Tânia Filipa Fernandes dos Santos Alves Custódio

Bachelor of Science in Biochemistry

”Staying under the radar” - The Multifunctional LANA Protein

Dissertation submitted in fulfilment of the requirements for the degree of Master of Science in Biochemistry

Supervisor: Colin McVey, Principal Investigator, ITQB AX-UNL
Co-Supervisor: Rajesh Ponnusamy, Post-doctoral fellow, ITQB, AX-UNL

Júri:

President: Dr. Carlos Alberto Gomes Salgueiro
Examiner: Dr. Tiago Miguel Guerra Miranda Bandeiras
Members: Dr. Colin McVey

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

September, 2015

Tânia Filipa Fernandes dos Santos Alves Custódio

Bachelor of Science in Biochemistry

”Staying under the radar” - The Multifunctional LANA Protein

Dissertation submitted in fulfilment of the requirements for the degree of Master of Science in Biochemistry

Supervisor: Colin McVey, Principal Investigator, ITQB AX-UNL
Co-Supervisor*:* Rajesh Ponnusamy, Post-doctoral fellow, ITQB, AX-UNL

Júri:

President: Dr. Carlos Alberto Gomes Salgueiro
Examiner: Dr. Tiago Miguel Guerra Miranda Bandeiras
Members: Dr. Colin McVey

September, 2015

ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my **AWESOME** supervisors Dr. Colin McVey and Dr. Rajesh Ponnusamy. I could not have wished for better or friendlier advisors for my Master Thesis.

To Colin, whose expertise, understanding, and patience, added considerably to my graduate experience. I appreciate his vast knowledge and skill in many areas and his assistance in writing this thesis, especially on the use of correct grammar and consistent notation in my writings.

A very especially thanks to Rajesh for excellent guidance, caring and knowledge whilst allowing me the room to work in my own way. Rajesh taught me how to question thoughts and express ideas. His patience even when, sometimes, I did the opposite of what he said. Also for his friendship, motivation and encouragement on pursuing my research career abroad and help me with several issues of applications.

I thank my fellow Lab mates of the Structural Virology Lab for technical Support and patient, principally in the beginning of the Master.

I am also indebted to the members of the Crystallographic Unit of ITQB, with whom I have interacted during the course of my Master Thesis. In my daily work I have been blessed with a friendly and cheerful group that helped me regain some sort of fitness: healthy body, healthy mind.

Many friends have helped me stay sane through these difficult year. Their support and care helped me overcome setbacks and stay focused on my thesis. I greatly value their friendship and I deeply appreciate their belief in me.

Most importantly, none of this would have been possible without the love and patience of my family. My family to whom this dissertation is dedicated to, has been a constant source of love, concern, support and strength not only this year but many others. I would like to express my heart-felt gratitude to my family.

“Wisdom is not a product of schooling but of the lifelong attempt to acquire it.”

Albert Einstein

ABSTRACT

Many viruses have developed numerous strategies to recruit and take advantage of cellular protein degradation pathways to evade the cellular viral immune system. One such virus is the Kaposi´s Sarcoma associated herpesvirus (KSHV), first discovered in Kaposi´s Sarcoma lesions found in AIDS patients. Latency-Associated Nuclear Antigen (LANA) is a KSHV multifunctional protein responsible for tethering viral DNA to the chromosome ensuring maintenance and segregation of the viral genome during cell division. Besides its main role of viral maintenance, LANA also physically interacts with several host proteins to modulate cell functions. One such function is to recruit the EC_5S ubiquitin-ligase complex by interacting with Elongin BC complex and Cullin 5 protein, which in turn ubiquitinate substrates such as NF-κB and p53 to allow persistent viral infection. Like any other post-translation modifications, ubiquitination is reversible through deubiquitination enzymes (DUBs). LANA also interacts with ubiquitin specific protease 7 (USP7), a deubiquitination enzyme involved in regulation of several proteins including p53. Interaction with USP7 is made through a conserved peptide motif, which is also present in LANA. This work addresses the role of LANA in the recruitment and modulation of the ubiquitination and deubiquitination pathways. Despite the continued efforts in uncovering new LANA interacting partners to form a functional EC_5S ubiquitin-ligase complex, only MHV-68 LANA interacted directly with Elongin BC, other interactions were not direct and may require a linker protein. On the other hand, LANA interaction with USP7 was able to be analysed by X-ray structure determination. In addition to a conserved P/AxxS motif, a novel Glutamine (Gln) residue from KSHV LANA was shown to make a specific interaction with USP7. This Gln residue is also present in other herpesvirus protein and hence it might be a conserved motif within herpesviruses.

Keywords: Kaposi´s sarcoma herpesvirus; Latency-Associated Nuclear Antigen; Ubiquitination; EC_5S Ubiquitin complex; Ubiquitin Specific Protease 7; X-ray structure.

RESUMO

Vários vírus desenvolveram inúmeras estratégias de recrutamento de vias celulares de degradação proteíca, para eludir o sistema imunitário. Um exemplo, é o Herpesvírus associado ao sarcoma de kaposi (KSHV), tendo sido descoberto no sarcoma de kaposi em pacientes portadores de HIV. O Antigénio Nuclear Associado à Latência (LANA) é uma proteína multifuncional deste vírus, responsável pela manutenção e segregação do genoma viral durante a divisão celular. Por outro lado, LANA tem a capacidade de interagir físicamente com várias proteínas do hospedeiro, de forma a garantir infecção viral ininterrupta. Um exemplo concreto é o recrutamento do complexo Ubiquitina-Ligase EC_5S, através da interação com as proteínas Elongin BC e Cullin 5 que por sua vez têm a capacidade de ubiquitinar diversos substratos como NF-κB e p53, para posterior degradação. Tal como acontece em outras modições pós-traducionais, ubiquitinação é reversível. De forma equivalente, LANA interage com a protease especifica de ubiquitina 7 (USP7), responsável pela regulação de várias proteínas através da remoção de cadeias de ubiquitina. A interacção entre USP7 e diversos ligandos é feita através de um motivo conservado, também presente na sequência de LANA. Este trabalho teve como objetivo ilucidar o papel de recrutamento e as interacções estabelecidas entre LANA e o complexo EC_5S, como entre LANA e a proteína USP7. Através de ensaios de pull down, verificou-se que a proteína LANA do KSHV não é capaz de estabelecer uma interação direta com Elongin BC. Por outro lado, a proteína homologa de murine, mLANA, consegue interagir diretamente com Elongin BC. Foi ainda possível determinar a estrutura raios-X do complexo USP7-LANA e ilucidar as interacções estabelcidas entre estas proteínas. Apesar de USP7 ser caracterizado por reconhecer o motivo conservado P/AxxS, um novo motifo QPGPR foi identificado no complexo USP7-LANA.

Palavras-Chave: Herpesvirus Associado ao Sarcoma de Kaposi; Antigénio Nuclear Associado à Latência; Ubiquitinação; Complexo Ubiquitina-Ligase EC_5S Ubiquitin; Protease Específica de Ubiquitina 7; Estrutura raios-X.

CONTENTS

FIGURES CONTENTS

TABLE CONTENTS

ABBREVIATIONS

KS	Kaposi´s sarcoma
HIV-1	Human Immunodeficiency Virus
KSHV	Kaposi´s Sarcoma Associated Herpesvirus
HHV8	Human Herpesvirus 8
MHV-68	Murine gammaherpesvirus 68
RTA	Replication and Transcription Activator
ORF50	Open Reading Frame 50
LANA	Latency-Associated Nuclear Antigen
v-Cyclin	Viral Cyclin
vFLIP	Viral Fas-associated death domain interleukin 1L- protein
vIFN	Viral Interferon Regulatory Factors
kLANA	KSHV LANA
ORF73	Open Reading Frame 73
NLS	Nuclear Localization Signal
LBS	LANA Binding Site
TR	Terminal Repeat
DBD	DNA-Binding Domain
mLANA	MHV-68 LANA
ER	Endoplasmic Reticulum
UPS	Ubiquitn proteasome system
Ub	Ubiquitin
HPV	Human Papillomavirus
CRL	Cullin-RING ligase
SDS	Sodium Dodecyl Sulfate
TCL	Total Cell Lysate
FT	Flowthrough
aa	Amino Acid residue
EMSA	Electrophoretic Mobility Shift Assay
TRAF	TNF-receptor associated factors
MATH	Meprin And TRAF Homology
USP7-NTD	N-Terminal Domain of USP7
EBV	Epstein-Barr virus
DSF	Differential Scanning Fluorimetry
Tm	Melting Temperature
ECS	Elongin B and C–Cullin–SOCS box protein
Elongin BC	Elongin B and C complex
Cul2	Cullin 2
Cul5	Cullin 5
CBF-β	Core Binding Factor beta
VHL	Von Hippel-Lindau
DUB	Deubiquitinating enzyme
USP7	Ubiquitin-specific protease 7
HAUSP	Herpesvirus-Associated Ubiquitin-specific protease
TNF	Tumor Necrosis Factor
PCR	Polymerase Chain Reaction

1

INTRODUCTION

1.1. Herpesvirus

The name herpes originated from the Greek word herpein which means to creep, which reflects the spreading nature of the skin lesions caused by many herpesvirus types. Of the more than 130 known herpesviruses, eight are a leading cause of human viral disease: herpes simplex virus types 1 and 2, varicella-zoster virus, cytomegalovirus, Epstein-Barr virus, human herpesvirus 6 (variants A and B), human herpesvirus 7, Kaposi's sarcoma associated herpesvirus and B virus. Each virus within the family has the potential to establish latency in specific host cells, and the latent viral genome may be extra-chromosomal or integrated into host cell DNA. A latent virus may be reactivated and enter a replicative cycle at any point in time [1].

Herpesviruses are divided into three sub-families: The α-Herpesviruses characterized by a short replicative cycle (hours) and have a broad host range. This sub-family establish latency in neurons and consists of herpes simples virus types 1 and 2 and varicella-zoster virus; the β-Herpesviruses, that in contrast to α-Herpesviruses, have a long replicative cycle (days) and have a broad host range. Cytomegalovirus and human herpesvirus belongs to the β-Herpesviruses and they establish latency in macrophages and lymphocytes. Lastly, γ-herpesviruses, such as Epstein-Barr virus and Kaposí´s sarcoma associated herpesvirus, have the most restricted range host and establish latency only in lymphocytes [1,2].

1.2. Kaposí´s Sarcoma associated Herpesvirus

Kaposí´s sarcoma (KS) is an infectious cancer that mainly occurs in immunocompromised patients, characterized by neoplastic cells and abnormal growth of blood vessels (Figure 1.1). This type of tumour mainly affects the skin, mouth and lymph nodes. It is named after the Hungarian dermatologist, Moritz Kaposi who discovered it in 1872 [3]. KS occurs in Europe and the Mediterranean countries (classic KS), Africa (endemic KS), immunosuppressed patients (post-transplant KS) and those with acquired immune deficiency syndrome (AIDS), and is especially prevalent among those who are infected by human immunodeficiency virus (HIV-1) [4,5].

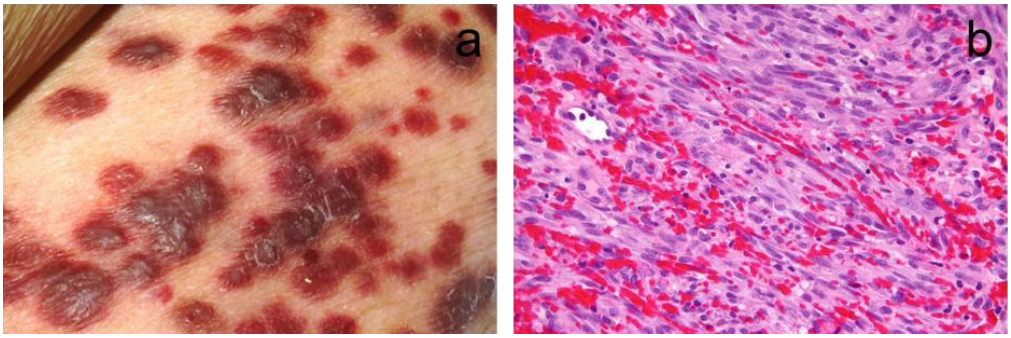

Figure 1.1- Kaposi's sarcoma oncogenesis. *(a) example of a typical kaposi's sarcoma skin lesion; (b) micrograph of a kaposi sarcoma (hematoxylin and eosin stain) showing the characteristic spindle cells, high vascularity and intracellular hyaline globs.*

In 1994 KS was discovered to be caused by a virus from the herpesvirus family, and named Kaposi´s sarcoma associated herpesvirus (KSHV) [6]. KSHV is the most recently discovered human tumour virus. It is the eighth human herpesvirus isolated to date and is thus also named human herpesvirus 8 (HHV8) [6]. It is a large double-stranded DNA virus from the γ-herpesvirus family. Besides being responsible for KS lesions, it is also associated with other lymphoproliferative disorders, such as primary effusion lymphoma and Multicentric Castleman´s disease [7–9].

