Dissertation zur Erlangung des Doktorgrades

der Fakultät für Chemie und Pharmazie

der Ludwig-Maximilians-Universität München

# The Schlafen Core Domain:

# From Structure to Function

Elisabeth Huber, geb. Krause

aus

Ebersberg, Deutschland

**2018**

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Karl-Peter Hopfner betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den 20.09.2018

.....
Elisabeth Huber

Dissertation eingereicht am: 21.09.2018
1. Gutachter: Prof. Dr. Karl-Peter Hopfner
2. Gutachter: Prof. Dr. Marc Schmidt-Supprian
Mündliche Prüfung am: 06.11.2018

This thesis was prepared from October 2014 to September 2018 in the laboratory of Prof. Dr. Karl-Peter Hopfner at the Gene Center of the Ludwig-Maximilians-Universität (LMU) München.

Parts of this thesis will be submitted for publication:

Huber E., Hopfner K.-P., Lammens K. (2019)
**Structural and Functional Analysis of the Schlafen Core Domain of human Schlafen 5**
Manuscript in preparation.

Parts of this thesis have been presented at international meetings:

Krause E., Hopfner K.-P., Lammens K. (2017)
**Structural and Functional Analysis of the Schlafen Protein Family**
Poster Presentation at the International School of Crystallography: 2.- 11.06.2017 in Erice, Italy.

Table of Contents

# 1. Summary

The immune system is a complex network of processes and structures to protect the organism against infectious diseases. All processes, which include the maturation of individual immune cells, must be strictly regulated to prevent malfunction. The schlafen protein family was recently reported to play key regulatory roles in T cell maturation and different subsets of individual schlafen family members are gradually upregulated during T cell maturation. The schlafen family belongs to the interferon-stimulated genes and various members play important roles in a plethora of cellular processes: regulation of the cell cycle, T cell quiescence, differentiation and proliferation of various cell types, tumorigenesis and inhibition of viral replication.

All schlafen protein family members share the schlafen core domain, which is a highly conserved N- terminal region of approximately 340 amino acids. Based on sequence similarities, this region has been predicted to contain a divergent ATPase domain. Additionally, the longer schlafen protein members harbor a C-terminal domain with sequence homologies to superfamily 1 DNA/RNA helicases.

Although the individual schlafen members play key roles in diverse cellular processes, they are poorly characterized *in vitro* and the underlying molecular mechanisms remain unknown. Since neither biochemical nor structural information were available, the overall goal of this study was to biochemically and structurally describe the schlafen proteins, and specifically their highly conserved schlafen core domain.

Crystallographic studies performed during this thesis, revealed the structures of full length murine Schlafen 2 (Slfn2) and N-terminal human Schlafen 5 ($SLFN5^{1-336}$), which is the first high resolution structure of a human schlafen protein ever reported. Both crystal structures show a unique horseshoe shape with a novel fold harboring a highly conserved zinc finger.

Interestingly, the schlafen core domain resembles no ATPase like fold and indeed, neither ATP-binding nor hydrolysis could be verified experimentally. Thus, structural information combined with experimental data prove that the schlafen core is no ATPase domain.

Biochemical characterization discovered nucleic acid binding properties of the schlafen core domain. Binding affinities of $SLFN5^{1-336}$ increased in a length dependent manner for both, single-stranded DNA and RNA, although no preference for DNA or RNA could be identified. The highest affinity towards RNA was detected for a stem-loop structure and no binding for a RNA duplex, indicating that single-stranded regions are necessary for the interaction. The highest affinity of $SLFN5^{1-336}$ was measured with double-stranded DNA and the residues identified to be responsible for DNA binding were mapped to the zinc finger region by generating DNA-binding deficient mutants. Slfn2 was capable to bind both, DNA and RNA, with a 10 times higher affinity for both substrates than $SLFN5^{1-336}$. The 5S rRNA was identified by co-purification with Slfn2 as a potential nucleic acid substrate.

Further biochemical investigations discovered a new function for $SLFN5^{1-336}$ and Slfn2, which is a 3'- 5' exonuclease activity on single-stranded DNA in a metal ion-dependent, but ATP-independent manner.

The importance of three conserved sites for DNA hydrolysis was confirmed by generating exonuclease-inactive mutants. However, the cleavage mechanism is not entirely solved and therefore a co-crystal structure of a substrate bound schlafen core domain would be of great benefit.

Besides investigations of the schlafen core domain, preliminary biochemical studies of full length SLFN5 revealed increased nucleic acid binding properties towards DNA compared to SLFN5$^{1-336}$, as well as ATP-binding. These results imply that the predicted C-terminal helicase domain significantly contributes to DNA binding and is responsible for ATP turnover. Furthermore, 5.8 S rRNA was determined as potential substrate for full length SLFN5, since it was co-purified with SLFN5 from insect cells.

Collectively, these results not only provide structural insights into the highly conserved slfn core domain, but also propose a new function on molecular basis for the schlafen protein family.

# 2. Introduction

## 2.1. The complex structure of the immune system

From birth to death – during the complete life of individuals, they are continually exposed to all different kinds of microorganisms. Many of those are commensal bacteria, which are harmless or even beneficial, because both, the host and the microorganism, can profit from each other in a symbiosis. However, there are also plenty disease-causing microorganisms or pathogens in the environment. Despite this constant exposure, human individuals become rarely ill. This is due to the immune system, a complex system of processes and structures to protect the human body against infectious diseases. It is classified into different subsystems: the innate immune system, which triggers a non-specific and short-term immunity, and the adaptive immune system, an antigen-specific response system, whose immunity is long lasting. Both systems work in close collaboration to fulfill the four major tasks of the immune system in order to protect the individual successfully. First, the immunological recognition, where white blood cells of the innate system and lymphocytes of the adaptive system detect the presence of the infection. Second, the immune effector function, in which the infection is contained or, if possible, eliminated. Key players of this task are the complement system of blood proteins, the antibodies produced by lymphocytes and the killing properties of lymphocytes. Third, the immune regulation, which is the ability of self-regulation of the immune system. Malfunction of this task leads to allergy and autoimmune disease. Last, the immunological memory, which is an essential ability of the adaptive immune system. The immune system memorizes infections and is able to trigger an immediate and stronger response against any subsequent events[reviewed in 1].

All cells from the immune system derive from hematopoietic stem cells of the bone marrow. Those cells are pluripotent, which means that they can give rise to all the different types of blood cells including erythrocytes, platelets and the two lineages of leukocytes, the lymphoid and myeloid lineages[reviewed in1] (Figure 2.1).

Macrophages, mast cells, dendritic cells, and granulocytes as part of the innate immune system as well as platelets and erythrocytes derive all from one common myeloid progenitor[2]. Macrophages are often the first immune cell type the pathogen gets in touch with[3]. Together with neutrophils and immature dendritic cells, they perform phagocytosis[3,4]. For that, they engulf pathogens and digest them in intracellular vesicles. After maturation, dendritic cells function together with macrophages as antigen presenting cells[5,6]. The antigens are presented to T lymphocytes, which following the recognition are activated and initiate the adaptive immune response[7]. Generally, the other myeloid cells such as the eosinophils, the basophils and mast cells help within the immune response as secretory cells[8]. After activation via antibodies, they release granules containing histamines and additional numerous inflammatory mediators[9]. Furthermore, they are important in allergic responses and kill antibody-coated parasites[10].

The common lymphoid progenitor is the precursor of the antigen-specific lymphocytes of the adaptive immune system and the non-antigen specific natural killer (NK) cells from the innate immune system[11].

The latter are large granular lymphocytes that kill abnormal cells, such as virus-infected cells or tumor cells. Furthermore, they play a role in controlling viral infections until the adaptive immune response is activated[12]. Although both B lymphocytes (B cells) and T lymphocytes (T cells) belong to the antigen-specific lymphocytes, they differ substantially in their individual roles and types of antigen receptors[13].

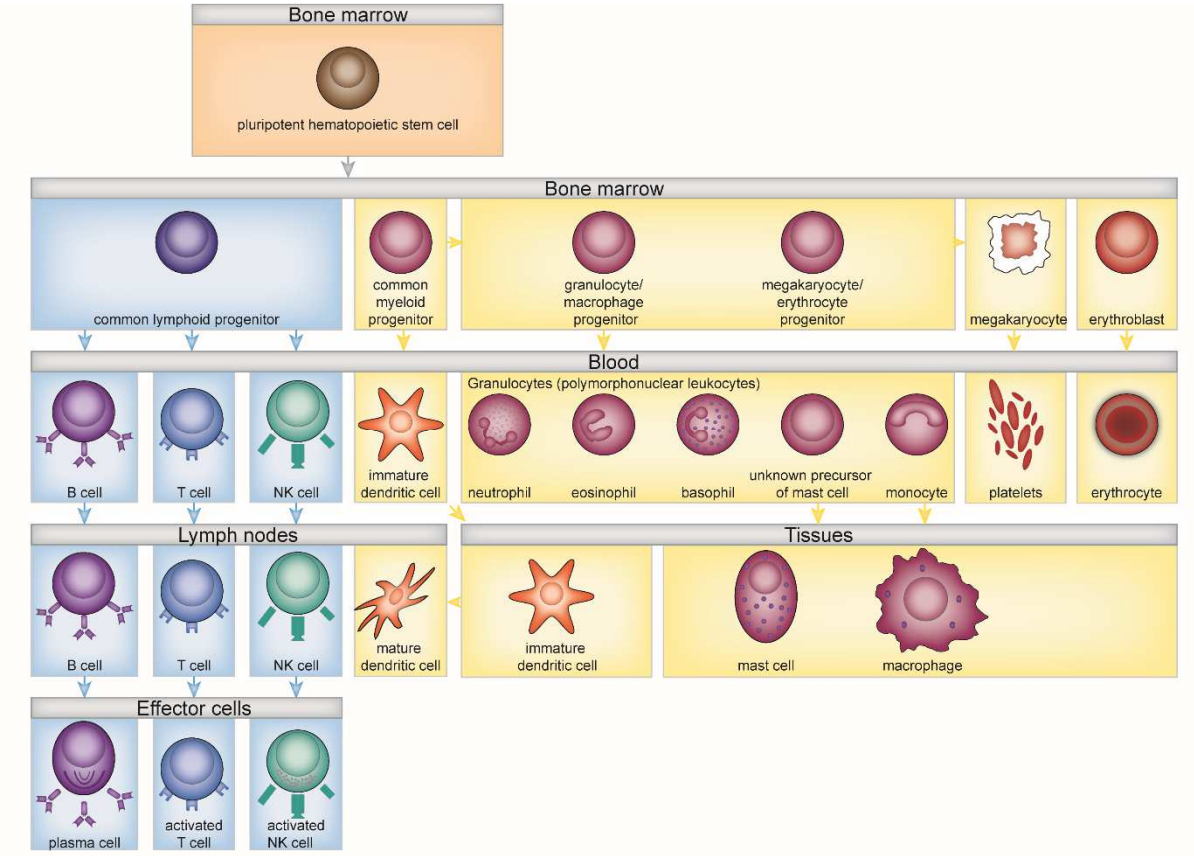



**Figure 2.1:** Schematic overview of the hematopoietic stem cell differentiation. The common lymphoid and myeloid progenitor cells derive from pluripotent hematopoietic stem cells in the bone marrow and give rise to further cell types, which mature at different sites, e.g. the lymph nodes, blood or tissues. B cells, T cells and natural killer cells, which are members of the leukocytes, belong to the lymphoid lineage (blue background) and all other leukocytes as well as erythrocytes and megakaryocytes derive from the myeloid lineage (yellow background). Adopted and modified from[1].

Upon antigen binding to B cell receptors, the B cells differentiate into plasma cells and produce antibodies with an identical antigen specificity[14]. Upon antigen binding to the T cell receptor, the T cell becomes activated and can proliferate into several effector T lymphocytes, which can be classified by their task[15]. First, cytotoxic T cells kill infected cells. Second, helper T cells activate other immune cells, such as B cells or macrophages. Third, regulatory T cells help to regulate the immune response by repressing the activity of other lymphocytes[15]. One additional feature of the immune system is the immunological memory. Some of the B cells and T cells differentiate into memory cells upon antigen activation. Those cells guarantee a livelong immunity. Upon a second exposure to their specific antigen, they directly differentiate into effector cells and trigger an immediate and more vigorous immune

response. However, in the absence of an infection, B and T cells are in an inactive state and display a small and featureless phenotype with few cytoplasmic organelles and an inactive nuclear chromatin, characterized by a condensed nucleus[16].

### 2.2. The newly identified schlafen proteins regulate thymocyte maturation and differentiation

T cells derive from a common lymphoid progenitor cell in the bone marrow, which belongs together with the thymus to the central lymphoid organs. For differentiation they migrate into the thymus[11]. During the maturation of T cells, differential expression or different combinations of cell-surface markers, such as the CD3-T cell receptor complex (TCR) or the co-receptor proteins cluster of differentiation (CD) 4 and CD8, characterize the various stages of thymocyte maturation (Figure 2.2). When the progenitor cells enter the thymus from the bone marrow, they still lack most of the surface molecules, which are specific for mature T cells, and the T cell receptor genes are not rearranged yet. During T cell maturation the different gene segments V (variable), D (diversity), J (joining) and C (constant) undergo rearrangements to guarantee variability of the various TCRs[15]. The immature progenitor cells can either develop to the major T cell population with α:β chains of the TCR, or to the minor population with a γ:δ TCR[17]. The latter population however misses the CD4 and CD8 markers, even when they are mature. In the first stage of α:β T cell maturation the cells still lack the CD3:TCR complex and the co-receptors CD4 and CD8 ($CD4^-$, $CD8^-$), thus they are called double-negative (DN) thymocytes. The double-negative stage is further separated into four phases: During the DN1 stage CD44, an adhesion molecule, but no CD25, the α-chain of the IL-2 receptor, are expressed. Furthermore, the genes encoding both chains of the T cell receptor are still in its germline configuration. The DN2 stage is characterized by additional expression of CD25 and the rearrangement of the T cell receptor β-chain locus begins with $D_\beta$ to $J_\beta$ rearrangements. During DN3, the expression of CD44 reduces, whereas the $V_\beta$ to $DJ_\beta$ rearrangement continues. Those DN4 cells with successful β-chain rearrangement lose CD25 expression again and can further develop. In both stages, DN3 and DN4, the expressed β-chains pair with a surrogate pre- T cell receptor α-chain (pTα), which allows assembly of a complete pre-T cell receptor[18]. Assembly of the pre-TCR together with the CD3 molecule induces ligand-independent receptor dimerization. This step is crucial for further cell proliferation, the arrest of the β-chain gene rearrangement and transition to double-positive (DP) thymocytes proliferation expressing both, CD4 and CD8 cell-surface markers. Throughout the T cell expression, but especially during the early stages, the Notch signaling pathway is important[19].

Those cells, expressing CD4 and CD8 on their surface ($CD4^+$ and $CD8^+$), are called double-positive thymocytes[17]. They further develop from large and active cells to small and resting cells, which start to rearrange the α-chain locus. In those cells, initially, only low levels of the α:β TCR are expressed and most of them fail positive selection, meaning that they cannot recognize self-peptides: self-major histocompatibility complex (MHC) complexes[20]. Cells passing the positive selection process continue with maturation and start expressing high TCR levels. They progress either to $CD4^+$ or to $CD8^+$ single positive (SP) T cells. However, before they can transfer from the thymus to the periphery they have to pass negative selection, as well[20]. Here, cells undergo self-programmed apoptosis, when they recognize

self-peptide: self-MHC too strong. The whole proliferation process takes approximately three weeks time to complete and happens mainly in the cortex region of the thymus. The last step of negative selection of single-positive T cells, however, takes place in the medulla of the thymus[17].

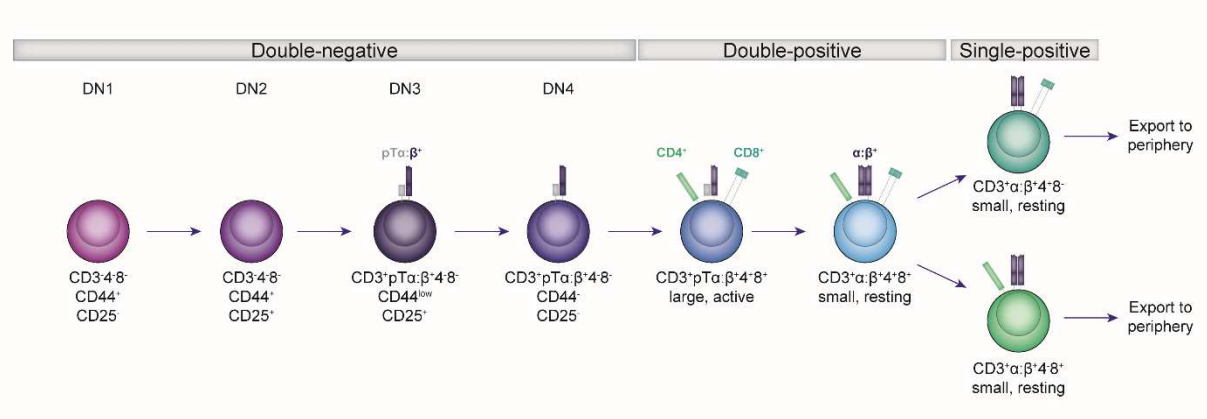



**Figure 2.2:** Schematic overview of the T cell development in the thymus. The double negative (DN) stage, when CD4 and CD8 are not expressed, are divided into four phases. In DN1 CD44, but not CD25 is expressed. DN2 is characterized by additional expression of CD25. During DN3 and DN4, the pre-T cell receptor is formed and CD3 is expressed. The expression levels of CD44 and CD25 decrease. Afterwards, the cells proliferate into double-positive cells, expressing both, CD4 and CD8. During this stage, the mature T cell receptor is formed. Before the cells are exported to the periphery, they lose either the CD4 or CD8 surface marker, resulting in single-positive cells. Adopted and modified from[1].

The surface markers CD4 and CD8 are important for the individual tasks of the T cells, because they help mediating the interaction between T cells and other cells. Up to this point, the whole development process happened independently of antigens. When naïve $CD8^+$ T cells leave the thymus and they find pathogen peptides presented by MHC class I molecules they differentiate into cytotoxic T cells, whose main task is to recognize and kill infected cells[21]. Naïve $CD4^+$ T cells, however, develop into different types of effector T cells, if they encounter pathogen peptides presented by MHC class II molecules: $T_H1$, $T_H2$, $T_H17$ and follicular helper T cells ($T_{FH}$)[22], which activate their target cells, or regulatory T cells with inhibitory activity[23].

A newly identified and largely uncharacterized protein family playing essential roles in the regulation of thymocyte differentiation is the schlafen protein family[24]. They were discovered in the late 1990s during studies investigating thymocyte development. The goal there was to identify genes that are upregulated during positive selection of T cells. For that purpose, thymic cDNA libraries from two different mice strains were compared with each other: On the one hand, AND.B6 mice, which mainly constitute $CD4^+$ and $CD8^+$ SP cells, and on the other hand, AND.4R mice, where the thymocyte maturation is halted at the $CD4^+CD8^+$ DP stage and almost no SP cells are maturing. Schlafen family member 1 (Slfn1) that was discovered first is preferentially expressed in AND.B6 mice and was highly upregulated in positively

selected, mature $CD4^+$ and $CD8^+$ SP thymocytes as compared to the immature progenitors cells. Ectopic expression of Slfn1 early in the T lineage in mice resulted in failed complete maturation of DP thymocytes by causing a cell cycle arrest at the G0/G1 stage[24]. Not only the number of thymocytes were reduced to 1 %- 30 % of normal, but also the thymus of heterozygous mice reduced in size to about 10 %- 30 %, and was in homozygous mice even only 1 %- 3 % of the usual size[24]. Since differentiation of mature naive lymphocytes is tightly interconnected with the cell cycle control machinery[25,26], it seems that Slfn1 inhibits the proliferation by acting on the cell cycle. Later these findings could be confirmed by demonstrating that Slfn1 blocks the cell cycle during the G1 phase by inhibiting the induction of the cell cycle regulator Cyclin D1[27-29]. Since this growth inhibition could be reversed by terminating the Slfn1 gene expression, it was named 'schlafen' (German for 'to sleep'). Besides Slfn1, also Slfn2, Slfn3 and Slfn4 proteins were discovered based on their sequence homology. All of them are mainly expressed in tissues of the immune system, especially in the thymus, except for Slfn3 that is mainly expressed in the testis. At this time point, no other proteins with similar sequences have been known, thus a new protein family participating in unknown regulatory mechanisms guiding both cell growth and T cell development was identified[24]. Later more schlafen protein family members have been identified, namely Slfn5, 8, 9 and 10. They seem to play significant roles in hematopoietic stem cell differentiation, based on their different regulation of various cells: in developing and activated T cells, in activated macrophages, in differentiated myeloblastic cells into macrophage-like cells and in splenocytes of mice infected with *Listeria*[30]. Generally, the distinct schlafen members seem to differentially express during the various stages of thymocyte development (Figure 2.3)[24,30]. After transition from CD4/ CD8 DP to SP stages Slfn1 is 100 times[24,30,31] and Slfn2 10 times more upregulated[24,30]. However, both expression levels decrease again upon T cell activation[30]. Slfn3 is well expressed throughout thymocyte development with the highest levels in SP T cells. Slfn4 is preferentially found in the early stages and decreases during development. However, both, Slfn3 and Slfn4, are upregulated in peripheral T cells after activation[24,30]. Interestingly, in another study Slfn3 was highly upregulated in $CD4^+CD25^+$ cells in the periphery and upon activation and proliferation Slfn3 is down-regulated in $CD4^+CD25^+$ $T_{regs}$ and up-regulated in $CD4^+CD25^-$ $T_{effs}$[32]. Slfn8 is less expressed during thymocyte development, but highly upregulated in mature, resting lymphocytes. Both Slfn9 and Slfn10 are upregulated during the early stages of thymocyte development, and again in peripheral T cells after stimulation[30] (Figure 2.3).

Not only murine, but also human schlafen family members could be linked to immune cell proliferation and especially to T cell maturation. All human schlafen proteins (SLFN5, SLFN11, SLFN12, SLFN12-L, SLFN13) except SLFN14 showed basal expression in monocytes, monocytes-derived dendritic cells (moDCs) and T cells[33]. SLFN11 expression is especially high in monocytes and moDCs, whereas SLFN5 is upregulated in T cells, suggesting a functional importance within these cell types. Like in the murine system, a combination of a subset of schlafen members seem to express simultaneously. For example, SLFN12-L and SLFN13 are gradually upregulated during the differentiation of monocytes to moDC, but SLFN12 levels decrease by differentiation stimuli. Moreover, during T cell activation SLFN5, SLFN12 and SLFN12-L are downregulated, indicating a role in the regulation of differentiation and development of T cells.

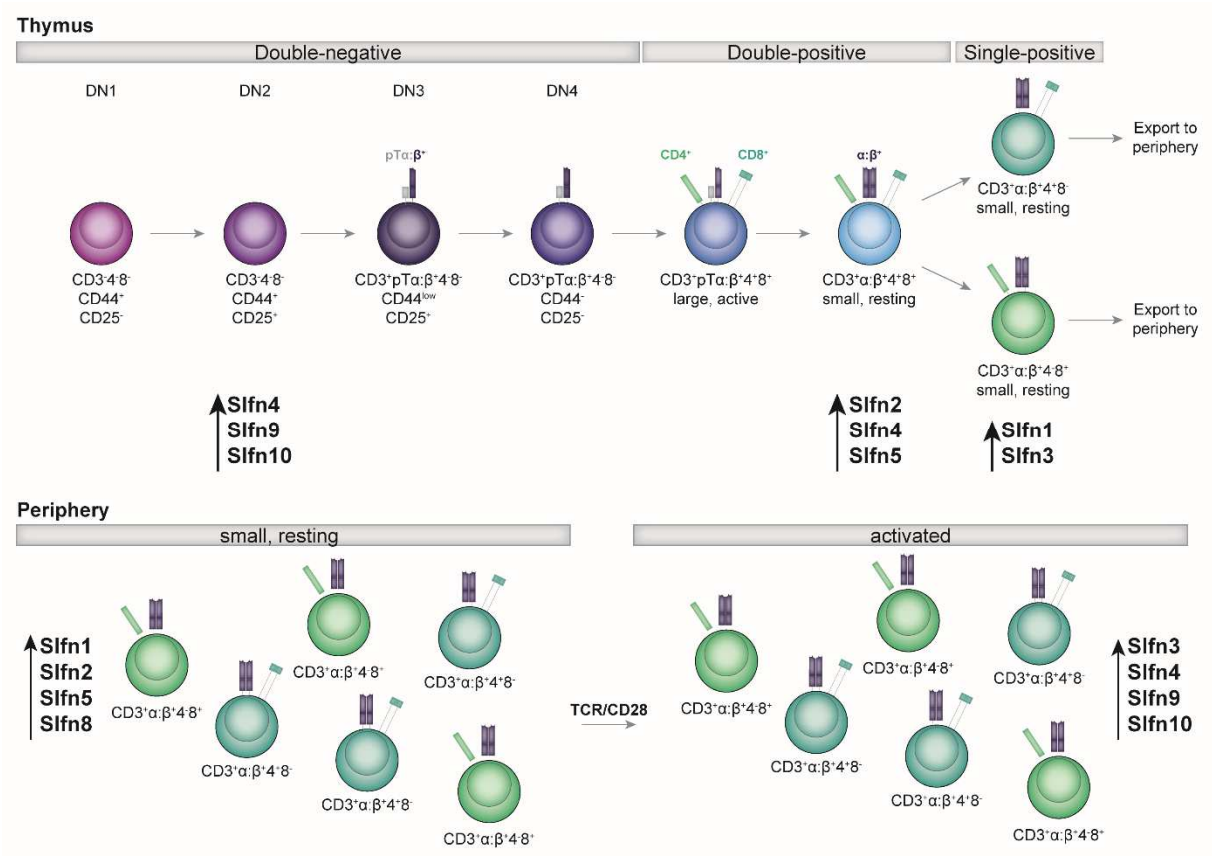



**Figure 2.3:** Schematic overview of the differential upregulation of murine slfn members during T cell development in the thymus and in resting and activated T cells in the periphery. During the DN phase Slfn4, 9 and 10 are upregulated, whereas in the DP stage Slfn2, 4, and 5 are mainly expressed. In both $CD4^+$ and $CD8^+$ SP thymocytes, Slfn1 and Slfn3 expression is increased. In resting T cells Slfn1, 2, 5 and 8 are upregulated. The expression pattern shifts upon TCR/CD28 stimulation to Slfn3, 4, 9 and 10[24,30].

### 2.3. The schlafen family belongs to the interferon-stimulated genes

Important signaling molecules in the immune system are cytokines. Generally, cytokines are produced by one cell and influence the behavior of another cell[1]. One group of important cytokines constitute interferons (IFN). The original definition of interferons, which were discovered 60 years ago, is the ability to 'interfere' with viral infections[34]. Interferons are divided into several types. Type I, which includes IFN- α, IFN-β, IFN-ε, IFN-κ and IFN-ω in humans, all bind a common cell-surface receptor, namely the type I IFN receptor[35-38]. Type II interferons exclusively consist of IFN-γ, which bind to type II IFN receptor[39,40]. A more recently identified group, are type III IFNs, consisting of IFNL1, IFNL2, IFNL3[41,42] and IFNL4 (also known as IFNλ1-4) [43,44].

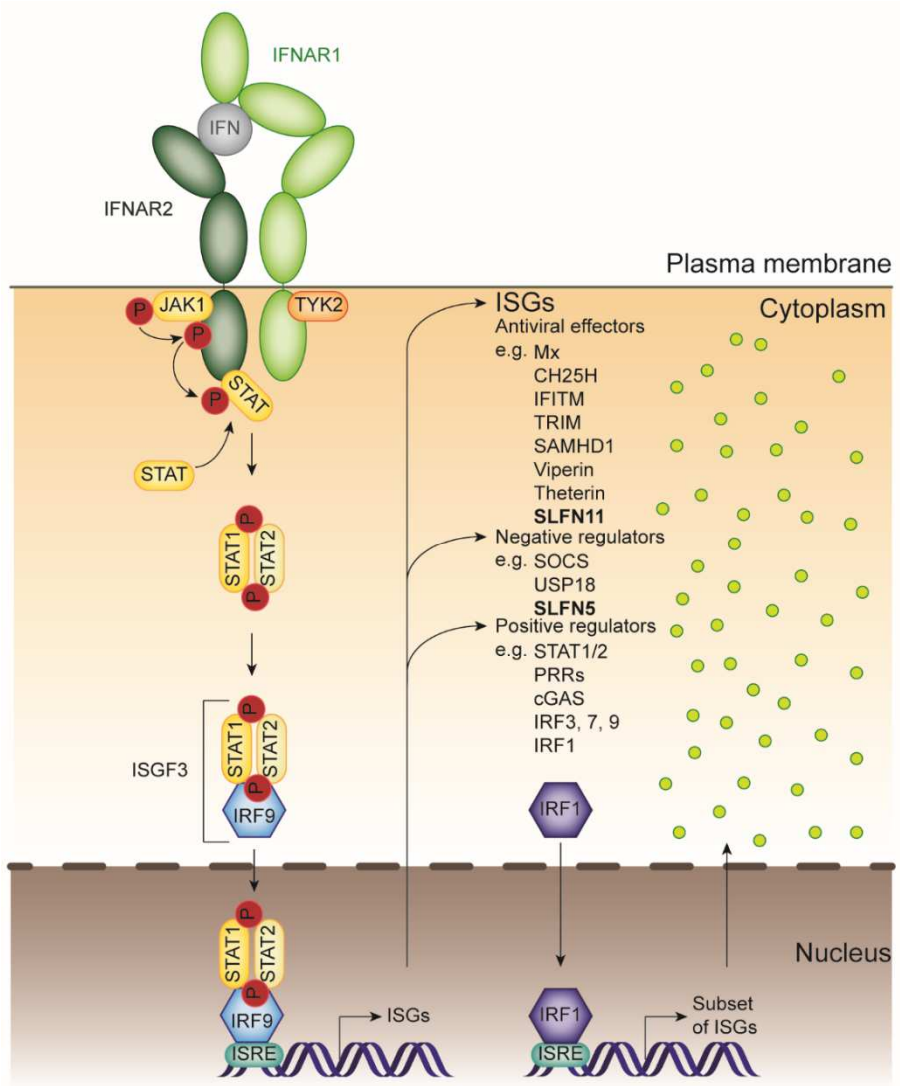



**Figure 2.4:** Schematic overview of type I IFN signaling. Upon IFN binding the two interferon receptor chains undergo conformational changes, thereby activating JAK1 and TYK2. Afterwards, the JAKs phosphorylate the two receptor chains, thus activating the STATs. The phosphorylation of the STATs triggers dimerization and together with IRF9, they translocate to the nucleus, where the expression of various ISGs is triggered. ISGs exhibit antiviral properties, or are positive or negative regulators of the JAK-STAT pathway. Adapted and modified from [45].

IFN signaling occurs via the Janus-activated kinase- Signal transducer and activator of transcription (JAK- STAT) pathway[45,46]. For type I IFN signaling the IFNAR1 and IFNAR2 receptor subunits are associated with the tyrosine kinase 2 (TKY2) and JAK1, respectively. Both kinases belong to the JAK-family. Upon IFN binding, the receptor subunits undergo large conformational changes and come in close proximity to each other. Those ligand-induced structural changes are required for activation since both JAK kinase domains juxtapose and undergo transphosphorylation[45,47]. After activation, the JAKs phosphorylate the two IFNAR1/2 receptor chains on highly conserved tyrosine residues, and STAT1 and 2 signaling molecules are repositioned or bound via their Src homology 2 (SH2) domain[48]. Next, the STATs are phosphorylated at conserved tyrosine residues and are then released from the receptor. Consequently, the STATs undergo heterodimerization and interact with the interferon regulatory factor 9 (IRF9), forming the interferon-stimulated-gene factor 3 (ISGF3) complex[49]. This complex translocates

to the nucleus and binds to IFN-stimulated response elements (ISREs) on the DNA upstream of IFN-stimulated genes (ISGs), thus initiating the transcription of a plethora of ISGs[45,50,51] (Figure 2.4).

The schlafen family belongs to the classical ISGs and is regulated by the STAT complexes[52]. All studied schlafen family members are highly upregulated in response to IFN-α treatment not only in lymphocytes, but also in other cell types. For example, murine Slfn1, 2, 3, 4, 5 and 8 are upregulated following IFN-α treatment in mouse melanoma cells and renal cell carcinomas[53] and SLFN5 in renal cell carcinomas[54] and malignant melanoma cells[55].

Some ISGs can regulate their own JAK-STAT signaling pathway in a positive or negative manner. Positive regulators, that are present at baseline levels under normal conditions, can reinforce the IFN response, e.g. JAK2, STAT1/2 and IRF9. Other positive regulators are cytosolic pattern-recognition receptors (PRRs), such as protein kinase R (PKR), melanoma differentiation-associated protein 5 (MDA5), retinoic acid-inducible gene I (RIG-I), AIM-2 or cyclic GMP-AMP synthase (cGAS). IRF1 represents a special case for positive regulators, because upon expression it can directly translocate into the nucleus to boost the expression of several ISGs reviewed in [45]. Negative regulators are important for cells to recover from IFN signaling, since dysregulation of IFN production leads to autoimmune disorders[56]. For example, suppressor of cytokine signaling (SOCS) proteins block phosphorylated tyrosines either on the IFN receptor or on JAKs, thereby preventing their activity and binding of STATs[57]. Another negative regulator is ubiquitin-specific peptidase 18 (USP18), which main task is the de-ISGylation, meaning the removal of ISG15, an ubiquitin-like protein, from proteins[58]. USP18 interacts in an ISG15-dependent manner with the IFNAR2 receptor, thereby inhibiting signaling. However, this inhibition seems to be restricted to IFN-α signaling[59,60]. Another positive regulator seem to be human SLFN5. By using mass spectrometry analysis, SLFN5 was found to specifically interact with STAT1. Further chromatin immunoprecipitation assays revealed that SLFN5 is enriched in the promoter region of STAT1 target genes, e.g. ISG15, suggesting that a potential SLFN5-STAT1 complex might bind to the ISRE in the promoters of ISGs. By using microarray experiments, it was observed that the expression of ISGs was increased in SLFN5 knockout cells compared with wild type cells, suggesting an ISRE repressor function for SLFN5. Taken together, SLFN5 is a potential transcriptional co-repressor of STAT1[61]. Since SLFN5 is not only an ISG, but also a repressor of ISG transcription, the existence of a negative-feedback regulatory loop is speculated[61].

Another essential function of ISGs are the antiviral properties, such as Mx, CH25H, IFITM, Trim, OAS/RNAse L, Viperin, Tetherin, SamHD or human SLFN11, each of those acting at a specific stage of the viral life cycle reviewed in [45,62].

For example human immunodeficiency virus type 1 (HIV-1) is an RNA virus that infects human T cells, macrophages and dendritic cells and causes the acquired immunodeficiency syndrome (AIDS)[63]. Upon infection, HIV binds to the host cell and infiltrates its RNA genome. By reverse transcription, a DNA copy of the viral genome is generated and incorporated into the host genome. The host's transcriptional and translational machinery is manipulated to generate new viral proteins and together with the viral RNA genome new viral particles are generated and released[63] (Figure 2.5). So-called host restriction factors can specifically block the replication at certain stages of the viral life cycle.

Mx1 (Myxovirus resistance) blocks the virus in an early post entry stage of its life cycle by trapping incoming viral components[64], whereas Mx2 hinders the reverse-transcribed genome in reaching its nuclear destination[65-67]. Both, cholesterol-25-hydroxylase (CH25H)[68], an enzyme that converts cholesterol into 25-hydroxycholesterol, and members of IFN-inducible transmembrane (IFITM) family, which are enriched in endosomes and lysosomes[69], are shown to inhibit viral entry by altering the membrane properties reviewed in [45]. TRIM5α and TRIM22, both belonging to the tripartite motif (TRIM) family of proteins, are two of those host restriction factors and play essential roles in inhibiting the replication of HIV-1reviewed in [70]. Trim5α acts on the early stage of the HIV life cycle by accelerating the dismantling of the capsid shell, leading to premature exposure of the viral nucleoproteins[71], whereas TRIM22 constrains trafficking of the viral Gag proteins to the host plasma membrane, resulting in a decline of viral particle release[72]. Another important HIV restriction factor is SAM and HD domain-containing protein 1 (SAMHD1), which decreases the available pool of dNTPS needed by the virus for the reverse transcriptase to generate the cDNA[73]. During the end of the viral life cycle, virus inhibitory protein, endoplasmic reticulum-associated, IFN-inducible (viperin) protein and tetherin, inhibit budding of viral particles, by blocking the activity of required enzymes[74,75] and by using membrane anchors to trap the virus on the host's plasma membrane[76], respectively.

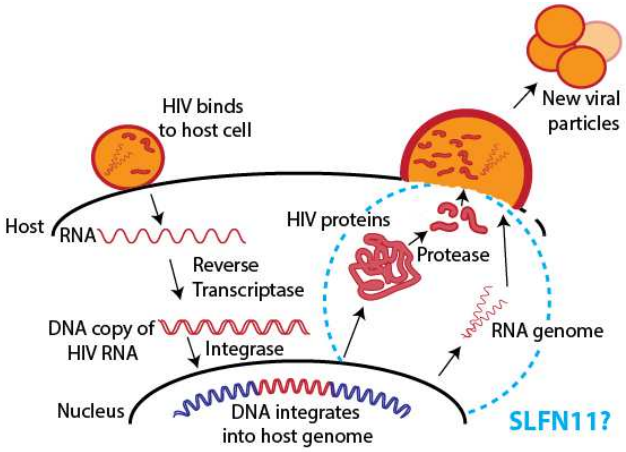



**Figure 2.5:** Schematic overview of the HIV life cycle. HIV binds to the host cell, injects the viral genome and generates a cDNA copy via reverse transcription. After integration of the cDNA into the host genome, new viral proteins are formed and together with the RNA genome, they are packed. Newly generated viral particles are then released. SLFN11 specifically abrogates the viral protein synthesis in a codon-usage dependent manner and acts in a late stage during the cycle. However, the exact working mechanism remains unknown. Adapted and modified from[77].

Another HIV-restriction factor delineates human SLFN11[78]. The protein specifically abrogates the replication of HIV by selectively inhibiting the expression of viral proteins in a codon-usage dependent manner. SLFN11 did not affect the viral life cycle at the early stages of reverse transcription, integration and transcription, but appears to bind tRNAs and thus inhibits the translation of viral proteins[78]. Usually, during HIV-infection, the host tRNA pool is shifted favoring the viral codons, which use predominantly A/U rich codons[79]. Little or no changes in the tRNA pool were observed upon overexpression of SLFN11, suggesting that SLFN11 in HIV-1-infected cells potentially counteracts the virus-induced changes in the

tRNA pool (Figure 2.5). Since SLFN5 fails to inhibit retrovirus production, this activity appears to be SLFN11-specific[78]. Indeed, SLFN11 mRNA expression levels were elevated in HIV-1 infected patients[80]. Furthermore, high SLFN11 expression levels were observed in $CD4^+$ T cells upon treatment with HIV elite controller[80] and elevated SLFN11 expression is strongly associated with reduced HIV RNA levels[81]. Polymorphisms in SLFN11 expression might partly explain why some HIV-positive patients never experience high viral titers or symptoms of AIDS[82]. The development of SLFN11 inducing drugs might thus help to fight the virus or SLFN11 could potentially serve as a biomarker to determine the likelihood of progression to AIDS[82]. Accordingly, equine Slfn11 restricts the production of equine infectious anemia virus via a similar codon-usage dependent mechanism[83]. Interestingly, nonhuman primate Slfn11 versions and human SLFN11 block the accumulation of viral proteins and non-viral as well as host transcripts. Thus, Schlafen 11 may rather be an interferon-stimulated gene with broad ability to inhibit viral and host protein production by creating an antiviral state in the cell, than being a classical host restriction factor[84].

### 2.4. The schlafen protein family is a versatile family with various cellular functions

The different schlafen members are not only upregulated by IFN-α, but also upon treatment with lipopolysaccharides (LPS), rhinovirus[33] and in mice after infection with *Listeria* or *Brucella*[30,85]. Generally, the various schlafen protein members play important roles in a plethora of proliferative processes: regulation of the cell cycle[27,28], T cell quiescence[33,86-89], hematopoiesis[52], osteoclastogenesis[90], arthritis-related genes[91], epithelial cell differentiation[92-97], myelopoiesis[98-100] and proliferation of endothelial[28] and epithelial cells[101,102]. Furthermore, both human and mouse schlafen members have been shown to play key roles in tumorigenesis and enhance the antineoplastic effects of IFN-α[52,53,55,93,101,103,104]. Some schlafen members sensitizes cancer cells to cytotoxic compounds[103-110] and some are shown to inhibit viral replication[78,111]. Further schlafen members are shown to be epigenetically suppressed by the methylation of their genes[104,110,112,113]. All studied functions of all murine, human and selected rat schlafen family members are summarized in Table 2.1.

**Table 2.1:** Summary of all murine, human and selected rat schlafen family members.

| | |
|---|---|
| **mSlfn1** | • Expression in hematopoietic stem cells, thymus, spleen, T cells, bone marrow, lymph nodes, macrophages, adipose tissue, aorta and veins, spinal cord[114]<br>• Expression in immature T cells in DP to SP transition and untreated T cells; decreased expression in activated T cells[24]<br>• Alters T cell growth and development[24]<br>• Upregulation upon treatment with IFN-α[53], LPS[115,116] and bleomycin[117], infection with *Listeria*[30]<br>• Epigenetic downregulation by gene methylation by Uhrf1[112]<br>• Arrests cell-cycle by inhibiting induction of cyclin D1[27]<br>• Inhibits proliferation and tube formation of endothelial progenitor cells[28] |
| **mSlfn2** | • Expression in thymus, spleen, dendritic cells, hematopoietic stem cells, macrophages, tumor cells, mammary gland, lung, fetus, joints, synovial fibroblasts, inner ear, taste buds[114] |

| | |
|---|---|
| | • Expression in immature T cells in DP to SP transition and untreated T cells; decreased expression in activated T cells[24]<br>• Alters cell growth and development[24] including mouse hematopoiesis[52]<br>• Upregulation upon treatment with IFN-α[53], LPS[116,118] and bleomycin[117], infection with *Listeria*[30] and *Brucella*[85]<br>• Upregulation upon Toll-like receptor triggered stimulation of AP-1 and NFκB[118]<br>• Negative regulator of ISGs by influencing NFκB signaling activation[119]<br>• Epigenetic downregulation by gene methylation by Uhrf1[112]<br>• Antineoplastic effects in melanoma and renal cell carcinoma cells by control of tumor cell proliferation and/or anchorage-independent growth[52,53]<br>• Regulator of osteoclastogenesis[90,120]<br>• Key role in T cell quiescence; loss-of-function Slfn2 point mutation in *elektra*** mice leads to immunosuppressed phenotype, decreased levels of $CD4^+$ and $CD8^+$ T cells and increased levels of apoptotic cells[86]<br>• Elevated sterol and lipid levels[89] and chronic ER stress[121] in *elektra* mice |
| **mSlfn3** | • Expression in thymus, spleen, dendritic cells, macrophages, testis, melanoma/melanocyte cells[114]<br>• Expression in mature SP T cells after antigen stimulation[24]<br>• Upregulation upon treatment with LPS[116]<br>• Downregulation by TGF-β[32]<br>• Antineoplastic effects in renal cell carcinoma and malignant melanoma cells[53]<br>• Inhibitory effects on differentiation and proliferation of FOLFOX-resistant cancer cells[93]<br>• Regulator of enterocytic differentiation[96,122], colonic mucosal growth during aging[94] and epithelial cell differentiation[97,123] |
| **mSlfn4** | • Expression in spleen, bone, lymph nodes, hematopoietic stem cells, vagina, embryo, mammary and lung tumors, taste buds, in unfertilized eggs and in 8-cell embryos[114]<br>• Expression in T cells decreases during development and increases in mature T cells upon stimulation[24]<br>• Expression during macrophage activation, but down-regulated during differentiation[98]<br>• Upregulation upon treatment with IFN-α[53], LPS[116] and bleomycin[117], infection with *Listeria*[30]<br>• Gli1-dependent inhibitory effects in myelopoiesis upon *Helicobacter* infection[99,100] |
| **mSlfn5** | • Expression in thymus, spleen, activated macrophages, bone, lung, kidney, embryonic Rathke's pouches, testis[114] and podocytes[124]<br>• Upregulation upon treatment with IFN-α[53] and infection with *Listeria*[30]<br>• Antineoplastic effects in renal cell carcinoma[53] |
| **mSlfn8** | • Expression in macrophages, dendritic cells, spleen and kidney[114]<br>• Upregulation upon treatment with IFN-α[53] and infection with *Listeria*[30]<br>• Regulation of T cell development and proliferation of peripheral T cells[30]<br>• Transgenic Slfn8 mice show reduced thymus size[30]<br>• $Slfn8^{-/-}$ mice are resistant to autoimmune encephalomyelitis (EAE)[125] |
| **mSlfn9** | • Expression in hematopoietic stem cells, activated macrophages, bone, embryonic nasal region, tumors, prostate, eye, embryos, mammary gland and embryonic stem cells[114]<br>• Upregulation upon treatment with IFN-α[53] and infection with *Listeria*[30] |
| **mSlfn10** | • Expression in activated macrophages, thymus, hematopoietic stem cells, dendritic cells, pancreatic islet, lung and mammary tumor[114]<br>• Upregulation upon treatment with IFN-α[53] and infection with *Listeria*[30] |
| **mSlfn14** | • Expression in spleen[114] |

** More detailed explanation about mSlfn2 and *elektra* mice in section 2.7

| | |
|---|---|
| **rSlfn13** | • Endoribonuclease cleaving tRNA/rRNA engaged in translational control[126] |
| **hSLFN5** | • High expression in T cells; downregulation upon T cell activation[33]<br>• Upregulation upon IFN-α treatment[55] and degradation of PDCD4 (programmed cell death protein)[127]<br>• Antineoplastic effects in malignant melanoma cells[55] and renal cell carcinoma by downregulation of metalloproteinases[54]<br>• Enhances tumor growth in glioblastoma cells[61] and gastric cancer cells[128]<br>• Transcriptional co-repressor of STAT1 through direct interaction with STAT1[61]<br>• Part of the NOTCH interactome[129] |
| **hSLFN11** | • Expression in monocytes and monocytes-derived dendritic cells[33]<br>• Upregulation upon treatment with IFN-α[33,55], LPS and rhinovirus[33]<br>• Part of the NOTCH interactome[129]<br>• Expression is influenced by treatment with HIV elite controllers[80] and levels of HIV RNA by antiviral therapy[81] in $CD4^+$ T cells[80,81]<br>• Inhibition of HIV-replication in a codon-usage dependent manner[78] (also equine Slfn11 restricts equine infectious anemia virus via a similar mechanism[83] and versions of nonhuman primate slfn11 have enhanced antiviral effect on HIV[84])<br>• Sensitizes cancer cells to DNA-damaging agents[103,104,106,107,113,130] and PARP-inhibitors[108,109,130,131]<br>• Downregulation by promoter methylation[107,110,113] by EZH2 in small-cell lung cancer[110]<br>• Direct interaction with DHX9[107,132] and RPA[132,133]<br>• Found at stalled replication forks[132,133] and involved in DNA-repair independently of ATR[132] |
| **hSLFN12** | • Basal expression in monocytes, monocytes-derived dendritic cells and T cells; expression decreases by differentiation stimuli[33]<br>• Upregulation upon treatment with IFN-α[33,55], LPS and rhinovirus[33]<br>• Upregulation together with KLF2 (a quiescence factor) in presence of ct-CD45 during T cell activation[87]<br>• Interaction with phosphodiesterase 3A (PDE3A) correlates with DNMDP (a cancer cytotoxic compound) sensitivity to cancer cells[105,134]<br>• Methylation of gene locus in PBMCs of allergic rhinitis patients[135]<br>• Regulation of differentiation of prostate epithelial cells[101] and enterocytic cells by Serpin B12 (deubiquitylase) interaction[102] |
| **hSLFN12L** | • Basal expression in monocytes, monocytes-derived dendritic cells (moDCs) and T cells and upregulation upon differentiation of monocytes into moDCs[33]<br>• Upregulation upon treatment with IFN-α[33,55], LPS and rhinovirus[33]<br>• Co-localization with cells expressing myeloid-derived suppressor cell (MDSC) markers[100] |
| **hSLFN13** | • Basal expression in monocytes, monocytes-derived dendritic cells (moDCs) and T cells and upregulation upon differentiation of monocytes into moDCs[33]<br>• Upregulation upon treatment with IFN-α, LPS and rhinovirus[33] |
| **hSLFN14** | • Upregulation upon influenza virus infection and inhibition of viral replication[111]<br>• Cleavage of rRNA and ribosome-associated mRNA in an $Mn^{2+}/Mg^{2+}$-dependent, NTP-independent and sequence non-specific manner[136]<br>• Mutations (K218E, K219N, V220D, R223W) cause thrombocytopenia in patients[137,138]<br>• Co-localization with ribosomes and mediation of rRNA endonucleolytic degradation[139] |
| **vSlfn** | • Orthopox virulence factor that affects host immune response to infection; infectious domain is the N-terminal virus-specific domain[140] |

### 2.5. Structure and domain architecture of the schlafen protein family

The schlafen protein family is mainly present in mammals, most abundantly in primates and rodents. For mammals, at least two schlafen members exist in every genome, except in platypus, where only one gene was discovered. Besides in mammals, schlafen members are also found in one fish (elephant fish; *Callorhinchus milii*) and one amphibian species (African clawed frog; *Xenopus laevis*), indicating an ancient origin of these genes. Within each genome all schlafen paralogues are clustered in tandem within orthologous regions, indicating that the schlafen clusters have been shaped by multiple tandem duplication events and also lineage-specific inversions[114]. Besides the presence in vertebrates, one copy of a schlafen family gene has been found in all genomes of orthopoxviruses, which was probably horizontally transferred from murine rodents. A suggested function of the viral schlafen protein is a viral defense mechanism to escape the host immune defense[114,140]. The cluster of schlafen genes is located on important sites in the genome. For example in mice, they are mapped in a 350 kb interval of chromosome 11 within a region that is connected with embryonic lethality and meiotic drive[114]. So far, eight schlafen members (*slfn1, 2, 3, 4, 5, 8, 9, 10*) in the murine genome have been discovered plus two additional schlafen copies that include some sequence similarities (*Slfn1-like, Slfn14*)[114]. In humans six schlafen members (*SLFN5, 11, 12, 12-like, 13, 14*) are described[30]. The schlafen family can either be divided into four phylogenetic clades[114], or into three subgroups according to the size of their encoded proteins[30]. Subgroup I harbors the shortest members (37 to 42 kDa), namely mSlfn1 and mSlfn2, subgroup II the intermediate sized proteins (59 to 68 kDa), including mSlfn3, mSlfn4, hSLFN12 and hSLFN12-L, and subgroup III includes members with molecular masses ranging from 100 kDa to 104 kDa, such as mSlfn5, mSlfn8, mSlfn9, mSlfn10, mSlfn14, hSLFN5, hSLFN11, hSLFN13 and hSLFN14[30] (Figure 2.6). According to the phylogenetic clades, group 1 harbors mSlfn1, mSlfn2, mSlfn3, mSlfn4, hSLFN12 and hSLFN12-L, group 2 includes mSlfn5 and hSLFN5, group 3 consists of mSlfn14 and hSLFN14, and group 4 harbors mSlfn8, mSlfn9, mSlfn10, hSLFN11 and hSLFN13[114]. The viral schlafen protein, which shares the highest homology to mSlfn1, is classified into subgroup I and phylogenetic group 1[114,140] (Figure 2.6).

All schlafen family members share a highly conserved N-terminal common core region of approximately 340 amino acids. Every schlafen family member contains at least this domain and subgroup I members even consist exclusively of this domain[30]. Based on sequence similarity, this region has been predicted to contain a divergent ATPase associated with various cellular activities (AAA 4 ATPase) domain[141]. Here, to this region will be referred as schlafen core region. Within the schlafen core one highly conserved stretch of 80-100 amino harbors a conserved domain signature, which is present in a variety of prokaryotic and eukaryotic proteins and is predicted to act as a transcriptional regulator or helicase in an ATP-dependent manner (NCBI-CDD: COG2865)[30]. Subgroup II and III members contain besides the schlafen core region an extra putative unstructured linker region including a conserved stretch consisting of Ser-Trp-Ala-Asp-Leu (SWADL) of unknown function. Subgroup III harbors an additional domain, which shares sequences homologies to SF1 DNA/RNA helicases[30] (Figure 2.6).

Generally, subgroup I and II schlafen members are predominantly located in the cytoplasm[141,142]. Except Slfn1 that was on the one hand found to be localized in the nucleus in $CD4^+/CD8^+$ DP and $CD4^+$ SP T cells, when endogenous Slfn1 levels were monitored by immunoblotting. On the other hand, it was

detected in the cytoplasm, when fused with GFP[31]. Subgroup III members were mainly detected in the nucleus as shown for murine schlafen members[142]. In particular, they show a “speckled” staining pattern in the nucleus[142], suggesting that they localize to the splicing machinery or to active sites of transcription, which often show similar patterns[143,144]. Human SLFN5 was also found to be expressed in the nucleus in malignant melanoma cells after treatment with IFN-α as detected by immunofluorescence[55]. SLFN11 was detected in the nucleus at stalled replication forks in leukemia cell lines[132]. Controversial, equine Slfn11 was detected in the cytoplasm of HEK 293T cells after infection with equine infectious anemia virus (EIAV)[83]. The viral schlafen member, which consists of the schlafen core region and an N-terminal viral specific domain of unknown function, is located in the cytoplasm of the host cell[140].

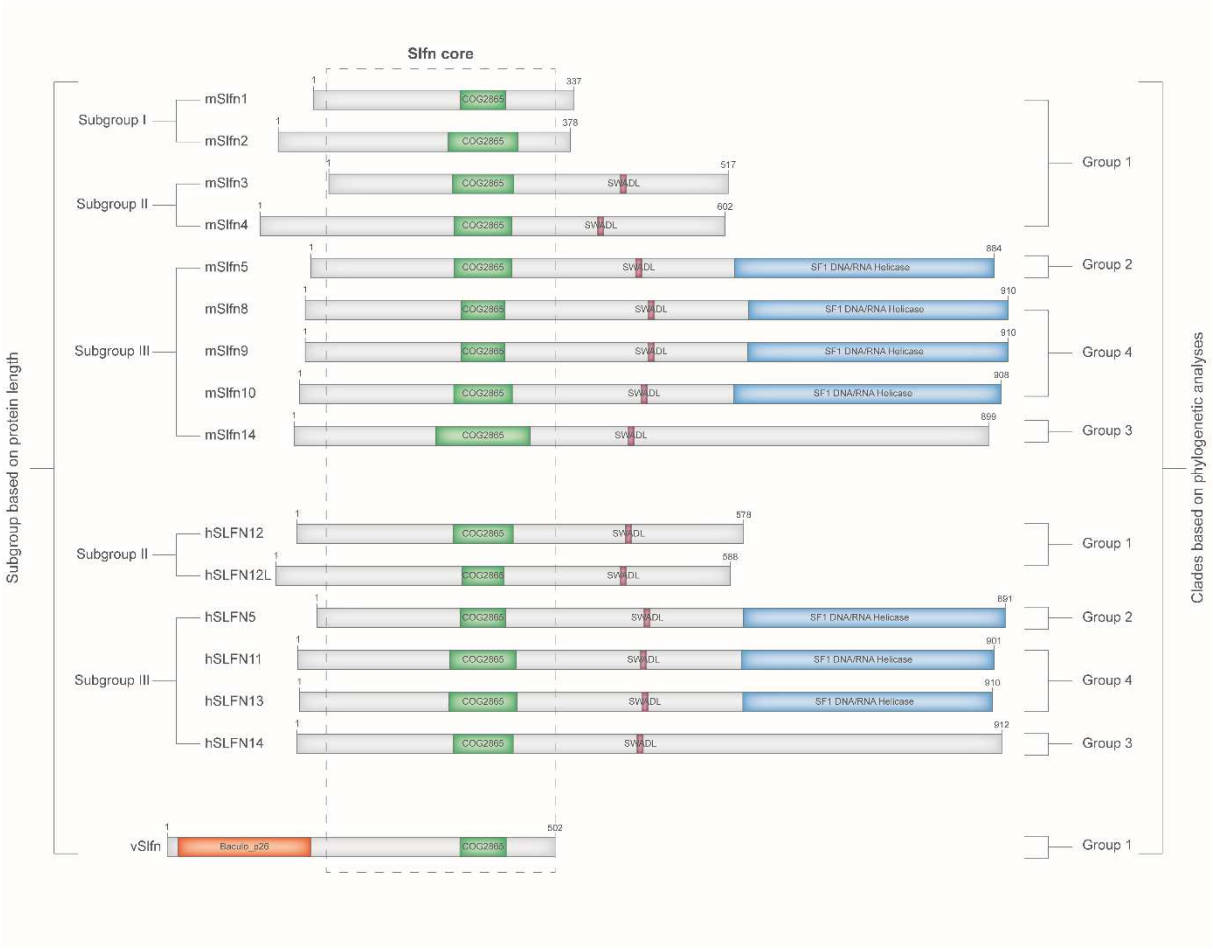



**Figure 2.6:** Up to date nine murine, six human and one viral schlafen protein members have been identified. The different members can either be divided into four different phylogenetic clades[114] or three subgroups according to the length of the encoded protein[30]. All schlafen family members share the slfn core region consisting of a highly conserved sequence motif that resembles similarities to a putative ATPase domain (NCBI-CDD: COG2865). Subgroup III members comprise an additional C- terminal extension, a predicted SF1 DNA/ RNA helicase domain[30]. Adopted and modified from[145].

### 2.6. From structure to putative function

For several schlafen members endoribonuclease activity has been observed[126,136]. Rabbit and human, C-terminally truncated SLFN14 (residues 1 to 400), were shown to bind ribosomes and cleave rRNA[136]. This is in line with another study that confirmed preferential tRNA and to a lesser extend rRNA cleavage by rat Slfn13 residues 14 to 353 (rSlfn13$^{14\text{-}353}$) and its human (hSLFN13$^{1\text{-}355}$) and mouse (mSlfn8$^{1\text{-}359}$) orthologues *in vitro*. No cleavage was observed for murine Slfn1$^{FL}$ and human SLFN5$^{12\text{-}334}$ [126]. Both, rSlfn13 and SLFN14 belong to subgroup III schlafen, and endonucleolytic activity has so far exclusively been observed with C-terminally truncated constructs in an $Mg^{2+}/Mn^{2+}$ dependent and NTP-independent manner[126,136]. By solving the crystal structure of rSlfn13$^{14\text{-}353}$, which resembles a U-pillow shape consisting of one N-lobe, one C-lobe and two bridging domains, a putative active site consisting of glutamate and aspartate residues (E205, E210 and D248) could be mapped to the C-lobe (Figure 2.7. A-C). The two lobes with the respective bridging domains are pseudo symmetric to each other. The tunnel between the pseudo-dimers reveals positive charged patches and may clamp base-paired RNAs[126]. The cleavage site within the tRNA could be mapped 11 nucleotides from the 3'-terminus of the tRNA acceptor stem (Figure 2.7. D). The authors propose binding of the RNA stem by the tunnel and subsequent nucleolytic cleavage in the putative active site (Figure 2.7. C).

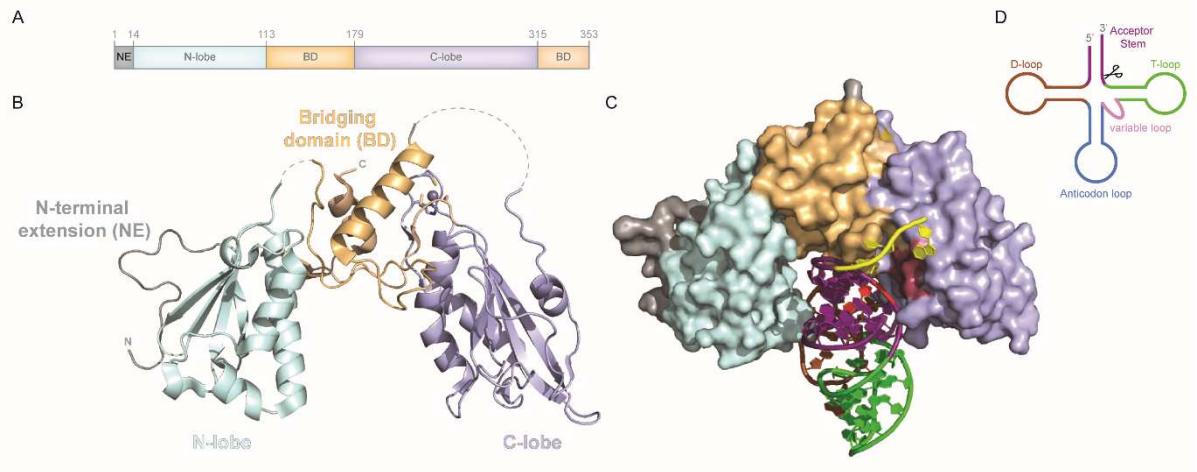



**Figure 2.7**: Overall structure of rat Slfn13$^{14\text{-}353}$ (rSlfn13$^{14\text{-}353}$) (PDB: 5YD0) and predicted mode of action for endonucleolytic tRNA cleavage. **A:** Domain architecture of rSlfn13$^{14\text{-}353}$, consisting of an artificial N-terminal extension (NE), an N-lobe, a C-lobe and two intraconnecting bridging domains (BD). **B:** Overall structure of rSlfn13$^{14\text{-}353}$ resembling a U-pillow shape. The domains are colored as in A. **C:** Model of predicted tRNA binding by rSlfn13$^{14\text{-}353}$ and endonucleolytic cleavage by the putative active site (colored in red). The tRNA (PDB: 5E6M, excerpted) is colored as in D and was modeled into rSlfn13$^{14\text{-}353}$. **D:** Secondary structure representation of a tRNA. The predicted nucleolytic cleavage site lies 11 nucleotides from the 3' end of the acceptor stem and is indicated by scissors.

### 2.7. Mouse Schlafen 2 plays key roles in T cell quiescence

Mature lymphocytes are usually found in a quiescent state, i.e. a resting and non-proliferating condition, until they are re-activated by foreign antigens. Lymphocyte quiescence prevents nonspecific widespread immune activation and destruction, and is characterized by a smaller cell phenotype that is metabolically less active and resistant to apoptosis. This state is reversible, that means, that the T cells are able to proliferate again after antigen stimulation[146]. Until recently, the quiescent state of T lymphocytes was thought to be due to a lack of stimulation by proliferation signals[147]. However, both, the activation of 'quiescence' genes and the inhibition of 'activation' genes are required[148-150]. Since Slfn1 and Slfn2 are mainly expressed in resting T cells it is speculated that those are required for maintaining the quiescent state[24]. Indeed, one of the key players in maintaining T cell quiescence in mice seems to be murine Slfn2. Berger et al., 2010[86], described the *elektra* mouse, which carries a point mutation in the s*lfn2* gene that was induced by treatment with N-ethyl-N-nitrosourea. This mutation leads to an isoleucine to asparagine substitution at position 135 within the Slfn2 protein. *Elektra* mice show an immune suppressed phenotype with a significantly decreased amount of peripheral $CD4^+$ and $CD8^+$ T cells, while the B lymphocytes are not affected. The existing T cells are in a semi-activated state and undergo apoptosis via the intrinsic apoptotic pathway upon antigen stimulation. Slfn2 might maintain the quiescent state by promoting the expression of 'quiescence' genes and inactivation of genes required for proliferation or differentiation[86] (Figure 2.8).

