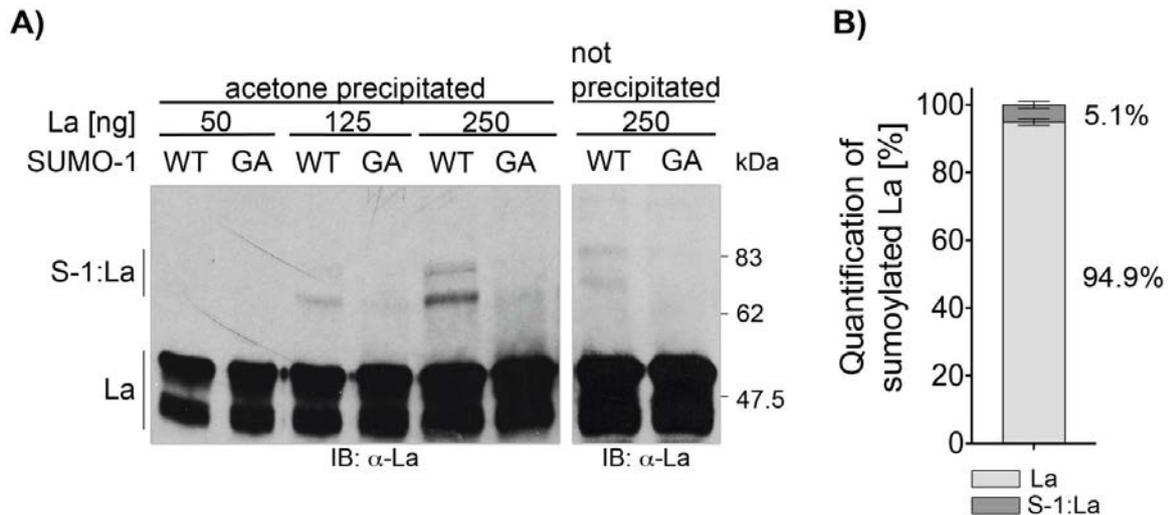
Figure 4.3.6: The detection of sumoylated La species is improved by acetone precipitation. To allow the analysis of the complete IVSA product and to enhance the immunodetection samples were precipitated with acetone overnight. **A)** IVSA was performed as described and aliquots of different hLa amounts (50, 125, and 250 ng) were taken and precipitated overnight and blotted as described. hLa modified by SUMO-1 (S-1:La) is indicated on the left. **B)** Immunoblot quantification of IVSA products reveals that on average 5% of total hLa is sumoylated. Chemiluminescence signals of four independent IVSA products were recorded (n=4) and quantified using the ImageQuant RT ECL instrument and ImageQuant TL software.

4.3.3 Recombinant human La is modified by all three SUMO isoforms *in vitro*

The modification of hLa by endogenous SUMO-1, exogenous GFP-tagged SUMO-1, and SUMO-2 has been described in sensory axons of rats. SUMO-2-modified hLa was also identified by tandem mass spectrometry in HeLa cells after heat shock [61]. Further, the modification of recombinant human La by covalent conjugation of SUMO-1 *in vitro* has been shown in figure 4.3.2. Although a target protein is specifically modified by either SUMO-1 or SUMO-2/-3 *in vivo*, it has been reported that many substrates can be sumoylated by all three SUMO paralogs *in vitro*. Taking this into consideration the next step in this investigation was to analyze whether SUMO-2 and SUMO-3 could be covalently conjugated to the recombinant human La protein.

His-tagged SUMO-2WT and SUMO-3WT were used in the optimized IVSA and the conjugation deficient His-SUMO-2GA and His-SUMO-3GA mutant proteins served as negative control in the following experiments. All reaction products were acetone precipitated overnight, separated by SDS-PAGE, and analyzed by immunoblotting for La. Figure 4.3.7A shows that all three His-SUMO-WT paralogs, but not the His-SUMO-GA mutants, were conjugated to La (S-1:La, S-2:La, and S-3:La). Intriguingly, mere modification by SUMO-1 resulted in two S-1:La bands, 67.5 and 77 kDa, suggesting that hLa is modified

by two SUMO-1 peptides, however, only one SUMO2 or SUMO-3 peptide was covalently attached to the hLa protein. This may result from one generalized main sumoylation site that is used by all three SUMO forms. In addition, IVSA products were analyzed by immunoblotting for SUMO specific antibodies in parallel to the La detection. Figure 4.3.7B displays specific SUMO-1 and SUMO-2/3 bands were detected over a wide molecular weight range, resulting from SUMO-modification of other proteins present in the IVSA reaction, those bands were detected only when His-SUMO-WT isoforms and not the conjugation-deficient His-SUMO-GA counterparts were used. It is most import to note that overlaying and aligning autoradiographs, figure 4.3.7A with 4.3.7B, revealed that S:La bands matched with bands detected by SUMO-1 and SUMO-2/-3 antibodies. Detection of the S-1:La, S-2:La, and S-3:La bands with the La antibody and the respective SUMO antibodies supports the conclusion that La is not only modified by SUMO-1 but also by SUMO-2 and SUMO-3 *in vitro*. Therefore, the previously described identification of human La modification by SUMO-1 or SUMO-2 in rat neurons and in HeLa cells [61] can also be achieved *in vitro*. Whether modification of hLa by SUMO-3 occurs also in living cells remains to be shown, however, *in vitro* SUMO-3 is conjugated to hLa.
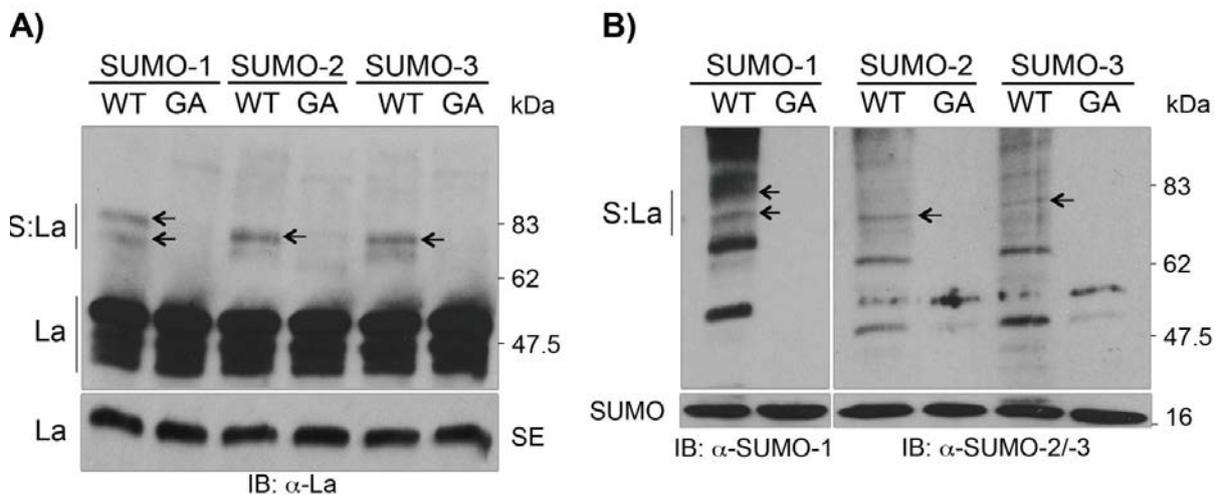
Figure 4.3.7: Human La is modified by all three SUMO isoforms *in vitro*. The modification of hLa by all three SUMO isoforms was performed to test the for specific La modification. The optimized IVSA was performed as described. The SUMO-GA mutants served as negative control. **A)** Immunoblot for La shows La-specific bands representing sumoylated La (S:La), as indicated on the left, in the presence of the wild type SUMO proteins, but not in the presence of the control SUMO proteins. **B)** IVSA products were separated by SDS-PAGE and immunoblotted on one membrane. The membrane was cut and exposed to either SUMO-1 antibody (left blot) or SUMO-2/-3 antibody (right blot). A wide range of SUMO-1 and SUMO-2/-3 specific bands were detected in the reactions with wild type SUMO isoforms. Arrows indicate matching bands detected by both La and SUMO antibodies as determined by overlaying the immunoblot for La and immunoblots for SUMO. SE = short exposure

4.3.4 La is modified by SUMO-2 at K200 and K208 *in vitro*

Two prerequisites for SUMO-modification *in vivo* have been described in the first reports on sumoylation, these requirements are the presence of a SUMO consensus sequence ΨKXE (Ψ representing a large hydrophobic amino acid, K the target lysine, X representing any amino acid, and E a glutamic acid) and a nuclear localization signal [183] [177] [184]. The SUMO consensus sequence is short, therefore, it is also found in proteins that are likely not SUMO-modified [158]. In addition, other proteins are SUMO-modified at non-consensus lysines [187] [188] [189] and some SUMO-modified proteins do not contain the consensus motif in their sequence [234] [181], thus the identification of the SUMO acceptor site of a protein may be challenging.

Van Niekerk *et al.* [60] have identified lysine 41 (K41) as the main sumoylation site for hLa, but they have also described two additional lysines, lysine 185 (K185) and lysine 208 (K208), as potential targets [60]. In order to identify the sumoylation site of hLa *in vitro*, a recombinant hLa mutant, LaK41R, with a lysine to arginine substitution at K41 was analyzed by IVSA. The expression vector encoding the human La lysine 41 to arginine substitution mutant (pet28b(+)LaK41R) was expressed in *E. coli* BL21 cells (3.2.1) and the recombinant protein was purified (3.2.2). The purified His-tagged LaK41R was then used in the optimized IVSA to determine modification *in vitro* by SUMO-1. The optimized IVSA was carried out as described in 3.2.6, the products were acetone precipitated overnight before separation by SDS-PAGE and immunoblotted for La. The conjugation deficient His-SUMO-1GA mutant served as a negative IVSA control. Figure 4.3.8A shows that LaK41R was modified by SUMO-1 *in vitro* forming the typical double band pattern at 67.5 and 77 kDa as often seen with hLaWT, hence, lysine 41 did not represent a major sumoylation site for SUMO-1 *in vitro*. It has been shown that mutations of known sumoylation sites facilitate sumoylation at other sites [235] [236] [237], therefore, the *in vitro* analysis of La mutants suggests that recombinant hLa is modified by SUMO at several lysine residues.

The additional lysines described by van Niekerk *et al.* [60] were K185 and K208, thus, a mutant La protein, LaSM123, with lysine to arginine substitutions for all three published putative sumoylation sites K41, K185, and K208, were created in order to identify the sumoylation site and investigate the function of SUMO conjugation to La in downstream experiments. The pet28b(+) La K185RK208R construct served as a template to create this triple lysine to arginine La mutant (LaSM123) by site directed mutagenesis (3.1.1). Upon

sequence confirmation, it was expressed in *E. coli* BL21 cells (3.2.1) and the recombinant protein was purified in a small-scale approach using Ni-NTA columns (refer to 3.2.2). The purified recombinant His-tagged LaSM123 protein was then used as a target for SUMO-1 in the optimized *in vitro* sumoylation assay (3.2.6). The conjugation deficient His-SUMO-1GA peptide served as a negative control. Immunoblotting the IVSA product for La, represented in figure 4.3.8B, showed a higher molecular weight species at 77 kDa, representing SUMO-1 modified LaSM123, which was not formed using the control SUMO-GA peptide. Since the sumoylation of LaSM123 resulted in only one S-1:LaSM123 species this result suggests that a single sumoylation site was successfully mutated.
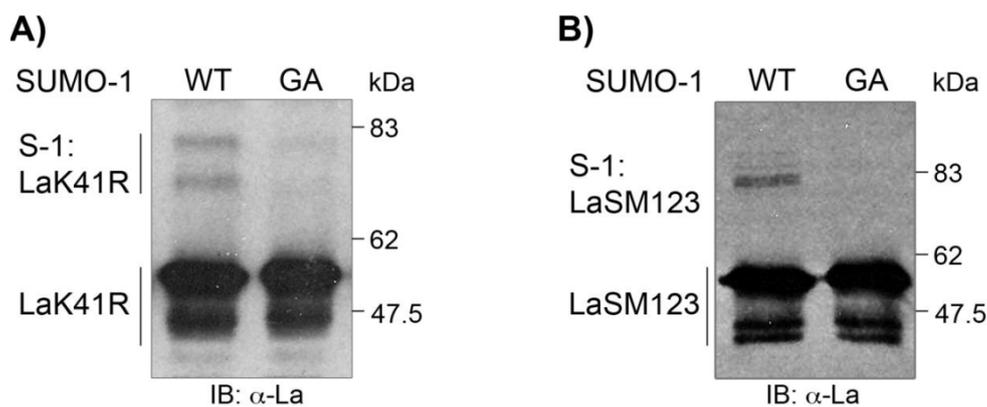
Figure 4.3.8: Sumoylation of lysine to arginine substitution mutants of hLa. Recombinant La mutants with lysine to arginine substitutions of putative sumoylation sites were analyzed for sumoylation to identify the SUMO-acceptor sites in hLa. IVSA were performed as described and reactions with the control SUMO-1GA mutant were run in parallel. Immunoblot analysis revealed both sumoylation site mutants are modified by SUMO-1 *in vitro*. **A)** Two La-specific high molecular weight species, S-1:LaK41R, were detectabled in the presence of SUMO-1WT. **B)** Merely one S-1:LaSM123 band was detected when SUMO-1WT was used in the IVSA.

Up to this point the IVSA analyses suggest that hLa is modified at a minimum of two lysine residues. The human La protein contains 51 lysines, which equals 12.5% of the protein, thus the preparation of recombinant proteins with single lysine substitutions and combinations of mutations were not reasonable. In order to identify additional sumoylation sites an i*n silico* approach using two different computer programs was used. The SUMOplot analysis program has been used by van Niekerk and coworkers [60] to predict sumoylation sites and was used herein as well; in addition, another *in silico* SUMO-prediction tool SUMOsp 2.0 [238] was published and utilized to predict sumoylation sites in the human La protein. The FASTA sequence of human La (GenBank reference P05455.2) was submitted to the software or the web server for analysis by SUMOsp 2.0 and SUMOplot analysis programs, respecitively. The SUMOsp2.0 program was able to identify consensus and non-consensus sequences for

sumoylation, however, note that a medium cutoff was chosen for SUMOsp 2.0 to exclude low confidence SUMO acceptor sites. A summary of each analysis can be found in table 4.3.1 for SUMOsp 2.0 and in table 4.3.2 for SUMOplot analysis program. The already established main sumoylation site lysine 41 identified by van Niekerk *et al.* [60] was predicted by both programs (compare table 4.3.1 with table 4.3.2), but was not predicted by SUMOsp 2.0 as a sumoylation site when the prediction stringency was increased by choosing the high cut-off (data not shown).

Position	Peptide	Score	Cutoff	Type
37	KFLKEQI	2.706	2.26	TypeII: Non-consensus
41	EQIKLDE	2.265	2.26	TypeII: Non-consensus
74	ALSKSKA	2.265	2.26	TypeII: Non-consensus
76	SKSKAEL	2.868	2.26	TypeII: Non-consensus
105	DEYKNDV	2.897	2.26	TypeII: Non-consensus
200	KQNKVEA	3.338	2.26	TypeII: Non-consensus
208	LRAKQEQ	2.142	0.1	TypeI: Ψ-K-X-E
214	QEAKQKL	2.691	2.26	TypeII: Non-consensus
216	AKQKLEE	2.324	2.26	TypeII: Non-consensus
224	AEMKSLE	2.426	2.26	TypeII: Non-consensus
383	GPVKRAR	2.309	2.26	TypeII: Non-consensus
391	ETDKEEP	2.676	2.26	TypeII: Non-consensus
397	PASKQQK	3.426	2.26	TypeII: Non-consensus
400	KQQKTEN	3.029	2.26	TypeII: Non-consensus

Table 4.3.1: Overview of sumoylation site predictions in the human La protein by the SUMOsp2.0 software [322]. The FASTA sequence of the human La protein sequence was entered and the medium threshold was chosen to predict SUMO acceptor sites in the human La protein. The peptide sequence around the acceptor lysine as well as the position of the predicted SUMO site is indicated. The score, Cutoff, and the type of the SUMO site sequence are listed. The type I SUMO site is represented by a lysine in the classical sumoylation consensus sequence. Type II sequences contain other non-canonical sites.
.

However, as mentioned above lysine 185 was predicted with a high probability (score = 0.85) to be SUMO-modified by the SUMOplot Analysis Program (table 4.3.2), but was not predicted by SUMOsp 2.0 (table 4.3.1). Interestingly, lysine 208 which was described by van Niekerk *et al.* [60] as a potential SUMO-site was not identified as a potential site by SUMOplot Analysis Program, but it was predicted by SUMOsp 2.0 as one with the highest probability to be SUMO-modified. Moreover, this lysine 208 residue is the only lysine embedded in the sumoylation consensus sequence. Surprisingly, both programs shared the

prediction of only five residues, namely lysine 41, 208, 216, 391, and 400, however, lysine 41 was predicted by both algorithms, but it is not the sumoylation site of hLa *in vitro* as experimentally determined.

Despite some common predicted sumoylation site the *in silico* analysis was not as promising, because of the large and varying number of potential SUMO acceptor sites.

Position	Peptide	Score
K41	FLKEQ **IKLD** EGWVP	**0.94**
K116	NRSVY **IKGF** PTDAT	0.59
K148	MRRTL **HKAF** KGSIF	0.17
K165	DSIES **AKKF** VETPG	0.44
K185	DLLIL **FKDD** YFAKK	**0.85**
K192	DDYFA **KKNE** ERKQN	0.48
K208	EAKLR **AKQE** QEAKQ	**0.79**
K216	EQEAK **QKLE** EDAEM	0.5
K229	MKSLE **EKIG** CLLKF	0.33
K269	DFVRG **AKEG** IILFK	0.62
K391	AREET **DKEE** PASKQ	0.5
K400	PASKQ **QKTE** NGAGD	0.5

Table 4.3.2: Overview of sumoylation site predictions in the human La protein by the SUMOplot Analysis Program. The FASTA sequence of the human La protein sequence was entered at http://www.abgent.com/tools/ to predict the probability for the SUMO-consensus sequence to be engaged in SUMO-conjugation of the human La protein. The positions as well as the sequence around the lysine acceptor sites are displayed in the table. The La protein specific SUMO-consensus sequence is in bold. The displayed score is based on the direct match of the amino acids to the consensus sequence and the substitution of residues exhibiting a similar hydrophobicity the higher the score the higher the probability of the lysine to serve as a SUMO-acceptor site (Abgent).

Recently the laboratory of P. Thibault (McGill University, Montreal, Canada) reported a new method for the identification of sumoylation sites [239], this approach utilized mass spectrometry, which is a much better approach in identifying SUMO-modified residues. Dr. Thibault was willing to perform the required experiments for the identification of the sumoylation sites used during *in vitro* sumoylation of hLa by SUMO-2.

Constructs containing the human La sequence were sent to Dr. Thibault's laboratory for tandem mass spectrometry analysis using SUMO-2 as a modifier. The analysis identified lysine 200 and lysine 208 as high confidence SUMO-2 acceptor sites, in addition, lysines in the C'-terminus of the La protein (K344, K352, K363, and K400) were also identified as lower confidence sumoylation sites (data not shown). Interestingly, only the high confidence assignments of K200 and K208 and the low confidence K400 were predicted by SUMOsp 2.0 (refer to table 4.3.1).

The main sumoylation sites K200 and K208 are located in the linker region between the RNA recognition motifs RRM1 and RRM2. The minimal La protein, RRM1+2, consisting of RRM1 and RRM2 as well as its interdomain peptide linker has been shown to be RNA-binding active (refer to 4.2.2). In order to confirm lysine 200 and 208 as the main SUMO acceptor sites in the La protein, a double K200R and K208R substitution La mutant, La K200R K208R, as well as a single K208R La mutant, La K208R, were created in the minimal La RRM1+2 context. The minimal protein context was chosen to eliminate the possibility of sumoylation at putative C'-terminal lysine residues that were identified as low confidence SUMO-acceptor sites by mass spectrometric analyses. The mutagenesis and cloning strategies are described in detail in sections 3.1.1 and 3.1.2. The sequence accuracy of the substitution mutants were confirmed by DNA-sequencing before the two constructs were expressed in *E. coli* BL21 cells (3.2.1) and purified in small-scale using Ni-NTA columns as described in section 3.2.2. The concentration was determined by Bradford and the proteins were analyzed for purity by SDS-PAGE and subsequent Coomassie staining (figure 4.3.9). The purified RRM1+2 mutants were of high purity and of the expected size (~ 25 kDa) according to their amino acid sequence.
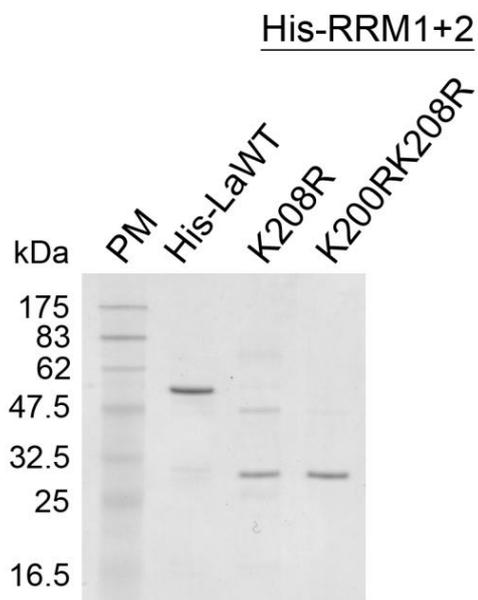
Figure 4.3.9: Coomassie staining of recombinant hLa proteins. Lysine to arginine substitutions were introduced in the minimal human La protein RRM1+2 context. The proteins were purified in small scale using Ni-NTA columns. The protein concentrations were determined by Bradford and equal molarities of the RRM1+2 mutants and the wild type hLa protein were separated on a 12.5% SDS-PAGE. The purity and correct size of the proteins were assessed by Coomassie staining. The protein marker (PM) is on the left with its corresponding molecular weights left of the gel. The RRM1+2 proteins are of correct size between 25 and 32.5 kDa.

The recombinant minimal RRM1+2 lysine substitution mutants as well as their wild type counterpart RRM1+2 were used as targets for *in vitro* sumoylation with His-SUMO-2WT and the negative control His-SUMO-2GA in 10-fold up scaled reactions to study the function of SUMO conjugation to La in downstream functional assays. Aliquots were acetone precipitated overnight, the reaction products were then separated by SDS-PAGE and the La

species were detected by immunoblotting using the La-specific mouse monoclonal SW5 antibody [240]. The previously used mouse monoclonal 3B9 antibody recognized an epitope in the NTD of hLa, thus it was not used for the immunodetection of RRM1+2. In figure 4.3.10 the wild type RRM1+2, compared to the control reaction with SUMO-2GA, displayed two higher molecular weight bands representing S-2:RRM1+2, these major band were slightly above 47 kDa and a weaker band above 62 kDa. As mentioned before hLaWT was modified by merely one SUMO-2 species (figure 4.3.7). Taken together, these findings implicate the efficient conjugation of one SUMO-2 peptide to hLa and RRM1+2, a secondary sumoylation site recognized weakly by SUMO-2. Interestingly, only the major S-2:RRM1+2 band was detectable when hLaK208R was sumoylated (Fig. 4.3.12 compare lanes 3 and 5), suggesting that lysine 208 is a weak acceptor site for SUMO-2. Intriguingly, the band intensity was a bit stronger compared to the wild type S-2:RRM1+2 bands, also suggesting that mutation K208R stimulates sumoylation at a different site. However, analysis of the double hLa mutant RRM1+2 K200RK208R demonstrates that major sumo acceptor sites were mutated. In conclusion, it was determined that the main SUMO acceptor sites for recombinant RRM1+2 protein were lysine 200 and 208.
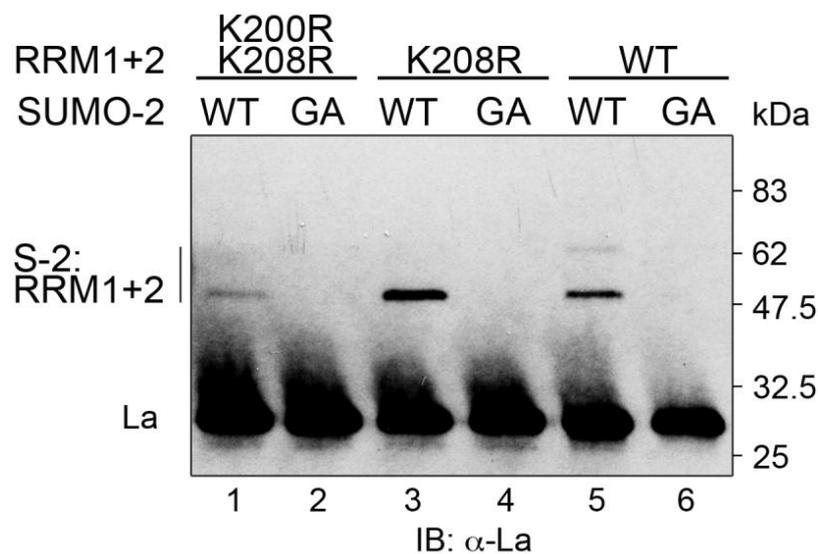
Figure 4.3.10: Lysine 200 and 208 are the main sumoylation sites of hLa *in vitro*. In order to study the effect of SUMO-modification on the La protein the main sumoylation sites were identified by IVSA. The recombinant RRM1+2 proteins were sumoylated with SUMO-2WT and SUMO-2GA, acetone precipitated overnight, separated by SDS-PAGE followed by immunoblotting for La. Sumoylated RRM1+2 (S-2:RRM1+2) and its respective mutants and native La are indicated on the left. The molecular weights are indicated in kDa on the right.

In summarization, the *in silico* prediction of putative SUMO attachments sites using two free programs, namely SUMOplot Analysis Program and SUMOsp 2.0, roved not to be a helpful tool in the case of human La. However, La has a high abundance of lysines that could serve as potential SUMO acceptor sites resulting in a high number of *in silico* predicted sumoylation sites. On the other hand, in rat neurons, La is SUMO-modified at lysine 41 [60], which could not be confirmed *in vitro* suggesting either a cell-type specific sumoylation at that site or the involvement of other lysine residues *in vitro*. The more practical mass spectrometral analysis of human La modification by SUMO-2 identified two main and four minor SUMO sites. These two main SUMO-conjugation sites at lysine 200 and lysine 208 were confirmed by *in vitro* sumoylation of a minimal La protein excluding the four minor SUMO acceptor sites in the C'-terminus of hLa.