KSHV infection is predominantly latent, and the virus persists in the nuclei of latently infected cells as a multiple-copy, covalently closed, extrachromosomal episome [5].

The murine gammaherpesvirus 68 (MHV-68 or murid herpesvirus 4) is a natural pathogen of small rodents (mice, voles and shrews), from the rhadinovirus family [10]. It is structurally and functionally related to human γ-herpesviruses and readily infects mice, thus providing an ideal mouse model for the investigation of γ-herpesvirus pathogenesis [11]. The sequence is 118 kb in length flanked by guainine cytosine (GC)-rich terminal repeats and exist in a circular episomal form during latency. It is predicted to encode 79 ORFs [12].

All γ-herpesviruses, including EBV, KSHV and MHV-68, encapsidate duplex linear DNA genomes in their particles. Upon infection of cells, this DNA is delivered to the nucleus, where it is circularized (largely by host enzymatic machinery), generating a closed-circular DNA form that can persist in the nucleus as a plasmid [12].

Similar to other herpesvirus family members, KSHV and MHV-68 display two alternate genetic life cycles: the establishment of latency and reactivation from latency, which lead to lytic replication [13,14]. The transition from latency to lytic replication is controlled by the KSHV replication and transcription activator (RTA), encoded by the KSHV gene open reading frame 50 (ORF50) [13,15].

Lytic replication leads to extensive viral gene expression, virion production, and death of the infected cell [13]. In contrast, latent infection is characterized by a limited gene expression to a small subset of viral latent genes and are often involved in viral episome maintenance, replication and the transformation of host cells [16].

Establishing latency enables KSHV to sustain a life-long infection and to escape immune surveillance [16,17]. At least, five viral genes are expressed during latency, the latency-associated nuclear antigen (LANA), viral cyclin (v-Cyclin), viral Fas-associated death domain interleukin 1L-converting enzyme inhibitory protein (vFLIP) encoded by K13, viral interferon regulatory factors encoded by K10 (vIFN3), and kaposin encoded by K12 [13,14].

1.3. LANA Protein

Latently expressed viral proteins have been shown to promote tumourigenesis by subverting the cellular mechanisms which would normally protect cells from aberrant proliferation; LANA is the predominant gene expressed in latent infection and exerts anti-apoptotic and anti-proliferative effects on cells by interacting with and inhibiting p53, retinoblastoma protein, and glycogen synthase kinase 3 [17]. The main function of LANA is to tether the viral genome to the chromosome during cell division and allows the virus to persist by segregation of the viral DNA to daughter cells.

LANA is multifunctional and capable of maintaining viral latency through repression of the transcriptional activity of RTA (replication and transcription activator) promoter [18]. It is also required for DNA replication, chromosome tethering, cell cycle regulatory and gene regulatory functions. LANA binds to sequences in the virus terminal repeats that are essential for replication of the episomal genomes during latency [19].

The KSHV LANA (kLANA) encoded by open reading frame 73 (ORF73), is a nuclear protein with 1162 amino acid in length and it was first identified by Western blot as a 222- to 234- kDa protein [20,21]. The N-terminal domain is separated from its C-terminal domain by a large internal repeat region (Figure 1.2). The N-terminal domain of kLANA binds to nucleosome core particles: the proline-rich region contains a nuclear localization signal (NLS) and a sequence motif that tethers LANA to chromosomes during mitosis.

The central domain contains different highly repetitive blocks of acidic amino acids. The acidic region close to the C-terminal domain contains a leucine zipper motif that has been demonstrated to be necessary for kLANA's interaction with the cellular proteins mentioned above [22].

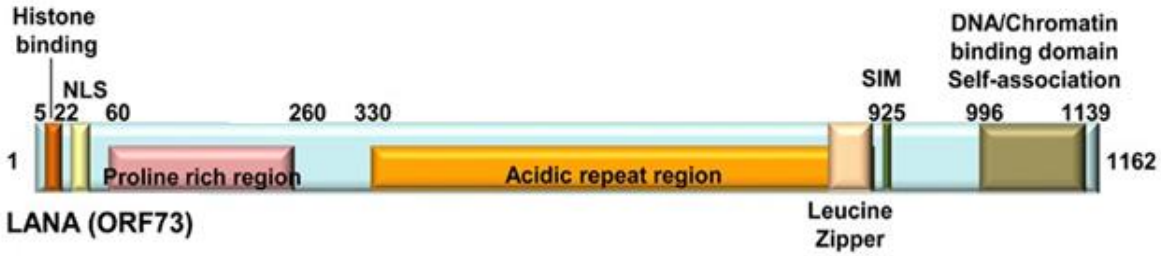

Figure 1.2- Schematic representation of KSHV LANA. *The N-terminal histone binding domain, nuclear localization signal (NLS), Proline rich region, Acidic repeat region, Leucine Zipper and SUMO interaction motif (SIM) are also indicated. Numbers indicate amino acid position. Adapted from [23].*

kLANA's C-terminal domain binds to the LANA binding sites (LBS) within the viral terminal repeat (TR) in a sequence-specific manner and thus this domain is referred to as DNA-binding domain (DBD) [24–27]. The N-terminal domain and the internal repeat region are predicted to be only poorly structured, in contrast the C-terminal domain comprises a stable 3D structure with a strong hydrophobic core. Therefore, only the LANA DBD structure was initially solved [28] and more recently the X-ray structure of LANA C-terminal domain bound to DNA was determined [24]. The DNA is bound to the kLANA DBD dimer with remarkable asymmetry [24]. Despite the great homology, kLANA differs from EBNA1 and E2, both of which bind palindromic target sequences in a symmetric fashion [29,30]. It was also found that LANA binds to a third binding site, the replication element [24], denoted LBS3, to distinguish from LBS1 and LBS2 tethering sites. All three sites are located in the KSHV terminal repeat subunit region [24].

The ORF73 from MHV-68 (mLANA) encodes a much smaller, 314 amino acid, 50 kDa nuclear protein which lacks the extensive internal acidic and glutamine-rich repeat region of kLANA. The C-terminal region of mLANA, comprising amino acid residues 140 to 263, which is similar to the kLANA DNA-binding domain. Similar to kLANA, mLANA was also recently shown to act on TR elements of the MHV-68 genome to mediate episome maintenance and to associate with mitotic chromosomes.

However, kLANA and mLANA DBD interact with its respective LBS regions very differently, with a bent and linear conformation, respectively, indicating divergence within KSHV and MHV-68 [31].

1.4. Ubiquitin Complex

All intracellular proteins and many extracellular proteins are continually "turning over;" i.e., they are being hydrolysed to their constituent amino acids and replaced by new synthesis. Although the continual destruction of cell proteins might seem wasteful, this process serves several important regulation functions, such as regulation of cell cycle, cell growth and development, transcription, DNA repair, oncogenesis, antigen processing and selective degradation of abnormally folded and damaged proteins [32]. Individual proteins in the nucleus and cytosol, as well as in the endoplasmic reticulum (ER) and mitochondria, are degraded at widely differing rates that vary from minutes for some regulatory enzymes to days or weeks for proteins such as actin and myosin in skeletal muscle or months for haemoglobin in the red cell. The Ubiquitin–proteasome system (UPS) is the principal mechanism for the turnover of short-lived proteins [33–35].

Ubiquitin (Ub) is a highly conserved protein in eukaryotes, of small size (8.5 kDa polypeptide) with 76 amino acids. Cellular proteins destined for degradation are covalently linked to ubiquitin through a multistep process known as ubiquitination. Ubiquitination consists of concerted actions of enzymes that link chains of the polypeptide co-factor, Ub, onto proteins to mark them for degradation. The presence of multiple substrate-linked ubiquitins recruits the 26S proteasome, a 2.5 MDa complex that uses energy derived from ATP hydrolysis to unfold the substrate polypeptide chain and translocate it into an interior chamber. Having arrived at this site, the substrate is hydrolysed to produce small peptides. Ubiquitin is spared from degradation through its release from the substrate (or a substrate fragment) by deubiquitinating enzymes [36,37].

The Ub–proteasome system comprises two phases. In the first phase (Figure 1.3), Ub is activated through the ATP-dependent formation of a thiol ester with a cysteine residue of Ub-activating enzyme (E1), then transferred to a cysteine residue of an Ub-conjugating enzyme (E2) before, finally being transferred to a lysine residue of the substrate in a reaction catalysed by a Ub-protein ligase (E3). Ub is linked to the substrate through an isopeptide bond. For substrates that will be targeted to proteasomes, additional Ub molecules are conjugated to the first one to form a branched chain that is also linked through isopeptide bonds [33,38].

In the second phase of the cycle, the proteasome recognizes the substrate via the poly-ubiquitinate chain; the substrate is degraded to small peptides in a processed manner, and Ub is regenerated by specific deubiquitinating enzymes. Like conjugation, proteasome-catalysed degradation is ATP dependent [33,38].

The E3 ligase usually determines the substrate specificity, although the E2-conjugating enzyme can also play a role in the substrate selection. Thus hundreds of E3/Ub ligases have been identified, while there are only few known E1 and E2 enzymes, in mammals. E3 enzymes are currently classified into three main classes, with different structural and functional characteristics: the HECT domain family of Ub ligases, the Cullin-RING family of Ub ligases, and the U-box containing Ub ligases [38]

Ub contains at least seven lysines (K) and additional residues that can be employed by the Ub ligases to generate different types of Ub chains on the target proteins, which, in turn, will interact with different downstream factors (Figure 1.3). For instance, it is well established that K-48-based linkages lead mainly to the proteasome-mediated degradation of the ubiquitinated protein, while K-63-based Ub chains control primarily protein endocytosis, as well as trafficking and enzyme activity [36,39]. Moreover, it has been demonstrated that Ub can also function to act independently from its proteolytic activity, by regulating protein function and protein/protein interaction[38,40].