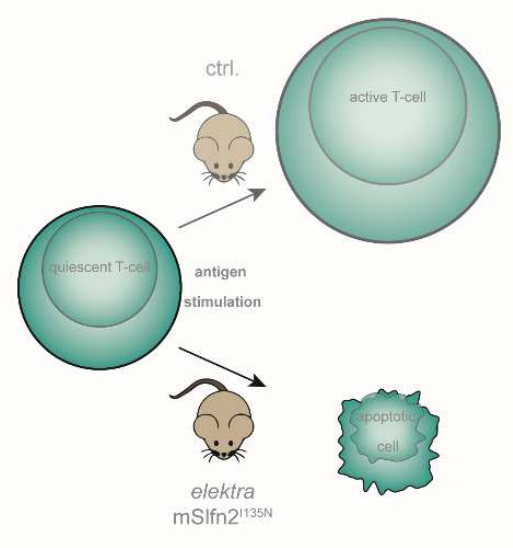



**Figure 2.8:** Quiescent T cells are small, metabolically less active cells that are able to proliferate after antigen stimulation. Quiescent T cells of *elektra* mice, which carry a point mutation in the *slfn2* gene leading to an isoleucine to asparagine substitution in the gene product, fail to proliferate after stimulation and undergo apoptosis[86].

More studies have linked the loss of quiescence in *elektra* mice to chronic endoplasmic reticulum (ER) stress[89] and to a dysregulated sterol homeostasis[121]. T cells undergo drastic metabolic changes upon the transition from the resting to the active state, for example rapid upregulation of lipid-biosynthesis

pathways[151]. T cells and monocytes from *elektra* mice displayed elevated levels of lipid rafts and lipid droplets. Furthermore, these cells showed increased levels of the enzyme HMG-CoA (3-hydroxy-3-methylglutaryl-Coenzyme A) reductase[121], an enzyme involved in sterol synthesis[152]. Consequently, *elektra* T cells contain elevated levels of free cholesterol and cholesteryl ester. These findings suggest that Slfn2 influences quiescence by the regulation of cholesterol and lipid levels through the prevention of ER stress[121]. Since the effect of the *elektra* mutation in Slfn2 seems to be specific for T cells and monocytes, targeting Slfn2 might possibly be harmless to other cell types. This could be exploited for the treatment of malignancies, such as T cell acute lymphoblastic leukemia (T-ALL), an aggressive malignancy of thymocytes. Indeed, it could be shown that inhibition of T cell quiescence by impairing the function of Slfn2 can prevent the development and proliferation of T cell leukemia by driving the malignant cells into a post-mitotic phase[88].

### 2.8. Human Schlafen 5 and its role in carcinogenesis

A subgroup III schlafen protein with nuclear localization[55] is human SLFN5. It exhibits essential roles in the proliferation and development of tumor cells. However, these actions seem to depend on the tumor type, since it can either inhibit [54,55] or stimulate[61,128] tumorigenesis. Many studies investigating the regulatory role in tumorigenesis of SLFN5 have been conducted, however, the exact molecular working mechanism remains unknown.

Compared to the other human schlafen proteins, SLFN5 appears to play a special role in tumorigenesis, because only SLFN5 and not SLFN11, SLFN12 or SLFN13 expression is decreased in malignant melanoma cells, suggesting a selection against SLFN5 in these tumor cells. Furthermore, only SLFN5 is upregulated in IFN-sensitive melanoma cells upon IFN treatment, whereas in healthy melanocytes SLFN11 and SLFN12 are also inducible. When stably knocking out SLFN5 in malignant melanoma cells, the tumor cells displayed a more aggressive, tumor-like behavior indicated by increased anchorage independent growth, enhanced colony formation in soft agar and increased invasion into three-dimensional collagen[55] (Figure 2.9. A). These findings potentially link SLFN5 to the enhancement of antineoplastic effects of IFN-α by regulating the invasion of the tumor cells[55]. In further studies, it was shown that SLFN5 plays key roles in renal cell carcinoma (RCC) tumor progression[54]. Again, of all human schlafen proteins, only SLFN5 was found to be upregulated in RCC and knockout of SLFN5 lead to drastic changes in tumor cell morphology, depicted in a spindle-shaped instead of epithelial-shaped cell form. Furthermore, upon knockout a dramatic decrease in the focal adhesion complex, loss of stress fibers, a reduction in cell size, increased cell motility and migration in a transwell migration assay were observed (Figure 2.9. A). Knockout of the other human schlafen proteins did not lead to those observations. RNA-sequencing analysis of SLFN5 knockout cells identified that cancer-related pathways and pathways involved in focal adhesion and cell junctions were affected most severely. Furthermore, the top upregulated genes are coding for metalloproteinase (MMP) 1 and 13. Consequently, if knocking down MMP-1 and -13, the phenotype reversed, suggesting that the increased motility observed upon SLFN5 targeting is dependent on the functions of these metalloproteinases.

Moreover, SLFN5 was significantly overexpressed in patients suffering from RCC and a higher expression of SLFN5 was significantly associated with a better overall survival[54].

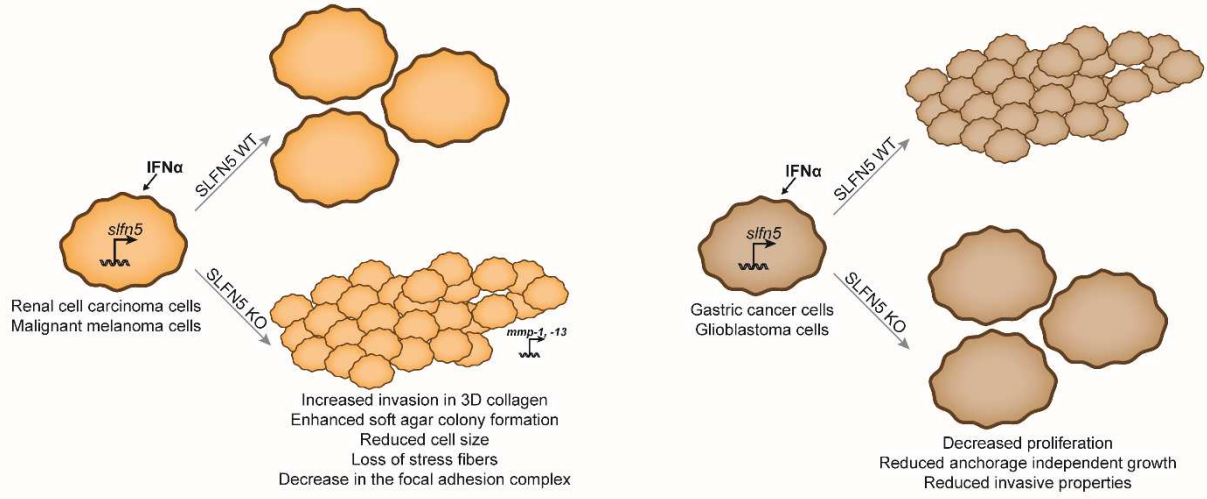



**Figure 2.9:** SLFN5 has contradictive functions in tumorigenesis. **A:** In renal cell carcinoma and malignant melanoma cells SLFN5 knockout results in increased proliferation and aggressiveness of the tumor cells. **B:** In gastric cancer and glioblastoma cells SLFN5 knockout leads to decreased proliferation and reduced motility.

Controversially, SLFN5 appears to have contrary effects in glioblastoma multiform and gastric cancer. SLFN5 was mainly expressed in cancer cells of patients suffering from intestinal metaplasia that progressed to gastric cancer. Since SLFN5 expression correlated with the histological diagnosis of intestinal metaplasia that developed to gastric cancer, it potentially can function as a future biomarker[128]. In glioblastoma multiform (GBM) high levels of SLFN5 expression correlate with high-grade gliomas and consequently, increased SLFN5 expression is associated with a shorter overall survival. Although here, not only SLFN5 but also SLFN11 and SLFN12 expression levels correlated with the progression of the glioma grade, particularly SLFN5 was overexpressed in malignant brain tumor cells. Upon knocking out SLFN5, the glioblastoma cell lines showed diminished proliferation, decreased anchorage independent growth in soft-agar and reduced invasive properties of the tumor cells (Figure 2.9. B)[61]. In mice lacking a thymus gland, SLFN5 knockout even results in delayed tumor formation and in decelerated tumor growth *in vivo*. Since SLFN5 interacts with STAT1 and functions not only as an ISG, but also as a repressor of ISG transcription, the existence of a negative-feedback regulatory loop is speculated that is responsible for the suppression of the antitumor response in glioblastoma[61]. The development of drugs that specifically target SLFN5 may potentially be successful to treat resistant glioblastoma and to eliminate glioma cancer stem cells[61].

### 2.9. Human Schlafen 11

Human SLFN11 plays not only important roles as HIV-restriction factor as described in section 2.3, but appears also to play a role at damaged DNA replication forks[109,132,133] upon treatment with poly (ADP-ribose) polymerase (PARP)- or topoisomerases (TOP) -inhibitors, alkylating agents or DNA synthesis inhibitors[104,106,107,153-155]. SLFN11 is as well biochemically largely uncharacterized and the exact molecular function remains speculative.

#### 2.9.1. SLFN11 increases the effects of chemotherapeutics to cancer cells

Prominent drugs for the use in cancer treatment are TOP inhibitors[156], alkylating agents[157], DNA synthesis inhibitors[158] or PARP inhibitors[159]. The activities of topoisomerases are indispensable for DNA maintenance and metabolism, because their main task is to relax DNA supercoiling generated by transcription, replication and chromatin remodeling. Topoisomerases are divided into two subgroups: TOP1, which cuts one strand of the DNA duplex, and TOP2, which induces double-strand breaks. Thus, TOP inhibitors display powerful anticancer drugs, since they can trap TOP-DNA complexes and stop essential cellular processes in the tumor cell[156]. Other important anticancer drugs are alkylating agents, which trigger the alkylation of the DNA during replication, thus introducing errors within new synthesized DNA[157], or DNA synthesis inhibitors, which affect the S-phase of the cell cycle for example by mimicking nucleic acid building blocks, thereby blocking synthesis[158]. PARP1, which is an abundant protein located in the nucleus, attaches poly (ADP-ribose) (PAR) to itself and to other target proteins at DNA lesions. This so-called PARylation, is one of the major functions of PARP1 in the repair of DNA lesions, amongst them single-strand and double-strand breaks (SSB and DSB). PARP1 is furthermore important in the stabilization of DNA replication forks and in chromatin remodeling. Due to those important cellular functions, PARP inhibitors, are an attractive way for the treatment of cancer cells, especially those that are deficient in the repair of DSBs by homologous-recombination[reviewed in 160]. Chemotherapeutics like PARP inhibitors or DNA damaging agents (DDA), including platinum derived DNA alkylating agents, can have severe side effects. Therefore, it would be from great advantage to directly choose the optimal treatment. Biological biomarkers could help choosing the optimal treatment for the respective patients[161,162]. SLFN11 could be one of those desired clinical biomarkers[163]. Recent bioinformatic analysis of cancer cell databases revealed the importance of SLFN11, sensitizing widely used DDAs, including TOP1 inhibitors (camptothecin, topotecan, and irinotecan), TOP2 inhibitors (etoposide, mixantrone, and doxorubicin), alkylating agents (cisplatin and carboplatin), DNA synthesis inhibitors (gemcitabine and cytarabine)[103,104,106,107,113] and PARP inhibitors (temozolomide)[108,109,130]. Analysis of cells from different tissues revealed that SLFN11 was the only human schlafen protein that showed a significant positive correlation in expression levels and cytotoxicity of DDAs. In high SLFN11 expressing cells, the camptothecin-induced cell cycle arrest stopped in the early S-phase, whereas SLFN11 silenced cells progressed further into S and G2/M phase. Furthermore, SLFN11 silenced cells exhibited a significant reduced sensitivity to DDAs, demonstrating, that SLFN11 confers sensitivity to TOP1 and TOP2 inhibitors, as well as alkylating agents and DNA synthesis inhibitors[104]. Moreover, SLFN11 also appears to play key roles in cell cycle arrest and induction of apoptosis in response to irinotecan-induced DNA damage in colorectal carcinoma cells (CRC)[103]. The clinical application for SLFN11 as a biomarker upon DDAs treatment is, however, exclusively suitable for cells with high endogenous expression

levels[163], since sarcoma cell lines expressing low SLFN11 levels showed no effect on camptothecin treatment[106].

As gene silencing by increased CpG promoter methylation is a common event in cancer cells[164], the so-called hypermethylation of promoters could serve as a suitable biomarker for cancer progression or drug response[165,166]. Indeed, it was shown in various cancer cell types that the human *SLFN11* gene is frequently methylated, and the expression of *SLFN11* is regulated by hypermethylation of CpG islands in the promoter region[104,110,113]. The hypermethylation-associated silencing of *SLFN11* could also be directly linked to increased resistance to DDAs and PARP inhibitors[104,113]. Consequently, patients exhibiting *SLFN11* epigenetic inactivation revealed a poor response to DDAs[107]. An example are small-cell lung cancer patients, which are primarily treated with cisplatin and etoposide and who usually develop chemoresistance within one year[167]. By using patient-derived xenografts, it could be shown that *SLFN11* expression is suppressed in chemoresistant models. Further, silencing of *SLFN11* is highly associated with a histone methylation H3K27me3 at the *SLFN11* open reading frame[110]. This epigenetic modification is generated by Enhancer of Zeste homolog 2 (EZH2), a polycomb group protein that serves as a recruitment platform for DNA methylases for epigenetic gene silencing[168]. Consequently, the local chromatin condenses and the genes in this area, including *SLFN11*, are silenced, which in turn leads to the development of a resistance against the chemotherapeutics. By applying EZH2 inhibitor *SLFN11* would not be silenced and the chemoresistance could be circumvented[110].

### 2.9.2. SLFN11 and its role at damaged replication forks

Since TOP or PARP inhibitors damage DNA by trapping TOP1 cleaved complexes[156] or PARP1/2-DNA complexes[169], respectively, SLFN11-enhanced cancer cell killing might function via a replication stress-associated mechanism[132]. Specifically, upon treatment with PARP or TOP inhibitors an S-phase arrest is induced by SLFN11. Consequently, apoptosis is triggered causing hypersensitivity to the inhibitors[109]. The irreversible and lethal replication inhibition is induced by SLFN11 independently of Ataxia telangiectasia and Rad3-related (ATR), because ATR inhibition overcomes the resistance of SLFN11-deletion cells to those inhibitors[109,132]. ATR is a critical S-phase checkpoint protein kinase that induces a global shutdown of DNA replication initiation (origin firing) throughout the genome upon its activation and slows down replication fork speed. ATR is activated by the single-stranded DNA- replication protein A (ss DNA-RPA) complex at stalled replication forks and DNA end-resection sites[170]. ATR then triggers the activation of checkpoint kinase 1 (CHK1) by phosphorylation, which in turn inactivates cyclin-dependent kinase (CDK). Under normal conditions CDK promotes the loading of replication factors such as cell division cycle 45 (Cdc45)[171] and GINS, which recruits the replication helicase minichromosome maintenance complex 2-7 (MCM)[172]. Phosphorylation of CHK1 by ATR prevents unscheduled replication origin firing[173]. During replication stress, SLFN11 is recruited to the stressed replication fork to ss DNA via RPA, where it inhibits replication[132,133]. SLFN11 blocks the replication by binding to the nascent DNA and interacts with the helicase subunit MCM3. Interestingly, SLFN11 neither blocks the initiation of replication nor interacts with Cdc45[132]. Crucial for SLFN11 activity is the C-terminal helicase domain, because cells expressing SLFN11 mutated in the Walker B motif (E699Q) of the helicase domain did not show hypersensitivity to chemotherapeutics compared to cells expressing the wild type protein. Cells expressing either SLFN11$^{E699Q}$ or wild type SLFN11 displayed the same replication stop.

However, co-treatment with an ATR inhibitor enabled replication in the mutated cells, but not in the wild type version. This suggests that the C-terminal helicase domain is essential for drug-induced cell killing and replication block[132].

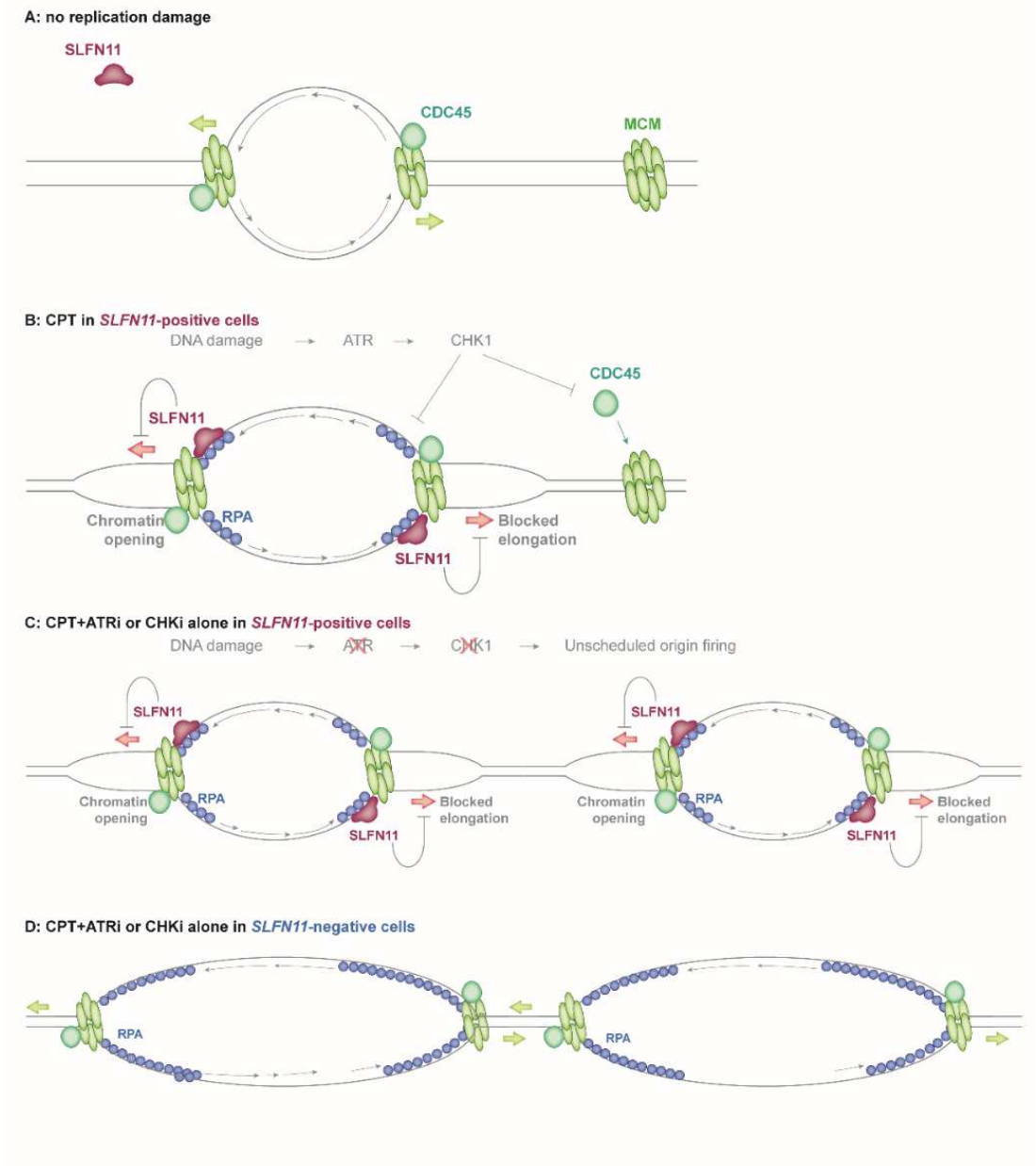



**Figure 2.10:** Proposed working model for SLFN11 at stalled replication forks. **A:** In stress-free conditions, the replication fork forms short ss DNA fragments coated with RPA and is inaccessible for SLFN11. **B:** Upon camptothecin (CPT) treatment in SLFN11 expressing cells ATR is activated and thus unscheduled DNA replication initiation (origin firing) is prevented. SLFN11 binds to resected DNA ends and to the replication fork together with RPA thereby opening the chromatin and blocking the elongation. **C:** Upon co-treatment of CPT and an ATR inhibitor (ATRi) or a CHK inhibitor (CHKi) alone, many stalled replication forks arise where SLFN11 binds and blocks elongation. **D:** Upon co-treatment of CPT and ATRi in SLFN11-deficient cells, long stretches of ssDNA coated with RPA are generated. Adopted and modified from[132].

Summarizing the findings a SLFN11 working model has been proposed (Figure 2.10): Under normal, stress-free conditions, during cell replication only short stretches of single-stranded DNA are accessible. They are coated by RPA and SLFN11 does not interact with the replication fork. Upon treatment of cells expressing SLFN11 with the TOP1 inhibitor camptothecin, DNA single strand breaks are introduced, leading to the activation of ATR as well as uncoupling of the MCM helicase complex from the DNA polymerase at stressed replication forks. ATR activates CHK1, which halts the replication initiation by inhibiting the loading of Cdc45 and thus prevents unscheduled origin firing. SLFN11 binds to the resected DNA ends and to stressed replication forks together with RPA, where it furthermore interacts with MCM3. This leads to chromatin opening and blocks elongation. When inhibiting ATR, unscheduled replication origin firing occurs, leading to many stressed replication forks upon camptothecin treatment. In SLFN11 expressing cells, SLFN11 binds to stressed replication forks and thus blocks fork progression. Since the helicase domain is required for the activity, SLFN11 might unwind the DNA ahead of the MCM helicase, which in turn blocks the MCM complex and fork progression. In SLFN11-deficient cells, the uncoupling of the replication helicases from the DNA polymerase continues, generating long stretches of ss DNA coated by RPA[132] (Figure 2.10). Taken together, since SLFN11 kills cells with defective replication, it emerges as a 'guardian of the genome'[132]. However, the exact molecular working mechanisms remain elusive.

# 3. Aim of the Thesis

The schlafen protein family is a versatile protein family that are key players in many important cellular processes, like regulation of the cell cycle, T- cell quiescence, differentiation and proliferation of various cell types, inhibition of viral replication and DNA repair. They belong to the interferon-stimulated genes and are also shown to enhance the antineoplastic effects of IFN-α. Furthermore, they could serve as potential biomarkers in clinical applications. Consequently, it is very important to understand the molecular working mechanisms of this interesting family. However, this protein family is very poorly characterized *in vitro* and the molecular working mechanism remains unknown. In the beginning of this work, neither biochemical studies nor structural information were available.

The overall goal of this study was the extensive biochemical and structural characterization of the schlafen protein family, in general, and of human schlafen 5 (SLFN5) and murine schlafen 2 (Slfn2) in particular.

Thereby, the project is divided into following key aims:

First, the structural characterization of the slfn core domain. All members of the schlafen protein family share a highly conserved N-terminal common core region of approximately 340 amino acids. Every schlafen family member at least contains this domain and subgroup I members even exclusively consist of this domain. Consequently, C-terminal truncated versions of SLFN5 with just the N-terminal slfn core domain and full length Slfn2 constructs were generated. A combination of structural methods including X-ray crystallography and small-angle X-ray scattering (SAXS) were chosen.

Second, the biochemical characterization of the slfn core domain. Based on the primary amino-acid sequence it was predicted to function as AAA ATPases. Therefore, after establishing a protocol for the purification of C-terminal truncated SLFN5 and full length Slfn2 biochemical assays investigating the ATPase activity, DNA binding and nuclease activities were performed. The generation of inactive mutants confirmed their biochemical function. Co-immunoprecipitation experiments and subsequent RNA-sequencing, as well as mass spectrometry analysis embedded the Slfn2 protein in its biological context.

Last, the biochemical investigation of full length subgroup III schlafen members. They contain besides the N-terminal schlafen core region a C-terminal helicase domain. Neither biochemical nor structural information about this domain are available so far. Thus, protocols for the recombinant expression trying different expression systems and purification protocols for full length Schlafen 5 and Schlafen 11 were generated. Full length Schlafen 5 could also be partly characterized biochemically.

# 4. Results

## 4.1. Structural and functional analysis of the schlafen core domain

The N-terminal schlafen (slfn) core domain is a highly conserved domain of unknown function and based on the primary amino-acid sequence it is predicted to be an ATPase[141]. However, when this thesis was initiated, neither the exact biochemical role nor the three-dimensional structure of this domain have been known. To address this issue, C-terminal truncated constructs of subgroup III schlafen containing only the N-terminal slfn core domain were generated and characterized. In particular N-terminal human schlafen 5 (SLFN5) was expressed, purified, crystallized and characterized. The same was done for full length murine schlafen 2 (Slfn2), which is a subgroup I schlafen protein and consists exclusively of the slfn core domain. Both crystal structures revealed a novel and so far uncharacterized fold.

### 4.1.1. The slfn core domain of human Schlafen 5

Different C-terminal truncated SLFN5 constructs comprising the slfn core domain were generated, expressed and purified (Table 4.1). Human SLFN5 has a predicted second alternative splicing isoform consisting of the N-terminal 338 amino acids of the full length 891 amino acids protein. A construct cloned in a pET21 vector with amino acids 1-336 could soluble be expressed in *E. coli* and a purification strategy could be established. Despite the good solubility of SLFN5$^{1-336}$, no other construct with additional parts of the linker region and the SWADL motif could be soluble expressed in bacterial expression systems.

**Table 4.1:** Overview of the N-terminal SLFN5 constructs used for expression and purification. The molecular weight is given in kDa and the protein length is indicated in numbers of amino acids (aa). Furthermore, the isoelectric points (pI), the kind of affinity tag, the expression system used and the solubility of SLFN5 are indicated.

| Construct | Description | Length [aa] | MW [kDa] | Theore-tical pI | Affinity tag | Expression system | Solubility and achievements |
|---|---|---|---|---|---|---|---|
| **pET21-SLFN5$^{1-336}$** | Slfn core domain; 2nd alternative splicing isoform | 343 | 38.9 | 6.94 | C-terminal $His_6$ | *E. coli* (Rosetta) | Soluble (crystal structure; DNA exonuclease activity) |
| **pET21-SLFN5$^{1-376}$** | Slfn core + linker | 384 | 43.5 | 7.87 | C-terminal $His_6$ | *E. coli* (Rosetta) | Not soluble |
| **pET21-SLFN5$^{1-447}$** | Slfn core + linker + SWADL | 455 | 51.7 | 7.24 | C-terminal $His_6$ | *E. coli* (Rosetta) | Not soluble |
| **pET21-SLFN5$^{1-502}$** | Slfn core + linker + SWADL | 510 | 57.8 | 8.39 | C-terminal $His_6$ | *E. coli* (Rosetta) | Not soluble |

4.1.1.2. Preparation of N-terminal human Schlafen 5[1-336]

For the biochemical characterization and crystallization, sufficient amounts of highly purified target protein are indispensable. Therefore, the N-terminal first 336 amino acids of SLFN5[1-336], which corresponds approximately to the second alternative splicing isoform, were cloned into a pET21 vector. The encoded proteins carry a C-terminal plasmid-encoded hexahistidinetag (Figure 4.1.1. A) and were expressed in *E. coli* BL21 DE3 (Rosetta) cells.

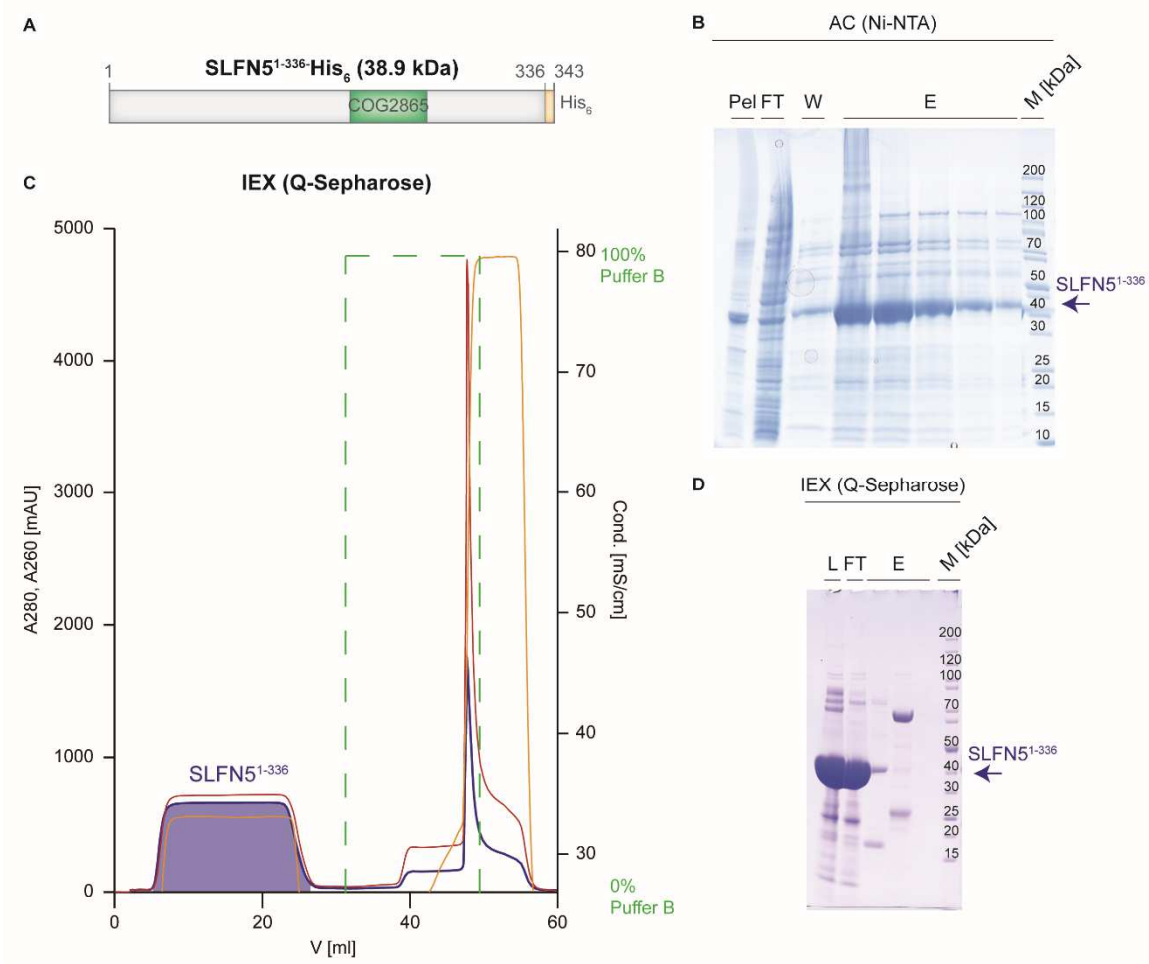



**Figure 4.1.1:** Purification of SLFN5[1-336] from *E. coli* Rosetta. **A:** Schematic overview of the N-terminal SLFN5[1-336] construct carrying a plasmid encoded C-terminal hexahistidine tag. The construct comprises the slfn core domain including the predicted ATPase domain (COG2865; colored in green). **B:** Coomassie-stained SDS-PAGE of fractions collected during affinity chromatography (AC) purification. The wash and the elution fractions contained the SLFN5[1-336] protein (indicated by an arrow). **C-D:** Chromatogram (C) and subsequent Coomassie-stained SDS-PAGE (D) of fractions collected during an ion exchange chromatography (IEX) using Q-Sepharose. SLFN5[1-336] remained in the flow-through, whereas nucleic acid contaminations and other negatively charged impurities bound to the Q-Sepharose. The blue and red lines indicate the absorption at λ=280nm and λ=260nm, respectively. The brown and green lines indicate the conductivity and concentration of IEX B buffer, respectively. E: elution, Sup: supernatant, Pel: pellet, FT: flow-through, W: wash fractions, E: elution fractions, L: load, M: marker in kDa.

The hexahistidine tag was used to capture the protein via Ni-NTA beads and SLFN5$^{1-336}$ could be detected in the wash and elution fractions (Figure 4.1.1. B). SLFN5$^{1-336}$ has a theoretical molecular weight of 38.9 kDa, thus it runs slightly below the 40 kDa marker protein. The elution fractions had an $A_{260}/A_{280}$ ratio of 1, indicating a severe contamination with nucleic acids. The nucleic acid contaminations and a 70 kDa impurity were removed by applying the sample onto a Q-Sepharose column. SLFN5$^{1-336}$ stayed in the flow-through and the unspecifically bound DNA interacted with the positively charged matrix (Figure 4.1.1. C-D). Next, remaining protein impurities and aggregates were separated from SLFN5$^{1-336}$ by size exclusion chromatography (SEC). SLFN5$^{1-336}$ elutes as one single peak at about 90 ml from a S200 16/60 column (Figure 4.1.2. A). A subsequent SDS-PAGE of the fractions collected during SEC, revealed high amounts of SLFN5$^{1-336}$ with a minor impurity running at approximately 20 kDa (Figure 4.1.2. B). Finally, 4 mg of pure SLFN5$^{1-336}$ per liter *E. coli* expression culture were obtained. The $A_{260}/A_{280}$ ratio of 0.53 indicates that the sample was not contaminated with DNA. The SLFN5$^{1-336}$ constructs carrying point mutations were expressed and purified following the same strategy.

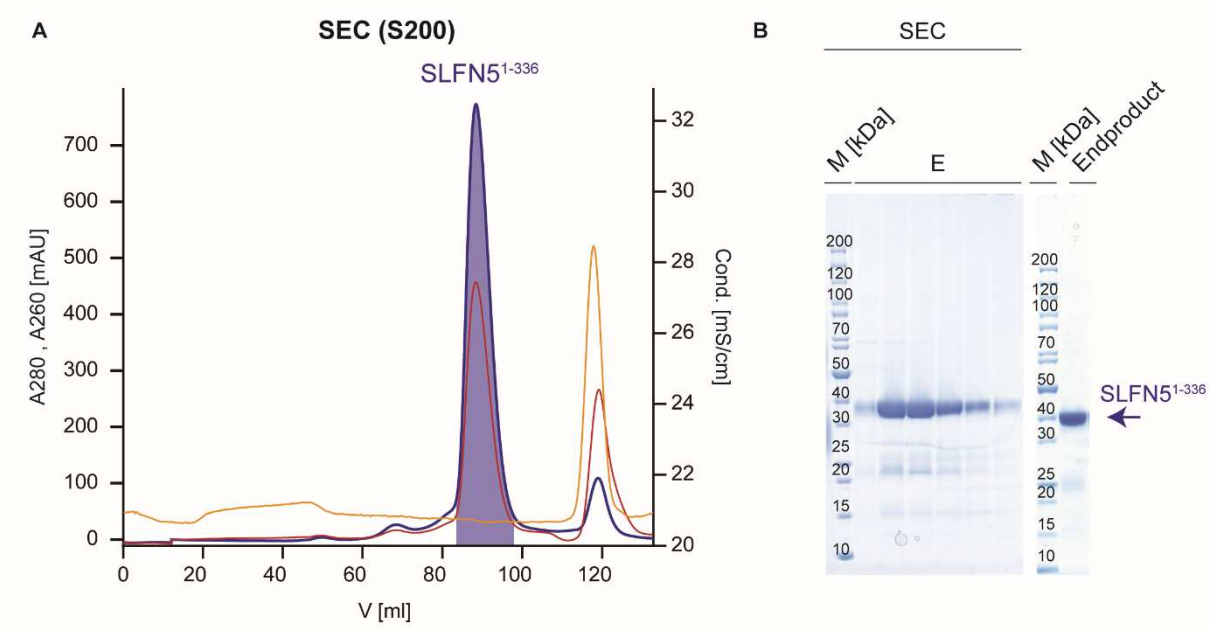



**Figure 4.1.2:** Size exclusion chromatography of SLFN5$^{1-336}$. **A-B:** Chromatogram (A) and subsequent Coomassie-stained SDS-PAGE (B) of the fractions collected after size exclusion chromatography (SEC) using a S200 16/60 column. SLFN5$^{1-336}$ eluted as one single peak at approximately 90 ml. The blue and red lines indicate the absorption at λ=280nm and λ=260nm, respectively. The brown line indicates the conductivity. E: elution.

4.1.1.3. Structural characterization of Schlafen $5^{1-336}$

After successful protein purification, the integrity of $SLFN5^{1-336}$ was verified by analytical size exclusion chromatography (SEC) on a S200 10/300 column and SEC-coupled right angle light scattering (RALS). $SLFN5^{1-336}$ elutes at one single peak at approximately 15 ml (Figure 4.1.3. A). Subsequent RALS-analysis revealed, that $SLFN5^{1-336}$ has an observed molecular weight of 43.7 kDa (Figure 4.1.3. B). Since the theoretical molecular weight of $SLFN5^{1-336}$ is 38.9 kDa, the protein appears to be in a monomeric state in solution. The quotient of molecular weight and number average molecular weight (MW/ Mn)[174] is an indicator of the polydiversity of a sample. For $SLFN5^{1-336}$ it is 1, indicating that the solution is monodisperse.

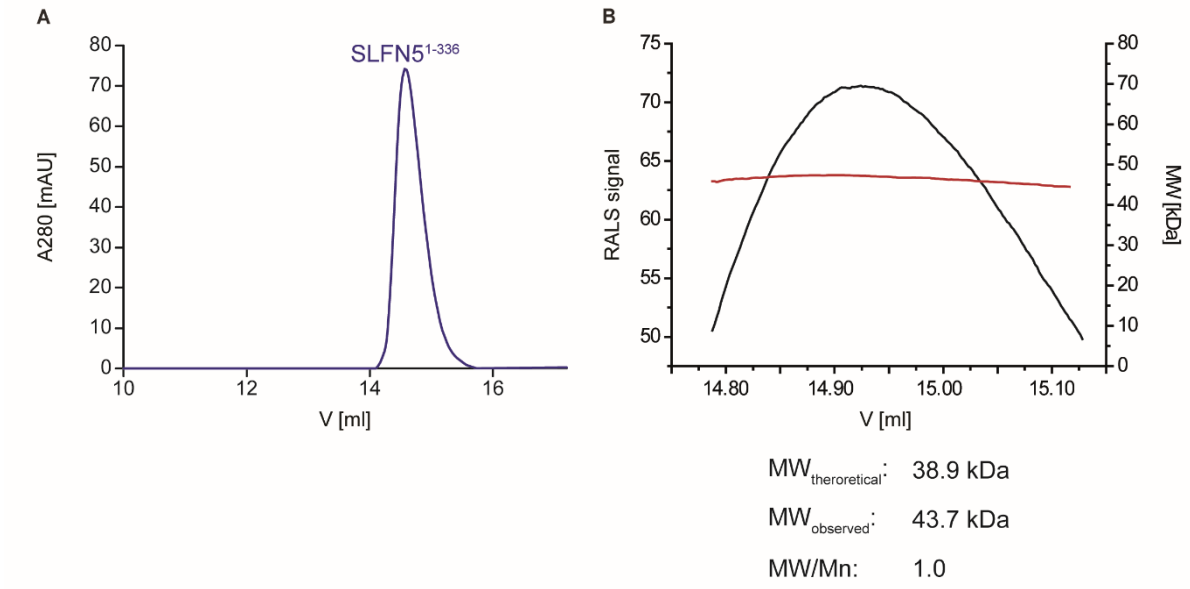



**Figure 4.1.3:** Molecular weight determination of $SLFN5^{1-336}$. **A:** Analytical size exclusion chromatography using a S200 10/300 column. $SLFN5^{1-336}$ elutes as one single peak at approximately 15 ml. The blue line indicates the absorption at λ=280 nm. **B:** Right-angle light scattering (RALS) coupled to SEC. The $SLFN5^{1-336}$ sample is monodisperse and the observed molecular weight is 43.7 kDa, which corresponds to a monomer in solution. MW: molecular weight, Mn: number average molecular weight[174].

For further verification that $SLFN5^{1-336}$ is indeed monodisperse in solution, small-angle X-ray scattering analysis (SAXS) was performed. The primary scattering curve of $SLFN5^{1-336}$ at a concentration of 2.2 g/l was the best obtained, showing proper folded protein in solution with minor aggregations (Figure 4.1.4. A). The scattering curve of the buffer subtracted raw data already suggests a monomeric protein in solution (Figure 4.1.4. A). Further information about the shape of the molecule can be obtained from a calculated pair distribution function, which reveals a parable shape with a second maximum at higher radiuses. This indicates that $SLFN5^{1-336}$ is a single-domain and globular protein with some minor extended regions (Figure 4.1.4. B). The Guinier plot analysis reveal that the sample indeed had some aggregates (Figure 4.1.4. C) and the corresponding data points were excluded from the linear regression to calculate the radius of gyration ($R_G$). The $R_G$ is an approximation of the size of the particle and could

be determined for SLFN5$^{1-336}$ to be 3.75 nm. The subsequent estimated molecular weight is 49 kDa. This indicates that SLFN5$^{1-336}$ is monomeric in solution, since its theoretical molecular weight is 38.9 kDa. A low-resolution, *ab initio* model could be generated, which indeed confirms the globular, single-domain shape with some minor extensions (Figure 4.1.4. D).

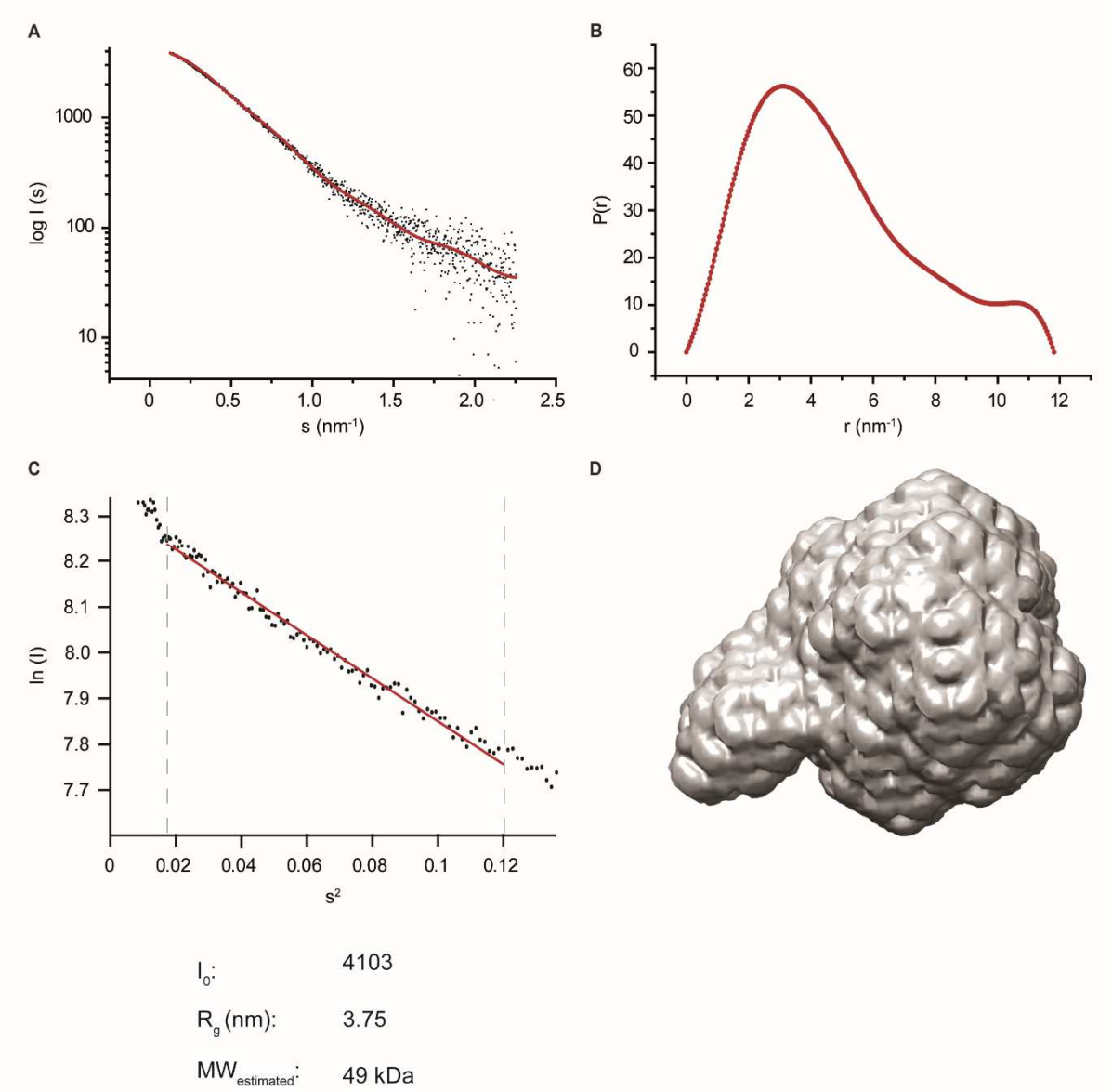



**Figure 4.1.4:** Analysis of SLFN5$^{1-336}$ small-angle X-ray scattering (SAXS) data. **A:** Buffer corrected SAXS curve of a 2.2 g/L SLFN5$^{1-336}$ sample. **B:** A pair distribution function (P(r)) analysis suggests a single domain protein. **C:** Guinier plot analysis. An $R_G$ of 3.75 nm was determined from the slope of the linear regression. The sample tends to aggregate, thus the estimated molecular weight of 49 kDa is only an approximation. **D:** Final averaged *ab initio* model of SLFN5$^{1-336}$ in solution.

After analyzing SLFN5$^{1-336}$ via SEC-RALS and SAXS in solution, the crystal structure of the slfn core domain of SLFN5 was furthermore tried to solve. Crystallization of SLFN5$^{1-336}$ resulted in two different crystal forms and for both the structures were determined. Crystal form 1 was growing within precipitate, had a cubic form of about 50 µm *50 µm *50 µm in size and polarized light (Figure 4.1.5. A-B). The crystals diffracted to 1.8 Å (Figure 4.1.5. C) and belonged to the monoclinic space group $P2_1$ (space group number 4).

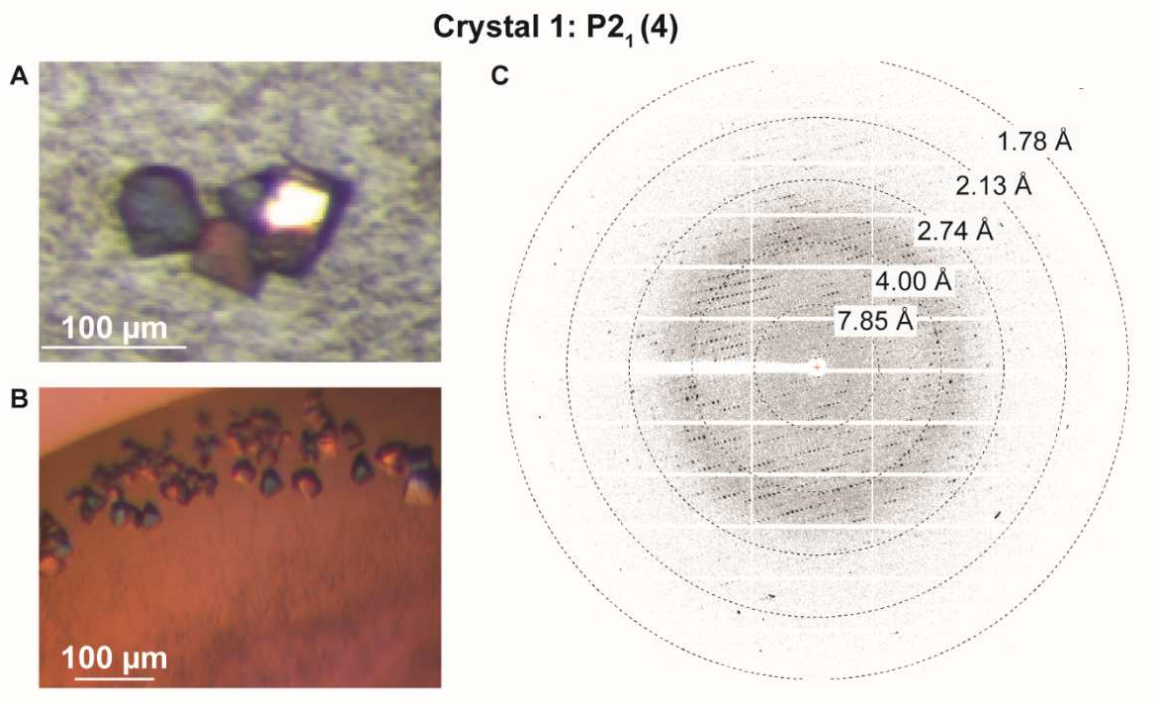



**Figure 4.1.5:** Crystallization of SLFN5$^{1-336}$ crystal form 1. **A-B:** Crystal form 1 growing in 200 mM NaCl, 100 mM MES, pH 5.8, 20 % (v/v) PEG 2000 MME. **C:** Diffraction pattern of crystal form 1.

Crystal form 2 was long needles of up to 300 µm length (Figure 4.1.6. A-B). They crystallized in the trigonal space group $P3_221$ (space group number 154) and diffracted to 3.2 Å (Figure 4.1.6. C). Here, anomalous signal at the zinc K-edge was detected, suggesting that SLFN5$^{1-336}$ intrinsically contains zinc (Figure 4.1.6. D). Subsequently, data collection was performed at λ=1.28 Å and the phases were obtained by performing a single-wavelength anomalous dispersion (SAD) experiment by using the intrinsically bound-zinc ion of crystal 2. Thus, an initial structural poly alanine model could be generated. The final model was obtained by molecular replacement with crystal 1 (1.8 Å) dataset using the poly alanine model as search model. The unit cell dimensions and detailed refinement statistics are summarized in table 4.2.

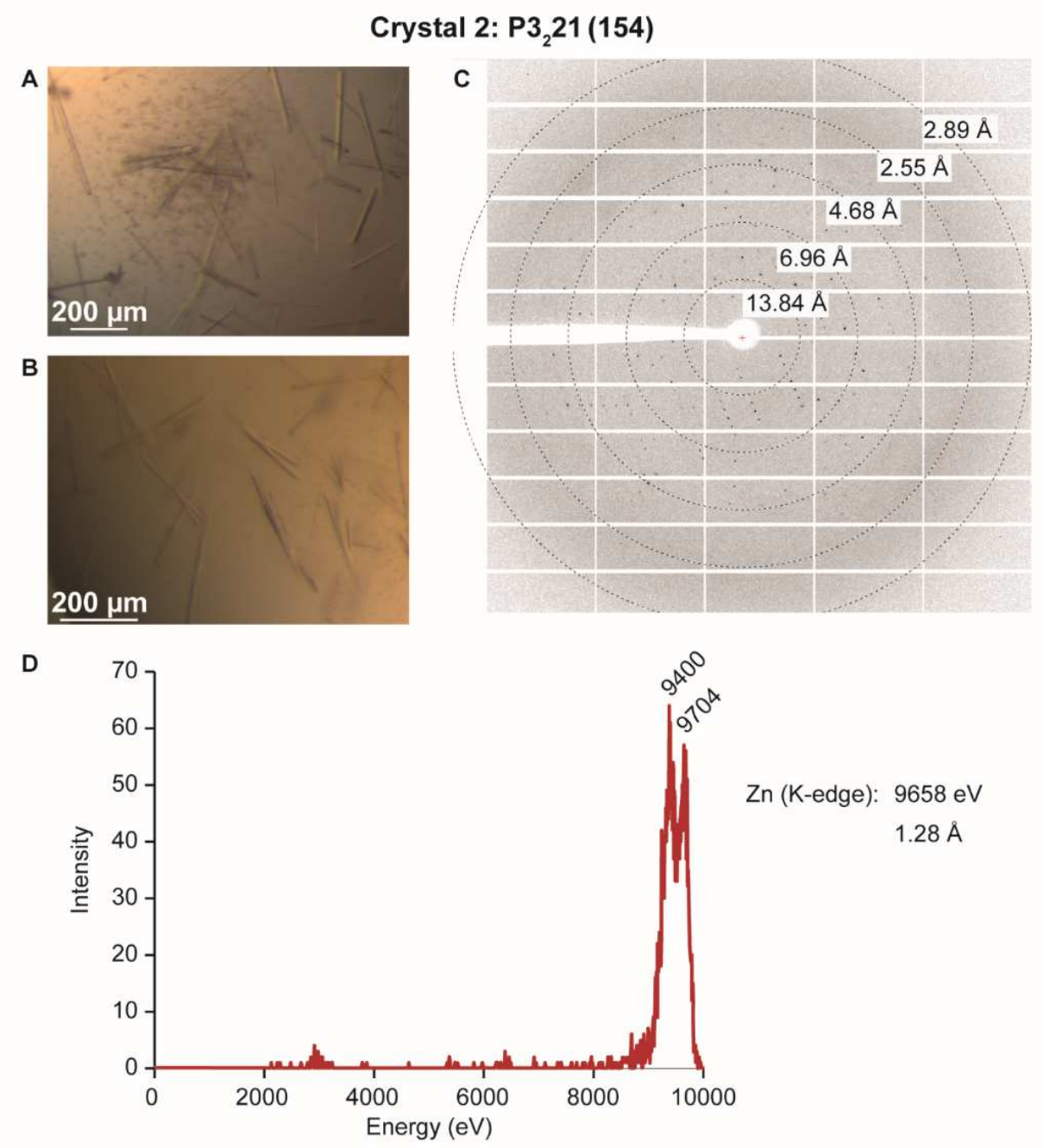



**Figure 4.1.6:** Crystallization of SLFN5$^{1-336}$ crystal form 2. **A-B:** Crystal form 2 growing in 0.1 M sodium acetate, pH 5.0, 1.5 M ammonium sulfate. **C:** Diffraction pattern of crystal form 2. **D:** X-ray fluorescence scan of the SLFN5$^{1-336}$ crystal 2.

**Table 4.2:** Crystal parameters of both crystal forms and refinement statistics.

| Data collection and processing | SLFN5$^{1-336}$ Crystal 1 | SLFN5$^{1-336}$ Crystal 2 |
|---|---|---|
| **Wavelength [Å]** | 1.0 | 1.28 |
| **Space group** | P 1 $2_1$ 1 | P $3_2$ 2 1 |
| **Unit cell dimensions** | | |
| **a, b, c [Å]** | 49.88, 64.04, 104.58 | 101.1, 101.1, 114.3 |
| **α, β, γ [°]** | 90, 93, 90 | 90, 90, 120 |
| **Resolution range [Å]** | 45.9 – 1.85 (1.92 – 1.85) | 50 – 3.4 |
| **$R_{meas}$** | 9.1 (129.9) | 13.0 (103) |
| **I/σI** | 8.98 (1.01) | 12.2 (1.89) |
| **Completeness [%]** | 97.9 (93.2) | 97.6 (86.8) |
| **Redundancy** | 3.4 (3.1) | 8.8 (7.9) |
| **Solvent content [%]** | 44 | 72 |
| **Matthews coefficient [$A^3$/Da]** | 2.2 | 4.44 |
| **Refinement** | | |
| **Number reflections** | 55695 (5454) | 9144 (779) |
| **$R_{work}/R_{free}$** | 0.18/0.21 | 0.21/0.27 |
| **Number of atoms** | | |
| **Protein** | 5025 | 2415 |
| **Ligands** | 11 | 29 |
| **Water** | 305 | |
| **Average B-factor [Å]** | 30.41 | 70.67 |
| **RMSD** | | |
| **Bond lengths [Å]** | 0.008 | 0.012 |
| **Bond angles [°]** | 1.18 | 1.54 |
| **Ramachandran plot** | | |
| **Favored [%]** | 96.85 | 87.41 |
| **Allowed [%]** | 2.65 | 11.22 |
| **Outliers [%]** | 0.5 | 1.36 |

Values in parentheses are for the highest-resolution shell.

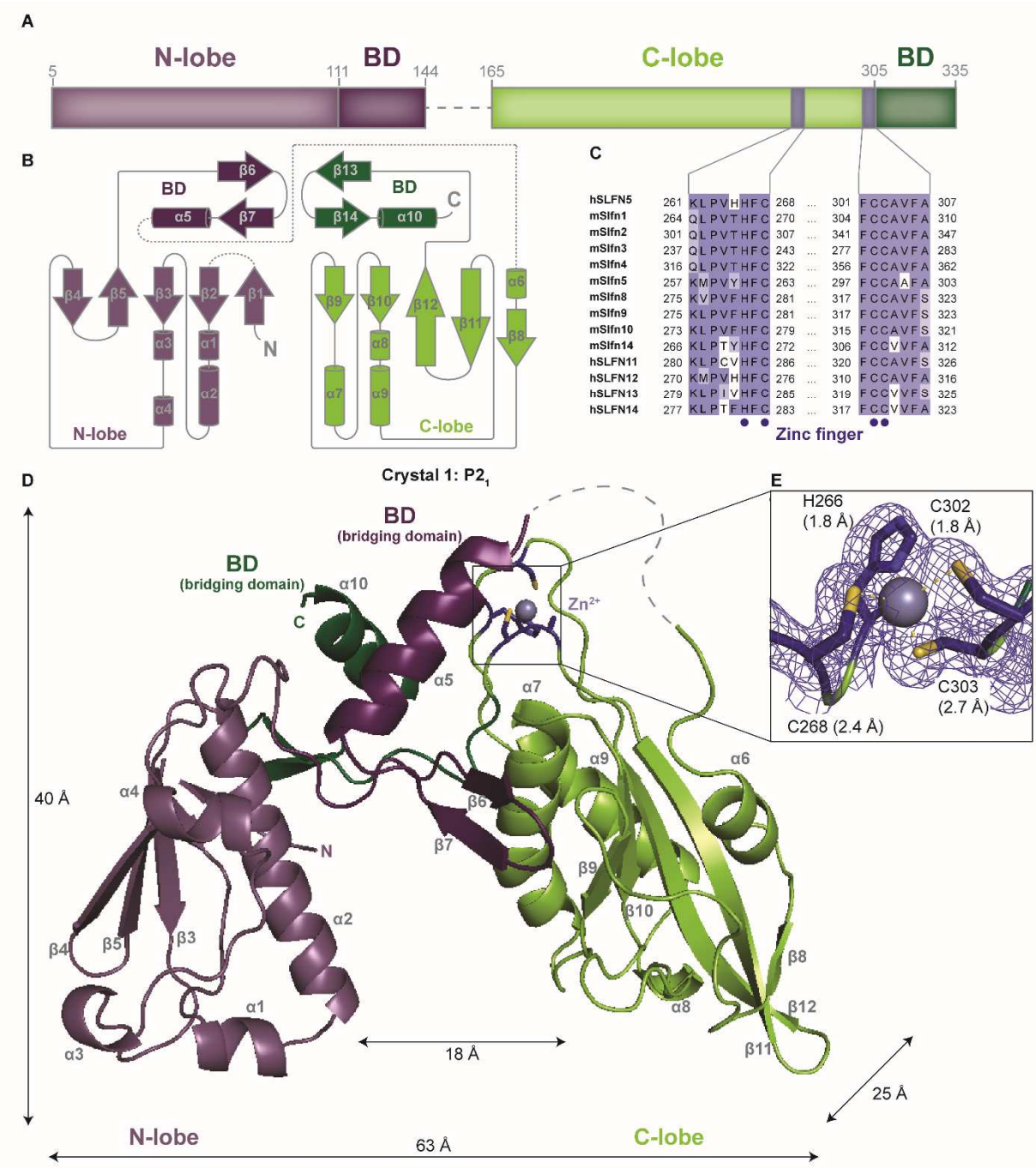



**Figure 4.1.7:** Crystal structure of SLFN5$^{1-336}$. **A:** Schematic overview and domain architecture of the crystallized construct. **B:** Topology diagram of SLFN5$^{1-336}$. **C:** Multiple sequence alignment of the zinc finger region of all human and murine schlafen members. The four residues forming the zinc finger are depicted with blue dots. **D:** Overall structure of SLFN5$^{1-336}$ crystal 1. **E:** Close-up view of the zinc finger motif. Residue numbers and the distance of atoms interacting with the zinc ion are indicated. The $2F_0$-$F_c$ electron density map is colored in dark blue and contoured to σ=1.

Throughout the figure, the following color code was chosen: N-lobe: purple; N-bridging domain: dark purple, C-lobe: green, C-bridging domain: dark green, zinc finger: dark blue. BD: bridging domain.

Electron density was observed for residues 5 to 144 and 165 to 335, in which the three-dimensional structure could be fitted. This suggests, that residues 145 and 164 are not ordered in the crystal. The overall structure of the SLFN5$^{1-336}$ monomer depicts a horseshoe shape with a mixed α/β topology, consisting of ten α-helices and 14 β-sheets (Figure 4.1.7. B). The approximate molecular dimensions are 63x 40x 25 Å. The inner tunnel of the horseshoe has a width of approximately 18 Å (Figure 4.1.7. D). SLFN5$^{1-336}$ can be subdivided into different domains: an N-lobe, a C-lobe and two bridging domains (Figure 4.1.7. A). Each lobe consists of four α-helices and five β-sheets, and each bridging domain of one helix and two sheets (Figure 4.1.7. D). Further, SLFN5$^{1-336}$ possesses a zinc finger motif located in the C-lobe region. The zinc finger consists of the three cysteines C268, C302 and C303 and the histidine H266. The zinc-ion is tightly coordinated within these four residues with distances between 1.8 Å and 2.7 Å (Figure 4.1.7. E). The four residues are highly conserved throughout the schlafen protein family (Figure 4.1.7. C), assuming that every schlafen member contains the identified zinc finger.

The second crystal structure is resolved worse with a resolution of 3.2 Å (Figure 4.1.8. A), but the overall shape is identical to crystal structure 1 (Figure 4.1.8. D). Electron density is missing for residues 23-28, 73-76 and 145-168. However, additional electron density was detected in the C-lobe region in which a sulfate was fitted since the crystallization condition contained 1.5 M ammonium sulfate (Figure 4.1.8. C). The sulfate is mainly coordinated by the amide side chain of arginine 211 and via backbone interactions with the amide group of methionine 199. Sulfates in crystal structures has been found to mimic phosphates of nucleic acids[175-177], suggesting, that this region could be a putative active site. The calculated anomalous map confirmed the presence of a bound zinc ion in crystal 2 (Figure 4.1.8. B). The residues of the zinc finger region and the putative active site are identical in both crystal structures with only slight differences in residue positions within the border of resolution (Figure 4.1.8. D-E).

Taken together, the schlafen core domain represents a novel fold that has not been described before. Further, the C-lobe region contains a putative active site, whose function is not yet defined.