4.3.5 Sumoylation of La enhances cyclin D1 RNA-binding

The two identified sumoylation sites K200 and K208 are located within the interdomain peptide linker between RRM1 and RRM2, which are both required for D1-FL binding. It is likely that binding of an RNA molecule by RRM1 and RRM2 induces structural changes in the La protein, thus it is reasonable to speculate that a bound RNA may prevent or increase sumoylation of hLa *in vitro*. In order to test this, the La protein was pre-incubated with either the D1-FL RNA or the smaller L22 RNA oligoribonucleotide which were both bound by hLa (figure 4.3.11) before it was subjected to sumoylation. Further, parallel reactions in the presence of 0.5 μg heparin, which is a non-specific RNA competitor, were prepared to study whether La is sumoylated at a higher efficiency as a monomer or multimer. The hLa was pre-incubated for 10 minutes at room temperature with 10 nM RNA in the presence or absence of 0.5 μg heparin prior to performing the IVSA with SUMO-1 or the conjugation deficient SUMO1-GA mutant (3.2.6). Subsequently, sumoylation products were acetone precipitated, separated by SDS-PAGE and visualized by immunoblotting for La. As shown (figure 4.3.11A) pre-incubation of hLa with the long (212 nts) D1-FL RNA and the small L22 RNA (37 nts) almost completely prevented sumoylation (figure. 4.3.11A lanes 1 and 3). The lower signal intensity in lane 1 in figure 4.3.11B is likely due to the lower amount of hLa loaded, however, in the presence of heparin sumoylation of hLa was more efficient than in the absence of heparin (figure 4.3.11). Interestingly, the addition of heparin strongly restored sumoylation of hLa in the presence of RNA (lanes 2 and 4, figure 4.3.11A and B). Note that

in the presence of RNA, heparin had a strong positive effect on the lower S-1:La band whereas the positive effect on the higher molecular band of S-1:La was not as strong. In conclusion, this experiment suggests that binding of a long RNA molecule impedes sumoylation of hLa and that the addition of heparin facilitates and restores sumoylation of the La:RNA complex. This findings suggest that only monomeric La:RNA complexes are sumoylated because it is shown that heparin prevents the formation of multimeric La:RNPs (figure 4.2.5).
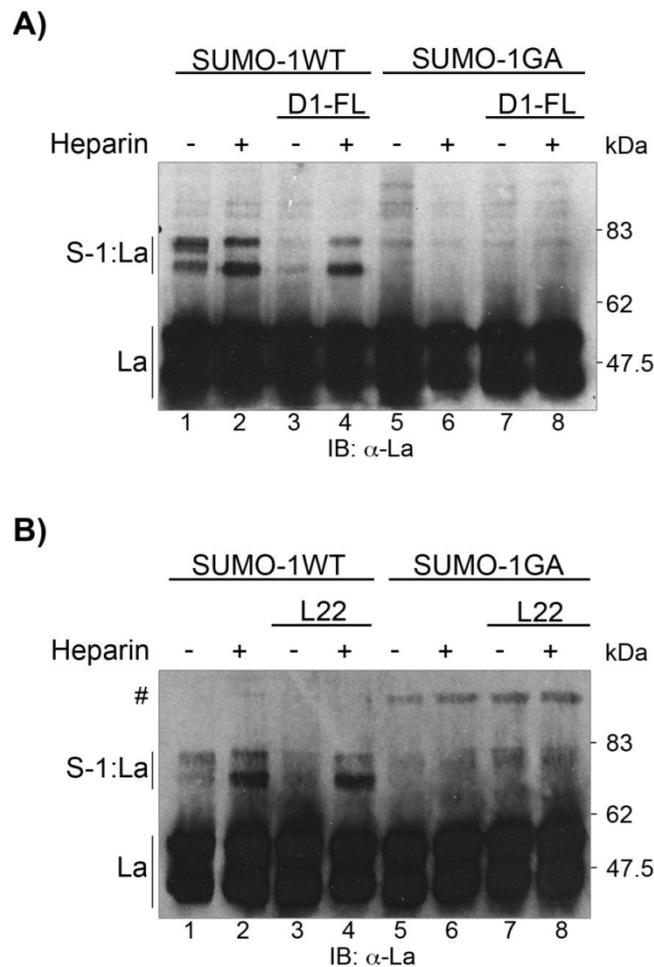
Figure 4.3.11: The non-specific competitor heparin facilitates SUMO-modification of hLa. Recombinant La protein was incubated with either 10 nM D1-FL RNA **(A)** or Cy3-labeled L22 RNA **(B)** to determine whether RNA-binding of the La protein improves modification of La by SUMO-1. The unspecific competitor heparin was used to study whether La is sumoylated preferentially as monomer or multimer. The IVSA was performed in presence or absence of RNA and heparin (as indicated). The sumoylation was controlled with parallel reactions using the conjugation deficient SUMO-1GA mutant. SUMO-conjugated La (S-1:La) and native La are indicated on the left. Unrelated La specific bands are indicated by the # sign.

The posttranslational modification of a protein by SUMO may have consequences on the functionality of the substrate as discussed in detail in the introduction (section 1.3), one consequence is the conformational change of the substrate, which directly regulates its function [172]. In the case of hLa sumoylation between the two RRMs may induce structural changes affecting its RNA binding activity.
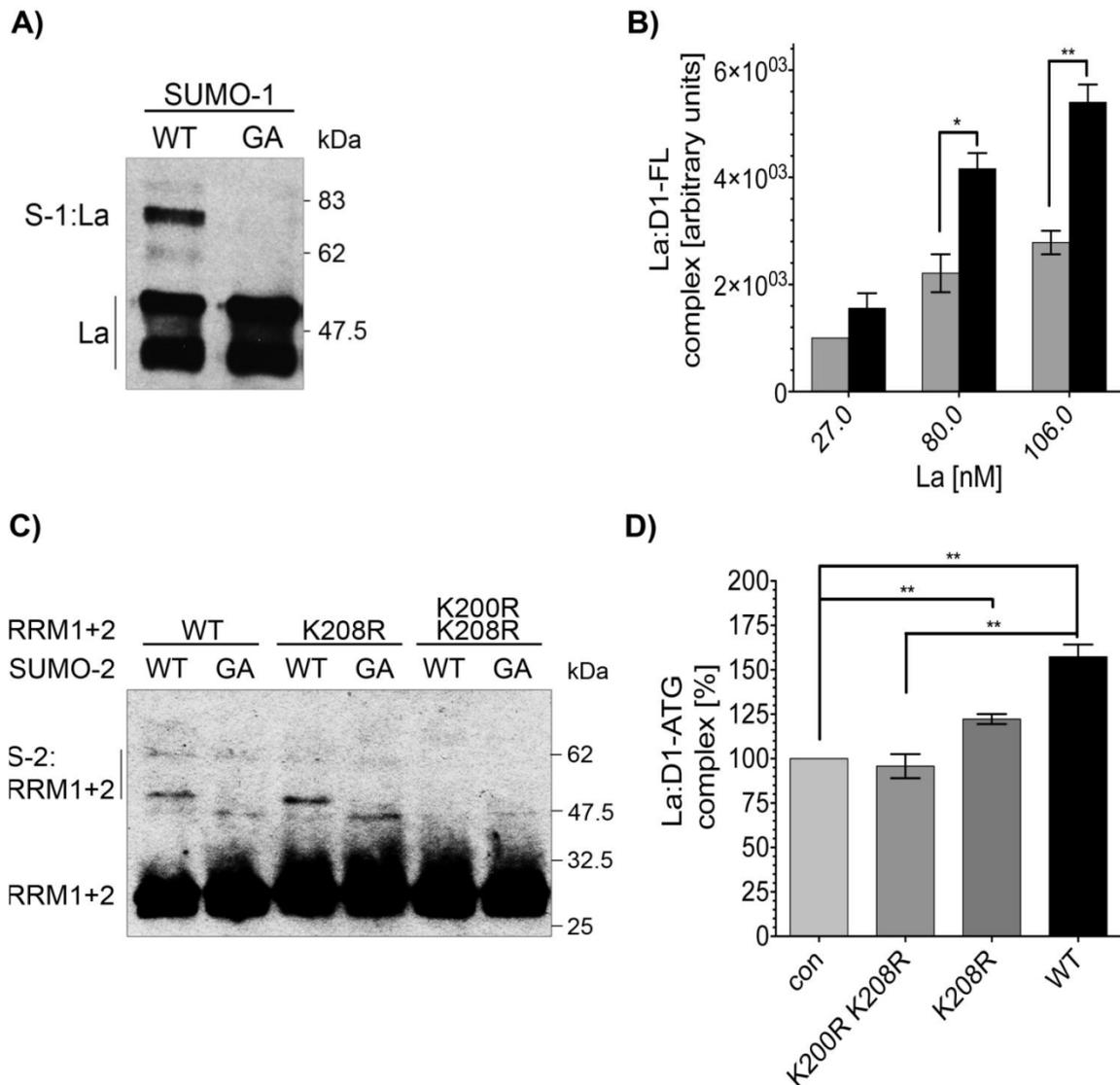
Figure 4.3.12: Modification of hLa with SUMO results in an increased binding activity to cyclin D1 RNA. In order to analyze whether sumoylation of La affects RNA-binding activity, wild type La and RRM1+2, including the substitution mutants of RRM1+2, were sumoylated and used for RNA-binding studies by native EMSAs. Three independent IVSA reactions were performed (n=3). **A)** and **C)** Representative immunoblots of IVSA used for downstream EMSA studies. Native La species were quantified to allow concentration adjustments for the downstream RNA-binding studies. **B)** EMSAs were carried out as described in 3.3.6.2 and quantified. The EMSA quantification revealed that significantly more RNA was shifted into complexes when 80 nM and 106 nM SUMO-1 modified La (S-1:La, black bar) was used compared to the control reaction (con La, grey bar). n=4 **D)** 80 nM of the IVSA products were used to analyze the binding to [32P]-labeled cyclin D1-ATG RNA by native EMSA. Quantification of those EMSAs showed a significantly higher RNA:La complex formation for sumoylated RRM1+2 and RRM1+2 K208R. The double mutant RRM1+2 K200RK208R had a similar RNA-binding activity as the control La (con, light grey bar). The paired two-tailed Student's t test was used to determine the *p*-value and displayed above the bars. Standard deviations from the mean values are indicated by error bars. n=3

In order to test the activity of La to bind cyclin D1 RNA regulated by SUMO-modification, electrophoretic mobility shift assays were performed to compare the RNA-binding of native La with SUMO-modified La. The RNA binding-activity of SUMO-modified wild type La to radiolabeled D1-FL RNA was analyzed. The La protein was modified by SUMO-1WT and a parallel reaction with conjugation-deficient SUMO-1GA mutant as a control (3.2.6). Since the buffering conditions for the IVSA differ from those required for the binding studies, the buffer was exchanged to EMSA buffer (refer to 3.3.6.2). An aliquot of the IVSA reaction in the RNA-binding buffer was taken, acetone precipitated and detected by immunoblotting for La (refer to 3.2.7 and 3.2.9). The sumoylated and control hLa were generated fresh for each EMSA, with the efficiency and concentration of La validated each time by immunoblotting, a representative example is shown in figure 4.3.12A. EMSAs were performed and quantified as described in sections 3.3.6.2 and 3.3.6.7 Three concentrations (27, 80, and 106 nM) of S-1:hLaWT were titrated in the presence of 10 nM of $[^{32}P]$-labeled D1-FL RNA, the same was done in parallel for control La. A representative EMSA is shown in appendix figure A1. These EMSA studies in figure 4.4.12B revealed an overall enhanced binding activity of SUMO-modified La (black bar) versus control La (grey bar) in three independent experiments. The difference in binding is noted as increasing with the La concentration and significant when 80 and 106 nM protein was used for the native EMSA.

The mutational elimination of the SUMO-conjugation sites in the minimal RRM1+2 context directly allows for drawing conclusions of the function of SUMO-conjugation on RNA-binding of the human La protein. Therefore, RRM1+2, the single mutant RRM1+2 K208R, and the double mutant RRM1+2 K200RK208R were subjected to *in vitro* sumoylation using SUMO-2. Parallel reactions with the conjugation deficient SUMO-2GA mutant were performed as a control. As described above, the buffer was exchanged to the RNA-binding buffer and an aliquot of each sample was analyzed for sumoylation by immunoblotting for La (refer to 3.2.9). The La amounts were quantified using the ImageQuant RT ECL instrument and ImageQuant TL software. This information was then used for adjusting the La concentrations for subsequent EMSA studies. The EMSAs in figure 4.3.12D were carried out as described before (3.3.6.2) using 10 nM $[^{32}P]$-labeled D1-ATG RNA as substrate for 80 nM SUMO-modified and control hLa (light grey bar). RNA-binding studies were performed in three independent experiments and quantified using a Storm phosphorimager and ImageQuant TL software as described in section 3.3.6.7. The control reactions were set to

100% and the RNA-binding by the sumoylated proteins were normalized in relation to the control reactions.

As noted in figure 4.3.12D sumoylation of the wild type RRM1+2 results in a significantly higher D1-ATG RNA-binding activity compared to the control. Interestingly, RRM1+2 K208R binds less RNA compared to sumoylated hLaWT, implying that the removal of one SUMO molecule affects its RNA-binding activity. Strikingly, the RNA-bindning activity of RRM1+2 K200RK208R is very similar to native hLa, strongly suggesting that SUMO-modification at lysines 200 and 208 facilitate the RNA-binding activity of hLa. This finding has potential significant functional implications if similar observations would be made with endogenous hLa in living cells.

In short, SUMO-modification of the minimal RRM1+2 La protein increases the RNA-binding activity. This RNA-binding advantage is lost upon eliminating the main sumoylation sites in RRM1+2 K200R K208R.

4.3.6 Sumoylation of La enhances 5'-TOP RNA-binding

Cytoplasmic La protein is involved in the translational regulation of a class of mRNAs containing a 5'-terminal oligopyrimidine track, referred to as 5'-TOP mRNAs [241] [104] [242] [93]. These 5'-TOP mRNAs encode for ribosomal proteins and other factors required for the translation machineries including elongation factors [243] [244].

In order to confirm that sumoylation of La facilitates the RNA-binding activity binding of the hLa protein to three RNA oligoribonucleotides (oligos) were analyzed by native fluorescent EMSA. Two of those RNA oligo represent the 5'-terminal 37 nucleotides of two rodent TOP mRNAs, namely L37 and L22, both encoding for the rodent ribosomal proteins L37 and L22. Further, the 37 nts long oligoribonucleotide representing the 5'-terminus of the unrelated GAP43 mRNA, which encodes for the growth associated protein 43 (GAP43), was included in the EMSA studies as well. The ribosomal protein mRNAs were chosen because these mRNAs localize to the axonal compartment of rat sensory neurons [60] [245] and could be exposed to sumoylated La proteins *in vivo*. The RNA oligoribonucleotides were 3'-end labeled with the fluorescence dye C3 cyanine to allow for visualization by native EMSAs, note the advantage of fluorescence vs. radiolabeled RNA probes discussed in section 4.2.1.

A binding curve using increasing hLa concentration was established for each RNA oligo and a RNA concentration of 22 nM was allowed to bind to the hLa protein, hLa-RNP formation

was then analyzed by native EMSA (figure 4.3.12A). The RNA-binding by hLa was quantified and the dissociation constants for the L37 and L22 RNP formation was determined at approximately 100 nM (data not shown), thus the affinity of the La protein for those two targets is similar to the D1-FL RNA. No saturation was reached for the GAP43 binding by hLa (data not shown).

The effect of SUMO-modification of the hLa protein on its TOP RNA-binding activity was analyzed next by IVSA, which was followed by native fluorescent EMSA. Recombinant hLa was sumoylated by the optimized IVSA and SUMO-1GA at 10-fold reactions. A parallel reaction with the conjugation deficient SUMO-1GA mutant was used to control the sumoylation reaction (i.e. con La). The IVSA product was exchanged to the RNA-binding buffer for an optimal EMSA (3.3.6.2) and an aliquot of the reaction was then acetone precipitated overnight and subjected to SDS-PAGE with subsequent immunoblotting for hLa to determine the sumoylation efficiency. Figure 4.3.13B displays a representative immunoblot of IVSA products used for downstream RNA-binding studies. The hLa protein is modified by SUMO-1WT, which is represented by the intense higher molecular weight band at ~ 70 kDa and the less intense ~ 80 kDa band, but not the conjugation deficient mutant. Further, the buffer exchanged IVSA products were checked for comparable hLa levels (short exposure SE, figure 4.3.13B). The validated IVSA products were then used in increasing amounts (27, 80, and 106 nM) for L37 RNA or L22 RNA-binding EMSA studies. Figure 4.3.13C shows a representative native fluorescent EMSA with L37 RNA as a La target, two RNP complexes are formed, but both complexes have a very similar mobility. Comparing the upper band with the lower band of free RNA in figure 4.3.13C displays that the free RNA contains two distinct RNA species of different mobility, but mainly the lower mobility RNA was shifted into a complex with the La protein. Quantification of the native fluorescent EMSAs using *in vitro* sumoylated La protein revealed that significantly more L37 and L22 RNA oligoribonucleotides were shifted into a complex with 80 nM S-1:La (black bar), and 106 nM in the case of the L37 RNA this was in comparison to the same concentrations of the control La (grey bar) (figure 4.3.13D). In conclusion, the SUMO-1 modified hLa protein binds RNA oligoribonucleotides representing strong and uridine-rich TOP elements *in vitro* more efficiently than the native protein. Currently, this is the first report demonstrating that SUMO-modification of a RNA-binding protein enhances its RNA binding activity.
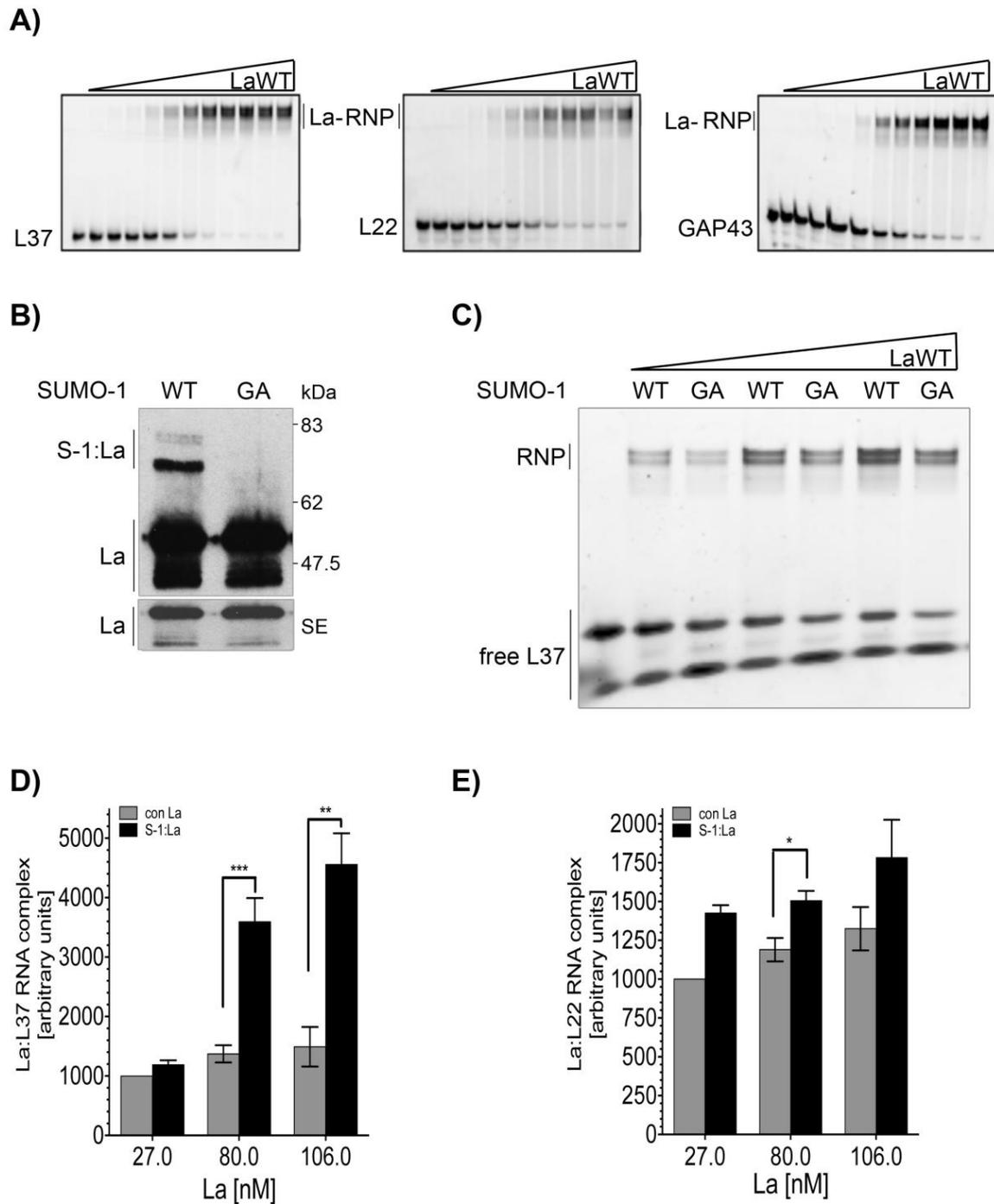
Figure 4.3.13: SUMO-modification of hLa enhances its 5'-TOP mRNA-binding activity. The RNA-binding of La to 5'-terminal oligopyrimidine (TOP) elements was studied by native EMSAs. An unrelated Cy3-labeled RNA oligoribonucleotide, GAP43, was included in these studies. **A)** The fluorescence-based EMSA was performed as described in 3.3.6.6. **B)** Representative immunoblot of IVSA used for downstream EMSA studies. Native La species were quantified to allow concentration adjustments for the downstream RNA-binding studies when La protein levels of S-1:La and con La were not comparable. **C)** Representative fluorescence-based EMSA of La binding L37 RNA oligoribonucleotide. **D)** and **E)** Quantification of EMSAs of La with the L37 RNA **(D)** and L22 RNA **(E)**. **D)** At high La concentration of 80 and 106 nM SUMO-1 modified La (black bar) shifted significantly more L37 RNA into a ribonucleoprotein complex than control La (grey bar). The *p*-value was determined from ten independent experiments (n=10) in Prims5 using the two-tailed paired Student's t-test. **D)** More of the L22 RNA oligoribonucleotide was shifted into La-RNP complex with S-1:La (black bar) than with control La (grey bar). The *p*-value is indicated. n=4

5. Discussion

5.1 The protein binds in close proximity to the cyclin D1 mRNA translational start site

One of the major functions of La is to facilitate translation of positive sense viral RNAs and of cellular mRNAs in the cytoplasm, e.g. viral RNAs of HIV [102], HCV [62] [99] [67], and poliovirus [98] [246] [67], cellular mRNAs, e.g. XIAP [91], Mdm2 [10], laminin B1 [94], and cyclin D1 [89] in the cytoplasm. Interestingly, those cellular mRNAs are translated into proteins with tumor-promoting activities such as inhibition of apoptosis (XIAP), the negative regulation of p53 tumor suppressor (Mdm2), and supporting epithelial-mesenchymal transition (EMT; laminin B1). Although hLa binding sites for cellular and viral mRNAs were mapped [94] [102] [91] [10], a consensus sequence for hLa binding sites in mRNAs has not been determined. Binding of hLa in close proximity to the translational start site has been reported for HCV [62] [113], poliovirus [98] [246], and CCND1 as described in more detail below.

Little is known about the molecular mechanism by which the La protein stimulates translation of mRNAs with internal binding sites. The development of novel cancer therapeutics targeting the La protein, which is overexpressed in solid tumors [11] and tumor cell lines [127], would require molecular details into how hLa is interacting with mRNAs of tumor-promoting factors and how they stimulate their translation. A major goal of this work was to characterize the interaction of human La with the mRNA of the cooperative oncogene cyclin D1 at the molecular level, and to develop novel strategies to reduce cyclin D1 expression in cancer cells by targeting the interactions of hLa with the CCND1 mRNA.

Different EMSA techniques were used to identify the La-binding site in the cyclin D1 mRNA *in vitro*. As expected, the human La protein recognized the entire cyclin D1 5'-UTR (D1-FL) consisting of the 209 nucleotides from the 5'-UTR and AUG start codon. Further mapping identified the major binding site within the 3'-terminal 113 nucleotides (figure 4.1.3B, part B) and a minor binding site within the 5'-terminus (figure 4.1.3B, part A, nts -209 to -110). A combination of different approaches to include blocking of hLa's binding to D1-FL RNA by DNA oligonucleotides (figure 4.1.4) and competitive EMSA (figures 4.1.5 to 4.1.6) allowed for the identification of the hLa binding site between nts -8 and +3 (figure 4.1.6D). RNA oligoribonucleotides (figure 4.1.7, figure 4.1.9) were used to illuminate the hLa binding site in more detail and revealed that the uridine of the AUG is not critical for binding (figure

4.1.9). However, changes mainly located in the Kozak consensus sequence demonstrate that nucleotides – 3 to -1 and +3 to +4 are critical for La binding (figure 4.1.9).

It has been shown that La binds to the translational start site context of the HCV 5'-non coding region (5'-NCR) as shown by [62]. In addition, earlier work suggested La binding to the highly conserved polypyrimidine tract/AUG motif in the 5'-NCR of poliovirus [246]. Furthermore, a previous study revealed that the sequence context of the translational start site is critical for La binding [111] suggesting that hLa binds preferentially to the translational start site when it is embedded in a strong Kozak consensus sequence. What needs to be addressed is how hLa stimulate IRES-mediated translation by binding at or in close proximity of the translational start site.

IRES-mediated translation is achieved by directly recruiting the 43S ribosomal subunit to the vicinity of the translational start site codon AUG is achieved, this is circumventing the necessity of the 5'-cap structure and the associated initiation factors [247] [248]. The 43S subunit scans along the 5'-UTR to find the authentic AUG start codon to initiate translation. Upon start codon recognition 48S complex formation at the translational start site AUG is induced. As shown by the Sonenberg laboratory overexpression of a *trans*-dominant mutant of La reduces 48S complex formation during HCV and polioviurs IRES-mediated translation suggesting that La supports the recognition of the start site [67]. The hLa protein has been found to associate with 40S ribosomal fractions when analyzed by sucrose gradient centrifugation [249] [250] and speculated to facilitate the landing of the 43S preinitiation complex on mRNA [67] [250]. However, a direct interaction with the 80S subunit has not been shown. This implicates only a transient association of hLa with the small 40S ribosomal subunit and its release before the elongation competent 80S ribosome is assembled.