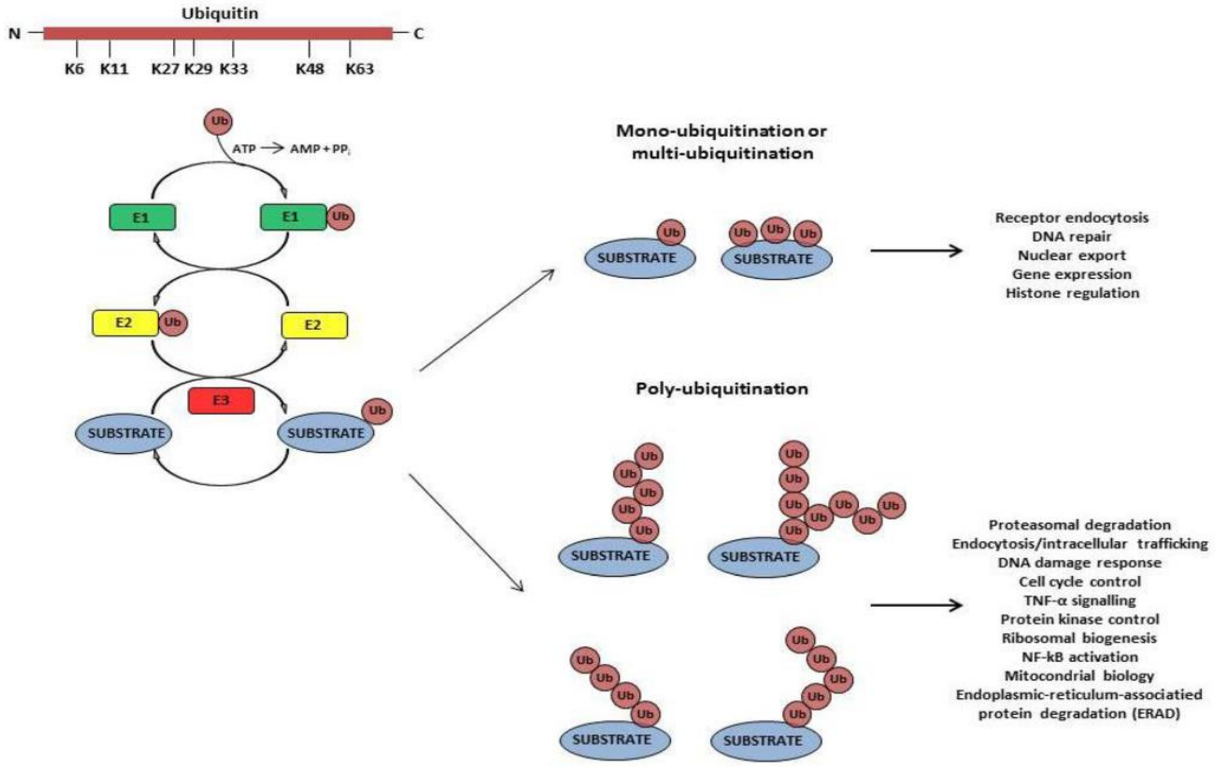

Figure 1.3- Schematic representation of the Ub molecule and of the enzymatic cascade leading to protein ubiquitination. *The seven lysines (K) involved in the process, the ubiquitin-activating enzyme E1, the ubiquitin-conjugating enzyme E2 and the ubiquitin ligase enzyme E3 are highlighted, along with the main fate of the target proteins. Adpated from [33].*

Innate immunity represents the first line of defence employed by our immune system to fight against pathogenic microorganisms. Human primary cells express restriction factors to block the replication and spread of viruses. To react to this immune response, viruses have evolved in order to counteract this effect. One such strategy used by viruses is to hijack cellular proteasomal degradation pathways, including UPS, to degrade the host restriction factors, so they can gain entry into target cells and replicate [33].

In 1990, it was demonstrated that the E6 protein of the oncogenic human papillomavirus (HPV) promotes p53 degradation through the ubiquitin proteasome system [41]. This was the first evidence supporting the ability of viruses to exploit the host immune system for their own purposes, in particular to utilize the UPS to interfere with the regulation of the cell cycle. Since then, it has become clear that members of almost all viral families subvert or exploit both the cellular Ub-conjugating and -deconjugating machineries in different phases of their replication cycle [33,42–46].

1.5. EC_5S Ubiquitin-ligase complex

The largest E3 superfamily consists of multi-subunit Cullin-RING ligases (CRLs). CRLs are nucleated by an extended Cullin scaffold interacting with a catalytic RING-containing protein, either RBX1 or RBX2. The Elongin B and C–Cullin–SOCS box protein (ECS) family also belongs to the Cullin RING ligase (CRL) superfamily. Elongin B and C complex (Elongin BC) is used as an adaptor in the ECS complex. In this family, the scaffold protein can either be Cullin 2 (Cul2) or Cullin 5 (Cul5). In the case of the systems using Cul5 protein it is named the EC_5S Ubiquitin-ligase complex (Figure 1.4) [47,48].

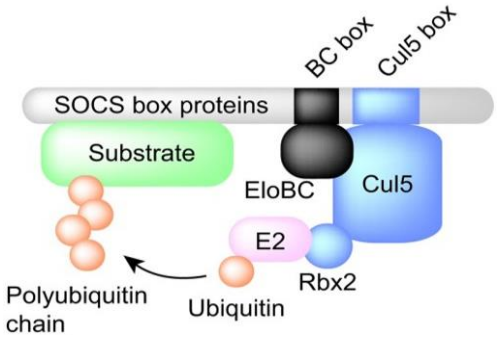

Figure 1.4- Schematic representation of EC_5S Ubiquitin-ligases family. Cul5 is used as a scaffold protein. Note that the Rbx2 is used for the recruitment of E2 enzyme. Adapted from [48].

HIV-1 Vif protein is one example of a viral protein recruiting the EC_5S Ubiquitin-ligase complex for its own purpose. Vif recruits a multi-subunit E3 Ub ligase complex composed of a scaffold protein, Cul5, RING-box protein, a SOCS box binding protein complex, Elongin BC, as well as the core binding factor beta (CBF-β) [42,49–55]. Vif achieves this by direct binding to Cullin 5. The ultimate goal of Vif-CBF-β-ElonginB-ElonginC-Cullin5-Rbx E3 complex is to polyubiquitinate APOBEC3G (host restriction factor) leading to its proteasomal degradation. As a consequence, APOBEC3G is not incorporated into the viral particle and HIV-1 can replicate in de novo infected cells. It has to be mentioned that additional APOBEC molecules (i.e. APOBEC3DE/3F) have been characterized for their ability to restrict Vif-defective HIV-1. The underlying mechanism and molecular details behind the recruiting process of Vif is already well characterized through the recently solved crystal structure of Vif-CBF-β-ElonginB-ElonginC-Cullin5 complex [49].

There is another viral protein that recruits this complex. Apart from its main role of tethering viral episome to chromosome, LANA is also shown to physically interact with many host proteins and manipulate them, including the EC_5S ubiquitin-ligase complex by interacting with Elongin BC and Cul5. This has been shown for kLANA where substrates such as NF-κB, p53 and VHL (von Hippel-Lindau) proteins are target for proteasomal degradation [43], as well for mLANA that was also demonstrated, targeting the transcription factor p65 (also known as nuclear factor NF-κB p65 subunit, involved in NF-κB heterodimer formation) for degradation through this complex [44].

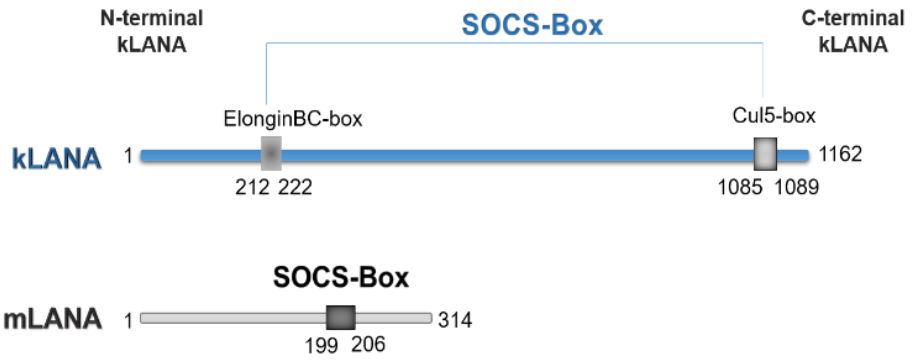

Figure 1.5– Difference in the localization of the SOCS-box element between kLANA and mLANA proteins.

Although kLANA and mLANA are homologues proteins that recruit the same Ubiquitin-ligase complex, the SOCS-box motif are differently arranged between them (Figure 1.5), and possibly might act with a different molecular mechanism and physical interactions with the EC_5S Ubiquitin-complex. The well characterized SOCS-box motif in the kLANA protein is composed of an Elongin BC box and a Cullin box, which is spatially located at its amino and carboxyl termini. This motif is necessary for LANA interaction with the Cul5–Elongin BC complex, respectively [43]. On the other hand, mLANA contains an unconventional SOCS-box-like motif, where the Elongin B and C box and Cullin box are spatially together [44].

Molecular details of LANA (both human and murine homologs) recruiting the EC_5S ubiquitin-ligase complex are unknown. One of the goals of the present study was to understand the mechanism behind the recruiting process.

1.6. Ubiquitin specific protease 7 (USP7)

Ubiquitination, like other post-translational modifications, is reversible. Reversal of ubiquitination, or deubiquitination, is carried out by deubiquitinating enzymes (DUBs), which belong to the metallo and cysteine families of proteases. By reversing the actions of ubiquitin ligases, DUBs offer a way to fine tune the effects of ubiquitination as a post-translational modification [56].

Ubiquitin-specific protease 7 (USP7), also known as Herpesvirus-associated ubiquitin-specific protease (HAUSP), is a deubiquitinating enzyme found in all eukaryotes that catalyses the removal of ubiquitin from specific target proteins, preventing its degradation [57,58]. USP7 was first discovered as a protein in complex with the herpes simplex virus protein ICP0, a viral E3 ligase (15), and was termed HAUSP for herpes-associated ubiquitin-specific protease, to cooperatively facilitate viral replication [59]. Since then, further studies have shown that HAUSP can bind to various other substrates and is involved in the stress response pathway, epigenetic silencing, neurodegenerative disorders, and progression of infections by DNA viruses [60]. USP7 is 1102 amino acid residues long (128.3 kDa). It contains an N-terminal TRAF/MATH domain (from residue 54 to 208), a catalytic domain also called USP domain (from residue 214 to 521), and a 64 kDa C-terminal region with five Ub-like domains (Figure 1.6) [61].

Tumour Necrosis Factor (TNF)-receptor associated factors (TRAF) are a family of proteins that were initially identified for their capability of interacting with and regulating different members of the TNF receptors family. TRAFs are characterized by a C-terminal region encompassing about 180 amino acids, forming a 7-8 anti-parallel (3-sheets fold (TRAF-C domain), which is preceded by a coiled coil in theTRAF-N-terminal domain. Meprins, a family of extracellular metalloproteases, also contain a C-terminal domain with high sequence homology to the TRAF-C domain5 which is predicted to form a similar anti-parallel β-sheet fold. Therefore, the TRAF-C domain is also known as the meprin and TRAF Homology (MATH) domain [62].

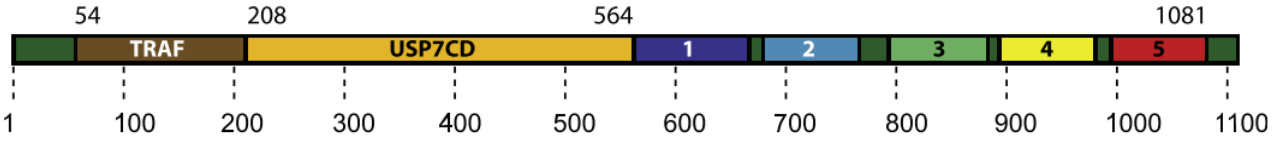

Figure 1.6- Schematic representation of USP7. *USP7 contains a TRAF substrate binding domain (brown), a catalytic domain (orange), and five Ub-like domains (rainbow). Adapted from* [63].

The only DUB/TRAF domain containing protein found in the human and mouse genomes is USP7. The crystal structure of USP7 revealed an eight anti-parallel β-sheet fold, very similar to the TRAF-C domain of TRAFs. The USP7 -TRAF domain (USP7-TRAF) has been implicated in substrate recognition [62].

Previous studies shown that the TRAF domain directly interacts with the substrates of USP7 such as p53, vIRF4, UbE2E1 and MDM2 [64–68]. The N-terminal domain of USP7 (USP7-NTD) recognizes a sequence motif ((P/ A)XXS) found in its interaction partners.

Another interesting interacting partner is LANA that has been shown to interact with USP7 through its TRAF domain [69]. This interaction involves a short sequence (amino acids residue 971 to 986) within the C-terminal domain of LANA with strong similarities to the USP7 binding site of the Epstein-Barr virus (EBV) EBNA-1 protein.