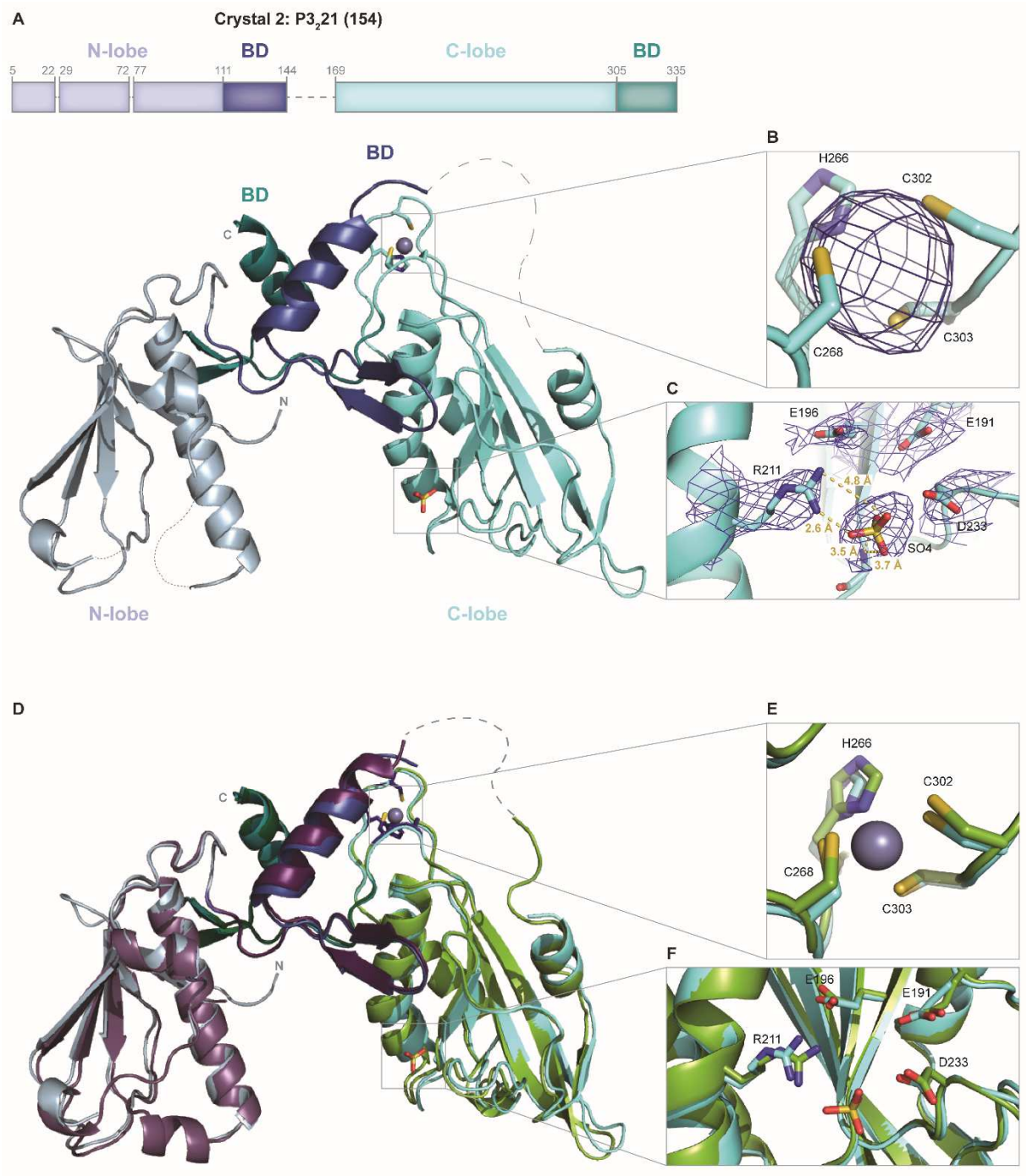



**Figure 4.1.8:** Crystal structure of SLFN5$^{1-336}$ crystal 2 and comparison of both structures. **A:** Schematic overview, domain architecture and overall structure of SLFN5$^{1-336}$ crystal form 2. N-lobe: light blue, N-BD: blue, C-lobe: cyan, C-BD: dark cyan. **B:** Close-up view of the zinc finger motif. The anomalous electron density map is colored in dark blue and contoured to σ=5. **C:** Close-up view of the bound sulfate and neighboring residues. The $2F_0$-$F_c$ electron density map is colored in dark blue and contoured to σ=1. The interaction distances are indicated in yellow. **D:** Superposition of both SLFN5$^{1-336}$ crystal structures. Crystal 1 is colored as in 4.1.7 and crystal 2 as in A. **E:** Superposition of the zinc finger region of both crystals structures. **F:** Superposition of the residues located in the sulfate-binding region.

4.1.1.4. N-terminal human Schlafen $5^{1-336}$ is no ATPase

Based on sequence similarity, the N-terminal schlafen domain was predicted to function as an ATPase[141], thus $SLFN5^{1-336}$ was tested experimentally in terms of ATP binding and ATP hydrolysis. ATP binding was analyzed by performing the radioactive differential radial capillary action of ligand assay (DRaCALA)[36] in absence and presence of a sequence unspecific single-stranded poly (dT) DNA. However, in all tested conditions the spotted [$\alpha^{32}$-P] ATP, moved with the solvent away from $SLFN5^{1-336}$, which remains in the center (Figure 4.1.9. A). This indicates no binding regardless if the substrate is present or absent.

To investigate if $SLFN5^{1-336}$ is able to hydrolyze ATP, a radioactive assay using [$\gamma$- $^{32}$P] ATP was chosen. $SLFN5^{1-336}$ was incubated in absence and presence of the single-stranded 60- mer poly (dT) DNA substrate and ATP hydrolysis was monitored by thin layer chromatography (TLC). In this assay the released $^{32}P_i$ would migrate on the TLC plate. However, again neither in presence nor absence of nucleic acid, an ATP hydrolysis was observed (Figure 4.1.9. B).

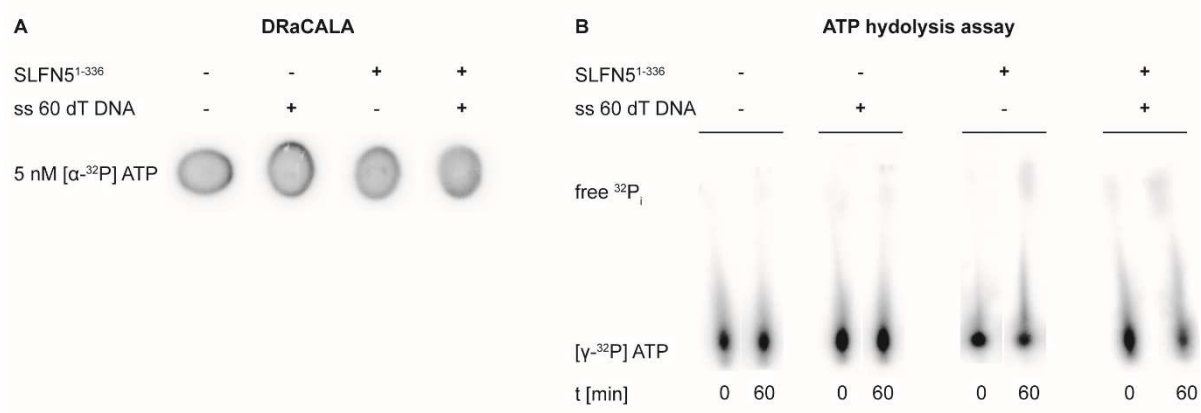



**Figure 4.1.9:** Functional analysis of the predicted ATPase domain. **A:** ATP binding assay of $SLFN5^{1-336}$ by using differential radial capillary action of ligand assay (DRaCALA) in presence or absence of a single-stranded 60- mer poly (dT) DNA substrate. **B:** ATP hydrolysis assay of $SLFN5^{1-336}$ in presence or absence of a single-stranded 60- mer (dT) DNA substrate.

4.1.1.5. Nucleic acid binding properties of Schlafen $5^{1-336}$

Since zinc finger proteins are described to be involved in DNA and RNA binding[178] nucleic acid binding properties of $SLFN5^{1-336}$ with various nucleic acid substrates were tested by electrophoretic mobility shift assay (EMSA) and subsequently analyzed by fluorescence anisotropy (FA) (Figure 4.1.10). First, single-stranded poly (dT) DNA substrates of various lengths were tested in EMSAs. No shift was detected with a single-stranded 10 or 20- mer poly (dT) DNA (Figure 4.1.10. A-B). Slight smearing occurred in EMSAs with the single-stranded 30- mer poly (dT) DNA (Figure 4.1.10. C). Only the single-stranded 60- mer poly (dT) DNA substrate resulted in a DNA shift that however appeared smeary. Additionally, an unspecific band was detected, which runs slightly higher than the free DNA (Figure 4.1.10. D). Nevertheless, the decrease in intensity of the free DNA bands with increasing $SLFN5^{1-336}$ concentrations indicates binding of $SLFN5^{1-336}$ to single-stranded 60- mer poly (dT) DNA. The EMSA of $SLFN5^{1-336}$ with a double-stranded DNA showed a distinct shift starting at 2.5 µM $SLFN5^{1-336}$ protein concentration. Here, no unspecific bands and a clear decrease of free DNA with increasing $SLFN5^{1-336}$ concentrations were observed (Figure 4.1.10. E). The EMSAs indicate that $SLFN5^{1-336}$ binds double-stranded DNA with a higher affinity than single-stranded DNA, since in the EMSA for the double-stranded DNA, the free DNA band vanished completely between the 2.5 µM and 5 µM protein concentration, whereas in the assay with single-stranded DNA even at 10 µM protein concentration a slight free DNA band is still present. FA could confirm the findings. A modest increase in anisotropy signal could be observed with increasing $SLFN5^{1-336}$ concentrations and single-stranded 5, 10 or 20- mer poly (dT) DNA, respectively. For those substrates, the equilibrium dissociation constant ($K_d$-value) could not be calculated (Figure 4.1.10. F). FA analysis revealed better binding of $SLFN5^{1-336}$ to single-stranded 30- mer and 60- mer poly (dT) DNA with $K_d$s of 11 µM and 7.8 µM, respectively. In agreement with the EMSAs, FA also confirmed the increased binding of $SLFN5^{1-336}$ to double-stranded DNA ($K_d$: 2.5 µM) compared to single-stranded DNA (Figure 4.1.10. F).

Since no data of the preferred substrate of $SLFN5^{1-336}$ were available, not only binding to DNA, but also RNA was tested. First, an EMSA was performed with a single-stranded 60- mer poly (rU) RNA (Figure 4.1.11. A). Binding was also detected starting from 2.5 µM $SLFN5^{1-336}$ concentration. Here as well, an unspecific band and some smear were observed. FA analysis confirmed binding of the single-stranded poly (rU) RNA and a $K_d$-value of 3.4 µM could be determined (Figure 4.1.11. B). Furthermore, the length dependency also holds true for RNA molecules. $SLFN5^{1-336}$ binds worse to a shorter single-stranded 30- mer poly (rU) RNA with a calculated $K_d$ of 10 µM than to the single-stranded 60- mer poly (rU) RNA substrate (Figure 4.1.11. B). Next, a DNA/RNA hybrid consisting of a 60- mer poly (dA) DNA forward strand and a poly (rU) RNA reverse strand was tested by EMSA. Binding was observed starting from 2.5 µM $SLFN5^{1-336}$ concentration (Figure 4.1.11. C). Interestingly, here two distinct bands appeared indicating that several molecules of $SLFN5^{1-336}$ might bind to the DNA/RNA hybrid. However, it cannot be excluded that free poly (rU) RNA, that is labeled at 5' end with FAM, contributes to the double shift.

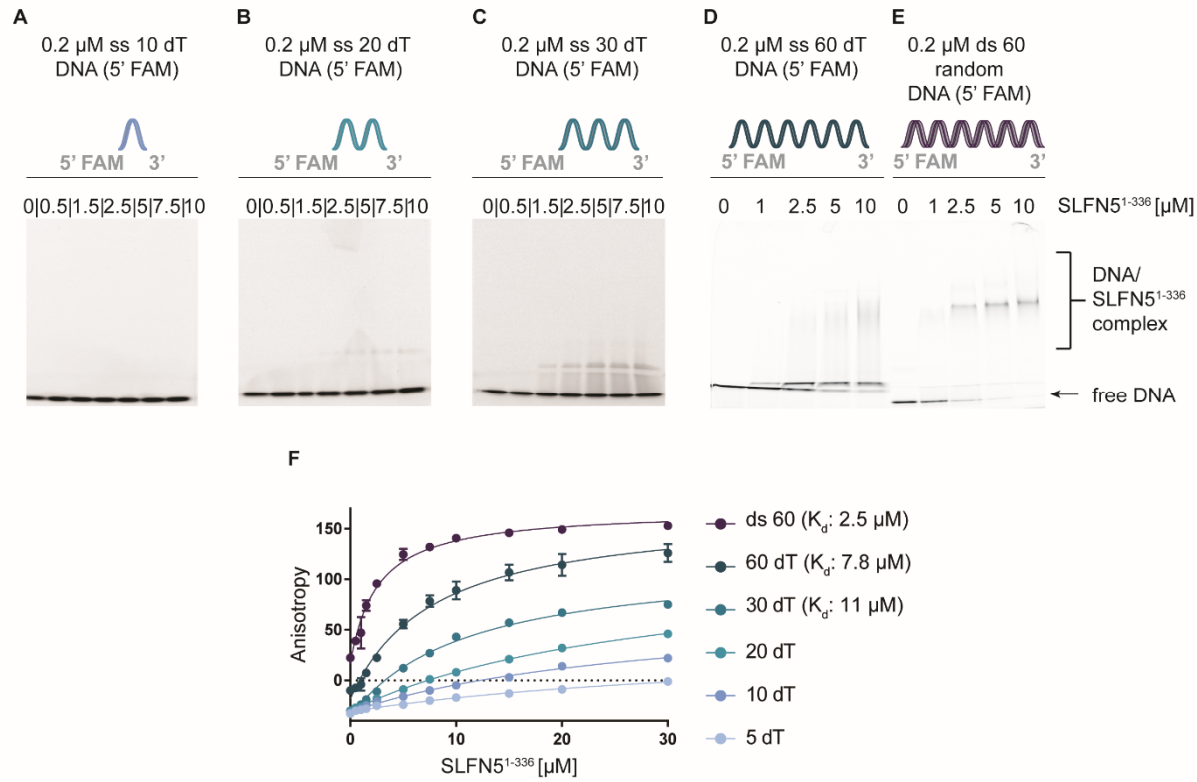



**Figure 4.1.10:** Analysis of DNA binding properties by electrophoretic mobility shift assay (EMSA) and fluorescence anisotropy (FA) of SLFN5$^{1-336}$. **A-E:** EMSA of SLFN5$^{1-336}$ with (A) single-stranded 10, (B) 20, (C) 30, (D) 60- mer poly (dT) DNA and (E) double-stranded 60- mer DNA. **F:** Fluorescence anisotropy (FA) assay to verify DNA binding. Data points represents the change in fluorescence anisotropy (arbitrary units). The various nucleic acid ligands are depicted in different colors. The solid line represents the least square fit to a Hill binding model. All nucleic acid substrates were labeled with 6-carboxyfluorescin at the 5' end, indicated in the figure as (5' FAM).

Next, SLFN5$^{1-336}$ binding to single-stranded, double-stranded and stem-loop RNAs that all share a common 41 bases forward sequence was tested. The stem loop RNA was designed to partially pair with itself thus forming base-paired as well as bulged stretches and a hairpin loop. EMSA experiments of SLFN5$^{1-336}$ to single-stranded RNA show a faint shift at 1 µM SLFN5$^{1-336}$ concentration (Figure 4.1.11. D). A distinct shift is visible at 5 µM and 10 µM concentration. Interestingly, the EMSA of double-stranded RNA with SLFN5$^{1-336}$ showed no shift (Figure 4.1.11. E). However, here as well an unspecific band appears. Band shifts with stem-loop RNA (Figure 4.1.11. F) revealed a distinct shift starting from 2.5 µM SLFN5$^{1-336}$ concentration and indicate binding. However, here as well, an unspecific band appears. Subsequent FA analysis confirmed the findings. A weak anisotropy signal was detected for the double-stranded RNA, which is however insufficient to determine a $K_d$-value. The $K_d$ of SLFN5$^{1-336}$ to the single-stranded RNA was calculated to be 10.3 µM. Thus, SLFN5$^{1-336}$ binds the single-stranded 41- mer RNA with a similar affinity than the single-stranded 30- mer poly (rU) RNA. SLFN5$^{1-336}$ shows the highest affinity for stem-loop RNA with a $K_d$ of 2.1 µM. Additionally, binding of SLFN5$^{1-336}$ to a tRNA was tested (Figure 4.1.11. G) and a $K_d$ of 4.3 µM was calculated. This indicates that SLFN5$^{1-336}$

might bind RNAs containing secondary structure elements with a higher affinity compared to similar single- and double-stranded RNAs.

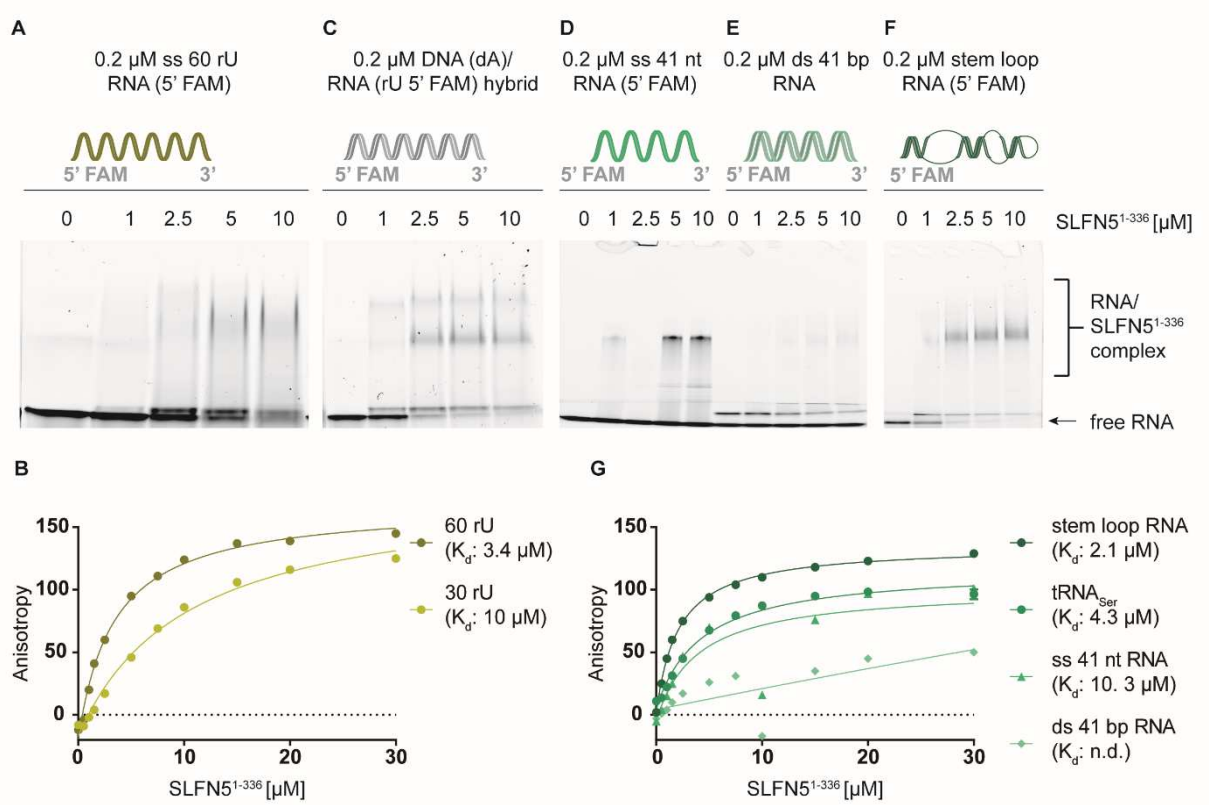



**Figure 4.1.11:** Analysis of RNA binding properties by electrophoretic mobility shift assay (EMSA) and fluorescence anisotropy (FA) of SLFN5$^{1-336}$. **A:** EMSA to visualize SLFN5$^{1-336}$ binding to a single-stranded 60- mer poly (rU) RNA. **B:** Fluorescence anisotropy (FA) assay to verify binding of SLFN5$^{1-336}$ to single-stranded 30 and 60- mer poly (rU) RNAs. Data points represents the change in fluorescence anisotropy (arbitrary units). The various nucleic acid ligands are depicted with different colors. The solid line represents the least square fit to Hill binding model. **C:** EMSA of SLFN5$^{1-336}$ to a 60- mer DNA/RNA hybrid consisting of a poly (rU) 5'FAM labeled RNA strand and a poly (dA) complementary DNA strand. **D-F:** EMSA to visualize SLFN5$^{1-336}$ binding to (D) single-stranded 41- mer RNA, (E) double-stranded 41- mer RNA and (F) stem-loop RNA. **G:** FA assay to confirm binding of SLFN5$^{1-336}$ to various RNA substrates with different secondary structure elements.
All nucleic acid substrates were labeled with 6-carboxyfluorescin at the 5' end, indicated in the figure as (5' FAM).

Taken together, SLFN5$^{1-336}$ is capable to bind both, DNA and RNA. Affinities for both, single-stranded DNA and RNA elevate with increasing nucleic acid length, even though no significant preference for DNA or RNA could been seen. Regarding binding to DNA, the affinity for double-stranded DNA unambiguously increases. Although the highest affinity of SLFN5$^{1-336}$ to RNA could be observed with a stem-loop RNA, the protein does not bind RNA with perfect double stranded regions.

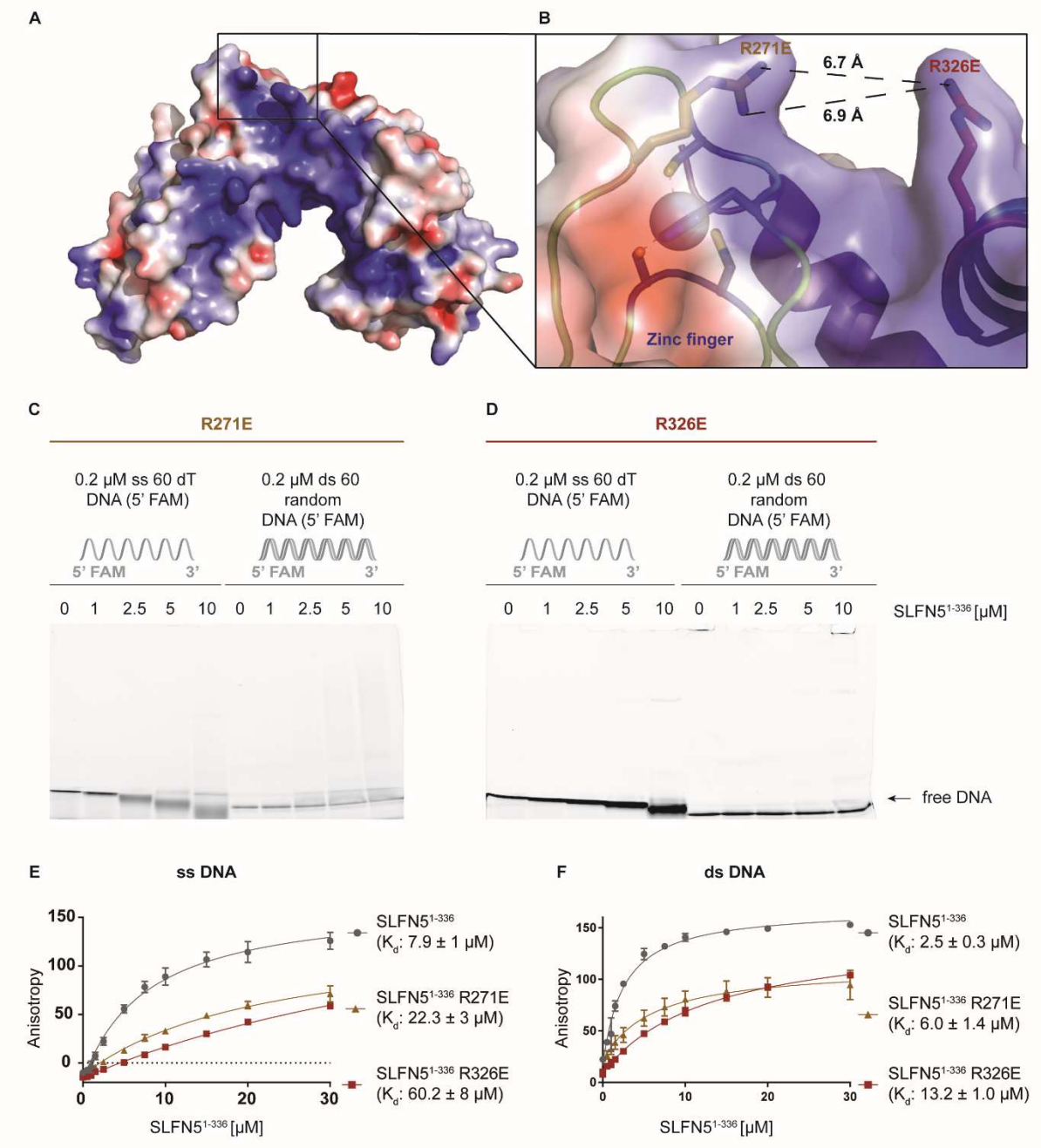



**Figure 4.1.12:** Comparison of SLFN5$^{1-336}$ wild type with DNA-binding mutants. **A:** Electrostatic surface potential of SLFN5$^{1-336}$ colored from red (-4kT/e) to blue (4kT/e). **B:** Close-up view of R271 (yellow) and R326 (red), which are located in close proximity to the zinc finger. **C-D:** EMSA to visualize binding of SLFN5$^{1-336}$R271E (C) and R326E (D) to single-stranded 60- mer poly (dT) DNA and double-stranded 60- mer DNA. **E-F:** Fluorescence anisotropy (FA) assay to monitor nucleic acid binding. Binding of SLFN5$^{1-336}$WT (grey), R271E (yellow) and R326E (red) was tested to single-stranded 60- mer poly (dT) DNA (E) and to double-stranded 60- mer DNA (F). Data points represent the change in fluorescence anisotropy (arbitrary units). Data point and error bars represent the means ± s.d. from three independent experiments.

All nucleic acid substrates were labeled with 6-carboxyfluorescin at the 5' end, indicated in the figure as (5' FAM).

In order to verify binding of SLFN5$^{1-336}$ to DNA several nucleic acid binding mutants were generated. Since no ligand bound co-crystal structure of SLFN5$^{1-336}$ is available so far, the electrostatic surface potential was calculated based on the SLFN5$^{1-336}$ crystal structure (Figure 4.1.12. A). Several positively charged patches could be observed. In one patch, two arginine residues in close proximity to the zinc finger region emerge, suggesting an essential role of these side chains in nucleic acid binding (Figure 4.1.12. B). Consequently, two single-point mutants of SLFN5$^{1-336}$ with substitutions of both arginines to glutamates (R271E and R326E) were generated. First, binding was analyzed by EMSA. Both mutants did not shift single-stranded DNA. However, the free DNA appears and seems to be degraded with increasing protein concentrations. Interestingly, for R326E (Figure 4.1.12. D) the DNA band seem to be less smeary and more distinct than for the R271E mutant (Figure 4.1.12. C). Furthermore, both mutants do not bind double-stranded DNA any more (Figure 4.1.12. C-D). These findings are in accordance with the analysis by FA. In both mutants, the $K_d$ values were unambiguously increased compared to the wild type SLFN5$^{1-336}$, confirming reduced binding. Both showed highly increased $K_d$s with 22.3 µM for R271E and 60.2 µM for R326E for single-stranded 60- mer poly (dT) DNA (Figure 4.1.12. E). The $K_d$s were increased as well for double-stranded 60- mer DNA, but with 6.0 µM for R271E and 13.2 µM for R326E not that severe (Figure 4.1.12. F), indicating weak binding. In summary, the zinc finger region is essential for DNA binding, since two mutations of positively charged residues in close proximity can negatively influence the binding.

#### 4.1.1.6. Schlafen 5$^{1-336}$ exhibits DNA-exonuclease activity

Although a crystal structure of the N-terminal slfn core domain is available, the function could not explicitly be derived from it. In the EMSAs of the single-stranded DNA binding assays frequently a smear was observed. Thus, some exonuclease activity was assumed. Indeed, exonuclease activity on a 60 bases, single-stranded poly (dT) DNA that carries a fluorescein fluorophore at the 5' terminus was observed (Figure 4.1.13. A). At 25°C, the DNA substrate was partially digested after 15 minutes and completely after 30 minutes. During the reaction, a ladder is visible on the gel, assuming successive digestion starting from the 3' end (Figure 4.1.13. A). By adding EDTA to the reaction mixture, the exonuclease activity was fully abolished, indicating a dependency on metal ions (Figure 4.1.13. A). Since the slfn core domain was predicted to function as an ATPase, the nuclease activity was also tested in presence of ATP. There was no detectable difference in activity in presence or absence of ATP (Figure 4.1.13. A). By adding various metal ions, SLFN5$^{1-336}$ seems most active in the presence of magnesium (Figure 4.1.13. B). When adding manganese or zinc to the reaction mixture, the DNA was not fully digested after 60 minutes, indicating an inhibitory function for manganese and zinc (Figure 4.1.13. B). SLFN5$^{1-336}$ was not active in the presence of calcium, implying an inhibitory role of calcium towards SLFN5$^{1-336}$ (Figure 4.1.13. B). Furthermore, SLFN5$^{1-336}$ cannot digest substrates labeled at the 3' terminus, (Figure 4.1.13. C), implying that the exonuclease activity functions in a 3'-5' directionality. Exonuclease activity seems to be restricted to single-stranded DNA. No digestion was observed on a double-stranded 60- mer DNA (Figure 4.1.13. D), although SLFN5$^{1-336}$ binding to this substrate was confirmed in section 4.1.1.5. Furthermore, neither digestion of single-stranded RNA nor double-stranded RNA by SLFN5$^{1-336}$ could be observed (Figure 4.1.13. E). Since Slfn13$^{14-353}$ was shown

to cleave $tRNA_{Ser}^{126}$, nuclease activity was tested with $tRNA_{Ser}$, as well. Interestingly, $SLFN5^{1-336}$ does not cleave $tRNA_{Ser}$ (Figure 4.1.13. F). Taken together, $SLFN5^{1-336}$ is a newly described exonuclease acting on single-stranded DNA. The activity seems to be dependent on metal ions, especially magnesium and follows a 3'-5' directionality.

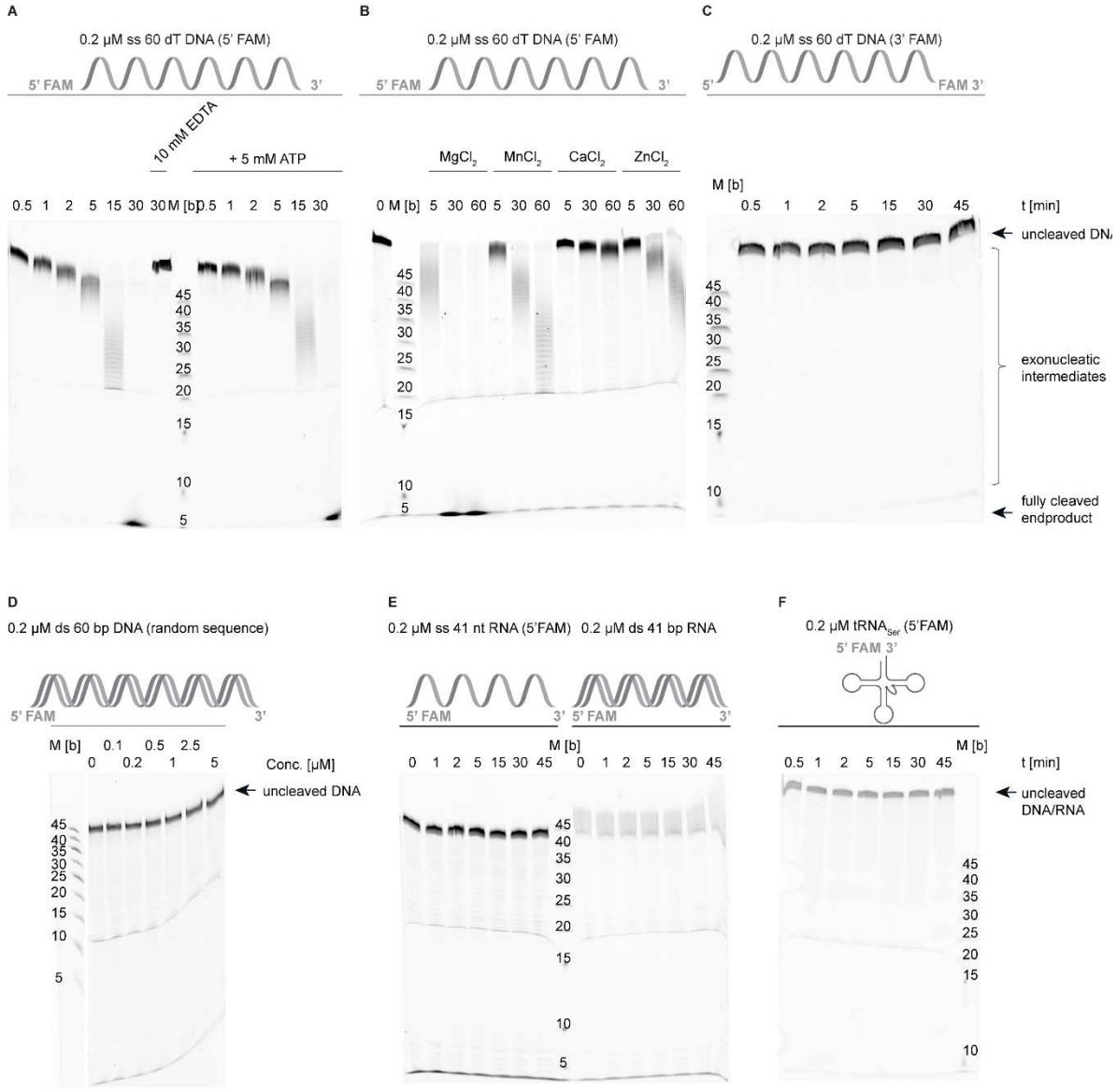



**Figure 4.1.13:** Analysis of exonuclease activity of $SLFN5^{1-336}$. Exonuclease assay of $SLFN5^{1-336}$ incubated with (A) a single-stranded 60- mer poly (dT) DNA (5' FAM) in absence and presence of ATP and EDTA, (B) a single-stranded 60- mer poly (dT) DNA (5' FAM) in presence of 2 mM $MgCl_2$, 1 mM $MnCl_2$, 2 mM $CaCl_2$ or 0.1 mM $ZnCl_2$, (C) single-stranded 60- mer poly (dT) DNA (3' FAM), (D) double-stranded 60- mer DNA, (E) single-stranded 41- mer RNA (5' FAM) and double-stranded 41- mer RNA, and (F) $tRNA_{Ser}$.

Nucleic acid substrates were labeled with 6-carboxyfluorescin either at the 5' end or at the 3' end, indicated in the figure as (5' FAM) or (3' FAM), respectively.

Different SLFN5$^{1-336}$ wild type batches purified with the same protocol (Ni-NTA, Q-Sepharose, S200; see section 4.1.1.2), showed differences in activity on single-stranded 60- mer poly (dT) DNA. For wild type batch 'Q1', digestion is completed after 45 minutes, whereas for wild type batch 'Q2' and 'Q3' the substrate is just partly digested after 45 minutes (Figure 4.1.14. A). The wild type batch 'Q+H', which was purified with an additional heparin column after the Q-Sepharose step, seems completely inactive (Figure 4.1.14. A). The SDS-PAGE of the endproducts of the differently purified wild type batches reveals that all of them have roughly the same degree of purity (Figure 4.1.14. B), making it improbable to account an impurity for the discrepancy in exonuclease activity. By adding different metal ions to the wild type batches 'Q2' and 'Q+H' the original exonuclease phenotype was tried to restore. However, the batch 'Q2' is only marginally active with magnesium (Figure 4.1.14 C), whereas the batch 'Q+H' remains inactive. Further investigations to identify potentially essential co-factors that might have been removed during SLFN5$^{1-336}$ purification, need to be performed directly.

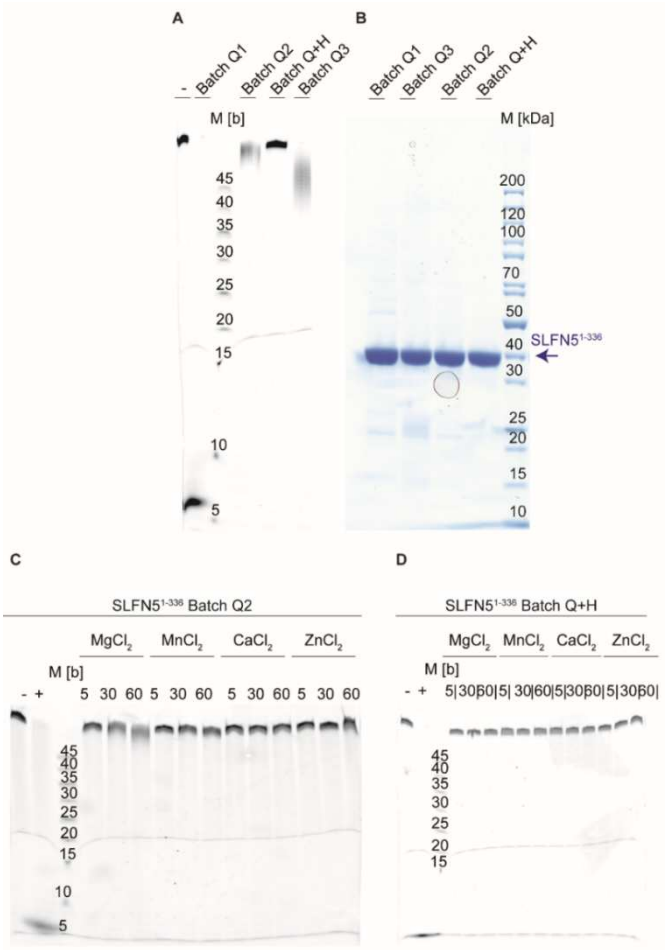


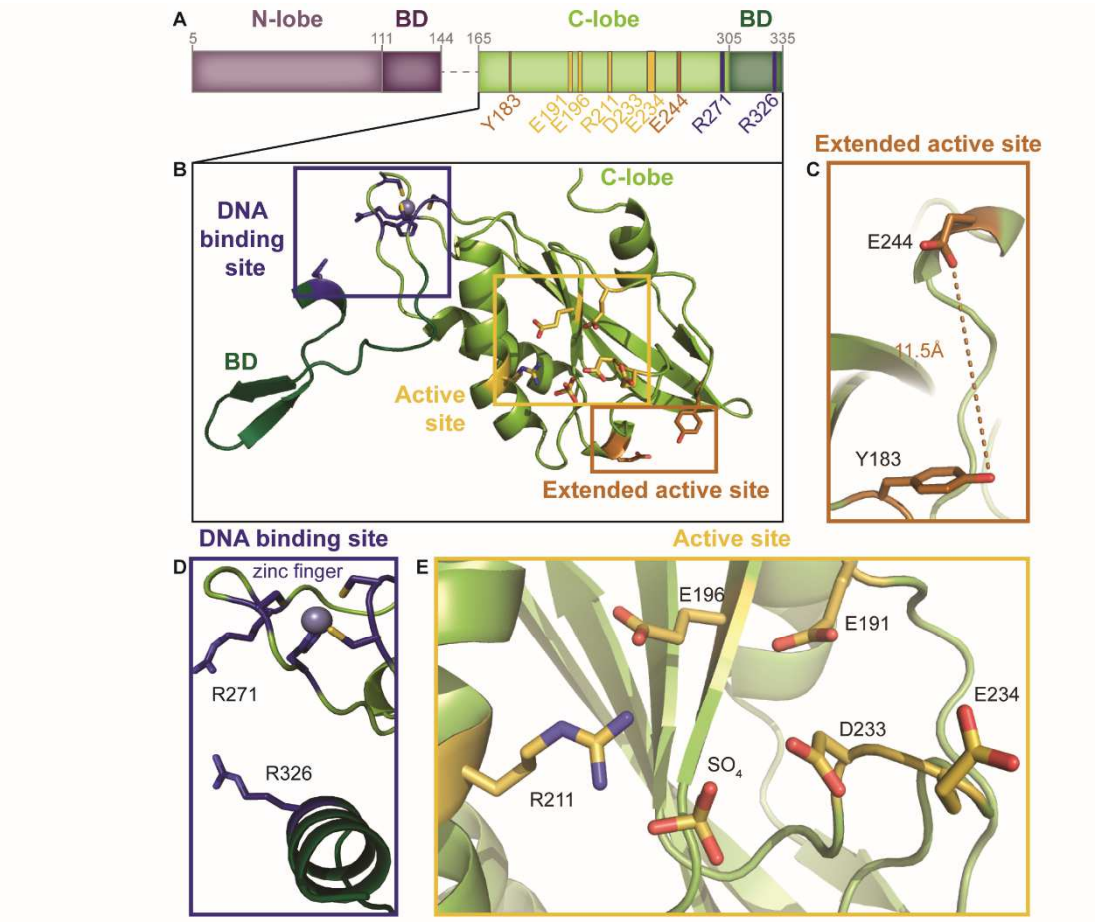

**Figure 4.1.14:** Comparison of different SLFN5$^{1-336}$ batches. **A:** Exonuclease assay of different SLFN5$^{1-336}$ batches with a single-stranded 60- mer poly (dT) DNA, which is labeled with FAM at the 5' end (5' FAM). **B:** Coomassie-stained SDS-PAGE of the corresponding SLFN5$^{1-336}$ batches. **C-D:** Analysis of exonuclease activity of SLFN5$^{1-336}$ Q2 batch (C) and SLFN5$^{1-336}$ Q+H batch (D) with a single-stranded 60- mer poly (dT) DNA (5' FAM) in presence of 2 mM $MgCl_2$, 1 mM $MnCl_2$, 2 mM $CaCl_2$ or 0.1 mM $ZnCl_2$. Q: purified via Ni-NTA, Q-Sepharose and S200; Q+H: Purified via Ni-NTA, Q-Sepharose, Heparin and S200.

To understand the exonuclease mechanism and to eliminate the possibility that the exonuclease is due to an impurity, point mutations were generated. The first site tested was the putative active site including the sulfate coordinating residues located in the C-lobe. Upon mutation of the partly conserved glutamate 191 (Figure 4.1.15. E, F) to either alanine or glutamine, the activity was in both constructs marginal or completely inactive (Figure 4.1.16. A and Figure 4.1.17. A, B), suggesting an important role for E191 in catalysis. Next, the highly conserved glutamate 196 (Figure 4.1.15. E, F) was mutated to either alanine or glutamine. Upon mutation to alanine, SLFN5$^{1-336}$E196A was inactive (Figure 4.1.16. A and Figure 4.1.17. C). However, when mutating it to glutamine it still showed activity (Figure 4.1.16. A and Figure 4.1.17. D). The partly conserved aspartate 233 (Figure 4.1.15. E, F) was mutated to either alanine or asparagine. When mutating aspartate 233 to alanine the activity was comparable good to wild type SLFN5$^{1-336}$ (Figure 4.1.16. A and Figure 4.1.17. E). However, upon mutation to asparagine the activity was completely abolished (Figure 4.1.16. A and Figure 4.1.17. F). Consequently, it was speculated that glutamate 234 could resume the role of aspartate 233 when mutated to alanine. Thus, a double mutant D233A and E234A as well as the single substitution of glutamate 234 to alanine were generated. Both constructs were exonuclease deficient (Figure 4.1.16. A and Figure 4.1.17. G-H). According to the SLFN5$^{1-336}$ crystal structure the non-conserved arginine 211 (Figure 4.1.15. E, F) seems to mainly coordinate the sulfate. Consequently, arginine 211 was mutated to alanine. Here, as well the activity was completely abolished (Figure 4.1.16. A and Figure 4.1.17. I). Taken together, essential residues of the catalytic site of the exonuclease were identified based on the SLFN5$^{1-336}$ crystal structure.

Tyrosines can play important roles in performing base-stacking interactions with the nucleic acid substrate[179]. Tyrosine 183, which is highly conserved within the schlafen protein family (Figure 4.1.15. F) and is in close proximity to the active site at the C-lobe region (Figure 4.1.15. B, C), could perform base stacking with the DNA and form a putative DNA entry site together with the conserved glutamate 244 (Figure 4.1.15. F). Those two residues are in distance of 11.5 Å from each other (Figure 4.1.15. C), which corresponds to the width of a single-stranded DNA. When mutating tyrosine 183 to phenylalanine, the activity is completely abolished (Figure 4.1.16. A and Figure 4.1.17. J), however, when mutating tyrosine 183 to alanine it still shows activity (Figure 4.1.17. K). Upon mutations of glutamate 244 to either alanine or glutamine, the activity is also completely abolished for E244A and partly for E244Q (Figure 4.1.16. E and Figure 4.1.17. L. M). Since the exact function of this site is not completely understood and it is located in close proximity to the active site, it is referred to as extended active site.

Since the DNA-binding deficient SLFN5$^{1-336}$ variants R271E and R326E did not show smear in the EMSA, they were tested for exonuclease activity, as well. Arginine 271 is partly conserved, whereas arginine 326 is not conserved (Figure 4.1.15. F). Interestingly, R271E showed exonuclease activity, (Figure 4.1.16. A and Figure 4.1.17 N), whereas R326E is not active (Figure 4.1.16. E and Figure 4.1.17 O).

The subsequent SDS-PAGE of the purified SLFN5$^{1-336}$ mutants after Ni-NTA, Q-Sepharose and S200 purification showed that all mutants have roughly the same degree of purity. No correlation of any impurity and exonuclease activity can be derived. Thus, surely SLFN5$^{1-336}$ is responsible for the exonuclease activity rather than an impurity (Figure 4.1.16. B).




F

Active site

| | | E191 | E196 | R211 | D233 E234 | |
|---|---|---|---|---|---|---|
| hSLFN5 | 191 | ESTHVEFVMFSTD-VS--HCVKDRLPKCVSAFANTEGGYVFFGVHDET | | | | 235 |
| mSlfn1 | 193 | ESKNVEYKSFETKKLL--QRVKEILPRTVSAFANTDGGYLFIGLDEKK | | | | 238 |
| mSlfn2 | 230 | RSRYVEVTLLSAKRLR--KRIKELLPQTVSAFANTDGGFLFIGLDGKT | | | | 275 |
| mSlfn3 | 164 | KSIHAEVKLTPKEKIFPKEKILELLPQTVSAFANTDGGFLFIGLDGKT | | | | 211 |
| mSlfn4 | 243 | KSTHAEVKLTPKEKISPKEKILELLPQTVSAFANTDGGYLFIGLDGKT | | | | 290 |
| mSlfn5 | 187 | KSSHVEFQMFSAD-LS--QGIRERLPKCVSALANSEGGYVFFGVHDET | | | | 231 |
| mSlfn8 | 205 | ESTSIEFKQFSTKHVQ--AYMKNIIPEYISAFANTQGGYLFIGVDDKR | | | | 250 |
| mSlfn9 | 205 | ESTSIEFKQFDTENAQ--KYMKDIIPEYISAFANTQGGYLFIGVDDKS | | | | 250 |
| mSlfn10 | 203 | ESPSIEFKQFSTKHVQ--KYIKNIIPEYISAFANTQGGYLFIGVNDKS | | | | 248 |
| mSlfn14 | 195 | ESTHVEFKRFTTKKIV--PRIKETLAHYVSAFANTQGGYII IGVDDKS | | | | 240 |
| hSLFN11 | 209 | ESQLVEFKQFSTKHFQ--EYVKRTIPEYVPAFANTGGGYLFIGVDDKS | | | | 254 |
| hSLFN12 | 200 | ESTHVEIKNFSTEKLL--QRIKEILPQYVSAFANTDGGYLFIGLNED- | | | | 244 |
| hSLFN13 | 208 | ESPSIEFKQFSTKHIQ--QYVENIIPEYISAFANTEGGYLFIGVDDKS | | | | 253 |
| hSLFN14 | 206 | ESTHVEFKRFTTKKVI--PRIKEMLPHYVSAFANTQGGYVLIGVDDKS | | | | 251 |

| | | Extended active site | | | | | DNA binding site | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Y183 | | | E244 | | zinc finger | R271 | | R326 | | |
| hSLFN5 | 181 | LQYLE | 185 | 242 | EKEK | 245 | 265 HHFCTQ--R | 271 | 325 | TREW | 328 | |
| mSlfn1 | 184 | LTKKE | 189 | 245 | EAKN | 248 | 268 THFCEE--R | 274 | 328 | IEEW | 331 | |
| mSlfn2 | 220 | FQYEE | 224 | 282 | EAEK | 285 | 305 THFCEE--K | 311 | 365 | TEEW | 368 | |
| mSlfn3 | 154 | LAYEE | 158 | 218 | EAEK | 221 | 241 THFCGE--K | 247 | 301 | TEEW | 304 | |
| mSlfn4 | 233 | LAYEE | 237 | 297 | EAEK | 300 | 320 THFCEE--K | 326 | 380 | TEEW | 383 | |
| mSlfn5 | 177 | LQYLE | 181 | 238 | EKEK | 241 | 261 YHFCGQ--N | 267 | 321 | ANDW | 324 | |
| mSlfn8 | 195 | LEYGQ | 199 | 256 | PKDN | 259 | 279 FHFCSSKDK | 287 | 342 | TEEW | 345 | |
| mSlfn9 | 195 | LEYGQ | 199 | 256 | PKDN | 259 | 279 FHFCSSKDK | 287 | 342 | TEEW | 345 | |
| mSlfn10 | 193 | LEYGQ | 197 | 254 | PKDN | 257 | 277 FHFCSSKDK | 286 | 340 | TEEW | 343 | |
| mSlfn14 | 185 | LNFKE | 189 | 247 | KKEK | 250 | 270 YHFCRE--K | 276 | 330 | VQQW | 333 | |
| hSLFN11 | 199 | LEYGE | 202 | 261 | AKEN | 264 | 284 VHFCQP--Q | 290 | 343 | TEKW | 346 | |
| hSLFN12 | 190 | LDRKE | 194 | 251 | KAEM | 254 | 274 HHFCME--K | 280 | 334 | RKEW | 337 | |
| hSLFN13 | 198 | IEYGE | 202 | 260 | AKEQ | 263 | 283 VHFCSS--K | 289 | 342 | TEEW | 345 | |
| hSLFN14 | 196 | LMYKE | 200 | 258 | KWEK | 261 | 281 FHFCCE--K | 287 | 341 | AEQW | 344 | |

**Figure 4.1.15:** Structural framework of the exonuclease activity. **A:** Schematic overview of the residues important for exonuclease activity. **B:** Structural overview of the three catalytically important sites of SLFN5$^{1-336}$ located in the C-lobe. Essential side-chains are presented as sticks and colored respectively; Yellow: active site, Brown: extended active site; Blue: DNA-binding region. **C:** Close-up view of the extended active site harboring Y183 and E244. **D:** Close-up view of the DNA-binding region harboring R271 and R326. **E:** Close-up view of the active site of SLFN5$^{1-336}$. **F:** Multiple sequence alignment of all murine and human schlafen members of the three respective sites.

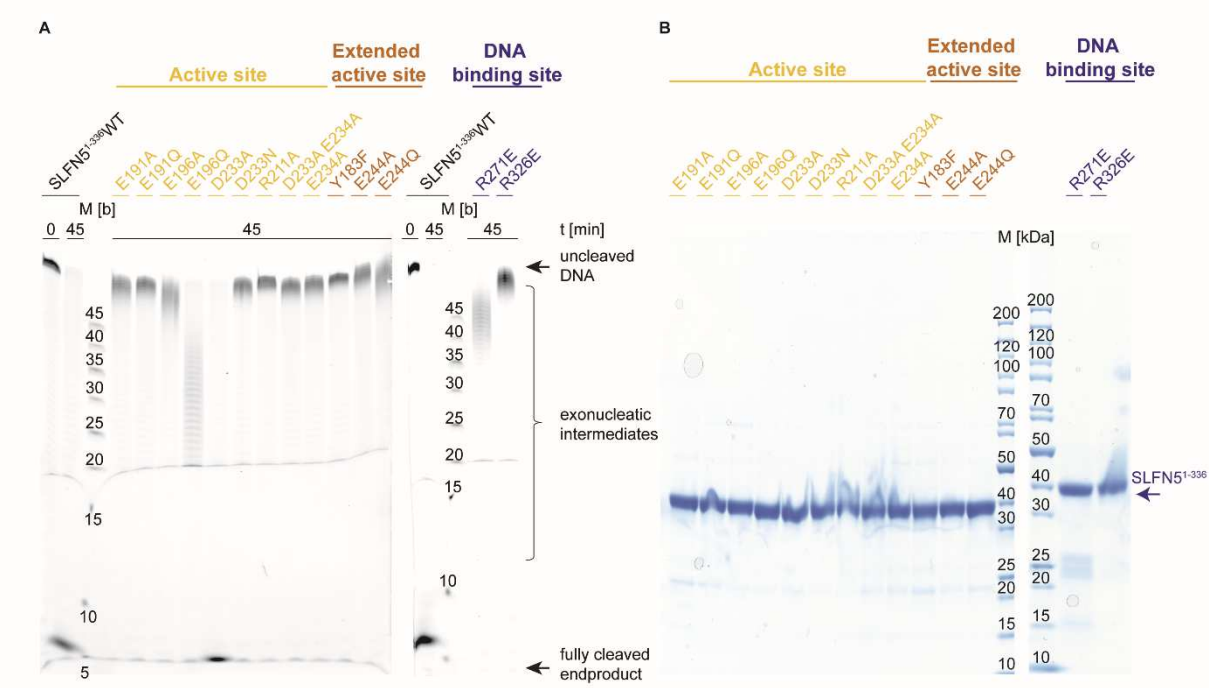



**Figure 4.1.16:** Analysis of SLFN5$^{1-336}$ exonuclease deficient mutants. **A:** Exonuclease activity assay of all SLFN5$^{1-336}$ mutants generated with single-stranded 60- mer poly (dT) DNA, which is labeled with FAM at the 5' end (5' FAM). **B:** Coomassie-stained SDS-PAGE of the endproducts of the purification of the SLFN5$^{1-336}$ mutants.

Taken together, three sites, which are important for the exonuclease activity have been identified (Figure 4.1.15. B). The active site, where the sulfate is coordinated, seems very critical for activity. All residues mutated within this region, except E196Q and D233A, were exonuclease deficient (Figure 4.1.15. E). Next, the conserved tyrosine 183 and glutamate 244 located in the extended active site, seem to play essential roles in catalysis, as well (Figure 4.1.15. B, C). Last, the region already shown to be important for DNA-binding located at close proximity of the zinc finger, contributes to the exonuclease activity (Figure 4.1.15. D).

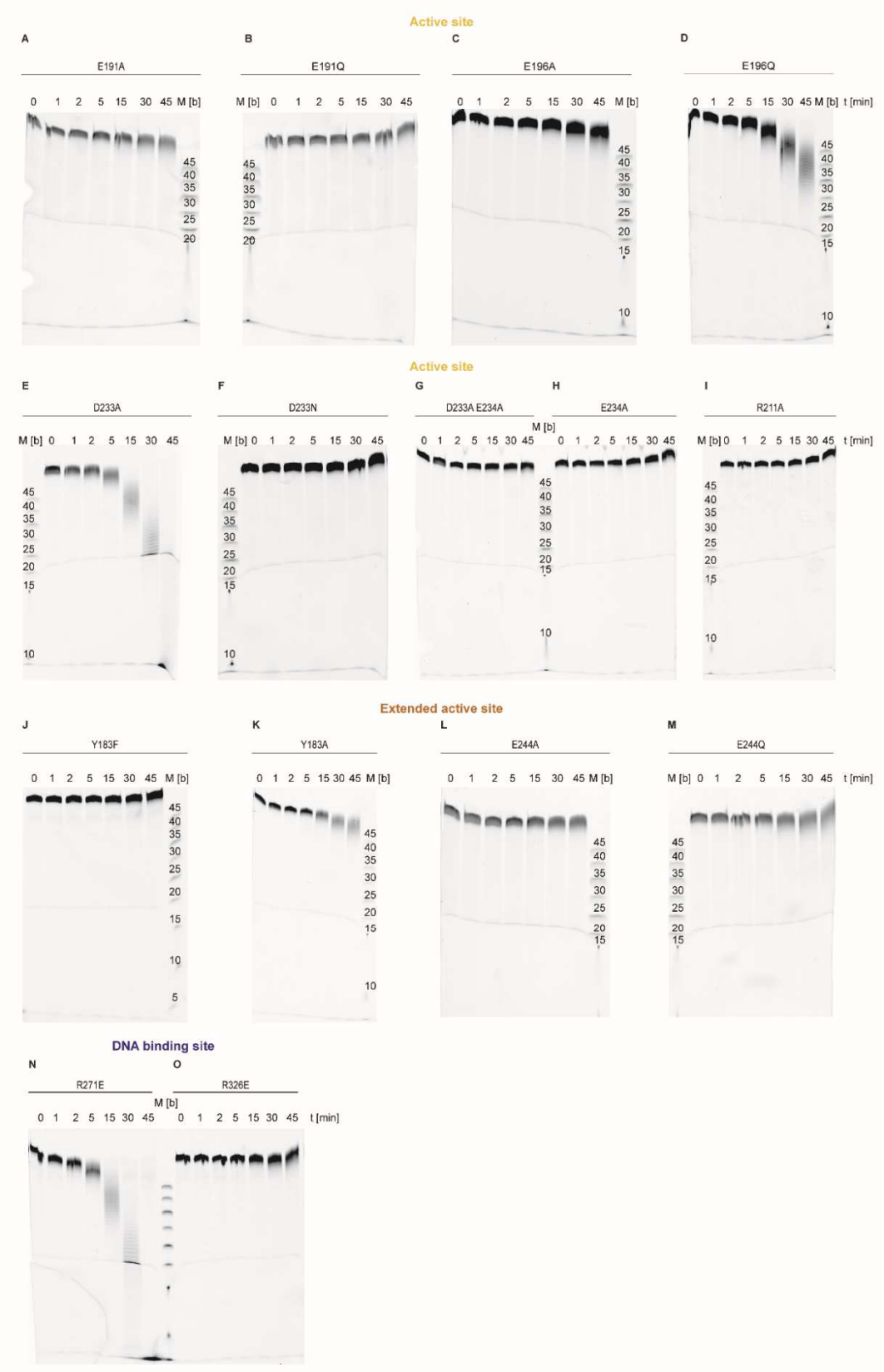



**Figure 4.1.17:** Exonuclease activity assay of all SLFN5$^{1-336}$ mutants in a time course (0-45 minutes) experiment with single-stranded 60- mer poly (dT) DNA, which is labeled with FAM at the 5' end (5' FAM).

### 4.1.2. The slfn core domain of murine Schlafen 2

Murine Schlafen 2 (Slfn2), which belongs to schlafen subgroup I, exclusively consists of the slfn core domain. A full length construct harboring all 378 amino acids of Slfn2 was generated and could solubly be expressed in *E. coli* Rosetta. Generally, all Slfn2 constructs express fine in bacterial expression systems and no problems in terms of solubility were observed. In addition, a Slfn2 version lacking a potentially exposed hydrophobic loop was generated, which improved crystallization of Slfn2.

**Table 4.3:** Overview of the Slfn2 constructs used for expression and purification. The molecular weight is given in kDa and the protein length is indicated in numbers of amino acids (aa). Furthermore, the isoelectric points (pI), the kind of affinity tag, the expression system used and the solubility of Slfn2 are indicated.

| Construct | Description | Length [aa] | MW [kDa] | Theore-tical pI | Affinity tag | Expression system | Solubility and achievements |
|---|---|---|---|---|---|---|---|
| **pET21-Slfn2$^{FL}$** | Full length; Slfn core domain; | 386 | 43.5 | 6.2 | C-terminal $His_6$ | *E. coli* (Rosetta) | Soluble (DNA-exonuclease) |
| **pET21-Slfn2$^{\Delta C116\text{-}W122}$** | Deleted hydrophobic loop | 379 | 42.6 | 6.2 | C-terminal $His_6$ | *E. coli* (Rosetta) | Soluble (Crystal structure) |

#### 4.1.2.1. Purification and characterization of Schlafen 2

Full length Slfn2$^{FL}$ was cloned into a pET21 vector and the encoded gene product was carrying a plasmid encoded C-terminal hexahistidinetag (Figure 4.1.18. A). The initial, non optimized purification protocol included an affinity chromatography step followed by a size exclusion chromatography (Figure 4.1.18. B-C). The C-terminal hexahistidine tag was used to capture Slfn2$^{FL}$ from the cell lysate on Ni-NTA beads. In the subsequent size exclusion chromatography four peaks appeared (Figure 4.1.18. B). A SDS-PAGE of the fractions collected during SEC revealed that peak I to III contained protein impurities running between 60 and 70 kDa (Figure 4.1.18. C). Slfn2$^{FL}$ could be detected in peak IV eluting at approximately 210 ml from a S200 26/60 column (Figure 4.1.18. C). Slfn2$^{FL}$ has a theoretical molecular weight of 43.5 kDa, thus, on the SDS-PAGE it runs slightly above the 40 kDa marker protein band (Figure 4.1.18. C). Interestingly, Slfn2$^{FL}$ runs as a double band on the SDS-PAGE. Moreover, a significant part seems to stay in the well and does not migrate into the gel (Figure 4.1.18. C). After pooling the fractions corresponding to peak IV, an $A_{260}/A_{280}$ ratio of 0.7 was measured. This indicates that Slfn2$^{FL}$ is contaminated with nucleic acid. Finally, 8 mg per liter *E. coli* expression culture were obtained.

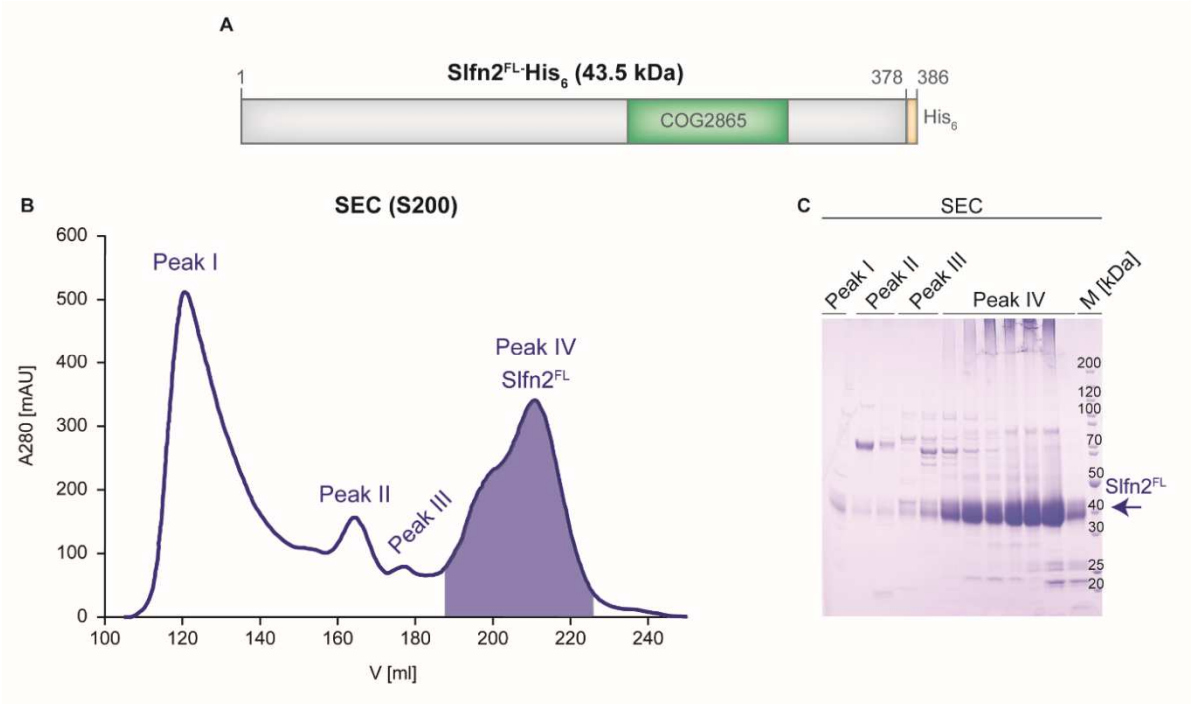



**Figure 4.1.18:** Purification of Slfn2$^{FL}$ from *E. coli* Rosetta**. A:** Schematic overview of the full length Slfn2$^{FL}$ construct carrying a plasmid encoded C-terminal hexahistidine tag. **B-C:** Chromatogram (B) and subsequent Coomassie-stained SDS-PAGE (C) of fractions collected after size exclusion chromatography (SEC) using a S200 26/60 column. Peak I to III harbor various impurities, whereas Slfn2$^{FL}$ eluted in peak IV at approximately 210 ml. The blue line indicates the absorption at λ=280nm.

Next, the Slfn2$^{FL}$ endproduct was analyzed by an analytical size exclusion chromatography using a Superose 6 10/300 column (Figure 4.1.19. A). The first appeared peak eluted at approximately 17 ml and showed a high signal for A260. The second peak eluted at 18.5 ml and the majority of the sample absorbed at 280 nm. This indicates that the Slfn2$^{FL}$ sample contains at least two protein species, and that one co-purifies with a nucleic acid. Consequently, the peak containing the nucleic acid was further analyzed on an agarose gel (Figure 4.1.19. B). When applying it, a faint band at approximately 100 bp appeared, however the sample looked smeary (Figure 4.1.19. B). Thus, the sample was treated with proteinase K to eliminate all proteins interacting with the nucleic acid. Subsequent analysis on the agarose gel showed a distinct band at 100 bp (Figure 4.1.19. B). Next, Slfn2$^{FL}$ peak I fractions were treated with DNase I or RNaseI. Upon treatment with DNase I the 100 bp band on the agarose gel remained (Figure 4.1.19. B). However, upon treatment with RNase I the band vanished (Figure 4.1.19. B). Subsequently, the sample was applied to an Agilent Bioanalyzer to determine the exact size of the co-purified RNA. Analysis revealed that the sample consists of one single peak with the size of 113 nucleotides (Figure 4.1.19. C).