It has to be addressed if hLa is recruited to the mRNA as a 40S:La or 43S:La complex or if it interacts with mRNA independently of ribosomal complexes. The *in vitro* data presented herein clearly demonstrate cyclin D1 RNA binding by hLa in the absence of ribosomal complexes. However, these findings do not exclude the possibility of La association with ribosomal complexes during 5'-UTR scanning for the translational start site AUG codon in living cells. The mRNA is pulled through a channel within the 43 subunit during 5'-UTR scanning. This process may be impeded by secondary and tertiarty mRNA structures [251], which need to be unfolded for an efficient 5'-UTR scanning and start codon recognition.

RNA recognition by hLa is suggested to be shape-directed rather than sequence specific [113] [90]. Sequence-independent recognition of hLa to short double stranded stretches flanked by single stranded extensions has been demonstrated for 5'-NCR HCV recognition

[113]. Recent evidence suggests that RRM-containing hnRNP proteins bind double-stranded structured RNAs [252] [253] [254] [76] the binding of an RRM to double-stranded RNA cannot be explained [21], but this needs to be structurally confirmed.

In order to support the idea of structural element-driven recognition by hLa the secondary structures of La binding sites in XIAP, Mdm2, HCV, and CCND1 were predicted by the software *mfold* [255]. The predicted XIAP RNA forms a stem comprised of 10 base pairs, similarly to the CCND1 La binding site, and a loop of five single stranded nucleotides. The La binding site in the Mdm2 RNA forms a shorter seven base pair stem followed by a five-nucleotide loop and an additional short stem-loop. The D1-ATG RNA forms a stem of eleven base pairs. The AUG and the embedding Kozak consensus sequence are located in double stranded as well as single stranded stem-loop regions (figure 5.1). Similarly, the hLa binding site in the AUG start codon context of HCV is represented by five base pairs and a loop of seven nucleotides. The structure of HCV RNA has been experimentally confirmed, however, the 5'- and 3'-ends are forming a bulge in the context of the entire 5'-non coding region. The loop comprises the AUG and the entire Kozak consensus sequence. The requirement of a strong Kozak sequence has been discussed as an important feature of La binding sites. The HCV and CCND1 structures share two locations in the loop, +1 A of the start codon AUG and postion -1 from the Kozak consensus sequence. The other Kozak consensus sequence nucleotides are in different structural contexts in those two RNAs, as well as the +2 U of AUG, which has been shown to be redundant for La recognition. RRM-containing RNA-binding proteins (RBPs) commonly interact with single-stranded nucleic acids [256] [21] suggesting that the single-stranded -1 C and +1 A are preferentially recognized by hLa. Interestingly, the Mdm2 RNA also harbors the single stranded CA-motif (nts 8 and 9 in figure 5.1), which may represent the La recognition site.

In conclusion, these structural analyses suggest that La recognizes stem-loop structures and binds single stranded nucleotides located in looped and bulged areas. Nucleotide binding within loops has also been shown for other RRM-containing proteins, such as for the splicosomal protein U1A, which directly interacts with nucleotides in the 10-nt loop of its target U1 hairpin II [257] [258].

If the looped and bulge structures indeed exist in the mRNAs and endogenous La is binding those structures in living cells then an interesting possibility emerges. It has to be addressed if hLa refolds secondary and tertiary structures in mRNAs and if it facilitates the scanning and recognition of the AUG start codon by the 43S complex. This scenario would also explain why La was suggested to be required for efficient 48S subunit formation [67]. Intriguingly,

the hLa protein has been proposed to have RNA-chaperone activity [115]. This activity allows the protein to facilitate the proper folding of RNA in an ATP-independent manner. RNA-chaperones are considered to contain intrinsically disordered regions [117] [126]. Preliminary data from RNA-chaperone assays performed in the laboratory revealed that hLa is indeed able to destabilize structures in the D1-ATG RNA. This strongly supports speculation that hLa facilitates translation when the start codon AUG is embedded in a structural context, which impedes efficient AUG recognition by the 43S subunit. Therefore, the following model can be envisioned. The hLa protein binds a structure in the cyclin D1 mRNA. This structure comprises the translational start site. La then stimulates translation by refolding this region which would otherwise impede the recognition of the start site by the 43S preinitiation complex.
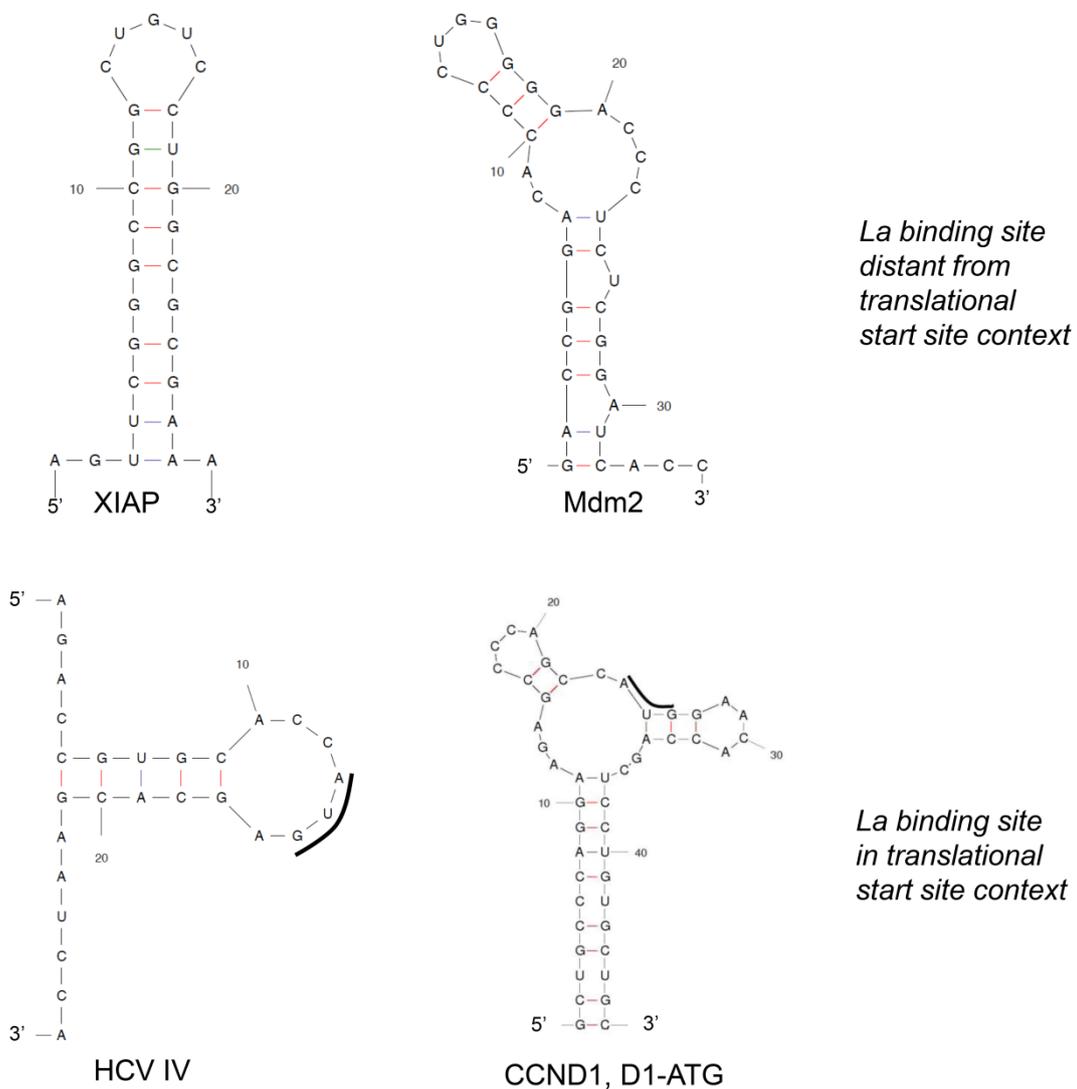
Figure 5.1: Secondary structure predictions of La binding sites in selected mRNAs by *mfold*. All four La binding site comprise stem-loop structures of varying lengths and sequences. The AUG start codon in the HCV IV and D1-ATG RNA are indicated by a black line. The thermodynamically most stable structures are displayed. For sequence information refer to the appendix.

5.2 Interaction of the human La protein with cyclin D1 mRNA

The human La protein is a modular RNA-binding protein containing three RNA-binding domains, the La motif, the RNA recognition motif 1 (RRM1), and the RNA recognition motif 2 (RRM2). The interaction of nuclear La with its main targets, RNA polymerase III transcripts and other RNAs ending in poly(U)-sequences, has been well studied and even a co-crystal structure of a small UUU-OH mimicking RNA bound to the N'-terminal domain (NTD) of the La protein has been solved [31]. The cytoplasmic La protein contributes to cap- and IRES-dependent translation, however, the La domains mediating the internal binding to those cellular mRNAs have not been identified yet.

In order to identify the RNA-binding domains of hLa a variety of La mutants were analyzed for their RNA-binding activity to the cyclin D1 RNA by two different *in vitro* techniques, EMSA and fluorescence polarization. Utilizing EMSA and FP recombinant hLaWT was shown to bind the entire CCND1 5'-UTR (D1-FL) with an affinity of K_D as 45 nM (figure 4.1.2D) and D1-ATG RNA (harboring the translational start site) with an affinity of approximately 80 nM (figure 4.1.7).

The RNA-binding affinity to the full length 5'-UTR is approximately 10-fold lower as the high affinity binding of the La protein to fragments of the poliovirus 5'-UTR, which has an affinity of 4 nM [98], and the HBV RNA ($K_D = 1$ nM) [112] [90]. The La protein also has a high affinity to RNA polymerase III transcripts, the dissociation constant are in the lower nanomolar range between 3-19 nM [40] [259]. The affinity of hLa for the 5'-UTR of cyclin D1 is lower compared to most other described hLa RNA targets. Preliminary data for the binding of 5'-TOP RNA oligoribonucleotides, L22 and L37, suggest a comparable affinity of $K_D \approx 100$ nM for those synthetic RNAs.

Although the affinities determined herein are lower compared to other described RNA targets of hLa, the loss of affinity is not within magnitudes, but still in the nanomolar range reflecting an affine binding [260]. The RNA-binding affinities were determined in different laboratories by varying techniques, which may have resulted in altered affinities. The preparation of recombinant hLa may also be crucial for comparable RNA-binding studies. The hLa preparation used herein is only 10-20% RNA-binding active, which may cause the observed decreased RNA-binding affinity of hLa, discussed later in 5.5.4. The affinity of hLa to cellular and viral mRNAs could be reduced, compared to terminal poly(U) sequences, because the recognition of internal La binding sites within secondary mRNA structures may

challenge hLa recognition. However, the RNA-binding affinity of hLa to a predicted stem loop in the HBV-RNA [112] [90] and the HCV 5'-NCR [98] are comparable to targets harboring terminal poly(U) sequences [40] [259] implying that hLa binds terminal poly(U) sequences with a similar affinity as structured regions.

Electrophoretic mobility shift and fluorescence polarization assays clearly show that the N'-terminal La motif (LaΔ1, aa 11-99) of the La protein is not required for the D1-ATG RNA-binding (figure 4.2.3). Since the RNA-binding affinity of LaΔ1, $K_D \approx 68$ nM, is comparable to the affinity of hLaWT the deletion of the La motif is not likely to induce a inducing disadvantage, structural changes for RNA-binding. This La motif-independent recognition has also been shown for HBV RNA-binding [35]. Intriguingly, this recognition mechanism differs from the interaction of La with RNA polymerase III transcripts, where the La motif is essential for RNA-recognition [31], as reported for 3'-terminal poly(U) motif binding in pre-tRNAs by the La protein [73]. In contrast, RRM1 and RRM2 are the main domains mediating hLa-binding of cyclin D1 RNA (D1-FL or D1-ATG (figures 4.2.4, 4.2.5). Hallmarks of RRMs are the RNP-1 and RNP-2 consensus sequences, which are directly mediating the interaction with the RNA typically by position 2 in the RNP-2, position 5 in RNP-1, and position 3 in RNP-1 [15] [261], furthermore they are located in the β1- and β3-strand, respectively (figure 5.2). The deletion of the RNP-2 in RRM2 resulted in a total loss of D1-ATG RNA-binding in EMSA studies with the LaΔ4 mutant (figure 4.2.3F). The deletion of the RNP-2 consensus

RNP1: [RK]-G-[FY]-[GA]-[FY]-[ILV]-X-[FY]

RNP2: [ILV]-[FY]-[ILV]-X-N-L

Figure 5.2: Schematic representation of the four-stranded β-sheet [21]. RNP-1 and RNP-3 consensus sequences of RRMs are displayed. The approximate location of the main conserved aromatic residues in RNP-1 and RNP-2 are indicated (circled). X = any amino acid residue

sequence in RRM1 (LaΔ2) did not abrogate D1-ATG RNA-binding in EMSAs; however, it reduced the complex formation with the RNA dramatically as seen in figure 4.2.3.F. Although the RNP-2 sequence in RRM1 is contributing substantially to RNA-binding, the FP-assays suggest that binding still occurs but the reaction kinetics are different. This assumption would explain the very weak complex formation in EMSA and the moderate decrease in affinity ($K_D \approx 223$ nM). Hence other parts of La may contribute to binding but

RNP-2 of RRM1 is absolutely required for La:D1-ATG RNA complex formation. Similarly, the LaΔ4:D1-ATG RNA complexes are almost invisible compared to a less dramatic affect on the RNA-binding affinity dropped ($K_D \approx 303$ nM), again suggesting that RNP-2 of RRM2 is absolutely required for stable La:D1-ATG RNA complex formation.

However, the dramatic reduction or loss of D1-ATG RNA-binding in the EMSA is not as clearly reflected by the dissociation constants, which were rather expected to be in the high nanomolar to low micromolar range, as shown for the RNA-binding activity of C'-terminal deletion mutants of the RNA-binding protein AUF1 to c-*fos* A+U-rich element RNA [262]. Therefore, suggesting other factors are involved in RNA-binding, such as the RNP-1 consensus sequences. Preliminary data suggest that the RNP-1 in RRM2 is involved in D1-FL RNA binding (data not shown).

These findings strongly suggest that both RRMs must be functionally active, since the loss of one RNP consensus sequence is not compensated by the other RNP sequences. The RRMs are likely acting cooperatively to establish a stable La:D1-ATG RNA complex as recently postulated for HBV RNA [35] and HCV RNA binding [113].

The RNA binding mechanisms for RBPs with more than one RRM have been shown for various proteins (reviewed in [21], such as poly(A)-binding protein (PABP; [263], nucleolin [264] [265], and DNA, e.g. hnRNP A1 [256] to facilitate higher binding affinity and specificity [21]. Structural studies of the interaction of the ATP-dependent helicase-related protein 1 (Hrp1) with RNA revealed that two RRMs are arranged in an almost parallel orientation to each other, which accommodates the single stranded RNA molecule shown in figure 5.3 [266]. The RRM-RRM interaction is stabilized by a salt bridge and hydrogen bonds between the inter-RRM peptide linker and the RRM2 [266]. This superposition of RRMs in complex with RNAs has also been shown for the proteins Sex-lethal (Sxl) [267] and HuD [260] and is shown in figure 5.3. The hydrogen bond formation between the linker and the RRM2 were observed for the approximately 58 distant amino acids arginine 202 and aspartic acid residue 206 for Sxl and arginine 116 and aspartatic acid residue 174 for HuD [266]. Intriguingly, the arginine 196 located in the hLa inter-RRM and aspartic acid 255 in the RRM2 are 59 amino acids apart, similar to the 58 amino acid spanning hydrogen bond in Sxl and HuD [266], suggesting that a hydrogen bond between those two residues is formed in hLa. The importance of the hLa inter-RRM peptide linker was analyzed by EMSA.
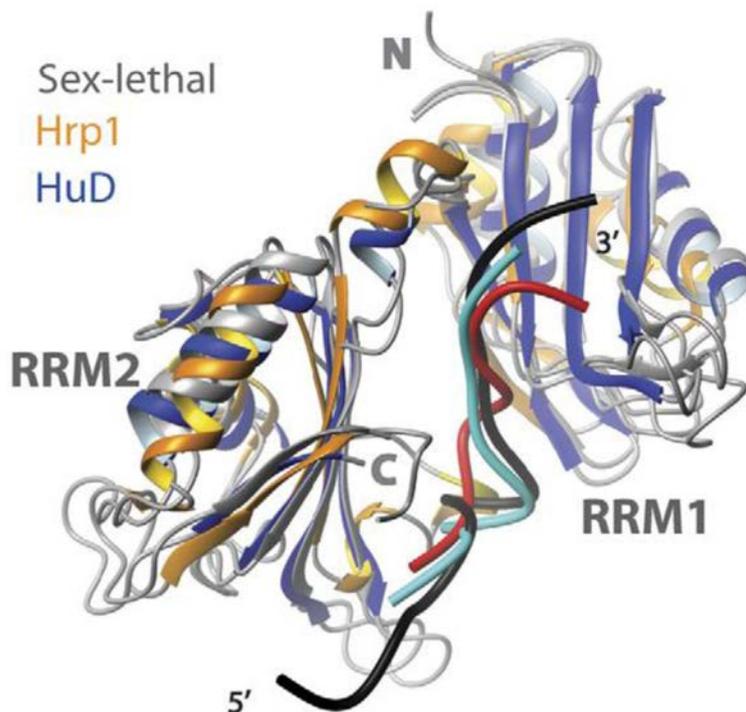
Figure 5.3: Superposition of protein structures of Sex-lethal, Hrp1, and HuD associated with their ssRNA targets [21]. The RRM ribbons of Sex-lethal, Hrp1, and HuD are indicated in grey, yellow, and blue, respectively. Their ssRNA targets are respectively represented in black, red, and cyan.

A hLa protein mutant, in which the linker region was deleted (60 aa between aa 170-230), did not form a La-RNP complex with D1-ATG RNA (appendix figure A2). Those preliminary data imply that the inter-RRM linker is required for the conformational position of the tandem RRMs, presumably similar to the superposition of Sxl and HuD [266], thus are required for RNA binding.

This hypothesis is supported by the *in silico* analysis of the RRM1+2 protein (amino acids 114-326), which revealed that the linker region is mainly intrinsically disordered towards the RRM2 (figure 5.4). Due to this lack of a tertiary structure the inter-RRM linker region is flexible and may cause a structural change of the La protein to assist the superposition of the RRM1 and RRM2 for efficient RNA-binding.

The RNA-binding affinity of the minimal La protein RRM1+2 is approximately 30% lower compared to the hLaWT as reflected by the higher K_D of 111 nM (figure 4.2.4B) suggesting that the N'-terminal and C'-terminal amino acids are not essential for D1-ATG binding, but are contributing directly to RNA-binding or indirectly by modulating the protein structure. Intermolecular interactions between the RRMs and other domains within the same protein are proposed to affect the specificity of the RRM-RNA interaction [21]. Structural analysis of the hLa NTD (La motif and RRM1) bound to RNA revealed that the La motif interacts broadly

with the RNA, which resulted in merely a few established non-canonical contacts between RRM1 and the RNA [76] [31].

The C'-terminal region of the La protein is suggested to play an important role in RNA-binding [40] [268] [269] [218]. The C'-terminus is mostly intrinsically disordered [224] and punctuated by a short basic motif. The C'-terminal amino acids (aa 327-408), as well as the basic and aromatic amino acids between aa 332-339 and aa 357-364 (table 4.2.1) may interact with RNA.
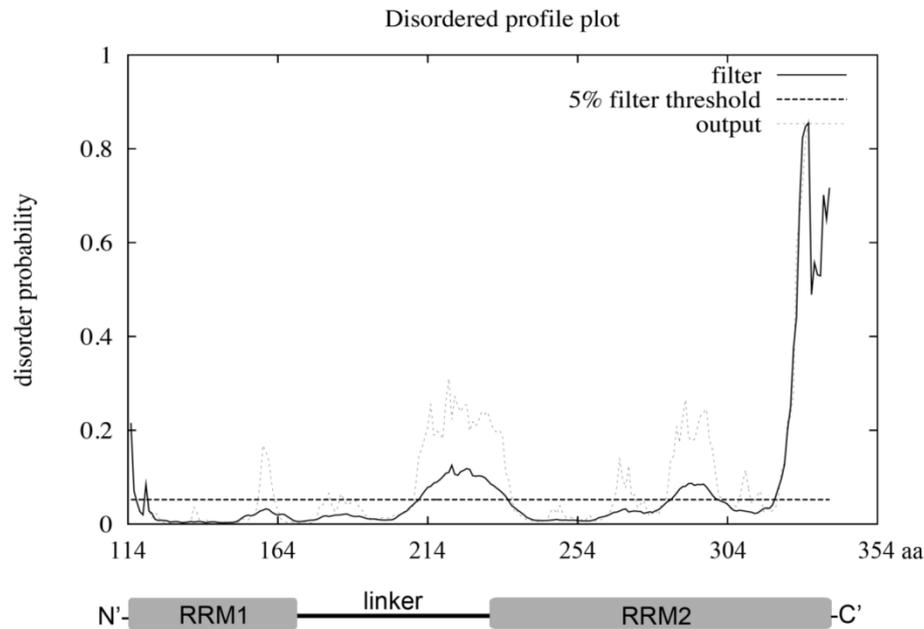
Figure 5.4: Disordered profile plot of the minimal hLa protein mutant RRM1+2. The FASTA sequence of RRM1, RRM2, and inter-RRM peptide linker (aa 114-326) was submitted to the DISOPRED2 server. The probability of disorder is plotted as a function of the amino acid sequence. A schematic representation of the minimal RRM1+2 protein is displayed as a reference below the plot. The black filter curve is the DISOPRED2 output for the disorder predictions, while the dotted output line represents the lower confidence disorder predictions. The algorithm's default fault positive rate threshold of 5% is indicated by the horizontal line. (http://cms.cs.ucl.ac.uk/ fileadmin/bioinf/Disopred/disopredhelp.html).

Whereas the RNA-binding affinity of RCD1, carrying basic and aromatic amino acid substitutions (aa 332-339), is similar to the wild type dissociation constant, the mutation of basic and aromatic amino acids in RCD2 (aa 357-364) resulted in a dramatically higher RNA-binding affinity of $K_D \approx 15$ nM. This is similar to the described affinities of La to RNA polymerase III transcripts. The dramatic reduction of the dissociation constant of RCD2 suggests that the basic and aromatic amino acids between aa 357 through 364 are impeding RNA-binding. It is possible that hydrogen bonds or salt bridges, as shown for the interaction of RRM1 and RRM2 of Hrp1 with its substrate [266], are formed between the basic and aromatic amino acid stretch (357-364) and the RRM2. Interestingly, the amino acid stretch

between aa 357-364 comprises three lysines (K360, K361, K363) which may form a salt bridge with one of the seven aspartic acid residues in RRM2. A cluster of three aspartic acid residues (D239, D241, D242), are surrounding the RNP-2 consensus sequence (aa 235-240) in RRM2. A formation of salt bridges may not impede binding, but slightly reduce the RNA-binding activity by covering the RRM2.

Another explanation for the reduced RNA-binding activity, as a result of basic and aromatic stretches, may result from the lack of posttranslational modification in the close vicinity to the amino acid stretch between aa 357-364. It has been shown that phosphorylation can perturb the energy landscape of a protein resulting in conformational changes to the protein [270] [271]. Human La is phosphorylated at threonine 362 (*in silico* prediction) [272], threonine 389 by Akt [42], and serine 366 by casein kinase II (CK2) [37] [58]. The phosphorylation sites for threonine 362 and serine 366 are located in or in close proximity of the RNA-binding inhibitory region, located between aa 357-364. Hence, the posttranslational modification of either or both residues may induce a structural change in this region resulting in the inhibition of the negative RNA-binding effects mediated by the basic and aromatic amino acids (aa 357-364) located in this intrinsically disordered region. Additional posttranslational modifications sites may also regulate the inhibitory effects of aa 357-364, lysine acetylation sites K320 and K360 [38], and sumoylation sites K344, K352, K363, and K400 are located in the CTD.

The minimal La protein RRM1+2 is RNA-binding active, as in the case of cyclin D1 RNA, making it a potential candidate for structural analyses, e.g. by NMR, which would lead to understanding the direct interaction of the binding domains as well as their interpeptide linker for RNA-binding.

In conclusion the following model for cellular mRNA binding by hLa is suggested. The RRM1 and RRM2 are positioned in a parallel and planar configuration creating an accessible RNA-binding surface which allows for the cooperative binding of D1-ATG RNA. This structural arrangement is possible because of the flexible, intrinsically disordered inter-RRM linker.
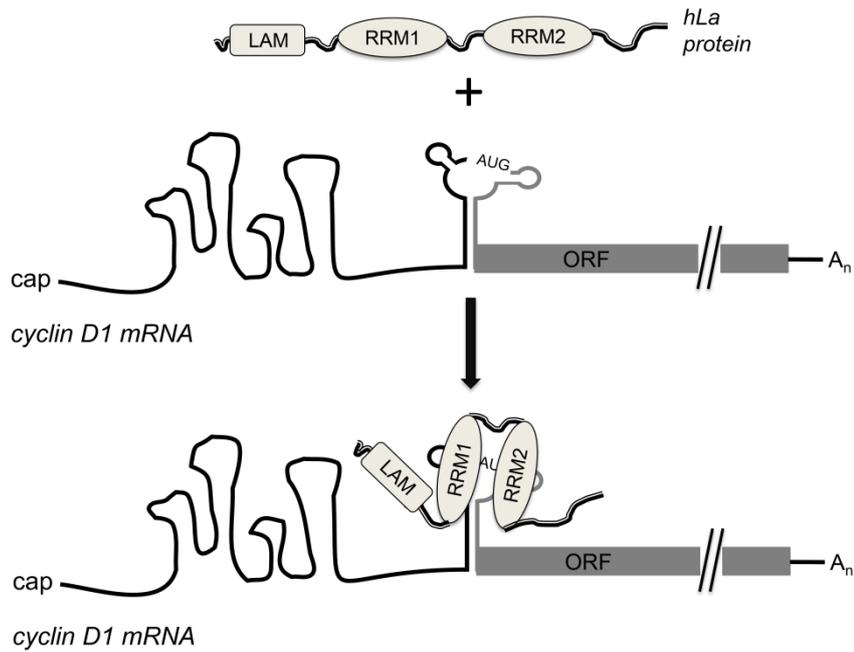
Figure 5.5: Model of La binding to the cyclin D1 mRNA. The hLa protein recognizes the translational start site context, which is located in a stem-loop structure, and binds the RNA via its RRM1 and RRM2.