In response to genotoxic stress, USP7 binds and deubiquitylates p53 thereby protecting it from proteasome-mediated degradation [70]. Previous studies on the EBNA1-USP7 interaction have shown that EBNA1 binds the USP7-NTD, which is distinct from the catalytic domain, p53 is also shown to interact with the same domain. EBNA1 and p53 bind the same pocket in this domain but EBNA1 does so with an affinity that is approximately 10-fold higher than that of p53. As a result, EBNA1 interferes with the binding and stabilization of p53 by USP7 and with p53-mediated apoptosis in response to DNA damage [66]. The same authors also found that USP7 had a large stimulatory effect on the DNA-binding activity of EBNA1 in vitro and can form a ternary complex with DNA-bound EBNA1 [71].

This work reports the crystal structure of USP7-kLANA complex and also discusses the structural insights derived from this complex. Since EBV has a higher homology to KSHV, of the known human herpesviruses, the possibility of USP7 enhancing LANA DNA binding was also investigated.

2

MATERIALS AND EXPERIMENTAL PROCEDURES

2.1. Constructs Design

To facilitate the structural studies and assays of kLANA interactions, several constructs were designed. For the interaction studies of kLANA with Elongin BC, three kLANA constructs (kLANA$_{207-321}$, kLANA$_{207-295}$ and kLANA$_{188-295}$) were designed in the N-terminal domain of the protein. To design this constructs, was used the DisMeta server using the N-terminal domain, of kLANA sequence. DisMeta server, runs several programs simultaneously, including GlobPlot [72], ANCHOR [73], DISOPRED [74], among others, that mainly predicts disordered/unstructured regions and secondary structure elements. The two regions that were important to be incorporated in the constructs were the Elongin BC binding region (residues 210-222) and a helix (residue 281-293) for structural stability. Figure 2.1 is a schematic illustration of all LANA constructs used in this work. All mLANA constructs were kindly provided by Dr Pedro Simas, Viral Pathogenesis Unit, IMM and the others kLANA constructs (kLANA$_{51-321}$, kLANA$_{4-333^{\wedge}929-1162}$ and kLANA$_{4-32^{\wedge}888-1162}$) were cloned by Dr. Rajesh Ponnusamy, Structural Virology Lab, ITQB.

CBF-β was previously shown to be important for Elongin BC interaction with the Vif protein. [52,55]. For the interaction studies of LANA with Elongin BC and CBF-β, three versions of CBF-β (only vary on sequence size) and one construct of the Elongin BC complex were used. Full length Vif was also used in the assays as a positive control.

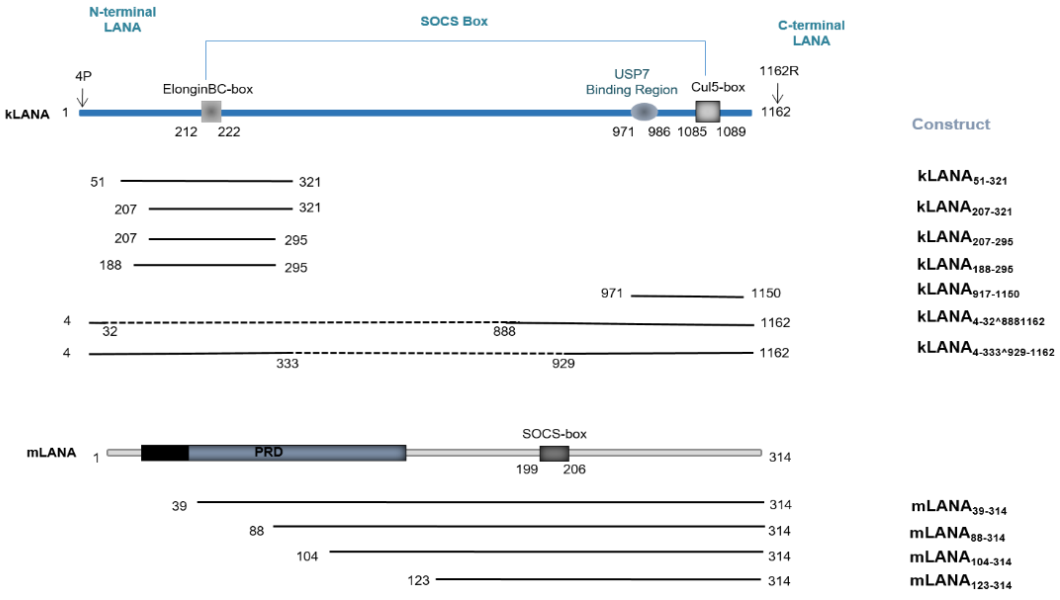

Figure 2.1- Schematic illustration of kLANA and mLANA constructs used in this study. *The numbers indicate the amino acid residue and the dashed lines represent deleted sequences.*

For the interaction studies of kLANA with USP7 protein, the construct kLANA$_{971-1150}$ (Figure 2.2) in the C-terminal of the protein that simply contains the USP7 binding region (residues 971-986) and the DNA binding domain. Regarding to USP7 constructs used, it contains the N-terminal tumour necrosis factor-receptor associated factor (TRAF)-like domain (residues 53-208) of USP7, showed to interact specifically with C-terminal sequences of kLANA, p53 and EBNA1 [65,66,69,75]. The first one, was from residues 54-205 and the second construct was simply to make the construct shorter, from residues 54-198, also illustrated in Figure 2.2.

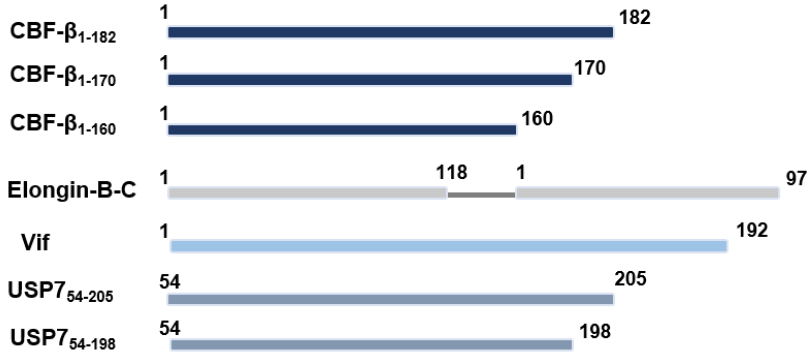

***Figure 2.2-Schematic illustration of the constructs used in this study**. The numbers indicate the amino acid residue.*

2.2. Cloning and Transformation

All constructs were cloned using the sequence and ligation independent cloning (SLIC) technique. This cloning strategy is based on the 3´-to-5´exonuclease activity of T4 DNA Polymerase. PCR products with complementary overhangs are created by incorporating appropriate extensions in the primers and treating them with T4 DNA polymerase. The annealing of the insert and the vector is made in the absence of ligase by simple mixing of the DNA fragments. More details about this technique are described elsewhere [76,77]. Figure 2.3 illustrates the procedure of SLIC. First, the vectors were linearized with the desired restriction enzymes. In most cases, digestion of 2 µg of the vector for 1h at 37°C, in a 50 µl reaction was performed. Then, linearized vectors were analysed with 1% agarose (Carl Roth) gel electrophoresis with SYBR safe DNA gel stain (Invitrogen), run at 80 V for 45min, follow by gel purification using QIAquick Gel Extraction Kit from Qiagen. Linearized vectors were treated with Shrimp Alkaline Phosphatase (SAP) enzyme.

Second, the insert is prepared by PCR using primers with 15-17-bp extensions (depending on the GC content) homologous to each end of the linearized vector. Primers were designed with Oligo Analyzer. All the primers used in this study were synthesized by SIGMA-ALDRICH.

Table 2.1- PCR cycling conditions.

Step	Temperature (°C)	Time	Cycles
Initial Denaturation	98	30 sec	1
Denaturation	98	10 sec	25-30
Annealing	Tm - 3°C	30 sec	
Extension	72	30 sec	
Final Extension	72	10 min	1

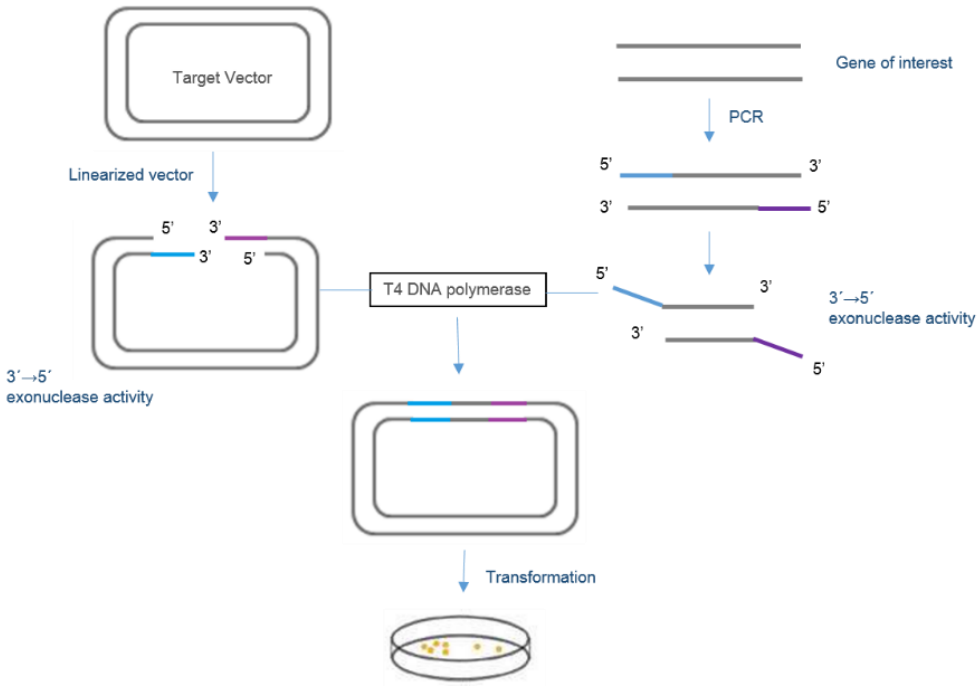

Figure 2.3- Schematic Illustration of the Sequence-Ligation-Independent Cloning (SLIC) technique

The PCR reaction components were: 100 µl total volume, 1µl Phusion DNA polymerase, 20 µl of 5× Phusion HF buffer; 2 µl of 10 mM dNTPs; 50 ng of DNA template and 5 pmol of each primer. The PCR cycling parameters are listed in Table 2.1. For the SLIC Reaction, 40 ng of vector and 80 ng of Insert are mixed and incubated at room temperature for 10 min with T4 DNA polymerase and 1x NEB Buffer 2.1 in a 10 µl solution. All enzymes and reaction components described here are from New England Biolabs.

Third, the reaction mixture is placed on ice for 10 min for single-strand annealing and then the DNA is transformed directly in DH5α, *Escherichia coli* competent cells. This mixture is incubated for 30 min on ice, and after heat shock at 42 °C for 45 sec, 900 µl of SOC medium was added to the mixture. The cells were grown in an incubator shaker, at 37ºC for 60 min at 180 rpm. The entire content of cell suspension is then plated onto a LB agar plate containing the appropriate antibiotics (50 µg/ml kanamycin, 50 µg/ml spectinomycin, 100 µg/ml Ampicillin and 34 µg/ml of chloramphenicol). The plates were then incubated at 37 °C, overnight.

Next day, 3 to 4 colonies from each construct were picked for colony PCR confirmation of each construct using GoTag DNA polymerase (Promega) and for inoculation in LB medium, with appropriate antibiotics for overnight culture of each clone for mini-prep. The DNA mini-prep was performed using the Minipreps DNA Purification Systems kit (Wizards Plus- Promega). Alternatively to PCR screening, confirmation was also done by restriction enzyme double digestion of 500 ng of DNA, in a 10 µl reaction at 37 °C for 30 min, analyzed by agarose gel electrophoresis for confirming positive clones. All the cloned sequences were finally confirmed through DNA sequencing by GATC Biotech labs.

2.3. USP7 chimeric protein

To generate the kLANA-USP7 chimeric protein, the complementary DNA sequence of kLANA peptide (residues 975-982) was inserted into the sequence of the forward primer, after the complementary sequence of the vector for SLIC cloning (5´-vector Sequence----LANA-peptide ---- USP7 sequence-3´).Then followed by the SLIC protocol described above.