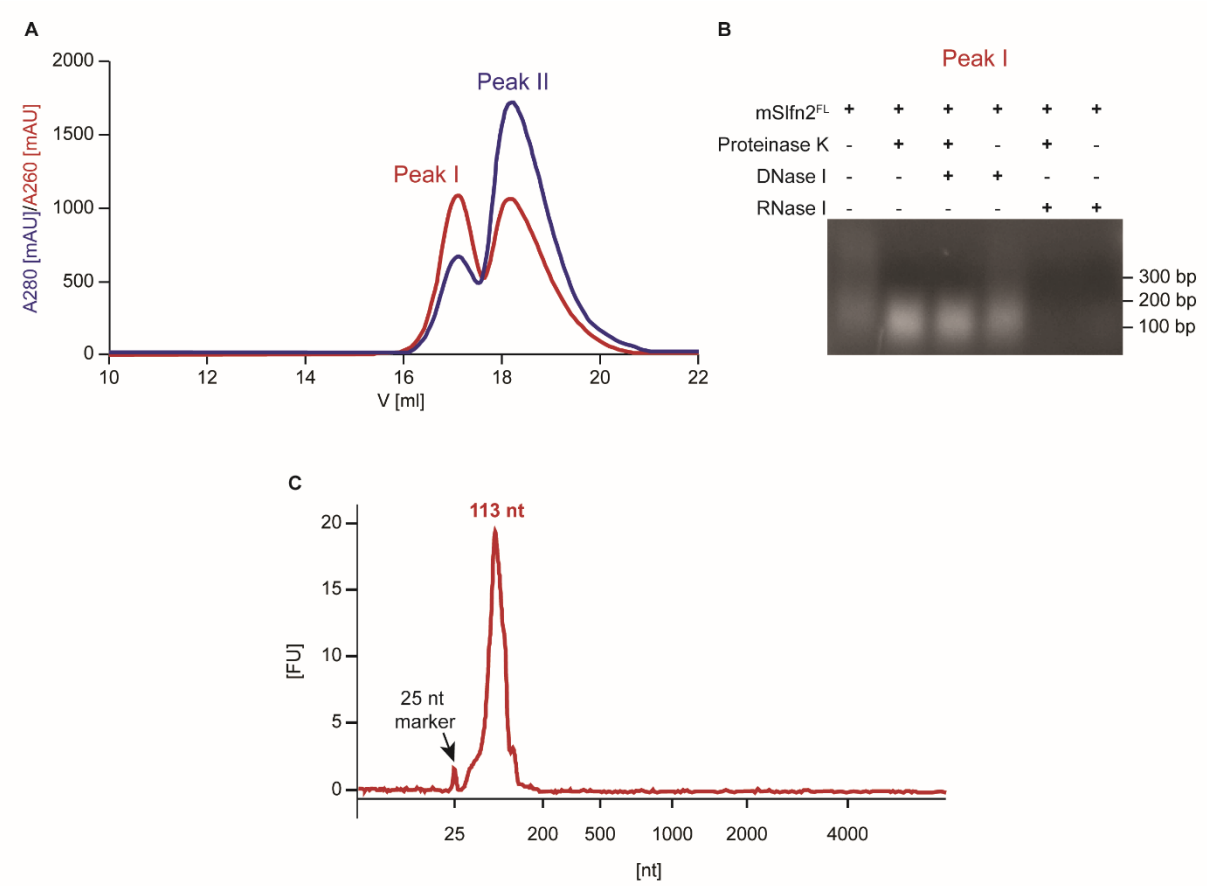



**Figure 4.1.19:** Analysis of the nucleic acid that co-purifies with Slfn2$^{FL}$. **A:** Analysis of Slfn2$^{FL}$ on a Superpose 6 10/300 column. Two peaks eluting at approximately 17 ml and 18.5 ml with high and low $A_{260}/A_{280}$ ratio, respectively, appear. **B:** Analysis of the high $A_{260}/A_{280}$ ratio peak on a 1% agarose gel after treatment with proteinase K, DNase I or RNase I. **C:** Analysis of Slfn2$^{FL}$-peak I on an Agilent Bioanalyzer revealing one single peak, which corresponds to 113 nucleotides in size. FU: Fluorescence units; nt: nucleotides.

Subsequently, the co-purified RNA was analyzed by Illumina RNA deep sequencing and the results were mapped against the genome of *Escherichia coli* K-12 (substrain: MG1655). The most abundant sequences were located in the 5 S rRNA operons: rrnH, rrnG, rrnD+rrnF, rrnC, rrnB and rrnE. Interestingly, no sequencing reads for operon rrnA were detected (Figure 4.1.20. A). 5S rRNAs are part of the big ribosomal subunit and are 120 nucleotides long. This length would also fit the size measured with the bioanalyzer. Although the 5S rRNA operons had in average 4000 mapped sequencing reads, some other genes enriched as well with 200-500 reads (Figure 4.1.20. B). The 4.5S rRNA, a part of the signal recognition particle, as well as many tRNAs were detected. Interestingly, the most abundant tRNA was $tRNA_{Ser}$, which was recently found to be the preferred substrate for N-terminal rSlfn13[126]. Furthermore, two copies of 16S rRNAs and one for 23S rRNA were identified. Taken together, Slfn2$^{FL}$ might bind 5S rRNA as deep sequencing of the co-purified RNA detected six out of in total seven 5S rRNA operons. Besides 5S rRNA also minor hits for tRNAs, 4.5S rRNA, 16S rRNA and 23S rRNA had been identified, indicating that Slfn2 prefers binding to RNAs with stem loop structures.

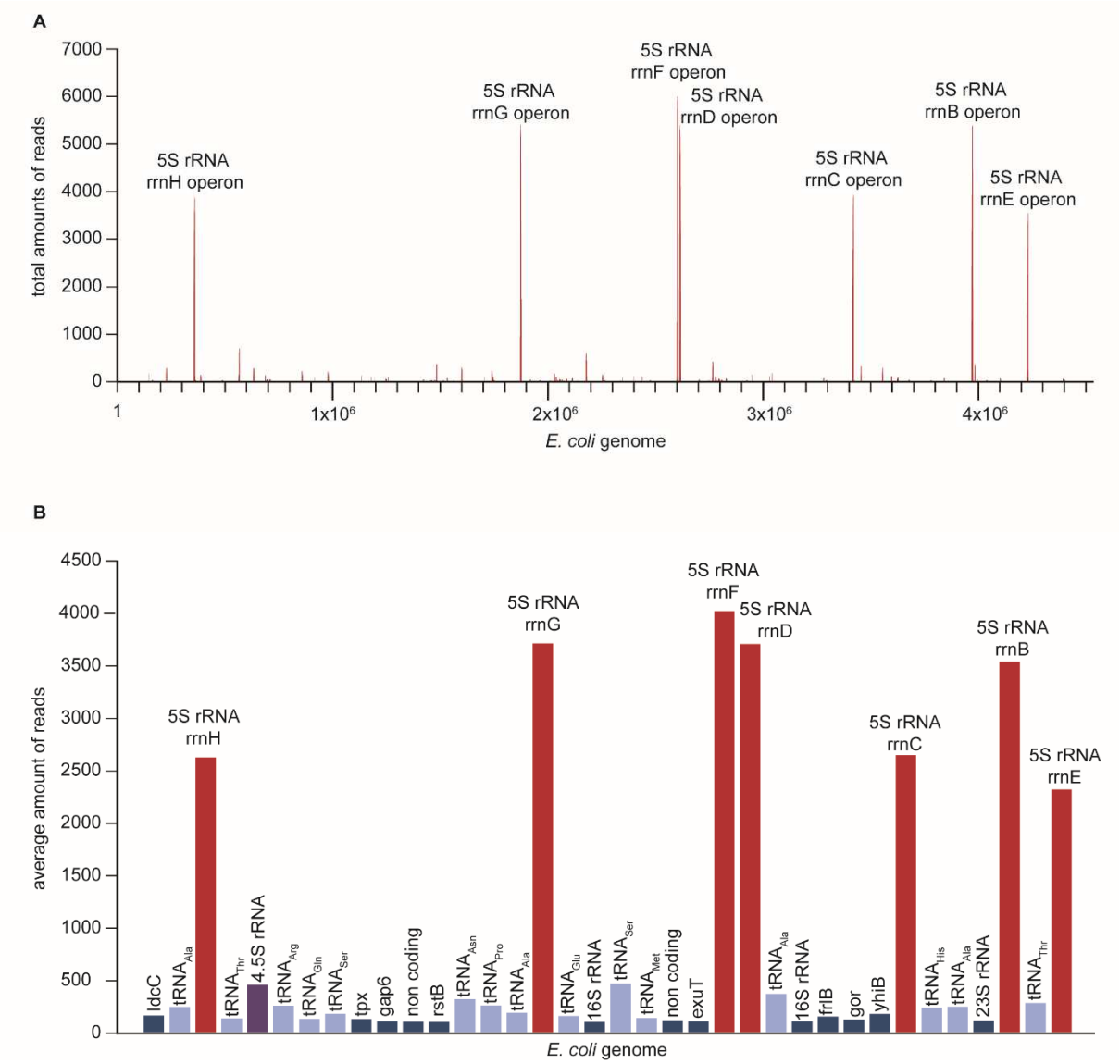



**Figure 4.1.20:** Analysis of Illumina RNA deep sequencing of RNA that co-purifies with Slfn2$^{FL}$. **A:** Mapping of the sequencing reads to the genome of *Escherichia coli* K-12 (substrain: MG1655). The most abundant reads are located in 5S rRNA operons. **B:** Average amount of reads per gene and subtraction of the background revealed besides the 5S rRNA operons many reads for tRNAs and some other genes.

For further characterization of Slfn2$^{FL}$ small-angle X-ray scattering analysis (SAXS) was performed. The primary scattering curve of Slfn2$^{FL}$ at concentration of 1.57 g/l dilution shows decent folded protein in solution (Figure 4.1.21. A). The calculated pair distribution function indicates a rather extended shape (Figure 4.1.21. B). The Guinier plot analysis reveals that the sample had some aggregates (Figure 4.1.21. C). Thus, these data points were omitted for the determination of the linear regression to calculate the radius of gyration ($R_G$), which could be determined for Slfn2$^{FL}$ to be 3.37 nm. The subsequent estimated molecular weight is 70 kDa, indicating that Slfn2$^{FL}$ would be a dimeric protein in solution. However, since the $A_{260}/A_{280}$ ratio of 0.7 indicate nucleic acid contaminations and a co-

purified RNA in the Slfn2$^{FL}$ sample was identified, this discrepancy in size could also derive from interacting RNA. Furthermore, the aggregation of Slfn2$^{FL}$ hinders the precise calculation of the molecular weight. A low-resolution, *ab initio* model could be generated. Indeed, the model shows a partly globular shape with a major elongated extension (Figure 4.1.21. D).

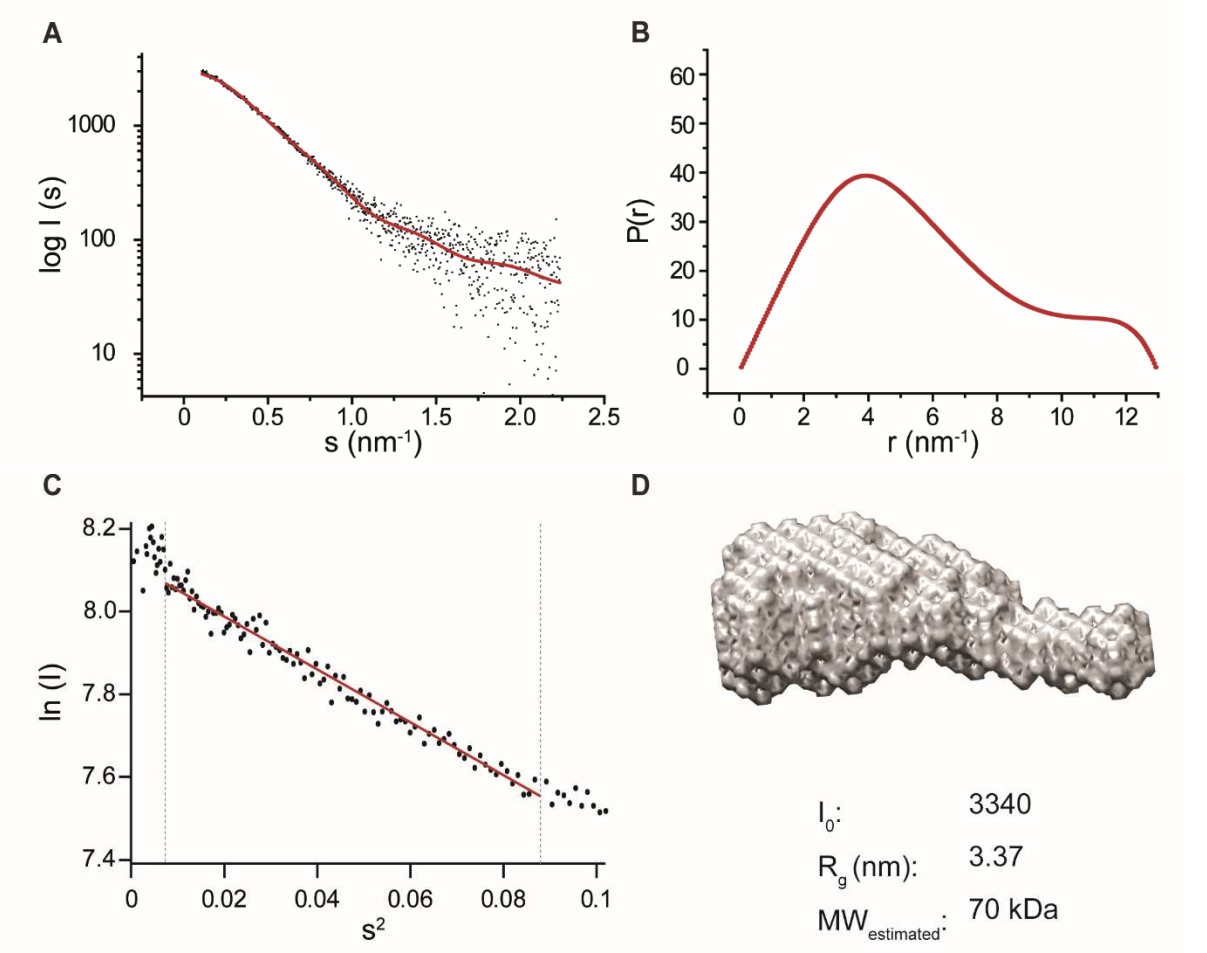



**Figure 4.1.21:** Analysis of Slfn2$^{FL}$ SAXS data. **A:** Buffer corrected SAXS curve of a 1.57 g/L Slfn2$^{FL}$ sample. **B:** A pair distribution function (P(r)) suggests a single domain protein. **C:** Guinier plot analysis determined an $R_G$ of 3.37 nm from the slope of the linear regression. The sample tends to aggregate, thus the estimated molecular weight of 70 kDa is only an approximation. **G:** Final averaged *ab initio* model of Slfn2$^{FL}$.

Taken together, full length Slfn2$^{FL}$ could be successfully expressed and purified. An initial and non-optimized purification strategy co-purified RNAs with Slfn2$^{FL}$, which is mostly likely 5S rRNA. Furthermore, SAXS data suggests that Slfn2 is either dimeric in solution or that it interacts with a ligand, which is similar in size and could be 5S rRNA.

4.1.2.2. Optimized preparation and structural characterization of Schlafen 2

For further biochemical characterization and crystallization, an optimized Slfn2$^{FL}$ protocol was established including an additional ion-exchange chromatography (Q-Sepharose) purification step.

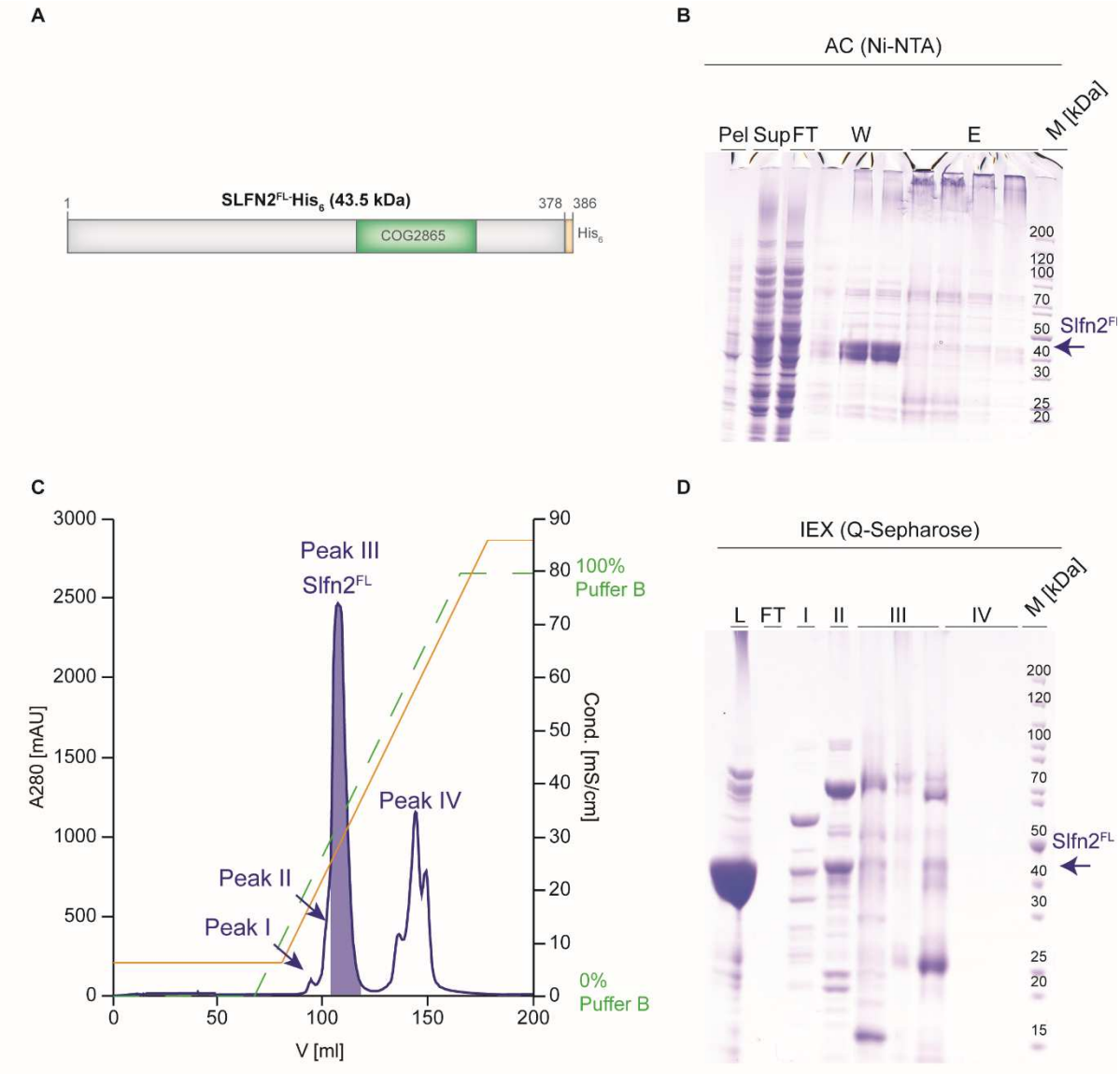



**Figure 4.1.22:** Optimized purification of Slfn2$^{FL}$ from *E. coli* Rosetta. **A:** Schematic overview of the full length Slfn2$^{FL}$ construct carrying a plasmid encoded C-terminal hexahistidine tag. **B:** Coomassie-stained SDS-PAGE of the fractions collected during affinity chromatography (AC) purification of Slfn2$^{FL}$. The wash and the elution fractions contained the Slfn2$^{FL}$ (indicated by an arrow). However, in the elution fractions the Slfn2$^{FL}$ protein stayed in the wells and did not migrate into the gel. **C-D:** Chromatogram (C) and subsequent Coomassie-stained SDS-PAGE (D) of the ion exchange chromatography (IEX) purification step using Q-Sepharose. Slfn2$^{FL}$ bound to the column, but eluted earlier than the nucleic acid contaminations. The blue line indicates the absorption at λ=280nm. The brown and green lines indicate the conductivity and the concentration of IEX B buffer, respectively. Sup: supernatant, Pel: pellet, FT: flow-through, W: wash, E: elution, L: load, M: marker in kDa.

First, Slfn2$^{FL}$ was captured after cell lysis using the C-terminal encoded hexahistidine tag. A significant amount of Slfn2$^{FL}$ already eluted in the wash fractions (Figure 4.1.22. B). Interestingly, in the elution fractions, no Slfn2$^{FL}$ protein can be detected on the SDS-PAGE. However, a smear shortly below the wells indicates that Slfn2$^{FL}$ did not migrate into the gel. Indeed, the determination of the protein concentration indicated 1-2 g/l protein concentration per 10 ml elution fraction. Consequently, the elution fractions were pooled, dialysed and applied onto an ion-exchange column (Q-Sepharose). Slfn2$^{FL}$ bound to the Q-Sepharose and was eluted by applying a high salt gradient (Figure 4.1.22. C). Analysis via SDS-PAGE again indicated that Slfn2$^{FL}$ sticks to the wells and does not migrate into the gel (Figure 4.1.22. D). Other protein impurities (peak I and II of the IEX) and nucleic acid contaminations (peak IV), however, could be separated from Slfn2 (Figure 4.1.22. C-D).

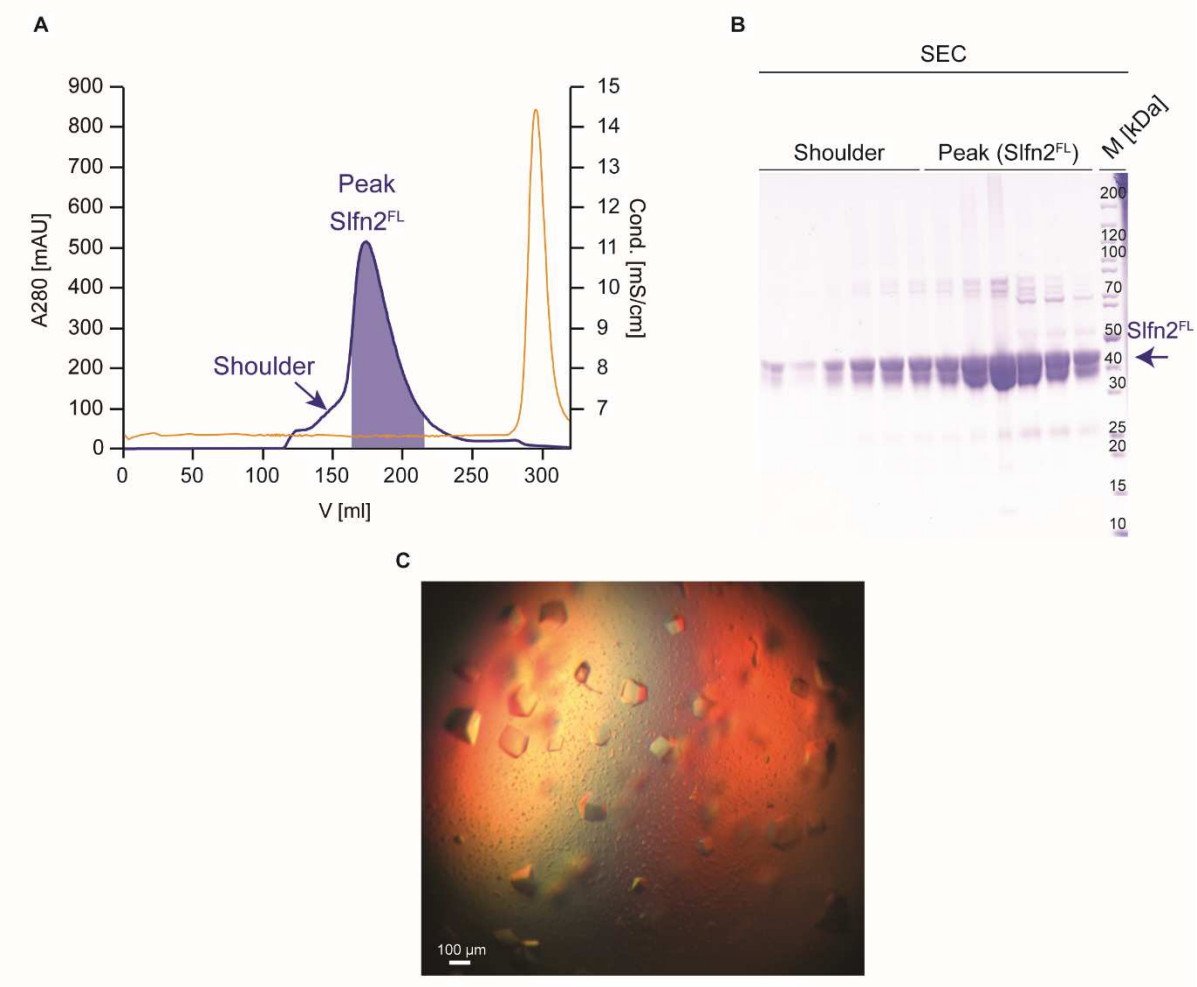



**Figure 4.1.23:** Optimized purification of Slfn2$^{FL}$ from *E.coli* Rosetta. **A-B:** Chromatogram (A) and subsequent Coomassie-stained SDS-PAGE (B) of the fractions collected during size exclusion chromatography (SEC) using a S200 26/60 column. Although, Slfn2$^{FL}$ eluted as one single peak at approximately 180 ml, the peak has a shoulder eluting before the Slfn2$^{FL}$ peak. The blue and brown lines indicate the absorption at λ=280nm and the conductivity, respectively. **C:** Crystallization of Slfn2$^{FL}$ in 0.1 M Bis-Tris pH 6.5, 0.2 M ammonium sulfate and 20% PEG3350.

To separate remaining impurities and aggregates from Slfn2$^{FL}$ a size exclusion chromatography (S200 26/60) was carried out (Figure 4.1.23. A). Slfn2$^{FL}$ elutes at approximately 180 ml in a single peak that has however a shoulder at approximately 150 ml. SDS-PAGE analysis shows that the fractions belonging to the shoulder contain Slfn2$^{FL}$ (Figure 4.1.23. B), indicating higher oligomeric forms or still contaminations with nucleic acids. The peak fractions contain mainly Slfn2$^{FL}$ and minor impurities running at 70 kDa. Again, Slfn2$^{FL}$ runs as a double-band on the SDS-PAGE. Finally, 1.5 mg pure protein per liter *E. coli* expression culture could be obtained. The $A_{260}/A_{280}$ ratio of the Slfn2$^{FL}$ was 0.57, indicating no contamination with nucleic acids.

To solve the crystal structure, crystallization trials with full length Slfn2, which was purified according to the optimized purification strategy, was performed. Wedge-shaped crystals that were about 50 µm in size and did not polarize light, were obtained (Figure 4.1.23. C). The crystals diffracted to 8 Å.

In order to improve the crystal quality of Slfn2, a structural model of Slfn2$^{FL}$ based on SLFN5$^{1-336}$ crystal structure was calculated. The Slfn2 model revealed that a hydrophobic loop insertion between amino acid 116 and 122 (CLPFLCQW) emerges from the N-lobe (Figure 4.1.24). Since the hydrophobic insertion could potentially disturb the crystal contacts in the Slfn2$^{FL}$ crystal, a deletion mutant lacking C116-W122 (Slfn2$^{\Delta C116\text{-}W122}$) was generated, expressed, purified and crystallized.

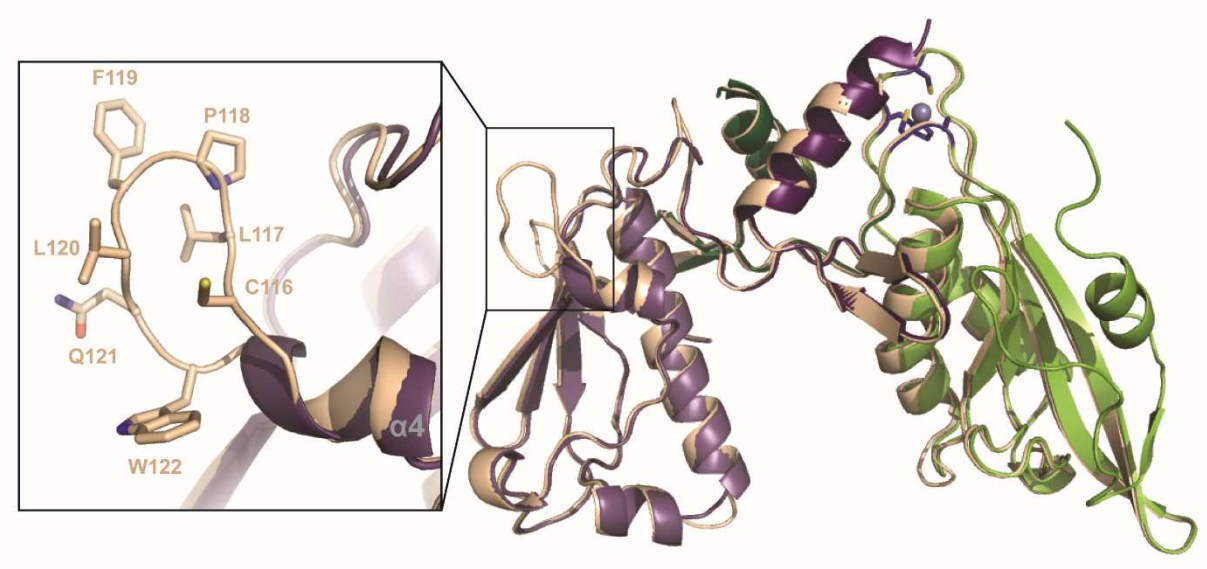



**Figure 4.1.24:** Towards the Slfn2$^{\Delta C116\text{-}W122}$ crystal structure. **A:** Superposition of a Slfn2 structural model and the SLFN5$^{1-336}$ crystal structure. Slfn2 contains an additional hydrophobic insertion (C116- W122). SLFN5$^{1-336}$ is colored as in 4.1.7 and Slfn2 is shown in wheat. The hydrophobic insertion is depicted as sticks.

Slfn2$^{\Delta C116\text{-}W122}$ was purified according to the initial purification protocol and the integrity of the protein was analyzed on a Superose 6 (10/300) column (Figure 4.1.25. A). Slfn2$^{\Delta C116\text{-}W122}$ elutes as a single peak from the Superpose 6 column at 18.5 ml, which corresponds to peak II (low $A_{260}/A_{280}$ ratio peak) of Slfn2$^{FL}$. A small shoulder eluting at 17 ml corresponds to peak I. This indicates not only that the

deletion mutant co-purifies less RNA than Slfn2$^{FL}$ wild type, but also that the hydrophobic insertion might be important for RNA binding.

Next, Slfn2$^{\Delta C116\text{-}W122}$ was used for crystallization. Large needle-shaped crystals with up to 2 mm size were obtained (Figure 4.1.25. B). They crystallized in the trigonal space group $P3_221$ (space group number 154) and diffracted to 4 Å (Figure 4.1.25. C). The crystal structure could be solved by molecular replacement with the SLFN5$^{1\text{-}336}$ crystal structure as search model. The asymmetric unit contains one Slfn2 $^{\Delta C116\text{-}W122}$ monomer and the crystal packing shows large solvent channels (Figure 4.1.25. D). The solvent content could be determined to 83 % (Table 4.4), possibly explaining why the crystals diffract not well despite the huge size.

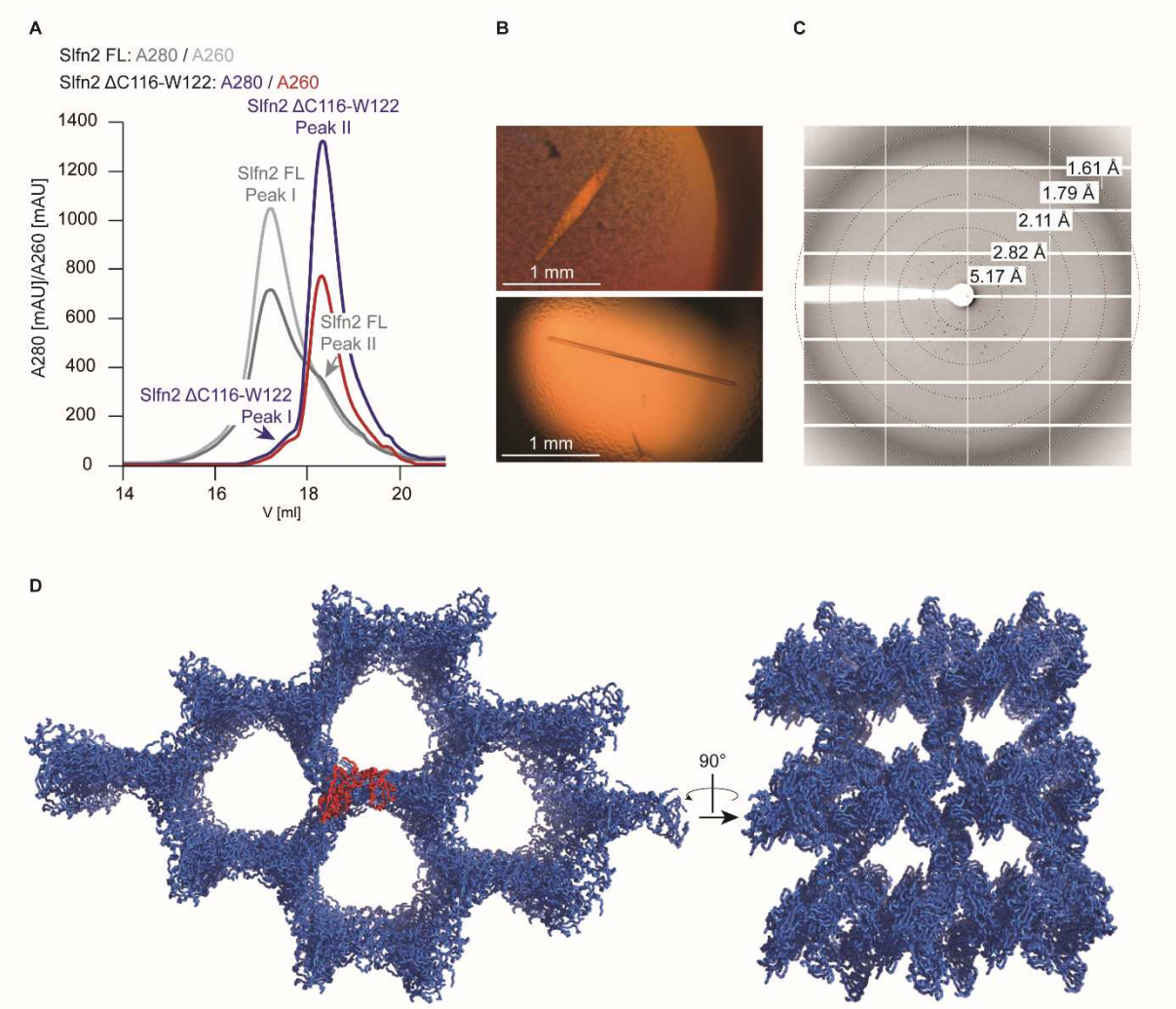



**Figure 4.1.25:** Towards the Slfn2$^{\Delta C116\text{-}W122}$ crystal structure. **A:** Analytical size exclusion chromatography using Superose 6 of Slfn2$^{\Delta C116\text{-}W122}$ (A280: blue, A260: red) and Slfn2$^{FL}$ (A280: dark grey, A260: light grey). **B:** Crystallization of Slfn2$^{\Delta C116\text{-}W122}$ in 0.15 M KCl, 10 mM $MgCl_2$, 50 mM MES pH 6.5, 2.5% PEG 8000. Crystals were up to 2 mm in size. **C:** Diffraction pattern of Slfn2$^{\Delta C116\text{-}W122}$ crystals. **D:** Crystal packing of Slfn2$^{\Delta C116\text{-}W122}$ crystals. One molecule (colored in red) is in an asymmetric unit. The crystals have a solvent content of 83% and contain large solvent channels.

**Table 4.4:** Crystal parameters of the Slfn2$^{\Delta C116\text{-}W122}$ crystal and refinement statistics.

| **Data collection and processing** | **Slfn2$^{\Delta C116\text{-}W122}$** |
|---|---|
| **Wavelength [Å]** | 0.98 |
| **Space group** | P $3_2$ 2 1 |
| **Unit cell dimensions** | |
| **a, b, c [Å]** | 143.2, 143.2, 104.2 |
| **α, β, γ [°]** | 90, 90, 120 |
| **Resolution range [Å]** | 48 - 4 (4.15 - 4.0) |
| **$R_{meas}$** | 7.6 (174) |
| **I/σI** | 7.1 (0.7) |
| **Completeness [%]** | 97.1 (97.3) |
| **Redundancy** | 2.5 (2.6) |
| **Solvent content [%]** | 83 |
| **Matthews coefficient [$A^3$/Da]** | 7.25 |
| **Refinement** | |
| **No. reflections** | 10482 (998) |
| **$R_{work}$/$R_{free}$** | 0.33/0.38 |
| **Number of atoms** | |
| **Protein** | 2417 |
| **Ligands** | 1 |
| **Average B-factor [Å]** | 220.43 |
| **RMSD** | |
| **Bond lengths [Å]** | 0.006 |
| **Bond angles [°]** | 1.38 |
| **Ramachandran plot** | |
| **Favored [%]** | 69.26 |
| **Allowed [%]** | 21.28 |
| **Outliers [%]** | 9.46 |

Values in parentheses are for the highest-resolution shell.

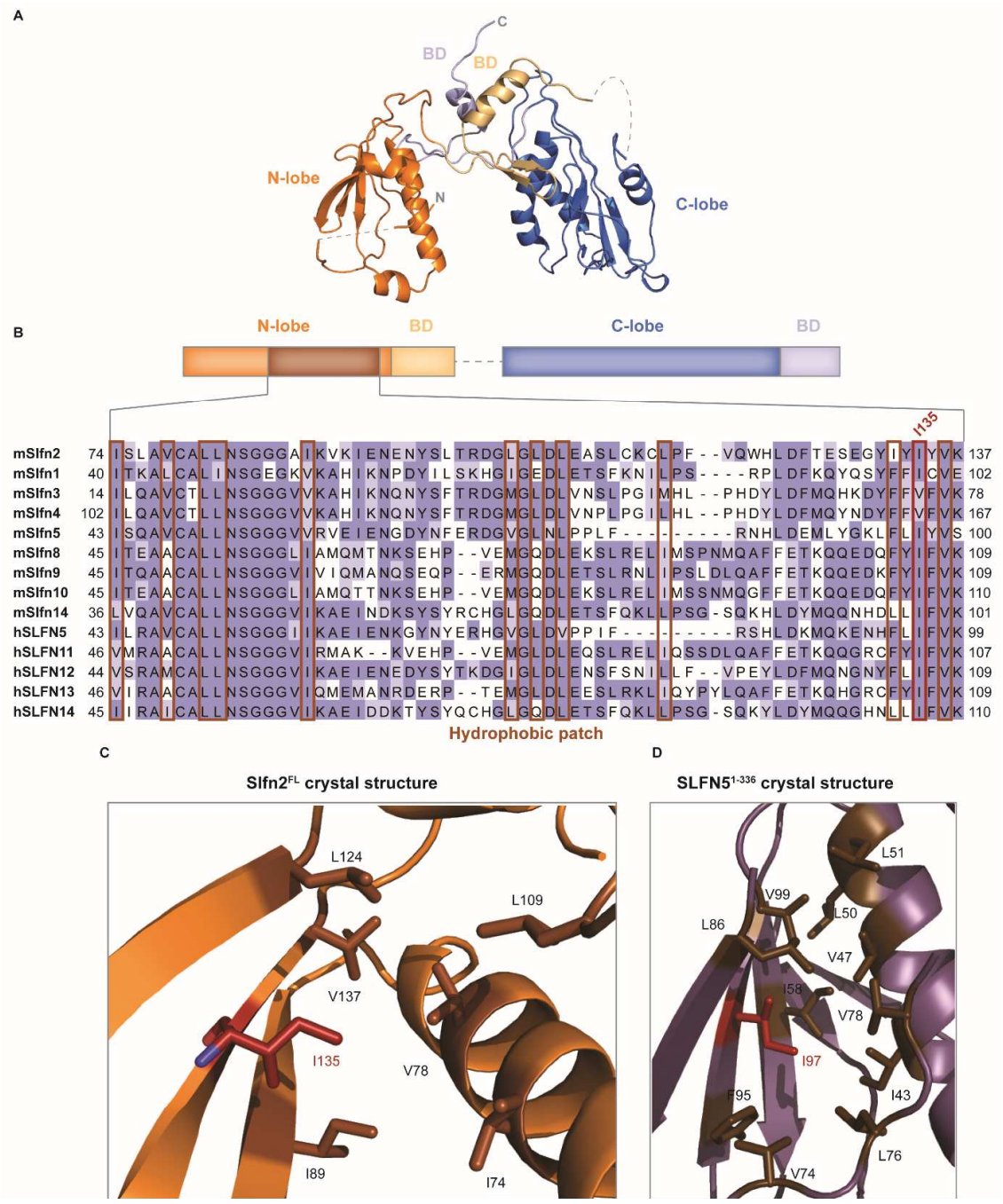



**Figure 4.1.26:** Crystal structure of Slfn2$^{\Delta C116\text{-}W122}$**. A:** Overall structure of Slfn2$^{\Delta C116\text{-}W122}$. **B:** Schematic overview and domain architecture of the Slfn2$^{\Delta C116\text{-}W122}$ crystal structure and multiple sequence alignment of all human and murine schlafen members of the N-lobe region, where the *elektra* mutation site (I135) is located. Hydrophobic residues are framed brownish. **C:** Close-up view of the hydrophobic patch, where the *elektra* mutation I135N would be located. **D:** Close-up view of the hydrophobic patch of the SLFN5$^{1\text{-}336}$ N-lobe region.

Throughout the figure, the following color code was chosen for Slfn2 $^{\Delta C116\text{-}W122:}$ N-lobe: orange; N-bridging domain: light orange, C-lobe: blue, C-bridging domain: light blue. I135 of Slfn2$^{FL}$ and the homologous residue in SLFN5$^{1\text{-}336}$ I97 are colored in red. Hydrophobic patch residues are colored in brown and presented as sticks. BD: bridging domain.

The crystal structure of Slfn2$^{\Delta C116\text{-}W122}$ could be solved and refined to $R_{work}/R_{free}$ of 0.33/0.38, respectively, indicating that the generated model might still include errors. Nevertheless, the overall structure depicts the same overall fold than SLFN5$^{1\text{-}336}$ (Figure 4.1.26. A).

The so called *elektra* mice are highly immunosuppressed due to a point mutation in the *slfn2* gene that leads to a substitution of isoleucine to asparagine at position 135[86]. I135 is located in the N-lobe region of Slfn2 (Figure 4.1.26. B). The structure reveals that I135 is located in a hydrophobic patch and is surrounded by several hydrophobic residues (Figure 4.1.26. C). I135 is conserved in all murine and human schlafen members, except for Slfn3 and Slfn4 that contain a valine instead of an isoleucine at this position (Figure 4.1.26. B). Moreover, the residues surrounding I135 are also conserved throughout the schlafen protein family (Figure 4.1.26. B), indicating that the hydrophobic patch region is present in all schlafen proteins as well. Indeed, SLFN5$^{1\text{-}336}$ has a similar hydrophobic core in the N-lobe region (Figure 4.1.26. D). Mutating isoleucine 135 to asparagine, thus inserting a polar side chain, would lead to misfolding and aggregation. Indeed, Slfn2$^{FL}$I135N cloned into a pET21 vector could not be solubly expressed in *E. coli* Rosetta, probably due to misfolding and aggregation (data not shown).

4.1.2.3. Nucleic acid binding properties of Schlafen 2

Since Slfn2$^{FL}$ contains the zinc finger motif and was also shown to co-purify RNA, nucleic acid binding properties were tested by EMSA and FA. The first EMSA showed binding to both, single-stranded 60- mer poly (dT) DNA (Figure 4.1.27. A) and double-stranded 60- mer DNA (Figure 4.1.27. B). A decrease in the intensity of free DNA indicates a clear binding of Slfn2$^{FL}$ to the substrates. For both substrates, the shift starts from 1 µM Slfn2$^{FL}$ concentration. Both EMSAs reveal several bands, indicating that two or more Slfn2$^{FL}$ molecules might bind to one DNA molecule. However, smear was observed, as well.

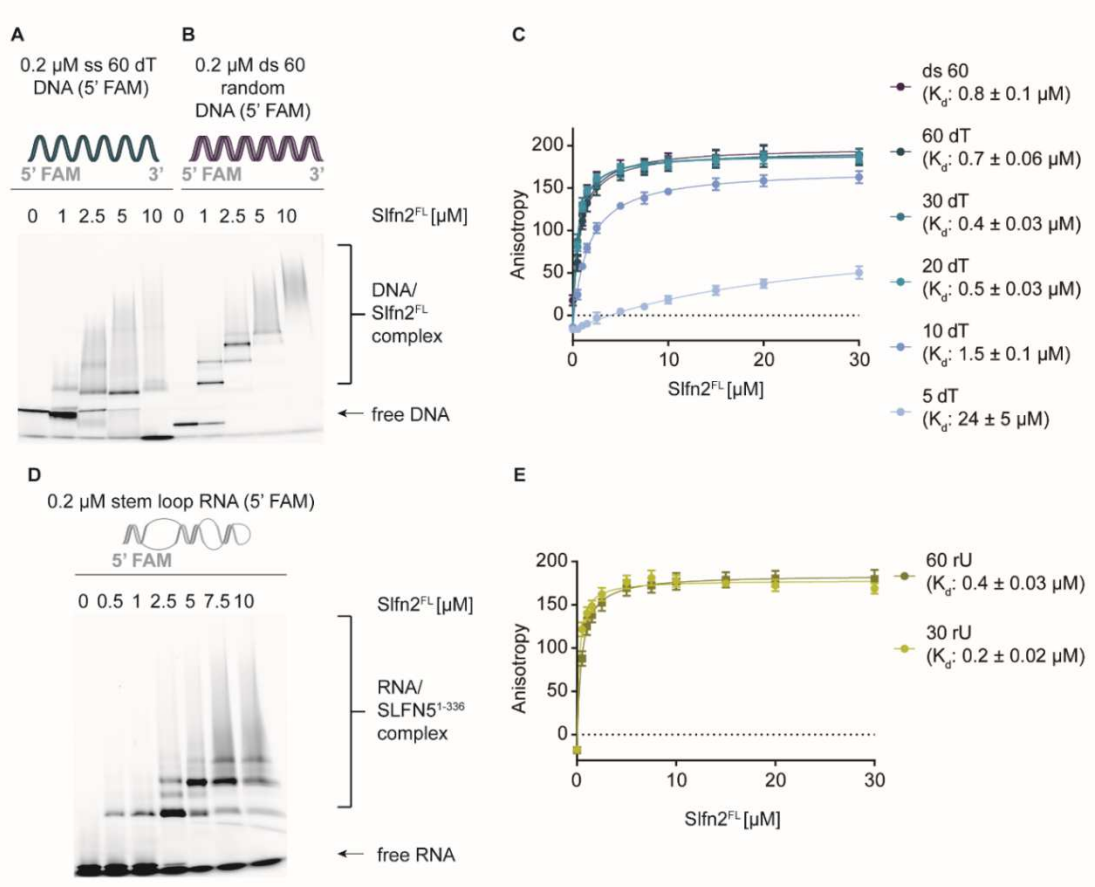



**Figure 4.1.27:** Analysis of nucleic acid binding properties by fluorescence anisotropy (FA) and electrophoretic mobility shift assay (EMSA) of Slfn2$^{FL}$. **A-B:** EMSA of Slfn2$^{FL}$ to single-stranded 60- mer poly (dT) (A) and double-stranded 60- mer DNA (B). **C:** Analysis of binding of Slfn2$^{FL}$ to single-stranded 60- mer poly (dT) DNA of various lengths and a double-stranded 60- mer DNA by FA. Data points represents the change in fluorescence anisotropy (arbitrary units). The various nucleic acid ligands are depicted with different colors. Data points represent the change in fluorescence anisotropy (arbitrary units). Data point and error bars represent the means ± s.d. from three independent experiments. The solid line represents the least square fit to Hill binding model. **D:** EMSA to visualize Slfn2$^{FL}$ binding to the head domain of 5S rRNA. **E:** Analysis of binding of Slfn2$^{FL}$ to single-stranded 30- and 60- mer poly (rU) RNA. All nucleic acid substrates were labeled with 6-carboxyfluorescin at the 5' end, indicated in the figure as (5' FAM).

Following FA analysis confirmed the findings, since a clear change in anisotropy signal was measured for single-stranded 60- mer poly (dT) DNA and double-stranded 60- mer DNA and the $K_d$s were determined to be 0.7 µM and 0.8 µM, respectively (Figure 4.1.27. B). Thus, it appears that $Slfn2^{FL}$ binds to single- and double-stranded DNA of the same length with a similar affinity. Generally, the affinity of $Slfn2^{FL}$ is approximately 3 times higher for double-stranded and 10 times higher for single-stranded DNA than the affinity of $SLFN5^{1-336}$ to the same substrates. The affinity of $Slfn2^{FL}$ correlates with the length of the DNA and increased from 24 µM to 1.5 µM for a 5- and 10- mer single-stranded poly (dT) DNA, respectively (Figure 4.1.27. C). All substrates longer than 20 bases are bound with a similar affinity. $Slfn2^{FL}$ binds single-stranded 20- and 30- mer poly (dT) DNA with $K_d$-values of 0.5 µM and 0.4 µM, respectively (Figure 4.1.27. C).

Since $Slfn2^{FL}$ was shown to co-purify a RNA substrate, the binding properties to RNA substrates were tested, as well. Thus, a stem loop RNA, which corresponds to the head domain of a 5S rRNA, was tested in EMSAs with $Slfn2^{FL}$. A shift can already be detected starting from 0.5 µM $Slfn2^{FL}$ concentration and already at 5 µM $Slfn2^{FL}$ concentration all free RNA has vanished (Figure 4.1.27. D). Here, as well, several bands per shift appear and also smear can be detected. The subsequent FA was carried out with single-stranded 30- and 60- mer poly (rU) RNA (Figure 4.1.27. E). For both substrates the $K_d$-values could be calculated to be 0.2 µM and 0.4 µM for single-stranded 30- and 60- mer poly (rU) RNA, respectively. Indeed, the $K_d$-values of the RNA substrates are slightly lower than the one for the respective DNA substrates, indicating that $Slfn2^{FL}$ might have a preference for RNA substrates. Interestingly, the best bound DNA substrate was a single-stranded 30- mer poly (dT) DNA and the best RNA substrate a single-stranded 30- mer poly (rU) RNA. This might indicate that 30 bases is the preferred size for $Slfn2^{FL}$.

Taken together, $Slfn2^{FL}$ is capable to bind both, DNA and RNA, but has marginally higher affinities to RNA *in vitro*. Furthermore, it has a much higher affinity to the tested substrates than $SLFN5^{1-336}$, although both constructs consist exclusively of the slfn core domain.

In order to verify the binding mode of $Slfn2^{FL}$ to the nucleic acid substrates several potential binding mutants were generated based on the calculated electrostatic surface potential of the Slfn2 model. It showed many stretches of positively charged residues in the origin of the horseshoe tunnel (Figure 4.1.28. A). The positively charged residues chosen for point mutations (H65, R67, K249 and K285) lie within this region and are partly conserved within the schlafen family (Figure 4.1.28. B). An initial assay to analyze the effect of the single-point mutations was the purification per se. Thus, all mutants were purified according to the initial purification protocol, following subsequent analysis on a Superose 6 (10/300) column and the peak fractions were applied to agarose gel electrophoresis to analyze co-purified RNA.

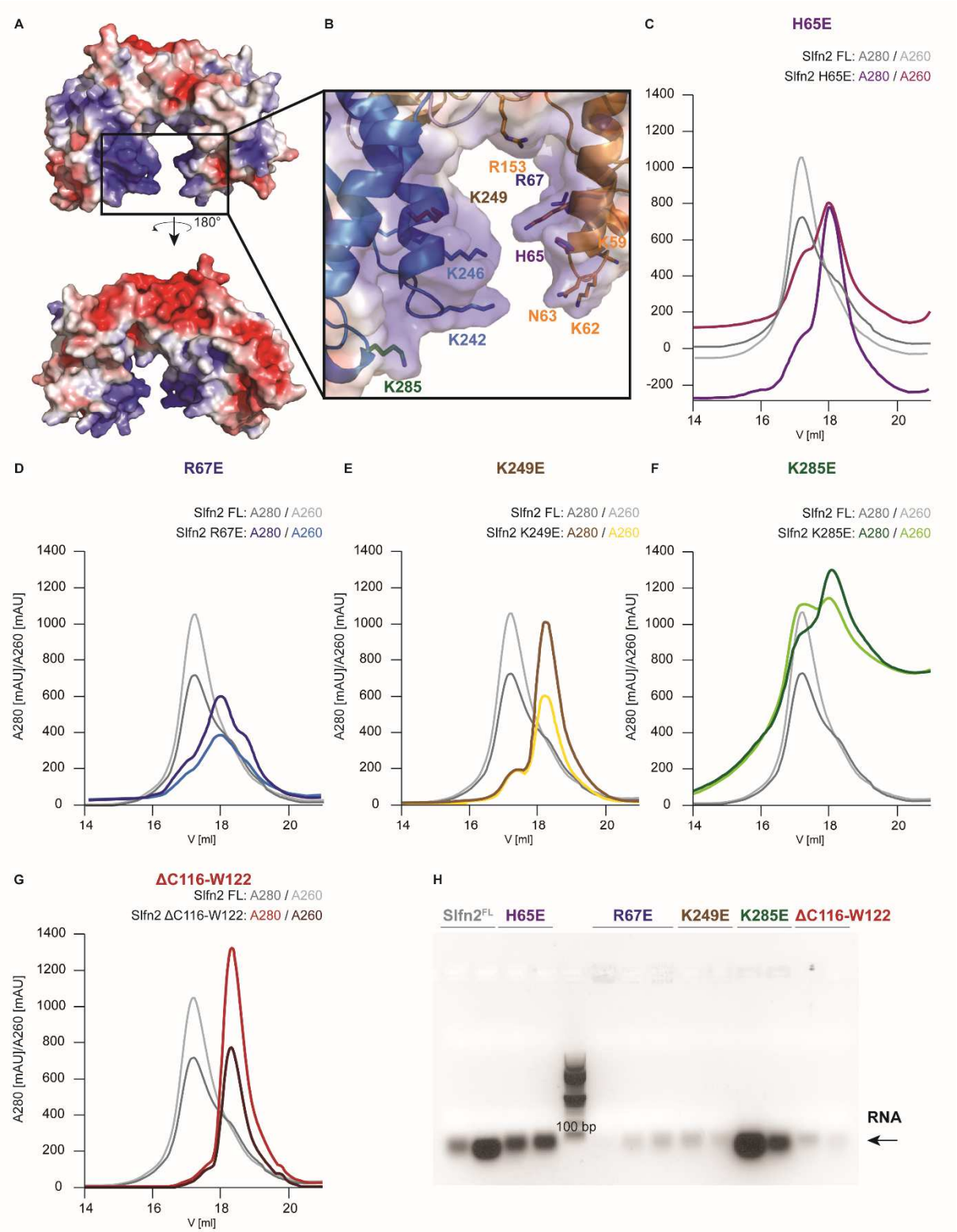



**Figure 4.1.28:** Generation and analysis of putative RNA-binding mutants of Slfn2$^{FL}$. **A:** Electrostatic surface potential of a Slfn2$^{FL}$ model colored from red (-4kT/e) to blue (4kT/e). **B:** Close-up view of the positive charged patch at the origin of the inner tunnel. **C-G:** Analytical size exclusion chromatography of different Slfn2 mutants using a Superose 6 10/300 column: Slfn2$^{H65E}$ colored in purple (C), Slfn2$^{R67E}$ colored in blue (D), Slfn2$^{K249E}$ colored in brown (E), Slfn2$^{K285E}$ colored in green (F) and the deletion mutant Slfn2$^{\Delta C116\text{-}W122}$ colored in red (G). **H:** Analysis of the size exclusion peaks on an agarose gel.

When mutating histidine 65 to glutamate in Slfn2, still two peaks were visible in analytical size exclusion chromatography analysis of the purified protein, however, the majority of the sample elutes in peak II with a low $A_{260}/A_{280}$ ratio (Figure 4.1.28. C). On an agarose gel, RNAs can be detected, however, it seems to be less compared to the wild type (Figure 4.1.28. H). Next, arginine 67 and lysine 249 of Slfn2 were mutated to glutamate each. For both constructs, no peak I could be detected during analytical size exclusion chromatography and the majority of sample eluted at approximately 18 ml with a low $A_{260}/A_{280}$ ratio (Figure 4.1.28. D, E). Indeed, analysis of co-purified RNA with Slfn2 R67E and Slfn2 K249E on agarose gels only depicts faint bands indicating only minor contaminations (Figure 4.1.28. H). Further, Slfn2 K285E, where the mutated lysine lies more distant from the tunnel entry site than the other mutated residues (Figure 4.1.28. B), eluted together with a significant amount of RNA from the size exclusion column at 17 ml (Figure 4.1.28. F). Consistently, the agarose gel depicts prominent bands for RNA within this sample (Figure 4.1.28. H). Taken together, arginine 67 and lysine 249, which lie opposite of each other in the tunnel, seem to play major roles in the co-purification and interaction of the RNA substrate. Histidine 65 appears to play a minor role in binding RNA, whereas lysine 285 seems to play no role in RNA binding at all. Probably, it is located too distant from the tunnel. Interestingly, also the hydrophobic insertion C116-W122 seems to play a role in RNA binding, because no peak I with high $A_{260}/A_{280}$ was detected on analytical size exclusion chromatography following protein purification (Figure 4.1.28. G). Thus, on the agarose gel just a faint RNA band could be detected for this deletion mutant $Slfn2^{\Delta C116\text{-}W122}$ (Figure 4.1.28. H). However, when testing those Slfn2 mutants on the EMSA, no differences in binding compared to the wild type protein could be detected (data not shown).

4.1.2.4. Schlafen 2 exhibits DNA-exonuclease activity

Next, exonuclease activity of Slfn2$^{FL}$ was tested, since in EMSAs with single-stranded DNA a smear was frequently observed and the slfn core domain of SLFN5$^{1-336}$ was also found to exhibit exonuclease activity.

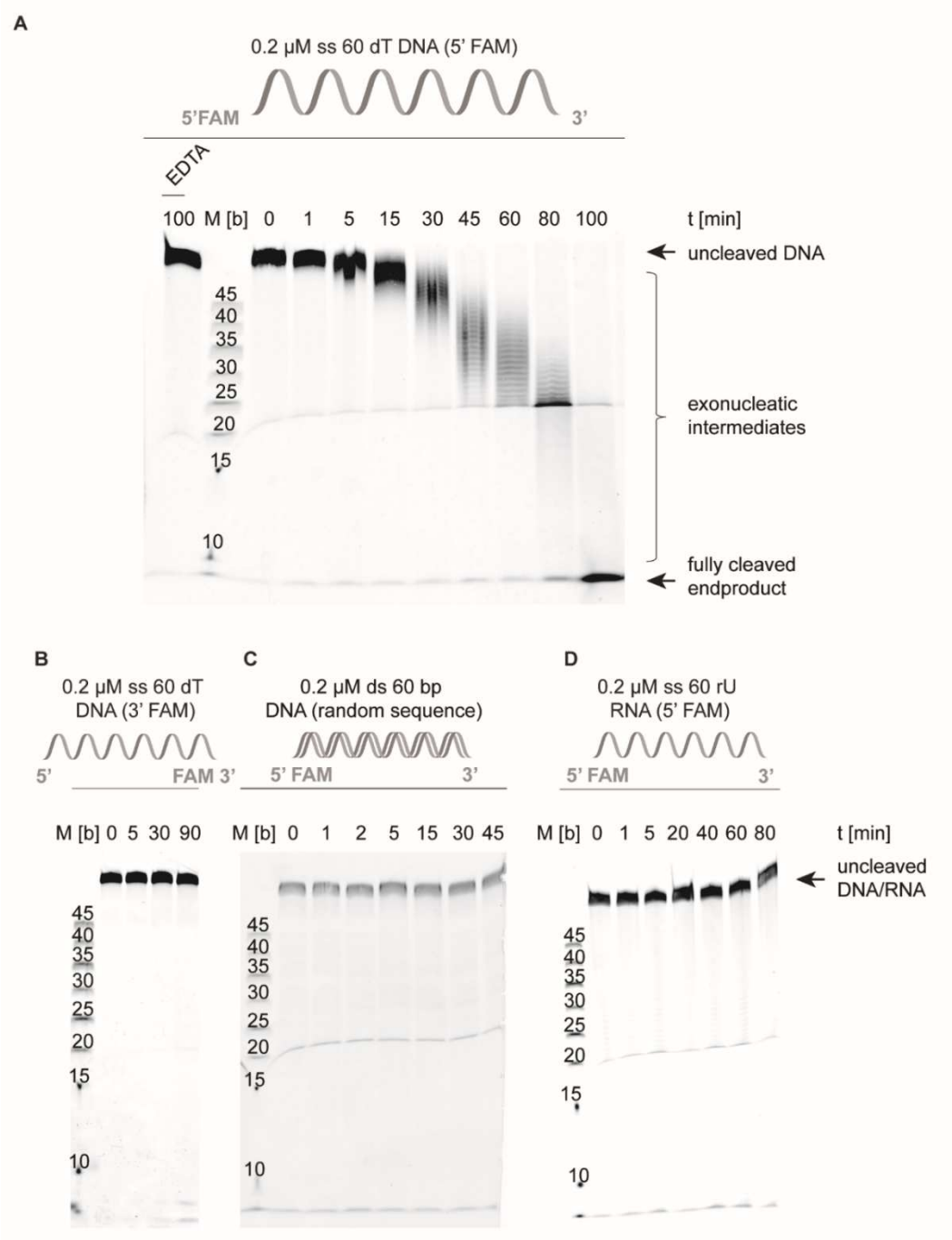



**Figure 4.1.29:** Analysis of exonuclease activity of Slfn2$^{FL}$. Exonuclease activity upon incubation with (A) single-stranded 60- mer poly (dT) DNA (5' FAM) in absence or presence of EDTA, (B) single-stranded 60- mer poly (dT) DNA (3' FAM), (C) double-stranded 60- mer DNA and (D) single-stranded 60- mer poly (rU) RNA. Nucleic acid substrates were labeled with 6-carboxyfluorescin either at the 5' end or at the 3' end, indicated in the figure as (5' FAM) or (3' FAM), respectively.

Similar to SLFN5$^{1-336}$ exonuclease activity of Slfn2 was observed on a single-stranded 60- mer poly (dT) DNA substrate that carries a fluorescein fluorophore at the 5' terminus (Figure 4.1.29. A). At 25°C, digestion starts after 5 minutes and the substrate is completely digested after 100 minutes. Since, a ladder is visible on the gel, a successive digestion starting from the 3' end can be assumed during the reaction (Figure 4.1.29. A). Addition of EDTA to the reaction mixture completely abolishes the exonuclease activity indicating a dependency on metal ions (Figure 4.1.29. A). Furthermore, there was no exonuclease activity detectable when the DNA substrate was labeled with fluorescein at the 3' terminus (Figure 4.1.29. C). This implies a 3'-5' directionality similar to SLFN5$^{1-336}$. Further, exonuclease activity of Slfn2 seems to be restricted to single-stranded DNA as no digestion was observed for double-stranded 60- mer DNA (Figure 4.1.29. D). Moreover, no digestion of a single-stranded 60- mer poly (rU) RNA could be observed, even though Slfn2 seems to have a preference for binding RNA than DNA (Figure 4.1.29. D).