5.3 "Regulating the regulator": SUMO-modification enhances the RNA-binding activity of La to translational regulatory elements

5.3.1 Characterization and identification of SUMO-acceptor sites *in vitro*

Posttranslational modifications (PTM) of proteins are critical mechanisms for regulating protein activity. In recent years, several studies reported the sumoylation of RNA-binding proteins [201] [202] and other proteins involved in RNA-processing [210] [212]. SUMO-modification of the RNA-binding protein hnRNP C1 has been shown to impair its binding to single stranded DNA [201]. Furthermore, a negative functional alteration of the RNA-editing activity upon sumoylation of ADAR1 has been reported by Desterro *et al*. [210]. Interestingly, the sumoylation of the RNA-binding protein Sam68 does not alter its RNA-binding activity, but has been shown to enhance its repressive effects on cyclin D1 expression [202] possibly due to alteration of Sam68 affinity for transcriptional coactivators such as acetyltransferase CBP [273] [274]. Sumoylation of proteins involved in posttranscriptional gene expression processes have been shown to positively alter their function, e.g. sumoylation of the translation initiation factor eIF4E activates mRNA translation [211] and SUMO-modification of 3'-end mRNA processing factors symplekin and CPSF-73 enhances

the activity of the 3'-mRNA processing complex [212]. SUMO-modification of eIF4E is proposed to modulate eIF4E protein-protein interactions, because sumoylation is suggested to induce a conformational change altering the interaction surfaces and resulting in an increased binding affinity for other proteins. eIF4E is another factor involved in CCND1 RNA expression by promoting the nuclear export of CCND1 mRNA [275]. It has been shown that SUMO-conjugation can alter the subcellular localization of their targets, e.g. unmodified RanGAP1 is localized to the cytoplasm [276], and it also alters protein-protein interactions, e.g. sumoylated RanGAP1 is tightly binding to the RanBP2/Nup358 nuclear pore complex [277] [278].

In conclusion, RNA-binding proteins are reported to be SUMO-modified and affected differently upon this posttranslational modification. Although the effects of sumoylation on the RNA-binding activity are not well studied, sumoylation has been shown to positively influence mRNA processing and translation suggesting an important involvement of sumoylation on protein expression.

The human La protein has been described to be modified both SUMO-2 as in HeLa cells upon heat shock or by SUMO-1 and SUMO-2 as in rat neurons [60]. In those sensory rat neurons sumoylation of La triggered the retrograde transport of La from the axons to the neuronal cell body [60].

A hLa-specific *in vitro* sumoylation assay (IVSA) was established to study if sumoylation has an effect beyond motor protein transport in axons, and on RNA-binding activities of the protein.

The critical parameters for an optimal La-specific IVSA were identified as molar Ubc9:SUMO ratios between 1:9 and 1:37 (figure 4.3.3A), an incubation time of two hours (figure 4.3.3B), and an incubation temperature of 30°C (figure 4.3.3B). Whereas a parallel reaction without the addition of ATP is a solid control (figure 4.3.4) the more appropriate negative sumoylation control was a parallel reaction with the conjugation deficient SUMO-GA mutant (figure 4.3.2).

The observation of two bands at approximately 67.5 kDa and 77 kDa represented SUMO-conjugated hLa protein, thus suggesting sumoylation at two sites. Those two bands were not only La-specific but also SUMO-specific, as detected by a SUMO-specific antibody (figure 4.3.7), further suggesting the bands are SUMO-modified hLa.

In some instances a larger third sumoylation band was detected implying sumoylation at a third site (figure 4.3.7). Interestingly, using SUMO-2 and SUMO-3 as modifiers resulted in SUMO-conjugation of hLa at only one site (figure 4.3.7) compared to SUMO-1 modification of at least two La lysine residues. The SUMO-2 and SUMMO-3 peptides contain lysine 11 as SUMO-acceptor site, which would allow for the formation of poly-SUMO chains [168], however, the sumoylation of one site by SUMO-2/-3 and two sites by SUMO-1 suggests that SUMO-2/-3 are recognizing only one site and not forming poly-SUMO chains on hLa *in vitro*. This result led to investigating specific sumoylation sites.

Therefore, to identify the SUMO-acceptor sites in hLa a number of La mutants with mutations in potential and confirmed sumoylation sites were tested *in vitro* (figure 4.3.8) However lysine 41, which was identified as SUMO-acceptor site in rat dendritic cells [60], and the other suggested lysines were not modified *in vitro* (figure 4.3.8A). These findings imply that recognition of sumoylation sites in cultured cells differs from *in vitro* recognition, further suggesting a cell-type specific sumoylation of La. To address the difference between *in vitro* and *in vivo* SUMO-acceptor site recognition SUMO-sites of hLa in different cell systems has to be identified and compared to already established *in vitro* identified SUMO-acceptor sites of recombinant hLa.

The identification of sumoylation sites was initially attempted by *in silico* analysis, however, the *in silico* predictions resulted in a large number of possible SUMO-sites so an alternative method was chosen to identify the sumoylation sites within hLa.

As an alternative, mass spectral analysis was used to identify the sumoylation sites in recombinant hLa. A collaboration with Dr. P. Thibault identified lysine 200 and lysine 208 as the main sumoylation sites and lysines K344, K352, K363, and K400 as minor sumoylation sites. The main SMO-acceptor sites were confirmed by mutagenesis in the context of the minimal RNA-binding competent RRM1+2 La protein. The *in vitro* sumoylation of the minimal La protein RRM1+2 resulted in two sumoylation events (figure 4.3.10).

Numerous proteins can be either modified by SUMO-1 and SUMO-2/-3 [279], however, some substrates, e.g. RanGAP1 and Sp100, are preferentially modified by either SUMO-1 or SUMO-2/-3, respectively [280] [279] [172]. Although the hLa protein is SUMO-conjugated by all three SUMO paralogs the observed efficient sumoylation of two sumoylation sites only by SUMO-1 suggests a preferential modification by SUMO-1 *in vitro*. However, the modification of hLa by SUMO1 and SUMO-2 has been shown in rat dorsal root ganglions

[60] and by SUMO-2 in U2OS cells upon heat shock [61]. These data suggest a complex modification of human La by SUMO in response to different cellular conditions (e.g. heat shock), in different cell types, and *in vitro*.

The identified *in vitro* sumoylation sites differ from the SUMO-acceptor lysine 41 identified by mutational analyses of overexpressed GFP-tagged hLa protein in rat neuronal cells. Sumoylation in cultured cells likely differs from *in vitro* sumoylation because other factors that regulate target specific sumoylation are absent *in vitro*. The interaction between sumoylation and other posttranslational modifications (PTMs) has been described [281], e.g. sumoylation can act antagonistically or synergistically with ubiquitination [177] and SUMO-modification has been shown to compete with phosphorylation [282]. Therefore, it is possible that the hLa protein is SUMO-modified at a different lysine because of the cross-talk with other PTMs. Those PTMs may induce either structural change within the protein, which may block the canonical sumoylation sites, or trigger a signal for the sumoylation apparatus resulting in a non-modified target. This aspect may be of critical importance in respect to the minor SUMO-acceptors sites (K344, K352, K363, and K400), which are located in close proximity to reported phosphorylation (threonine 362 and 389, serine 366) [272] [42] [37] and acetylation sites (K320 and K360) [38] in the CTD of hLa.

Because hLa was modified by SUMO-1 in the distal parts of axons in neuronal cells, hLa may be SUMO-modified in the cytoplasm. In context of the regulation, it needs to be addressed in which cellular compartments hLa gets modified and demodified. The localization of the E3 ligase RanBP2 on the cytoplasmic and SENP2 on the nuceloplasmic side of the nuclear pore complex suggests a rapid sumoylation-desumoylation upon nuclear import of proteins [277] [283] [284] However, sumoylation of the ets-related transcriptional repressor TEL and *Dictyostelium* MEK1 triggered nuclear export of those targets [285] [286].

It can be concluded that the identification of lysines 200 and 208 as SUMO-acceptor sites supports the hypothesis that sumoylation modulates the RNA-binding activity of hLa because the sumoylation sites are located within the 60 amino acid inter-RRM linker region (aa 170-230) between the RNA-binding mediating domains RRM1 (aa 112-169) and RRM2 (231-327).
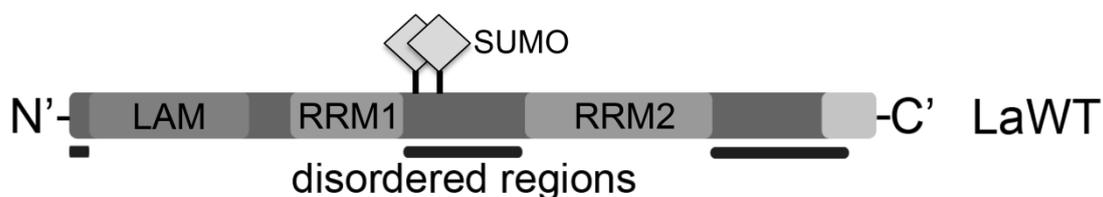
Figure 5.6: Recombinant human La is modified at lysine residues 200 and 208. Schematic representation of hLa sumoylation sites located in the peptide linker between RRM1 and RRM2.

5.3.2 SUMO enhances the RNA-binding activity of hLa *in vitro*

The location of La sumoylation sites, lysines 200 and 208, in the inter-RRM linker region and in close proximity to the internal mRNA-binding domains RRM1 and RRM2 strongly suggests that sumoylation at those sites may modulate the RNA-binding activity of hLa. This has been experimentally shown and explained in the following section.

The RNA-binding activity of hLa was significantly stimulated upon SUMO-modification (figure 4.3.12B and D). The SUMO-modification of hLaWT and wild type RRM1+2 proteins had positive effects on the RNA-binding activity to cyclin D1 RNA. The stimulatory effect was decreased but not abrogated upon abolishing the K208 sumoylation site suggesting that the sumoylation at lysine 200 alone cannot stimulate the RNA-binding activity of hLa to the same extend as sumoylation at both lysines.

There was a marginal increase the RNA-binding activity by 25% compared to a 50% stimulation by the sumoylated wild type RRM1+2 represented in the S-2:RRM1+2 K208R which approximately equaled the intensity of both S-2:RRM1+2 bands together (figure 4.3.12C). These findings suggest that the sumoylation of both lysines, K200 and K208, is required for efficient stimulation in RNA-binding activity of the hLa protein.

The RNA-binding activity was reduced to normal wild type levels upon mutation of both sumoylation sites, although sumoylated hLa species were detected in some IVSAs (figure 4.3.12D). These observations conclude that additional SUMO acceptor sites exist, as described earlier, but are not modulating the RNA-binding activity to a detectable level. Therefore, the SUMO-modification at the main SUMO-acceptor sites, lysines 200 and 208, stimulate the D1-FL RNA-binding activity of hLa.

As discussed earlier the inter-RRM linker is critical for hLa RNA-binding. This is possibly due to the inter-RRM linker enabling a structural conformation of the RRMs optimizing them for RNA-binding. The SUMO-modification of the intrinsically disordered inter-RRM peptide linker may support proper inter-RRM-linker induced positioning of the tandem RRMs.

It is shown herein (figure 4.3.11) that the accessibility of SUMO-acceptor sites might be affected by prior RNA-binding to hLa. The binding of the RNA to hLa prior to IVSA virtually abrogated SUMO-modification. Many SUMO-consensus sites, such as K208, are located in intrinsically disordered regions or extended loops [163]. The hLa protein is proposed to bind RNA synergistically by its RRM1 and RRM2, and forms multimeric ribonucleoprotein complexes at specific hLa protein concentrations. The multimerization of hLa upon RNA-binding may induce a structural conformation resulting in shielding of the SUMO-acceptor lysines in the interdomain peptide linker from SUMO-modification by Ubc9.

The addition of the unspecific competitor heparin restored sumoylation (figure 4.3.11B), which implies accessibility of the SUMO-acceptor sites in monomeric RNA-bound hLa (figure 4.3.11). This emphasizes the importance of establishing functional significance of hLa multimerization *in vitro* and *in vivo*.

Although much work is needed to illuminate how the RNA-binding activity of hLa is regulated upon sumoylation, it is apparent from the presented data that this may represent an important regulatory mechanism.

The literature hints to a stress-induced modification of target proteins by SUMO-2/-3. There is little free SUMO-1 available under resting conditions, whereas there is a pool of free SUMO-2/-3, which gets increasingly conjugated during oxidative, genotoxic, osmotic, ethanol and mild heat stress [280]. This stress-dependent response of increased SUMO-2/-3 conjugation was detected in various cell systems [287] [288] [289] suggests that sumoylation acts as a protective stress-response. However, substrates are modified at different levels upon stress, some substrates get less sumoylated, however, an overall net increase of sumoylation was observed upon stress-induction [61] [290].

The *in vitro* sumoylation of hLa with SUMO-2/-3 appears less efficient than the conjugation of SUMO-1. Further, SUMO-2 modification of La has been described upon inducing heat stress in HeLa cells [61]. Those two findings hint to stress-induced SUMO-2/-3 conjugation of hLa signaled externally or internally. Under these stress conditions cap-dependent translation is impaired and IRES-mediated translation is activated. This correlation supports the following model. In normal cells hLa is mainly modified by SUMO-1. Sumoylation induces structuring of the inter-RRM peptide linker, which then promotes optimal positioning of RRM1 and RRM2 facilitating the binding of cyclin D1 mRNA in close proximity to the translational start site. The proposed RNA-chaperone activity of hLa restructures the RNA, which then supports translation initiation resulting in cyclin D1 expression. However, in

cancerous and stressed cells the sumoylation pathway is activated, which leads to modification of hLa by SUMO-2/3. This modification further stabilizes the positioning of the tandem RRMs thus enhancing the RNA-binding activity of hLa further. The translation rate of the cooperative oncogene cyclin D1 is augmented because sumoylated hLa facilitates translational initiation.

5.3.1 SUMO-modification enhances the RNA-binding activity of hLa to 5'-TOP RNAs

The enhancement of RNA-binding activity on a protein upon sumoylation has not been previously reported, however it is demonstrated herein for the cyclin D1 RNA-binding activity by sumoylated hLa. The effect of sumoylation of hLa on its RNA-binding activity was analyzed for RNA oligoribonucleotides representing the 5'-terminal oligopyrimidine tract (5'-TOP) element of RNAs encoding rat ribosomal proteins L37 and L22.

The La protein has is known to be involved in the translational regulation of mRNAs containing translational regulatory 5'-TOP motifs [241] [104] [242] [93], which encode for ribosomal proteins and other factors involved in the translational machinery [243] [291]. The translation of 5'-TOP mRNAs is stimulated by growth or mitogenic stimuli via both the mTOR and PI3-kinase pathway [292] [293] resulting in the expression of ribosomal proteins and ultimately in the activation of protein synthesis and cell growth [293]. A hallmark of TOP mRNAs is a sequence of cytosine followed by 4-14 pyrimidines in the context of an unstructured 5'-UTR, which is in close proximity to the 5'-cap [294] [295] [291] [244] [105] [58].

Fluorescence EMSAs were used to study the RNA-binding activity of native hLa to 5'-TOP representing oligoribonucleotides (oligo) followed by RNA-binding studies with sumoylated La.

The RNA-binding activity of native La to 5'-TOP RNAs was similar to cyclin D1 RNA-binding. However, La binding to the unrelated GAP43 RNA oligoribonucleotide representing the 5'-terminus of the mRNA encoding for the growth associated protein 43 did not result in saturation. The preferential binding of the L37 and L22 RNA oligos coincides with high pyrimidine abundance in the 5'-terminal 15 nucleotides of the L22 (86%) and L37 (79%) RNAs compare to GAP43 (53%) shown in figure 5.7. Although the overall pyrimidine abundance in the RNA oligoribonucleotides is similar, this suggests that hLa specifically recognizes and binds RNAs with a high abundance of pyrimidines in their 5'-terminus.
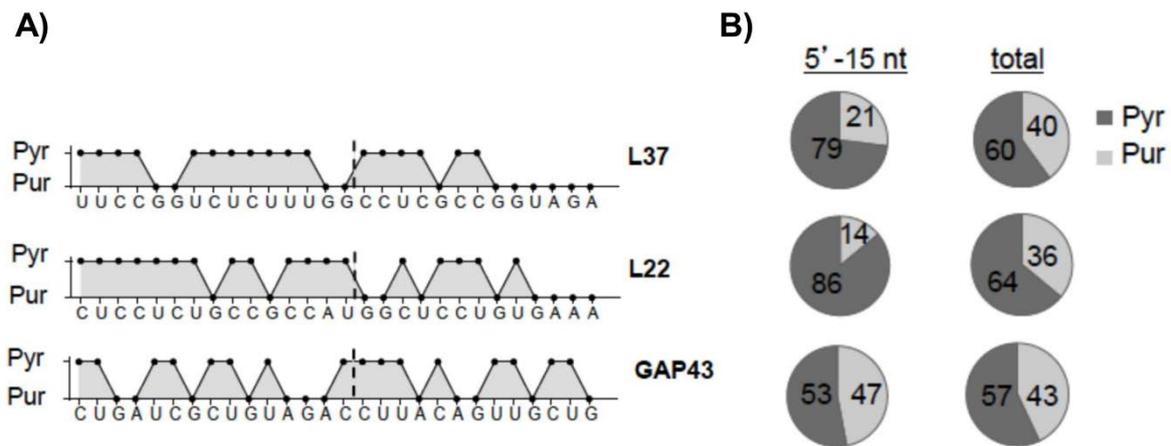
Figure 5.7 Pyrimidine abundance analyses of the RNA oligoribonucleotides used in EMSA studies. L22 and L37 RNAs display a high abundance of terminal pyrimidines. **B)** Quantification of pyrimidine abundance. The overall pyrimidine abundance is similar in all three RNAs, but differs in the 5'-terminal nucleotides. Pyr = pyrimidine residues (cytosine and uracil), Pur = purine residies (adenine and guanine)

Sequence and structure driven recognition of internal RNAs by hLa, as discussed earlier, suggested an important role for the location of the start codon AUG embedded in the Kozak consensus sequence (partially) in a loop. (refer to 5.1).

Secondary structure analyses by *mfold* revealed that all three predicted RNA structures, L22, L37, and GAP43, comprise stem-loops, however, the stem and loop lengths differ as well as single and double stranded nucleotide amounts. Note, L22 and L37 do not share similar structural features (figure 5.8). The high pyrimidine content within the first 15 nts of L22 and L37, not structural aspects, could explain the higher RNA-binding affinity of La to L22 and L37 compared to GAP43. Additional work would be required to understand the foundation of this preferential La to L22 and L37 binding.

It is unknown whether the 5'-TOP RNA-binding mechanism is similar to D1-FL RNA- or D1-ATG RNA-binding, therefore 5'-TOP mRNA-binding may be mediated by the synergistic interaction of RRM1 and RRM2 or by the La motif. However, the binding of hLa to 5'-TOP RNAs was strongly enhanced upon SUMO-1 modification of hLa (figure 4.3.13) implying a similar RNA-binding mechanism. La sumoylation resulted in a significant 2.5- to 3-fold increase in L37 RNA-binding activity compared to an approximately 1.5-fold of L22 RNA-binding activity upon La sumoylation. The higher increase in L37 RNA-binding activity upon sumoylation may be a result of the already increased RNA-binding affinity of native La to L37. In conclusion, from RNA-binding studies using different RNA targets and native and sumoylated La it appears that sumoylation of hLa facilitates its RNA-binding efficiency.

Figure 5.8: Secondary structure predictions of RNA oligos L22, L37, and GAP43 by *mfold*. For sequence information and accession numbers refer to section 2.11.

Biochemical evidences indicates the binding of La to 5'-TOP mRNAs. The La protein has been shown to co-sediment with 5'-TOP mRNAs in polysomes [93] [104] [59], has been reported to bind synthetic oligoribonucleotides representing 5'-TOP motifs in competition studies [296], and *Xenopus laevis* La has been demonstrated to associate with 5'-TOP motif containing mRNAs [297] [93] [104] [59]. However, whereas *X. laevis* La is proposed to have a stimulatory effect on translation [93] [104] [59], a negative effect of La on the translation of the elongation factor 1A (EF1A) was shown [242]. Those studies were all done *in vitro* lacking functional analyses to prove the role of the La protein in the regulation of 5'-TOP RNAs. Polysome distribution profiles of La-overexpressing human embryonic kidney cells revealed that non-phosphorylated La is associated with L37 mRNA [58]. Interestingly, yeast mRNAs lack a canonical 5'-TOP motif, however, the yeast La homolog Lhp1p has been shown to be associated with ribosomal protein mRNAs [74] suggesting that La plays a role as a *trans*-acting factor in the translational regulation of 5'-TOP mRNAs.

It would be important to investigate La-mediated stimulation of 5'-TOP mRNA translation, because the tumor suppressor protein p53 is regulated by those mRNAs.

It has been shown that ribosomal proteins L5, L11, and L23 bind Mdm2, which is the negative regulator of p53, thus preventing binding of Mdm2 to p53 and the Mdm2-mediated

degradation of p53 [298] [299]. The role of hLa in this process would suggest an important cellular function.

5.4 The brave new world: The role of cytoplasmic La in cell proliferation
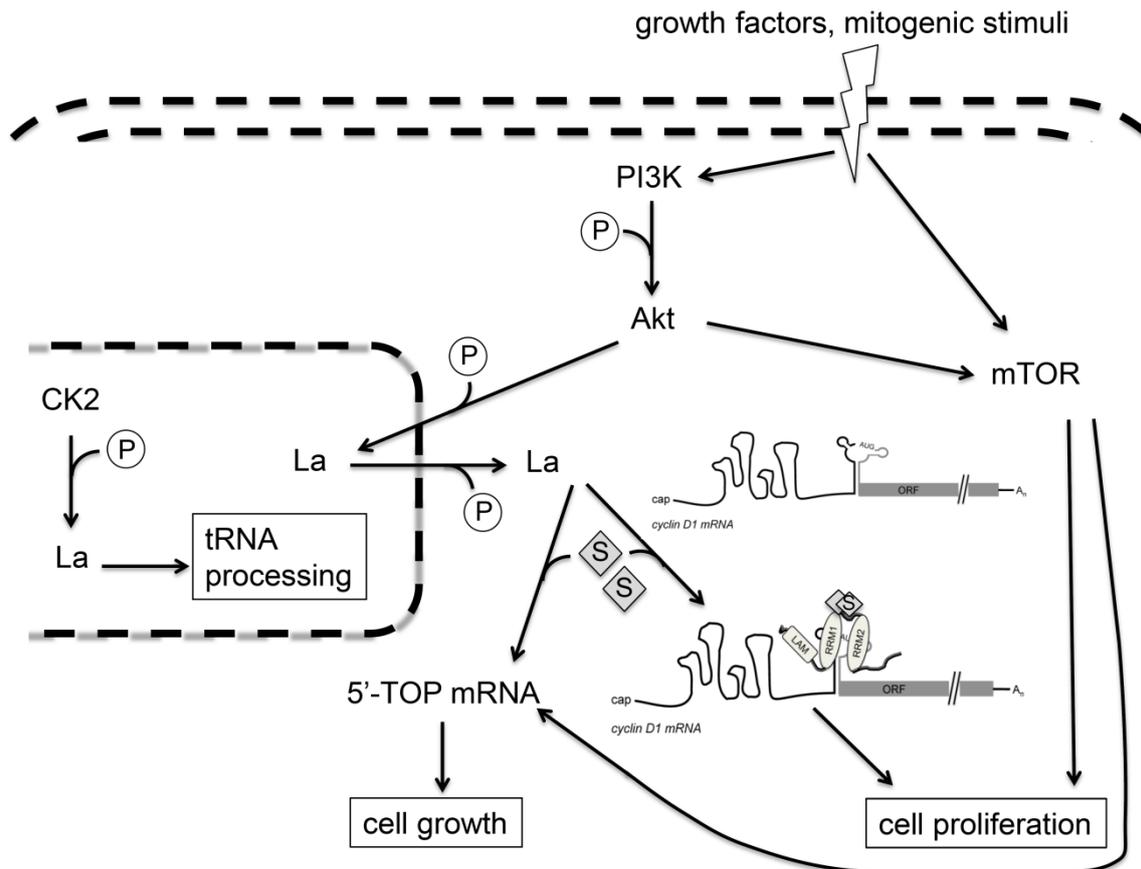
Figure 5.9: Cytoplasmic La contributes to oncogenesis by stimulating cell growth and cell proliferation. Growth factors ignite a signaling cascade by activating mTOR and PI3K. Phosphorylation of La at threonine 389 by activated Akt downstream of PI3K mediates nuclear export. The cytoplasmic hLa is de-phosphorylated which triggers SUMO-modification of hLa. The hLa sumoylation enabled the enhanced RNA-binding activity to 5'-TOP mRNAs and cyclin D1 mRNA by inducing an optimal structural change in hLa. La recognizes the translational start site context located in a stem-loop structure. The C'-terminal located RNA-chaperone domain opens the stem structure, allowing the stalled 43S preinitiation complex to scan the mRNA for the translational start site to initiate translation. Cyclin D1 IRES-dependent translation occurred in an Akt-dependent manner, by which the La protein was specifically activated by Akt. Phosphorylation of La by CK2 on the other hand stimulates the function of nuclear La in tRNA processing.

The aim of this study was to unravel the mechanism by which the hLa protein is stimulating cell proliferation. La may promote cell proliferation by upregulating the translation of cyclin D1. Human La has been shown to regulate the expression of this cooperative oncogene [89], which is often overexpressed in cancerous cells [129] [131] [300]. The aberrant expression of cyclin D1 is not only regulated on the transcriptional level [301] [302] [303] but also by its gene amplification [128] [304] and mRNA translation. IRES-mediated translation of CCND1 has been shown to be Akt-dependent [146], proposing that factors involved in IRES-mediated translation, such as ITAFs [247], are regulated by Akt-phosphorylation. Human La [89] and hnRNP A1 [213] have been identified as ITAFs regulating IRES-mediated translation of cyclin D1. Phosphorylation of hnRNP A1 by Akt inactivates ITAF functions [213], however, molecular and regulatory mechanisms by which hLa influences cyclin D1 translation are unknown. Findings presented in this study illuminate some important aspects. Sumoylation of hLa at lysines 200 and 208 facilitated hLa binding to the translational start site context of cyclin D1 *in vitro*. Therefore, the following is supporting evidence in how molecular and regulatory mechanisms stimulate hLa-mediated cyclin D1 translation.