The specific sequences of primers, vectors and restriction enzymes used in this study, for each construct, are listed in Table 2.2 and Table 2.3.

Table 2.2- Cloning details of kLANA constructs.

Protein (region)	Vector/ Resistance	Restriction site	Tag/ cleavage site	SLIC Primers (Tm)
kLANA$_{971-1150}$	pET47 b (+) / kanamycin	BamHI/NotI	His-3C protease	FW: 5-TCAGGGACCCGGGTACCAGGATCCGgacgatgacccacaacctggc-3´ (Tm=63 ºC) RW: 5'-GGCACCAGAGCGAGCTCTGCGGCCGCTTAttcccctggctgggttaatgg-3' (Tm=61 ºC)
kLANA$_{971-1150}$	pET49 b (+) / kanamycin	BamHI/NotI	GST-His-3Cprotease	FW:5-TCAGGGACCCGGGTACCAGGATCCGgacgatgacccacaacctggc-3´ (Tm=63 ºC) RW:5'-GGCACCAGAGCGAGCTCTGCGGCCGCTTAttcccctggctgggttaatgg-3' (Tm=61 ºC)
kLANA$_{207-321}$	pET47 b(+) / kanamycin	BamHI/NotI	His-3C-protease	FW: 5'-TCAGGGACCCGGGTACCAGGATCCGcccaaggtccctctacacta 3' (Tm=58 °C) RW:5'-GGCACCAGAGCGAGCTCTGCGGCCGCTTActtggatgcttcttctgcaatc 3´ (Tm=57 °C)
kLANA$_{207-321}$	pET49 b (+) / kanamycin	BamHI/NotI	GST-His-3C-protease	FW: 5'-TCAGGGACCCGGGTACCAGGATCCGcccaaggtccctctacacta 3' (Tm=58 °C) RW:5'-GGCACCAGAGCGAGCTCTGCGGCCGCTTActtggatgcttcttctgcaatc 3´ (Tm=57 °C)
kLANA$_{207-295}$	pET47 b(+) / kanamycin	BamHI/NotI	His-3C-protease	FW: 5'-TCAGGGACCCGGGTACCAGGATCCGcccaaggtccctctacacta 3' (Tm=58 °C) RW: 5´-GGCACCAGAGCGAGCTCTGCGGCCGCTTActtggatgcttcttctgcaatc 3´ (Tm=57 °C)
kLANA$_{207-295}$	pET49 b (+) / kanamycin	BamHI/NotI	GST-His-3C protease	FW: 5'-TCAGGGACCCGGGTACCAGGATCCGcccaaggtccctctacacta 3' (Tm=58 °C) RW: 5´-GGCACCAGAGCGAGCTCTGCGGCCGCTTActtggatgcttcttctgcaatc 3´ (Tm=57 °C)
kLANA$_{188-295}$	pET47 b(+) / kanamycin	BamHI/NotI	His-3C-protease	FW: 5´-TCAGGGACCCGGGTACCAGGATCCGccagactctttagctccgtctac-3´ (Tm=58 °C) RW: 5´-GGCACCAGAGCGAGCTCTGCGGCCGCTTActtggatgcttcttctgcaatc 3´ (Tm=57 °C)
kLANA$_{188-295}$	pET49 b (+) / kanamycin	BamHI/NotI	GST-His-3C-protease	FW: 5´-TCAGGGACCCGGGTACCAGGATCCGccagactctttagctccgtctac-3´ (Tm=58 °C) RW: 5´-GGCACCAGAGCGAGCTCTGCGGCCGCTTActtggatgcttcttctgcaatc 3´ (Tm=57 °C)

Table 2.3- Cloning details of CBF-β, Vif and USP7 constructs.

Protein (region)	Vector/ Resistance	Restriction site	Tag/ cleavage site	SLIC Primers (Tm)
CBF-β_{1-160}	pNEATH / Ampicillin	NdeI / BamHI	His-3C-protease	FW: 5´-GTGCCGCGCGGCAGCCATATGCTTGAAGTCCTCTTTCAGGGACCCatgccgcgcgtcgtgc-3´ (Tm=55 °C) RW: 5´-GCTAGCTCTAGACTATTAGGATCCTTAgtctctatcttcaaattcgcgtg-3´(Tm=57 °C)
CBF-β_{1-170}	pNEATH / Ampicillin	NdeI / BamHI	His-3C-protease	FW: 5´-GTGCCGCGCGGCAGCCATATGCTTGAAGTCCTCTTTCAGGGACCCatgccgcgcgtcgtgc-3´ (Tm=55 °C) RW: 5´-CTAGCTCTAGACTATTAGGATCCTTActgtgaaactctcacctccatttc-3´(Tm=57 °C)
CBF-β_{1-182}	pNEATH / Ampicillin	NdeI / BamHI	His-3C-protease	FW: 5´-GTGCCGCGCGGCAGCCATATGCTTGAAGTCCTCTTTCAGGGACCCatgccgcgcgtcgtgc-3´ (Tm=55 °C) RW: 5´-GCTAGCTCTAGACTATTAGGATCCTTAgggtcttgttgtcttcttgcc-3´ (Tm=58 °C)
Vif_{1-192}	pET47 b(+) / Kanamycin	BamHI / XhoI	His-3C-protease	FW: 5´-GACCCGGGTACCAGGATCCTatggaaaacagatggcaggtgatg-3´ (Tm=61 °C) RW: 5´-AGGTTAATTAAGCCTCGAGctagtgtccattcattgtatggctcc-3´ (Tm=59 °C)
USP7-$TRAF_{54-206}$	pET47 b(+) / Kanamycin	BamHI / NotI	His-3C-protease	FW: 5´-TCAGGGACCCGGGTACCAGGATCCGaccgcggaggaggacatggag-3´(Tm=65 °C) RW: 5´-GGCACCAGAGCGAGCTCTGCGGCCGCTTAcacgcaactccatggggagc-3´ (Tm=65 °C)
$USP7_{54-198}$	pNEATH / Ampicillin	NdeI / BamHI	His-3C-protease	FW: 5´-CATCATCACAGCAGCGGTACCCTTGAAGTCCTCTTTCAGGGACCCCATATGaccgcggaggaggacatg-3´(Tm=62 °C) RW: 5´-GCTAGCTCTAGACTATTAGGATCCTTAatccgcctgtacaaagacttc-3´ (Tm=57 °C)
USP7-kLANA chimeric Protein	pET47 b(+) / Kanamycin	BamHI / XhoI	His-3C-protease	FW: 5´-TCAGGGACCCGGGTACCAGGATCCGaccgcggaggaggacatggag-3´(Tm=65 °C) RW: 5´-AGCCTAGGTTAATTAAGCCTCGAGTTATTCGCGAGATGGGCCAGGTTGgggagcatccgcctgtac-3´(Tm=60 °C)

2.4. Small Scale Expressions and Pull down assay´s

Sequence-verified clones were transformed into chemically competent BL21 Star (DE3) pRARE2 cells for protein expression. These cells have chloramphenicol resistance. Freshly grown colonies were used to inoculate 3 ml overnight cultures of ZY-0.8G medium. All overnight cultures were grown in 50 ml Falcon tubes, in an incubator shaker at 37 °C and approximately 180 rpm. In the next day, this overnight culture was used to set up a small scale cultures in 5-7 ml ZYP-5052 auto induction medium with a 0.07 starting OD. The cells were grown at 37 °C for 2h followed by 18 °C overnight. All media were supplemented with the appropriated antibiotics; cells were then harvested by centrifugation at 5000 rpm, for 15 min at 4 °C.

Small scale expression trials are normally used to determine the best conditions to express a particular protein. It´s normally tested in different media, with different temperatures, different concentrations of inducing agent and incubation times. In most cases of the proteins used in this study, the procedure was already optimized, where auto induction medium turn out to be the best condition for expression.

For pull down assays, a co-transformation of 2 or 3 vector constructs was required. When using more than 3 antibiotics, the concentration of each is reduced by half to enable cell growth. The protocol used for co-expression is the same as described above.

For small scale expression trials, 2 ml of cells were pellet and resuspended in 1 ml of homemade BugBuster solution (25 mM Na/k Pi, 150 mM NaCl, 25% Sucrose, 5 mM $MgCl_2$, 1% Triton-X, pH 8.5) for cell lysis, supplemented with 0.25 mg/ml of lysozyme and 5 unit/ml of Omicleave nuclease (Epicentre). The resuspended cells were incubated with agitation, for 30 min at 4 ºC and cell lysates were clarified by centrifugation a16000g for 15 min at 4 °C.

Depending on the construct in question, either Ni-NTA agarose beads (Qiagen) for His-tag proteins or Gluthathione agarose beads (Thermo Scientifics) for GST-tag proteins were used. For affinity purification, 40 µl of beads was added to clarified lysates and incubated with agitation at 4 °C. The beads slurry was previously washed 3 times with cold PBS. After 1h incubation the beads were washed 2 times with 1 ml of cold PBS. In the case of Ni-NTA agarose beads, the PBS was supplemented with 20 mM of imidazole. After the washing steps, the beads pellet was boiled at 95 °C for 10 min in 50 µl of SDS sample loading buffer. The solution was centrifuged at maximum speed for 10 min, and then 10 µl loaded in a 15 % tricine-SDS-PAGE gel [78] for electrophoresis analysis.

For total cell lysate (TCL) analysis, 50 µl of culture pellet was resuspended in 45 µl of Tris-HCl pH 7.5, and 5 µl of 20% SDS, boiled at 95 °C for 10 min and centrifuged at maximum speed for 10 min and 5 µl loaded in a tricine-SDS-PAGE gel. Protein expression was confirmed by western blot analysis [79] using the appropriate antibody.

2.5. Large Scale Expressions and Purification

All proteins were expressed in the same conditions as identified by small-scale trials: Cultures of ZY-0.8G medium supplemented with the appropriate antibiotics were prepared with freshly grown colonies. The cells were grown overnight in an incubator shaker, 180 rpm at 37 °C. ZYP-5052 auto-induction medium was freshly prepared and supplemented with the appropriated antibiotics. 500 ml of culture, with a starting OD of 0.07 was grown, in 2L Flasks, at 37 °C for 2h follow by 18 °C overnight in an incubator shaker with a cooling system. To harvest the cells, the culture was centrifuge at 8000 rpm for 20 min at 4 °C. The cell pellets were withier used immediately, or frozen and store at -20 °C.

2.6. Purification of kLANA truncations

The cell pellet of 1L culture (~13 g of dry pellet) of the GST-tagged version of $kLANA_{971-1150}$ was resuspended by vigorous stirring for 30 min at 4 °C, in 150 ml of buffer A (25 mM Na/K Phosphate, 500 mM NaCl, 10 % (v/v) glycerol, pH 7.5) supplemented with 1 M NDSB, 1 mM DTT, 10 mM $MgCl_2$, 0.25 mg/ml of lysozyme, 5 Units/ml of omicleave (epicentre) and one tablet of EDTA-free protease inhibitor cocktail (Roche). Cells were then lysed using a Branson 450D sonicator with 10% of amplitude, 30 sec of pulse on followed by 30 sec of pulse off, for 2-3 min. Cell lysates were clarified by centrifugation at 19000 rpm, for 40 min at 4 °C.

All the purification steps were performed using an ÄKTA Explorer 10 FPLC System at room temperature (GE Healthcare). Clarified lysates were applied to 2 x 5 ml GSTrap™ 4B columns (GE Healthcare), connected in tandem, using a peristaltic P-1 pump (GE Healthcare), at 4 °C. Unbound proteins were washed with buffer A and the protein was eluted in buffer B (25 mM Na/K Phosphate, 500 mM NaCl, 10 % (v/v) glycerol, 20 mM Glutathione, pH 7.5), using a two steps gradient.