Taken together, Slfn2$^{FL}$ shows exonuclease with similar characteristics than SLFN5$^{1-336}$. Both cleave single-stranded DNA in a 3'-5' directionality. Furthermore, the successive cleavage patterns seem to look similar for both proteins. This indicates similar functions and a common working mechanism for the slfn core region within the schlafen protein family.

4.1.2.5. Analysis of the Schlafen 2 interactome by affinity enrichment mass spectrometry

In order to determine the biological context of Slfn2$^{FL}$ affinity enrichment mass spectrometry experiments were conducted. For that purpose E.G7, a mouse T lymphocyte cell line, and Panc02, a mouse pancreatic ductal adenocarcinoma cell line, were used. Cells were transiently transfected with vectors carrying *slfn2$^{FL}$* and *slfn2$^{I135N}$* (*elektr*a mutation) genes. Cells transfected with the empty vector served as a control. Co-immunoprecipitation experiments were performed using anit-His$_{10}$ antibody and samples were subsequently measured using LC-MS. The t-test difference provide information about the enrichment of the respective proteins and the p-value marks the significance of a specific interaction. Furthermore, the higher the enrichment of the respective proteins compared to the control, the more right the protein can be found on the volcano plot (Figure 4.1.30). Slfn2$^{FL}$ could be enriched successfully and some potential interaction partners could be determined, whereas Slfn2$^{I135N}$ could not be enriched successfully, probably due to aggregation. Here the most enriched interaction partners determined by using the E.G7 cells are presented (Figure 4.1.30). The complete list of the whole interactome of Slfn2$^{FL}$ in E.G7 and Panc02 cells can be found in the appendix section (Table 7.1 and 7.2).

Highly enriched is the fatty acid desaturase 3 (Fads3), which has oxidoreductase activity and plays a role in fatty acid biosynthetic processes. It is potentially associated with the endoplasmic reticulum[180]. More ER-associated and in lipid metabolism participating proteins that interact with Slfn2$^{FL}$, are steryl-sulfatase (STS)[181], phospholipase D3 (Pld3)[182], extended synaptotagmin-1 (Esyt1)[183] and glycosylphosphatidylinositol anchor attachment 1 protein (Gpaa1)[184]. The pre-B cell leukemia transcription factor-interacting protein 1 (Pbxip1) is highly enriched, as well. On the one hand, it is localized in the nucleus and regulates pre B cell leukemia transcription factors, on the other hand it is associated with the cytoskeleton and plays a role in estrogen signaling[185]. Other cytoskeleton associated

Slfn2$^{FL}$-interactors are actin (Actc1)[186], CAP-Gly domain-containing linker protein 2 (Clip2)[187], calponin (Cnn2)[188] and tensin (Tns1)[188]. Although, Slfn2 was proposed to be localized in the cytoplasm[142] only two of the most enriched proteins are localized in the cytoplasm namely the E3 ubiquitin/ISG15 ligase tripartite motif protein 25 (Trim25)[189] and 5-AMP-activated protein kinase subunit beta-2 (Prkab2)[190]. Most interactors are membrane associated either at the ER, the plasma membrane or at endosomes. Slfn2$^{FL}$-interactors at the plasma membrane are unconventional myosin-1c (Myo1c)[191], extended synaptotagmin-1 (Esyt1)[183], serine/ threonine-protein kinase (PAK1)[192], netrin receptor UNC5B (Unc5b)[193] and metalloreductase 2 (Steap2)[194]. Localized at endosomes are metalloreductase 1 (Steap1)[195] and 2 (Steap2)[194]. More transcriptional regulators, which are zinc finger proteins as well, and are localized in the nucleus are Gata6[196] and four and a half LIM domains protein 1 (Fhl1)[197]. Two mitochondria associated proteins, namely cytochrome C1 (Cyc1)[198] and ferrochelatase (Fech)[199] had also been identified. However, the verification of the direct interactions would have to be confirmed.

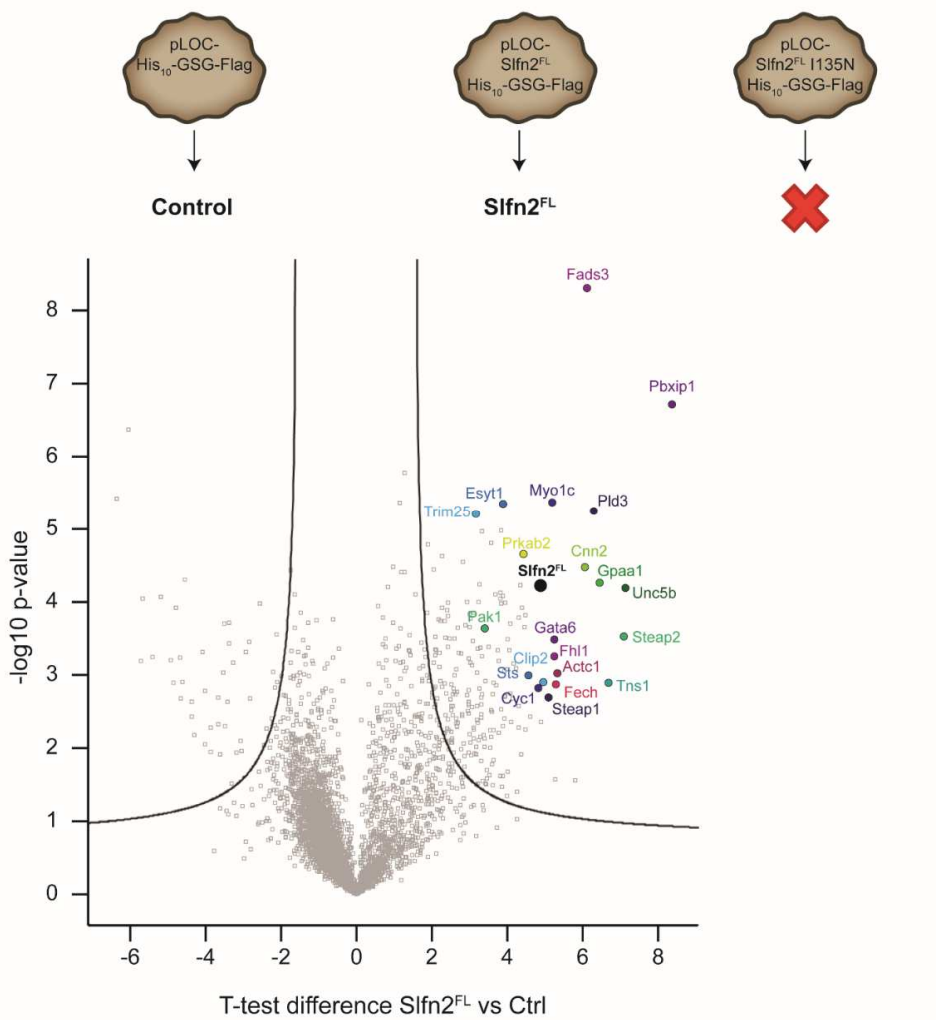



**Figure 4.1.30:** Mouse E.G7 cell supernatant was subjected to affinity enrichment mass spectrometry experiments using anti-His$_{10}$ antibody. For the control E.G7 cells carrying an empty vector were used. Slfn2$^{I135N}$ could not be enriched successfully and no interactome could be established, whereas SLFN2$^{FL}$ could be enriched successfully (right side, colored in black). Assays were performed in triplicates and a two-sided and two-sample t-test shows significant enrichment of Slfn2$^{FL}$ interactors (volcano plot, right side).

### 4.2. Purification and characterization of full length subgroup III schlafen protein members

Not only the investigation in the N-terminal domain is important, but also the characterization of full length schlafen proteins is crucial. However, so far structural and biochemical information about the full length subgroup III schlafen members including the C-terminal predicted helicase domain are lacking. Therefore, different constructs of human schlafen 11 and schlafen 5 were generated and expressed using different bacterial and eukaryotic expression systems.

#### 4.2.1. Purification and biochemical characterization of full length human Schlafen 5

Compared to the N-terminal truncated version of SLFN5, the other domains could not be expressed solubly. Neither the full length protein nor N-terminal truncated versions harboring the helicase domain could be soluble expressed using *E. coli*. The isolated helicase domain could not even be expressed using *S. cerevisiae*, whereas full length SLFN5 was soluble in *S. cerevisiae* and insect cells (High Five).

**Table 4.5:** Overview of the SLFN5 constructs used for expression and purification. The molecular weight of the proteins are given in kDa and the respective length in numbers of amino acids (aa). Furthermore, the isoelectric points (pI), the kind of affinity tag plus additional protease cleavage sites, the expression system used and the solubility of SLFN5 are indicated.

| Construct | Description | Length [aa] | MW [kDa] | Theoretical pI | Affinity/ Solubility tag + Cleavage site | Expression system | Solubility - Achievements |
|---|---|---|---|---|---|---|---|
| **pET21-SLFN5$^{FL}$** | Full length SLFN5 | 899 | 102 | 8.5 | C-terminal $His_6$ | *E. coli* (Rosetta, BL21 lemo) | Not soluble |
| **pRS426-SLFN5$^{FL}$** | Full length SLFN5 | 915 | 104 | 8.49 | C-terminal $His_8$ + TEV cleavage site | *S. cerevisiae* | Soluble (partly characterized) |
| **pFBDM-SLFN5$^{FL}$** | Full length SLFN5 | 910 | 103 | 8.49 | N-terminal $His_6$ + PreScission site | High Five Insect cells | Not soluble |
| **pFBDM-SLFN5$^{FL}$** | Full length SLFN5 | 1303 | 147 | 7.5 | N-terminal $His_6$ + MBP + TEV cleavage site C-terminal StrepII | High Five Insect cells | Soluble |
| **pET21-SLFN5$^{424-891}$** | Linker + SWADL + Helicase domain | 477 | 54.2 | 8.69 | C-terminal $His_6$ | *E. coli* (Rosetta) | Not soluble |
| **pET21-SLFN5$^{542-891}$** | Helicase domain | 359 | 41.0 | 8.55 | C-terminal $His_6$ | *E. coli* (Rosetta) | Not soluble |
| **pRS426-SLFN5$^{424-891}$** | Linker + SWADL + Helicase domain | 491 | 55.7 | 8.77 | C-terminal $His_8$ + TEV cleavage site | *S. cerevisiae* | Not soluble |
| **pRS426-SLFN5$^{542-891}$** | Helicase domain | 373 | 42.5 | 8.69 | C-terminal $His_8$ + TEV cleavage site | *S. cerevisiae* | Not soluble |

4.2.1.1. Purification of full length Schlafen 5 expressed in *S. cerevisiae*

Since full length SLFN5$^{FL}$ could not be expressed in bacterial expression systems, a *S. cerevisiae* expression strategy in a fermenter was established. The yeast expression construct carries a C-terminal, plasmid encoded octahistidine tag and consists overall of 915 amino acids corresponding to a molecular weight of 104 kDa (Figure 4.2.1. A).

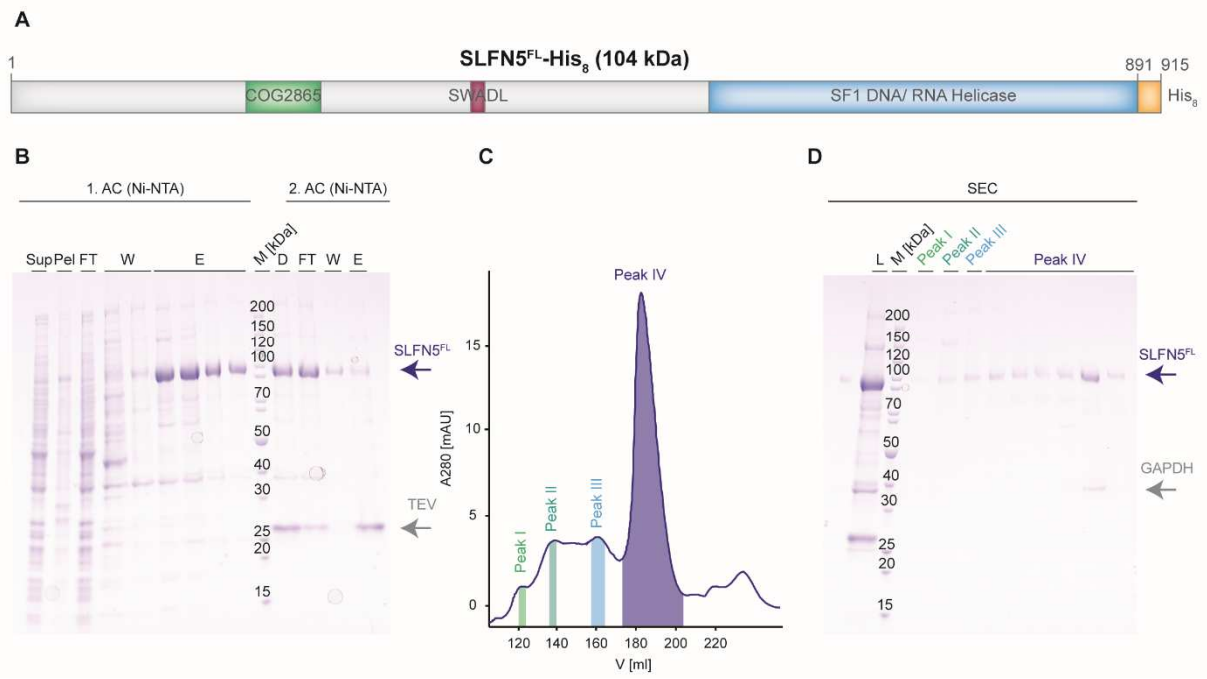



**Figure 4.2.1:** Purification of SLFN5$^{FL}$ from *S. cerevisiae* BCY123. **A:** Schematic overview of the full length SLFN5 construct carrying a plasmid encoded C-terminal octahistidine tag. **B:** Coomassie-stained SDS-PAGE of the fractions collected during the first and second affinity chromatography (AC) steps of the SLFN5$^{FL}$ purification. The elution fractions of the first AC, as well as the flow-through and washing fractions of the second AC contained SLFN5$^{FL}$ (indicated by an arrow). **C:** Chromatogram of the size exclusion chromatography (SEC) of SLFN5$^{FL}$ using a S200 26/60 column. Peak II to IV contain SLFN5$^{FL}$ in different oligomeric states. The blue line indicates the absorption at λ=280nm. **D:** Subsequent Coomassie-stained SDS-PAGE of the fractions collected during SEC. The arrow indicates SLFN5$^{FL}$ and GAPDH. Sup: supernatant, Pel: pellet, FT: flow-through, W: wash, E: elution, D: dialysate, L: load, M: marker in kDa.

This C-terminal octahistidine tag was used to capture SLFN5$^{FL}$ by loading the cell lysate onto Ni-NTA beads. The elution fractions contained besides SLFN5$^{FL}$ several impurities. All elution fractions were pooled, dialyzed and the octahistidine tag was cleaved with the TEV-protease (25 kDa), before applying the sample again onto Ni-NTA beads. SLFN5$^{FL}$ was not entirely cleaved and thus partially bound to the Ni-NTA resin (Figure 4.2.1. B). The flow-through and washing fractions were pooled, concentrated and applied onto a size exclusion column (S200 26/60) (Figure 4.2.1. C). Of the four appeared peaks, peak II to IV contain SLFN5$^{FL}$ protein (Figure 4.2.3.1. D). Peak IV contains the monomeric protein, whereas

peak II and III possibly correspond to dimers or higher oligomeric forms and are possibly contaminated with DNA. Furthermore, peak IV contains besides assumedly monomeric protein an impurity of 30-40 kDa, which could be identified as Glycerinaldehyd-3-phosphat-Dehydrogenase (GAPDH) (36 kDa) by subsequent mass spectrometry analysis. Finally, 376 µg protein of 15 liter yeast expression culture were obtained. The $A_{260}/A_{280}$ ratio was 0.7, indicating the presence of nucleic acid contaminations.

Taken together, this purification strategy of $SLFN5^{FL}$ is the best so far. However, the yield was very little and contained protein impurities as well as nucleic acid contaminations. Furthermore, the protein is prone to precipitation and 500 mM NaCl throughout the purification had been used. Nevertheless, the protein was used for subsequent characterization.

#### 4.2.1.2. Biochemical characterization of full length Schlafen 5

To determine the oligomeric state of $SLFN5^{FL}$ an analytical size exclusion chromatography on a S200 (10/300) column was conducted. $SLFN5^{FL}$ eluted as one single peak at approximately 15 ml (Figure 4.2.2. A). This peak overlays with the 158 kDa marker protein peak, indicating a monomeric state, since $SLFN5^{FL}$ has a calculated molecular weight of 104 kDa. So far no stable DNA-$SLFN5^{FL}$ complex could be reconsitituted by size exclusion chromatography, regardless if single-stranded or double-stranded DNA was used.

Nevertheless, nucleic acid binding properties were tested in electrophoretic mobility shift assay (EMSA). A clear shift could be observed when incubating $SLFN5^{FL}$ with single-stranded 30- mer poly (dT) DNA, regardless, if the DNA is labeled at the 3' or 5' terminus (Figure 4.2.2. B). At lower $SLFN5^{FL}$ concentrations one shift is detectable and upon increasing protein concentration, several shifts are detectable. This suggest that several protein molecules bind to one DNA molecule. Binding takes place from a $SLFN5^{FL}$ concentration of 50 nM. Thus, $SLFN5^{FL}$ binds the substrate with a higher affinity than $SLFN5^{1-336}$, where binding occurs starting at a concentration of 2 µM. This indicates that binding does not only take place at the N-terminal region, but that also the helicase domain significantly contributes to the DNA binding.

Interestingly, no binding occurs with single-stranded 30- mer poly (dA) DNA as substrate (Figure 4.2.2. C), whereas binding with a single-stranded 30- mer poly (dT) DNA takes place. Binding to a single-stranded 30- mer poly (rU) RNA could be observed starting also from 50 nM $SLFN5^{FL}$ concentration, but the pattern of the shift looks differently (Figure 4.2.2. D). Next, a DNA/RNA hybrid consisting of poly (dA) forward DNA strand aligned with a poly (rU) reverse RNA strand was tested. No binding was observed here either (Figure 4.2.2. E). However, since $SLFN5^{FL}$ does not bind poly (dA) DNA, which was used as the forward strand, it might still bind to a different DNA/RNA hybrid aligned with a different forward strand.

$SLFN5^{FL}$ does not only bind single-stranded DNA, but also double-stranded, since a shift can be observed with a single-stranded 60- mer poly (dT) DNA and double-stranded 60- mer DNA (Figure 4.2.3. E). For both a shift starting from 50 nM $SLFN5^{FL}$ can be observed. In the assay, using single-

stranded DNA smear appears with increasing protein concentrations. This might also indicate exonuclease activity. However, for the double-stranded assay, it also seems to smear, which might be due to potential helicase activity.

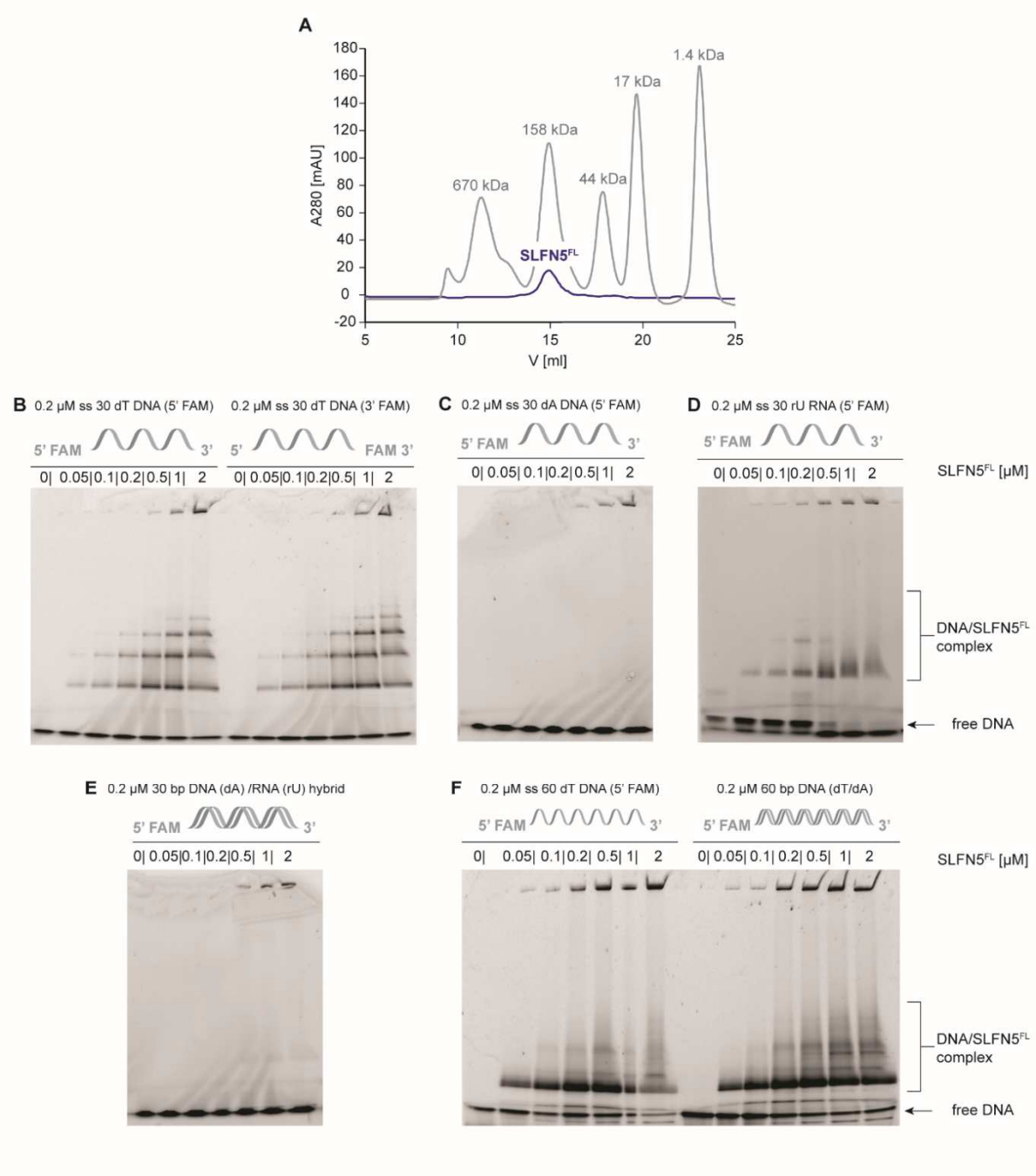



**Figure 4.2.2:** Analysis of nucleic acid binding properties of SLFN5$^{FL}$ by size exclusion chromatography (SEC) and electrophoretic mobility shift assay (EMSA). **A:** Chromatogram of the SEC of SLFN5$^{FL}$ using a S200 10/300 column. The blue and the grey lines indicate the absorption at λ=280 nm of SLFN5$^{FL}$ and marker proteins with the respective sizes indicated. **B-F:** Analysis of nucleic acid binding properties of SLFN5$^{FL}$ by EMSAs. Different SLFN5$^{FL}$ concentrations were titrated to (B) single-stranded 30- mer poly (dT) DNA either fluorescently labeled at the 5' or 3' terminus, (C) single-stranded 30- mer (dA) DNA, (D) single-stranded 30- mer poly (rU) RNA, (E) DNA/RNA hybrid, (F) single-stranded 60- mer poly (dT) and double-stranded 60- mer DNA.

Helicases are proteins that unwind DNAs/RNAs under the consumption of ATP[200]. Therefore, full length SLFN5 was tested for ATP binding performing the differential radial capillary action of ligand assay (DRaCALA). ATP-binding of SLFN5$^{FL}$ was tested in absence or presence of nucleic acid substrates known to bind SLFN5$^{FL}$: single-stranded 30- mer poly (dT) DNA and single-stranded 30- mer poly (rU) RNA. In both cases, in presence or in absence of nucleic acid, the [α-$^{32}$P] ATP remained in the center together with SLFN5$^{FL}$ (Figure 4.2.3. A). This indicates that ATP binding occurs, regardless if the substrate is present or absent. Next, ATP hydrolysis was tested by incubating SLFN5$^{FL}$ in presence or absence of single-stranded 30- mer poly (dT) DNA using ATP, which is radioactively labeled at the γ-phosphate ([γ-$^{32}$P] ATP). If ATP would be hydrolyzed, the free $^{32}P_i$ would migrate on the TLC plate, indicating hydrolysis of ATP. However, for SLFN5$^{FL}$ no ATP hydrolysis could be observed, since the radioactive labeled ATP stayed at the bottom together with the protein (Figure 4.2.3. B). Furthermore, with a linearized 5000 bp plasmid DNA no hydrolysis could be observed. Next, different buffer conditions were tested, however, neither in a HEPES/NaCl/$MgCl_2$ composition nor in a Tris/KCl/$MgCl_2$ solution hydrolysis could be observed.

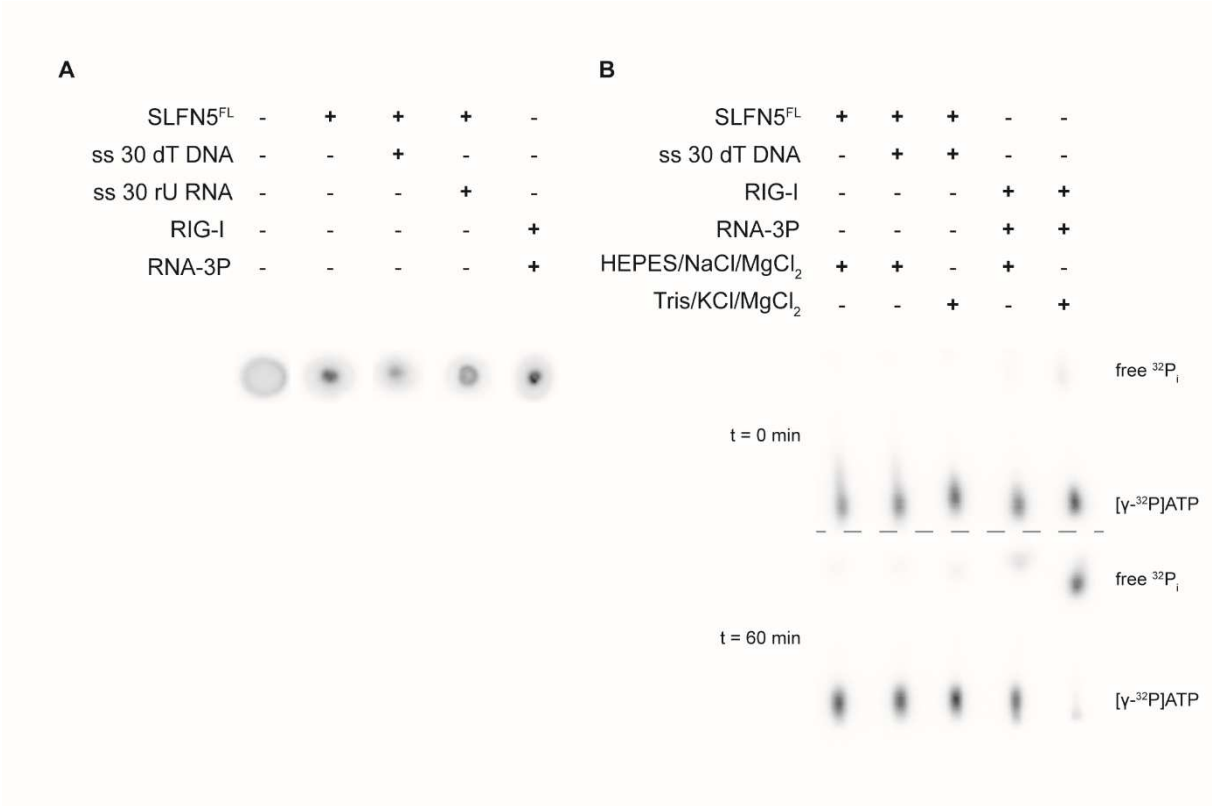



**Figure 4.2.3:** Analysis of ATP binding by differential radial capillary action of ligand assay (DRaCALA) and ATP hydrolysis of SLFN5$^{FL}$. **A:** DRaCALA ATP-binding assay of SLFN5$^{FL}$ in presence or absence of single-stranded 30- mer poly (dT) DNA or single-stranded 30- mer poly (rU) RNA substrates. **B:** ATP-hydrolysis assay of SLFN5$^{FL}$ in presence or absence of single-stranded 30- mer poly (dT) DNA in HEPES or Tris-buffer. **A-B:** RIG-I in presence with its substrate (tri-phosphorylated RNA) served as positive control[201].

Taken together, SLFN5$^{FL}$ is a monomeric protein in solution, binding to both, single-stranded and double-stranded DNA with a higher affinity than SLFN5$^{1-336}$ estimated by EMSAs. So far ATP binding, but no hydrolysis could be detected for SLFN5$^{FL}$.

4.2.1.3. Preparation and analysis of the co-purified nucleic acid substrate of full length Schlafen 5

Since the preparation of SLFN5$^{FL}$ expressed in *S. cerevisiae* yielded only little protein, the expression strategy was changed to insect cells High Five. Two different constructs were generated: His$_6$-SLFN5$^{FL}$ carrying a N-terminal hexahistidine tag with a theoretical molecular weight of 103 kDa and His$_6$-MBP-SLFN5$^{FL}$-Strep carrying a N-terminal hexahistidine and maltose-binding-protein (MBP) tag and a C-terminal Strep-tag with 147 kDa (Figure 4.2.4. A). For both constructs quite some SLFN5$^{FL}$ stayed in the insoluble pellet fraction after lysis. However, for the construct carrying the MBP tag the solubility was significantly increased. For both constructs, the hexahistidine-tag was used to capture SLFN5$^{FL}$ from the lysate. Afterwards the proteins were applied onto a size exclusion chromatography (Figure 4.2.4. B-D). For both proteins, four peaks were visible. His$_6$-SLFN5$^{FL}$ was just detectable in peak III, running at approximately 75 ml (Figure 4.2.4. B-D). His$_6$-MBP-SLFN5$^{FL}$-Strep is visible in peak I-III, probably in different oligomeric states or contaminated with DNA (Figure 4.2.4. C-F). Indeed, His$_6$-MBP-SLFN$^{FL}$ Peak I-II show an $A_{260}/A_{280}$ ratio of 0.86, indicating high contaminations of DNA. In His$_6$-SLFN5$^{FL}$ Peak III and His$_6$-MBP-SLFN5$^{FL}$ Peak III the $A_{260}/A_{280}$ ratio were just 0.4 and 0.53, respectively, assuming that less contaminations with nucleic acids. Interestingly, when analyzing the different peak fractions on an agarose gel, also in each of the peak III fractions an approximately 100 bp DNA band appears (Figure 4.2.4. E-G). This band has roughly the same size as the 5S rRNA co-purified from Slfn2 (Figure 4.2.4. G). Subsequently, those peak III bands were analyzed on a bioanalyzer. For both single peaks with a size of 113 nt to 148 nt were observed (Figure 4.2.4. H-I). RNA-sequencing data revealed, that SLFN5$^{FL}$ also co-purifies ribosomal RNA, however mainly 16 S and 5.8S rRNA and almost no 5S rRNA.

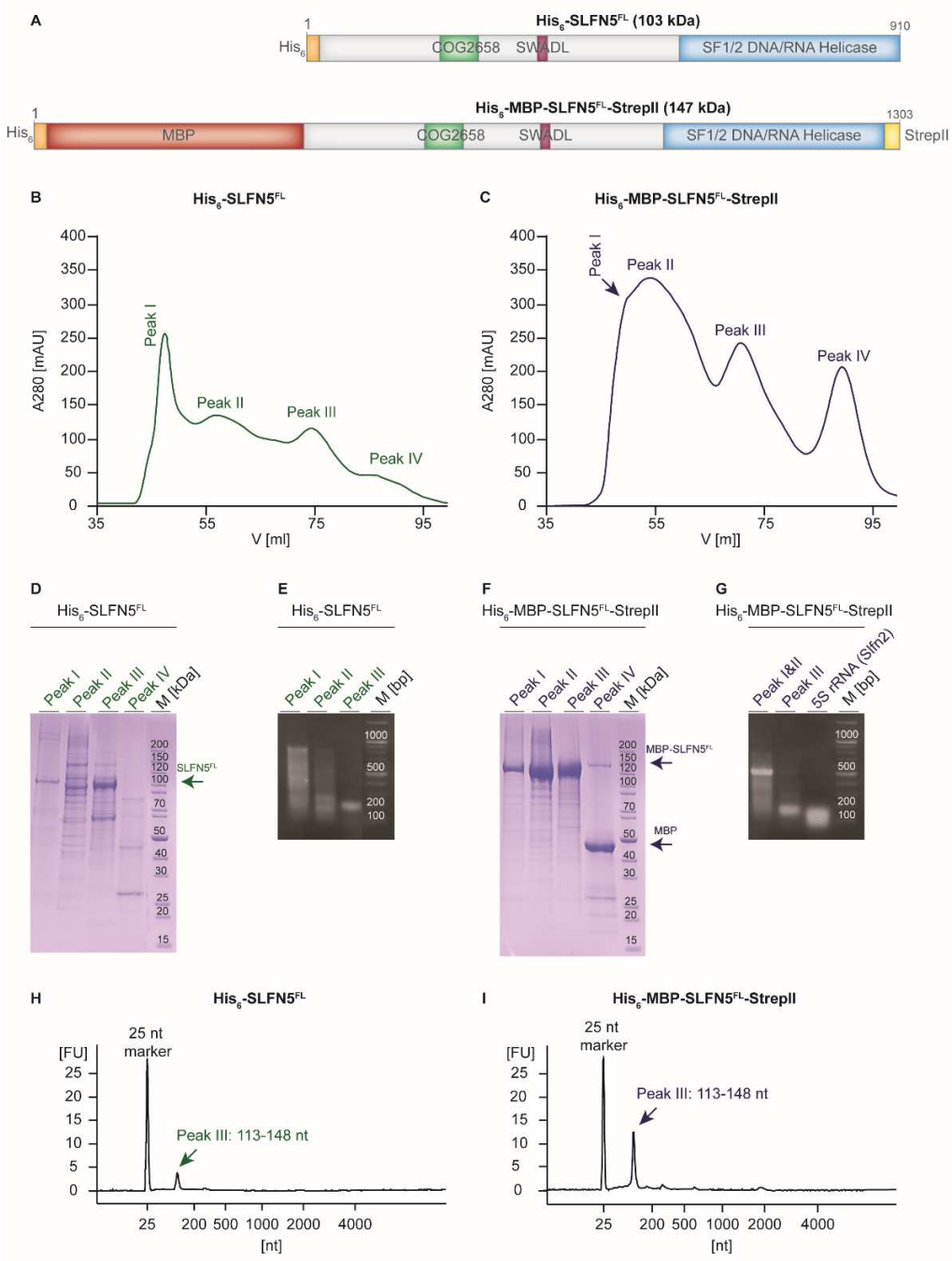



**Figure 4.2.4:** Purification and subsequent nucleic acid substrate analysis of SLFN5$^{FL}$ and MBP-SLFN5$^{FL}$ expressed in High Five. **A:** Schematic overview of the full length SLFN5 constructs His$_6$-SLFN5, carrying an N-terminal hexahistidine tag, and His$_6$-MBP-SLFN5-StrepII with N-terminal hexahistidine and maltose-binding-protein-tag (MBP) and C-terminal StrepII-tag. **B-C:** Chromatogram of the size exclusion chromatography using a S200 16/60 column of His$_6$-SLFN5$^{FL}$ and His$_6$-MBP-SLFN5$^{FL}$-StrepII, respectively. **D-F:** Coomassie-stained subsequent SDS-PAGE of the fractons collected during the size exclusion step. **E-G:** Agarose gel of the different peak fractions analyzing the co-purified nucleic acid contaminations of His$_6$-SLFN5$^{FL}$ and His$_6$-MBP-SLFN5$^{FL}$-StrepII, respectively. **H-I:** Analysis of the nucleic acid substrates by using a bioanalyzer. FU: Fluorescence units.

4.2.2. Towards full length human Schlafen 11

Since SLFN11 seems to play important roles in DNA-repair, tumorigenesis and has a potential application as a clinical biomarker, the biochemical and structural characterization of SLFN11 would be of scientific interest. Consequently, initially different SLFN11 constructs of different lengths and various expression systems using different *E. coli* strains and *S. cerevisiae* have been tested. (Table 4.6). Generally, SLFN11 was not soluble expressed in *E. coli*, regardless if the full length protein or the C-terminal truncated version of SLFN11$^{1-357}$ were tried. Soluble full length SLFN11, however, could be obtained by expression in *S. cerevisiae*.

**Table 4.6:** Overview of the SLFN11 constructs used for expression and purification. The molecular weight is indicated in kDa and the length of the protein in numbers of amino acids (aa). Furthermore, the isoelectric points (pI), the kind of affinity tag plus additional protease cleavage sites, the expression system used and the solubility of SLFN11 are indicated.

| Construct | Description | Length [aa] | MW [kDa] | Theore-tical pI | Affinity/ Solubility tag + Cleavage site | Expression system | Solubility |
|---|---|---|---|---|---|---|---|
| **pET21-SLFN11$^{FL}$** | Full length SLFN11 | 912 | 104 | 7.86 | C-terminal $His_6$ | *E. coli* (Rosetta, Arctic, Shuffle) | Not soluble |
| **pET28M-SUMO1-SLFN11$^{FL}$** | Full length SLFN11 | 1002 | 114 | 7.69 | N-terminal $His_6$-SUMO1 + SenP2 cleavage site | *E. coli* (Rosetta) | Not soluble |
| **pRS424-SLFN11 FL** | Full length SLFN11 | 923 | 105 | 8.00 | C-terminal $His_8$ + TEV cleavage site | *S. cerevisiae* | Soluble |
| **pET21-SLFN11$^{1-357}$** | Slfn core domain | 588 | 67.7 | 6.00 | N-terminal GST-tag + PreScission site | *E. coli* (Rosetta) | Not soluble |

The full length SLFN11 version cloned in pRS424 and expressed in *S. cerevisiae* BCY123 under the control of a galactose-inducible promotor, carries a C-terminal octahistidine tag, thus encoding for a 923 amino acid (105 kDa) protein (Figure 4.2.5. A). Although, soluble protein fractions after cell lysis could be detected, the subsequent purification did not reach sufficient yield and purity of SLFN11 protein. The cell lysate was first applied onto Ni-NTA beads. Besides the 105 kDa SLFN11, a multitude of unspecifically bound proteins eluted from the beads as well. After loading the eluted fractions onto a heparin column, many impurities remained. Following preoteolytic cleavage with TEV-protease the octahistidine-tag should be removed. Theoretically, the protease and the impurities should bind to the column with the hexahistidine tag and with unspecific interactions, respectively, and the cleaved SLFN11 should stay in the flow-through. However, no SLFN11 could be detected in the flow-through, but a faint band of approximately 105 kDa still in the eluted fractions (Figure 4.2.5. B). Taken together, with this

expression and purification strategy no sufficient amounts of SLFN11 with the required purity were obtained. Furthermore, *in vitro* the protein is prone for aggregation and impurities are hard to separate.

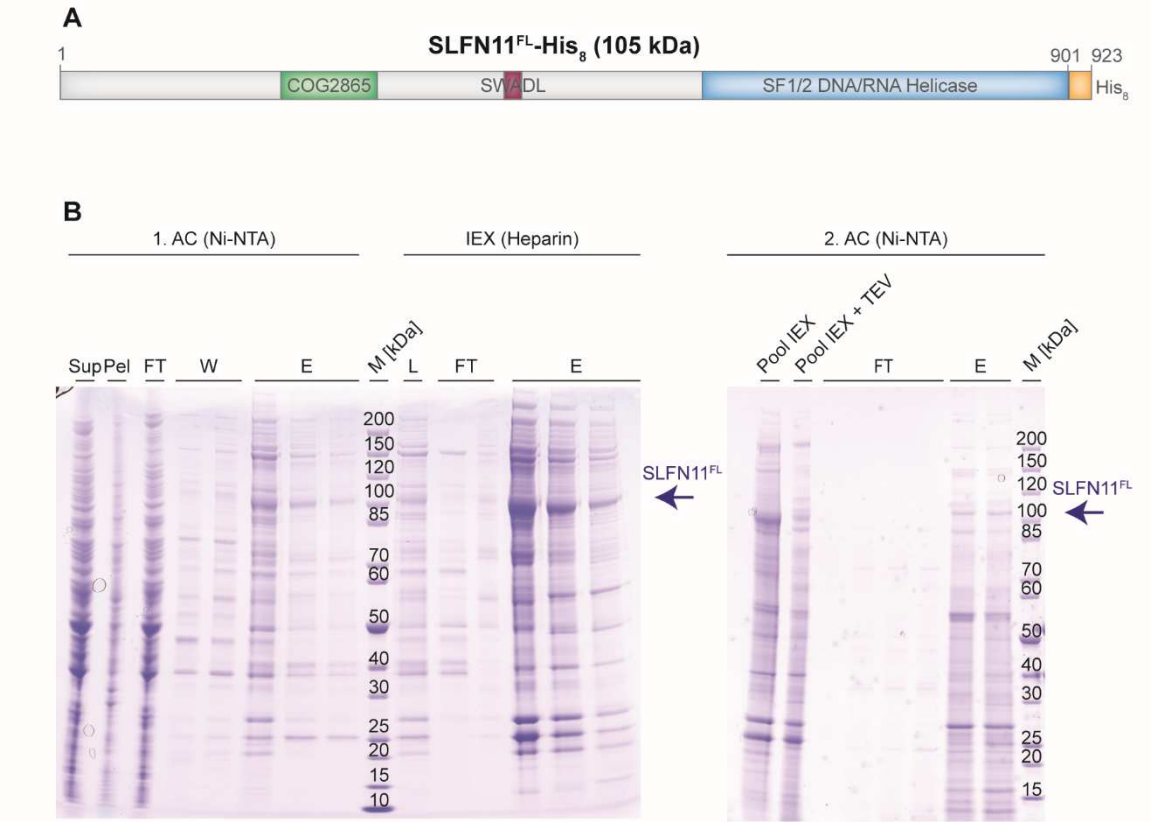



**Figure 4.2.5:** Purification of full length human SLFN11 from *S. cerevisiae* BCY123. **A:** Schematic overview of the full length SLFN11 construct carrying a C-terminal octahistidine tag. **B:** Coomassie-stained SDS-PAGE of the fractions collected during affinity chromatography (AC) using Ni-NTA beads, ion exchange chromatography (IEX) and subsequent AC after TEV-protease cleavage. The arrow indicates SLFN11. Sup: supernatant, Pel: pellet, FT: flow-through, W: wash fractions, E: elution fractions, L: load, M: marker in kDa.

# 5. Discussion

## 5.1. The slfn core domain

### 5.1.1. The slfn core domain is a highly conserved domain

The schlafen protein family is a highly conserved protein family with various cellular functions, but of so far unknown molecular working mechanism. The individual family members of human and murine schlafen share high sequence similarities (up to 89.9 %, Figure 7.2). When this thesis was initiated, no structural information was available for neither the full length protein nor the slfn core domain. For structural characterization of the slfn core domain, crystal structures of N-terminal human SLFN5 and full length murine Slfn2 were solved. Here, for the first time a high-resolution structure of a human schlafen protein is reported. Both proteins show a similar overall structure that was also observed for the recently published slfn core domain of rat Slfn13[126] (Figure 5.2. A). The slfn core domain depicts a horseshoe shape, consisting of an N-terminal lobe, a C-terminal lobe and two bridging domains (Figure 5.1).

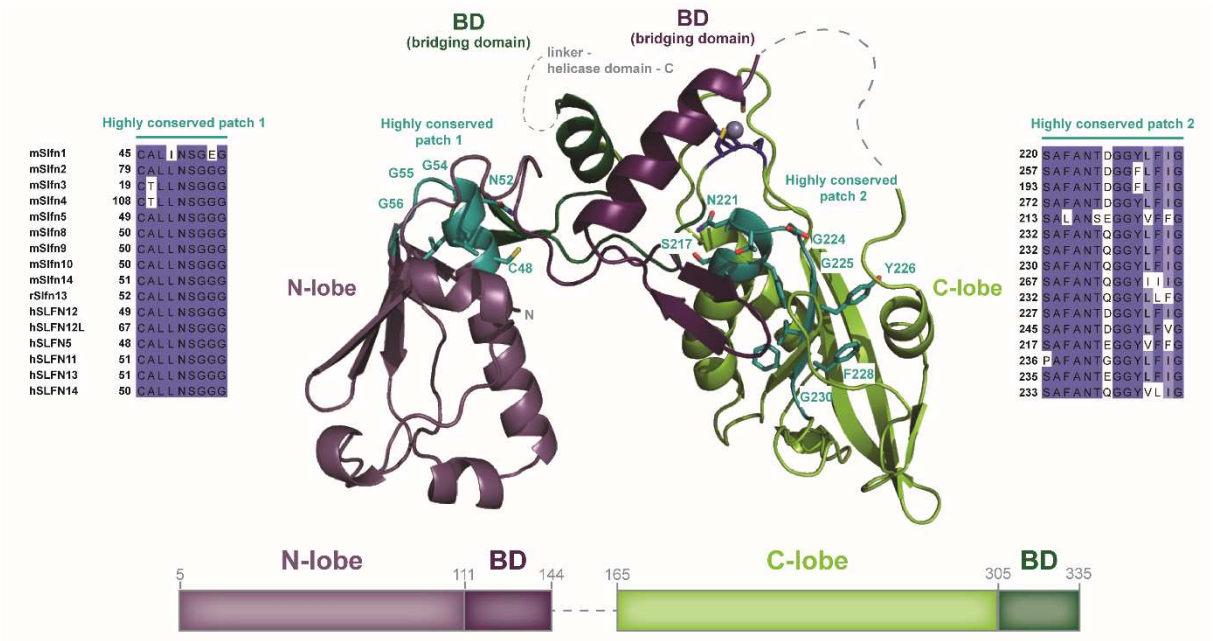



**Figure 5.1:** Overall structure of the slfn core domain represented by SLFN5[1-336] crystal structure and multiple sequence alignment of highly conserved patches. SLFN5[1-336] is colored as in 4.1.7 and the conserved stretches are depicted as sticks and colored dark turquois.

Within the schlafen core domain, several highly conserved sequence motifs emerge among them CALLNSGGG and SAFANTxGGYhFhG, which are located in the N-lobe and C-lobe, respectively (Figure 5.1) (x and h denote either D, E or Q, and hydrophobic residues, respectively). However, for

both patches no involvement in enzymatic function could be identified, suggesting a rather structural role. Interestingly, the highly conserved motif CDALL recurs in the following linker domain, as well. The repetition might be a consequence of multiple tandem duplication events and lineage specific inversions that were described for the *slfn* clusters of mouse, human and rat genomes[114]. The C-terminal lobe harbors furthermore the highly conserved zinc finger motif that consists in every schlafen of one histidine residue and three cysteine residues. A detailed multiple sequence alignment of all murine and human as well as rSlfn13 can be found in the appendix section (Figure 7.1).

Besides the high structural similarities of the slfn core domains, some secondary structure elements differ between N-terminal SLFN5$^{1-336}$ and rSlfn13$^{14-353}$.

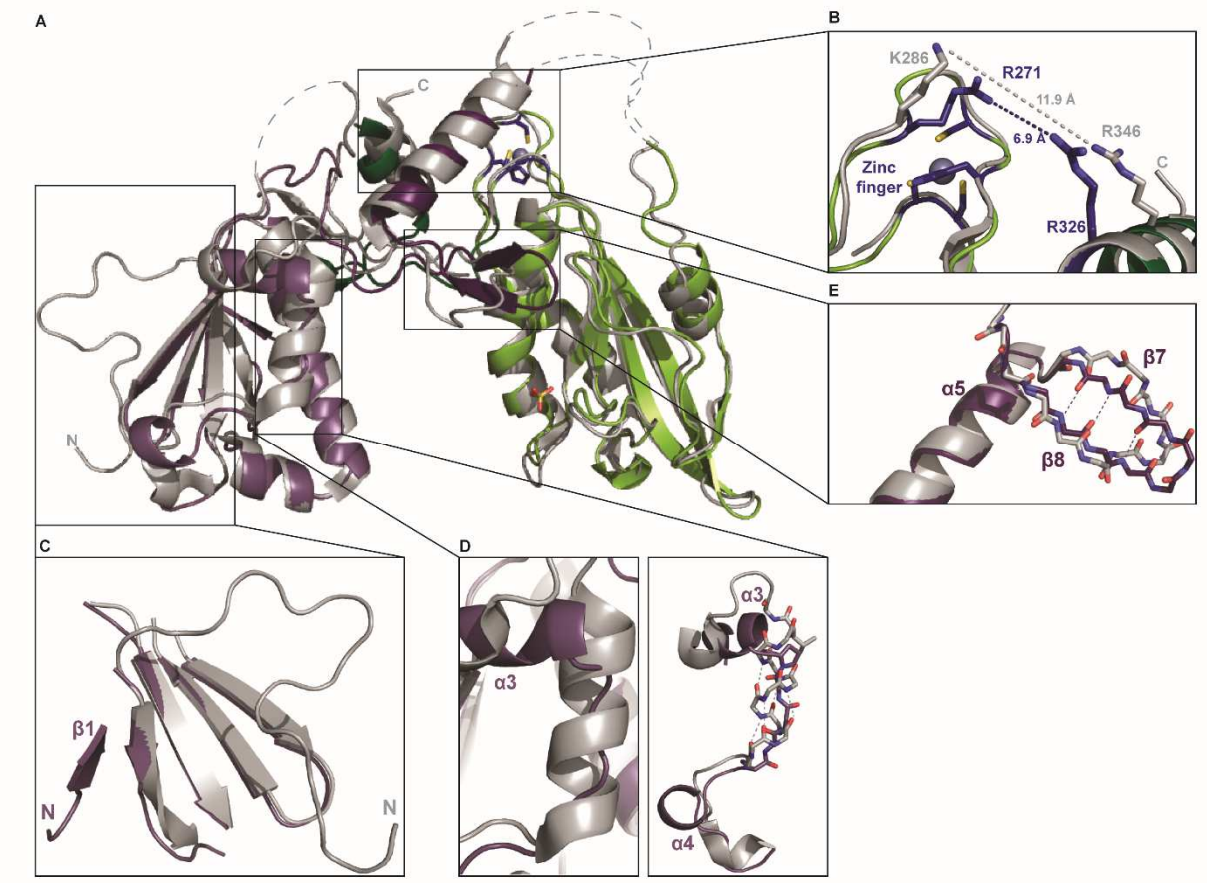



**Figure 5.2:** Structural comparison of SLFN5$^{1-336}$ and rSlfn13$^{14-353}$. **A:** Superposition of SLFN5$^{1-336}$ (colored as in 4.1.7) and rSlfn13$^{14-353}$ (PDB: 5YD0; colored in grey) crystal structure. **B:** Close-up view of the DNA-binding region of SLFN5$^{1-336}$ superpositioned with rSlfn13$^{14-353}$. SLFN5$^{1-336}$ and Slfn13 residues involved in DNA binding are presented as sticks and are colored in blue and grey, respectively. **C:** Close-up view of the extra N-terminal β-sheet of SLFN5$^{1-336}$ (colored in purple) compared to the artificial N-terminus of rSlfn13$^{14-353}$ (colored in grey). **D:** Close-up view of the extra α-helix of rSlfn13$^{14-353}$ (grey) compared to SLFN5$^{1-336}$ (purple), presented in cartoon mode (left) and as sticks (right). Hydrogen bonds are indicated as dashed lines. **E:** Close-up view of the extra two β-sheets in the first bridging domain of SLFN5$^{1-336}$ (dark purple) compared to rSlfn13$^{14-353}$ (grey). The residues are presented as sticks and the hydrogen bonds are indicated as dashed lines.

In rSlfn13$^{14-353}$ a pseudo symmetry between the N- and C-lobe with four helices and four β-sheets each was proposed[126], whereas in SLFN5$^{1-336}$ no such pseudo-symmetry was found. Here each lobe consists of four helices and five β-sheets, however, the sequential order between the secondary structure elements of each lobe differ. The pseudo-dimers of rSlfn13$^{14-353}$ that form the tunnel of the horseshoe are suggested to clamp tRNA and position it for endonucleolytic cleavage[126]. In turn, SLFN5$^{1-336}$ interacts with the DNA by residues located at the zinc finger region, since two arginines (R271 and R326) were DNA binding deficient upon mutation to glutamate. rSlfn13 depicts also two positively charged residues in this region, namely R346 and K286 (Figure 5.2. B), which might be involved in nucleic acid binding, as well. It would be of great interest to test R346 and K286 of rSlfn13 for nucleic acid binding.

Next, in the SLFN5$^{1-336}$ crystal structure the whole N-terminus is resolved, displaying an additional β- sheet (β1) that interacts with the β2-sheet (Figure 5.2. C). This sheet could not be displayed in rSlfn13$^{14-353}$, because the rSlfn13$^{14-353}$ construct is carrying an additional artificial N-terminus. Supposedly, Slfn13 would also depict this β-sheet, because the N-terminal residues forming the β-sheet are partly conserved between SLFN5 and rSlfn13. Consequently, it remains speculative, if the two lobes are really pseudo symmetric, since rSlfn13 harbors an additional N-terminal β-sheet. Furthermore, SLFN5$^{1-336}$ is missing a helix between α3 and α4 that is present in rSlfn13$^{14-353}$ (Figure 5.2. D). Hydrophobic residues of the additional helix contribute to the conserved hydrophobic patch. It also harbors negatively charged residues pointing into the tunnel, e.g. E84. The residue, however, did not show an effect in endonuclease activity, when mutating it to alanine[126]. This suggests a rather structural than an enzymatic role of this additional helix in rSlfn13. Moreover, in SLFN5$^{1-336}$ the arrangement of β7 and β8 is more structured forming hydrogen-bonds with each other, whereas in rSlfn13$^{14-353}$ the β-sheets are not formed (Figure 5.2. E). Taken together, the overall structure and the zinc finger motif of SLFN5$^{1-336}$ are similar to rSlfn13$^{14-353}$, however there are significant differences in some secondary structure elements.

### 5.1.2. The slfn core is a novel fold

The DALI server[202], an algorithm that searches for published structures with similar folds, was applied for the SLFN5$^{1-336}$ crystal structure. Some bacterial proteins were found to be structurally related, like the N- terminal domain of a putative ATP-dependent DNA helicase RecG-related protein from *Nitrosomas europaea* (PDB: 2KYY) or DIP2311 protein from *Corynebacterium diphtheria* (PDB: 3LMM). Yet the function of both proteins are unknown. By aligning those two structures, just one lobe of SLFN5$^{1-336}$ fits. No protein has been identified so far, that aligns with both lobes simultaneously.

The ATP-dependent DNA helicase RecG-related protein domain from *Nitrosomas europaea* (PDB: 2KYY) aligns very well with the C- lobe region of SLFN5$^{1-336}$ (Figure 5.3. A). Remarkably, even the sequential order of the secondary structure elements overlap, except for two missing β-sheets in SLFN5$^{1-336}$ (SLFN5$^{C\text{-lobe}}$: $^{N}$**α- β- α- β- α- β- β**$^{C}$; RecG-related: $^{N}$**α- β- α- β-** (β- β-) **α- β- β**$^{C}$). This bacterial domain is homologous to the C- terminal section of RecG, which is an ATP-dependent helicase playing important roles in the processing of stalled replication forks[203].

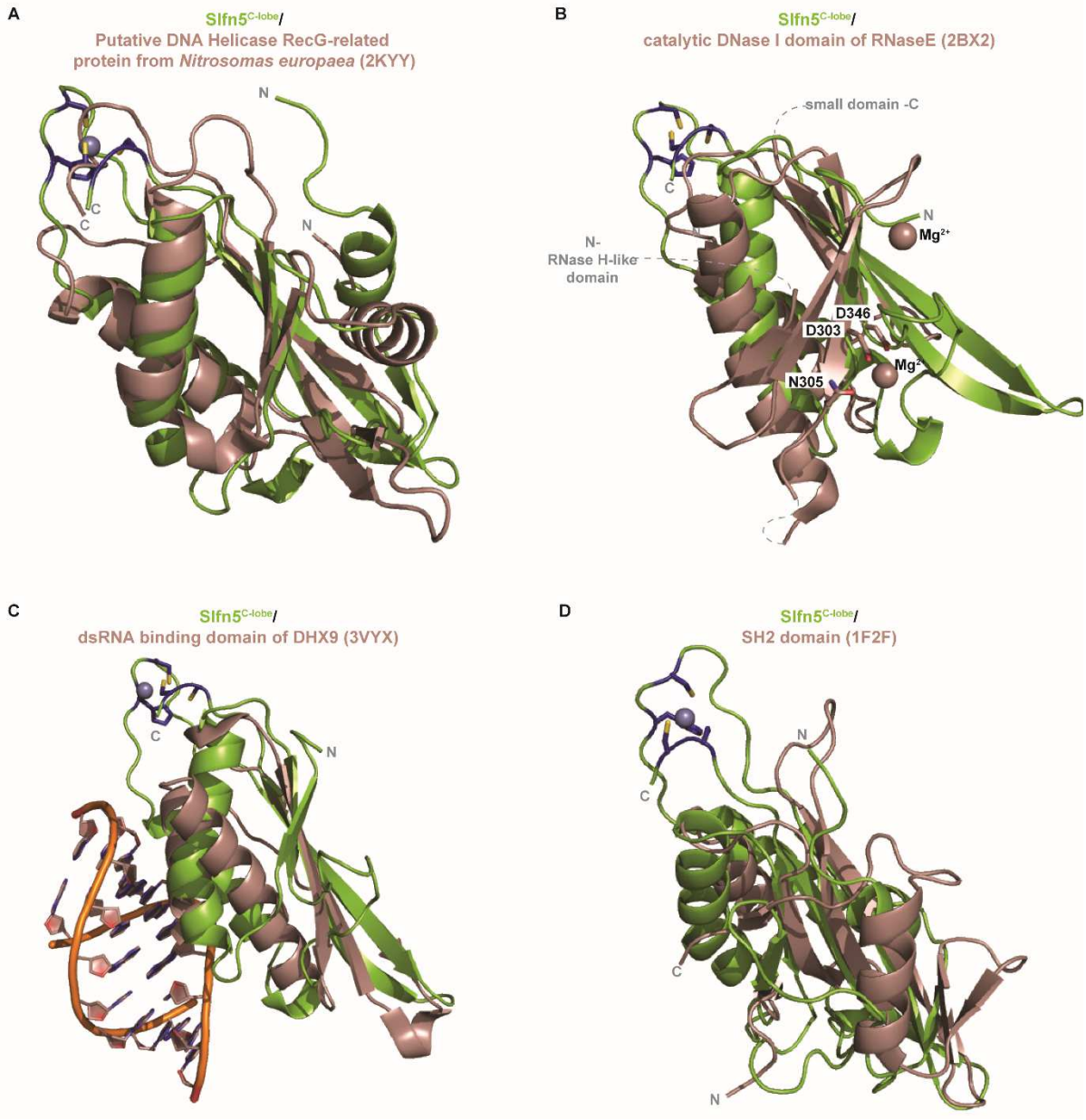



**Figure 5.3:** Superposition of SLFN5$^{C\text{-}lobe}$ with a (A) putative DNA helicase RecG-related protein from *Nitrosomas europaea* (PDB: 2KYY), (B) DNase I domain of RNase E (PDB: 2BX2)[204], (C) dsRNA binding domain of DHX9 (PDB: 3VYX)[205] and (D) SH2 domain (PDB: 1F2F)[206]. SLFN5$^{C\text{-}lobe}$ is colored in green and the other domains in brown.

Besides the proteins found in the DALI search, more proteins partly aligning with SLFN5$^{1\text{-}336}$ were identified. Among them are DNase I domain of RNase E (2BX2)[204], that aligns very well with the C-lobe of SLFN5$^{1\text{-}336}$ (Figure 5.3. B). They depict the same orientation and a similar order of the secondary structure elements, except for an additional N-terminal helix in SLFN5 and two N-terminal sheets for RNase E (SLFN5$^{C\text{-}lobe}$: $^{N}$α- **β- α- β- α- β- β**$^{C}$; RNase E: $^{N}$ (β- β-) **β- α- β- α- β- β**$^{C}$). RNase E is a bacterial endonuclease that plays essential roles in RNA processing and degradation including the maturation of 16S and 5S rRNA[207]. However, the active site residues of RNase E consisting of the conserved D303, N305 and D346[204] point rather to the outer rim of the fold, whereas the active site residues of SLFN5$^{1\text{-}336}$

point to the tunnel of the horseshoe. D303 and D346 of RNaseE are involved in coordination of a magnesium ion. The proposed catalytic mechanism of RNAse E is a nucleophilic attack of the scissile phosphate by a hydroxyl group generated from a water molecule interacting with the magnesium. The two aspartates might function as general bases, thereby activating the attacking water[204]. Since RNase E shows a 3' to 5' nuclease directionality[208] similar to SLFN5$^{1-336}$ some parallels of the working mechanisms might be assumed.

Moreover, similarities of the C-lobe of SLFN5$^{1-336}$ were found with the dsRNA binding domain of DExH-box helicase 9 (DHX9)[205]. Although, one can align the three, identically oriented central β-sheets as well as the two helices, the latter are oriented with different angles (Figure 5.3. C). Also the sequential order of the secondary structure elements differ (SLFN5$^{C\text{-}lobe}$: $^{N}$β- **α- β- α- β- β**$^{C}$; DHX9: $^{N}$**α- β- β- β-α**$^{C}$). DHX9 is an ATP-dependent nucleic acid helicase that unwinds both, DNA and RNA in a 3' to 5' directionality[209]. DHX9 plays important roles in many cellular processes, such as DNA replication, transcriptional activation[210], post-transcriptional gene regulation[211] and RNA-mediated gene silencing[205]. Interestingly, a direct interaction of DHX9 and SLFN11 was described[107,132].

SLFN5 was described to directly interact with STAT-1 and is predicted to form a negative-feedback regulatory loop for transcriptional regulation[61]. STATs normally interact with each other via their Src homology 2 (SH2) domain[212]. Thus, the C-lobe of SLFN5$^{1-336}$ was compared to a SH2 domain and indeed the fold showed similarities (Figure 5.3. D). The central β-sheets of the two proteins depict the same orientation and aligned well to each other. However, SLFN5$^{1-336}$ has an additional helix and the two common helices are differently oriented. Moreover, the order of the secondary structure elements differ (SLFN5$^{C\text{-}lobe}$: $^{N}$**β- α- β- α- β- β**$^{C}$; SH2: $^{N}$**β- α- β- β- β- (β- β-) α**$^{C}$). Since the C-lobe of SLFN5$^{1-336}$ resembles a SH2 domain and STATs interact with each other via their SH2 domains, it is tempting to speculate that SLFN5 and STAT interact via the C-lobe and SH2 domain, respectively, thus regulating transcription.

Taken together, the slfn core domain possesses a novel fold, since no protein was found that superposes with both lobes simultaneously, even though some protein domains aligned with one lobe, mainly with the C-lobe.