The hLa protein is mainly localized in the nucleus, but also shuttles to the cytoplasm [53]. It is unclear to which extend phosphorylation by CK2 [58] and Akt [42] are involved in this subcellular localization, although the literature strongly suggests a critical involvement of phosphorylation in the intracellular distribution of hLa. Because Akt activity is often increased in cancerous cells [305] an Akt-dependent phosphorylation may trigger the localization of hLa to the cytoplasm, assuming that hLa behaves similarly to murine La [41]. The result of higher hLa abundance in the cytoplasm supports the speculation of hLa promoting IRES-dependent translation of tumor promoting factors such as cyclin D1, XIAP, and laminin B1.

The human La protein is not only regulated by Akt- and CK2-phosphorylation, but also by SUMO-modification [60] [61], which promotes the RNA-binding activity of hLa. The sumoylation pathway is involved in tumorigenesis [306] [193], because it is involved in tumor promoting pathways such as cell growth, differentiation, senescence, apoptosis, and autophagy [194] [195]. SUMO-2/-3 peptides are conjugated from a free pool to their targets induced by a variety of stresses [280]. It has been shown that in some cases phosphorylation co-regulates SUMO-modification [307] [308] [309]. Therefore, it may be speculated that hLa is modified by SUMO-2/-3 in cancerous cells, which results in a higher binding activity and ultimately facilitates IRES-mediated translation of tumor promoting factors.

The hLa protein may be contributing to the regulation of cell growth, as it also regulates cyclin D1 as part of the cell cycle machinery, with both cell growth and division ultimately leading to cell proliferation. It is controversially discussed in the literature whether hLa stimulates the translation of 5'-TOP mRNAs [93] [104] [59] [242] [310], however, it has been shown that hLa is not associated with 5'-TOP mRNAs when CK2-mediated phosphorylation is inhibited in living cells [58]. The sumoylation-dependent promotion of hLa 5'-TOP mRNAs binding implies a correlation between sumoylation and the expression of ribosomal proteins and elongation factors during cell growth.

According to the model, summarized in figure 5.9, hLa is proposed to stimulate cell proliferation and cell growth by stimulating the translation of cyclin D1 and 5'-TOP mRNAs in response to Akt phosphorylation and SUMO-modification.

5.5 Discussion of technical observations, problems, and solutions

5.5.1 Comparison of EMSA and fluorescence polarization assay

The difference in La:D1-ATG RNA interactions determined by EMSA and FP assay are in the methodology of the assays. La-RNP complexes may not be detected in EMSAs due to the rapid dissociation of those complexes during electrophoresis. On the other hand, a slow dissociation may result in data misinterpretation by underestimating the intensity of the La:D1-ATG RNA complex. In contrast, the RNA and La protein are not electrophoretically separated in the fluorescence polarization assay, but kept in solution. This allows the protein to constantly bind, release, and re-bind RNA, thus allowing for analysis of the RNA-protein interaction and kinetics when measured over time. However, the formation of either one or several protein:RNA complexes cannot be distinguished by the fluorescence polarization assay, but can be distinguished by native EMSA. Hence, the fluorescence assay is a time saving method for monitoring the hLa-D1-ATG RNA interaction and kinetics over time, but does not allow conclusions about protein stoichiometry. The fluorescence polarization assay is of higher accuracy compared to the EMSA; because, no subjective interpretation of the signal area for quantification has to be assessed, inter alia. However, this work shows that the RNA-binding affinities of hLa to the cyclin D1 translational start site context determined by both assays were similar.

A limitation of the FP assay is the requirement of a small labeled ligand, usually less than 10 kDa and/or less than 30 nucleotides [220]. The D1-ATG RNA oligoribonucleotide used

herein consisted of 47 nucleotides, which equals a mass of 15.1 kDa, or 15.6 kDa including the 6-FAM dye. Although the RNA was 50% larger than recommended [220], the fluorescence polarization assay worked with this oligoribonucleotide, which may have resulted from folding of the RNA into a more compact structure. Indeed, the *mfold* software [255] only predicts one folding for the D1-ATG RNA into a stem-loop structure as shown in figure 5.1 and later discussed in detail. Ultimately, using a combination of both methods made for a great advantage in understanding the interactions of RNA-binding proteins and their target RNA.

5.5.2 Oligomerization of the human La protein *in vitro*

The dimerization domain of hLa was proposed to be located between aa 298 and 348 [35]. Interestingly, the La mutant RRM+2 forms multimers in EMSAs (figures 4.2.4 and 4.2.5) suggesting that the dimerization domain is located between amino acids 298 to 326. It is known that RRMs can mediate protein:protein interactions [311] [21] allowing for the possibility that La multimerization may be mediated via RRM:RRM contacts between two different La molecules. However, the most critical question is if potential multimerization are of functional relevance. The La mutant (La(236-348)) [67] is obviously not able to interact with La using Far-Western blotting and also proposed to inactivate endogenous La when expressed in cells. This mutant was used as a *trans*-dominant La mutant to study the role of La in translation [67].

However, *in vitro* higher order La-RNP complexes are suggested to result from homodimerization, which are formed after the formation of saturated primary complexes and [47]. These complexes have been shown to have lower target specificity [63]. Native EMSA binding studies revealed that the La protein and most of the La protein mutants, except RCD2, form not only a primary but also secondary and tertiary RNP complexes with the D1-ATG RNA.

In order to determine if hLa binds RNA as a monomer, EMSAs with the minimal La protein RRM1+2 were performed in the presence of the unspecific competitor molecule heparin, which mimics the negatively charged RNA phosphate backbone. Monomeric La binding was induced upon the addition of excessive amounts of heparin; however, it could not be concluded if more than one copy of hLa binds the cyclin D1 RNA in the absence of heparin or if La is forming homodimers.

The collected data suggested that monomeric hLa binds cyclin D1 RNA *in vitro*. However, it is still unclear if multimeric La-RNP complexes are a result of protein-protein interaction of RNA-bound La or due to the association of more than one La copy with one RNA molecule. The *in vivo* dimerization of hLa remains elusive. The overexpression of a La dominant negative mutant in HeLa cells abrogated the function of endogenous La in poliovirus and HCV IRES-dependent translation, hinting at functional dimerization of hLa in HeLa cells [67]. The homodimerization of hLa could not be confirmed by chemical shift analyses, ^{15}N backbone dynamics, or by analytical ultracentrifugation [66], therefore, further studies are required to understand the functional aspects of La dimerization *in vivo*.

5.5.3 Considerations for establishing an *in vitro* sumoylation assay

In optimizing the IVSA the ratio of the conjugating enzyme Ubc9 to SUMO appears to be one of the most critical parameters for efficient sumoylation as shown in figure 4.3.3A. The molar ratio of Ubc9 to SUMO (1:18) resulted in the most efficient conjugation of SUMO-1 to hLa protein. Similar results were achieved with Ubc9:SUMO-1 ratios of 1:9 and 1:37. The requirement of such an excessive molar amount of SUMO-1 compared to Ubc9 implies that (a) a large number of recombinant SUMO proteins may not be in a functional folding due to the recombinant protein expression, (b) the catalytic domain of the Ubc9 protein is folded incorrectly and does not allow the release of the SUMO-1 protein to the La protein and no recycling of Ubc9 can occur, or (c) a ligase is required to facilitate the conjugation of the small peptide to the target. In order to optimize the IVSA two catalytically active fragments, IR1+M and BP2ΔFG [231], of the E3 ligase RanBP2 were included in the IVSA assay. However, the presence of either truncated ligase did not improve the IVSA efficiency but decreased the efficiency (figure 4.3.5).The requirement of a ligase for an optimal transfer of SUMO to the La protein cannot be excluded, because several other SUMO described ligases [179] [180] [181] may target-specifically, facilitating the sumoylation of the hLa protein *in vitro*.

However, determining the optimal incubation parameters increased the efficiency of the IVSA. SUMO-modification of the hLa protein at the physiological relevant temperature of 37°C resulted in an overall less efficient sumoylation, and the sumoylation product also appeared later in comparison to the sumoylation performed at 30°C (figure 4.3.3B). This was

optimization of an *in vitro* system, therefore it does not allow conclusions regarding biological sumoylation conditions.

The conjugation deficient SUMO-GA mutant represents a vigorous control for the IVSA because the substitution of the terminal glycine to alanine in this mutant inhibits the trans-esterification and ligation of the SUMO peptide to the target protein [229]. The non-hydrolysable ATP analog phosphorothioate ATP-γ-S was tested as a secondary negative sumoylation control. In the absence of any energy source no sumoylation products were detected (figure 4.3.4), unexpectedly sumoylation occurred in the presence of the ATP analog ATP-γ-S. However, the efficiency of the SUMO-modification was reduced compared to the parallel reaction with ATP. A contamination of the reaction with ATP during the reaction assembly was possible, but is not very like because in parallel prepared negative controls without ATP did not result in sumoylation of the La protein. A possible contamination of the protein preparations could be discarded for the same reason. It is rather likely that the ATP-γ-S solution was contaminated with ATP. However, Peck and Herschlag [312] have demonstrated that the eukaryotic initiation factor 4A (eIF4A) efficiently uses ATP-γ-S due to its ATPase activity. It is suggested to be cautious when using an ATP analog as a mechanistic probe.

The human La protein has been proposed to contain a Walker A nucleotide binding motif in the CTD to render ATP dependent-helicase activity [114] [313] or to bind 5'-triphosphate ends of nascent tRNAs [37]. Hence, the sumoylation of La in the presence of the putatively non-hydrolysable ATP-γ-S may also be a result of the ATPase activity in the La protein.

In conclusion, the SUMO-conjugation deficient SUMO-GA mutant appears to be the most vigorous and appropriate sumoylation control.

5.5.4 Efficiency of IVSA

A sumoylation of an average 5% hLa protein resulted in a two-fold increased D1-FL RNA-binding activity and a 50% increased D1-ATG RNA-binding at higher La concentrations (figure 4.3.12B and D). The performed RNA-binding studies and additional observations in the laboratory suggest that only ~ 10% to 20% of the hLa preparation is RNA-binding active. This would assume that a majority of the protein may be misfolded, as often observed during recombinant protein expression. The dramatic effects of sumoylation on the RNA-binding activity of hLa protein may result from the sumoylation-induced structural change of the

inter-RRM peptide linker, which transforms RNA-binding inactive into RNA-binding active hLa proteins and ultimately results in an increased RNA-binding activity.

SUMO-modification has been suggested to mask or add interaction surfaces and result in altered protein–protein interactions of target proteins by inducing conformational changes [314]. In addition, it has been suggested that disordered regions may adopt an ordered structure upon target binding [122] [123] [124] [125]. Therefore, sumoylation at those unstructured regions may induce (partial) structuring of the disordered inter-RRM peptide linker. It is well established that the fusion of the SUMO peptide to the N'-terminus of a protein of interest promotes correct folding of its fusion partner and enhances its effects on recombinant protein expression and solubility [315] [316]. The expression of SUMO-target proteins in living cells is often used to study the function of sumoylated proteins in a cellular context. Therefore, it would be interesting to investigate if the SUMO-La fusion protein would display a larger fraction of RNA-binding competent La protein. Ultimately, it would be critical to establish if sumoylation of hLa promotes correct folding or if it is a regulatory modification that regulates the RNA-binding activity of the La protein (as discussed earlier, 5.3.2).

On the other hand, the strong effects on RNA-binding by sumoylation may be explained if sumoylation actually occurs preferentially at the 10-20% RNA-binding competent and La proteins resulting in a positive alteration of the RNA-binding activity, e.g. by facilitating the positioning of the tandem RRMs. Indeed, more work is required to clarify this problem and may be approached by enriching SUMO-La (e.g. by gel filtration) to compare the RNA-binding affinity of SUMO-modified La versus native La.

5.5.5 *In silico* prediction of SUMO-acceptor sites

Two different programs were used as an *in silico* approach to predict potential lysine residues as SUMO-acceptor sites.

The FASTA sequence of hLa (refer to appendix) was submitted to SUMOsp2.0 and SUMOplot Analysis Program to predict sumoylation sites of hLa. The algorithms in both programs identified several different sumoylation site of different stringency (table 4.3.1 and 4.3.2). The sites with the highest score in SUMOplot were determined for K41, K185 and K208. The SUMOsp2.0 program predicted 14 potential SUMO-acceptor sites at the default medium cut-off. The SUMOplot program predicted lysine 208 as a highly scored

sumoylation site; however, experimentally lysine 208 was definitively identified as a sumoylation site. The *in silico* approach helped strengthening the hypothesis that lysine 208 may be serving as a sumoylation site, but the outputted list was too large to experimentally handle the identification of the additional sites by mutagenesis of hLa.

In conclusion, *in silico* prediction of sumoylation sites may be useful for some substrates; however, a high abundance of lysines, as in the case of hLa, may make it difficult to predict a small amount of potential sumoylation residues. Combining *in silico* predictions with structural information in regards of the location of putative residues in exposed or disordered regions would certainly narrow down possible sumoylation sites.

5.6 Future directions

The role and function of cytoplasmic La in stimulating translation of cellular mRNAs has become more consolidated, however, future studies are required to fully comprehend the distinct role of the human La protein.

Addressing the oligomerization of the human La protein *in vivo* is important. The functional loss of endogenous La by a dominant negative acting overexpressed La mutant has been shown, however, the multimerization has not been shown directly *in vivo*. A mammalian-two-hybrid assay would determine if the human La protein is forming dimers or multimers *in vivo*. Another method based on fluorescence recovery after energy transfer (FRET) could show the oligomerization of two populations of the hLa protein, labeled with two different fluorescence dyes, when expressed in one cell and analyzed by microscopy.

The described cooperative RRM1 and RRM2 binding mechanism of La to the cyclin D1 mRNA needs to be proven for other cellular mRNAs to exclude a cyclin D1 specific RNA-binding mechanism. The requirement of RRM1 and RRM2 for La-binding to XIAP mRNAs and Mdm2 mRNAs by EMSA, as representative internally recognized mRNAs, would further support the role of La in the translational stimulation of those tumor-promoting factors. *In vitro* XIAP and Mdm2 RNA-binding studies including SUMO-modified La would support the presented model by which sumoylation enhances its RNA-binding activity to cellular mRNAs.

The binding site of the La protein in cyclin D1 mRNA has not been fully mapped. In order to identify the contact sites of the human La protein, RNA-footprinting experiments using specific riboucleases to digest La bound and unbound RNA would need to be performed. Using the RNases T1, cleaving the 3'-ends of single stranded guanines, and V1, which cuts specifically base-paired sequences, would aid in identify the La binding site and also provide structural information about the RNA.

A comprehensive study of structure-driven binding rather than sequence-specific binding would be needed to understand how La recognizes substrates and is capable of targeting the interaction. In order to better understand the sequence/structure-driven La recognition, the RNA-binding affinity of La to all described La targets could be compared by competitive EMSAs or FP assays. The novel PAR-CLIP (photoreactive-ribonucleoside-enhanced clossslinking and immunoprecipitation) method would help identify comprehensively and specific La binding sites in the proximity of a translational start site codon in living cells.

In addition, RNA-binding studies of sumoylated La with RNA oligoribonucleotides comprised of different structural features, such as different stem lengths and/or loop sizes and altering sequences in those structural features, would clarify the recognition of mRNAs by the La protein and explain why there is no identified consensus sequence.

In order to show that the RNA-chaperone activity of the La protein is responsible for restructuring highly structured 5'-UTRs and to facilitate the scanning of 43S preinitiation complexs and/or assembly of the 80S initiation complex at the translational start site the required domain has to be mapped.

Recombinant La mutants can be tested in an RNA-chaperone assay based on a molecular beacon. An RNA oligoribonucleotide representing the RNA of interest harboring a stem-loop structure would be labeled with a fluorophore at one end and on the other with a quencher. Upon excitation the fluorophore does not emit fluorescence light because of the quencher, which is in close proximity to the fluorophore as a result of the formed stem. As La binds, a fluorescence signal would be detected if La were opening the stem because the quencher would be too far from the fluorophore to quench the signal.

To show the requirement of RNA-chaperone activity for translation, the RNA-chaperone deficient La mutants need to be overexpressed in La-depleted cells and cell extracts analyzed for the expression of the RNA of interest.

In order to understand the regulation and effects of sumoylation on the La protein *in vivo*, the sumoylation of La needs to be studied in detail in living cells. Sumoylation in rat neurons trigged the retrograde transport of La and herein was demonstrated that the RNA-binding activity of La is enhanced upon sumoylation. However, the sumoylation of La in human cells needs to be analyzed for the localization, SUMO paralog specificity, and the effect on translation. A robust system has to be established to show sumoylation in living cells. This could be achieved by overexpressing the conjugating enzyme Ubc9, SUMO, and using a fluorescence labeled La protein as well as inducing stress in a cellular system.

To achieve a high impact in support of this work a co-crystalized structure or NMR structure of human La bound to a cellular mRNA would need to be solved, such as cyclin D1. The minimal La protein RRM1+2 has been shown to be cyclin D1 RNA-binding active with a similar but slightly less affinity, compared to the wild type hLa. However, the smaller size of the RRM1+2 protein makes the protein a reasonable candidate for crystallization and/or NMR analyses. This would allow understanding into the interactions of La to the RNA and would allow a targeted approach to design anti-cancer drugs.

References

2.   Mattioli M & Reichlin M (1974) Heterogeneity of RNA protein antigens reactive with sera of patients with systemic lupus erythematosus. Description of a cytoplasmic nonribosomal antigen. *Arthritis & Rheumatism* **17**, 421-429.

3.   Alspaugh MA & Tan EM (1975) Antibodies to cellular antigens in Sjogren's syndrome. *J Clin Invest* **55**, 1067-1073, doi: 10.1172/JCI108007.

4.   Chambers JC & Keene JD (1985) Isolation and analysis of cDNA clones expressing human lupus La antigen. *Proc Natl Acad Sci U S A* **82**, 2115-2119.

5.   Park JM, Kohn MJ, Bruinsma MW, Vech C, Intine RV, Fuhrmann S, Grinberg A, Mukherjee I, Love PE, Ko MS, et al. (2006) The multifunctional RNA-binding protein La is required for mouse development and for the establishment of embryonic stem cells. *Mol Cell Biol* **26**, 1445-1451.

6.   Troster H, Metzger TE, Semsei I, Schwemmle M, Winterpacht A, Zabel B & Bachmann M (1994) One gene, two transcripts: isolation of an alternative transcript encoding for the autoantigen La/SS-B from a cDNA library of a patient with primary Sjogrens' syndrome. *J Exp Med* **180**, 2059-2067.

7.   Grolz D & Bachmann M (1997) The nuclear autoantigen La/SS-associated antigen B: one gene, three functional mRNAs. *Biochemical Journal* **323**, 151-158.

8.   Carter MS & Sarnow P (2000) Distinct mRNAs that encode La autoantigen are differentially expressed and contain internal ribosome entry sites. *J Biol Chem* **275**, 28301-28307.

9.   Chambers JC, Kenan D, Martin BJ & Keene JD (1988) Genomic structure and amino acid sequence domains of the human La autoantigen. *J Biol Chem* **263**, 18043-18051.

10.   Trotta R, Vignudelli T, Candini O, Intine RV, Pecorari L, Guerzoni C, Santilli G, Byrom MW, Goldoni S, Ford LP, et al. (2003) BCR/ABL activates mdm2 mRNA translation via the La antigen. *Cancer Cell* **3**, 145-160.

11.   Sommer G, Rossa C, Chi AC, Neville BW & Heise T (2011) Implication of RNA-binding protein La in proliferation, migration and invasion of lymph node-metastasized hypopharyngeal SCC cells. *PloS one* **6**, e25402, doi: 10.1371/journal.pone.0025402.

12.   Chen Y & Varani G (2005) Protein families and RNA recognition. *Febs J* **272**, 2088-2097, doi: 10.1111/j.1742-4658.2005.04650.x.

13.   Lunde BM, Moore C & Varani G (2007) RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol* **8**, 479-490, doi: 10.1038/nrm2178.

14.   Lee MH & Schedl T (2006) RNA-binding proteins. *WormBook*, 1-13, doi: 10.1895/wormbook.1.79.1.

15.   Burd CG & Dreyfuss G (1994) Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265**, 615-621.

16.   Swanson MS, Nakagawa TY, LeVan K & Dreyfuss G (1987) Primary structure of human nuclear ribonucleoprotein particle C proteins: conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA, and pre-rRNA-binding proteins. *Mol Cell Biol* **7**, 1731-1739.

17.   Maraia RJ & Intine RV (2001) Recognition of nascent RNA by the human La antigen: conserved and divergent features of structure and function. *Mol Cell Biol* **21**, 367-379.

18.   Maraia RJ & Intine RV (2002) La protein and its associated small nuclear and nucleolar precursor RNAs. *Gene Expr* **10**, 41-57.

19.   Jacks A, Kelly G, Curry S & Conte MR (2002) Resonance assignment and secondary structure determination of a C-terminal fragment of the lupus autoantigen (La) protein containing a putative RNA recognition motif (RRM). *J Biomol NMR* **22**, 387-388.

20.   Wolin SL & Cedervall T (2002) The La protein. *Annu Rev Biochem* **71**, 375-403.

21.   Clery A, Blatter M & Allain FH (2008) RNA recognition motifs: boring? Not quite. *Curr Opin Struct Biol* **18**, 290-298.

22.   Maris C, Dominguez C & Allain FH (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *Febs J* **272**, 2118-2131.

23.   Dreyfuss G, Philipson L & Mattaj IW (1988) Ribonucleoprotein particles in cellular processes. [Review] [100 refs]. *Journalof Cell Biology* **106**, 1419-1425.

24.   Wittekind M, Gorlach M, Friedrichs M, Dreyfuss G & Mueller L (1992) 1H, 13C, and 15N NMR assignments and global folding pattern of the RNA-binding domain of the human hnRNP C proteins. *Biochemistry* **31**, 6254-6265.

25.   Nagai K, Oubridge C, Jessen TH, Li J & Evans PR (1990) Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature* **348**, 515-520.

26.   Birney E, Kumar S & Krainer AR (1993) Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res* **21**, 5803-5816.

27.   Gorlach M, Burd CG & Dreyfuss G (1994) The determinants of RNA-binding specificity of the heterogeneous nuclear ribonucleoprotein C proteins. *Journalof BiologicalChemistry* **269**, 23074-23078.

28.   Bousquet-Antonelli C & Deragon JM (2009) A comprehensive analysis of the La-motif protein superfamily. *Rna* **15**, 750-764, doi: 10.1261/rna.1478709.

29.   Gajiwala KS & Burley SK (2000) Winged helix proteins. *Curr Opin Struct Biol* **10**, 110-116.

30.   Dong G, Chakshusmathi G, Wolin SL & Reinisch KM (2004) Structure of the La motif: a winged helix domain mediates RNA binding via a conserved aromatic patch. *Embo J* **23**, 1000-1007.

31.   Teplova M, Yuan YR, Phan AT, Malinina L, Ilin S, Teplov A & Patel DJ (2006) Structural basis for recognition and sequestration of UUU(OH) 3' temini of nascent RNA polymerase III transcripts by La, a rheumatic disease autoantigen. *Mol Cell* **21**, 75-85.

32.   Maraia RJ & Bayfield MA (2006) The La protein-RNA complex surfaces. *Mol Cell* **21**, 149-152.

33.   Horke S, Reumann K, Schweizer M, Will H & Heise T (2004) Nuclear trafficking of La protein depends on a newly identified nucleolar localization signal and the ability to bind RNA. *J Biol Chem* **279**, 26563-26570. Epub 22004 Apr 26561.

34.   Craig A, Svitkin Y, Lee H, Belsham G & Sonenberg N (1997) The La autoantigen contains a dimerization domain that is essential for enhancing translation. *Mol Cell Biol* **17**, 163-169.

35.   Horke S, Reumann K, Rang A & Heise T (2002) Molecular characterization of the human La protein.hepatitis B virus RNA.B interaction in vitro. *J Biol Chem* **277**, 34949-34958.

36.   Broekhuis CH, Neubauer G, van Der Heijden A, Mann M, Proud CG, van Venrooij WJ & Pruijn GJ (2000) Detailed Analysis of the Phosphorylation of the Human La (SS-B) Autoantigen. (De)phosphorylation Does Not Affect Its Subcellular Distribution. *Biochemistry* **39**, 3023-3033.

37.   Fan H, Sakulich AL, Goodier JL, Zhang X, Qin J & Maraia RJ (1997) Phosphorylation of the human La antigen on serine 366 can regulate recycling of RNA polymerase III transcription complexes. *Cell* **88**, 707-715.

38.   Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, Walther TC, Olsen JV & Mann M (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325**, 834-840, doi: 10.1126/science.1175371.

39.   Topfer F, Gordon T & McCluskey J (1993) Characterization of the mouse autoantigen La (SS-B). Identification of conserved RNA-binding motifs, a putative ATP binding site and reactivity of recombinant protein with poly(U) and human autoantibodies. *J Immunol* **150**, 3091-3100.

40.   Goodier JL, Fan H & Maraia RJ (1997) A carboxy-terminal basic region controls RNA polymerase III transcription factor activity of human La protein. *Mol Cell Biol* **17**, 5823-5832.

41.   Brenet F, Socci ND, Sonenberg N & Holland EC (2009) Akt phosphorylation of La regulates specific mRNA translation in glial progenitors. *Oncogene* **28**, 128-139.

42.   Moritz A, Li Y, Guo A, Villen J, Wang Y, MacNeill J, Kornhauser J, Sprott K, Zhou J, Possemato A, et al. (2010) Akt-RSK-S6 kinase signaling networks activated by oncogenic receptor tyrosine kinases. *Sci Signal* **3**, ra64, doi: 10.1126/scisignal.2000998.

43.   Simons FH, Broers FJ, Van Venrooij WJ & Pruijn GJ (1996) Characterization of cis-acting signals for nuclear import and retention of the La (SS-B) autoantigen. *Experimental Cell Research* **224**, 224-236.

44.   Rosenblum JS, Pemberton LF & Blobel G (1997) A nuclear import pathway for a protein involved in tRNA maturation. *Journal of Cell Biology* **139**, 1655-1661.

45.   Rosenblum JS, Pemberton LF, Bonifaci N & Blobel G (1998) Nuclear import and the evolution of a multifunctional RNA-binding protein. *J Cell Biol* **143**, 887-899.