The main fractions were pooled and applied to a 5 ml HiTrap Heparin HP columns (GE Healthcare); the column was washed in buffer C (25mM Na/K Phosphate, 200 mM NaCl, pH 7.5) and the protein eluted using a linear gradient in buffer D (25mM Na/K Phosphate, 2 M NaCl, pH 7.5). The main fraction peak, was pooled, concentrated and injected onto a Superdex 200 Increase 10/300 GL column (GE Healthcare) which was pre equilibrated with buffer E (25 mM Na/K Phosphate, 500 mM NaCl, 5% (v/v) glycerol, pH 7.0). Purity of the protein was monitored at all stages of the purification process using a 12% SDS-PAGE gel electrophoresis, run at 200 V for 50 min. All kLANA protein truncations were purified in a similar manner to $kLANA_{971-1150}$.

2.7. USP7-TRAF domain Purification

The cell pellet of 1L culture of the His-tagged version of $USP7_{54-205}$ was resuspended in 150 ml of buffer F (25 mM Na/K Phosphate, 500 mM NaCl, 10 % (v/v) glycerol, 20 mM imidazole, pH 8.0), supplemented with 10 mM $MgCl_2$, 0.25 mg/ml of lysozyme, 5 Units/ml of omicleave and one tablet of EDTA-free protease inhibitor cocktail and incubated for 30 min, at 4 ºC with constant agitation. Cells were lysed using a Z Basic cell disruptor (Constant Systems Ltd.). Cell lysates were clarified by centrifugation at 19000 rpm, for 40 min at 4 ºC.

Clarified lysates were applied to two 5 ml HisTrap HP columns (GE Healthcare), with a peristaltic pump at 4 °C. The columns were washed with buffer F and the protein was eluted using a linear gradient with buffer G (25 mM Na/K Phosphate, 500 mM NaCl, 500 mM imidazole, pH 7.5). The main fractions were pooled and 3C protease was added in a 1:100 molar ratio with an addition of 2 mM DTT and 1 mM EDTA. The protein was cleaved under agitation, overnight at 4 ºC. In the next day, to separate cleaved protein from uncleaved protein, the protein was passed through a 5 ml GSTrap HP column (GE Healthcare). Then, the protein was concentrated and injected into a HiLoad™ 16/600 Superdex™ 75 pg column (GE Healthcare), which was pre equilibrated in buffer H (20 mM Tris-HCl, 100 mM NaCl, pH 7.5). Purity of the protein was monitored at all stages of the purification process using a 12% SDS-PAGE gel electrophoresis, run at 200 V for 50 min.

All USP7 protein truncations were purified in a similar manner to $USP7_{54-205}$.

2.8. Thermofluor experiments

The fluorescence-based thermal shift assay or differential scanning fluorimetry (DSF) is a technique commonly use to access the optimal stabilizing buffer components to achieve high purification yields [80]. It is also widely used to identify the potential buffers and additives necessary to crystallize proteins [81,82].

The thermofluor experiment requires the presence of a fluorescent dye that will interact with exposed hydrophobic regions. It´s principle is based on the detection of changes in the exposure of a protein´s hydrophobic core upon heat denaturation. For many proteins the gradual increase of temperature has little effect on the protein fold, until a temperature is reached where it will quickly unfold. At this point, the unfolded protein will expose its hydrophobic core and the dye will become fluorescent. In an ideal case, it´s possible to draw a sigmoidal curve that allows the calculation of a melting temperature (T_m), which corresponds to the temperature where 50% of the protein is unfolded, and this allows the determination of the buffer more stable for a protein, by comparing the several melting temperatures of the protein in different buffers [80].

Thermofluor experiments were performed with USP7 N-terminal TRAF domain, using the $USP7_{54-205}$ truncation, to access the protein stability at different pH´s, and when incubated with three different peptides. The peptides used were: two from the N-terminal part of kLANA protein, k10 with the amino acid sequence: ^{975}DPQPGPSREY^{984} and k16 with the amino acid sequence: ^{971}PDDDPQPGPSREYRYV^{986}; the last peptide is from p53 protein with the amino acid sequence: ^{358}EPGGSRAHSS^{367}. There are several studies

about the interaction of USP7 with p53, so this protein will be used as a control in the experiment. All peptides were synthesized by SIGMA-ALDRICH.

This USP7 protein is very soluble, which hampers the growth of protein crystals. With the goal of producing high-quality crystals, buffer screening was performed to find the optimal condition where the protein is less soluble, in a way that favours the formation of crystals. Protein was purified by Ni-NTA and the tag was cleaved using 3C protease followed by gel-filtration purification. The protein was in 10 mM Tris-HCl, 100 mM NaCl, pH 7.5.

For the thermofluor assay a real-time PCR Detection System iQ5™ (Bio-Rad) was used. The assay was performed using a 96-well thin-wall PCR plate (Bio-Rad). The total reaction volume was 20 µl and tested 24 different buffers for each sample, from the JBS Solubility kit (Jena Biosciences). The peptides were incubated with the protein in a 2:1 molar ratio, overnight at 4 °C. The final amount of protein in the assay was 2.5 µg and a 5-fold (final concentration) of SYPRO Orange (Invitrogen) was used. For dilution buffer it was used 25mM HEPES pH 7.5.

After pipetting of samples, the plates were sealed with optical quality sealing tape (Bio-Rad) and centrifuged at 4000 rpm for 1 min. For the thermofluor experiment the plates were heated from 20 °C to 90 °C in increments of 1 °C.

From this experiment, six buffers were chosen in a pH range from 4.0 to 8.5, and another thermofluor assay was performed, following the protocol described above. Nonlinear regression analysis in the curve-fitting program GraphPad Prism was used to construct melting curves and calculate Tm values for each sample condition. All experiments were carried out in triplicate.

2.9. Electrophoretic Mobility Shift Assay

The electrophoretic mobility shift assay (EMSA) is a commonly used method to identify DNA binding proteins. It is based on the fact that labelled DNA with bound protein migrates more slowly through a polyacrylamide gel than the corresponding protein-free DNA [83]. In theory, an EMSA is very easy and simply to perform, but in reality a clean and successful gel shift requires the optimization of a number of parameters. Crucial parameters, include binding buffers (salt, glycerol concentration and pH), non-specific competitor DNA (it is important to titrate an amount of nonspecific DNA, to avoid non-specific binding) and non-denaturing gel conditions [84].

For this EMSA assay, the $kLANA_{4\text{-}32\text{^}888\text{-}1162}$ and $mLANA_{123\text{-}314}$ proteins were used, against $USP7_{54\text{-}205}$ truncation. DNA probes for LANA EMSAs were generated by 5´-end fluorescein (Flc) labeling with and annealing to its complementary sequence. The Flc-DNA was synthesized by SIGMA-ALDRICH. The probes used were:
kLBS1-2 (5´-TTTGACGCCGCCGGGGCCTGCGGCGCCTCCCGCCCGGGCATGGGGCCGC-3') for kLANA and mLBS1-2 (5´- TATAATGGATCCGGCGCCATGCGCCCGCGCCTGGGGCGCCACGCGGCG-3´) for mLANA, correspond to the LBS1-2 site of the terminal repeat (TR) DNA sequences of KSHV and MHV-68, respectively.

The assays contained LANA and either USP7 or BSA (used as a negative control). LANA was incubated with USP7 or BSA at room temperature for 30 min, prior to adding the labelled DNA. For both USP7 and BSA, the protein concentration used was 40 µM. LANA concentration was titrated from 0.05 µM to 0.7 µM.

Protein mixtures were incubated with 0.2 µM of DNA probe, at room temperature for 30 min, in the presence of 0.63 µg of PolydI-dC (SigmaAldrich), in binding buffer (20 mM Na/K phosphate, 300 mM potassium chloride, 10 mM magnesium chloride, 1 mM EDTA, 10 % (v/v) glycerol, 0.1 % (v/v) Tween-20).

IEF sample buffer (50% (w/w) Glycerol, Bio-Rad) was added to incubations prior to loading on a 5% non-denaturing TBE-polyacrylamide gel. The gel was run in TBE buffer for 3 hour at 200 V at 4 °C, and revealed using a Fuji Scanner.

2.10. Crystallization and Structure determination

X-ray crystallography is a high quality approach to obtain the three-dimensional structure of macromolecules with atomic resolution, providing essential information regarding the molecular mechanism underlying the function. It also provides powerful insights about the way two molecules interact with each other. This technique is possible through the crystallization of the protein(s) of interest. A crystal is an orderly three-dimensional array of molecules, held together by non-covalent interactions. In a very simply way, when X-ray interact with the crystal, most of the X-ray light will be scattered by the electrons in the molecule, giving rise to a diffraction pattern. Since the crystal is formed by the repetition of individual unit cells through its entire volume, displayed by symmetry operations, it can be treated as three-dimensional diffraction. From the diffraction pattern we can determine the structure [85,86].

To determine a 3D structure of a protein by X-ray crystallography there are several steps involved, from cloning, expression and purification of the protein to the first attempt on crystallizing the protein, where we use sparse matrix screens, which are in fact a collection of diverse conditions of buffers, precipitants and other additives to search for initial crystallization conditions; After identifying "initial hit" conditions, the second step is the crystal optimization; Once diffraction-quality-crystals are produced, a X-ray diffraction data set is collected; The last step is the structure determination and refinement of the 3D model [85,87]

Co-crystallization screens of $USP7_{54-205}$ (29 mg/ml) and $USP7_{54-198}$ (26 mg/ml) with peptides k10 and k16, and crystallization screens of USP7-Chimera (29 mg/ml) were performed. $USP7_{54-205}$ and $USP7_{54-198}$ were incubated with the peptides, in a 1:2 molar ratio, for 1h at 4 °C. The crystallization screens were performed using a Cartesian MiniBee robot (Genomics Solutions), using two sparse matrix formulations from commercial screens (JCSG-plus and PACT-Premier) from molecular dimensions. The sitting-drop vapor-diffusion method was used, with three different drops size: 100 nl of protein mixed with equal volume of the reservoir screening solution, 100 nl of protein using the double of the volume of the reservoir screening solution and 200 nl of protein using half of the volume of the reservoir screening solution. The crystallization drops were imaged regularly using Minstrel DT UV imager (Rigaku).

For optimization, plates were set-up manually using the hanging-drop vapour-diffusion method. While the co-crystallizations of USP7 failed, for USP7-Chimera crystals grown to full size in two to three days, at 20 °C in two different conditions: 0.1 M Tris-HCl 7.5, 0.2 M calcium chloride and 9% (w/v) PEG 2000 MME (Crystals also grown in the same condition, in pH 7.0 and pH 7.5 and in 10% and 11% (w/v) of PEG) and 0.1 M Tris-HCl pH 9.0, 0.2 M magnesium chloride and 14% (w/v) PEG 1500; For this last condition, crystals were further optimized by serial streak-seeding in increasing precipitant concentrations. Crystals grown to an optimal size in 7 days.

In another attempt, to co-crystallize USP7 with k10 and k16 peptides, $USP7_{Chimera}$ crystals were used as seeds for streak-seeding, in clear drops. After several attempts, $USP7_{54-198}$ co-crystallized with the k16 peptide, as small needle shaped in 0.1M Tris-HCl pH 8.5, 0.2 M lithium sulphate and 26% (w/v) PEG 3350. Further optimization to produce diffraction-quality-crystals failed.

For USP7-chimera crystals, several cryoprotectant solutions and different cryo-soaking techniques were tested; The optimal way of cryo-cool-down the crystals, was using a two-steps approach, where the crystal is first soak in half of the cryosolution concentration, and then transferred to the final cryosolution concentration; The best cryosolution found was the solution with an PEG concentration increase of 3% (w/v) and the addition of 30% (v/v) Ethylene Glycol.

All diffraction datasets were collected at Diamond Light Source (Harwell campus, Oxford), in beamline I04. The data were indexed and integrated using iMOSFLM [88], then merged and converted with POINTLESS and scaled in SCALA [89]. The structure was solved by molecular replacement using PHASER [90], with the TRAF-like domain of USP7 (PDB ID: 2F1W) as a search model; Model building was performed using COOT [91] and model refinement with REFMAC5 [92]. Structural illustrations were prepared with PyMOL (v.1.7).