### 5.1.3. The slfn core domain is no ATPase

Based on sequence similarity, the N-terminal schlafen domain has been predicted to be an ATPase domain[141]. Within the schlafen core one highly conserved stretch of 80-100 amino acids harbors a conserved domain signature, which is predicted bind ATP (NCBI-CDD: COG2865)[30]. Furthermore, two conserved aspartate residues D248 and D249 in SLFN14 were predicted to be part of the Walker B motif[136]. In ATPases essential sequence motifs are Walker A and Walker B motifs, which are responsible for interaction and hydrolysis of the nucleoside triphosphate as well as the coordination of the metal ion, respectively. The Walker A consensus sequence consists of GxxxxGK(T/S) and the Walker B motif of hhhhD(D/E)[213], where X and h denote any amino acid and hydrophobic amino acids, respectively. A Walker A motif is characterized by a glycine-rich loop, preceded by a beta sheet and followed by an

alpha helix[213]. The predicted ATPase region within the schlafen proteins show sequence similarities with a Walker A motif, because they contain two conserved glycines, then any four residues followed by another glycine (Figure 5.4. A). However, the lysine that interacts with the phosphate of the nucleoside triphosphate and a hydrophilic residue (either T or S) that ligates the metal ion[213] are either completely missing in some schlafen members or in others they appear to be shifted by three residues. The Walker B motif, consisting of any four hydrophobic residues followed by two negatively charged residues (D and D or E)[213], seems to be inserted into the Walker A motif. However, when monitoring those residues in the SLFN5$^{1-336}$ crystal structure (Figure 5.4. B) and comparing them to the active sites of other ATPases, they differ structurally from each other (Figure 5.4. C-E).

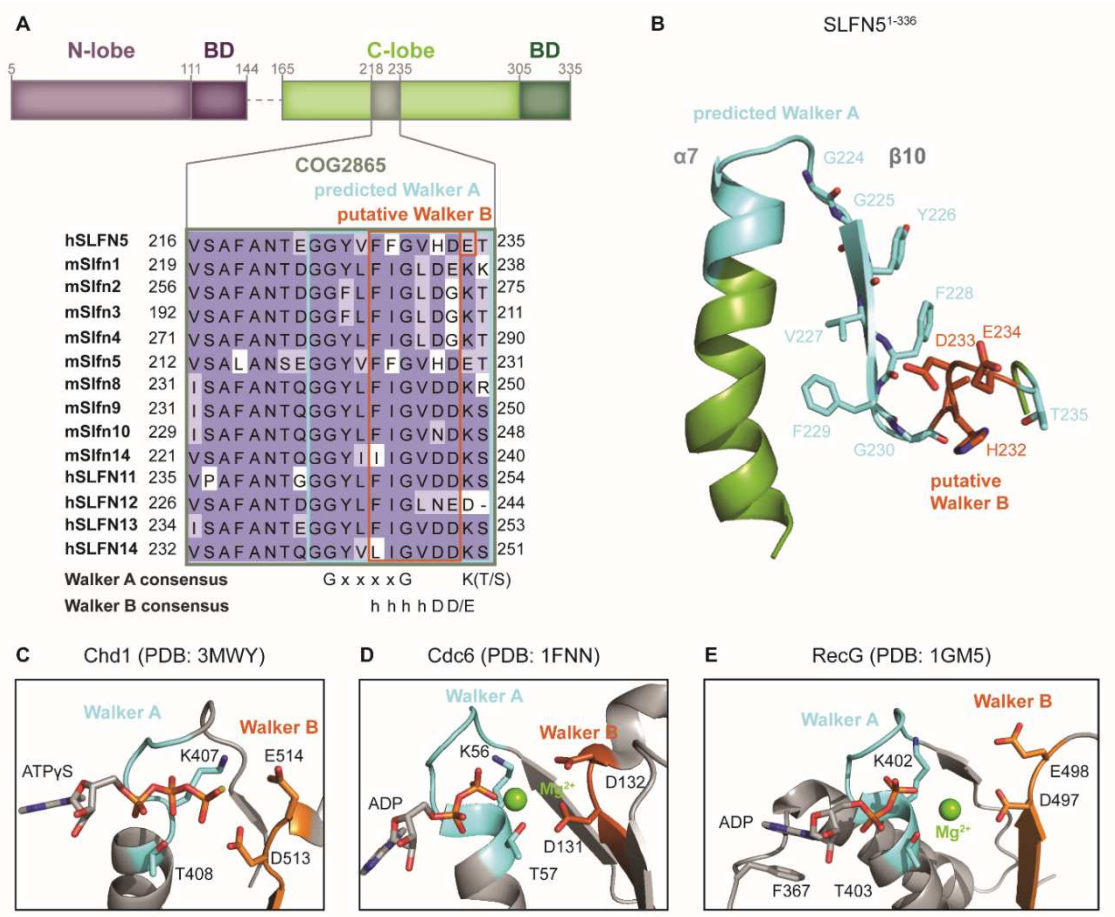



**Figure 5.4:** Structural analysis of the predicted ATPase domain. **A:** Schematic overview of the domain organization of SLFN5$^{1-336}$ with the predicted COG2865 domain including the putative Walker A/Walker B motif sequences and multiple sequence alignment of all murine and human schlafen members for this region. X and h denote any and hydrophobic amino acids, respectively. **B:** Structural framework of the predicted Walker A/Walker B motifs. The domains are colored as in 4.1.7 and the predicted Walker A and B motifs are colored in aquamarine and orange, respectively. **C-E:** Comparison of the ATP-binding sites containing Walker A and Walker B motifs with other ATPases: Chd1 (PDB: 3MWY)[214] (C), Cdc6 (PDB: 1FNN)[215] (D) and RecG (1GM5)[203] (E). ATP, ADPγS and important side chains are presented as sticks.

In SLFN5$^{1-336}$ the putative Walker A and B motifs are located in series, but in functional ATPases they are rather facing each other. Moreover, the positively charged residue of the Walker A motif that would interact with the phosphate of the ATP, is missing in SLFN5$^{1-336}$ (Figure 5.4. B). Furthermore, the secondary structure elements are in a reverse order than in other compared to described ATPases. The glycine-rich loop is preceded by a helix and followed by a β-sheet in SLFN5$^{1-336}$. The crystal structure of already described ATPases look different from the N-terminal schlafen fold, e.g. Chd1 in an ATP-bound state[214] (Figure 5.4. C), Cdc6[215] in an ADP bound state (Figure 5.4. D) and RecG in an ADP bound state[203] (Figure 5.4. E). For RecG the ATP binding and hydrolysis is a prerequisite for structural rearrangements of the DNA binding domains and sliding on the DNA for fork reversal during DNA replication[203]. The ATP binding domain of RecG is located upstream of the domain, which is similar to the C-lobe of SLFN5$^{1-336}$. It was confirmed experimentally, that neither SLFN5$^{1-336}$ nor Slfn2$^{FL}$ (data not shown) bind or hydrolyze ATP regardless if a nucleic acid substrate was present or absent.

Structural studies of the slfn core domain show that the slfn core resembles no ATPase like fold, although sequence based predictions indicated ATPase activity. Further, negative ATP- binding and hydrolysis assays confirmed experimentally that the slfn core domain functions in an ATP-independent manner, thus concluding that the slfn core domain is no ATPase.

### 5.1.4. The nucleic acid binding properties of the schlafen core domain

One of the highly conserved sequence motifs of the slfn core domain is the zinc finger motif. Generally, a zinc finger motif consist of approximately 30 amino acids forming an elongated loop, which is stabilized by a single $Zn^{2+}$ ion. Usually four residues coordinate the zinc ion and several hydrophobic side chains further stabilize the core of fold. A single zinc finger binds the nucleic acid substrate typically weak, thus, multiple zinc fingers considerably enhance binding by interacting simultaneously with the nucleic acid substrate[216]. The first zinc finger protein identified was the transcription factor TFIIIA of *Xenopus*, whose zinc fingers consist of two cysteine and two histidine residues (CCHH)[217]. Besides the classical CCHH zinc finger also CCCC or CHCC zinc finger motifs were discovered[218]. All schlafen members contain one zinc finger motif consisting of CHCC. Since many zinc finger proteins are interacting with DNA or RNA, the slfn core region of SLFN5$^{1-336}$ and Slfn2$^{FL}$ were analyzed concerning their nucleic acid binding capabilities.

SLFN5$^{1-336}$ binds preferentially double-stranded DNA with a determined $K_d$ of 2.5 µM for a 60- mer DNA, whereas binding to a single-stranded DNA of the same length resulted in a decreased affinity ($K_d$: 7.8 µM). Furthermore, SLFN5$^{1-336}$ shows a bad affinity for short single-stranded DNA and RNA substrates (5 to 30- mer), whereas longer single-stranded substrates significantly increase the affinity. Moreover, SLFN5$^{1-336}$ binds single-stranded RNA and stem loop RNAs, however, no perfectly aligned RNA duplexes, indicating that single-stranded stretches are necessary for the interaction. The binding affinities of SLFN5$^{1-336}$ to DNAs and RNAs were determined to be in a µM range, which is comparably bad to typical affinites determined for other zinc finger proteins. For example tristetraprolin, which consists as well of CHCC zinc fingers, binds AU-rich RNAs with an affinity of $K_d$< 20 nM[219].

Slfn2 shows a 10 times higher affinity to both, single-stranded 60- mer poly (dT) DNA ($K_d$: 700 nM) and a double-stranded DNA duplex ($K_d$: 800 nM) compared to SLFN5$^{1-336}$. Furthermore, binding of Slfn2$^{FL}$ to single-stranded RNA was considerably increased compared to SLFN5$^{1-336}$ ($K_d$: 400 nM), as well. Slfn2$^{FL}$ binds both, DNA and RNA, substrates well. However, it seems to have a preference for RNA substrates, indicated by fluorescence anisotropy experiments and by co-purification of 5S rRNA with Slfn2$^{FL}$. Illumina deep sequencing of co-purified nucleic acid substrates of Slfn2$^{FL}$, identified besides 5S rRNA also minor hits for tRNAs, 4.5S rRNA, 16S rRNA and 23S rRNA, thus indicating that Slfn2 prefers RNAs containing stem loop structures.

Since Slfn2$^{FL}$ is the physiological protein and SLFN5$^{1-336}$ a C-terminally truncated version, in SLFN5$^{1-336}$ some elements, which are located downstream of the slfn core and are potentially involved in nucleic acid binding, might be missing. This might lead to decreased nucleic acid binding properties compared to Slfn2. Indeed, full length SLFN5$^{FL}$ showed increased DNA binding properties compared to SLFN5$^{1-336}$.

To gain insights into the DNA binding mode, an electrostatic surface potential of SLFN5$^{1-336}$ was calculated, which revealed several positively charged patches. One patch that is in close proximity to the zinc finger and harbors two prominent arginines. Upon mutation of R271 and R326 to a glutamate each, they abolished DNA binding of SLFN5$^{1-336}$. This indicates that the zinc finger region substantially contributes to nucleic acid binding of SLFN5$^{1-336}$.

Slfn2$^{FL}$ also contains two positively charged residues in a similar position, namely K311 and K370. At least for K311 no decrease in affinity was observed in EMSAs for stem loop RNA as well as single- and double-stranded DNA (data not shown). The electrostatic surface potential presentation of Slfn2$^{FL}$ reveals several positively charged patches. Some prominent positively charged residues emerge at the origin of the tunnel. So far single point mutations of selected residues were not enough to efficiently diminish RNA binding. Four conserved residues have been tested, but only the co-purification of RNA was diminished in three out of four and DNA binding properties by EMSAs were not affected. Since Slfn2$^{FL}$ has a considerably higher affinity to the nucleic acid substrates than SLFN5$^{1-336}$, probably the mutation of one single residue might not be sufficient. This indicates that in Slfn2$^{FL}$ several interaction sites are simultaneously participating in nucleic acid binding.

Since Slfn2$^{FL}$ binds both, DNA and RNA including 5S rRNA, it would be interesting to elucidate the binding mode of Slfn2$^{FL}$ to the individual substrates. The first identified zinc finger protein was the transcription factor TFIIIA, which, interestingly, binds besides the 5S rRNA gene also the 5S rRNA itself[220]. TFIIIA contains nine zinc fingers, in which fingers 4 to 7 are required to bind mainly the central loops of 5S rRNA[221,222]. Whereas zinc fingers 1 to 3 wrap smoothly round the major groove of the DNA[223,224]. This means that different subsets of zinc fingers are responsible for RNA or DNA binding implying two different modes of binding. However, Slfn2$^{FL}$, which also binds both substrates, contains only a single zinc finger. This implies that the binding to RNA or DNA must follow different rules than the TFIIIA binding to both substrates. A co-crystal structure of either Slfn2$^{FL}$ or SLFN5$^{1-336}$ with either RNA or DNA would be of great interest to fully elucidate the DNA/RNA binding mode of the slfn core domain.

5.1.5. The slfn core domain shows 3'-5' exonuclease activity on single-stranded DNA

Nucleases are phosphodiesterases that cleave one of the two bridging P-O bonds in nucleic acid substrates and they play important roles in various cellular processes, such as DNA replication, recombination, repair, RNA maturation, processing and interference. Nucleases can either function in a endo- or exonuclease manner, cut DNA or RNA or function as topoisomerases, recombinases or RNA splicing enzymes [reviewed in 225]. Generally, the reaction mechanism starts with a nucleophilic attack. In enzymes cleaving DNA often serines, tyrosines and histidines have been observed to serve as a nucleophile. A water molecule can serve as a nucleophile, as well, but in this case a general base is required for deprotonation and efficient cleavage. A covalent DNA phosphoryl-protein intermediate is then formed, which is resolved by hydrolysis in a two-step cleavage reaction[reviewed in 225]. Most nucleases with 3'-5' exonuclease activity, which include for example the biggest superfamily DnaQ-like 3'-5' exonuclease, follow the two-metal-ion dependent mechanism. DnaQ-like proteins have an absolutely conserved sequence motif consisting of DEDD. The two metal ions are coordinated between the first D of the motif and the oxygen of the scissile phosphate. Each of the remaining three carboxylate groups coordinates one of the two-metal-ions. The nucleophilic water is in those enzymes coordinated and activated by conserved tyrosine or histidine residues[225].

Biochemical studies discovered exonuclease activity towards single-stranded DNA by $SLFN5^{1-336}$ and $Slfn2^{FL}$, however, they do not digest any tested RNA nor double-stranded DNA. The activity is dependent on metal ions, especially magnesium, but independent of ATP. Furthermore, $SLFN5^{1-336}$ and $Slfn2^{FL}$ require an accessible 3' DNA terminus for exonuclease activity, thus indicating 3'-5' directionality.

Crystallographic investigations of $SLFN5^{1-336}$ identified three important sites within the C-lobe, which are responsible for the catalysis. Since sulfates can mimic the phosphates of nucleic acids[175-177], the first site tested contained the residues coordinating the sulfate (Figure 5.5. A). Within this region, four conserved negatively charged residues exist: EEDE. Indeed, most mutations of residues E191, E196, D233 or E234 in $SLFN5^{1-336}$ lead to a nucleolytic inactive protein. E191 was inactive regardless, if mutating it to A or Q, whereas E196 was inactive just for A and not Q. This suggests that in E196Q the hydroxyl group of glutamine might still be sufficient for the activity. Furthermore, E191 seems to be more essential for catalysis than E196, although E196 is highly conserved throughout the schlafen family, and E191 is not present in murine Slfn1-5. Although D233 mutated to alanine was still active, D233 mutated to asparagine was inactive. Presumably, the functional role of D233A could be resumed by E234, whereas D233N might be too bulky to allow for structural rearrangement. Indeed, the double mutant D233A E234A was nucleolytic inactive. E234A alone was inactive, as well, indicating a more important role for E234 in catalysis than D233. Interestingly, two successively arranged, negatively charged residues are conserved within the schlafen protein family. The central located arginine 211 is mainly responsible for the interactions with the sulfate and was catalytically inactive upon mutation to alanine. Two possible roles of R211 could either be a functional task or the positioning of the nucleic acid, since R211 is substantially responsible for coordinating the sulfate.

Besides the active site, two other important catalytic sites have been identified. First, R326, which was shown to be DNA-binding deficient upon mutation to alanine, was also exonuclease inactive. This

suggests a tight interplay of the DNA binding site including the zinc finger motif and the active site. Second, the mutations of the highly conserved residues Y183 and E244 to alanine abolished the activity completely. They are located in close proximity to the active site at the tip of the horseshoe. Tyrosines play important roles in nuclease activity. They can either function as a nucleophile[226,227] or perform base stacking with the DNA substrate and coordinate the DNA for digestion[179]. Tyrosine 183 together with glutamate 244 could function as an entry or sensing site, because they lie 11.5 Å apart from each other, which corresponds to the width of a single-stranded DNA. Both residues are highly conserved within the schlafen family (Figure 5.5. F).

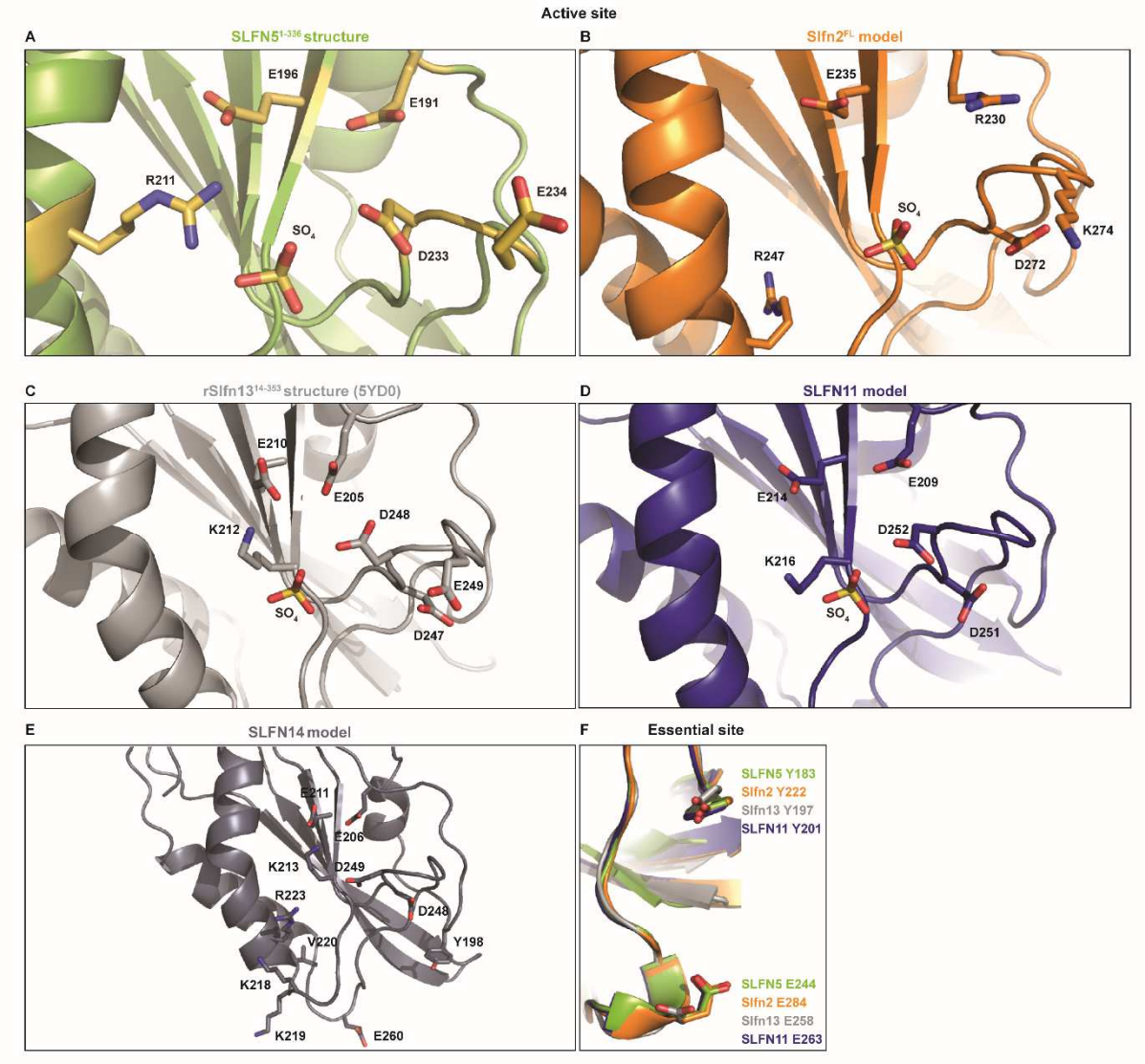



**Figure 5.5:** Structural framework of the active sites located in the C-lobe of various schlafen members. **A:** SLFN5$^{1-336}$ crystal structure. **B:** Slfn2 model based on SLFN5$^{1-336}$ crystal structure. **C:** rSlfn13$^{14-353}$ crystal structure. **D:** SLFN11 model based on SLFN5$^{1-336}$ crystal structure. **E:** SLFN14 model based on rSlfn13 crystal structure. **F:** Superposition of the second extended active site residues of SLFN5$^{1-336}$, Slfn2$^{FL}$ model, Slfn13$^{14-353}$ and SLFN11 model.

Slfn2$^{FL}$ shows exonuclease activity on single-stranded DNA in a 3'-5' directionality similar to SLFN5$^{1-336}$. The active site of Slfn2$^{FL}$ looks however slightly different (Figure 5.5. B). E235, which is homologous to SLFN5 E196 is present, indicating an important role in catalysis. An equivalent residue to SLFN5 E191 is missing. Furthermore, Slfn2 lacks a homologous residue of SLFN5 R211, but R247 is located in close proximity and could resume the task of SLFN5 R211, which might be the coordination of the nucleic acid substrate. Thus, the generation of Slfn2$^{FL}$ constructs carrying E235A or R247A would be of great interest. Furthermore, D272 of Slfn2 is a homologous residue to D233 of SLFN5, indicating also essential roles for D272 in catalysis. Although, the second negatively charged residue is missing in Slfn2, SLFN5$^{1-336}$ and Slfn2$^{FL}$ resemble the same activity indicated by a similar cleavage pattern.

rSlfn13 has nuclease activity, however, the solely substrates are tRNAs and rRNAs. rSlfn13 performs an endocut 11 bases distant of the 3' end in the acceptor stem of the tRNA. The active site consists of the residues E205, E210 and D248, which are homologous to residues to E191, E196 and D233 of SLFN5, respectively. Mutation of E205 and E210 to alanine completely abolishes the endoribonuclease activity, whereas mutation of D248 only reduces the activity[126]. Slfn13 has no homologous residues to R211 of SLFN5, however K212 is positioned similarly and might be involved in the correct orientation of the nucleic acid in the site (Figure 5.5. C). Although the active sites of Slfn13$^{14-353}$ and SLFN5$^{1-336}$ coincide, the specificity for the nucleic acid substrates differs. It remains elusive how the different schlafen members discriminate between their respective substrates despite the similarities of the fold. Co-crystal structures of a substrate bound slfn core domain would provide important insights into the mode of action.

Human SLFN14 was described to be involved in ribosome degradation and cleavage of ribosomal RNA[136]. The two negatively charged residues D248 and D249, which are homologous to D233 and E234 of SLFN5, showed a decreased activity upon mutation[136]. Furthermore, mutations in the *slfn14* gene in patients suffering from thrombocytopenia lead to a substitution of K218E, K219E, V220D and R223W[137-139]. Those residues are located in close proximity to the active site residues, which are also preserved in SLFN14 (Figure 5.5. E). K218 and K219 point to the inner tunnel of the horseshoe. In Slfn13 the residues for RNA binding were proposed to be located in this tunnel. Thus, upon mutation of K218 and K219 the RNA binding properties might be disturbed. Furthermore, by insertion of a charged amino acid instead of a hydrophobic valine at position 220, the protein might be misfolded and aggregate, since V220 is located in an otherwise hydrophobic region. Indeed, aggregation was observed experimentally[139]. Arginine 233, which is the homologous residue to R247 of Slfn2 and points into the active site, might position the nucleic acid for cleavage. By substitution of the arginine to a rather bulky tryptophan residue, positioning of the nucleic acid might be disturbed, thereby resulting in malfunction of SLFN14. The existence of the patient mutations, confirmed that this site functions as active site and illustrates the importance of this region within the slfn core domain.

Based on the crystal structure of SLFN5$^{1-336}$, a model of SLFN11 was generated. In this model, the active site is highly similar to the rSlfn13 and SLFN5 active sites. E209 and E214, which are homologous to E191 and E196 of SLFN5 and E205 and E210 of rSlfn13 and which confer essential roles for the nuclease activity, are present as well. Furthermore, it contains the two negatively charged residues

D251 and D252 and one positively charged residue K216 that might position the nucleic acid. Due to the high similarity of the active sites and the high degree of conservation, nuclease activity can be assumed for SLFN11. Yang et al, 2018, propose endonuclease activity for SLFN11, as well[126], thus explaining the antiviral effect of SLFN11. In their model, the putative tRNA degradation activity by SLFN11 may negatively influence the translation of newly synthesized viral proteins[126]. SLFN11 has not only antiviral functions, but is also involved at stalled replication forks[132]. SLFN11 induces an S-phase arrest upon treatment with PARP- or TOP inhibitors, leading to an irreversible and lethal replication inhibition, presumably by opening the chromatin and blocking the elongation, mainly by its C-terminal helicase domain[132]. Since SLFN11 binds to the nascent DNA ends together with RPA, SLFN11 might act by destabilizing the RPA-ss DNA complex[132]. It might be possible that the N-terminal domain of SLFN11, which has potential DNA exonuclease activity on single-stranded DNA, plays a role in the destabilization of the ssDNA-RPA complex by digesting DNA ends. Indeed, in SLFN11-deficient cells single-stranded DNA was persistent at DNA damage sites, whereas it was significantly degraded in SLFN11-deficient cells[133].

Since the nuclease active site is conserved within the schlafen protein family and the schlafen members have been described to play roles in T cell maturation[24], a potential regulation by exonuclease activities might be speculated. This might be possible during the V, D, J gene segment rearrangements, where the introduction and repair of double-strand breaks is an essential step[15]. In mammals, RAG1/2 recombinases initiate at first a single-strand nick and in a second step, the phosphodiester bond of the opposing strand is attacked. Randomly chosen coding segments are finally arranged and ligated by the non-homologous end joining (NHEJ) pathway [228]. The schlafen proteins could potentially regulate T cell maturation by participating in the various DNA-repair steps during V, D, J gene segment rearrangements.

Taken together, several schlafen members function as a nuclease, however, they differ in their preferred substrates. SLFN5$^{1-336}$ and Slfn2$^{FL}$ show 3'-5' exonuclease activity on single-stranded DNAs, whereas Slfn13 cuts preferentially tRNAs[126]. Independent studies determined the active site to be at the C-lobe region and to contain four partially conserved negatively charged residues. Since the active site residues are partially conserved and the fold is similar within the so far analyzed schlafen protein members, generally nuclease activity is proposed for the slfn core domain. A substrate bound co-crystal structure would be of great advantage to elucidate the nuclease working mechanism.

### 5.1.6. Slfn2 might be involved in ribosomopathies and lipid homeostasis

Rat Slfn13 and human SLFN14 are described to be associated with the ribosome and to degrade ribosomal RNA[126,136,139]. Although Slfn2$^{FL}$ did also show exonuclease activity, it did not degrade ribosomal RNA. However, tight binding between Slfn2$^{FL}$ and 5S rRNA is assumed due to co-purification of the RNA with Slfn2$^{FL}$. The interaction of Slfn2$^{FL}$ with the co-purified 5S rRNA can also be assumed due to SAXS experiments, since a Slfn2$^{FL}$ monomer can be fitted into the rather elongated *ab initio* model besides some other unknown molecule (Figure 5.6. A). A 5S rRNA could potentially fit into the

model together with Slfn2$^{FL}$. However, the arrangement and orientation of the molecules to each other remains speculative. On the other side, SAXS data of SLFN5$^{1-336}$, which is a monomer in solution and does not co-purify a substrate, suggest a rather globular model that fits one monomer of the crystal structure (Figure 5.6. B).

Interestingly, co-purification and RNA deep sequencing of a Slfn2$^{FL}$ RNA substrate identified the genes for all 5S rRNA operons except for *rrnA*. The only difference between *rrnA* and the other operons is a thymidine at position 92 instead of a cytosine (Figure 5.6. C). 5S rRNA transcribed from *rrnA* thus contains a U92-G98 wobble base pair, whereas in all other 5S rRNAs a C92-G98 Watson-Crick basepair is formed (Figure 5.6. D), indicating, that Slfn2$^{FL}$ could potentially interact with 5S rRNA at the stem of the molecule and a proper Watson-Crick basepair of C92 and G98 might be essential. However, EMSAs of Slfn2$^{FL}$ with a 5S rRNA stem revealed binding to both rRNA stems regardless if U92-G98 or C92-G98 were present (data not shown).

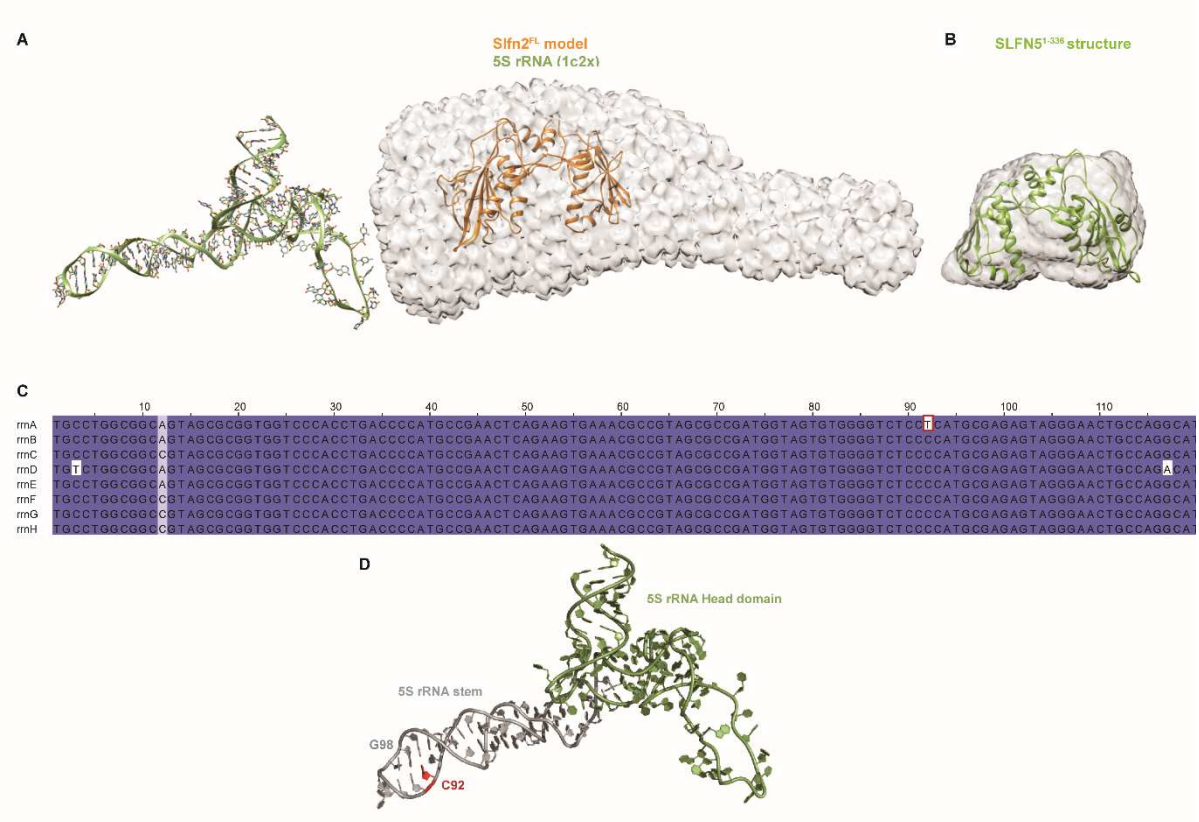



**Figure 5.6:** Analysis of the Slfn2$^{FL}$ and co-purified 5S rRNA. **A:** Into the SAXS derived *ab initio* model of Slfn2$^{FL}$ another molecule besides Slfn2$^{FL}$ could be fitted. A 5S rRNA would fit in the model by size, however, the arrangement of the molecules remains speculative. The Slfn2$^{FL}$ model is colored in orange and the 5S rRNA in green (PDB: 1C2X). **B:** As a control, the *ab initio* model of SLFN5$^{1-336}$ shows a rather globular form, into which one monomer of the SLFN5$^{1-336}$ fit. **C:** Multiple sequence alignment of all 5S rRNA operons. The *rrnA* 5S rRNA gene has a thymidine at position 92 instead of a cytosine. **D:** Crystal structure of a 5S rRNA of *E. coli* (PDB: 1C2X). C92 is located in the stem of 5S rRNAs and forms a basepair with guanidine 98.

Although the 5S rRNA is part of the ribosome, it also plays a crucial role in cellular stress induced ribosomopathies and in p53 homeostasis. Under non-stressful conditions, the 5S rRNA forms together with the 60 S ribosomal protein 5 (RPL5) and RPL11 the 5S ribonucleoprotein particle, which is an essential subcomplex of the large ribosomal subunit. Specifically, during ribosome biogenesis, RPL5 binds the 5S rRNA and the 5S rRNA/ RPL complex translocates to the nucleolus, where it interacts with RPL11 and is integrated into the ribosome[229]. If ribosome biogenesis is blocked due to cellular stress the free 5S rRNA/ RPL complex associates with mouse double minute 2 homolog (Mdm2)[230,231]. Mdm2 is an E3 ubiquitin ligase that inhibits the tumor suppressor p53 by ubiquitination-mediated proteosomal degradation[232]. Binding of Mdm2 by the 5S/ RPL complex thus prevents p53 degradation and induces growth arrest and apoptosis[233,234]. Interestingly, the p53 tumor suppressor is also involved in the apoptotic death of Slfn2-deficient T cells[88]. This suggests that Slfn2 might disturb the formation of the 5S/ RPL/ Mdm2 complex by interacting with 5S rRNA in Slfn2 proficient T cells. This would lead to p53 degradation by Mdm2 and survival of the cells. However, in Slfn2- deficient cells, the 5S/ RPL/ Mdm2 complex could be formed. Thus, p53 would not be degraded and would induce cell growth and apoptosis. The occurrence of apoptotic, Slfn2-deficient T cells resembles the immune-suppressed phenotype of the *elektra* mouse[86]. In *elektra* mice, the *slfn2* gene contains a point mutation that leads to an isoleucine to asparagine substitution at position 135 in the Slfn2 protein[86]. It could be shown that the point mutation lies in a conserved hydrophobic path and that the insertion of the hydrophilic asparagine causes a loss-of-function mutation due to misfolding and aggregation of the protein. Aggregation of Slfn2$^{I135N}$ could also be observed *in vitro* (data not shown).

Slfn2 cannot only be linked to the Mdm2-p53 axis by its direct interaction with 5S rRNA, but also by its potential direct interaction with phospholipase D3 (Pld3) and tripartite motif protein 25 (Trim25), whose interactions could be identified by affinity enrichment mass spectrometry. Increased levels of Pld3, which acts as a endonuclease[235], leads to elevated levels of Mdm2 and suppresses DNA-damage-induced increase in p53[236]. Furthermore, the Slfn2-interacting E3 ubiquitin ligase Trim25 has been described as a novel regulator for p53 and Mdm2[237]. Increased estrogen levels causes a higher abundance of p53 and Mdm2[238-240] and this is dependent on Trim25, which in turn is activated by estrogen[237]. Interestingly, one of the most enriched Slfn2 interactors is the sterol-synthetase (STS). STS, which triggers the conversion of sulfated steroid precursors to estrogens[181], could be regulated by Slfn2. The direct interactions of Slfn2 with STS and Trim25 suggest an indirect regulation of p53 by Slfn2 through influencing the sterol levels.

In general, several proteins involved in lipid homeostasis that are associated with the endoplasmic reticulum (ER), could be identified within the Slfn2 interactome, namely the fatty acid desaturase 3 (Fads3), extended synaptotagmin-1 (Esyt1) and glycosylphosphatidylinositol anchor attachment 1 protein (Gpaa1). This implies that Slfn2 could potentially regulate proteins involved in lipid metabolism and homeostasis. Furthermore, recent studies discovered chronic ER stress[89] and elevated levels of lipids[121] in *elektra* T cells. Indeed, T cell quiescence and lipid metabolism are tightly interconnected, because upon T cell activation, cells undergo massive reprogramming of their metabolism including the activation of lipid/sterol biogenesis pathways[241]. There are increasing evidences that Slfn2 regulates anabolic lipid synthesis pathways, which are required during the proliferation of quiescent to activated

T cells. The hydrophobic insertion (C116-W122) within Slfn2 that is not present in SLFN5 could serve as a potential interaction site with the ER membrane or with ER-associated proteins. However, this remains speculative and the direct interaction between Slfn2 and the potentially identified interactors have to be verified.

Taken together, the cell-proliferative functions of Slfn2 could be caused by its link to the Mdm2-p53 axis via a direct interaction with 5S rRNA. Furthermore, Slfn2 seems to interact with ER-associated proteins and is involved in lipid homeostasis. Thereby, Slfn2 might regulate T cell quiescence by influencing the metabolic changes of a T cell upon activation. The exact mode of actions, however, needs to be investigated in more detail.

### 5.2. The C-terminal domain shows all typical superfamily 1 DNA/RNA helicase motifs

Subgroup III schlafen protein, which are about 100 kDa in size, harbor a C-terminal region that is exclusively present in these “long” schlafen proteins. Experimentally, the helicase domain alone could neither be expressed from *E. coli* nor from *S. cerevisiae*, as it remained in insoluble fractions probably due to aggregation. Expression of the full length SLFN5 protein increased the solubility, but nevertheless, the sample tends to precipitate in solution and high salt concentrations had to be used throughout the purification. However, full length SLFN5 was shown to bind ATP even though ATP hydrolysis was not observed. Since the N-terminal slfn core domain alone does not bind ATP, the C-terminal domain is likely responsible for ATP binding. Indeed, the C-terminal domain of subgroup III schlafen proteins harbors all sequence motifs that are characteristic for ATP-dependent superfamily (SF) 1 DNA/RNA helicases[30,242,243].

Motif I, which is also called Walker A motif, is a characteristic motif in many nucleotide-binding proteins. Usually, it consists of a glycine-rich loop preceded by a beta sheet and followed by an alpha helix[213]. Residues of the Walker A motif typically interact with the alpha and beta phosphates of the nucleoside triphosphate [244]. The classical consensus sequence consists of GxxxxGK(T/S) (x denotes any amino acid). In subgroup III schlafen proteins the predicted Walker A motif is GLPGSGKT. It is conserved in all subgroup III schlafen, except for murine and human SLFN14, in which the first and the third glycine residues are missing. The lysine and serine or threonine are important residues for interaction with the phosphates of the nucleoside triphosphate and the ligation of the metal co-factor, respectively[244,245]. Indeed, a mutation within the Walker A motif of SLFN11, namely K506M, abolishes the ATPase activity of SLFN11, however without affecting its role in DNA damage response[133].

A further downstream-situated motif of the Walker A motif is motif II, which is also called Walker B motif. It is generally hydrophobic and forms a putative beta sheet followed by aspartate and glutamate residues[246]. The consensus sequence is hhhhDE, where h denotes any amino acids. The conserved aspartate co-ordinates the metal ion and the glutamate plays an essential role in ATP hydrolysis[247]. The predicted Walker B motif in subgroup III schlafen proteins is HIIIDEAQ. Walker B residues, especially the aspartate and glutamate, are indispensable for the activity for the enzymes. Indeed, mutation of D668Q in SLFN11 resulted in abolished ATPase activity, without affecting the DNA damage

response[133]. In another study E669 in SLFN11 was mutated to glutamine resulting in abolished ATPase activity and additionally in an abolished ability to inhibit DNA replication[132].

Besides the Walker A and Walker B motifs, additional other motifs are shown to be characteristic for SF 1 DNA/RNA helicases. The Q-motif and motifs IIIa and VI are also involved in ATP binding and hydrolysis, whereas motifs III and Va are involved in the coordination between the nucleic aicd and NTP binding sites. Furthermore, motifs Ia, Ic, IV, IVa, V and Vb are responsible for nucleic acid binding[200]. All sequence motifs could be assigned to conserved sequence stretches in the subgroup III schlafen proteins (Figure 7.1. in the appendix section).

The assignments of all characteristic SF1 sequence motifs, suggests ATP-dependent helicase activity. It could already be shown that SLFN11 might block DNA-replication at stalled replication forks by unwinding the DNA ahead of the replication forks, thereby causing a replication stop and cell death[132]. Mutations in the Walker B motif confirmed the ATP-dependency[132,133]. However, the exact molecular working mechanism remains unknown. Therefore, an established purification protocol and crystallization studies of full length subgroup III schlafen members would be of great benefit to elucidate the mode of action.

### 5.3. The schlafen proteins are ribosome associated

There is increasing evidences that the schlafen protein family is associated with ribosomes. Both, rat Slfn13 and human SLFN14 are functionally linked to ribosomes through the endonucleolytic cleave of 18S and 28S rRNA[126,136]. Furthermore, SLFN14 was recently shown to directly interact and co-localize with 5.8 S rRNA[139]. The present study shows that not only Slfn2, but also $SLFN5^{FL}$ co-purifies with 5S rRNA from bacterial cells and 5.8 S rRNA from insect cells, respectively. The 5.8 S rRNA is part of the big subunit of the ribosome[248]. During ribosome biogenesis the 5.8S, 18S and 28S rRNA are transcribed as one single 47S transcript, processed and modified in the nucleolus and afterwards transported to the nucleoplasm, where the individual subunits are assembled[249]. In this process several endo- and exonucleases are involved[249]. Since SLFN5 localizes to the nucleus[55], it is tempting to speculate, if SLFN5 might be involved in 5.8S rRNA processing. However, the direct link between the potential 5.8S rRNA interaction of SLFN5 and the suggested transcriptional regulation by directly interacting with STAT1 as well as the discussed negative-feedback regulatory loop in cancer cells[61] remains to be directly studied.

# 6. Materials and Methods

## 6.1. Materials

Chemicals were purchased from Sigma-Aldrich (Deisenhofen), Carl Roth (Karlsruhe), or Merck (Darmstadt). Generally, the purity grade *pro analysis* was chosen. Restriction enzymes were ordered at New England BioLabs (Frankfurt) or Thermo Scientific (St. Leon-Rot). Mini, Midi and Maxi kits for plasmid isolation were purchased from Machery&Nagel (Düren). For protein purification, Ni-NTA beads from GE Healthcare (Munich) or Machery&Nagel (Düren) were used.

### 6.1.1. Materials used in molecular biology

#### 6.1.1.1. Oligonucleotides

**Table 6.1:** DNA oligonucleotides used for molecular cloning and site-directed mutagenesis. The primers were purchased from Metabion (Planegg).

| Description | Sequence (5'-3' direction) | Restriction site/ Method |
|---|---|---|
| **pET21-SLFN5$^{1-336}$Y183F_for** | GAT AGA AAG CGG CTT CAG TTT CTG GAA AAA CTC AAC | Quick Change |
| **pET21-SLFN5$^{1-336}$Y183F_rev** | GTT GAG TTT TTC CAG AAA CTG AAG CCG CTT TCT ATC | Quick Change |
| **pET21-SLFN5$^{1-336}$Y183A_for** | GAT AGA AAG CGG CTT CAG GCT CTG GAA AAA CTC AAC | Quick Change |
| **pET21-SLFN5$^{1-336}$Y183A_rev** | GTT GAG TTT TTC CAG AGC CTG AAG CCG CTT TCT ATC | Quick Change |
| **pET21-SLFN5$^{1-336}$E191A_for** | CTC AAC CTT CCT GCG TCC ACA CAT GTT G | Quick Change |
| **pET21-SLFN5$^{1-336}$E191A_rev** | CAA CAT GTG TGG ACG CAG GAA GGT TGA G | Quick Change |
| **pET21-SLFN5$^{1-336}$E191Q_for** | GAA AAA CTC AAC CTT CCT CAG TCC ACA CAT GTT G | Quick Change |
| **pET21-SLFN5$^{1-336}$E191Q_rev** | CAA CAT GTG TGG ACT GAG GAA GGT TGA GTT TTT C | Quick Change |
| **pET21-SLFN5$^{1-336}$E196A_for** | GAG TCC ACA CAT GTT GCA TTT GTA ATG TTC TCG AC | Quick Change |
| **pET21-SLFN5$^{1-336}$E196A_rev** | GTC GAG AAC ATT ACA AAT GCA ACA TGT GTG GAC TC | Quick Change |
| **pET21-SLFN5$^{1-336}$E196Q_for** | GAG TCC ACA CAT GTT CAA TTT GTA ATG TTC TCG ACA G | Quick Change |
| **pET21-SLFN5$^{1-336}$E196Q_rev** | CTG TCG AGA ACA TTA CAA ATT GAA CAT GTG TGG ACT C | Quick Change |
| **pET21-SLFN5$^{1-336}$R211A_for** | CAC TGT GTT AAA GAC GCA CTT CCG AAG TGT G | Quick Change |
| **pET21-SLFN5$^{1-336}$R211A_rev** | CAC ACT TCG GAA GTG CGT CTT TAA CAC AGT G | Quick Change |
| **pET21-SLFN5$^{1-336}$D233A_for** | GGA TAT GTA TTT TTT GGT GTG CAT GCT GAG ACT TGT CAA GTG ATT G | Quick Change |
| **pET21-SLFN5$^{1-336}$D233A_rev** | CAA TCA CTT GAC AAG TCT CAG CAT GCA CAC CAA AAA ATA CAT ATC C | Quick Change |
| **pET21-SLFN5$^{1-336}$D233N_for** | GTA TTT TTT GGT GTG CAT AAT GAG ACT TGT CAA GTG ATT G | Quick Change |
| **pET21-SLFN5$^{1-336}$D233N_rev** | CAA TCA CTT GAC AAG TCT CAT TAT GCA CAC CAA AAA ATA C | Quick Change |
| **pET21-SLFN5$^{1-336}$E234A_for** | GGT GTG CAT GAT GCG ACT TGT CAA GTG | Quick Change |
| **pET21-SLFN5$^{1-336}$E234A_rev** | CAC TTG ACA AGT CGC ATC ATG CAC ACC | Quick Change |
| **pET21-SLFN5$^{1-336}$ D233A, E234A_for** | GGT GTG CAT GCT GCG ACT TGT CAA G | Quick Change |
| **pET21-SLFN5$^{1-336}$ D233A, E234A_rev** | CTT GAC AAG TCG CAG CAT GCA CAC C | Quick Change |
| **pET21-SLFN5$^{1-336}$E244A_for** | GAT TGG ATG TGA AAA AGC GAA AAT AGA CCT TAC GAG | Quick Change |
| **pET21-SLFN5$^{1-336}$E244A_rev** | CTC GTA AGG TCT ATT TTC GCT TTT TCA CAT CCA ATC | Quick Change |
| **pET21-SLFN5$^{1-336}$E244Q_for** | GAT TGG ATG TGA AAA ACA GAA AAT AGA CCT TAC GAG C | Quick Change |
| **pET21-SLFN5$^{1-336}$E244Q_rev** | GCT CGT AAG GTC TAT TTT CTG TTT TTC ACA TCC AAT C | Quick Change |
| **pET21-SLFN5$^{1-336}$R271E_for** | CAT CAT TTC TGC ACA CAG GAG CCT GAG ATA AAA TAT GTC | Quick Change |
| **pET21-SLFN5$^{1-336}$R271E_rev** | GAC ATA TTT TAT CTC AGG CTC CTG TGT GCA GAA ATG ATG | Quick Change |
| **pET21-SLFN5$^{1-336}$R326E_for** | GAC AAT TGC CCA CAG AAG AAT GGA CTG CTT G | Quick Change |
| **pET21-SLFN5$^{1-336}$R326E_rev** | CAA GCA GTC CAT TCT TCT GTG GGC AAT TGT C | Quick Change |
| **pET21-SLFN5$^{1-376}$_for** | AAG TCG ATA TAC ATA TGA GTC TTA GGA TTG AT | NdeI |
| **pET21-SLFN5$^{1-376}$_rev** | GTG GTG GTG GTG CTC GAG GCG TTT CTG CTG CTC | NotI |
| **pET21-SLFN5$^{1-447}$_for** | AAG TCG ATA TAC ATA TGA GTC TTA GGA TTG AT | NdeI |
| **pET21-SLFN5$^{1-447}$_rev** | GTG GTG GTG GTG CTC GAG GTT CTG GGA AAT TAG AAG | NotI |
| **pET21-SLFN5$^{1-502}$_for** | AAG TCG ATA TAC ATA TGA GTC TTA GGA TTG AT | NdeI |
| **pET21-SLFN5$^{1-502}$_rev** | GTG GTG GTG GTG CTC GAG TTT TCT ATC AGA ATT CAG C | NotI |
| **pET21-SLFN5$^{424-891}$_for** | GAG ATA TAC ATA TGA GCT GGG CTG TGG ATT TAG | NdeI |
| **pET21-SLFN5$^{424-891}$_rev** | GTG GTG GTG CTC GAG CAC AGA AGC CTT CAG AAT ATA C | XhoI |
| **pET21-SLFN5$^{542-891}$_for** | GAG ATA TAC ATA TGA GCT GGG CTG TGG ATT TAG | NdeI |

| | | |
|---|---|---|
| **pET21-SLFN5$^{542-891}$_rev** | GTG GTG GTG CTC GAG CAC AGA AGC CTT CAG AAT ATA C | XhoI |
| **pRS426-SLFN5$^{424-891}$_for** | CAT AAA TAA ACC ATG AGC TGG GCT GTG GAT TTA G | Infusion |
| **pRS426-SLFN5$^{424-891}$_rev** | ACC CAG ATC TGG ATC CCA CAG AAG CCT TCA GAA TAT AC | Infusion |
| **pRS426-SLFN5$^{542-891}$_for** | CAT AAA TAA ACC ATG AGC TGG GCT GTG GAT TTA G | Infusion |
| **pRS426-SLFN5$^{542-891}$_rev** | ACC CAG ATC TGG ATC CCA CAG AAG CCT TCA GAA TAT AC | Infusion |
| **pRS426-SLFN5$^{FL}$_for** | CAT AAA TAA ACC ATG AGT CTT AGG ATT GAT GTG G | Infusion |
| **pRS426-SLFN5$^{FL}$_rev** | ACC CAG ATC TGG ATC CCT CGA GCA CAG AAG CCT | Infusion |
| **pFBDM-SLFN5$^{FL}$_for** | GGG GCC CCA TAT GAG TCT TAG GAT TGA TGT GGA TAC | NdeI |
| **pFBDM-SLFN5$^{FL}$_rev** | CAA GCT TGC GGC CGC TCA CAC AGA AGC CTT CAG AAT ATA C | NotI |
| **pFBDM-SLFN5$^{FL_MBP}$_MBP_for** | CTT TCG AAT CTA GAA TGA AAC ATC ACC ATC ACC | XbaI |
| **pFBDM-SLFN5$^{FL_MBP}$_MBP_rev** | GTA TCC ACA TCA ATC CTA AGA CTC ATA TGG GCG CCC TGA AAA TAA AGA TTC CAG | NdeI |
| **pFBDM-SLFN5$^{FL_MBP}$_SLFN5_for** | GAA TCT TTA TTT TCA GGG CGC CCA TAT GAG TCT TAG GAT TGA TGT GGA TAC | NdeI |
| **pFBDM-SLFN5$^{FL_MBP}$_MBP_rev** | CTC GAC AAG CTT TCA TTT TTC GAA CTG CGG ATG GCT CCA ACT CCC CAC AGA AGC CTT CAG AAT ATA CAG | HindIII |
| **pET21-SLFN11$^{1-357}$_for** | GGG GCC CCT GGG ATC CAT GGA GGC AAA TCA G | BamHI |
| **pET21-SLFN11$^{1-357}$_rev** | GAT GCG GCC GCT CGA GCA CAG ATC CAG ATT GA | XhoI |
| **pET28M-SUMO1-SLFN11$^{FL}$_for** | CAG GAA CAA ACC GGT GGA ATG GAG GCA AAT CAG TGC CCC | Infusion |
| **pET28M-SUMO1-SLFN11$^{FL}$_rev** | GTG GTG GTG CTC GAG TCA ATG GCC ACC CCA CGG AAA AAT ATA C | Infusion |
| **pRS424-SLFN11$^{FL}$_for** | CAT AAA TAA ACC ATG GAG GCA AAT CAG TGC | Infusion |
| **pRS424-SLFN11$^{FL}$_rev** | ACC CAG ATC TGG ATC CAT GGC CAC CCC ACG G | Infusion |
| **pET21-SLFN2$^{FL}$H65E_for** | AAA ATG AAA AAT AGT GAG CTG CGT AAA CAA GAG | Quick Change |
| **pET21-SLFN2$^{FL}$H65E_rev** | CTC TTG TTT ACG CAG CTC ACT ATT TTT CAT TTT | Quick Change |
| **pET21-SLFN2$^{FL}$R67E_for** | GAA AAA TAG TCA CCT GGA GAA ACA AGA GAA CGC G | Quick Change |
| **pET21-SLFN2$^{FL}$R67E_rev** | CGC GTT CTC TTG TTT CTC CAG GTG ACT ATT TTT C | Quick Change |
| **pET21-SLFN2$^{FL}$K249E_for** | CTGCGTAAACGTATCGAAGAACTTTTGCCAC | Quick Change |
| **pET21-SLFN2$^{FL}$K249E_rev** | GTG GCA AAA GTT CTT CGA TAC GTT TAC GCA G | Quick Change |
| **pET21-SLFN2$^{FL}$K285E_for** | GGCTTTGAAGCGGAAGAGAGCGACCTGGTTC | Quick Change |
| **pET21-SLFN2$^{FL}$K285E_rev** | GAA CCA GGT CGC TCT CTT CCG CTT CAA AGC C | Quick Change |
| **pET21-SLFN2$^{\Delta C\text{-}116\text{-}W122}$_for** | GCCAGCCTGTGTAAGCACCTGGATTTTACC | Quick Change |
| **pET21-SLFN2$^{\Delta C\text{-}116\text{-}W122}$_rev** | GGT AAA ATC CAG GTG CTT ACA CAG GCT GGC | Quick Change |
| **pLOC-Slfn2_for** | GTA CAA AAA AGC AGG CTC CAC CAT GGG CAC GCG TCT TGA GG | Infusion |
| **pLOC-Slfn2_rev** | GTA CAA GAA AGC TGG GTC CAA GCC GGA AGG CGC GTT CAT C | Infusion |

**Table 6.2:** DNA and RNA oligonucleotides used for EMSA, FA and nuclease assay. Fluorescent-labeled oligonucleotides were ordered in an HPLC purity grade and dissolved in dd$H_2O$. DNA and RNA oligonucleotides were purchased from Metabion (Planegg) or Biomers (Ulm), respectively. For annealing, labeled oligonucleotides were mixed with equal molarity. The mixed oligonucleotides were incubated for 10 minutes at 95°C and slowly (for approximately 3 hours) cooled down to room temperature. 6'-FAM: 6-Carboxyfluorescin.

| Name | Direction | Sequence (5'-3' direction) | 3' label | 5' label |
|---|---|---|---|---|
| **ss 5- mer poly (dT) DNA** | | TTT TT | - | 6' FAM |
| **ss 10- mer poly (dT) DNA** | For | TTT TTT TTT T | - | 6' FAM |
| **ss 15- mer poly (dT) DNA** | For | TTT TTT TTT TTT TTT | - | 6' FAM |
| **ss 20- mer poly (dT) DNA** | For | TTT TTT TTT TTT TTT TTT TT | - | 6' FAM |
| **ss 30- mer poly (dT) DNA** | For | TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT | - | 6' FAM |
| **ss 30- mer poly (dT) DNA** | For | TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT | 6' FAM | - |
| **ss 30- mer poly (dA) DNA** | For | AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA | | |
| **ss 60- mer poly (dT) DNA** | For | TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT | - | 6' FAM |
| **ds 60- mer poly (dT) GC DNA** | For | TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TGC TTT TTT TTT | - | 6' FAM |
| **ds 60- mer poly (dA) CG DNA** | Rev | AAA AAA AAA GCA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA | - | - |
| **ds 60- mer poly (dT) GC DNA** | Rev | AAA AAA AAA GCA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA | - | - |
| **ss 60- mer DNA (random sequence)** | For | CAC CTC CTA TGA GAG GGT AAT GTA CTT GAC CTC GGT GTC CCT GCG CTC ATC TAC TAC CCG | 6' FAM | - |
| **ds 60- mer DNA (random sequence)** | For | CGG GTA GTA GAT GAG CGC AGG GAC ACC GAG GTC AAG TAC ATT ACT CTC TCA TAG GAG GTG | - | 6' FAM |
| **ds 60- mer DNA (random sequence)** | Rev | CAC CTC CTA TGA GAG AGT AAT GTA CTT GAC CTC GGT GTC CCT GCG CTC ATC TAC TAC CCG | - | - |
| **ss 30- mer poly (rU) RNA** | For | UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU | - | 6' FAM |
| **ds 60- mer poly (rU) GC RNA** | For | UUU UUU UUU GCU UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU | - | 6' FAM |
| **ds 60- mer poly (dA) GC (for DNA/RNA hybrid)** | Rev | AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AGC AAA AAA AAA | - | - |
| **Stem loop RNA** | For | GAC GGC CAU ACC ACC CUG AAC GCG CCC GAU CUC GUC UGA UCU CGA GGA CUA AGC AGG GUC GGG CCG AAA GGU GCU GUC UU | - | 6' FAM |
| **ds 41- mer RNA** | For | GAC GGC CAU ACC ACC CUG AAC GGC CCG AUC UUG UCU GAU CC | - | 6' FAM |
| **ds 41- mer RNA** | Rev | GGA UCA GAC AAG AUC GGG CCG UUC AGG GUG GUA UGG CCG UC | - | - |
| **$tRNA_{Ser}$** | For | GUA GUC GUG GCC GAG UGG UUA AGG CGA UGG ACU UGA AAU CCA UUG UGG UUU CCC CGC GCA GGU UCG AAU CCU GCC GAC UAC GCC A | - | 6' FAM |

6.1.1.2. Plasmids

**Table 6.3:** Plasmids used for recombinant protein expression.

| Construct | Description | Restriction sites/ Method | Affinity tag | Expression system |
|---|---|---|---|---|
| **pET21-SLFN5$^{1-336}$** | Slfn core domain of human SLFN5; second alternative splicing isoform | NdeI/ XhoI | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-336}$Y183F** | Slfn core domain | Quick change | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-336}$Y183A** | Slfn core domain | Quick change | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-336}$E191A** | Slfn core domain | Quick change | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-336}$E191Q** | Slfn core domain | Quick change | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-336}$E196A** | Slfn core domain | Quick change | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-336}$E196Q** | Slfn core domain | Quick change | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-336}$R211A** | Slfn core domain | Quick change | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-336}$D233A** | Slfn core domain | Quick change | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-336}$D233N** | Slfn core domain | Quick change | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-336}$E234A** | Slfn core domain | Quick change | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-336}$ D233A, E234A** | Slfn core domain | Quick change | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-336}$E244A** | Slfn core domain | Quick change | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-336}$E244Q** | Slfn core domain | Quick change | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-336}$R271E** | Slfn core domain | Quick change | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-336}$R326E** | | Quick change | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-376}$** | Slfn core domain of human SLFN5 plus linker | NdeI/ XhoI | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-447}$** | Slfn core domain plus linker and SWADL sequence of human SLFN5 | NdeI/ XhoI | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{1-502}$** | Slfn core domain plus linker and SWADL sequence of human SLFN5 | NdeI/ XhoI | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{424-891}$** | Linker region plus SWADL plus helicase domain of human SLFN5 | NdeI/ XhoI | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN5$^{542-891}$** | Helicase domain of human SLFN5 | NdeI/ XhoI | C-terminal $His_6$ | *E. coli* |
| **pRS426-SLFN5$^{424-891}$ via pENTRY_KND vector** | Linker region plus SWADL plus helicase domain of human SLFN5 | Infusion/ Gateway | C-terminal $His_8$ + TEV cleavage site | *S. cerevisiae* |
| **pRS426-SLFN5$^{542-891}$ via pENTRY_KND vector** | Helicase domain of human SLFN5 | Infusion/ Gateway | C-terminal $His_8$ + TEV cleavage site | *S. cerevisiae* |
| **pET21-SLFN5$^{FL}$** | Full length human SLFN5 | | C-terminal $His_6$ | *E. coli* |
| **pRS426-SLFN5$^{FL}$ via pENTRY_KND vector** | Full length human SLFN5 | Infusion/ Gateway | C-terminal $His_8$ + TEV cleavage site | *S. cerevisiae* |
| **pFBDM-SLFN5$^{FL}$** | Full length human SLFN5 | NdeI/ NotI | N-terminal $His_6$ + PreScission site | High Five Insect cells |
| **pFBDM-SLFN5$^{FL_MBP}$** | Full length human SLFN5 | HindIII/ XbaI | N-terminal $His_6$ + MBP + TEV cleavage site; C-terminal StrepII | High Five Insect cells |
| **pET21-SLFN11$^{1-357}$** | Slfn core domain of human SLFN11 | BamHI/ XhoI | C-terminal $His_6$ | *E. coli* |
| **pET21-SLFN11$^{FL}$** | Full length human SLFN11 | NdeI/ NotI | C-terminal $His_6$ | *E. coli* |

| | | | | |
|---|---|---|---|---|
| **pET28M-SUMO1-SLFN11$^{FL}$** | Full length human SLFN11 | Infusion | N-terminal $His_6$-SUMO1 + SenP2 cleavage site | *E. coli* |
| **pRS424-SLFN11$^{FL}$ via pENTRY_KND vector** | Full length human SLFN11 | Infusion/ Gateway | C-terminal $His_8$ + TEV cleavage site | *S. cerevisiae* |
| **pET21-Slfn2$^{FL}$** | Full length murine Slfn2 | NdeI/ XhoI | C-terminal $His_6$ | *E. coli* |
| **pET21-Slfn2$^{FL}$H65E** | Full length murine Slfn2 | Quick Change | C-terminal $His_6$ | *E. coli* |
| **pET21-Slfn2$^{FL}$R67E** | Full length murine Slfn2 | Quick Change | C-terminal $His_6$ | *E. coli* |
| **pET21-Slfn2$^{FL}$K249E** | Full length murine Slfn2 | Quick Change | C-terminal $His_6$ | *E. coli* |
| **pET21-Slfn2$^{FL}$K285E** | Full length murine Slfn2 | Quick Change | C-terminal $His_6$ | *E. coli* |
| **pET21-Slfn2$^{\Delta C\text{-}116\text{-}W122}$** | Full length murine Slfn2, but deleted hydrophobic loop C116-W122 | Quick Change | C-terminal $His_6$ | *E. coli* |
| **pLOC-Slfn2-$His_{10}$-GSG-Flag** | Full length murine Slfn2 for Co-IP | Infusion | $His_{10}$, GSG linker and Flag tag | *E.G7/Panc02* |
| **pLOC-Slfn2-$His_{10}$-GSG-Flag** | Full length murine Slfn2 carrying the *elektra* mutation I135N for Co-IP | Infusion | $His_{10}$, GSG linker and Flag tag | *E.G7/Panc02* |

pET21-SLFN5$^{1\text{-}336,}$ pET21-SLFN5$^{FL,}$ pET21-SLFN11$^{FL,}$ pET21-SLFN2$^{FL}$ were kindly provided by Dr. Florian Schlauderer.
pET21-SLFN5$^{FL,}$ pET21-SLFN11$^{FL,}$ pET21-SLFN2$^{FL}$ were purchased from GeneArt (Regensburg).
pET28-SUMO1 was kindly provided by EMBL.
pKND_ENTRY, pRS424 and pRS426 were kindly provided by the group of Kiyoshi Nagai.
Consturcts in pLOC were generated by Annika Frauenstein and Felix Meissner, MPI for Biochemistry (Munich).