46.   Intine RV, Dundr M, Misteli T & Maraia RJ (2002) Aberrant nuclear trafficking of La protein leads to disordered processing of associated precursor tRNAs. *Mol Cell* **9**, 1113-1123.

47.   Fan H, Goodier JL, Chamberlain JR, Engelke DR & Maraia RJ (1998) 5' processing of tRNA precursors can Be modulated by the human La antigen phosphoprotein. *Mol Cell Biol* **18**, 3201-3211.

48.   Gottlieb E & Steitz JA (1989) Function of the mammalian La protein: evidence for its action in transcription termination by RNA polymerase III. *Embo J* **8**, 851-861.

49.   Hendrick JP, Wolin SL, Rinke J, Lerner MR & Steitz JA (1981) Ro small cytoplasmic ribonucleoproteins are a subclass of La ribonucleoproteins: further characterization of the Ro and La small ribonucleoproteins from uninfected mammalian cells. *Mol Cell Biol* **1**, 1138-1149.

50.   Elenitsas R, Bair LW, Jr., Medsger TA, Jr. & Deng JS (1986) Discordance of SSA/Ro and SSB/La cellular antigens in synchronized cells. *J Invest Dermatol* **87**, 504-509.

51.   Deng JS, Takasaki Y & Tan EM (1981) Nonhistone nuclear antigens reactive with autoantibodies. Immunofluorescence studies on distribution in synchronized cells. *J Cell Biol* **91**, 654-660.

52.   Graus F, Cordon-Cardo C, Bonfa E & Elkon KB (1985) Immunohistochemical localization of La nuclear antigen in brain. Selective concentration of the La protein in neuronal nucleoli. *J Neuroimmunol* **9**, 307-319.

53.   Fok V, Friend K & Steitz JA (2006) Epstein-Barr virus noncoding RNAs are confined to the nucleus, whereas their partner, the human La protein, undergoes nucleocytoplasmic shuttling. *J Cell Biol* **173**, 319-325.

54.   Rutjes SA, Utz PJ, van der Heijden A, Broekhuis C, van Venrooij WJ & Pruijn GJ (1999) The La (SS-B) autoantigen, a key protein in RNA biogenesis, is dephosphorylated and cleaved early during apoptosis. *Cell Death Differ* **6**, 976-986.

55.   Ayukawa K, Taniguchi S, Masumoto J, Hashimoto S, Sarvotham H, Hara A, Aoyama T & Sagara J (2000) La Autoantigen Is Cleaved in the COOH Terminus and Loses the Nuclear Localization Signal during Apoptosis. *J Biol Chem* **275**, 34465-34470.

56.   Shiroki K, Isoyama T, Kuge S, Ishii T, Ohmi S, Hata S, Suzuki K, Takasaki Y & Nomoto A (1999) Intracellular redistribution of truncated La protein produced by poliovirus 3Cpro-mediated cleavage. *J Virol* **73**, 2193-2200.

57.   Bachmann M, Falke D, Schroder HC & Muller WE (1989) Intracellular distribution of the La antigen in CV-1 cells after herpes simplex virus type 1 infection compared with the localization of U small nuclear ribonucleoprotein particles. *J Gen Virol* **70**, 881-891.

58.   Schwartz EI, Intine RV & Maraia RJ (2004) CK2 Is Responsible for Phosphorylation of Human La Protein Serine-366 and Can Modulate rpL37 5'-Terminal Oligopyrimidine mRNA Metabolism. *Mol Cell Biol* **24**, 9580-9591.

59.   Intine RV, Tenenbaum SA, Sakulich AL, Keene JD & Maraia RJ (2003) Differential phosphorylation and subcellular localization of La RNPs associated with precursor tRNAs and translation-related mRNAs. *Mol Cell* **12**, 1301-1307.

60.   van Niekerk EA, Willis DE, Chang JH, Reumann K, Heise T & Twiss JL (2007) Sumoylation in axons triggers retrograde transport of the RNA-binding protein La. *Proc Natl Acad Sci U S A* **104**, 12913-12918, doi: 10.1073/pnas.0611562104.

61.   Golebiowski F, Matic I, Tatham MH, Cole C, Yin Y, Nakamura A, Cox J, Barton GJ, Mann M & Hay RT (2009) System-wide changes to SUMO modifications in response to heat shock. *Sci Signal* **2**, ra24, doi: 2/72/ra24 [pii] 10.1126/scisignal.2000282.

62.   Ali N & Siddiqui A (1997) The La antigen binds 5' noncoding region of the hepatitis C virus RNA in the context of the initiator AUG codon and stimulates internal ribosome entry site-mediated translation. *Proc Natl Acad Sci U S A* **94**, 2249-2254.

63.   Long KS, Cedervall T, Walch-Solimena C, Noe DA, Huddleston MJ, Annan RS & Wolin SL (2001) Phosphorylation of the Saccharomyces cerevisiae La protein does not appear to be required for its functions in tRNA maturation and nascent RNA stabilization. *Rna* **7**, 1589-1602.

64.   Ohndorf UM, Steegborn C, Knijff R & Sondermann P (2001) Contributions of the individual domains in human La protein to its RNA 3'-end binding activity. *J Biol Chem* **276**, 27188-27196.

65.   Kuhn U & Pieler T (1996) Xenopus poly(A) binding protein: functional domains in RNA binding and protein-protein interaction. *J Mol Biol* **256**, 20-30, doi: 10.1006/jmbi.1996.0065.

66.   Jacks A, Babon J, Kelly G, Manolaridis I, Cary PD, Curry S & Conte MR (2003) Structure of the C-terminal domain of human La protein reveals a novel RNA recognition motif coupled to a helical nuclear retention element. *Structure* **11**, 833-843.

67.   Costa-Mattioli M, Svitkin Y & Sonenberg N (2004) La Autoantigen Is Necessary for Optimal Function of the Poliovirus and Hepatitis C Virus Internal Ribosome Entry Site In Vivo and In Vitro. *Mol Cell Biol* **24**, 6861-6870.

68.   Bai C, Li Z & Tolias PP (1994) Developmental characterization of a Drosophila RNA-binding protein homologous to the human systemic lupus erythematosus-associated La/SS-B autoantigen. *Molecular & Cellular Biology* **14**, 5123-5129.

69.   Yoo CJ & Wolin SL (1994) La proteins from Drosophila melanogaster and Saccharomyces cerevisiae: a yeast homolog of the La autoantigen is dispensable for growth. *Molecular & Cellular Biology* **14**, 5412-5424.

70.   Scherly D, Stutz F, Lin-Marq N & Clarkson SG (1993) La proteins from Xenopus laevis. cDNA cloning and developmental expression. *J Mol Biol* **231**, 196-204.

71.   Pellizzoni L, Cardinali B, Lin-Marq N, Mercanti D & Pierandrei-Amaldi P (1996) A Xenopus laevis homologue of the La autoantigen binds the pyrimidine tract of the 5' UTR of ribosomal protein mRNAs in vitro: implication of a protein factor in complex formation. *J Mol Biol* **259**, 904-915, doi: S0022-2836(96)90368-X [pii] 10.1006/jmbi.1996.0368.

72.   Yoo CJ & Wolin SL (1997) The yeast La protein is required for the 3' endonucleolytic cleavage that matures tRNA precursors. *Cell* **89**, 393-402.

73.   Intine RV, Sakulich AL, Koduru SB, Huang Y, Pierstorff E, Goodier JL, Phan L & Maraia RJ (2000) Control of transfer RNA maturation by phosphorylation of the human La antigen on serine 366. *Mol Cell* **6**, 339-348.

74.   Inada M & Guthrie C (2004) Identification of Lhp1p-associated RNAs by microarray analysis in Saccharomyces cerevisiae reveals association with coding and noncoding RNAs. *Proc Natl Acad Sci U S A* **101**, 434-439, doi: 10.1073/pnas.0307425100.

75.   Stefano JE (1984) Purified lupus antigen La recognizes an oligouridylate stretch common to the 3' termini of RNA polymerase III transcripts. *Cell* **36**, 145-154.

76.   Huang Y, Bayfield MA, Intine RV & Maraia RJ (2006) Separate RNA-binding surfaces on the multifunctional La protein mediate distinguishable activities in tRNA maturation. *Nat Struct Mol Biol* **13**, 611-618.

77.   Kufel J, Allmang C, Chanfreau G, Petfalski E, Lafontaine DL & Tollervey D (2000) Precursors to the U3 small nucleolar RNA lack small nucleolar RNP proteins but are stabilized by La binding. *Mol Cell Biol* **20**, 5415-5424.

78.   Xue D, Rubinson DA, Pannone BK, Yoo CJ & Wolin SL (2000) U snRNP assembly in yeast involves the La protein [published erratum appears in EMBO J 2000 Jun 1;19(11):2763]. *Embo J* **19**, 1650-1660.

79.   Maraia RJ, Kenan DJ & Keene JD (1994) Eukaryotic transcription termination factor La mediates transcript release and facilitates reinitiation by RNA polymerase III. *Mol Cell Biol* **14**, 2147-2158.

80.   Pannone BK, Xue D & Wolin SL (1998) A role for the yeast La protein in U6 snRNP assembly: evidence that the La protein is a molecular chaperone for RNA polymerase III transcripts. *Embo J* **17**, 7442-7453.

81.   Lin-Marq N & Clarkson SG (1998) Efficient synthesis, termination and release of RNA polymerase III transcripts in Xenopus extracts depleted of La protein. *EMBO Journal* **17**, 2033-2041.

82.   Rinke J & Steitz JA (1982) Precursor molecules of both human 5S ribosomal RNA and transfer RNAs are bound by a cellular protein reactive with anti-La lupus antibodies. *Cell* **29**, 149-159.

83.   Rinke J & Steitz JA (1985) Association of the lupus antigen La with a subset of U6 snRNA molecules. *Nucleic Acids Res* **13**, 2617-2629.

84.   Heise T, Guidotti LG, Cavanaugh VJ & Chisari FV (1999) Hepatitis B virus RNA-binding proteins associated with cytokine-induced clearance of viral RNA from the liver of transgenic mice. *J Virol* **73**, 474-481.

85.   Heise T, Guidotti LG & Chisari FV (2001) Characterization of nuclear RNases that cleave hepatitis B virus RNA near the La protein binding site. *J Virol* **75**, 6874-6883.

86.   Spangberg K, Wiklund L & Schwartz S (2001) Binding of the La autoantigen to the hepatitis C virus 3' untranslated region protects the RNA from rapid degradation in vitro. *J Gen Virol* **82**, 113-120.

87.   Brenet F, Dussault N, Borch J, Ferracci G, Delfino C, Roepstorff P, Miquelis R & Ouafik L (2005) Mammalian peptidylglycine alpha-amidating monooxygenase mRNA expression can be modulated by the La autoantigen. *Mol Cell Biol* **25**, 7505-7521.

88.   Vazquez-Pianzola P, Urlaub H & Rivera-Pomar R (2005) Proteomic analysis of reaper 5' untranslated region-interacting factors isolated by tobramycin affinity-selection reveals a role for La antigen in reaper mRNA translation. *Proteomics* **5**, 1645-1655.

89.   Sommer G, Dittmann J, Kuehnert J, Reumann K, Schwartz PE, Will H, Coulter BL, Smith MT & Heise T (2011) The RNA-binding protein La contributes to cell proliferation and CCND1 expression. *Oncogene* **30**, 434-444, doi: 10.1038/onc.2010.425.

90.	Ehlers I, Horke S, Reumann K, Rang A, Grosse F, Will H & Heise T (2004) Functional Characterization of the Interaction between Human La and Hepatitis B Virus RNA. *J Biol Chem* **279**, 43437-43447. Epub 42004 Aug 43439.

91.	Holcik M & Korneluk RG (2000) Functional characterization of the X-linked inhibitor of apoptosis (XIAP) internal ribosome entry site element: role of La autoantigen in XIAP translation. *Mol Cell Biol* **20**, 4648-4657.

92.	Kim YK, Back SH, Rho J, Lee SH & Jang SK (2001) La autoantigen enhances translation of BiP mRNA. *Nucleic Acids Res* **29**, 5009-5016.

93.	Cardinali B, Carissimi C, Gravina P & Pierandrei-Amaldi P (2003) La protein is associated with terminal oligopyrimidine mRNAs in actively translating polysomes. *J Biol Chem* **278**, 35145-35151, doi: 10.1074/jbc.M300722200 M300722200 [pii].

94.	Petz M, Them N, Huber H, Beug H & Mikulits W (2012) La enhances IRES-mediated translation of laminin B1 during malignant epithelial to mesenchymal transition. *Nucleic Acids Res* **40**, 290-302, doi: 10.1093/nar/gkr717.

95.	King HA, Cobbold LC & Willis AE (2010) The role of IRES trans-acting factors in regulating translation initiation. *Biochem Soc Trans* **38**, 1581-1586, doi: 10.1042/BST0381581.

96.	Lopez-Lastra M, Rivas A & Barria MI (2005) Protein synthesis in eukaryotes: the growing biological relevance of cap-independent translation initiation. *Biol Res* **38**, 121-146.

97.	Meerovitch K, Pelletier J & Sonenberg N (1989) A cellular protein that binds to the 5'-noncoding region of poliovirus RNA: implications for internal translation initiation. *Genes Dev* **3**, 1026-1034.

98.	Meerovitch K, Svitkin YV, Lee HS, Lejbkowicz F, Kenan DJ, Chan EK, Agol VI, Keene JD & Sonenberg N (1993) La autoantigen enhances and corrects aberrant translation of poliovirus RNA in reticulocyte lysate. *J Virol* **67**, 3798-3807.

99.	Ali N, Pruijn GJ, Kenan DJ, Keene JD & Siddiqui A (2000) Human La antigen is required for the hepatitis C virus internal ribosome entry site-mediated translation. *J Biol Chem* **275**, 27531-27540.

100.	Pudi R, Srinivasan P & Das S (2004) La protein binding at the GCAC site near the initiator AUG facilitates the ribosomal assembly on the hepatitis C virus RNA to influence internal ribosome entry site-mediated translation. *J Biol Chem* **279**, 29879-29888.

101.	Cordes S, Kusov Y, Heise T & Gauss-Muller V (2008) La autoantigen suppresses IRES-dependent translation of the hepatitis A virus. *Biochem Biophys Res Commun* **368**, 1014-1019.

102.	Svitkin YV, Pause A & Sonenberg N (1994) La autoantigen alleviates translational repression by the 5' leader sequence of the human immunodeficiency virus type 1 mRNA. *J Virol* **68**, 7001-7007.

103.	Chang YN, Kenan DJ, Keene JD, Gatignol A & Jeang KT (1995) Direct interactions between autoantigen La and human immunodeficiency virus leader RNA. *J Virol* **69**, 618-619.

104.	Crosio C, Boyl PP, Loreni F, Pierandrei-Amaldi P & Amaldi F (2000) La protein has a positive effect on the translation of TOP mRNAs in vivo. *Nucleic Acids Res* **28**, 2927-2934.

105.	Meyuhas O (2000) Synthesis of the translational apparatus is regulated at the translational level. *Eur J Biochem* **267**, 6321-6330.

106.	Chang Y-N, Kenan DJ, Keene JD, Gatignol A & Jeang K-T (1994) Direct interactions between autoantigen La and human immunodeficiency virus leader RNA. *J Virol* **68**, 7008-7020.

107.	Grimm C, Lund E & Dahlberg JE (1997) In vivo selection of RNAs that localize in the nucleus. *Embo J* **16**, 793-806.

108.   Kozak M (1987) At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J Mol Biol* **196**, 947-950.

109.   Kozak M (1989) Context effects and inefficient initiation at non-AUG codons in eucaryotic cell-free translation systems. *Mol Cell Biol* **9**, 5073-5080.

110.   De Angioletti M, Lacerra G, Sabato V & Carestia C (2004) Beta+45 G --> C: a novel silent beta-thalassaemia mutation, the first in the Kozak sequence. *Br J Haematol* **124**, 224-231.

111.   McBratney S & Sarnow P (1996) Evidence for involvement of trans-acting factors in selection of the AUG start codon during eukaryotic translational initiation. *Molecular & Cellular Biology* **16**, 3523-3534.

112.   Heise T, Guidotti LG & Chisari FV (1999) La autoantigen specifically recognizes a predicted stem-loop in hepatitis B virus RNA. *J Virol* **73**, 5767-5776.

113.   Martino L, Pennell S, Kelly G, Bui TT, Kotik-Kogan O, Smerdon SJ, Drake AF, Curry S & Conte MR (2012) Analysis of the interaction with the hepatitis C virus mRNA reveals an alternative mode of RNA recognition by the human La protein. *Nucleic Acids Res* **40**, 1381-1394, doi: 10.1093/nar/gkr890.

114.   Bachmann M, Pfeifer K, Schroder HC & Muller WE (1990) Characterization of the autoantigen La as a nucleic acid-dependent ATPase/dATPase with melting properties. *Cell* **60**, 85-93.

115.   Belisova A, Semrad K, Mayer O, Kocian G, Waigmann E, Schroeder R & Steiner G (2005) RNA chaperone activity of protein components of human Ro RNPs. *Rna* **11**, 1084-1094. Epub 2005 May 1031.

116.   Naeeni AR, Conte MR & Bayfield MA (2012) RNA chaperone activity of human La protein is mediated by variant RNA recognition motif. *J Biol Chem* **287**, 5472-5482, doi: 10.1074/jbc.M111.276071.

117.   Herschlag D (1995) RNA chaperones and the RNA folding problem. *J Biol Chem* **270**, 20871-20874.

118.   Schroeder R, Grossberger R, Pichler A & Waldsich C (2002) RNA folding in vivo. *Curr Opin Struct Biol* **12**, 296-300.

119.   Gottlieb E & Steitz JA (1987) The mammalian La protein is an RNA polymerase III trasncription termination factor. *Molecular Biology Reports* **12**.

120.   Gottlieb E & Steitz JA (1989) The RNA binding protein La influences both the accuracy and the efficiency of RNA polymerase III transcription in vitro. *EMBO Journal* **8**, 841-850.

121.   Chakshusmathi G, Kim SD, Rubinson DA & Wolin SL (2003) A La protein requirement for efficient pre-tRNA folding. *Embo J* **22**, 6562-6572, doi: 10.1093/emboj/cdg625.

122.   Wright PE & Dyson HJ (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol* **293**, 321-331, doi: 10.1006/jmbi.1999.3110.

123.   Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM & Obradovic Z (2002) Intrinsic disorder and protein function. *Biochemistry* **41**, 6573-6582.

124.   Fink AL (2005) Natively unfolded proteins. *Curr Opin Struct Biol* **15**, 35-41, doi: 10.1016/j.sbi.2005.01.002.

125.   Sugase K, Dyson HJ & Wright PE (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021-1025, doi: 10.1038/nature05858.

126.   Rajkowitsch L & Schroeder R (2007) Dissecting RNA chaperone activity. *Rna* **13**, 2053-2060, doi: 10.1261/rna.671807.

127.   Al-Ejeh F, Darby JM & Brown MP (2007) The La autoantigen is a malignancy-associated cell death target that is induced by DNA-damaging drugs. *Clin Cancer Res* **13**, 5509s-5518s.

128.   Sherr CJ (1995) D-type cyclins. *Trends Biochem Sci* **20**, 187-190.
129.   Sherr CJ (1996) Cancer cell cycles. *Science* **274**, 1672-1677.
130.   Sherr CJ & Roberts JM (2004) Living with or without cyclins and cyclin-dependent kinases. *Genes Dev* **18**, 2699-2711.
131.   Sherr CJ (2000) The Pezcoller lecture: cancer cell cycles revisited. *Cancer Res* **60**, 3689-3695.
132.   Pestell RG, Albanese C, Reutens AT, Segall JE, Lee RJ & Arnold A (1999) The cyclins and cyclin-dependent kinase inhibitors in hormonal regulation of proliferation and differentiation. *Endocr Rev* **20**, 501-534.
133.   Fu M, Wang C, Li Z, Sakamaki T & Pestell RG (2004) Minireview: Cyclin D1: normal and abnormal functions. *Endocrinology* **145**, 5439-5447.
134.   Horstmann S, Ferrari S & Klempnauer KH (2000) Regulation of B-Myb activity by cyclin D1. *Oncogene* **19**, 298-306, doi: 10.1038/sj.onc.1203302.
135.   Inoue K & Sherr CJ (1998) Gene expression and cell cycle arrest mediated by transcription factor DMP1 is antagonized by D-type cyclins through a cyclin-dependent-kinase-independent mechanism. *Mol Cell Biol* **18**, 1590-1600.
136.   Lamb J & Ewen ME (2003) Cyclin D1 and molecular chaperones: implications for tumorigenesis. *Cell Cycle* **2**, 525-527.
137.   Neuman E, Ladha MH, Lin N, Upton TM, Miller SJ, DiRenzo J, Pestell RG, Hinds PW, Dowdy SF, Brown M, et al. (1997) Cyclin D1 stimulation of estrogen receptor transcriptional activity independent of cdk4. *Mol Cell Biol* **17**, 5338-5347.
138.   Zwijsen RM, Wientjens E, Klompmaker R, van der Sman J, Bernards R & Michalides RJ (1997) CDK-independent activation of estrogen receptor by cyclin D1. *Cell* **88**, 405-415.
139.   Herber B, Truss M, Beato M & Müller R (1994) Inducible regulatory elements in the human cyclin D1 promoter. *Oncogene* **9**, 1295-1304.
140.   Hilakivi-Clarke L, Wang C, Kalil M, Riggins R & Pestell RG (2004) Nutritional modulation of the cell cycle and breast cancer. *Endocr Relat Cancer* **11**, 603-622, doi: 10.1677/erc.1.00665.
141.   Castro-Rivera E, Samudio I & Safe S (2001) Estrogen regulation of cyclin D1 gene expression in ZR-75 breast cancer cells involves multiple enhancer elements. *J Biol Chem* **276**, 30853-30861, doi: 10.1074/jbc.M103339200.
142.   Hellen CU & Sarnow P (2001) Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev* **15**, 1593-1612, doi: 10.1101/gad.891101.
143.   Stoneley M & Willis AE (2004) Cellular internal ribosome entry segments: structures, trans-acting factors and regulation of gene expression. *Oncogene* **23**, 3200-3207, doi: 10.1038/sj.onc.1207551.
144.   Johannes G, Carter MS, Eisen MB, Brown PO & Sarnow P (1999) Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray. *Proc Natl Acad Sci U S A* **96**, 13118-13123.
145.   Qin X & Sarnow P (2004) Preferential translation of internal ribosome entry site-containing mRNAs during the mitotic cycle in mammalian cells. *J Biol Chem* **279**, 13721-13728.
146.   Shi Y, Sharma A, Wu H, Lichtenstein A & Gera J (2005) Cyclin D1 and c-Myc internal ribosome entry site (IRES)-dependent translation is regulated by AKT activity and enhanced by rapamycin through a p38 MAPK and ERK-dependent pathway. *J Biol Chem*, M407874200.
147.   Chu EC & Tarnawski AS (2004) PTEN regulatory functions in tumor suppression and cell biology. *Med Sci Monit* **10**, RA235-241.

148.   Sherr CJ, Rettenmier CW, Sacca R, Roussel MF, Look AT & Stanley ER (1985) The c-fms proto-oncogene product is related to the receptor for the mononuclear phagocyte growth factor, CSF-1. *Cell* **41**, 665-676.

149.   Diehl JA (2002) Cycling to cancer with cyclin D1. *Cancer Biol Ther* **1**, 226-231.

150.   Weinstein IB (1996) Relevance of cyclin D1 and other molecular markers to cancer chemoprevention. *J Cell Biochem Suppl* **25**, 23-28.

151.   Aaltomaa S, Eskelinen M & Lipponen P (1999) Expression of cyclin A and D proteins in prostate cancer and their relation to clinopathological variables and patient survival. *Prostate* **38**, 175-182.

152.   Hosokawa Y & Arnold A (1998) Mechanism of cyclin D1 (CCND1, PRAD1) overexpression in human cancer cells: analysis of allele-specific expression. *Genes Chromosomes Cancer* **22**, 66-71.

153.   Hinds PW, Dowdy SF, Eaton EN, Arnold A & Weinberg RA (1994) Function of a human cyclin gene as an oncogene. *Proc Natl Acad Sci U S A* **91**, 709-713.

154.   Lovec H, Sewing A, Lucibello FC, Muller R & Moroy T (1994) Oncogenic activity of cyclin D1 revealed through cooperation with Ha-ras: link between cell cycle control and malignant transformation. *Oncogene* **9**, 323-326.

155.   Grewal SI & Rice JC (2004) Regulation of heterochromatin by histone methylation and small RNAs. *Current opinion in cell biology* **16**, 230-238, doi: 10.1016/j.ceb.2004.04.002.

156.   Glozak MA, Sengupta N, Zhang X & Seto E (2005) Acetylation and deacetylation of non-histone proteins. *Gene* **363**, 15-23, doi: 10.1016/j.gene.2005.09.010.

157.   Hayashi T, Seki M, Maeda D, Wang W, Kawabe Y, Seki T, Saitoh H, Fukagawa T, Yagi H & Enomoto T (2002) Ubc9 is essential for viability of higher eukaryotic cells. *Experimental cell research* **280**, 212-221.

158.   Johnson ES (2004) Protein modification by SUMO. *Annu Rev Biochem* **73**, 355-382.

159.   Geiss-Friedlander R & Melchior F (2007) Concepts in sumoylation: a decade on. *Nat Rev Mol Cell Biol* **8**, 947-956, doi: nrm2293 [pii] 10.1038/nrm2293.

160.   Anckar J & Sistonen L (2007) SUMO: getting it on. *Biochem Soc Trans* **35**, 1409-1413, doi: 10.1042/BST0351409.

161.   Mukhopadhyay D & Dasso M (2007) Modification in reverse: the SUMO proteases. *Trends Biochem Sci* **32**, 286-295, doi: 10.1016/j.tibs.2007.05.002.

162.   Sarge KD & Park-Sarge OK (2011) SUMO and its role in human diseases. *International review of cell and molecular biology* **288**, 167-183, doi: 10.1016/B978-0-12-386041-5.00004-2.

163.   Gareau JR & Lima CD (2010) The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nat Rev Mol Cell Biol* **11**, 861-871, doi: 10.1038/nrm3011.

164.   Ulrich HD (2008) The fast-growing business of SUMO chains. *Mol Cell* **32**, 301-305, doi: 10.1016/j.molcel.2008.10.010.