3

RESULTS

3.1. LANA recruiting the EC_5S Ubiquitin Complex

3.1.1. Cloning

To identify strategies that could result in detecting LANA interaction with the EC_5S Ubiquitin complex components, a number of prokaryotic plasmid expression vectors were constructed and tested for LANA and it´s predicted EC_5S interaction partners.

LANA N-terminal constructs $kLANA_{207-321}$, $kLANA_{207-295}$ and $kLANA_{188-295}$ were cloned into two vectors: pET47 b(+), which incorporates a 6xHis-tag and pET49 b(+) which has both a GST and 6xHis-tag. All constructs were kanamycin resistant, and included a 3C protease cleavage site after the tag, for protein tag removal. pET47 b(+) and pET49 b(+) vector maps are illustrated in Figure S1A and S1B, respectively. For SLIC cloning, both vectors were digested with BamHI and NotI restriction enzymes, and then treated with SAP to prevent relegation. To confirm complete digestion and purification, the linear vectors were run on a 1 % agarose gel (Figure 3.1a).

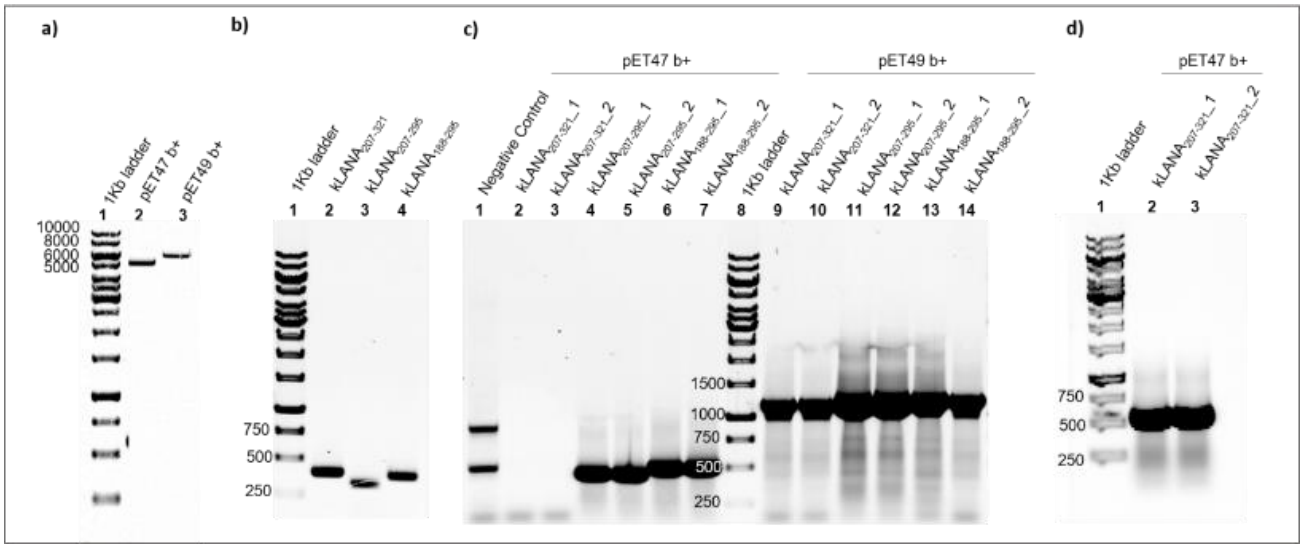

Figure 3.1- Cloning steps of kLANA constructs. *(a) Linearization of vectors using BamHI and NotI restriction enzymes; lane 2: pET47 b(+) (5203 bp) and lane 3: pET49 b(+) (5926 bp); b) PCR amplification of inserts with the specific SLIC primers; lane 2: $kLANA_{207-321}$ (345 bp); lane 3: $kLANA_{207-295}$ (267 bp); lane 4: $kLANA_{188-295}$ (324 bp); (c) Clone validation by PCR using GoTaq.*

Inserts were amplified by PCR (Figure 3.1b) with gene specific primers containing 10-15 bp of overlapping sequence from either the BamHI or NotI-cleaved vector end. Vectors and inserts were then mixed and treated with T4 DNA polymerase. After transformation, colony PCR was performed to identify positive clones. Figure 3.1c illustrates the positive colonies from the colony PCR as analysed by 1% agarose gel that show the presence of all inserts, except in the case of $kLANA_{207-321}$ construct cloned into pET-47 b(+). For this case, the SLIC reaction was successfully repeated (Figure 3.1d). All positive were confirmed by DNA sequencing.

The strategy planned to study the interaction between LANA and its ubiquitin ligase partners involves co-expression of genes and protein pull-down assays to confirm binding activity. The easiest approach for co-expression of multiple genes in the same *E. coli* host is to use vectors with different resistance markers, each vector bearing a single gene [93]. Thereby, the cloning of kLANA, CBF-β as well the experimental control protein, Vif, was carefully planned to allow for co-expression in compatible vectors.

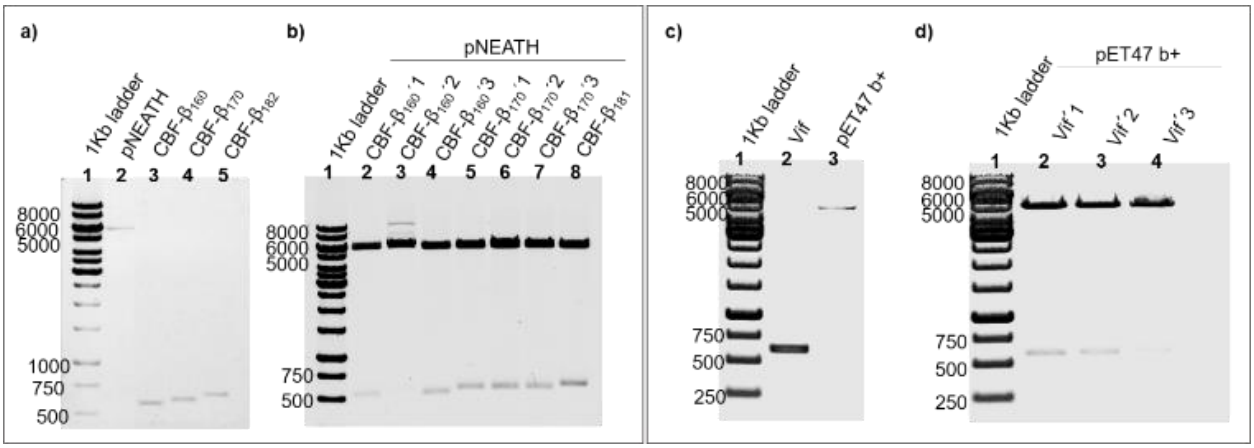

Figure 3.2- Cloning steps of CBF-β and Vif constructs*. (a) Vector linearization and PCR amplification of CBF- β truncations. Lane2: pNEATH (5763 bp) linearization with NdeI and BamHI restriction enzymes; Lane 3: PCR amplification of CBF-β160 (507 bp); Lane 4: PCR amplification of CBF-β170 (546 bp); Lane 5: PCR amplification of CBF-β182 (582 bp); b) CBF-β clone validation by restriction digestion; c) Vector linearization and PCR amplification of Vif. Lane 2: PCR amplification of Vif (579 bp); Lane 3: pET-47b(+)(5203 bp) linearization with BamHI and XhoI restriction enzymes. d) Vif clone validation by restriction digestion.*

kLANA truncations were cloned into pET vectors, (kanamycin resistant) and Elongin BC was already cloned, by a member of the lab, into pCDF-Duet (spectinomycin resistance). For CBF-β truncations, (CBF-β1-160, CBF-β1-170 and CBF-β1-182) the pNEATH vector was chosen and has ampicillin resistance, allowing the co-transformation of the three plasmids. pNEATH vector also includes a 6xHis-tag (Vector map illustrate in Figure S1C).

CBF-β cloning was performed in a similar way to kLANA constructs, where first the pNEATH vector was linearized with NdeI and BamHI restriction enzymes (Figure 3.2a). Clone validation was performed by restriction enzyme digestion (Figure 3.2b).

The Vif protein was used as a control experiment to validate the pull down assays performed with LANA truncations, as it has been previously shown to interact with Elongin BC and CBF- β [49,53–55]; Vif was cloned in the same vector as kLANA truncations, pET 47 b(+), using BamHI and XhoI as restriction enzymes. Figure 3.2c and d show Vif cloning steps. Both CBF-β and Vif constructs were successfully cloned.

3.1.2. Expression of kLANA N-terminal domain truncations

As a starting point of the study, expressions tests were performed, focusing only on the soluble fractions to evaluate the expression of the genes and also to access the purification ability with affinity beads in these constructs.

Three different media were tested: LB, TB and ZYP-5052 auto-induction medium. The GST version of kLANA (cloned into pET 49 b(+) vector) and the corresponding glutathione agarose beads were used. Figure 3.3a, b and c show the purification results in Tricine gels for GST- kLANA207-321, GST-kLANA207-295 and GST- kLANA188-295, respectively.

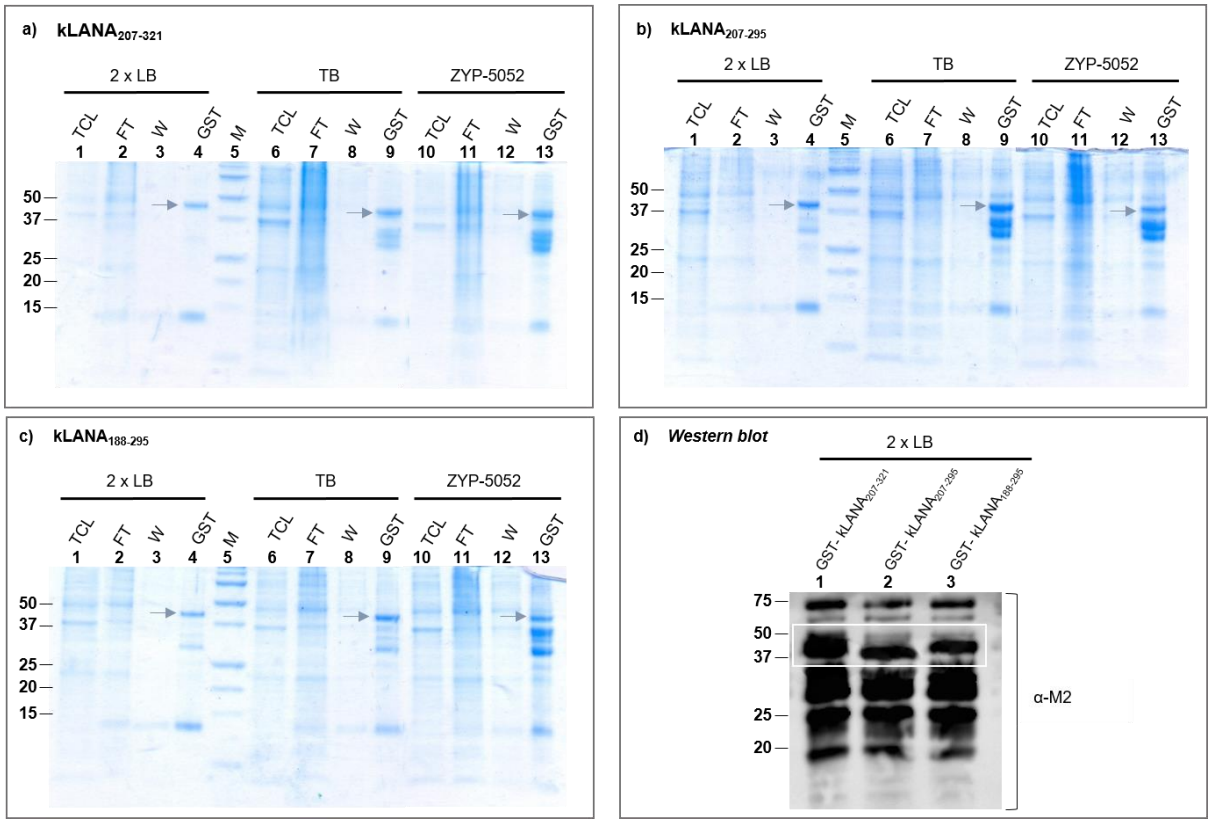

Figure 3.3- Expression test of kLANA N-terminal truncations*. Small scale expressions of GST-kLANA in three different mediums were prepared; the cells were lysed using a BugBuster solution, and GST-agarose beads were added to clarified lysates. (a), (b), (c) resulting fractions from pull down of GST- kLANA$_{207-321}$ (38 kDa), GST- kLANA$_{207-295}$ (35 kDa) and GST- kLANA$_{188--295}$ (37 kDa), respectively, were run on a 12% Tricine SDS-PAGE gel and stained with Coomassie blue. (d) Further confirmation of the proteins was performed by immunoblotting using α-M2 antibody that recognizes the DYKDDDDK peptide, present on the GST tag. Arrows indicate the expressed proteins. TCL- Total cell Lysate, FT- Flow through from beads purification, W-Wash of unbound proteins, GST- Beads purification.*

All genes were expressed, in each medium; the proteins were also successfully purified by the affinity agarose beads, of which 2xLB medium contained less contaminants. Hence, purified fractions from 2xLB medium was used for immunoblotting assay (Figure 3.3d); although there is a lot of unspecific binding of the antibody to other proteins, it was possible to confirm the presence of three kLANA truncations, indicated by the box, through its different molecular weights.