### 6.1.1.3. Strains

**Table 6.4:** *E. coli* strains and eukaryotic cell lines used for molecular cloning and recombinant protein expression.

| Strain | Genotype | Source |
|---|---|---|
| ***E. coli* XL1blue** | *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacIqZΔM15 Tn10 (Tet$^R$)]* | Stratagene (Heidelberg) |
| ***E. coli* BL21 (DE3)** | *F$^-$ ompT hsdSB ($r_B^-m_B^-$) gal dcm* (DE3) | Novagen (USA) |
| ***E. coli* Rosetta** | *F$^-$ ompT hsdSB ($r_B^-$ $m_B^-$) gal dcm* (DE3) *pRARE (CamR)* | Novagen (USA) |
| ***E. coli* ArcticExpress (DE3)** | *B F$^-$ompT hsdS($r_B^-m_B^-$) dcm$^+$ Tet$^r$ galλ*(DE3) *endA* Hte *[cpn10 cpn60 Gent$^r$]* | Agilent Technologies (USA) |
| ***E. coli* SHuffle T7** | *fhuA2 lacZ::T7 gene1 [lon] ompT ahpC gal λatt::pNEB3-r1-cDsbC (Spec$^R$,lacI$^q$) ΔtrxB sulA11 R(mcr-73::miniTn10--Tet$^S$)2 [dcm] R(zgb-210::Tn10 --Tet$^S$) endA1 Δgor Δ(mcrC-mrr)114::IS10* | New England BioLabs (USA) |
| ***E. coli* BL21 Lemo** | *fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS/* pLemo(Cam$^R$) *λ DE3 = λsBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5* pLemo = pACYC184-*PrhaBAD-lysY* | New England BioLabs (USA) |
| ***E. coli* DH10MultiBac** | *F– mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔ M15 lacX74 recA1 endA1 araD139 (ara, leu)7697 galU galK Δ– rpsL nupG/pMON14272/pMON7124* | Redbiotech (Schlieren) |
| ***S. cerevisiae* BCY123** | *MATα trp1 ura3 his3 leu2 pep4::HIS3 prb1::LEU2 bar1::HIS3 lys2::GAL1/10-GAL4* | kindly provided by the group of Kiyoshi Nagai |
| **IPLB-Sf21-AE (*Spodoptera frugiperda)*** | | Life Technologies (Thermo Fisher Scientific, USA) |
| **BTI-TN-5B1-4 (High Five) *(Trichoplusia ni)*** | | Life Technologies (Thermo Fisher Scientific, USA) |

| | | |
|---|---|---|
| ***Mus musculus* E.G7-OVA** | [derivative of EL4] ATCC® CRL-2113™ | Kindly provided by the group of Sebastian Kobold |
| ***Mus musculus* Panc02-OVA** | 266-6 ATCC® CRL-2151™ | Kindly provided by the group of Sebastian Kobold |

6.1.1.4. Media and antibiotics

LB medium was used to culture *E. coli* strains and were prepared as described in Table 6.5. For the generation of solid media, 1.5 % agar (w/v) was added. Antibiotics used for the selection of transformed *E. coli* are listed in Table 6.6. The stocks were used as 1: 1000 (v/v) dilutions.

**Table 6.5:** Media used for culturing *E. coli* strains.

| Compound | LB |
|---|---|
| Tryptone | 10 g |
| Yeast extract | 5 g |
| NaCl | 10 g |
| $H_2O$ (deionized) | Up to 1000 ml |

**Table 6.6:** List of antibiotics used for the maintenance of *E. coli* strains.

| Compound | Stock concentration [g/l] | Solvent |
|---|---|---|
| Ampicillin | 100 | $ddH_2O$ |
| Kanamycin | 50 | $ddH_2O$ |
| Chloramphenicol | 34 | Ethanol |
| Tetracyclin | 10 | Ethanol |
| Gentamicin | 7 | $ddH_2O$ |

Untransformed *S. cerevisiae* BCY123 were cultivated in YPD medium, whereas transformed cells were cultivated in minimal medium SD –Trp –Ura for expression and maintenance of transformed cells (Table 6.7.). For the generation of solid media 2 % agar (w/v) was added.

**Table 6.7:** Media used for culturing *S. cerevisiae* BCY123.

| Compound | YPD | Compound | SD –Ura -Trp |
|---|---|---|---|
| Peptone (or tryptone) | 20 | Yeast nitrogen base without amino acids (Formedium) | 6.9 g |
| Yeast extract | 10 | Complete supplement mixture (CSM) | 0.72 g |
| $H_2O$ (deionized) | Up to 900 ml | $H_2O$ (deionized) | Up to 950 ml |
| autoclave | | | |
| 20 % (w/v) glucose | 100 ml (2 % (w/v) endconcentration) | 20 % (w/v) raffinose | 50 ml (1 % (w/v) endconcentration) |

Sf21 and High five insect cells were cultivated in Sf-900 III serum-free medium purchased from Life Technologies (Thermo Fisher Scientific) and Express Five serum-free medium prepared according to the manufacturer's instructions (Life Technologies, Thermo Fisher Scientific), respectively. Insect cell media were supplemented with 10 µg/ ml gentamicin (Roth) and 10 mM L-glutamine (Invitrogen, Thermo Fisher Scientific).

### 6.1.2. Miscellaneous material

**Table 6.8:** Overview of the DNA and protein markers used.

| Name | Manufacturer |
|---|---|
| GeneRuler 1 kb DNA Ladder | Thermo Scientific (#SM0311) |
| Low molecular weight DNA Ladder | New England BioLabs (#N3233S) |
| PageRuler Unstained Protein Ladder | Thermo Scientific (#26614) |
| PageRuler Prestained Protein Ladder | Thermo Scientific (#26616) |
| Gel Filtration Standard (Thyroglobulin: 670 kDa, Bovine γ-globulin: 158 kDa, Chicken ovalbumin: 44 kDa, Equine myoglobin: 17 kDa, Vitamine $B_{12}$: 1.35 kDa) | BioRad (#1511901) |

**Table 6.9:** Overview of the crystallization screens used.

| Screen | Manufacturer |
|---|---|
| Additive Screen | Hampton Research |
| Ammonium Sulfate Screen | Qiagen |
| Crystal Screen 1, 2 | Hampton Research |
| Crystal Screen Cryo | Hampton Research |
| Index | Hampton Research |
| JBScreen Classic 1, 2, 4, 5, 6, 7, 8, 10 | Jena Bioscience |
| JBScreen Nuc-Pro 1 - 4 | Jena Bioscience |
| JBScreen PACT 1-4 | Jena Bioscience |
| JCSG-plus | Molecular Dimensions |
| MIDAS | Molecular Dimensions |
| Morpheus | Molecular Dimensions |
| Natrix HT | Hampton Research |
| PGA Screen | Molecular Dimensions |
| Pre Crystallization Test | Hampton Research |

## 6.2. Methods

### 6.2.1. Molecular biology methods

#### 6.2.1.1. Molecular cloning

The principles of molecular cloning were followed according to[250]. The genes encoding for Slfn2, SLFN5 and SLFN11 were amplified by polymerase chain reaction (PCR) using a self-made mixture with the identical composition as the Phusion® High-Fidelity PCR Master Mix (Thermo Fisher Scientific). To a typical PCR reaction 0.5 pmol of each primer and approximately 1 ng of the template DNA was added. A standard PCR program is shown in Table 6.10.

**Table 6.10:** Standard PCR program used for molecular cloning. The annealing temperatures are primer specific and were calculated for each primer using the OligoCalc online tool[251].

| Cycle step | Temperature [°C] | Time | Cycles |
|---|---|---|---|
| **Initial denaturation** | 98 | 5 min | 1 |
| **Denaturation** | 98 | 15 sec | 30 |
| **Annealing** | X (primer specific) | 15 sec | |
| **Extension** | 72 | 20 sec / 1kb | |
| **Final extension** | 72 | 10 min | 1 |

Amplified PCR products were either separated by an agarose gel electrophoresis, respective DNA fragments of the correct size were cut out and isolated using the NucleoSpin® Gel, or purified by PCR Clean-up Kit (Machery&Nagel). The amplified PCR products and the target vector DNA were either assembled with the In-Fusion® HD Cloning Kit (Takara Bio USA) or digested with the corresponding restriction enzymes and ligated. The assembled constructs were transformed into chemically competent *E. coli* XL1blue and the enriched plasmids were isolated using NucleoSpin Plasmid Easy Pure kit (Machery&Nagel). Proper assembly of the cloned constructs and the integrity of the open reading frame were verified by sequencing by GATC (Konstanz).

#### 6.2.1.2. Site-directed mutagenesis (Quick Change)

Point mutations of Slfn2 and SLFN5 and the Slfn2 deletion mutant were generated by site-directed mutagenesis (Quick Change)[252]. For a typical PCR reaction 0.05 pmol of each primer was used and 1 ng of the template DNA. The standard PCR program was performed as described in Table 6.8, however the amount of cycles were decreased to 15. Subsequently, DpnI (Fast Digest, Thermo Scientific) was added directly to the PCR reaction and incubated at 37°C for 1 hour. The reactions were transformed into chemically competent *E. coli* XL1blue. The presence of the mutation was verified by sequencing.

#### 6.2.1.3. Agarose gelelectrophoresis

The agarose-gelelectrophoresis was performed to separate and visualize DNA of various sizes. Normally a 1 % (w/v) agarose gel were used. The amount of agarose was dissolved in 1x TAE buffer by heating. Before pouring the agarose in the gel-chambers GelRed was added in a 1: 10 000 ratio. Before applying the samples onto the gel they were prepared by adding 1 x DNA loading dye. As running buffer

1 x TAE was chosen. The run was performed with 120 V and 400 mA. The DNA was visualized under UV-light (λ=254 nm).

| 1 x Tris/ Acetate/ EDTA (TAE) buffer |
|---|
| 40 mM Tris/HCl pH 7.6 |
| 20 mM Acetic acid |
| 1 mM EDTA |

| 6 x DNA loading dye |
|---|
| 0.5 % (w/v) Bromphenol blue |
| 0.5 % (w/v) Xylene cyanol |
| 60 % (v/v) Glycerol |

6.2.1.4. Transformation of *E. coli*

To one 50 µl aliquot of chemical competent *E. coli* cells, which were thawed on ice, 20 – 200 ng plasmid DNA were added. After 20 minutes, incubation on ice the heat shock was performed via incubation at 42°C for one minute, followed by a two minutes incubation on ice. After adding 900 µl of LB medium the cells were shaken (600 rpm) at 37°C for 45 minutes. Cells were further cultivated either as precultures in liquid LB or on LB-agar plates supplied with the respective antibiotics. Cultivation takes normally place at 37°C over night.

For the transposition of the pFBDM vector into the bacmid, *E. coli* DH10MultiBac cells were transformed. After performing the heat shock, cells recovered 6-16 h at 37°C. The cells were spread onto agar-plates containing kanamycin, gentamicin and tetracycline at 1:1000 dilutions, as well as X-Gal (100 µg/ml) and IPTG (0.5 mM). The plates were incubated for 2-3 days at 37°C. Afterwards white colonies were picked, enriched and isolated with the NucleoSpin Plasmid Easy Pure Midi kit (Machery&Nagel).

6.2.1.5. RNA preparation and cDNA generation for Ilumina RNA deep sequencing

Protein samples containing co-purified nucleic acid were treated with Proteinase K (Thermo Fisher Scientific) in 1x Proteinase K buffer at 50°C for 30 minutes in a total volume of 200 µl. The same volume of phenol/ chloroform/ isoamylalcohol (25: 24: 1; pH 4.5- 5) (Carl Roth) was added and transferred into Phase Lock Gel™ (VWR, USA) tubes. After centrifugation (25 000 x g, 1 min) the hydrophilic phase containing the RNA was transferred into a new tube. Next, the RNA solution was mixed in a 1: 0.4: 0.05: 5 ratio with 3 M sodium acetate, 10 LiCl and 100 % ethanol, respectively. Prior to incubation at -80°C for 2-3 hours, GlycoBlue™ Coprecipitant (15 g/l) (Thermo Fisher Scientific) was added to increase the visibility of the RNA in the precipitation followed. The sample was centrifuged at 25 000 x g for 30 minutes and the pellet was washed with 70 % (v/v) ethanol. After another centrifugation step (25 000 x g, 30 min), the sample was dried and resolved in dd$H_2O$. The quality of the prepared RNA samples were validated on a Bioanalyzer 2100 (Agilent) according to the manufacturer's instructions.

| 1 x Proteinase K buffer |
|---|
| 50 mM Tris-HCl, pH 7.5 |
| 75 mM NaCl |
| 6 mM EDTA |
| 1 % (w/v) SDS |

The cDNA library was generated using Encore® Complete RNA-Seq Library Systems (Nugen, USA) or SENSE™ mRNA-Seq Library Prep Kit V2 (Lexogen, Austria). When using the latter kit, samples were adenylated and purified by phenol precipitation prior to cDNA generation. Polyadenylation was performed using Poly(A) Polymerase (New England BioLabs) by following the manufacturer's

instructions. Sequencing was performed on the Illumina Genome Analyzer in the Gene Center sequencing facility (LAFUGA). cDNA library preparations, validation on the Agilent Bioanalyzer and Illumina RNA deep sequencing were performed by Stefan Krebs (LAFUGA).

### 6.2.2. Bioinformatic Methods

#### 6.2.2.1. Illumina deep sequencing data analysis

Obtained Illumina sequences were processed and analyzed using the GALAXY platform[253]. First, the sequences were demultiplexed for paired-end runs with barcodes at the beginning of the first read or with an index lane by using GBSX demultiplexer[254]. Next, adaptor sequences were removed by applying NGS (Galaxy)[253] and identical sequences were collapsed into one FASTA file. Remaining sequences were mapped to *Escherichia coli* str. K-12 substr. MG1655, complete genome (GenBank: U00096.3, NC_000913.2) and *Bombyx mori* genome for the Slfn2$^{FL}$ and SLFN5$^{FL}$ samples, respectively, by utilization of the Bowtie algorithm[255]. For Slfn2$^{FL}$, the number of reads were plotted against the *E. coli* K-12 genome for visualization.

#### 6.2.2.2. Sequence Alignment

Multiple Sequence alignments were generated with T-coffee[256] and visualized using Espript[257] or Jalview[258].

#### 6.2.2.3. Calculation of Protein Parameters

Physical and chemical parameters of the recombinant proteins, e.g. molecular weight, theoretical isoelectric point, extinction coefficients were calculated using ProtParam, part of the Expasy server[259].

The sequence identitiy was calculated using T coffee[256] and the nuclear localization signals were predicted using NLS mapper[260].

#### 6.2.2.4. Structure Visualization and Analysis

The visualization of crystal structures was carried out using PyMol 4.5.0 (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.) or the UCSF Chimera package[261]. The electrostatic surface was calculated using the APBS plugin for Pymol[262]. Data analyses were performed with Prism (GraphPad Software, La Jolla California, USA), OriginPro 8 (Origin Lab, Northampton, USA.) and Adobe Illustrator CS6 (Adobe Systems).

#### 6.2.2.5. Structural Homology Modeling

Comparative structural homology modeling was performed with Expasy SWISS-MODEL[263].

6.2.3. Protein biochemistry methods

6.2.3.1. Recombinant protein expression

**i Protein expression using *E. coli***

For recombinant protein expression in bacterial expression systems (*E. coli*, Table 6.4.), generally the protocols of the manufacturers were followed. In principle, the expression cultures were inoculated 1: 500 with an overnight preculture and incubated at 37°C under agitation (220 rpm). After the cultures have reached an $OD_{600}$ of 0.4 the temperature was set to the final expression temperature. Recombinant protein expression was induced by adding 0.2 mM IPTG after reaching $OD_{600}$ 0.8-1.0. Depending on the final expression temperatures, the expression continued for several hours (e.g. 3-4 h at 37°C) or overnight (18°C) before cell harvest.

**ii Protein expression in *S. cerevisiae* BCY123**

For recombinant protein expression in *S. cerevisiae* BCY123, cells were transformed with the respective plasmids. For that a *S. cerevisiae* preculture was grown until $OD_{600}$ reached 0.6-1.0. Cells were harvested by centrifugation (2000 rpm, 3-5 min), washed with sterile deionized water once and with LiTE twice. Finally, they were resuspended in 200 µl LiTE. For each transformation 2 µl salmon sperm carrier DNA (10 µg/µl), 2-5 µl of each plasmid (typically 0.2-1 µg) and 15 µl of LiTE-treated yeast cells were used. After incubation at 30°C for 15 minutes pre-warmed PEGTE were added to the transformation mix and again incubated at 30°C for 15 minutes. The cells were then transferred to a 42°C water bath for approximately 8 minutes. Afterwards the cells were chilled on ice for 1-2 minutes and harvested (12 000 x g, 10-15 s). The pellet was resuspended in 100 µl deionized water and spread on selective plate (SD –Trp –Ura agar). If the transformation was successful, colonies were visible after 3-5 days at 30°C. To perform recombinant protein expression transformed cells were twice precultured in SD –Trp –Ura + 2 % (w/v) raffinose at 30°C under agitation (155 rpm) over night. The expression cultures were inoculated with the precultures for an $OD_{600}$ of 0.2- 0.4 and incubated at 30°C and 155 rpm. Recombinant protein expression was either conducted in baffled flaskes (150 rpm, 30°C) or in a fermenter (360 rpm, 30°C, pH 5.0). Recombinant protein expression was induced by adding 50 ml of 40 % (w/v) galactose per liter of culture. Cells were harvested the next day by centrifugation, drop-freezed in liquid nitrogen and stored at -80°C until further use.

| LiTE |
|---|
| 0.1 M lithium acetate |
| 10 mM Tris-HCl pH7.4 |
| 1 mM EDTA |

| PEGTE |
|---|
| 40 % (w/v) PEG4000 |
| 10 mM Tris-HCl pH7.4 |
| 1 mM EDTA |

**iii Protein expression in insect cell expression system (High five)**

For transfection of Sf21 cells, cells were diluted to 0.4 x $10^6$/ml suspended in Sf-900 III SFM medium and incubated at 27.5°C for 30 minutes. A premixture was added to the cells consisting of 200 µl of Sf-900 III SFM medium, 3 µl of FuGene transfection reagent (Promega) and 1-3 µg of isolated bacmid DNA after 45 minutes of incubation at room temperature. Afterwards the transfected cells were incubated at

27.5°C for about 3 days. After 3 to 4 days the media containing P0 virus were collected and centrifuged at 3000 rpm for 15 minutes. The initial virus P0 was added to 10 ml Sf21 cells at 1 x $10^6$/ml and incubated in a shaker at 27.5°C for 3 days. After harvesting the cells by centrifugation at 3000 rpm for 15 minutes, the supernatant containing virus P1 was kept and stored at 4°C until further use. Again, P1-containing media was added to 20 ml Sf21 cells diluted to 0.4 x $10^6$/ml in Sf-900 III SFM medium. The volume of the P1 virus supernatant was dependent on the strength of the virus. After a 3 day incubation in a shaker at 27.5°C, the cells were counted and centrifuged at 3000 rpm for 15 minutes. The supernatant containing P2 viruses were stored at 4°C. For protein expression 0.5-1 l of High Five cells at 1 x $10^6$/ml cultured in Express Five medium was infected with the P2 virus. The ratio of expression culture to volume of virus added was dependent on the strength of the P2 virus, typically in a 1: 1000 ratio. After 72 hours, the cells were harvested by centrifugation (3000 rpm, 15 minutes). The pellets were frozen in liquid nitrogen and stored at -80°C until further use.

6.2.2.3. Protein purification

All buffers and their components used for protein purification are listed in Table 6.11.

*E.coli* cells were lysed in lysis buffer supplemented with a protease inhibitor cocktail (0.176 g/l phenylmethylsulfonyl fluoride, 0.316 g/L benzamidine hydrochloride, 1.372 mg/l pepstatin, 0.256 mg/l leupeptin, 0.2 mg/l chymostatin) by sonication (15 min, microtip level 6, 50 % output). Frozen *S. cerevisiae* cells were disrupted using the freezer mill and afterwards resuspended in lysis buffer. High Five insect cells were lysed in lysis buffer supplemented with cOmplete™ Protease Inhibitor Cocktail (Roche) and sonified for 30 seconds.

Cell debris was removed by centrifugation (30 000 x g at 4°C for 30 minutes) and the target protein was captured by applying the supernatant onto Ni-NTA Sepharose beads (affinity chromatography (AC)) packed in a gravity flow cartridge (BioRad, Munich).

SLFN5$^{1-336}$ was directly applied onto a 5 ml Q-HP column (GE Healthcare), but remained in the flow-through. It was further purified by performing a size exclusion chromatography (SEC) using Superdex S200 16/60 or 26/60 column (GE Healthcare) previously equilibrated with SEC buffer. All mutants were purified accordingly. For the SLFN5$^{1-336}$ "Q+H" batch an additional dialysis step and Heparin column (5 ml HiTrap Heparin HP, GE Healthcare) was interposed between the negative Q-Sepharose and SEC.

After the capturing step with Ni-NTA, SLFN5$^{FL}$ expressed in *S. cerevisiae* was dialyzed and simultaneously the octahistidine tag was cleaved by adding TEV-protease. The following day it was again applied onto Ni-NTA beads. The protein remaining in the flow-through was further purified by SEC using a S200 26/60 column.

SLFN5$^{FL}$ isolated from insect cells were directly loaded onto a SEC S200 26/60 column.

After the capturing step by Ni-NTA SLFN11$^{FL}$ expressed in *S. cerevisiae* was loaded to a 5 ml Heparin (QTrap Heparin HP, GE Healthcare). SLFN11$^{FL}$ bound to the column and was eluted by applying high salt. After dialysis and simultaneous cleavage of the octahistidine tag by applying TEV-protease SLFN11$^{FL}$ was again applied onto Ni-NTA beads.

The initial purification protocol for Slfn2$^{FL}$ consists of a SEC after the AC by using a Superdex S200 16/60 or 26/60 equilibrated with SEC buffer. The mutants were purified accordingly. In the optimized Slfn2$^{FL}$ protocol, Slfn2$^{FL}$ was dialyzed prior to application to a 5 ml Q-HP column. Slfn2$^{FL}$ bound to the

column and eluted by applying high salt buffer. Finally, SEC was performed using S200 16/60 or 26/60 preequilibrated with SEC buffer. The deletion mutant used for crystallization was purified following the optimized protocol.

Prior to SEC and prior to freezing the proteins were concentrated with centrifugal concentrators (Amicon® Ultra Centrifugal Filters, Merck, Darmstadt). After successful purification, the proteins were flash frozen and stored at -80°C until further use.

**Table 6.11:** Buffers used for protein purification.

| Buffer | Composition |
|---|---|
| **SLFN5$^{1-336}$ Lysis Buffer** | 50 mM HEPES pH 8.0, 200 mM NaCl, 2 mM $MgCl_2$, 10 mM imidazole, 4 mM β-ME |
| **SLFN5$^{1-336}$ Ni-NTA Wash** | 50 mM HEPES pH 8.0, 200 mM NaCl, 2 mM $MgCl_2$, 22 mM imidazole, 4 mM β-ME |
| **SLFN5$^{1-336}$ Ni-NTA Elution** | 50 mM HEPES pH 8.0, 200 mM NaCl, 2 mM $MgCl_2$, 250 mM imidazole, 4 mM β-ME |
| **SLFN5$^{1-336}$ IEX A** | 50 mM HEPES pH 8.0, 200 mM NaCl, 2 mM $MgCl_2$, 4 mM β-ME |
| **SLFN5$^{1-336}$ IEX B** | 50 mM HEPES pH 8.0, 1 M NaCl, 2 mM $MgCl_2$, 4 mM β-ME |
| **SLFN5$^{1-336}$ SEC** | 50 mM HEPES pH 8.0, 200 mM NaCl, 2 mM $MgCl_2$, 4 mM β-ME |
| **SLFN5$^{FL}$ Lysis Buffer** | 50 mM HEPES pH 6.5, 500 mM NaCl, 2 mM $MgCl_2$, 10 mM imidazole, 4 mM β-ME |
| **SLFN5$^{FL}$ Ni-NTA Wash I** | 50 mM HEPES pH 6.5, 500 mM NaCl, 2 mM $MgCl_2$, 58 mM imidazole, 4 mM β-ME |
| **SLFN5$^{FL}$ Ni-NTA Wash I** | 50 mM HEPES pH 6.5, 500 mM NaCl, 2 mM $MgCl_2$, 82 mM imidazole, 4 mM β-ME |
| **SLFN5$^{FL}$ Ni-NTA Elution** | 50 mM HEPES pH 6.5, 500 mM NaCl, 2 mM $MgCl_2$, 250 mM imidazole, 4 mM β-ME |
| **SLFN5$^{FL}$ SEC** | 50 mM HEPES pH 6.5, 500 mM NaCl, 2 mM $MgCl_2$, 4 mM β-ME |
| **SLFN11$^{FL}$ Lysis Buffer** | 50 mM HEPES pH 6.5, 300 mM NaCl, 10 mM Imidazol, 4 mM β-ME |
| **SLFN11$^{FL}$ Ni-NTA Elution** | 50 mM HEPES pH 6.5, 300 mM NaCl, 2 mM $MgCl_2$, 250 mM imidazole, 4 mM β-ME |
| **SLFN11$^{FL}$ IEX A** | 50 mM HEPES pH 6.5, 300 mM NaCl, 4 mM β-ME |
| **SLFN11$^{FL}$ IEX B** | 50 mM HEPES pH 6.5, 1 M NaCl, 4 mM β-ME |
| **SLFN11$^{FL}$ 2.AC** | 50 mM HEPES pH 6.5, 300 mM NaCl, 10 mM Imidazol, 4 mM β-ME |
| **SLFN11$^{FL}$ 2.AC Elution** | 50 mM HEPES pH 6.5, 300 mM NaCl, 2 mM $MgCl_2$, 250 mM imidazole, 4 mM β-ME |
| **Slfn2$^{FL}$ Lysis Buffer** | 50 mM HEPES pH 7.5, 300 mM NaCl, 2 mM $MgCl_2$, 10 mM imidazole, 4 mM β-ME |
| **Slfn2$^{FL}$ Ni-NTA Wash I** | 50 mM HEPES pH 7.5, 300 mM NaCl, 2 mM $MgCl_2$, 22 mM imidazole, 4 mM β-ME |
| **Slfn2$^{FL}$ Ni-NTA Wash II** | 50 mM HEPES pH 7.5, 300 mM NaCl, 2 mM $MgCl_2$, 37.5 mM imidazole, 4 mM β-ME |
| **Slfn2$^{FL}$ Dialysis/ IEX A/ SEC** | 50 mM HEPES pH 7.5, 50 mM NaCl, 2 mM $MgCl_2$, 4 mM β-ME |
| **Slfn2$^{FL}$ Dialysis/ IEX B** | 50 mM HEPES pH 7.5, 1 M NaCl, 2 mM $MgCl_2$, 4 mM β-ME |

6.2.3.3. Biochemical methods

**i Discontinuous SDS-polyacrylamide gelelectrophoresis**

The electrophoretic separation of proteins and other molecules was performed by discontinuous SDS-polyacrylamide gelelectrophoresis according to Laemmli[264]. Here, RunBlue 4-20 % gradient SDS polyacrylamide gels (Expedeon) using the vertical Mini-PROTEAN 3 System (BioRad) were used. Samples were prepared by adding 1 x Laemmli loading dye and heated to 95°C for 5 minutes. Electrophoresis was performed at 180V- 200V in RunBlue TEO-Tricine buffer (Expedon). Subsequently, the gels were stained with Coomassie or Instant Blue (Expedeon). For destaining water was used.

| 1 x Laemmli loading dye |
|---|
| 25 mM Tris pH 8.4 |
| 4 % (v/v) glycerol |
| 1 % (w/v) SDS |
| 1.25 % β- mercaptoethanol |
| 0.15 %(w/v) bromphenolblue |

| Coomassie |
|---|
| 7 % acetic acid |
| 50 % ethanol |
| 0.2 % (w/v) Coomassie Brilliant Blue R-250 |

**ii Quantification of protein concentration**

**Protein quantification by Bradford test**

Protein concentration determination was performed photometrically according to Bradford[265]. The Bradford reagent (Roti®-Quant, Carl Roth) was diluted 1: 5 with $ddH_2O$. Subsequently the protein solution was mixed 1: 50 with the diluted Bradford solution and incubated 10 minutes at room temperature. Afterwards the $OD_{595}$ was measured against one blank sample. The absorption in the range of 0.1 until 0.9 correspond to the protein concentration between 0.1 mg/ml and 0.9 mg/ml, respectively.

**Protein quantification by the molar extinction coefficient**

Protein quantification was determined photometrically by using the Lambert-Beer law. By using a Nanodrop Spectrophotometer (Thermo Fisher Scientific) the absorption at λ=260 nm and λ=280 nm were measured. With the calculated molar extinction coefficient, the concentration could be determined (c: sample concentration, d: distance to detector (1 cm), A: absorption at 280 nm, ε: molar extinction coefficient of the sample): $c = A / (d * \varepsilon)$

**iii Analytical size exclusion chromatography (SEC) and SEC coupled right-angle light scattering (RALS)**

For analytical size exclusion chromatography either a Superdex 200 10/300 Increase (GE Healthcare) or a Superose 6 10/300 Increase (GE Healthcare) was used, previously equilibrated with the respective SEC buffer (Table 6.11). To check the integrity of the column and for an estimation of the molecular weight of the target protein, a gel filtration standard (BioRad) consisting of five marker proteins was applied regularly.

For SEC coupled right-angle light scattering an ÄKTA micro (GE Healthcare) was used equipped with a VE3580 Refractive Index-device and TD270 right-angle laser light scattering detector (Viscotek/Malvern Instruments). A Superdex S200 10/300 Increase size-exclusion column (GE Healthcare) was chosen. Calibration was performed by using BSA (66 kDa) as a standard protein. Analysis of data was performed using Viscothek Software OmniSEC.

6.2.3.4. Bioanalytical methods

**i Electrophoretic mobility shift assay (EMSA)**

To constant concentrations of 0.2 µM 6'-FAM labeled DNA or RNA increasing amounts of proteins were titrated in EMSA buffer and incubated for 30 minutes at 37°C. Samples were separated using NativePAGE™ 4-16% Bis-Tris Protein Gels (Thermo Scientific). The fluorescence was measured at an excitation wavelength of 495 nm and an emission wavelength of 520 nm by using the Typhoon Laser Scanner (GE Healthcare).

| 1 x EMSA buffer |
|---|
| 50 mM HEPES pH 8.0 |
| 50 mM NaCl |
| 5 mM $MgCl_2$ |

### ii Affinity measurement by fluorescence anisotropy (FA)

Initial protein dilutions (0, 1, 2, 3, 5, 10, 15, 20, 30, 40 and 60 µM) were prepared in 2x assay buffer and then mixed with 6'-FAM labeled DNA or RNA (at a final concentration of 100 nM) in a 1:1 (v/v) ratio. The reaction was incubated for 30 minutes on ice and subsequently the fluorescence anisotropy was measured at an excitation wavelength of 495 nm and an emission wavelength of 520 nm. The data were analyzed with Prism (GraphPad Software, La Jolla California, USA) and fit to a Hill model.

| 2 x Assay buffer |
|---|
| 100 mM HEPES pH 8.0 |
| 100 mM NaCl |

### iii Urea-PAGE based exonuclease assay

Steady state nuclease assays were performed with 5 µM $SLFN5^{1-336}$ or $Slfn2^{FL}$ variants with 0.2 µM 6'- FAM labeled DNA/ RNA in 1x assay buffer in presence or absence of 5 mM ATP, 10 mM EDTA, 1 mM $MnCl_2$, 2 mM $CaCl_2$ or 0.1 mM $ZnCl_2$ as indicated. The reaction mixture was incubated for 1, 2, 5, 15, 30 or 45 minutes at 25°C. The zero time point was taken prior to the initiation of the reaction. The reaction was terminated by adding an equal volume of 2x Ficoll-loading buffer and incubation at 95°C for 10 minutes. Subsequently, the reactions were loaded onto a Urea-Acrylamide gel (Rotiphorese® DNA sequencing system, Carl Roth) and run at a constant power of 35 Watt for 105 minutes in 1xTBE buffer. FAM labeled substrates were scanned with a 495 nm laser and 520 nm filter by using a Thyphoon fluorescence imager (GE Healthcare).

| 1 x assay buffer | 2x Ficoll –loading dye | 1 x Tris/Borate/EDTA (TBE) buffer | 15 % Urea-Polyacrylamidegel |
|---|---|---|---|
| 50 mM HEPES pH 8.0 | 8 M Urea | 90 mM Tris/HCl pH 7.6 | 60 % (v/v) Rotiphorese® sequencing gel concentrate (25 % acrylamide/bisacrylamide, mixing ratio 19:1 and 50 % urea) |
| 50 mM NaCl | 125 mM EDTA | 90 mM Boric acid | 30 % (v/v) Rotiphorese® sequencing gel diluent (50 % urea) |
| 5 mM $MgCl_2$ | 7 % Ficol 400 | 2 mM EDTA | 10 % 10x TBE |
| | | | 0.1 % (w/v) APS |
| | | | 0.1 % (v/v) TEMED |

### iv Differential radial capillary action of ligand assay (DRaCALA)

To test ATP binding capabilities DRaCALA assays were performed as described in [266]. For $SLFN5^{1-336}$ 10 µM protein was incubated in absence or in presence of 10 µM single-stranded 60- mer poly (dT) DNA in 1x Tris-based ATPase buffer and 5 nM [$\alpha$-$^{32}$P] ATP (Hartmann Analytik, Germany) at room

temperature for 20 minutes. For $SLFN5^{FL}$ 1.8 µM protein was incubated in absence or in presence of 2 µM single-stranded 30- mer poly (dT) DNA or 2 µM single-stranded 30- mer poly (rU) RNA in 1x HEPES-based ATPase buffer and 5 nM [α-$^{32}$P] ATP (Hartmann Analytik, Germany) at room temperature for 20 minutes. As a positive control RIG-I was used, which was kindly provided by Dr. Charlotte Lässig[201]. For analysis 2.5 µl of the reaction mixture was spotted on nitrocellulose membrane (0.22 µM pores, GE Healthcare, Buckinghamshire, UK), air-dried and the [α-$^{32}$P] ATP was detected using Thyphoon phosphor imager (GE Healthcare).

| 1x Tris-based ATPase buffer |
|---|
| 50 mM Tris-HCl pH 7.5 |
| 150 mM KCl |
| 5 mM $MgCl_2$ |

| 1x HEPES-based ATPase buffer |
|---|
| 50 mM HEPES pH 6.5 |
| 300 M NaCl |
| 5 mM $MgCl_2$ |
| 5 mM $CaCl_2$ |

### v Radioactive based ATP hydrolysis assay

For the ATPase hydrolysis assay 5 µM $SLFN5^{1-336}$ was incubated in presence or absence of 0.2 µM single-stranded 60- mer poly (dT) DNA in 1x Tris-based ATPase buffer supplemented with 1.5 mM ATP and 10 nM [γ-$^{32}$P] ATP (Hartmann Analytik, Germany) at 37°C for 0 or 60 minutes. 1 µM $SLFN5^{FL}$ was incubated with 1 µM single-stranded 30- mer poly (dT) DNA, 3 nM pRS vector or 3 nM XhoI-linearised pRS vector with Tris- or HEPES-based ATPase buffer at room temperature for 20 minutes. As a positive control RIG-I was used, which was kindly provided by Dr. Charlotte Lässig[201]. For analysis, 1 µl of the reaction mixture was applied onto polyethylenimine cellulose TLC plates (Sigma-Aldrich, Germany) and free phosphate was separated from ATP by thin layer chromatography in TLC running buffer (1 M formic acid, 0.5 M LiCl). [γ-$^{32}$P] ATP was detected using a phosphor-imaging system (GE Healthcare, Germany).

| 1x Tris-based ATPase buffer |
|---|
| 50 mM Tris-HCl pH 7.5 |
| 150 mM KCl |

| 1x HEPES-based ATPase buffer |
|---|
| 50 mM HEPES pH 6.5 |
| 300 M NaCl |
| 5 mM $MgCl_2$ |
| 5 mM $CaCl_2$ |

| TLC running buffer |
|---|
| 1 M formic acid |
| 0.5 M LiCl |

### vi Co-immunoprecipitation and affinity enrichment mass spectrometry

E.G7 and Panc02 cells were cultivated in DMEM/ GlutaMAX medium supplemented with 10 % FBS (Thermo Fisher Scientific). Cells were lentiviral transfected with pLOC-$His_{10}$-GSG-Flag, pLOC-$Slfn2^{FL}$-$His_{10}$-GSG-Flag and pLOC-$Slfn2^{I135N}$-$His_{10}$-GSG-Flag according to[267]. Samples for affinity enrichment mass spectrometry measurement were prepared according to[268]. Briefly, 1 x $10^7$ cells were washed in 1 x PBS and lysed by an incubation with lysis buffer supplemented with benzonase (50 U/ ml), 1 x Protease Inhibitor tablet (Roche) and 1x Phospatase Inhibitor tablet (Roche) on ice for 15 minutes and a subsequent centrifugation at 25 000 x g for 2 minutes. The supernatant was mixed with Protein G beads coupled to anti-His antibody and washed extensively with washing buffer. Next, samples were

sequentially digested with LysC (Wako Chemicals USA) and trypsin (Promega), acidified with 0.1% TFA, desalted with C18 stage tips and analyzed by liquid chromatography coupled to Q Exactive instrument (Thermo Fisher Scientific)[269]. Initial data analysis was performed by using the MaxQuant software suite including the label-free algorithm for label-free quantification intensity calculation[270]. Further analysis was conducted by using the Perseus enviroment[271]. To reveal significant binders, a two-sample t-test was performed.

| Lysis Buffer |
|---|
| 50 mM HEPES pH 7.5 |
| 100 mM NaCl |
| 1 mM $MgCl_2$ |
| 1 % Triton-X100 |

| Wash Buffer |
|---|
| 50 mM HEPES pH 7.5 |
| 100 M NaCl |
| 1 mM $MgCl_2$ |

Cell cultivation, cell transduction, co-immunoprecipitation, affinity enrichment mass spectrometry measurement and analysis were performed by Annika Frauenstein and Felix Meissner (MPI of Biochemistry, Munich).

6.2.3.5. Structural biology methods

**i X-ray crystallography**

**Crystallization**

Initial high-throughput protein crystallization screening was performed using the sitting-drop vapor method using the screens listed in Table 6.9. Setups were pipetted by Crystal Phoenix Robot (Art Robbins Instruments) with 0.2 µl + 0.2 µl drop size. Promising crystallization conditions were optimized by alteration of protein concentration, reservoir conditions, pH, temperature, drop size, mixing ratio and by adding additive screen components. Refinements were set up in 24 well format either using sitting drop Cryschem plates (Hampton research) or hanging drop EasyXtal Tool plates (Crystalgen SuperClear™ Plates, Jena Bioscience). Prior to crystallization proteins were centrifuged at 25 000 x g for 10 minutes. The plates were stored at either 4°C or 20°C and checked regularly.

1 µl of $SLFN5^{1-336}$ and 1 µl of the reservoir solution were mixed and crystals were grown by the hanging drop vapor diffusion method. Crystal form 1 crystallized with 7 g/l concentration and 200 mM NaCl, 100 mM MES, pH 5.8, 20 % (v/v) PEG 2000 MME were used as reservoir conditions. Reservoir solution supplemented with 25% (v/v) ethylene glycol was used as cryo-protectant prior to flash-cooling in liquid nitrogen. The concentration of $SLFN5^{1-336}$ was set to 3.2 g/l for the second crystal form and 0.1 M sodium acetate, pH 5.0 and 1.5 M ammonium sulfate were used as reservoir. Reservoir solution supplemented with 25% (v/v) glycerol was used as cryo-protectant prior to flash cooling in liquid nitrogen.

1 µl of $Slfn2^{\Delta C116-W122}$ concentrated to 7 g/l and 1 µl of the reservoir solution containing 150 mM KCl, 10 mM $MgCl_2$, 50 mM MES pH 6.5 and 2.5% PEG 8000 were mixed and pipetted using the hanging drop vapor diffusion method. Prior to flash cooling in liquid nitrogen the reservoir solution supplemented with 15 % (v/v) 1, 3-butanediol was used as cryo-protectant.

**Data collection, processing and structure solution**

Diffraction data was collected at the beamlines X06SA (PXI) and X06DA (PXIII) (Swiss light source, Paul-Scherrer Institute, Villigen, Switzerland) at 100 K. Data was integrated and scaled with XDS[272]. The quality of the experimental data was verified by Xtriage[273] and the solvent content of the crystals were calculated according to Matthews[274].

The phases for SLFN5$^{1-336}$ were determined by performing a single-wavelength anomalous dispersion experiment using the 3.2 Å Zn-SAD dataset from the needle shaped crystals measured at 1.28 Å wavelength. For the generation of the Zn-substructure and an initial poly-alanine model HySS[275], AutoSol[276] and Autobuild[277] within the Phenix software[278] and Chainsaw[279] within the CCP4[280] package were used. The poly-alanine model was used as search model for molecular replacement using the 1.8 Å native dataset from the cubic crystals measured at 1.0 Å wavelength. This was done using Phaser[281,282] and the initial model was automatically rebuild using Autobuild[277]. The final structure of SLFN5$^{1-336}$ was solved at 1.8 Å by iterative refinement cycles in phenix.refine[283] or Refmac[284-286] and manual model building steps in COOT[287].

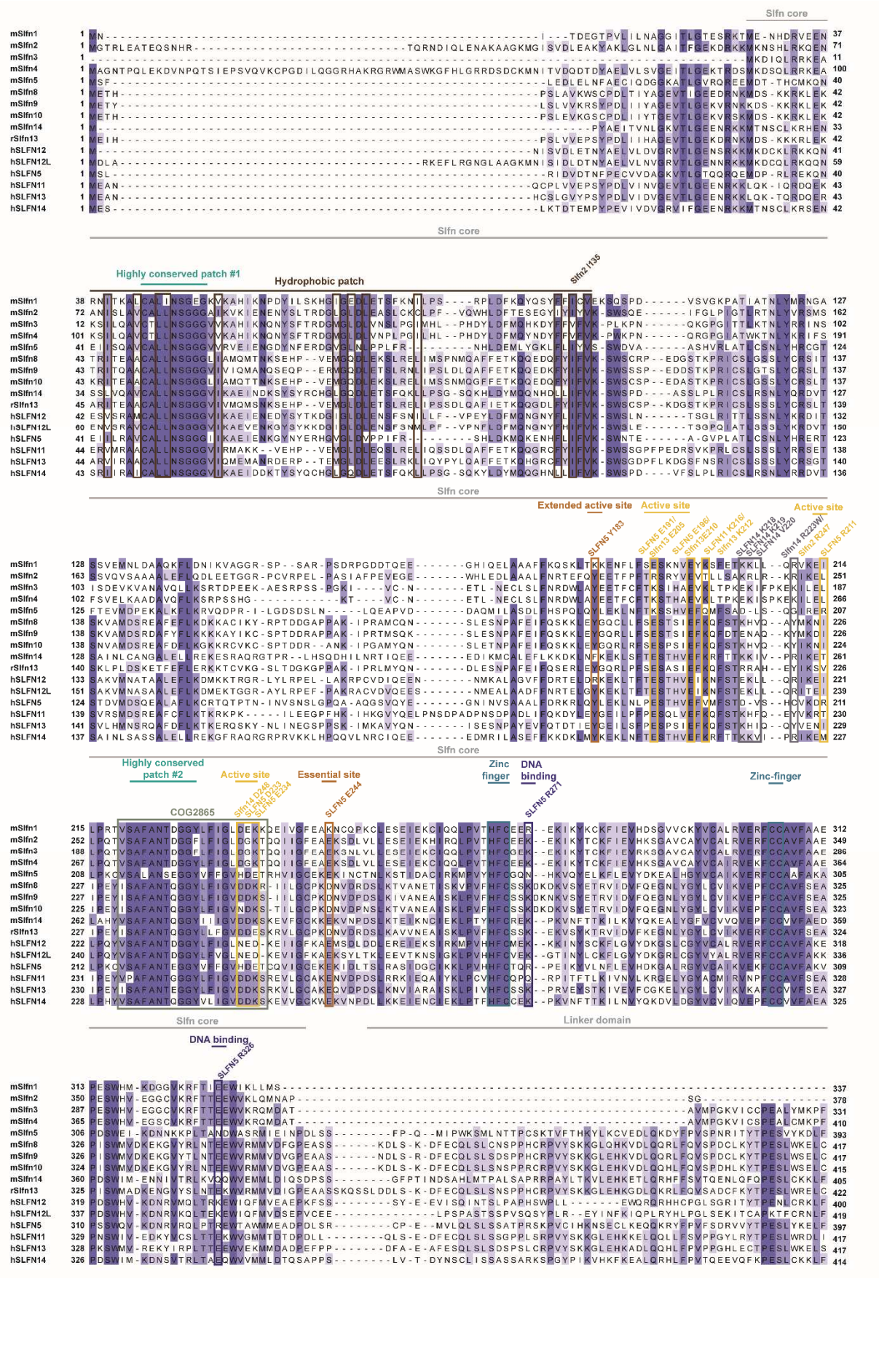
The Slfn2$^{\Delta C116-W122}$ crystal structure was solved by performing molecular replacement using SLFN5$^{1-336}$ as search model with Phaser[281,282]. The initial model was automatically rebuild using Autobuild[277]. The final structure of Slfn2$^{\Delta C116-W122}$ was solved at 4 Å by iterative refinement cycles in phenix.refine[283] or Refmac[284-286] and manual model building steps in COOT[287].

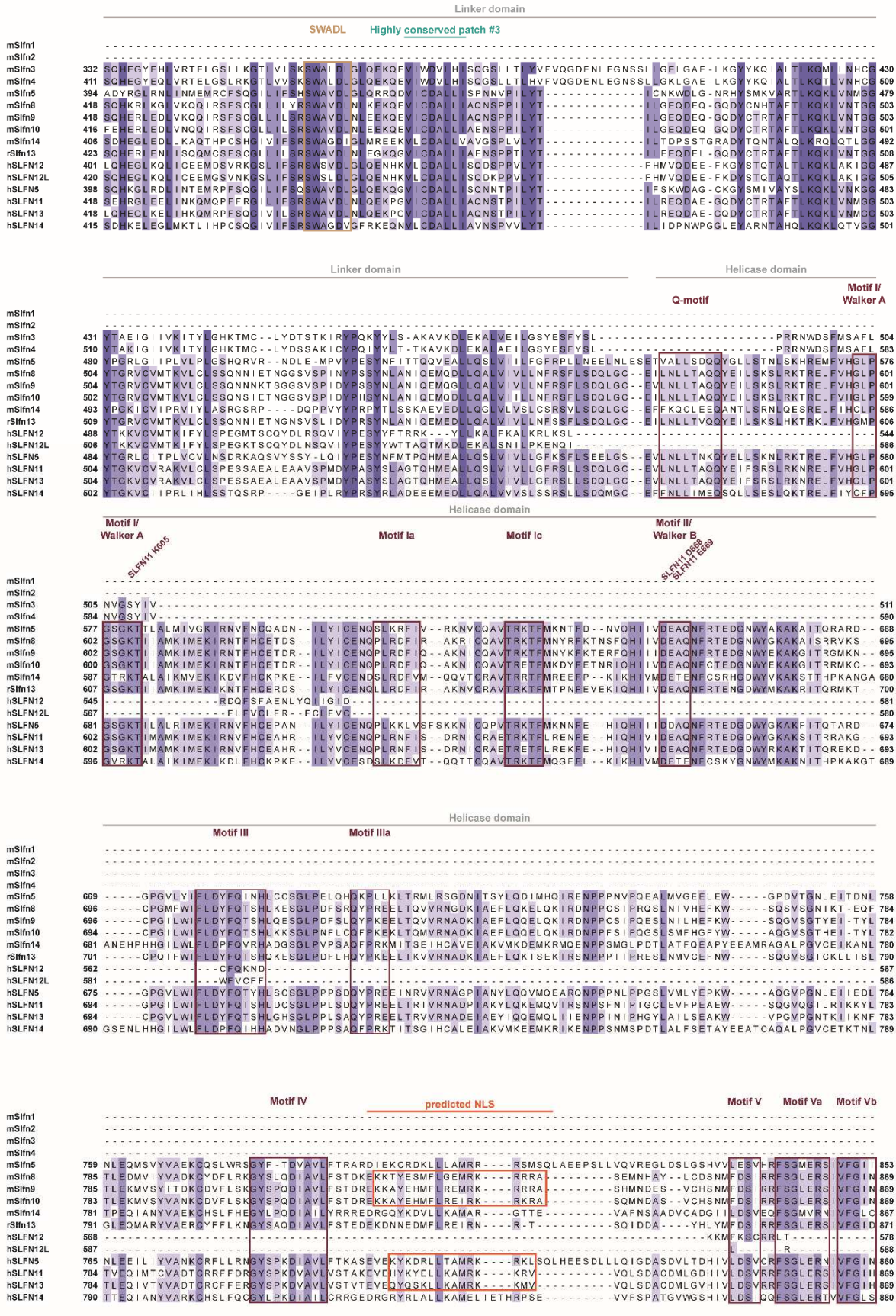
Prior to model building and refinement, 5 % of the reflections were randomly omitted for monitoring the free R value.

**ii Small-angle X-ray scattering (SAXS)**

Prior to sample measurement Slfn2$^{FL}$ was applied onto an analytical gel filtration column (Superose 6 10/300 Increase, GE Healthcare). Peak I fractions were measured at 5.5, 3.0 and 0.8 g/l and peak II fractions at 11.6, 5.2, 2.1, 1.6, 0.8 and 0.2 g/l. Slfn5$^{1-336}$ was measured with 9.2, 5.6, 2.2, 1 and 0.4 g/l. Samples were measured in batch mode and to remove potential aggregates, all samples were centrifuged prior to data collection. Measurement took place at the EMBL P12 beamline at PETRA III (DESY Hamburg, Germany). All data sets were analyzed using the ATSAS software package[288] according to[289]. *Ab initio* modeling was done with GASBORp[290]. 10 calculated models were aligned and averaged using DAMAVER[288]. Prior to model visualization a situs map was generated using situs.biomachina software[291] and visualized using UCSF Chimera[261].

# 7. Appendix


Slfn core
mSlfn1
mSlfn2
mSlfn3
mSlfn4
mSlfn5
mSlfn8
mSlfn9
mSlfn10
mSlfn14
rSlfn13
hSLFN12
hSLFN12L
hSLFN5
hSLFN11
hSLFN13
hSLFN14
Highly conserved patch #1
Hydrophobic patch
Slfn2 I135
Extended active site
Active site
SLFN5 Y183
SLFN5 E181/ Slfn13 E205
SLFN5 E196/ Slfn13E210
SLFN11 K216/ Slfn13 K212
SLFN14 K218/ SLFN14 K219/ SLFN14 V220
Slfn14 R223W/ Slfn2 R247
SLFN5 R211
Highly conserved patch #2
Active site
Essential site
Zinc finger
DNA binding
Zinc-finger
COG2865
Slfn14 D48/ SLFN5 D33/ SLFN5 E234
SLFN5 E244
SLFN5 R271
Linker domain
DNA binding
SLFN5 R326


Linker domain
SWADL
Highly conserved patch #3
Linker domain
Helicase domain
Q-motif
Motif I/
Walker A
Helicase domain
Motif I/
Walker A
SLFN11 K605
Motif Ia
Motif Ic
Motif II/
Walker B
SLFN11 D668
SLFN11 E669
Helicase domain
Motif III
Motif IIIa
Motif IV
predicted NLS
Motif V
Motif Va
Motif Vb
mSlfn1
mSlfn2
mSlfn3
mSlfn4
mSlfn5
mSlfn8
mSlfn9
mSlfn10
mSlfn14
rSlfn13
hSLFN12
hSLFN12L
hSLFN5
hSLFN11
hSLFN13
hSLFN14

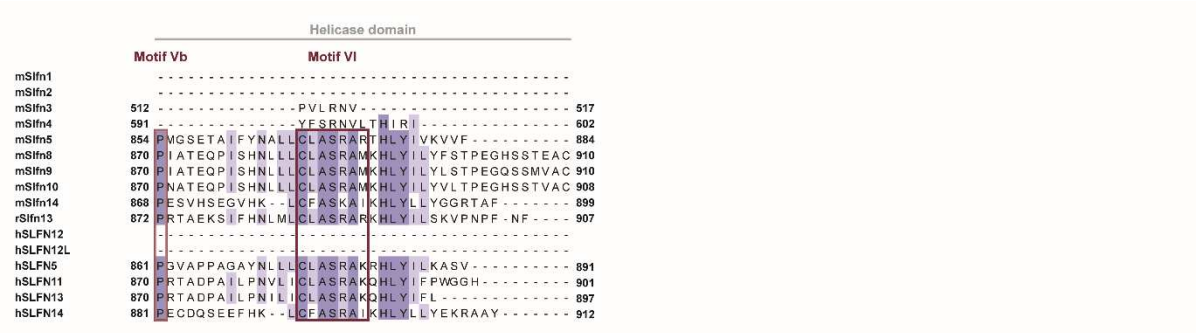



**Figure 7.1:** Multiple sequence alignments of all human and murine schlafen proteins and rat Slfn13. Conserved residues are colored to the percentage of their identity. Important residues and predicted motifs are marked respectively. The helicase motifs were assigned manually according to the consensus sequences described in[200].

% Identitiy of the respective slfn core domains

| | mSlfn1 | mSlfn2 | mSlfn3 | mSlfn4 | mSlfn5 | mSlfn8 | mSlfn9 | mSlfn10 | mSlfn14 | rSlfn13 | hSLFN12 | hSLFN12L | hSLFN5 | hSLFN11 | hSLFN13 | hSLFN14 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| mSlfn1 | 100 | 52.68 | 49.17 | 50.93 | 34.46 | 35.15 | 36.67 | 35.06 | 38.37 | 32.63 | 50.90 | 48.66 | 38.30 | 33.74 | 32.33 | 41.37 | mSlfn1 |
| mSlfn2 | 52.68 | 100 | 62.46 | 58.89 | 37.76 | 35.00 | 34.71 | 34.62 | 40.30 | 34.31 | 54.52 | 52.91 | 41.37 | 35.31 | 33.92 | 40.58 | mSlfn2 |
| mSlfn3 | 49.17 | 62.18 | 100 | 89.04 | 37.97 | 32.45 | 33.44 | 32.67 | 37.46 | 31.68 | 48.53 | 46.58 | 39.93 | 33.89 | 32.34 | 38.31 | mSlfn3 |
| mSlfn4 | 50.62 | 58.29 | 89.94 | 100 | 38.61 | 33.85 | 36.00 | 34.46 | 40.00 | 33.44 | 51.99 | 48.41 | 42.95 | 35.49 | 33.94 | 39.21 | mSlfn4 |
| mSlfn5 | 35.09 | 36.94 | 34.57 | 35.35 | 100 | 31.38 | 31.69 | 32.51 | 38.63 | 34.36 | 46.95 | 44.38 | 57.27 | 33.54 | 34.05 | 40.61 | mSlfn5 |
| mSlfn8 | 34.64 | 34.1 | 32.73 | 33.46 | 41.8 | 100 | 84.33 | 88.54 | 42.04 | 77.94 | 38.69 | 37.76 | 37.80 | 57.89 | 62.07 | 43.31 | mSlfn8 |
| mSlfn9 | 36.14 | 33.82 | 33.73 | 35.17 | 41.69 | 85.82 | 100 | 83.38 | 42.34 | 77.65 | 38.39 | 38.05 | 37.20 | 57.89 | 60.63 | 43.31 | mSlfn9 |
| mSlfn10 | 34.85 | 34.3 | 33.27 | 34.22 | 42.94 | 85.68 | 87.22 | 100 | 41.69 | 77.81 | 38.02 | 37.69 | 38.34 | 58.53 | 61.56 | 42.69 | mSlfn10 |
| mSlfn14 | 37.99 | 39.64 | 34 | 35.2 | 37.76 | 40 | 40.91 | 41.12 | 100 | 42.22 | 48.50 | 47.90 | 41.23 | 43.29 | 42.22 | 78.47 | mSlfn14 |
| rSlfn13 | 32.43 | 33.43 | 31.86 | 32.83 | 43.06 | 76.97 | 77.31 | 77.48 | 40.66 | 100 | 37.69 | 37.94 | 37.08 | 59.77 | 60.46 | 44.64 | rSlfn13 |
| hSLFN12 | 51.05 | 54.62 | 46.06 | 49.08 | 42.06 | 38.07 | 38.25 | 38.03 | 41.35 | 38 | 100 | 81.92 | 50.00 | 40.54 | 39.05 | 50.88 | hSLFN12 |
| hSLFN12L | 48.8 | 53.02 | 44.7 | 46.47 | 40 | 36.22 | 36.75 | 36.7 | 40.97 | 37.04 | 75.8 | 100 | 50.75 | 40.48 | 39.59 | 49.13 | hSLFN12L |
| hSLFN5 | 39.26 | 40.06 | 38.16 | 39.73 | 58.91 | 48.8 | 49.26 | 49.14 | 40.32 | 48.11 | 45.71 | 45.32 | 100 | 39.51 | 37.99 | 42.51 | hSLFN5 |
| hSLFN11 | 33.64 | 35.67 | 33.47 | 34.22 | 44.66 | 60.88 | 62.12 | 62.37 | 43.61 | 61.51 | 39.93 | 40.04 | 53.39 | 100 | 67.34 | 45.13 | hSLFN11 |
| hSLFN13 | 31.83 | 33.91 | 32.87 | 33.9 | 44.46 | 63.26 | 63.82 | 63.63 | 43 | 61.53 | 39.34 | 39.08 | 52 | 77.64 | 100 | 43.19 | hSLFN13 |
| hSLFN14 | 41.07 | 39.83 | 34.87 | 35.36 | 39.63 | 43 | 43.34 | 43.67 | 70.57 | 43.44 | 42.98 | 42.48 | 42.61 | 44.89 | 44.51 | 100 | hSLFN14 |
| | mSlfn1 | mSlfn2 | mSlfn3 | mSlfn4 | mSlfn5 | mSlfn8 | mSlfn9 | mSlfn10 | mSlfn14 | rSlfn13 | hSLFN12 | hSLFN12L | hSLFN5 | hSLFN11 | hSLFN13 | hSLFN14 | |

% Identitiy of the full length schlafen protein members

**Figure 7.2:** Percentage of the identity of the human and murine full length schlafen proteins (yellow) and the slfn core domains of the respective proteins (blue).

**Table 7.1:** Significantly enriched Slfn2$^{FL}$ interactors in the E.G7 cell line.

| Student's T-test Difference | Protein names | Gene names |
|---|---|---|
| 8.3799 | Pre-B-cell leukemia transcription factor-interacting protein 1 | Pbxip1 |
| 7.13899 | Netrin receptor UNC5B | Unc5b |
| 7.1084 | Metalloreductase STEAP2 | Steap2 |
| 6.74972 | | Tns1 |
| 6.46834 | Glycosylphosphatidylinositol anchor attachment 1 protein | Gpaa1 |
| 6.32674 | Phospholipase D3 | Pld3 |
| 6.14446 | Fatty acid desaturase 3 | Fads3 |
| 6.07635 | Calponin-2;Calponin | Cnn2 |
| 5.34218 | Actin. alpha cardiac muscle 1;Actin. alpha skeletal muscle;Actin. aortic smooth muscle;Actin. gamma-enteric smooth muscle | Actc1;Acta1;Acta2;Actg2 |
| 5.30529 | Ferrochelatase. mitochondrial;Ferrochelatase | Fech |
| 5.27462 | Transcription factor GATA-6 | Gata6 |
| 5.25462 | Lanosterol 14-alpha demethylase | Cyp51a1 |
| 5.14544 | Cell migration-inducing and hyaluronan-binding protein | Cemip |
| 4.98172 | Steryl-sulfatase | Sts |
| 4.9681 | CAP-Gly domain-containing linker protein 2 | Clip2 |
| 4.91423 | | Slfn2 |
| 4.8706 | Cytochrome c1. heme protein. mitochondrial | Cyc1 |
| 4.58655 | Plectin | Plec |
| 4.49572 | Sorting nexin-8 | Snx8 |
| 4.45062 | 5-AMP-activated protein kinase subunit beta-2 | Prkab2 |
| 4.41029 | Putative RNA-binding protein 15B | Rbm15b |
| 4.36436 | Protein Shroom4 | Shroom4 |
| 4.24788 | AP-2 complex subunit sigma | Ap2s1 |
| 4.09895 | 40S ribosomal protein S15 | Rps15 |
| 4.09034 | Myosin-9 | Myh9 |
| 4.08354 | Ataxin-1 | Atxn1 |
| 4.01151 | GATA zinc finger domain-containing protein 1 | Gatad1 |
| 3.96303 | | Nrg1 |
| 3.9502 | Serine/threonine-protein phosphatase PGAM5. mitochondrial | Pgam5 |
| 3.9217 | Extended synaptotagmin-1 | Esyt1 |
| 3.90007 | Angiomotin-like protein 2 | Amotl2 |
| 3.85541 | TIR domain-containing adapter molecule 1 | Ticam1 |
| 3.8536 | Interleukin enhancer-binding factor 3 | Ilf3 |
| 3.83471 | PDZ and LIM domain protein 5 | Pdlim5 |
| 3.82384 | GTP-binding protein 10 | Gtpbp10 |
| 3.79624 | Cytohesin-2 | Cyth2 |
| 3.72862 | Protein transport protein Sec61 subunit alpha isoform 1 | Sec61a1 |
| 3.70878 | Cleavage and polyadenylation specificity factor subunit 3 | Cpsf3 |
| 3.62339 | Sarcoplasmic/endoplasmic reticulum calcium ATPase 2;Calcium-transporting ATPase | Atp2a2 |
| 3.58581 | Tubulin beta-6 chain | Tubb6 |
| 3.52093 | Arf-GAP with SH3 domain. ANK repeat and PH domain-containing protein 3 | Asap3 |
| 3.42883 | Golgin subfamily A member 1 | Golga1 |
| 3.40622 | Zinc finger protein GLIS3 | Glis3 |
| 3.34584 | Ribosome-binding protein 1 | Rrbp1 |
| 3.34038 | Double-stranded RNA-binding protein Staufen homolog 1 | Stau1 |

| | | |
|---|---|---|
| 3.30172 | von Willebrand factor A domain-containing protein 9 | Vwa9 |
| 3.28862 | Peroxisome biogenesis factor 1 | Pex1 |
| 3.27998 | Rap guanine nucleotide exchange factor 3 | Rapgef3 |
| 3.23658 | Mitochondrial chaperone BCS1 | Bcs1l |
| 3.19431 | E3 ubiquitin/ISG15 ligase TRIM25 | Trim25 |
| 3.10434 | Kalirin | Kalrn |
| 3.07512 | Cyclin-dependent kinase 6 | Cdk6 |
| 3.04386 | Ubiquitin-conjugating enzyme E2 variant 3 | Uevld |
| 3.01038 | Non-specific serine/threonine protein kinase;Serine/threonine-protein kinase PAK 1 | Pak1 |
| 3.00707 | Mitochondrial 2-oxoglutarate/malate carrier protein | Slc25a11 |

**Table 7.2:** Significantly enriched Slfn2$^{FL}$ interactors in the Panc02 cell line.

| Student's T-test Difference | Protein names | Gene names |
|---|---|---|
| 6.08919 | Schlafen 2 | Slfn2 |
| 5.43593 | Steryl-sulfatase | Sts |
| 4.44549 | Ribosome production factor 2 homolog | Rpf2 |
| 4.37311 | Peptidyl-prolyl cis-trans isomerase FKBP3 | Fkbp3 |
| 4.32954 | Insulin receptor substrate 2 | Irs2 |
| 3.88687 | U1 small nuclear ribonucleoprotein 70 kDa | Snrnp70 |
| 3.77043 | U1 small nuclear ribonucleoprotein C | Snrpc |
| 3.69706 | Tyrosine--tRNA ligase. mitochondrial;Tyrosine--tRNA ligase | Yars2 |
| 3.64539 | WD repeat-containing protein 11 | Wdr11 |
| 3.40217 | Glycylpeptide N-tetradecanoyltransferase 1 | Nmt1 |
| 3.38173 | Cell division cycle and apoptosis regulator protein 1 | Ccar1 |
| 3.17901 | Protein KHNYN | Khnyn |
| 3.16921 | Na(+)/H(+) exchange regulatory cofactor NHE-RF1 | Slc9a3r1 |
| 3.1653 | Integrator complex subunit 11 | Cpsf3l |
| 3.09575 | Chromodomain-helicase-DNA-binding protein 2 | Chd2 |
| 3.05598 | Actin-related protein 5 | Actr5 |
| 3.03457 | Son of sevenless homolog 1 | Sos1 |
| 3.02374 | 28S ribosomal protein S27. mitochondrial | Mrps27 |
| 3.00621 | UPF2 regulator of nonsense transcripts homolog | Upf2 |
| 2.94632 | Structural maintenance of chromosomes protein 3 | Smc3 |
| 2.90687 | Nuclear factor related to kappa-B-binding protein | Nfrkb |
| 2.83567 | | Ddx23 |
| 2.82643 | SAP domain-containing ribonucleoprotein | Sarnp |
| 2.80921 | Transcriptional coactivator YAP1 | Yap1 |
| 2.79752 | Uncharacterized protein C19orf43 homolog | |
| 2.763 | U1 small nuclear ribonucleoprotein A | Snrpa |
| 2.61386 | Structural maintenance of chromosomes protein 1A | Smc1a |
| 2.54983 | MHC class II regulatory factor RFX1 | Rfx1 |
| 2.49592 | RNA-binding protein 10 | Rbm10 |
| 2.41059 | Bifunctional lysine-specific demethylase and histidyl-hydroxylase NO66 | No66 |

# 8. References


1. Murphy, K., Travers, P. & Walport, M. Janeway's immunobiology. 2012. *New York: Garland Science* **8**.
2. Akashi, K., Traver, D., Miyamoto, T. & Weissman, I.L. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* **404**, 193-7 (2000).
3. Aderem, A. & Underhill, D.M. Mechanisms of phagocytosis in macrophages. *Annu Rev Immunol* **17**, 593-623 (1999).
4. Savina, A. & Amigorena, S. Phagocytosis and antigen presentation in dendritic cells. *Immunol Rev* **219**, 143-56 (2007).
5. Banchereau, J. & Steinman, R.M. Dendritic cells and the control of immunity. *Nature* **392**, 245-52 (1998).
6. Unanue, E.R. Antigen-presenting function of the macrophage. *Annu Rev Immunol* **2**, 395-428 (1984).
7. Holt, P.G., Degebrodt, A., O'Leary, C., Krska, K. & Plozza, T. T cell activation by antigen-presenting cells from lung tissue digests: suppression by endogenous macrophages. *Clin Exp Immunol* **62**, 586-93 (1985).
8. Dvorak, A.M., Galli, S.J., Schulman, E.S., Lichtenstein, L.M. & Dvorak, H.F. Basophil and mast cell degranulation: ultrastructural analysis of mechanisms of mediator release. *Fed Proc* **42**, 2510-5 (1983).
9. Varricchi, G., Raap, U., Rivellese, F., Marone, G. & Gibbs, B.F. Human mast cells and basophils-How are they similar how are they different? *Immunol Rev* **282**, 8-34 (2018).
10. Kubo, M. Mast cells and basophils in allergic inflammation. *Curr Opin Immunol* **54**, 74-79 (2018).
11. Kondo, M., Weissman, I.L. & Akashi, K. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell* **91**, 661-72 (1997).
12. Vivier, E., Tomasello, E., Baratin, M., Walzer, T. & Ugolini, S. Functions of natural killer cells. *Nat Immunol* **9**, 503-10 (2008).
13. Lai, A.Y. & Kondo, M. T and B lymphocyte differentiation from hematopoietic stem cell. *Semin Immunol* **20**, 207-12 (2008).
14. LeBien, T.W. & Tedder, T.F. B lymphocytes: how they develop and function. *Blood* **112**, 1570-80 (2008).
15. Santana, M.A. & Esquivel-Guadarrama, F. Cell biology of T cell activation and differentiation. *Int Rev Cytol* **250**, 217-74 (2006).
16. Zinkernagel, R.M. et al. On immunological memory. *Annu Rev Immunol* **14**, 333-67 (1996).
17. Robey, E. & Fowlkes, B.J. Selective events in T cell development. *Annu Rev Immunol* **12**, 675-705 (1994).
18. Godfrey, D.I., Kennedy, J., Suda, T. & Zlotnik, A. A developmental pathway involving four phenotypically and functionally distinct subsets of CD3-CD4-CD8- triple-negative adult mouse thymocytes defined by CD44 and CD25 expression. *J Immunol* **150**, 4244-52 (1993).
19. Laky, K. & Fowlkes, B.J. Notch signaling in CD4 and CD8 T cell development. *Curr Opin Immunol* **20**, 197-202 (2008).