165.   Zhao J (2007) Sumoylation regulates diverse biological processes. *Cellular and molecular life sciences : CMLS* **64**, 3017-3033, doi: 10.1007/s00018-007-7137-4.

166.   Muller S, Hoege C, Pyrowolakis G & Jentsch S (2001) SUMO, ubiquitin's mysterious cousin. *Nat Rev Mol Cell Biol* **2**, 202-210, doi: 10.1038/35056591.

167.   Tatham MH, Jaffray E, Vaughan OA, Desterro JM, Botting CH, Naismith JH & Hay RT (2001) Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. *J Biol Chem* **276**, 35368-35374, doi: 10.1074/jbc.M104214200.

168.   Kroetz MB (2005) SUMO: a ubiquitin-like protein modifier. *The Yale journal of biology and medicine* **78**, 197-201.

169.   Matic I, van Hagen M, Schimmel J, Macek B, Ogg SC, Tatham MH, Hay RT, Lamond AI, Mann M & Vertegaal AC (2008) In vivo identification of human small ubiquitin-like modifier polymerization sites by high accuracy mass spectrometry and an in vitro to in vivo strategy. *Mol Cell Proteomics* **7**, 132-144, doi: 10.1074/mcp.M700173-MCP200.

170.   Ulrich HD (2009) The SUMO system: an overview. *Methods Mol Biol* **497**, 3-16, doi: 10.1007/978-1-59745-566-4_1.

171.   Bohren KM, Nadkarni V, Song JH, Gabbay KH & Owerbach D (2004) A M55V polymorphism in a novel SUMO gene (SUMO-4) differentially activates heat shock transcription factors and is associated with susceptibility to type I diabetes mellitus. *J Biol Chem* **279**, 27233-27238, doi: 10.1074/jbc.M402273200.

172.   Wilkinson KA & Henley JM (2010) Mechanisms, regulation and consequences of protein SUMOylation. *Biochem J* **428**, 133-145, doi: 10.1042/BJ20100158.

173.   Takahashi Y, Toh-e A & Kikuchi Y (2001) A novel factor required for the SUMO1/Smt3 conjugation of yeast septins. *Gene* **275**, 223-231.

174.   Gong L, Li B, Millas S & Yeh ET (1999) Molecular cloning and characterization of human AOS1 and UBA2, components of the sentrin-activating enzyme complex. *FEBS Lett* **448**, 185-189.

175.   Johnson ES & Blobel G (1997) Ubc9p is the conjugating enzyme for the ubiquitin-like protein Smt3p. *J Biol Chem* **272**, 26799-26802.

176.   Gong L, Kamitani T, Fujise K, Caskey LS & Yeh ET (1997) Preferential interaction of sentrin with a ubiquitin-conjugating enzyme, Ubc9. *J Biol Chem* **272**, 28198-28201.

177.   Desterro JM, Thomson J & Hay RT (1997) Ubch9 conjugates SUMO but not ubiquitin. *FEBS Lett* **417**, 297-300.

178.   Schwarz SE, Matuschewski K, Liakopoulos D, Scheffner M & Jentsch S (1998) The ubiquitin-like proteins SMT3 and SUMO-1 are conjugated by the UBC9 E2 enzyme. *Proc Natl Acad Sci U S A* **95**, 560-564.

179.   Hochstrasser M (2001) SP-RING for SUMO: new functions bloom for a ubiquitin-like protein. *Cell* **107**, 5-8, doi: S0092-8674(01)00519-0 [pii].

180.   Pichler A, Gast A, Seeler JS, Dejean A & Melchior F (2002) The nucleoporin RanBP2 has SUMO1 E3 ligase activity. *Cell* **108**, 109-120.

181.   Kagey MH, Melhuish TA & Wotton D (2003) The polycomb protein Pc2 is a SUMO E3. *Cell* **113**, 127-137.

182.   Melchior F, Schergaut M & Pichler A (2003) SUMO: ligases, isopeptidases and nuclear pores. *Trends Biochem Sci* **28**, 612-618, doi: 10.1016/j.tibs.2003.09.002.

183.   Sternsdorf T, Jensen K, Reich B & Will H (1999) The nuclear dot protein sp100, characterization of domains necessary for dimerization, subcellular localization, and modification by small ubiquitin-like modifiers. *J Biol Chem* **274**, 12555-12566.

184.   Rodriguez MS, Dargemont C & Hay RT (2001) SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. *J Biol Chem* **276**, 12654-12659, doi: 10.1074/jbc.M009476200.

185.   Sampson DA, Wang M & Matunis MJ (2001) The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. *J Biol Chem* **276**, 21664-21669, doi: 10.1074/jbc.M100006200 M100006200 [pii].

186.   Hietakangas V, Anckar J, Blomster HA, Fujimoto M, Palvimo JJ, Nakai A & Sistonen L (2006) PDSM, a motif for phosphorylation-dependent SUMO modification. *Proc Natl Acad Sci U S A* **103**, 45-50, doi: 10.1073/pnas.0503698102.

187.   Adamson AL & Kenney S (2001) Epstein-barr virus immediate-early protein BZLF1 is SUMO-1 modified and disrupts promyelocytic leukemia bodies. *J Virol* **75**, 2388-2399, doi: 10.1128/JVI.75.5.2388-2399.2001.

188.   Kamitani T, Kito K, Nguyen HP, Wada H, Fukuda-Kamitani T & Yeh ET (1998) Identification of three major sentrinization sites in PML. *J Biol Chem* **273**, 26675-26682.

189.   Chakrabarti SR, Sood R, Nandi S & Nucifora G (2000) Posttranslational modification of TEL and TEL/AML1 by SUMO-1 and cell-cycle-dependent assembly into nuclear bodies. *Proc Natl Acad Sci U S A* **97**, 13281-13285, doi: 10.1073/pnas.240315897.

190.   Rangasamy D, Woytek K, Khan SA & Wilson VG (2000) SUMO-1 modification of bovine papillomavirus E1 protein is required for intranuclear accumulation. *J Biol Chem* **275**, 37999-38004, doi: 10.1074/jbc.M007777200.

191.   Nacerddine K, Lehembre F, Bhaumik M, Artus J, Cohen-Tannoudji M, Babinet C, Pandolfi PP & Dejean A (2005) The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Dev Cell* **9**, 769-779, doi: 10.1016/j.devcel.2005.10.007.

192.   Bettermann K, Benesch M, Weis S & Haybaeck J (2012) SUMOylation in carcinogenesis. *Cancer Lett* **316**, 113-125, doi: 10.1016/j.canlet.2011.10.036.

193.   Seeler JS, Bischof O, Nacerddine K & Dejean A (2007) SUMO, the three Rs and cancer. *Curr Top Microbiol Immunol* **313**, 49-71.

194.   Bischof O & Dejean A (2007) SUMO is growing senescent. *Cell Cycle* **6**, 677-681.

195.   Meinecke I, Cinski A, Baier A, Peters MA, Dankbar B, Wille A, Drynda A, Mendoza H, Gay RE, Hay RT, et al. (2007) Modification of nuclear PML protein by SUMO-1 regulates Fas-induced apoptosis in rheumatoid arthritis synovial fibroblasts. *Proc Natl Acad Sci U S A* **104**, 5073-5078, doi: 10.1073/pnas.0608773104.

196.   McDoniels-Silvers AL, Nimri CF, Stoner GD, Lubet RA & You M (2002) Differential gene expression in human lung adenocarcinomas and squamous cell carcinomas. *Clin Cancer Res* **8**, 1127-1138.

197.   Mo YY, Yu Y, Theodosiou E, Ee PL & Beck WT (2005) A role for Ubc9 in tumorigenesis. *Oncogene* **24**, 2677-2683, doi: 10.1038/sj.onc.1208210.

198.   Lee JS & Thorgeirsson SS (2004) Genome-scale profiling of gene expression in hepatocellular carcinoma: classification, survival prediction, and identification of therapeutic targets. *Gastroenterology* **127**, S51-55.

199.   Mo YY & Moschos SJ (2005) Targeting Ubc9 for cancer therapy. *Expert Opin Ther Targets* **9**, 1203-1216, doi: 10.1517/14728222.9.6.1203.

200.   Martin SF, Hattersley N, Samuel ID, Hay RT & Tatham MH (2007) A fluorescence-resonance-energy-transfer-based protease activity assay and its use to monitor paralog-specific small ubiquitin-like modifier processing. *Anal Biochem* **363**, 83-90, doi: 10.1016/j.ab.2006.12.018.

201.   Vassileva MT & Matunis MJ (2004) SUMO modification of heterogeneous nuclear ribonucleoproteins. *Mol Cell Biol* **24**, 3623-3632.

202.   Babic I, Cherry E & Fujita DJ (2006) SUMO modification of Sam68 enhances its ability to repress cyclin D1 expression and inhibits its ability to induce apoptosis. *Oncogene* **25**, 4955-4964, doi: 10.1038/sj.onc.1209504.

203.   Arhin GK, Boots M, Bagga PS, Milcarek C & Wilusz J (2002) Downstream sequence elements with different affinities for the hnRNP H/H' protein influence the processing efficiency of mammalian polyadenylation signals. *Nucleic Acids Res* **30**, 1842-1850.

204.   Veraldi KL, Arhin GK, Martincic K, Chung-Ganster LH, Wilusz J & Milcarek C (2001) hnRNP F influences binding of a 64-kilodalton subunit of cleavage stimulation factor to mRNA precursors in mouse B cells. *Mol Cell Biol* **21**, 1228-1238, doi: 10.1128/MCB.21.4.1228-1238.2001.

205.	Konig J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM & Ule J (2010) iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat Struct Mol Biol* **17**, 909-915, doi: 10.1038/nsmb.1838.

206.	Eversole A & Maizels N (2000) In vitro properties of the conserved mammalian protein hnRNP D suggest a role in telomere maintenance. *Mol Cell Biol* **20**, 5425-5432.

207.	Ford LP, Suh JM, Wright WE & Shay JW (2000) Heterogeneous nuclear ribonucleoproteins C1 and C2 associate with the RNA component of human telomerase. *Mol Cell Biol* **20**, 9084-9091.

208.	Liu K, Li L, Nisson PE, Gruber C, Jessee J & Cohen SN (2000) Neoplastic transformation and tumorigenesis associated with sam68 protein deficiency in cultured murine fibroblasts. *J Biol Chem* **275**, 40195-40201, doi: 10.1074/jbc.M006194200.

209.	Taylor SJ, Resnick RJ & Shalloway D (2004) Sam68 exerts separable effects on cell cycle progression and apoptosis. *BMC Cell Biol* **5**, 5, doi: 10.1186/1471-2121-5-5.

210.	Desterro JM, Keegan LP, Jaffray E, Hay RT, O'Connell MA & Carmo-Fonseca M (2005) SUMO-1 modification alters ADAR1 editing activity. *Mol Biol Cell* **16**, 5115-5126, doi: 10.1091/mbc.E05-06-0536.

211.	Xu X, Vatsyayan J, Gao C, Bakkenist CJ & Hu J (2010) Sumoylation of eIF4E activates mRNA translation. *EMBO Rep* **11**, 299-304, doi: 10.1038/embor.2010.18.

212.	Vethantham V, Rao N & Manley JL (2007) Sumoylation modulates the assembly and activity of the pre-mRNA 3' processing complex. *Mol Cell Biol* **27**, 8848-8858, doi: MCB.01186-07 [pii]
10.1128/MCB.01186-07.

213.	Jo OD, Martin J, Bernath A, Masri J, Lichtenstein A & Gera J (2008) Heterogeneous nuclear ribonucleoprotein A1 regulates cyclin D1 and c-myc internal ribosome entry site function through Akt signaling. *J Biol Chem* **283**, 23274-23287, doi: M801185200 [pii]
10.1074/jbc.M801185200.

214.	Kozak M (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene* **234**, 187-208.

215.	Kozak M (1997) Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6. *Embo J* **16**, 2482-2492, doi: 10.1093/emboj/16.9.2482.

216.	Xu H, Wang P, Fu Y, Zheng Y, Tang Q, Si L, You J, Zhang Z, Zhu Y, Zhou L, et al. (2010) Length of the ORF, position of the first AUG and the Kozak motif are important factors in potential dual-coding transcripts. *Cell Res* **20**, 445-457, doi: 10.1038/cr.2010.25.

217.	Alfano C, Sanfelice D, Babon J, Kelly G, Jacks A, Curry S & Conte MR (2004) Structural analysis of cooperative RNA binding by the La motif and central RRM domain of human La protein. *Nat Struct Mol Biol* **11**, 323-329.

218.	Kenan DJ & Keene JD (2004) La gets its wings. *Nat Struct Mol Biol* **11**, 303-305.

219.	Royer CA & Scarlata SF (2008) Fluorescence approaches to quantifying biomolecular interactions. *Methods Enzymol* **450**, 79-106, doi: 10.1016/S0076-6879(08)03405-8.

220.	Pagano JM, Clingman CC & Ryder SP (2011) Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. *Rna* **17**, 14-20, doi: 10.1261/rna.2428111.

221.	Enright C & Sollner-Webb D (1994) Ribosomal RNA processing in vertebrates: analysis of protein-RNA associations in rRNA processing. In *RNA Processing: A practical approach* (Higgins SJ & Hames BD, eds), pp. 155-164. Oxford University press, Oxford.

222.	Wilson GM & Brewer G (1999) The search for trans-acting factors controlling messenger RNA decay. *Prog Nucleic Acid Res Mol Biol* **62**, 257-291.

223.   Thomson AM, Rogers JT, Walker CE, Staton JM & Leedman PJ (1999) Optimized RNA gel-shift and UV cross-linking assays for characterization of cytoplasmic RNA-protein interactions. *Biotechniques* **27**, 1032-1039, 1042.

224.   Kucera NJ, Hodsdon ME & Wolin SL (2011) An intrinsically disordered C terminus allows the La protein to assist the biogenesis of diverse noncoding RNA precursors. *Proc Natl Acad Sci U S A* **108**, 1308-1313, doi: 10.1073/pnas.1017085108.

225.   Bencsath KP, Podgorski MS, Pagala VR, Slaughter CA & Schulman BA (2002) Identification of a multifunctional binding site on Ubc9p required for Smt3p conjugation. *J Biol Chem* **277**, 47938-47945, doi: 10.1074/jbc.M207442200
M207442200 [pii].

226.   Zhong N, Kim CY, Rizzu P, Geula C, Porter DR, Pothos EN, Squitieri F, Heutink P & Xu J (2006) DJ-1 transcriptionally up-regulates the human tyrosine hydroxylase by inhibiting the sumoylation of pyrimidine tract-binding protein-associated splicing factor. *J Biol Chem* **281**, 20940-20948, doi: 10.1074/jbc.M601935200.

227.   Kroetz MB & Hochstrasser M (2009) Identification of SUMO-interacting proteins by yeast two-hybrid analysis. *Methods Mol Biol* **497**, 107-120, doi: 10.1007/978-1-59745-566-4_7.

228.   Poukka H, Karvonen U, Janne OA & Palvimo JJ (2000) Covalent modification of the androgen receptor by small ubiquitin-like modifier 1 (SUMO-1). *Proc Natl Acad Sci U S A* **97**, 14145-14150, doi: 10.1073/pnas.97.26.14145
97/26/14145 [pii].

229.   Chen WY, Lee WC, Hsu NC, Huang F & Chung BC (2004) SUMO modification of repression domains modulates function of nuclear receptor 5A1 (steroidogenic factor-1). *J Biol Chem* **279**, 38730-38735, doi: 10.1074/jbc.M405006200
M405006200 [pii].

230.   Wu SY & Chiang CM (2009) Crosstalk between sumoylation and acetylation regulates p53-dependent chromatin transcription and DNA binding. *Embo J* **28**, 1246-1259, doi: emboj200983 [pii]
10.1038/emboj.2009.83.

231.   Pichler A, Knipscheer P, Saitoh H, Sixma TK & Melchior F (2004) The RanBP2 SUMO E3 ligase is neither HECT- nor RING-type. *Nat Struct Mol Biol* **11**, 984-991, doi: 10.1038/nsmb834
nsmb834 [pii].

232.   Gocke CB, Yu H & Kang J (2005) Systematic identification and analysis of mammalian small ubiquitin-like modifier substrates. *J Biol Chem* **280**, 5004-5012, doi: 10.1074/jbc.M411718200.

233.   Hay RT (2005) SUMO: a history of modification. *Mol Cell* **18**, 1-12, doi: S1097-2765(05)01182-2 [pii]
10.1016/j.molcel.2005.03.012.

234.   Rangasamy D & Wilson VG (2000) Bovine papillomavirus E1 protein is sumoylated by the host cell Ubc9 protein. *J Biol Chem* **275**, 30487-30495, doi: 10.1074/jbc.M003898200
M003898200 [pii].

235.   Tian S, Poukka H, Palvimo JJ & Janne OA (2002) Small ubiquitin-related modifier-1 (SUMO-1) modification of the glucocorticoid receptor. *Biochem J* **367**, 907-911, doi: 10.1042/BJ20021085.

236.   Le Drean Y, Mincheneau N, Le Goff P & Michel D (2002) Potentiation of glucocorticoid receptor transcriptional activity by sumoylation. *Endocrinology* **143**, 3482-3489.

237.   Comerford KM, Leonard MO, Karhausen J, Carey R, Colgan SP & Taylor CT (2003) Small ubiquitin-related modifier-1 modification mediates resolution of CREB-dependent

responses to hypoxia. *Proc Natl Acad Sci U S A* **100**, 986-991, doi: 10.1073/pnas.0337412100.

238.   Ren J, Gao X, Jin C, Zhu M, Wang X, Shaw A, Wen L, Yao X & Xue Y (2009) Systematic study of protein sumoylation: Development of a site-specific predictor of SUMOsp 2.0. *Proteomics* **9**, 3409-3412, doi: 10.1002/pmic.200800646.

239.   Galisson F, Mahrouche L, Courcelles M, Bonneil E, Meloche S, Chelbi-Alix MK & Thibault P (2011) A novel proteomics approach to identify SUMOylated proteins and their modification sites in human cells. *Mol Cell Proteomics* **10**, M110 004796, doi: 10.1074/mcp.M110.004796.

240.   Smith PR, Williams DG, Venables PJ & Maini RN (1985) Monoclonal antibodies to the Sjogren's syndrome associated antigen SS-B (La). *Journal of Immunological Methods* **77**, 63-76.

241.   Schlatter S, Senn C & Fussenegger M (2003) Modulation of translation-initiation in CHO-K1 cells by rapamycin-induced heterodimerization of engineered eIF4G fusion proteins. *Biotechnol Bioeng* **83**, 210-225, doi: 10.1002/bit.10662.

242.   Zhu J, Hayakawa A, Kakegawa T & Kaspar RL (2001) Binding of the La autoantigen to the 5' untranslated region of a chimeric human translation elongation factor 1A reporter mRNA inhibits translation in vitro. *Biochim Biophys Acta* **1521**, 19-29.

243.   Shama S, Avni D, Frederickson RM, Sonenberg N & Meyuhas O (1995) Overexpression of initiation factor eIF-4E does not relieve the translational repression of ribosomal protein mRNAs in quiescent cells. *Gene Expr* **4**, 241-252.

244.   Iadevaia V, Caldarola S, Tino E, Amaldi F & Loreni F (2008) All translation elongation factors and the e, f, and h subunits of translation initiation factor 3 are encoded by 5'-terminal oligopyrimidine (TOP) mRNAs. *Rna* **14**, 1730-1736, doi: 10.1261/rna.1037108.

245.   Willis DE, van Niekerk EA, Sasaki Y, Mesngon M, Merianda TT, Williams GG, Kendall M, Smith DS, Bassell GJ & Twiss JL (2007) Extracellular stimuli specifically regulate localized levels of individual neuronal mRNAs. *J Cell Biol* **178**, 965-980, doi: 10.1083/jcb.200703209.

246.   Svitkin YV, Meerovitch K, Lee HS, Dholakia JN, Kenan DJ, Agol VI & Sonenberg N (1994) Internal translation initiation on poliovirus RNA: further characterization of La function in poliovirus translation in vitro. *J Virol* **68**, 1544-1550.

247.   Holcik M & Sonenberg N (2005) Translational control in stress and apoptosis. *Nature Reviews Molecular Cell Biology*
Nat Rev Mol Cell Biol **6**, 318-327.

248.   Xia Z, Knaak C, Ma J, Beharry ZM, McInnes C, Wang W, Kraft AS & Smith CD (2009) Synthesis and evaluation of novel inhibitors of Pim-1 and Pim-2 protein kinases. *J Med Chem* **52**, 74-86.

249.   Peek R, Pruijn GJ & Van Venrooij WJ (1996) Interaction of the La (SS-B) autoantigen with small ribosomal subunits. *Eur J Biochem* **236**, 649-655.

250.   Zhang J, Dinh TN, Kappeler K, Tsaprailis G & Chen QM (2012) La autoantigen mediates oxidant induced de novo nrf2 protein translation. *Mol Cell Proteomics* **11**, M111 015032, doi: 10.1074/mcp.M111.015032.

251.   Hinnebusch AG (2011) Molecular mechanism of scanning and start codon selection in eukaryotes. *Microbiol Mol Biol Rev* **75**, 434-467, first page of table of contents, doi: 10.1128/MMBR.00008-11.

252.   Hallay H, Locker N, Ayadi L, Ropers D, Guittet E & Branlant C (2006) Biochemical and NMR study on the competition between proteins SC35, SRp40, and heterogeneous nuclear ribonucleoprotein A1 at the HIV-1 Tat exon 2 splicing site. *J Biol Chem* **281**, 37159-37174, doi: 10.1074/jbc.M603864200.

253. Kalifa Y, Huang T, Rosen LN, Chatterjee S & Gavis ER (2006) Glorund, a Drosophila hnRNP F/H homolog, is an ovarian repressor of nanos translation. *Dev Cell* **10**, 291-301, doi: 10.1016/j.devcel.2006.01.001.

254. Mitchell SA, Spriggs KA, Bushell M, Evans JR, Stoneley M, Le Quesne JP, Spriggs RV & Willis AE (2005) Identification of a motif that mediates polypyrimidine tract-binding protein-dependent internal ribosome entry. *Genes Dev* **19**, 1556-1571, doi: 10.1101/gad.339105.

255. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**, 3406-3415.

256. Ding J, Hayashi MK, Zhang Y, Manche L, Krainer AR & Xu RM (1999) Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. *Genes Dev* **13**, 1102-1115.

257. Tsai DE, Harper DS & Keene JD (1991) U1-snRNP-A protein selects a ten nucleotide consensus sequence from a degenerate RNA pool presented in various structural contexts. *Nucleic Acids Res* **19**, 4931-4936.

258. Law MJ, Rice AJ, Lin P & Laird-Offringa IA (2006) The role of RNA structure in the interaction of U1A protein with U1 hairpin II RNA. *Rna* **12**, 1168-1178, doi: 10.1261/rna.75206.

259. Nashimoto M, Nashimoto C, Tamura M, Kaspar RL & Ochi K (2001) The inhibitory effect of the autoantigen La on in vitro 3' processing of mammalian precursor tRNAs. *J Mol Biol* **312**, 975-984.

260. Wang X & Tanaka Hall TM (2001) Structural basis for recognition of AU-rich element RNA by the HuD protein. *Nat Struct Biol* **8**, 141-145, doi: 10.1038/84131.

261. Kenan DJ, Query CC & Keene JD (1991) RNA recognition: towards identifying determinants of specificity. *Trends Biochem Sci* **16**, 214-220.

262. DeMaria CT & Brewer G (1996) AUF1 binding affinity to A+U-rich elements correlates with rapid mRNA degradation. *J Biol Chem* **271**, 12179-12184.

263. Deo RC, Bonanno JB, Sonenberg N & Burley SK (1999) Recognition of polyadenylate RNA by the poly(A)-binding protein. *Cell* **98**, 835-845.

264. Allain FH, Gilbert DE, Bouvet P & Feigon J (2000) Solution structure of the two N-terminal RNA-binding domains of nucleolin and NMR study of the interaction with its RNA target. *J Mol Biol* **303**, 227-241.

265. Johansson C, Finger LD, Trantirek L, Mueller TD, Kim S, Laird-Offringa IA & Feigon J (2004) Solution structure of the complex formed by the two N-terminal RNA-binding domains of nucleolin and a pre-rRNA target. *J Mol Biol* **337**, 799-816, doi: 10.1016/j.jmb.2004.01.056.

266. Perez-Canadillas JM (2006) Grabbing the message: structural basis of mRNA 3'UTR recognition by Hrp1. *Embo J* **25**, 3167-3178, doi: 10.1038/sj.emboj.7601190.

267. Handa N, Nureki O, Kurimoto K, Kim I, Sakamoto H, Shimura Y, Muto Y & Yokoyama S (1999) Structural basis for recognition of the tra mRNA precursor by the Sex-lethal protein. *Nature* **398**, 579-585, doi: 10.1038/19242.

268. Maraia RJ (1996) Transcription termination factor La is also an initiation factor for RNA polymerase III. *Proc Natl Acad Sci U S A* **93**, 3383-3387.

269. Bayfield MA, Yang R & Maraia RJ (2010) Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochim Biophys Acta* **1799**, 365-378, doi: 10.1016/j.bbagrm.2010.01.011.

270. Groban ES, Narayanan A & Jacobson MP (2006) Conformational changes in protein loops and helices induced by post-translational phosphorylation. *PLoS Comput Biol* **2**, e32, doi: 10.1371/journal.pcbi.0020032.

271.   Latzer J, Shen T & Wolynes PG (2008) Conformational switching upon phosphorylation: a predictive framework based on energy landscape principles. *Biochemistry* **47**, 2110-2122, doi: 10.1021/bi701350v.

272.   Gauci S, Helbig AO, Slijper M, Krijgsveld J, Heck AJ & Mohammed S (2009) Lys-N and trypsin cover complementary parts of the phosphoproteome in a refined SCX-based approach. *Anal Chem* **81**, 4493-4501, doi: 10.1021/ac9004309.

273.   Hong W, Resnick RJ, Rakowski C, Shalloway D, Taylor SJ & Blobel GA (2002) Physical and functional interaction between the transcriptional cofactor CBP and the KH domain protein Sam68. *Mol Cancer Res* **1**, 48-55.

274.   Babic I, Jakymiw A & Fujita DJ (2004) The RNA binding protein Sam68 is acetylated in tumor cell lines, and its acetylation correlates with enhanced RNA binding activity. *Oncogene* **23**, 3781-3789, doi: 10.1038/sj.onc.1207484.

275.   Culjkovic B, Topisirovic I, Skrabanek L, Ruiz-Gutierrez M & Borden KL (2005) eIF4E promotes nuclear export of cyclin D1 mRNAs via an element in the 3'UTR. *J Cell Biol* **169**, 245-256, doi: 10.1083/jcb.200501019.

276.   Matunis MJ, Wu J & Blobel G (1998) SUMO-1 modification and its role in targeting the Ran GTPase-activating protein, RanGAP1, to the nuclear pore complex. *J Cell Biol* **140**, 499-509.

277.   Mahajan R, Delphin C, Guan T, Gerace L & Melchior F (1997) A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell* **88**, 97-107, doi: S0092-8674(00)81862-0 [pii].

278.   Matunis MJ, Coutavas E & Blobel G (1996) A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *J Cell Biol* **135**, 1457-1470.

279.   Vertegaal AC, Andersen JS, Ogg SC, Hay RT, Mann M & Lamond AI (2006) Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics. *Mol Cell Proteomics* **5**, 2298-2310, doi: 10.1074/mcp.M600212-MCP200.

280.   Saitoh H & Hinchey J (2000) Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. *J Biol Chem* **275**, 6252-6258.

281.   Ulrich HD (2005) Mutual interactions between the SUMO and ubiquitin systems: a plea of no contest. *Trends Cell Biol* **15**, 525-532, doi: 10.1016/j.tcb.2005.08.002.

282.   Desterro JM, Rodriguez MS & Hay RT (1998) SUMO-1 modification of IkappaBalpha inhibits NF-kappaB activation. *Mol Cell* **2**, 233-239.

283.   Hang J & Dasso M (2002) Association of the human SUMO-1 protease SENP2 with the nuclear pore. *J Biol Chem* **277**, 19961-19966, doi: 10.1074/jbc.M201799200.

284.   Zhang H, Saitoh H & Matunis MJ (2002) Enzymes of the SUMO modification pathway localize to filaments of the nuclear pore complex. *Mol Cell Biol* **22**, 6498-6508.

285.   Wood LD, Irvin BJ, Nucifora G, Luce KS & Hiebert SW (2003) Small ubiquitin-like modifier conjugation regulates nuclear export of TEL, a putative tumor suppressor. *Proc Natl Acad Sci U S A* **100**, 3257-3262, doi: 10.1073/pnas.0637114100.

286.   Sobko A, Ma H & Firtel RA (2002) Regulated SUMOylation and ubiquitination of DdMEK1 is required for proper chemotaxis. *Dev Cell* **2**, 745-756.

287.   Sramko M, Markus J, Kabat J, Wolff L & Bies J (2006) Stress-induced inactivation of the c-Myb transcription factor through conjugation of SUMO-2/3 proteins. *J Biol Chem* **281**, 40065-40075, doi: 10.1074/jbc.M609404200.

288.   Yang W, Sheng H, Warner DS & Paschen W (2008) Transient focal cerebral ischemia induces a dramatic activation of small ubiquitin-like modifier conjugation. *J Cereb Blood Flow Metab* **28**, 892-896, doi: 10.1038/sj.jcbfm.9600601.

289.   Loftus LT, Gala R, Yang T, Jessick VJ, Ashley MD, Ordonez AN, Thompson SJ, Simon RP & Meller R (2009) Sumo-2/3-ylation following in vitro modeled ischemia is reduced in delayed ischemic tolerance. *Brain Res* **1272**, 71-80, doi: 10.1016/j.brainres.2009.03.034.

290.   Huang C, Han Y, Wang Y, Sun X, Yan S, Yeh ET, Chen Y, Cang H, Li H, Shi G, et al. (2009) SENP3 is responsible for HIF-1 transactivation under mild oxidative stress via p300 de-SUMOylation. *Embo J* **28**, 2748-2762, doi: 10.1038/emboj.2009.210.

291.   Hamilton TL, Stoneley M, Spriggs KA & Bushell M (2006) TOPs and their regulation. *Biochem Soc Trans* **34**, 12-16, doi: 10.1042/BST20060012.

292.   Tang H, Hornstein E, Stolovich M, Levy G, Livingstone M, Templeton D, Avruch J & Meyuhas O (2001) Amino acid-induced translation of TOP mRNAs is fully dependent on phosphatidylinositol 3-kinase-mediated signaling, is partially inhibited by rapamycin, and is independent of S6K1 and rpS6 phosphorylation. *Mol Cell Biol* **21**, 8671-8683, doi: 10.1128/MCB.21.24.8671-8683.2001.

293.   Stolovich M, Tang H, Hornstein E, Levy G, Cohen R, Bae SS, Birnbaum MJ & Meyuhas O (2002) Transduction of growth or mitogenic signals into translational activation of TOP mRNAs is fully reliant on the phosphatidylinositol 3-kinase-mediated pathway but requires neither S6K1 nor rpS6 phosphorylation. *Mol Cell Biol* **22**, 8101-8113.

294.   Avni D, Shama S, Loreni F & Meyuhas O (1994) Vertebrate mRNAs with a 5'-terminal pyrimidine tract are candidates for translational repression in quiescent cells: characterization of the translational cis-regulatory element. *Mol Cell Biol* **14**, 3822-3833.

295.   Davuluri RV, Suzuki Y, Sugano S & Zhang MQ (2000) CART classification of human 5' UTR sequences. *Genome Res* **10**, 1807-1816.

296.   Biberman Y & Meyuhas O (1999) TOP mRNAs are translationally inhibited by a titratable repressor in both wheat germ extract and reticulocyte lysate. *FEBS Lett* **456**, 357-360.

297.   Pellizzoni L, Lotti F, Maras B & Pierandrei-Amaldi P (1997) Cellular nucleic acid binding protein binds a conserved region of the 5' UTR of Xenopus laevis ribosomal protein mRNAs. *J Mol Biol* **267**, 264-275, doi: S0022-2836(96)90888-8 [pii] 10.1006/jmbi.1996.0888.

298.   Zhang Y, Wolf GW, Bhat K, Jin A, Allio T, Burkhart WA & Xiong Y (2003) Ribosomal protein L11 negatively regulates oncoprotein MDM2 and mediates a p53-dependent ribosomal-stress checkpoint pathway. *Mol Cell Biol* **23**, 8902-8912.

299.   Dai MS & Lu H (2004) Inhibition of MDM2-mediated p53 ubiquitination and degradation by ribosomal protein L5. *J Biol Chem* **279**, 44475-44482, doi: 10.1074/jbc.M403722200.

300.   Lung JC, Chu JS, Yu JC, Yue CT, Lo YL, Shen CY & Wu CW (2002) Aberrant expression of cell-cycle regulator cyclin D1 in breast cancer is related to chromosomal genomic instability. *Genes Chromosomes Cancer* **34**, 276-284, doi: 10.1002/gcc.10072.

301.   Guttridge DC, Albanese C, Reuther JY, Pestell RG & Baldwin AS, Jr. (1999) NF-kappaB controls cell growth and differentiation through transcriptional regulation of cyclin D1. *Mol Cell Biol* **19**, 5785-5799.

302.   Westerheide SD, Mayo MW, Anest V, Hanson JL & Baldwin AS, Jr. (2001) The putative oncoprotein Bcl-3 induces cyclin D1 to stimulate G(1) transition. *Mol Cell Biol* **21**, 8428-8436, doi: 10.1128/MCB.21.24.8428-8436.2001.

303.   Hinz M, Krappmann D, Eichten A, Heder A, Scheidereit C & Strauss M (1999) NF-kappaB function in growth control: regulation of cyclin D1 expression and G0/G1-to-S-phase transition. *Mol Cell Biol* **19**, 2690-2698.

304.   Palmero I & Peters G (1996) Perturbation of cell cycle regulators in human cancer. *Cancer Surv* **27**, 351-367.

305.   Testa JR & Tsichlis PN (2005) AKT signaling in normal and malignant cells. *Oncogene* **24**, 7391-7393, doi: 10.1038/sj.onc.1209100.

306.   Kim KI & Baek SH (2006) SUMOylation code in cancer development and metastasis. *Mol Cells* **22**, 247-253.

307.   Muller S, Matunis MJ & Dejean A (1998) Conjugation with the ubiquitin-related modifier SUMO-1 regulates the partitioning of PML within the nucleus. *Embo J* **17**, 61-70, doi: 10.1093/emboj/17.1.61.

308.   Yang SH, Jaffray E, Senthinathan B, Hay RT & Sharrocks AD (2003) SUMO and transcriptional repression: dynamic interactions between the MAP kinase and SUMO pathways. *Cell Cycle* **2**, 528-530.

309.   Hietakangas V, Ahlskog JK, Jakobsson AM, Hellesuo M, Sahlberg NM, Holmberg CI, Mikhailov A, Palvimo JJ, Pirkkala L & Sistonen L (2003) Phosphorylation of serine 303 is a prerequisite for the stress-inducible SUMO modification of heat shock factor 1. *Mol Cell Biol* **23**, 2953-2968.

310.   Ivanov P, Kedersha N & Anderson P (2011) Stress puts TIA on TOP. *Genes Dev* **25**, 2119-2124, doi: 10.1101/gad.17838411.

311.   Berglund JA, Abovich N & Rosbash M (1998) A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes Dev* **12**, 858-867.

312.   Peck ML & Herschlag D (2003) Adenosine 5'-O-(3-thio)triphosphate (ATPgammaS) is a substrate for the nucleotide hydrolysis and RNA unwinding activities of eukaryotic translation initiation factor eIF4A. *RNA* **9**, 1180-1187.

313.   Huhn P, Pruijn GJ, van Venrooij WJ & Bachmann M (1997) Characterization of the autoantigen La (SS-B) as a dsRNA unwinding enzyme. *Nucleic Acids Res* **25**, 410-416.

314.   Meulmeester E & Melchior F (2008) Cell biology: SUMO. *Nature* **452**, 709-711, doi: 10.1038/452709a.

315.   Marblestone JG, Edavettal SC, Lim Y, Lim P, Zuo X & Butt TR (2006) Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO. *Protein Sci* **15**, 182-189, doi: 10.1110/ps.051812706.

316.   Peroutka I RJ, Orcutt SJ, Strickler JE & Butt TR (2011) SUMO fusion technology for enhanced protein expression and purification in prokaryotes and eukaryotes. *Methods Mol Biol* **705**, 15-30, doi: 10.1007/978-1-61737-967-3_2.

317.   Bradford, MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry* **72,** 248-54

318.   Laemmli, UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227,** 680-5

319.   Meyer, TS, Lamberts, BL (1965) Use of coomassie brilliant blue R250 for the electrophoresis of microgram quantities of parotid saliva proteins on acrylamide-gel strips. *Biochimica et biophysica acta* **107,** 144-5

320.   Braman, J, Papworth, C, Greener, A (1996) Site-directed mutagenesis using double-stranded plasmid DNA templates. *Methods in Molecular Biology* **57,** 31-44

321.   Dorval, V., Fraser PE (2007) SUMO on the road to neurodegeneration. *Biochimica et biophysica acta* **1773**, 694-706; doi:10.1016/j.bbamcr.2007.03.017

322*.*   Ren, J,Gao, X, Jin, C, Zhu, M, Wang, X, Shaw, A, Wen L. Yao, X, Xue (2009) Systematic study of protein sumoylation: Development of a site-specific predictor of SUMOsp 2.0. Proteomics **9**, 3409-3412; doi: 10.1002/pmic.200800646

Appendix

human La amino acid sequence (GenBank reference P05455.2)

[1] MAENG DNEKM AALEA KICHQ IEYYF GDFNL PRDKF LKEQI
[41] KLDEG WVPLE IMIKF NRLNR LTTDF NVIVE ALSKS KAELM
[81] EISED KTKIR RSPSK PLPEV TDEYK NDVKN RSVYI KGFPT
[121] DATLD DIKEW LEDKG QVLNI QMRRT LHKAF KGSIF VVFDS
[161] IESAK KFVET PGQKY KETDL LILFK DDYFA KKNEE RKQNK
[201] VEAKL RAKQE QEAKQ KLEED AEMKS LEEKI GCLLK FSGDL
[241] DDQTC REDLH ILFSN HGEIK WIDFV RGAKE GIILF KEKAK
[281] EALGK AKDAN NGNLQ LRNKE VTWEV LEGEV EKEAL KKIIE
[321] DQQES LNKWK SKGRR FKGKG KGNKA AQPGS GKGKV QFQGK
[361] KTKFA SDDEH DEHDE NGATG PVKRA REETD KEEPA SKQQK
[401] TENGA GDQ

vector map of pet28b(+) (Novagen)
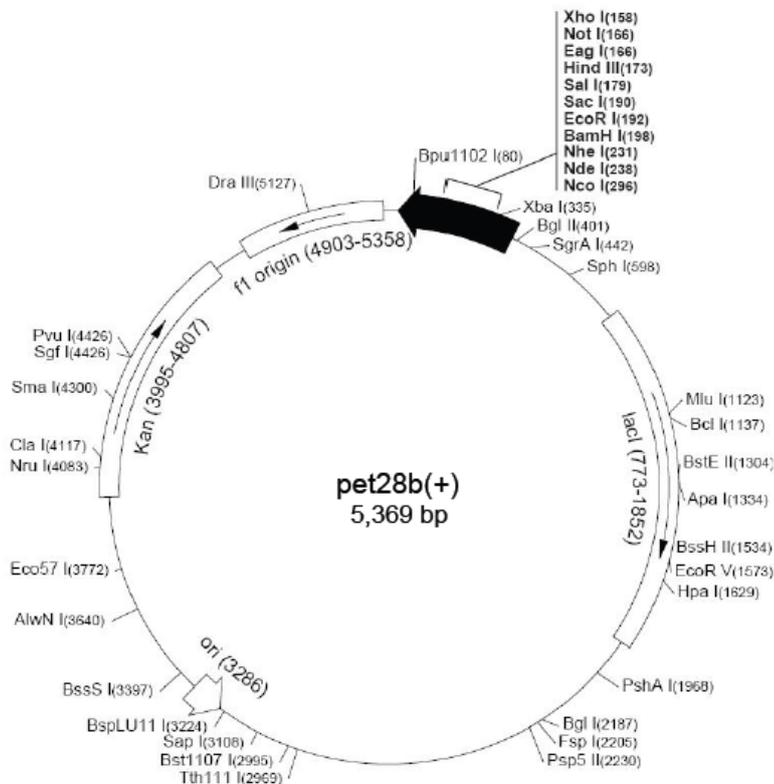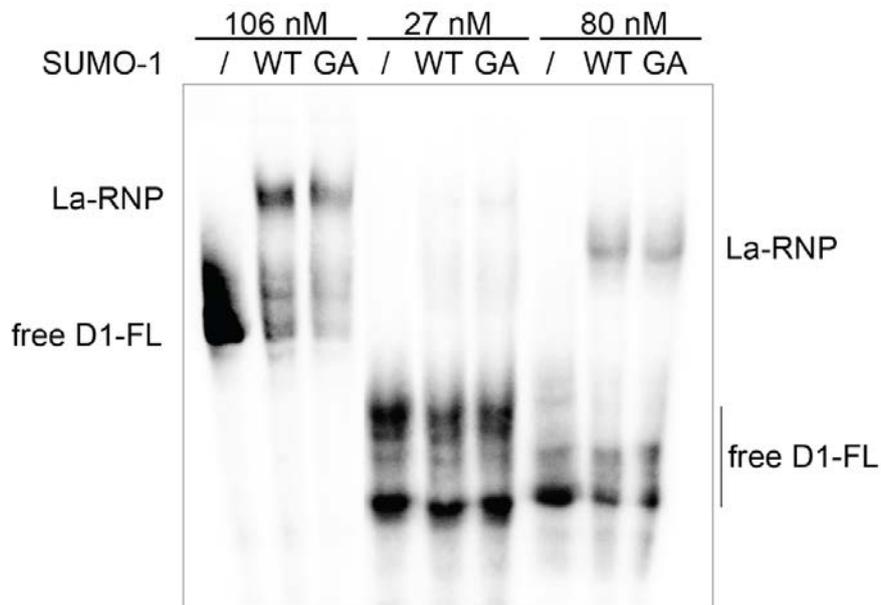
Figure A1: Modification of La with SUMO-1 results in an increased binding activity to cyclin D1 RNA. A standard IVSA was performed and products were used for EMSA. The utilization of SUMO-1WT or SUMO-1GA in the IVSA is indicated. Representative EMSA in the presence of 0.5 µg heparin. The La-RNP complex and free RNA are indicated.
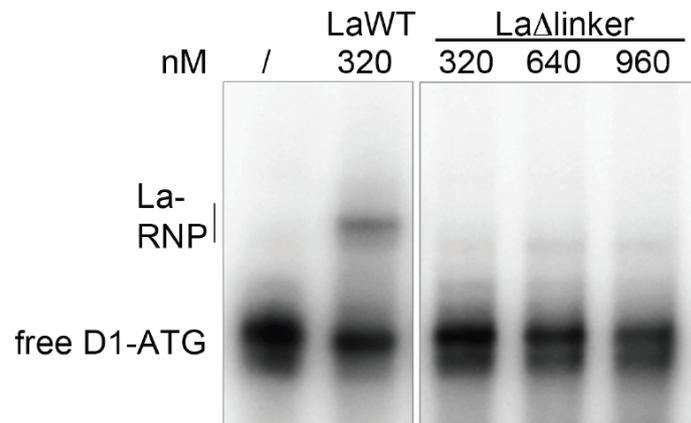
Figure A2: Cyclin D1 RNA-binding is abrogated upon deletion of the inter-RRM peptide linker. The 60 amino acid inter-RRM linker (LaΔlinker) was deleted between aa 170-230. Standard EMSA with D1-FL and the indicated protein concentrations were performed. La-RNP formation is abrogated upon deletion of the linker region.
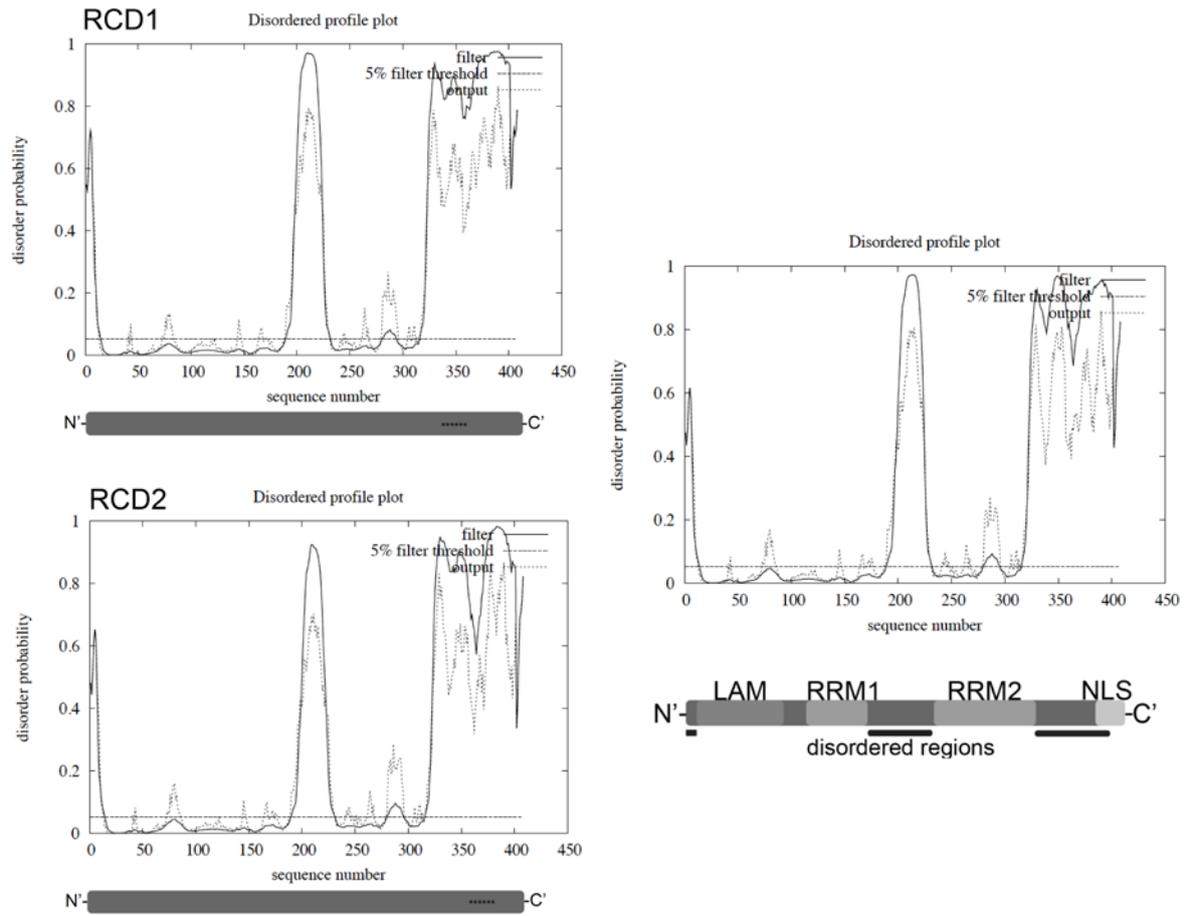
Figure A3: Disopred profile plot of hLaWT and its respective mutants RCD1 and RCD2.

Mt SybUC Reverse 23-Base Sequencing Primer

5'- AGGCGGA TAACAATTTC ACACAGB-3'

5'- AGCGGGA TAACAATTTC ACACAGGAAA CAGACC ATG TCG TAC TAG CAT CAC CAT CAC CAT TAC GAT GAT ATC CCA AGG ACC GAA AAC CTG TAT TTT CAG** GGC GCC ATG GGA TCC GGA ATT CAA AGG GGT CAA AGG TCG AGG AGC TCA CTA GTC

met ser tyr tyr his his his his his his tyr asp asp ile pro arg thr thr glu asn leu tyr phe gln gly ala met gly ser gly ile gln arg pro gln arg ser thr ser ser leu val

RBS   (His)₆   spacer region   rTEV protease cleavage site

Xma III

Not I   Nsp V   Xba I   Pst I   Xho I   Sph I*   Kpn I   Hind III

GCG GCC GCT TTC GAA TCT AGA DCC TGC AGT CTC GAG GCA TGC GGT ACC AAG CTT GGC TGT TTT GGC GGA TGA GAGA-3'
ala ala phe glu ser arg ala cys ser leu glu ala cys gly thr lys leu gly cys phe phe gly gly stop

* Sph I has 2 recognition sites in pPROEX HT.

pT7Htb description

size ~ 3.9 kbp
6xHis tag N-terminus
TEV protease cleavage site
Kanamycin resistance

PT7 vectors (example: N-terminal FLAG)

| RBS |─| ATG |─| FLAG |─| MCS |─| STOP | PT7
         start   N-term   protein
                  HisC.TEV

5'   3'

MCS: 5'→ Ncol/BamHI/EcoRI/Xbal/EcoRI/Xbal/PstI/KpnI/HindⅢ→3'

Eidesstattliche Versicherung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Danksagung –Acknowledgments

Mein Dank gilt insbesondere folgenden Personen, die zum Teil entscheidenden Einfluss auf das Gelingen dieser Arbeit haben:

Herrn Prof. Dr. Thomas Dobner danke ich herzlich für die Unterstützung und die Mühen der Begutachtung.

Herrn Dr. Dirk Warnecke möchte ich für die schnelle Bereiterklärung zur Begutachtung dieser Arbeit außerhalb seines Fachbereiches danken.

Herrn Dr. Tilman Heise danke ich nicht nur für die Überlassung des Themas, sondern auch für seine stete Aufmerksamkeit, Diskussionsbereitschaft, Betreuung und gemeinsamen Planung.

Frau Dr. Gunhild Sommer danke ich für professionelle und private Unterstützung sowie produktiven und anregenden Diskussionen.

I am thankful for the support given by Dr. Alena Fedarovich, who helped me setting up the FP assays, Dzmitry Fedarovich, who purified most of he recombinant proteins, and of course Dr. Pierre Thibault, who pushed the SUMO project forward by identifying the sumoylation sites in hLa.

Further, I would like to thank past and present members of the Sommer lab, especially Alex and Audra for creating a joyful work environment and for all the refreshing morning coffee breaks.

Mein besonderer Dank gilt meiner Familie, v.a. meinen Eltern, die mich in die große weite Welt ziehen ließen und es geschafft haben mich stets, trotz der Ferne, zu unterstützen. Meinem Bruder danke ich für die technische Unterstützung in der Endphase und den Glauben an meinen Erfolg.

I am truly thankful for having had the chance to meet the following people, who inspired and supported me throughout my life: Martin, Andrea, Dani, Isa, Jens, Chris, Lukasz, Jörg, and Koba.

I am deeply indebted to the loving support of Jason who made me push harder for my goals and made me laugh whenever I needed to.

Julia Kühnert

Lebenslauf und Wissenschaftlicher Werdegang

■ Persönliche Daten

Name: Julia Kühnert
Geburtsdaten: 10.05.1983
Geburtsort: Berlin, Deutschland

■ Schulbildung

1989 – 1995 18. Grundschule Berlin-Hohenschönhausen
1995 – 2002 Pestalozzi-Gymnasium, Berlin
Abschluss: Abitur

■ Hochschulbildung

WS 2002/2003 Beginn des Diplomstudienganges Biologie an der Humboldt-Universität zu Berlin

SS 2007 Beginn der externen Diplomarbeit am Max-Planck-Institut für Infektionsbiologie bei Prof. Dr. Arturo Zychlinsky unter Dr. Hedda Wardemann zum Thema „Screening of IgG+ B cell antibodies for reactivity with HIV-1 associated antigens"

06/2008 Graduierung zur Diplom-Biologin an der Humboldt-Universität zu Berlin

10/2008 Beginn der Promotion an der Universität Hamburg bei Prof. Dr. Thomas Dobner (Heinrich-Pette-Institut), Durchführung an der Medical University of South Carolina, USA

■ Praktika

Sommer 2006 Praktikum am Institut für Biologie/ Genetik bei Prof. Dr. Thomas Börner zum Thema Genexpression in Chloroplasten, Humboldt Universität zu Berlin

02/2008- 10/2008 Praktikum an dem Department of Biochemistry and Molecular Biology bei Dr. Tilman Heise zum Thema der Aufreinigung und biochemischen Charakterisierung des RNA-Bindungsproteins La, Medical University of South Carolina, SC, USA

■ Förderung

12/2007 Reisekostenzuschuss vom DAAD für das o.g. Praktikum