20. Klein, L., Kyewski, B., Allen, P.M. & Hogquist, K.A. Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). *Nat Rev Immunol* **14**, 377-91 (2014).

21. Fung-Leung, W.P. et al. CD8 is needed for development of cytotoxic T cells but not helper T cells. *Cell* **65**, 443-9 (1991).

22. Nurieva, R.I. et al. Generation of T follicular helper cells is mediated by interleukin-21 but independent of T helper 1, 2, or 17 cell lineages. *Immunity* **29**, 138-49 (2008).

23. Corthay, A. How do regulatory T cells work? *Scand J Immunol* **70**, 326-36 (2009).

24. Schwarz, D.A., Katayama, C.D. & Hedrick, S.M. Schlafen, a new family of growth regulatory genes that affect thymocyte development. *Immunity* **9**, 657-68 (1998).

25. Furukawa, Y. Cell cycle control genes and hematopoietic cell differentiation. *Leuk Lymphoma* **43**, 225-31 (2002).

26. Zebedee, Z. & Hara, E. Id proteins in cell cycle control and cellular senescence. *Oncogene* **20**, 8317-25 (2001).

27. Brady, G., Boggan, L., Bowie, A. & O'Neill, L.A. Schlafen-1 causes a cell cycle arrest by inhibiting induction of cyclin D1. *J Biol Chem* **280**, 30723-34 (2005).

28. Kuang, C.Y., Yang, T.H., Zhang, Y., Zhang, L. & Wu, Q. Schlafen 1 inhibits the proliferation and tube formation of endothelial progenitor cells. *PLoS One* **9**, e109711 (2014).

29. Baldin, V., Lukas, J., Marcote, M.J., Pagano, M. & Draetta, G. Cyclin D1 is a nuclear protein required for cell cycle progression in G1. *Genes Dev* **7**, 812-21 (1993).

30. Geserick, P., Kaiser, F., Klemm, U., Kaufmann, S.H. & Zerrahn, J. Modulation of T cell development and activation by novel members of the Schlafen (slfn) gene family harbouring an RNA helicase-like motif. *Int Immunol* **16**, 1535-48 (2004).

31. Zhang, Y. et al. The Hsp40 family chaperone protein DnaJB6 enhances Schlafen1 nuclear localization which is critical for promotion of cell-cycle arrest in T-cells. *Biochem J* **413**, 239-50 (2008).

32. Condamine, T. et al. Characterization of Schlafen-3 expression in effector and regulatory T cells. *J Leukoc Biol* **87**, 451-6 (2010).

33. Puck, A. et al. Expression and regulation of Schlafen (SLFN) family members in primary human monocytes, monocyte-derived dendritic cells and T cells. *Results Immunol* **5**, 23-32 (2015).

34. Isaacs, A. & Lindenmann, J. Virus interference. I. The interferon. *Proc R Soc Lond B Biol Sci* **147**, 258-67 (1957).

35. Pestka, S. The human interferon-alpha species and hybrid proteins. *Semin Oncol* **24**, S9-4-S9-17 (1997).

36. Pestka, S., Krause, C.D. & Walter, M.R. Interferons, interferon-like cytokines, and their receptors. *Immunol Rev* **202**, 8-32 (2004).

37. Pestka, S., Langer, J.A., Zoon, K.C. & Samuel, C.E. Interferons and their actions. *Annu Rev Biochem* **56**, 727-77 (1987).

38. Platanias, L.C. Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nat Rev Immunol* **5**, 375-86 (2005).

39. Bach, E.A., Aguet, M. & Schreiber, R.D. The IFN gamma receptor: a paradigm for cytokine receptor signaling. *Annu Rev Immunol* **15**, 563-91 (1997).

40. Pestka, S. et al. The interferon gamma (IFN-gamma) receptor: a paradigm for the multichain cytokine receptor. *Cytokine Growth Factor Rev* **8**, 189-206 (1997).

41. Kotenko, S.V. et al. IFN-lambdas mediate antiviral protection through a distinct class II cytokine receptor complex. *Nat Immunol* **4**, 69-77 (2003).

42. Sheppard, P. et al. IL-28, IL-29 and their class II cytokine receptor IL-28R. *Nat Immunol* **4**, 63-8 (2003).

43. Prokunina-Olsson, L. et al. A variant upstream of IFNL3 (IL28B) creating a new interferon gene IFNL4 is associated with impaired clearance of hepatitis C virus. *Nat Genet* **45**, 164-71 (2013).

44. Fox, B.A., Sheppard, P.O. & O'Hara, P.J. The role of genomic data in the discovery, annotation and evolutionary interpretation of the interferon-lambda family. *PLoS One* **4**, e4933 (2009).

45. Schneider, W.M., Chevillotte, M.D. & Rice, C.M. Interferon-stimulated genes: a complex web of host defenses. *Annu Rev Immunol* **32**, 513-45 (2014).

46. Stark, G.R. & Darnell, J.E., Jr. The JAK-STAT pathway at twenty. *Immunity* **36**, 503-14 (2012).

47. van Boxel-Dezaire, A.H., Rani, M.R. & Stark, G.R. Complex modulation of cell type-specific signaling in response to type I interferons. *Immunity* **25**, 361-72 (2006).

48. Heim, M.H., Kerr, I.M., Stark, G.R. & Darnell, J.E., Jr. Contribution of STAT SH2 groups to specific interferon signaling by the Jak-STAT pathway. *Science* **267**, 1347-9 (1995).

49. Levy, D.E., Kessler, D.S., Pine, R. & Darnell, J.E., Jr. Cytoplasmic activation of ISGF3, the positive regulator of interferon-alpha-stimulated transcription, reconstituted in vitro. *Genes Dev* **3**, 1362-71 (1989).

50. Levy, D., Larner, A., Chaudhuri, A., Babiss, L.E. & Darnell, J.E., Jr. Interferon-stimulated transcription: isolation of an inducible gene and identification of its regulatory region. *Proc Natl Acad Sci U S A* **83**, 8929-33 (1986).

51. Levy, D.E., Kessler, D.S., Pine, R., Reich, N. & Darnell, J.E., Jr. Interferon-induced nuclear factors that bind a shared promoter element correlate with positive and negative transcriptional control. *Genes Dev* **2**, 383-93 (1988).

52. Katsoulidis, E. et al. Role of Schlafen 2 (SLFN2) in the generation of interferon alpha-induced growth inhibitory responses. *J Biol Chem* **284**, 25051-64 (2009).

53. Mavrommatis, E. et al. Expression and regulatory effects of murine Schlafen (Slfn) genes in malignant melanoma and renal cell carcinoma. *J Biol Chem* **288**, 33006-15 (2013).

54. Sassano, A. et al. Human Schlafen 5 (SLFN5) Is a Regulator of Motility and Invasiveness of Renal Cell Carcinoma Cells. *Mol Cell Biol* **35**, 2684-98 (2015).

55. Katsoulidis, E. et al. Role of interferon {alpha} (IFN{alpha})-inducible Schlafen-5 in regulation of anchorage-independent growth and invasion of malignant melanoma cells. *J Biol Chem* **285**, 40333-41 (2010).

56. Ronnblom, L. The type I interferon system in the etiopathogenesis of autoimmune diseases. *Ups J Med Sci* **116**, 227-37 (2011).

57. Hong, X.X. & Carmichael, G.G. Innate immunity in pluripotent human cells: attenuated response to interferon-beta. *J Biol Chem* **288**, 16196-205 (2013).

58. Malakhov, M.P., Malakhova, O.A., Kim, K.I., Ritchie, K.J. & Zhang, D.E. UBP43 (USP18) specifically removes ISG15 from conjugated proteins. *J Biol Chem* **277**, 9976-81 (2002).

59. Francois-Newton, V. et al. USP18-based negative feedback control is induced by type I and type III interferons and specifically inactivates interferon alpha response. *PLoS One* **6**, e22200 (2011).

60. Makowska, Z., Duong, F.H., Trincucci, G., Tough, D.F. & Heim, M.H. Interferon-beta and interferon-lambda signaling is not affected by interferon-induced refractoriness to interferon-alpha in vivo. *Hepatology* **53**, 1154-63 (2011).

61. Arslan, A.D. et al. Human SLFN5 is a transcriptional co-repressor of STAT1-mediated interferon responses and promotes the malignant phenotype in glioblastoma. *Oncogene* **36**, 6006-6019 (2017).

62. Schoggins, J.W. Interferon-stimulated genes: roles in viral pathogenesis. *Curr Opin Virol* **6**, 40-6 (2014).

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

64. Gao, S. et al. Structural basis of oligomerization in the stalk region of dynamin-like MxA. *Nature* **465**, 502-6 (2010).

65. Kane, M. et al. MX2 is an interferon-induced inhibitor of HIV-1 infection. *Nature* **502**, 563-6 (2013).

66. Goujon, C. et al. Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection. *Nature* **502**, 559-62 (2013).

67. Liu, Z. et al. The interferon-inducible MxB protein inhibits HIV-1 infection. *Cell Host Microbe* **14**, 398-410 (2013).

68. Park, K. & Scott, A.L. Cholesterol 25-hydroxylase production by dendritic cells and macrophages is regulated by type I interferons. *J Leukoc Biol* **88**, 1081-7 (2010).

69. Li, K. et al. IFITM proteins restrict viral membrane hemifusion. *PLoS Pathog* **9**, e1003124 (2013).

70. Ozato, K., Shin, D.M., Chang, T.H. & Morse, H.C., 3rd. TRIM family proteins and their emerging roles in innate immunity. *Nat Rev Immunol* **8**, 849-60 (2008).

71. Campbell, E.M., Perez, O., Anderson, J.L. & Hope, T.J. Visualization of a proteasome-independent intermediate during restriction of HIV-1 by rhesus TRIM5alpha. *J Cell Biol* **180**, 549-61 (2008).

72. Barr, S.D., Smiley, J.R. & Bushman, F.D. The interferon response inhibits HIV particle production by induction of TRIM22. *PLoS Pathog* **4**, e1000007 (2008).

73. Goldstone, D.C. et al. HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. *Nature* **480**, 379-82 (2011).

74. Wang, X., Hinson, E.R. & Cresswell, P. The interferon-inducible protein viperin inhibits influenza virus release by perturbing lipid rafts. *Cell Host Microbe* **2**, 96-105 (2007).

75. Nasr, N. et al. HIV-1 infection of human macrophages directly induces viperin which inhibits viral production. *Blood* **120**, 778-88 (2012).

76. Perez-Caballero, D. et al. Tetherin inhibits HIV-1 release by directly tethering virions to cells. *Cell* **139**, 499-511 (2009).

77. Rehwinkel, J. Mouse knockout models for HIV-1 restriction factors. *Cell Mol Life Sci* **71**, 3749-66 (2014).

78. Li, M. et al. Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11. *Nature* **491**, 125-8 (2012).

79. van Weringh, A. et al. HIV-1 modulates the tRNA pool to improve translation efficiency. *Mol Biol Evol* **28**, 1827-34 (2011).

80. Abdel-Mohsen, M. et al. Expression profile of host restriction factors in HIV-1 elite controllers. *Retrovirology* **10**, 106 (2013).

81. Abdel-Mohsen, M. et al. Select host restriction factors are associated with HIV persistence during antiretroviral therapy. *AIDS* **29**, 411-20 (2015).

82. Razzak, M. Genetics: Schlafen 11 naturally blocks HIV. *Nat Rev Urol* **9**, 605 (2012).

83. Lin, Y.Z. et al. Equine schlafen 11 restricts the production of equine infectious anemia virus via a codon usage-dependent mechanism. *Virology* **495**, 112-21 (2016).

84. Stabell, A.C. et al. Non-human Primate Schlafen11 Inhibits Production of Both Host and Viral Proteins. *PLoS Pathog* **12**, e1006066 (2016).

85. Eskra, L., Mathison, A. & Splitter, G. Microarray analysis of mRNA levels from RAW264.7 macrophages infected with Brucella abortus. *Infect Immun* **71**, 1125-33 (2003).

86. Berger, M. et al. An Slfn2 mutation causes lymphoid and myeloid immunodeficiency due to loss of immune cell quiescence. *Nat Immunol* **11**, 335-43 (2010).

87. Puck, A. et al. The soluble cytoplasmic tail of CD45 (ct-CD45) in human plasma contributes to keep T cells in a quiescent state. *Eur J Immunol* **47**, 193-205 (2017).

88. Goldshtein, A. et al. Loss of T-cell quiescence by targeting Slfn2 prevents the development and progression of T-ALL. *Oncotarget* **7**, 46835-46847 (2016).

89. Omar, I., Lapenna, A., Cohen-Daniel, L., Tirosh, B. & Berger, M. Schlafen2 mutation unravels a role for chronic ER stress in the loss of T cell quiescence. *Oncotarget* **7**, 39396-39407 (2016).

90. Lee, N.K., Choi, H.K., Yoo, H.J., Shin, J. & Lee, S.Y. RANKL-induced schlafen2 is a positive regulator of osteoclastogenesis. *Cell Signal* **20**, 2302-8 (2008).

91. Fujikado, N., Saijo, S. & Iwakura, Y. Identification of arthritis-related gene clusters by microarray analysis of two independent mouse models for rheumatoid arthritis. *Arthritis Res Ther* **8**, R100 (2006).

92. Kovalenko, P.L. et al. Regulation of epithelial differentiation in rat intestine by intraluminal delivery of an adenoviral vector or silencing RNA coding for Schlafen 3. *PLoS One* **8**, e79745 (2013).

93. Oh, P.S. et al. Schlafen-3 decreases cancer stem cell marker expression and autocrine/juxtacrine signaling in FOLFOX-resistant colon cancer cells. *Am J Physiol Gastrointest Liver Physiol* **301**, G347-55 (2011).

94. Patel, B.B. et al. Schlafen 3, a novel gene, regulates colonic mucosal growth during aging. *Am J Physiol Gastrointest Liver Physiol* **296**, G955-62 (2009).

95. Patel, V.B., Yu, Y., Das, J.K., Patel, B.B. & Majumdar, A.P. Schlafen-3: a novel regulator of intestinal differentiation. *Biochem Biophys Res Commun* **388**, 752-6 (2009).

96. Walsh, M.F., Hermann, R., Lee, J.H., Chaturvedi, L. & Basson, M.D. Schlafen 3 Mediates the Differentiating Effects of Cdx2 in Rat IEC-Cdx2L1 Enterocytes. *J Invest Surg* **28**, 202-7 (2015).

97. Walsh, M.F., Hermann, R., Sun, K. & Basson, M.D. Schlafen 3 changes during rat intestinal maturation. *Am J Surg* **204**, 598-601 (2012).

98. van Zuylen, W.J. et al. Macrophage activation and differentiation signals regulate schlafen-4 gene expression: evidence for Schlafen-4 as a modulator of myelopoiesis. *PLoS One* **6**, e15723 (2011).

99. El-Zaatari, M. et al. Gli1 deletion prevents Helicobacter-induced gastric metaplasia and expansion of myeloid cell subsets. *PLoS One* **8**, e58935 (2013).

100. Ding, L. et al. Schlafen 4-expressing myeloid-derived suppressor cells are induced during murine gastric metaplasia. *J Clin Invest* **126**, 2867-80 (2016).

101. Kovalenko, P.L. & Basson, M.D. Schlafen 12 expression modulates prostate cancer cell differentiation. *J Surg Res* **190**, 177-84 (2014).

102. Basson, M.D. et al. Schlafen 12 Interaction with SerpinB12 and Deubiquitylases Drives Human Enterocyte Differentiation. *Cell Physiol Biochem* **48**, 1274-1290 (2018).

103. Tian, L. et al. Schlafen-11 sensitizes colorectal carcinoma cells to irinotecan. *Anticancer Drugs* **25**, 1175-81 (2014).

104. Zoppoli, G. et al. Putative DNA/RNA helicase Schlafen-11 (SLFN11) sensitizes cancer cells to DNA-damaging agents. *Proc Natl Acad Sci U S A* **109**, 15030-5 (2012).

105. de Waal, L. et al. Identification of cancer-cytotoxic modulators of PDE3A by predictive chemogenomics. *Nat Chem Biol* **12**, 102-8 (2016).

106. Barretina, J. et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603-7 (2012).

107. Nogales, V. et al. Epigenetic inactivation of the putative DNA/RNA helicase SLFN11 in human cancer confers resistance to platinum drugs. *Oncotarget* **7**, 3084-97 (2016).

108. Lok, B.H. et al. PARP Inhibitor Activity Correlates with SLFN11 Expression and Demonstrates Synergy with Temozolomide in Small Cell Lung Cancer. *Clin Cancer Res* **23**, 523-535 (2017).

109. Murai, J. et al. Resistance to PARP inhibitors by SLFN11 inactivation can be overcome by ATR inhibition. *Oncotarget* **7**, 76534-76550 (2016).

110. Gardner, E.E. et al. Chemosensitive Relapse in Small Cell Lung Cancer Proceeds through an EZH2-SLFN11 Axis. *Cancer Cell* **31**, 286-299 (2017).

111. Seong, R.K. et al. Schlafen 14 (SLFN14) is a novel antiviral factor involved in the control of viral replication. *Immunobiology* **222**, 979-988 (2017).

112. Chen, C. et al. Uhrf1 regulates germinal center B cell expansion and affinity maturation to control viral infection. *J Exp Med* **215**, 1437-1448 (2018).

113. He, T. et al. Methylation of SLFN11 is a marker of poor prognosis and cisplatin resistance in colorectal cancer. *Epigenomics* **9**, 849-862 (2017).

114. Bustos, O. et al. Evolution of the Schlafen genes, a gene family associated with embryonic lethality, meiotic drive, immune processes and orthopoxvirus virulence. *Gene* **447**, 1-11 (2009).

115. Lund, S. et al. The dynamics of the LPS triggered inflammatory response of murine microglia under different culture and in vivo conditions. *J Neuroimmunol* **180**, 71-87 (2006).

116. Welch, J.S., Ricote, M., Akiyama, T.E., Gonzalez, F.J. & Glass, C.K. PPARgamma and PPARdelta negatively regulate specific subsets of lipopolysaccharide and IFN-gamma target genes in macrophages. *Proc Natl Acad Sci U S A* **100**, 6712-7 (2003).

117. Jang, S. et al. Bleomycin Inhibits Proliferation via Schlafen-Mediated Cell Cycle Arrest in Mouse Alveolar Epithelial Cells. *Tuberc Respir Dis (Seoul)* (2018).

118. Sohn, W.J. et al. Novel transcriptional regulation of the schlafen-2 gene in macrophages in response to TLR-triggered stimulation. *Mol Immunol* **44**, 3273-82 (2007).

119. Fischietti, M. et al. Slfn2 Regulates Type I Interferon Responses by Modulating the NFkB Pathway. *Mol Cell Biol* (2018).

120. Omar, I. et al. Schlafen2 mutation in mice causes an osteopetrotic phenotype due to a decrease in the number of osteoclast progenitors. *Sci Rep* **8**, 13005 (2018).

121. Omar, I. et al. Slfn2 mutation-induced loss of T-cell quiescence leads to elevated de novo sterol synthesis. *Immunology* **152**, 484-493 (2017).

122. Kovalenko, P.L. & Basson, M.D. The correlation between the expression of differentiation markers in rat small intestinal mucosa and the transcript levels of schlafen 3. *JAMA Surg* **148**, 1013-9 (2013).

123. Yuan, L., Yu, Y., Sanders, M.A., Majumdar, A.P. & Basson, M.D. Schlafen 3 induction by cyclic strain regulates intestinal epithelial differentiation. *Am J Physiol Gastrointest Liver Physiol* **298**, G994-G1003 (2010).

124. Rodriguez, P.Q. et al. Knockdown of Tmem234 in zebrafish results in proteinuria. *Am J Physiol Renal Physiol* **309**, F955-66 (2015).

125. Nakagawa, K. et al. Schlafen-8 is essential for lymphatic endothelial cell activation in experimental autoimmune encephalomyelitis. *Int Immunol* **30**, 69-78 (2018).

126. Yang, J.Y. et al. Structure of Schlafen13 reveals a new class of tRNA/rRNA- targeting RNase engaged in translational control. *Nat Commun* **9**, 1165 (2018).

127. Kroczynska, B., Sharma, B., Eklund, E.A., Fish, E.N. & Platanias, L.C. Regulatory effects of programmed cell death 4 (PDCD4) protein in interferon (IFN)-stimulated gene expression and generation of type I IFN responses. *Mol Cell Biol* **32**, 2809-22 (2012).

128. Companioni Napoles, O. et al. SCHLAFEN 5 expression correlates with intestinal metaplasia that progresses to gastric cancer. *J Gastroenterol* **52**, 39-49 (2017).

129. Yatim, A. et al. NOTCH1 nuclear interactome reveals key regulators of its transcriptional activity and oncogenic function. *Mol Cell* **48**, 445-58 (2012).

130. Allison Stewart, C. et al. Dynamic variations in epithelial-to-mesenchymal transition (EMT), ATM, and SLFN11 govern response to PARP inhibitors and cisplatin in small cell lung cancer. *Oncotarget* **8**, 28575-28587 (2017).

131. Pietanza, M.C. et al. Randomized, Double-Blind, Phase II Study of Temozolomide in Combination With Either Veliparib or Placebo in Patients With Relapsed-Sensitive or Refractory Small-Cell Lung Cancer. *J Clin Oncol*, JCO2018777672 (2018).

132. Murai, J. et al. SLFN11 Blocks Stressed Replication Forks Independently of ATR. *Mol Cell* **69**, 371-384 e6 (2018).

133. Mu, Y. et al. SLFN11 inhibits checkpoint maintenance and homologous recombination repair. *EMBO Rep* **17**, 94-109 (2016).

134. Nazir, M. et al. Targeting tumor cells based on Phosphodiesterase 3A expression. *Exp Cell Res* **361**, 308-315 (2017).

135. North, M.L. et al. Blood and nasal epigenetics correlate with allergic rhinitis symptom development in the environmental exposure unit. *Allergy* (2017).

136. Pisareva, V.P., Muslimov, I.A., Tcherepanov, A. & Pisarev, A.V. Characterization of Novel Ribosome-Associated Endoribonuclease SLFN14 from Rabbit Reticulocytes. *Biochemistry* **54**, 3286-301 (2015).

137. Fletcher, S.J. et al. SLFN14 mutations underlie thrombocytopenia with excessive bleeding and platelet secretion defects. *J Clin Invest* **125**, 3600-5 (2015).

138. Marconi, C. et al. SLFN14-related thrombocytopenia: identification within a large series of patients with inherited thrombocytopenia. *Thromb Haemost* **115**, 1076-9 (2016).

139. Fletcher, S.J. et al. Role of the novel endoribonuclease SLFN14 and its disease-causing mutations in ribosomal degradation. *RNA* **24**, 939-949 (2018).

140. Gubser, C. et al. Camelpox virus encodes a schlafen-like protein that affects orthopoxvirus virulence. *J Gen Virol* **88**, 1667-76 (2007).

141. Chaturvedi, L., Sun, K., Walsh, M.F., Kuhn, L.A. & Basson, M.D. The P-loop region of Schlafen 3 acts within the cytosol to induce differentiation of human Caco-2 intestinal epithelial cells. *Biochim Biophys Acta* **1843**, 3029-37 (2014).

142. Neumann, B., Zhao, L., Murphy, K. & Gonda, T.J. Subcellular localization of the Schlafen protein family. *Biochem Biophys Res Commun* **370**, 62-6 (2008).

143. Osborne, C.S. et al. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet* **36**, 1065-71 (2004).

144. Spector, D.L., Fu, X.D. & Maniatis, T. Associations between distinct pre-mRNA splicing components and the cell nucleus. *EMBO J* **10**, 3467-81 (1991).

145. de la Casa-Esperon, E. From mammals to viruses: the Schlafen genes in developmental, proliferative and immune processes. *Biomol Concepts* **2**, 159-69 (2011).

146. Horton, M.R. & Powell, J.D. Quieting T cells with Slfn2. *Nat Immunol* **11**, 281-2 (2010).

147. Hua, X. & Thompson, C.B. Quiescent T cells: actively maintaining inactivity. *Nat Immunol* **2**, 1097-8 (2001).

148. Modiano, J.F., Johnson, L.D. & Bellgrau, D. Negative regulators in homeostasis of naive peripheral T cells. *Immunol Res* **41**, 137-53 (2008).

149. Tzachanis, D. & Boussiotis, V.A. Tob, a member of the APRO family, regulates immunological quiescence and tumor suppression. *Cell Cycle* **8**, 1019-25 (2009).

150. Tzachanis, D., Lafuente, E.M., Li, L. & Boussiotis, V.A. Intrinsic and extrinsic regulation of T lymphocyte quiescence. *Leuk Lymphoma* **45**, 1959-67 (2004).

151. Bensinger, S.J. et al. LXR signaling couples sterol metabolism to proliferation in the acquired immune response. *Cell* **134**, 97-111 (2008).

152. Brown, M.S. & Goldstein, J.L. Cholesterol feedback: from Schoenheimer's bottle to Scap's MELADL. *J Lipid Res* **50 Suppl**, S15-27 (2009).

153. Sousa, F.G. et al. Alterations of DNA repair genes in the NCI-60 cell lines and their predictive value for anticancer drug activity. *DNA Repair (Amst)* **28**, 107-15 (2015).

154. Tang, S.W. et al. Overcoming Resistance to DNA-Targeted Agents by Epigenetic Activation of Schlafen 11 (SLFN11) Expression with Class I Histone Deacetylase Inhibitors. *Clin Cancer Res* **24**, 1944-1953 (2018).

155. Deng, Y. et al. High SLFN11 expression predicts better survival for patients with KRAS exon 2 wild type colorectal cancer after treated with adjuvant oxaliplatin-based treatment. *BMC Cancer* **15**, 833 (2015).

156. Pommier, Y. Topoisomerase I inhibitors: camptothecins and beyond. *Nat Rev Cancer* **6**, 789-802 (2006).

157. Fu, D., Calvo, J.A. & Samson, L.D. Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat Rev Cancer* **12**, 104-20 (2012).

158. Daher, G.C., Harris, B.E. & Diasio, R.B. Metabolism of pyrimidine analogues and their nucleosides. *Pharmacol Ther* **48**, 189-222 (1990).

159. Lin, K.Y. & Kraus, W.L. PARP Inhibitors for Cancer Therapy. *Cell* **169**, 183 (2017).

160. Ray Chaudhuri, A. & Nussenzweig, A. The multifaceted roles of PARP1 in DNA repair and chromatin remodelling. *Nat Rev Mol Cell Biol* **18**, 610-621 (2017).

161. McKeage, M.J. Comparative adverse effect profiles of platinum drugs. *Drug Saf* **13**, 228-44 (1995).

162. Puisset, F., Schmitt, A. & Chatelut, E. Standardization of chemotherapy and individual dosing of platinum compounds. *Anticancer Res* **34**, 465-70 (2014).

163. Ballestrero, A. et al. Report on the first SLFN11 monothematic workshop: from function to role as a biomarker in cancer. *J Transl Med* **15**, 199 (2017).

164. Esteller, M., Corn, P.G., Baylin, S.B. & Herman, J.G. A gene hypermethylation profile of human cancer. *Cancer Res* **61**, 3225-9 (2001).

165. Moutinho, C. et al. Epigenetic inactivation of the BRCA1 interactor SRBC and resistance to oxaliplatin in colorectal cancer. *J Natl Cancer Inst* **106**, djt322 (2014).

166. Heyn, H. & Esteller, M. DNA methylation profiling in the clinic: applications and challenges. *Nat Rev Genet* **13**, 679-92 (2012).

167. Kalemkerian, G.P. et al. Small cell lung cancer. *J Natl Compr Canc Netw* **11**, 78-98 (2013).

168. Vire, E. et al. The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**, 871-4 (2006).

169. Murai, J. et al. Trapping of PARP1 and PARP2 by Clinical PARP Inhibitors. *Cancer Res* **72**, 5588-99 (2012).

170. Branzei, D. & Foiani, M. Regulation of DNA repair throughout the cell cycle. *Nat Rev Mol Cell Biol* **9**, 297-308 (2008).

171. Szambowska, A. et al. Cdc45-induced loading of human RPA onto single-stranded DNA. *Nucleic Acids Res* **45**, 3217-3230 (2017).

172. Fragkos, M., Ganier, O., Coulombe, P. & Mechali, M. DNA replication origin activation in space and time. *Nat Rev Mol Cell Biol* **16**, 360-74 (2015).

173. Feijoo, C. et al. Activation of mammalian Chk1 during DNA replication arrest: a role for Chk1 in the intra-S phase checkpoint monitoring replication origin firing. *J Cell Biol* **154**, 913-23 (2001).

174. Heinzmann, G. GPC/SEC mit Dreifachdetektion Tipps & Tricks Ausgabe Nr. 16; https://www.analytik-news.de/Fachartikel/Volltext/GPC_SEC16.pdf. (2008; opened: 15.09.2018).

175. Jeong, S.Y., Kumagai, A., Lee, J. & Dunphy, W.G. Phosphorylated claspin interacts with a phosphate-binding site in the kinase domain of Chk1 during ATR-mediated activation. *J Biol Chem* **278**, 46782-8 (2003).

176. Zhang, X. et al. The cytosolic DNA sensor cGAS forms an oligomeric complex with DNA and undergoes switch-like conformational changes in the activation loop. *Cell Rep* **6**, 421-30 (2014).

177. Tanaka, Y. & Chen, Z.J. STING specifies IRF3 phosphorylation by TBK1 in the cytosolic DNA signaling pathway. *Sci Signal* **5**, ra20 (2012).

178. Klug, A. & Rhodes, D. 'Zinc fingers': a novel protein motif for nucleic acid recognition. *Trends in Biochemical Sciences* **12**, 464-469 (1987).

179. Schroeder, L.A. et al. Evidence for a Tyrosine–Adenine Stacking Interaction and for a Short-lived Open Intermediate Subsequent to Initial Binding of Escherichia coli RNA Polymerase to Promoter DNA. *Journal of Molecular Biology* **385**, 339-349 (2009).

180. UniProtKB - Q9JJE7 (FADS3_MOUSE); https://www.uniprot.org/uniprot/Q9JJE7. (opened: 14.09.2018).

181. UniProtKB - P50427 (STS_MOUSE); https://www.uniprot.org/uniprot/P50427. (opened: 14.09.2018).

182. UniProtKB - O35405 (PLD3_MOUSE); https://www.uniprot.org/uniprot/O35405. (opened: 14.09.2018).

183. UniProtKB - Q3U7R1 (ESYT1_MOUSE); https://www.uniprot.org/uniprot/Q3U7R1. (opened: 14.09.2018).

184. UniProtKB - Q9WTK3 (GPAA1_MOUSE); https://www.uniprot.org/uniprot/Q9WTK3. (opened: 14.09.2018).

185. UniProtKB - Q3TVI8 (PBIP1_MOUSE); https://www.uniprot.org/uniprot/Q3TVI8. (opened: 14.09.2018).

186. UniProtKB - P68033 (ACTC_MOUSE); https://www.uniprot.org/uniprot/P68033. (opened: 14.09.2018).

187. UniProtKB - Q9Z0H8 (CLIP2_MOUSE); https://www.uniprot.org/uniprot/Q9Z0H8. (opened: 14.09.2018).

188. UniProtKB - Q08093 (CNN2_MOUSE); https://www.uniprot.org/uniprot/Q08093. (opened: 14.09.2018).

189. UniProtKB - Q61510 (TRI25_MOUSE); https://www.uniprot.org/uniprot/Q61510. (opened: 14.09.2018).

190. UniProtKB - Q6PAM0 (AAKB2_MOUSE); https://www.uniprot.org/uniprot/Q6PAM0. (opened: 14.09.2018).

191. UniProtKB - Q9WTI7 (MYO1C_MOUSE); https://www.uniprot.org/uniprot/Q9WTI7. (opened: 14.09.2918).

192. UniProtKB - O88643 (PAK1_MOUSE); https://www.uniprot.org/uniprot/O88643. (opened: 14.09.2018).

193. UniProtKB - Q8K1S3 (UNC5B_MOUSE); https://www.uniprot.org/uniprot/Q8K1S3. (opened: 14.09.2018).

194. UniProtKB - Q8BWB6 (STEA2_MOUSE); https://www.uniprot.org/uniprot/Q8BWB6. (opened: 14.09.2018).

195. UniProtKB - Q9CWR7 (STEA1_MOUSE); https://www.uniprot.org/uniprot/Q9CWR7. (opened: 14.09.2018).

196. UniProtKB - Q61169 (GATA6_MOUSE); https://www.uniprot.org/uniprot/Q61169. (opened: 14.09.2018).

197. UniProtKB - P97447 (FHL1_MOUSE); https://www.uniprot.org/uniprot/P97447. (opened: 14.09.2018).

198. UniProtKB - Q9D0M3 (CY1_MOUSE); https://www.uniprot.org/uniprot/Q9D0M3. (opened: 14.09.2018).

199. UniProtKB - P22315 (HEMH_MOUSE); https://www.uniprot.org/uniprot/P22315. (opened: 14.09.2018).

200. Fairman-Williams, M.E., Guenther, U.-P. & Jankowsky, E. SF1 and SF2 helicases: family matters. *Current Opinion in Structural Biology* **20**, 313-324 (2010).

201. Lässig, C. et al. ATP hydrolysis by the viral RNA sensor RIG-I prevents unintentional recognition of self-RNA. *Elife* **4**, e10859 (2015).

202. Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545-9 (2010).

203. Singleton, M.R., Scaife, S. & Wigley, D.B. Structural analysis of DNA replication fork reversal by RecG. *Cell* **107**, 79-89 (2001).

204. Callaghan, A.J. et al. Structure of Escherichia coli RNase E catalytic domain and implications for RNA turnover. *Nature* **437**, 1187 (2005).

205. Fu, Q. & Yuan, Y.A. Structural insights into RISC assembly facilitated by dsRNA-binding domains of human RNA helicase A (DHX9). *Nucleic Acids Res* **41**, 3457-70 (2013).

206. Kimber, M.S. et al. Structural Basis for Specificity Switching of the Src SH2 Domain. *Molecular Cell* **5**, 1043-1049 (2000).

207. Mackie, G.A. RNase E: at the interface of bacterial RNA processing and decay. *Nat Rev Microbiol* **11**, 45-57 (2013).

208. Feng, Y., Vickers, T.A. & Cohen, S.N. The catalytic domain of RNase E shows inherent 3' to 5' directionality in cleavage site selection. *Proc Natl Acad Sci U S A* **99**, 14746-51 (2002).

209. Chakraborty, P. & Grosse, F. Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair (Amst)* **10**, 654-65 (2011).

210. Lee, T. & Pelletier, J. The biology of DHX9 and its potential as a therapeutic target. *Oncotarget* **7**, 42716-42739 (2016).

211. Aktas, T. et al. DHX9 suppresses RNA processing defects originating from the Alu invasion of the human genome. *Nature* **544**, 115-119 (2017).

212. Chen, X. et al. Crystal Structure of a Tyrosine Phosphorylated STAT-1 Dimer Bound to DNA. *Cell* **93**, 827-839 (1998).

213. Walker, J.E., Saraste, M., Runswick, M.J. & Gay, N.J. Distantly related sequences in the alpha-and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *The EMBO journal* **1**, 945-951 (1982).

214. Hauk, G., McKnight, J.N., Nodelman, I.M. & Bowman, G.D. The Chromodomains of the Chd1 Chromatin Remodeler Regulate DNA Access to the ATPase Motor. *Molecular Cell* **39**, 711-723 (2010).

215. Liu, J. et al. Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control. *Mol Cell* **6**, 637-48 (2000).

216. Nelson, D.L., Lehninger, A.L. & Cox, M.M. *Lehninger principles of biochemistry*, (Macmillan, 2008).

217. Miller, J., McLachlan, A.D. & Klug, A. Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. *EMBO J* **4**, 1609-14 (1985).

218. Laity, J.H., Lee, B.M. & Wright, P.E. Zinc finger proteins: new insights into structural and functional diversity. *Curr Opin Struct Biol* **11**, 39-46 (2001).

219. Brewer, B.Y., Malicka, J., Blackshear, P.J. & Wilson, G.M. RNA sequence elements required for high affinity binding by the zinc finger domain of tristetraprolin: conformational changes coupled to the bipartite nature of Au-rich MRNA-destabilizing motifs. *J Biol Chem* **279**, 27870-7 (2004).

220. Pelham, H.R. & Brown, D.D. A specific transcription factor that can bind either the 5S RNA gene or 5S RNA. *Proc Natl Acad Sci U S A* **77**, 4170-4 (1980).

221. Clemens, K.R. et al. Molecular basis for specific recognition of both RNA and DNA by a zinc finger protein. *Science* **260**, 530-3 (1993).

222. Bogenhagen, D.F. Proteolytic footprinting of transcription factor TFIIIA reveals different tightly binding sites for 5S RNA and 5S DNA. *Mol Cell Biol* **13**, 5149-58 (1993).

223. Foster, M.P. et al. Domain packing and dynamics in the DNA complex of the N-terminal zinc fingers of TFIIIA. *Nat Struct Biol* **4**, 605-8 (1997).

224. Wuttke, D.S., Foster, M.P., Case, D.A., Gottesfeld, J.M. & Wright, P.E. Solution structure of the first three zinc fingers of TFIIIA bound to the cognate DNA sequence: determinants of affinity and sequence specificity. *J Mol Biol* **273**, 183-206 (1997).

225. Yang, W. Nucleases: diversity of structure, function and mechanism. *Quarterly Reviews of Biophysics* **44**, 1-93 (2011).

226. Yang, W. An equivalent metal ion in one- and two-metal-ion catalysis. *Nature Structural &Amp; Molecular Biology* **15**, 1228 (2008).

227. Larkin, C. et al. Inter- and Intramolecular Determinants of the Specificity of Single-Stranded DNA Binding and Cleavage by the F Factor Relaxase. *Structure* **13**, 1533-1544 (2005).

228. Schatz, D.G. & Swanson, P.C. V(D)J Recombination: Mechanisms of Initiation. *Annual Review of Genetics* **45**, 167-202 (2011).

229. Chakraborty, A., Uechi, T. & Kenmochi, N. Guarding the 'translation apparatus': defective ribosome biogenesis and the p53 signaling pathway. *Wiley Interdiscip Rev RNA* **2**, 507-22 (2011).

230. Sloan, K.E., Bohnsack, M.T. & Watkins, N.J. The 5S RNP couples p53 homeostasis to ribosome biogenesis and nucleolar stress. *Cell Rep* **5**, 237-47 (2013).

231. Donati, G., Peddigari, S., Mercer, C.A. & Thomas, G. 5S ribosomal RNA is an essential component of a nascent ribosomal precursor complex that regulates the Hdm2-p53 checkpoint. *Cell Rep* **4**, 87-98 (2013).

232. Li, M. et al. Mono- versus polyubiquitination: differential control of p53 fate by Mdm2. *Science* **302**, 1972-5 (2003).

233. Kruse, J.P. & Gu, W. Modes of p53 regulation. *Cell* **137**, 609-22 (2009).

234. Vousden, K.H. & Prives, C. Blinded by the Light: The Growing Complexity of p53. *Cell* **137**, 413-31 (2009).

235. Stuckey, J.A. & Dixon, J.E. Crystal structure of a phospholipase D family member. *Nat Struct Biol* **6**, 278-84 (1999).

236. Hui, L. et al. Phospholipase D Elevates the Level of MDM2 and Suppresses DNA Damage-Induced Increases in p53. *Molecular and Cellular Biology* **24**, 5677-5686 (2004).

237. Zhang, P. et al. TRIM25 has a dual function in the p53/Mdm2 circuit. *Oncogene* **34**, 5729 (2015).

238. Okumura, N. et al. Estradiol stabilizes p53 protein in breast cancer cell line, MCF–7. *Japanese Journal of Cancer Research* **93**, 867-873 (2002).

239. Kinyamu, H.K. & Archer, T.K. Estrogen receptor-dependent proteasomal degradation of the glucocorticoid receptor is coupled to an increase in mdm2 protein expression. *Molecular and cellular biology* **23**, 5867-5881 (2003).

240. Hurd, C., Dinda, S., Khattree, N. & Moudgil, V.K. Estrogen-dependent and independent activation of the P1 promoter of the p53 gene in transiently transfected breast cancer cells. *Oncogene* **18**, 1067 (1999).

241. MacIver, N.J., Michalek, R.D. & Rathmell, J.C. Metabolic regulation of T lymphocytes. *Annu Rev Immunol* **31**, 259-83 (2013).

242. Gorbalenya, A.E. & Koonin, E.V. Helicases: amino acid sequence comparisons and structure-function relationships. *Current opinion in structural biology* **3**, 419-429 (1993).

243. Caruthers, J.M. & McKay, D.B. Helicase structure and mechanism. *Curr Opin Struct Biol* **12**, 123-33 (2002).

244. Henriksen, U., Gether, U. & Litman, T. Effect of Walker A mutation (K86M) on oligomerization and surface targeting of the multidrug resistance transporter ABCG2. *J Cell Sci* **118**, 1417-26 (2005).

245. Hopfner, K.P. et al. Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. *Cell* **101**, 789-800 (2000).

246. Linder, P. Dead-box proteins: a family affair—active and passive players in RNP-remodeling. *Nucleic acids research* **34**, 4168-4180 (2006).

247. Hanson, P.I. & Whiteheart, S.W. AAA+ proteins: have engine, will work. *Nature Reviews Molecular Cell Biology* **6**, 519 (2005).

248. Walker, T.A. & Pace, N.R. 5.8 s ribosomal RNA. *Cell* **33**, 320-322 (1983).

249. Lafontaine, D.L.J. Noncoding RNAs in eukaryotic ribosome biogenesis and function. *Nature Structural &Amp; Molecular Biology* **22**, 11 (2015).

250. Maniatis, T., Fritsch, E.F. & Sambrook, J. *Molecular cloning: a laboratory manual*, (Cold spring harbor laboratory Cold Spring Harbor, NY, 1982).

251. Kibbe, W.A. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res* **35**, W43-6 (2007).

252. Liu, H. & Naismith, J.H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol* **8**, 91 (2008).

253. Afgan, E. et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic acids research* **46**, W537-W544 (2018).

254. Herten, K., Hestand, M.S., Vermeesch, J.R. & Van Houdt, J.K. GBSX: a toolkit for experimental design and demultiplexing genotyping by sequencing experiments. *BMC Bioinformatics* **16**, 73 (2015).

255. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**, R25 (2009).

256. Notredame, C., Higgins, D.G. & Heringa, J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205-17 (2000).

257. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320-4 (2014).

258. Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M. & Barton, G.J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-91 (2009).

259. Gasteiger, E. et al. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* **31**, 3784-8 (2003).

260. Kosugi, S., Hasebe, M., Tomita, M. & Yanagawa, H. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proceedings of the National Academy of Sciences* **106**, 10171-10176 (2009).

261. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12 (2004).

262. Jurrus, E. et al. Improvements to the APBS biomolecular solvation software suite. *Protein Sci* **27**, 112-128 (2018).

263. Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296-W303 (2018).

264. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *nature* **227**, 680 (1970).

265. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-54 (1976).

266. Roelofs, K.G., Wang, J., Sintim, H.O. & Lee, V.T. Differential radial capillary action of ligand assay for high-throughput detection of protein-metabolite interactions. *Proc Natl Acad Sci U S A* **108**, 15528-33 (2011).

267. Tiscornia, G., Singer, O. & Verma, I.M. Production and purification of lentiviral vectors. *Nature Protocols* **1**, 241 (2006).

268. Keilhauer, E.C., Hein, M.Y. & Mann, M. Accurate Protein Complex Retrieval by Affinity Enrichment Mass Spectrometry (AE-MS) Rather than Affinity Purification Mass Spectrometry (AP-MS). *Molecular & Cellular Proteomics* **14**, 120-135 (2015).

269. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols* **2**, 1896 (2007).

270. Cox, J. et al. Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Molecular & Cellular Proteomics* **13**, 2513-2526 (2014).

271. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* **13**, 731 (2016).

272. Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-32 (2010).

273. Zwart, P.H., Grosse-Kunstleve, R.W. & Adams, P.D. Xtriage and Fest: automatic assessment of X-ray data and substructure structure factor estimation. *CCP4 Newsl* **43**, 27-35 (2005).

274. Matthews, B.W. Solvent content of protein crystals. *Journal of molecular biology* **33**, 491-497 (1968).

275. Grosse-Kunstleve, R.W. & Adams, P.D. Substructure search procedures for macromolecular structures. *Acta Crystallographica Section D: Biological Crystallography* **59**, 1966-1973 (2003).

276. Terwilliger, T.C. et al. Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallographica Section D: Biological Crystallography* **65**, 582-601 (2009).

277. Terwilliger, T.C. et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* **64**, 61-9 (2008).

278. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-21 (2010).

279. Stein, N. CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *Journal of Applied Crystallography* **41**, 641-643 (2008).

280. Winn, M.D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-42 (2011).

281. McCoy, A.J. et al. Phaser crystallographic software. *Journal of applied crystallography* **40**, 658-674 (2007).

282. McCoy, A.J., Nicholls, R.A. & Schneider, T.R. SCEDS: protein fragments for molecular replacement in Phaser. *Acta Crystallographica Section D: Biological Crystallography* **69**, 2216-2225 (2013).

283. Afonine, P.V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**, 352-67 (2012).

284. Vagin, A.A. et al. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallographica Section D: Biological Crystallography* **60**, 2184-2195 (2004).

285. Murshudov, G., Vagin, A. & Dodson, E. Application of Maximum Likelihood Refinement. Refinement of Protein Structures. *Proceedings of Daresbury Study Weekend* (1996).

286. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D* **53**, 240-255 (1997).

287. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography* **66**, 486-501 (2010).

288. Franke, D. et al. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *Journal of applied crystallography* **50**, 1212-1225 (2017).

289. Mertens, H.D.T. & Svergun, D.I. Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *Journal of Structural Biology* **172**, 128-141 (2010).

290. Svergun, D.I., Petoukhov, M.V. & Koch, M.H. Determination of domain structure of proteins from X-ray solution scattering. *Biophys J* **80**, 2946-53 (2001).

291. Wriggers, W. Conventions and workflows for using Situs. *Acta Crystallographica Section D: Biological Crystallography* **68**, 344-351 (2012).

# 9. List of abbreviations

| | |
|---|---|
| **- Trp** | Minus tryptophan |
| **- Ura** | Minus uracil |
| **(v/v)** | Volume per volume |
| **(w/v)** | Weight per volume |
| **(w/w)** | Weight per weight |
| **°C** | Degree celsius |
| **6'-FAM** | 6-carboxyfluorescin |
| **A** | Absorption |
| **A** | Adenine |
| **Å** | Angstrom |
| **aa** | Amino acid |
| **AAA ATPase** | Atpase associated with various cellular activities |
| **AC** | Affinity chromatography |
| **Actc1** | Actin |
| **AIDS** | Acquired immunodeficiency syndrome |
| **AP-1** | Activator protein 1 |
| **APS** | Ammonium persulfate |
| **ATP** | Adenosine triphosphate |
| **ATPγS** | Adenosine 5´-[γ-thio]triphosphate |
| **ATR** | Ataxia telangiectasia and Rad3-related |
| **BD** | Bridging domain |
| **bp** | Base pairs |
| **BSA** | Bovine serum albumin |
| **c** | Concentration |
| **C** | Cytosine |
| **Ca** | Calcium |
| **CD** | Cluster of differentiation |
| **Cdc45** | Cell division cycle 45 |
| **CDK** | Cyclin-dependent kinase |
| **cDNA** | Complementray DNA |
| **cGAS** | Cyclic GMP-AMP synthase |
| **CH25H** | Cholesterol-25-hydroxylase |
| **ChIP** | Chromatin immunoprecipitation |
| **CHK1** | Checkpoint kinase 1 |
| **Clip2** | CAP-Gly domain-containing linker protein 2 |
| **C-lobe** | C-terminal lobe |
| **Cnn2** | Calponin |
| **Cond** | Conductivity |
| **CpG** | 5'- C- phosphate- G- 3' |
| **CPT** | Camptothecin |
| **CRC** | Colorectal carcinoma cells |
| **ct-CD45** | Cytoplasmic tail of CD45 |
| **C-terminus** | Carboxy terminus |
| **ctrl** | Control |

| | |
|---|---|
| **CV** | Column volume |
| **Cyc1** | Cytochrome C1 |
| **d** | Distance |
| **d** | Distance |
| **D** | Dialysate |
| **DDA** | DNA damaging agents |
| **dd$H_2O$** | Double-destilled water |
| **DN** | Double-negative (T cell) |
| **DNA** | Deoxyribonucleic acid |
| **DNMDP** | (4R)-3-[4-(diethylamino)-3-nitrophenyl]-4-methyl-4,5-dihydro-1H-pyridazin-6-one |
| **DP** | Double-positive (T cell) |
| **DRaCALA** | Differential radial capillary action of ligand assay |
| **ds DNA** | Double stranded DNA |
| **DSB** | Double-strand break |
| **E** | Elution fractions |
| ***E. coli*** | *Eschericia coli* |
| **e.g.** | Exempli gratia (for example) |
| **EAE** | Autoimmune encephalomyelitis |
| **EDTA** | N, N, N', N'-Ethylenediaminetetraacetic acid |
| **EIAV** | Equine infectious anemia virus |
| **EMSA** | Electrophoretic mobility shift assay |
| **ER** | Endoplasmic reticulum |
| **Esyt1** | Extended synaptotagmin-1 |
| **et al.** | Et altera |
| **EZH2** | Enhancer of Zeste homolog 2 |
| **FA** | Fluorescence anisotropy |
| **Fads3** | Fatty acid desaturase 3 |
| **FBS** | Fetal bovine serum |
| **Fech** | Ferrochelatase |
| **FF** | Fast flow |
| **Fhl1** | Four and a half LIM domains protein 1 |
| **FL** | Full length |
| **FOLFOX** | Folinic acid (calcium folonic acid), 5-fluorouracil, oxaliplatin |
| **FT** | Flow-through |
| **FU** | Fluorescence units |
| **G** | Guanine |
| **GBM** | Glioblastoma multiform |
| **GC** | Guanine - cytidine base pair |
| **Gli1** | Glioma-associated oncogene |
| **Gpaa1** | Glycosylphosphatidylinositol anchor attachment 1 |
| **GST** | Glutathione-S-transferase |
| **GU** | Guanine - uracil base pair |
| **h** | Hour |
| **HEPES** | 2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethansulfonsäure |
| **Hisx** | Polyhistidinetag consisting of X histidines |
| **HIV-1** | Human immunodeficiency virus type 1 |
| **HMG-CoA** | 3-hydroxy-3-methylglutaryl-Coenzyme A |
| **HPLC** | High performance liquid chromatography |

| | |
|---|---|
| **hSLFNX** | Human schlafen X |
| **i.e.** | Id est (that is to say) |
| **IEX** | Ion exchange chromatography |
| **IFITM** | IFN-inducible transmembrane |
| **IFN** | Interferon |
| **IFNAR1/ 2** | Interferon alpha and beta receptor subunit 1/ 2 |
| **IL** | Interleukin |
| **IPTG** | Isopropyl-ß-D-thiogalactopyranosid |
| **IRF1** | Interferon regulatory factor |
| **ISG** | Interferon stimulated genes |
| **ISGF** | Interferon-stimulated-gene factor |
| **ISREs** | IFN-stimulated response elements |
| **JAK** | Janus-activated kinase |
| **kbp** | Kilobasepairs |
| **kb** | Kilo base |
| **$K_d$** | Dissociation constant |
| **kDa** | Kilo Dalton |
| **KLF2** | Krüppel-like factor |
| **L** | Load |
| **LB** | Lysogeny broth |
| **LPS** | Lipopolysaccharides |
| **M** | Marker |
| **MBP** | Maltose-binding-protein |
| **MCM** | Minichromosome maintenance complex |
| **MDA5** | Melanoma differentiation-associated protein 5 |
| **Mdm2** | Mouse double minute 2 homolog |
| **MDSC** | Myeloid-derived suppressor cell |
| **Mg** | Magnesium |
| **MHC** | Major histocompatibility complex |
| **min** | Minute |
| **MMP** | Metalloproteinase |
| **Mn** | Manganese |
| **moDC** | Monocytes-derived dendritic cell |
| **mSlfnX** | Murine schlafen X |
| **Mx1** | Myxovirus resistance |
| **Myo1c** | Unconventional myosin-1c |
| **NE** | N-terminal extension |
| **NFκB** | Nuclear factor 'kappa-light-chain-enhancer' of activated B-cells |
| **NHEJ** | Non-homologous end joining |
| **Ni** | Nickel |
| **NK cell** | Natural killer cells |
| **N-lobe** | N-terminal lobe |
| **NLS** | Nuclear localization signal |
| **nt** | Nucleotide |
| **NTA** | Nitrilotriacetic acid |
| **N-terminus** | Amino terminus |
| **$OD_{xy}$** | Optical denisty at wavelength xy |
| **P(r)** | Pair-distance distribution function |

| | |
|---|---|
| **PAR** | Poly (ADP-ribose) |
| **PARP** | Poly (ADP-ribose) polymerase |
| **PBMC** | Peripheral blood mononuclear cell |
| **PBS** | Phosphate buffered saline |
| **Pbxip1** | Pre-B cell leukemia transcription factor-interacting protein 1 |
| **PCR** | Polymerase chain reaction |
| **PDB** | Protein data bank |
| **PDCD4** | Programmed cell death protein |
| **PDE3A** | Phosphodiesterase 3A |
| **PEG** | Polyethylen glycol |
| **PEG** | Polyethylene glycol |
| **PEG MME** | Poly(ethylene glycol) methyl ether |
| **Pel** | Pellet |
| **pH** | Potential of hydrogen |
| **$P_i$** | Inorganic phosphate |
| **pI** | Isoelectric point |
| **PKR** | Protein kinase R |
| **Pld3** | Phospholipase D3 |
| **poly (dA)** | Poly desoxyadenine |
| **poly (dT)** | Poly desoxythymidine |
| **poly (rU)** | Poly uracile |
| **Prkab2** | 5-AMP-activated protein kinase subunit beta-2 |
| **PRRs** | Pattern-recognition receptors |
| **pTα** | Pre T cell receptor α-chain |
| **RALS** | Right-angle light scattering |
| **RCC** | Renal cell carcinoma |
| **RIG-I** | Retinoic acid-inducible gene I |
| **RNA** | Ribonucleic acid |
| **RPA** | Replication protein A |
| **RPL** | Large ribosomal subunit protein |
| **rpm** | Rounds per minute |
| **rRNA** | Ribosomal RNA |
| **rSlfn X** | Rat schlafen X |
| **s** | Second |
| **S** | Svedberg unit |
| ***S. cerevisiae*** | *Sacchromyces cerevisiae* |
| **s.d.** | Standard deviation |
| **S200** | Superdex 200 |
| **SAD** | Single-wavelength anomalous dispersion |
| **SAMHD1** | SAM and HD domain-containing protein 1 |
| **SD** | Synthetic defined |
| **SDS** | Sodium dodecyl sulfate |
| **SDS-PAGE** | SDS-Polyacrylamide gelelectrophoresis |
| **SEC** | Size exclusion chromatography |
| **SenP2** | Sentrin-specific protease 2 |
| **Sf** | *Spodoptera frigiperda* |
| **SF** | Superfamily |
| **SH2** | Scr homology 2 |

| | |
|---|---|
| **Slfn** | Schlafen |
| **SLFN12-L** | Schlafen 12- like |
| **SO4** | Sulfate |
| **SOCS** | Suppressor of cytokine signaling |
| **SP** | Single positive (T cell) |
| **ss DNA** | Single-stranded DNA |
| **SSB** | Single-strand break |
| **STAT** | Signal transducer and activator of transcription |
| **STEAP** | 6-transmembrane epithelial antigen of prostate |
| **Strep** | Streptavidine |
| **STS** | Steryl-sulfatase |
| **SUMO1** | Small Ubiquitin-related Modifier |
| **Sup** | Supernatant |
| **T** | Thymine |
| **TAE** | Tris/Acetate/EDTA buffer |
| **T-ALL** | T cell acute lymphoblastic leukemia |
| **TBE** | Tris/Borate/EDTA buffer |
| **TCR** | T cell receptor |
| **TEMED** | N, N, N', N'-Ethylenediaminetetraacetic acid |
| **TEV-protease** | Protease from Tobacco Etch Virus |
| **TFA** | Trifluoroacetic acid |
| **TFH** | Follicular helper T cell |
| **TGF-β** | Transforming Growth Factor beta |
| **TH** | T helper cell |
| **TKY2** | Tyrosine kinase 2 |
| **TLC** | Thin-layer chromatography |
| **Tns1** | Tensin |
| **TOP** | Topoisomerases |
| **Treg** | Regulatory T cell |
| **TRIM** | Tripartite motif |
| **Tris** | Tris(hydroxymethyl)aminomethane |
| **tRNA** | Transfer RNA |
| **U** | Uracil |
| **Uhrf1** | Ubiquitin-like, containing PHD and RING finger domains 1 |
| **Unc5b** | Uncoordinated netrin receptor 5 |
| **USP18** | Ubiquitin-specific peptidase 18 |
| **UV** | Ultraviolet |
| **V, D, J and C** | Variable, diversity, joining and constant gene segments |
| **viperin** | Virus inhibitory protein, endoplasmic reticulum-associated, IFN- |
| **W** | Wash fractions |
| **WT** | Wild type |
| **x g** | Multiple of the acceleration of gravity |
| **X-Gal** | 5-Brom-4-chlor-3-indoxyl-β-D-galactopyranosid |
| **YPD** | Yeast extract peptone dextrose |
| **Zn** | Zinc |
| **β-ME** | 2- mercaptoethanol |
| **ε** | Molar extinction coefficient |
| **λ** | Wavelength |

# 10. Acknowledgments

First, I would like to thank Dr. Katja Lammens for her infinite support, the great supervision, the funding of my project, the daily discussions including theoretical and practical advices, good ideas, the trust and the friendship that has evolved during the last four years. Thank you Katja for the enjoyable time, in which I learned a lot.

Next, I would like to thank Prof. Karl-Peter Hopfner who gave me the opportunity to join his laboratories, for being my “Doktorvater” and for reviewing this thesis. I am most thankful for his supervision, the regular and efficient discussions and the right balance of supervision and experimental freedom.

I am also very thankful to Prof. Marc Schmidt-Supprian for being second reviewer and being a member of my TAC. Furthermore, I thank Prof. Daniel Krappmann, PD Gregor Witte, Prof. Dr. Klaus Förstemann and Prof. Julian Stingele for accepting to be co-reviewers.

I am very grateful to Katja Lammens, Charlotte Lässig, Laia Pascual Ponce, David Drexler and Beatrice Laudenbach for proofreading of this thesis and for their helpful advices.

Moreover, I am most thankful for Gregor Witte for the scientific advices and IT support.

I am thankful to Stefan Krebs from LAFUGA as well as to Annika Frauenstein and Felix Meissner from the MPI for the successful collaborations.

I am deeply thankful to the whole Hopfner group for the great and friendly working atmosphere. Furthermore, I want to thank my old lab 4.60 and my new lab 4.20 for the enjoyable times and the support. I am deeply thankful that colleagues developed to friends. I am grateful to all former and current members of the Schlafen subgroup: Florian and Felix. I also want to acknowledge Alexandra Till, Brigitte Keßler, Manuela Moldt and Olga Fettscher for the excellent organization of the laboratory.

I am very grateful to my students Enrica Zanuttigh and Hendrik Bulthaupt for their work and enthusiasm.

Furthermore, I am most grateful to my graduate schools IMPRS-LS and IRTG (SFB1054) for the financial and scientific support and the opportunity of building up a scientific network.

Most importantly, I want to thank my family and friends. My deepest thanks go to my parents and my sister Johanna for their love, good advices and unlimited support throughout my life. Especially, I am very grateful to my beloved husband Georg for his endless love, support and patience.