3.1.3. CBF-β does not bind to N-terminal kLANA truncations

The core-binding factor subunit beta, CBF-β, is a host protein required for preserving HIV-1 infectivity [52]. It is hijacked by HIV-1 Vif to form the Vif-CBF-β-CUL5-ELOBC complex, for Vif-mediated degradation of APOBEC3G [49,52,55]. It is possible that LANA may recruit CUL5-EloBC, similar to HIV-1 Vif protein, by interacting with CBF-β. Since HIV-1 Vif has a similar function to LANA, a simple pull down experiment was performed in order to investigate the possible interaction between LANA and CBF-β.

His-CBF-β was co-transformed and expressed with GST-His-kLANA. , Glutathione agarose beads were used to bind GST-His-LANA for pull down experiments. Ni-NTA agarose beads were used to identify the expression of LANA and CBF-β (Figure 3.4). As another control, CBF-β$_{1-170}$ was expressed alone and purified with Ni-NTA beads (Figure 3.4a, lane 5 and Figure 3.4b, lane 10).

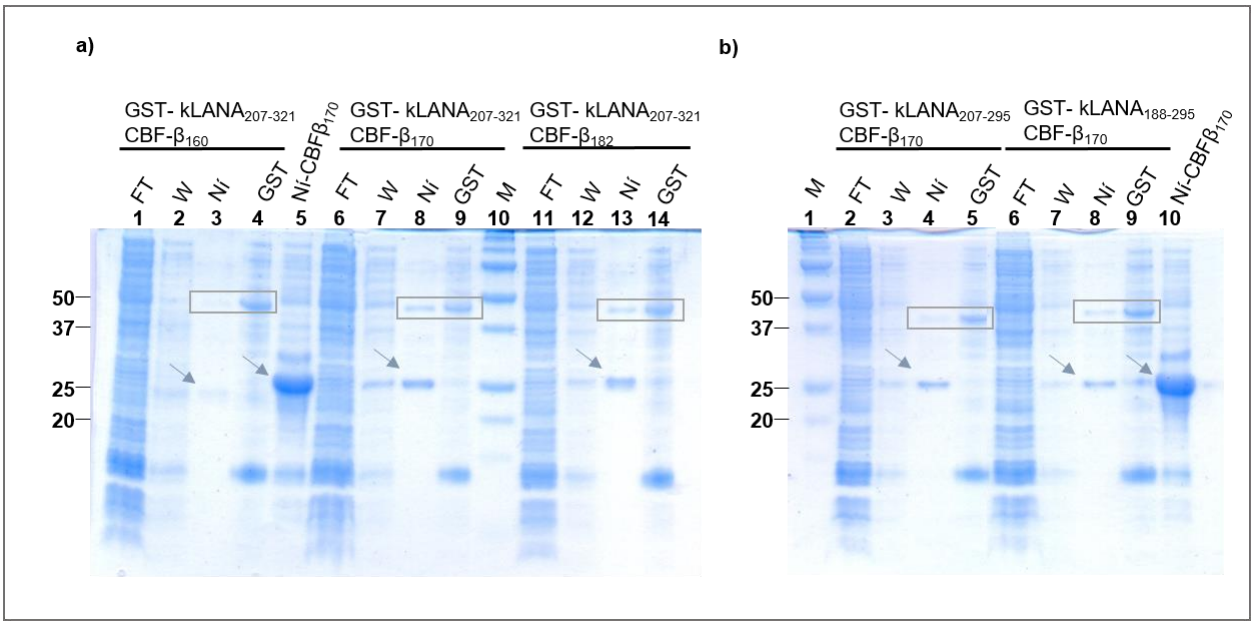

Figure 3.4- CBF-β do not interact directly with kLANA N-terminal truncations. *Co- expressions of GST-kLANA with CBF-β were performed. The cells were lysed using a BugBuster solution, and GST-agarose beads were added to clarified lysates. (a) GST- kLANA$_{207-321}$ (38 kDa) assay with three CBF-β truncations: His-CBF-β$_{1-160}$ (19 kDa), His-CBF-β$_{1-170}$ (21 kDa) and His-CBF-β$_{1-182}$ (22 kDa); the resulting fractions from pull down of GST- kLANA$_{207-321}$, were run on a 12% Tricine SDS-PAGE gel and stained with Coomassie blue. Both proteins were possible to be identified in the gel, i.e. they were expressed, but LANA was not able to pull down CBF- β. The same result was obtained with the other two constructs of LANA, (b) GST- kLANA$_{207-295}$ (35 kDa) and GST- kLANA$_{188-295}$ (37 kDa). Arrows indicate the localization of CBF-β truncations and the boxes the localization of LANA truncations. FT- Flow through from beads purification, W-Wash of unbound proteins, GST agarose beads purification, Ni-NTA agarose beads purification.*

As a first trial, kLANA$_{207-321}$ was used to perform the assay with all CBF-β truncations (Figure 3.4a). In Figure 3.4a lanes 3, 8 and 13 it is possible to identify kLANA$_{207-321}$ and CBF-β from the Ni-NTA agarose beads pull down, indicating that both proteins were expressed. For the GST-pull down the same result does not occur. From lanes 4, 9 and 14, kLANA$_{207-321}$ was clearly bound to the beads, but CBF-β was not. The second assay (Figure 3.4b) was performed using two other N-terminal kLANA truncations, co-expressed along with CBF-β$_{1-170}$. The same conditions from the previous assay were used, and similar results were obtained. Again, both proteins (kLANA and CBF-β) are identified in the Ni- purification, (Figure 3.4b, lanes 4 and 8) but not in the (GST-pull down, Figure 3.4b, lanes 5 and 9).

These assays were repeated, with all possible combinations between LANA and CBF-β truncations, and similar results were obtained.

3.1.4. kLANA does not bind to CBF-β and Elongin BC complex

Without positive results from pull down assays of kLANA with CBF-β, Elongin BC co-expression was added to the expression trials. The rationale being that CBF-β may only interact with LANA in the presence of other Ubiquitin complex components, like the adaptor Elongin BC, already shown to interact with the N-terminal part of LANA [43].

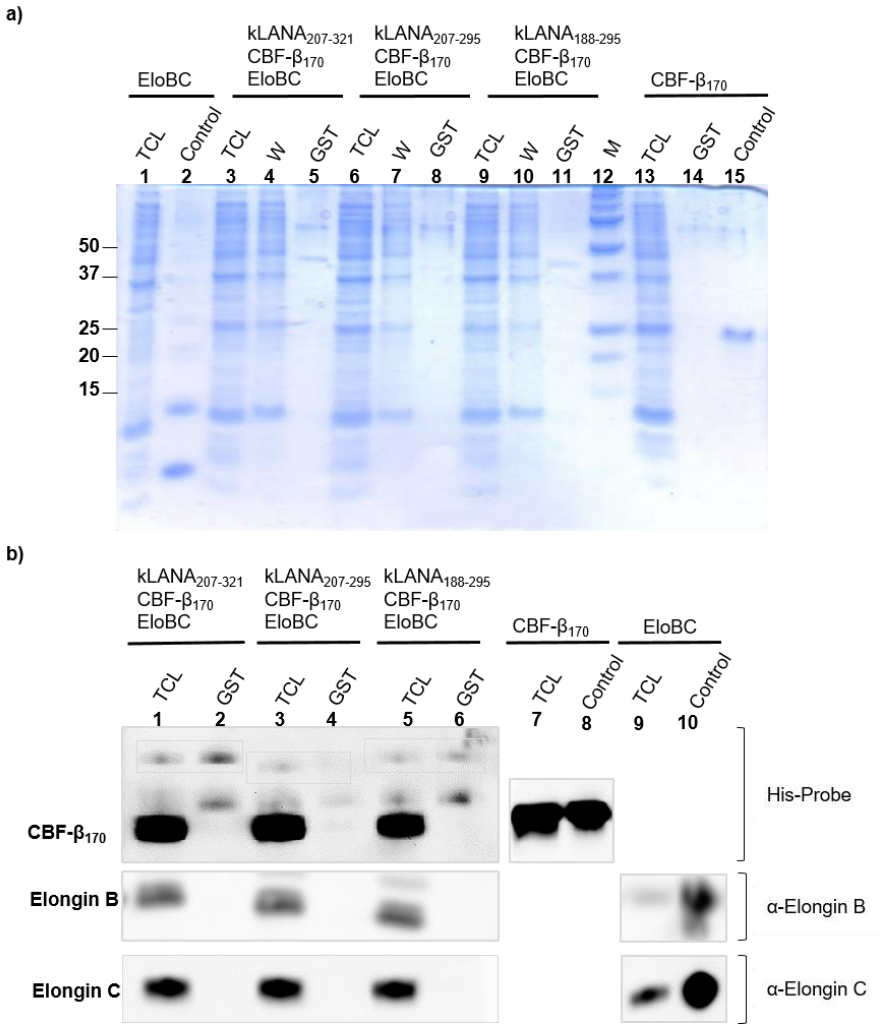

Figure 3.5- N-Terminal kLANA truncations do not interact directly with CBF-β neither with Elongin BC proteins. *GST-kLANA truncations were co-expressed with His-CBF-$β_{1-170}$ (21 kDa) and Elongin-B (13 kDa) and Elongin C (11 kDa). The cells were lysed using a BugBuster solution, and GST-agarose beads were added to clarified lysates. (a) GST- kLANA$_{207-321}$ (38 kDa), GST- kLANA$_{207-295}$ (35 kDa), GST- kLANA$_{188-295}$ (37 kDa) fractions from the beads purification were run on a 12 % Tricine SDS-PAGE gel, and (b) the panel represents the immunoblotting of the same samples. The LANA truncations are the only samples which had a GST-tag and they were successfully bound to the glutathione beads. TCL- total cell lysate, W-Wash of unbound proteins, GST- GST agarose beads purification.*

The four proteins, LANA, CBF-β and Elongin BC were co-expressed, of which LANA is the only component of the complex that has a GST tag. The samples from the pull down assays were run on a SDS-gel (Figure 3.5a), and the same samples were used for Immunoblotting assays (Figure 3.5b); His-probe was used to reveal LANA and CBF-β proteins, and specific antibodies to reveal the Elongin BC proteins. Assays were performed in the same manner for all three kLANA truncations.

All four proteins, LANA, CBF-β and Elongin BC were expressed (Figure 3.5b, lanes 1, 3 and 5), but only kLANA protein was purified and neither CBF-β nor Elongin BC was pull down (Figure 3.5b, lanes 2, 4 and 6). This experiment was performed more than one occasion, with reproducible results.

3.1.5. mLANA interacts with Elongin BC

The well characterized SOCS-box motif in the kLANA protein is composed of an Elongin BC box and a Cullin box, which is spatially located at its amino and carboxyl termini. This motif is necessary for LANA interaction with the Cul5–Elongin BC complex, respectively [43].

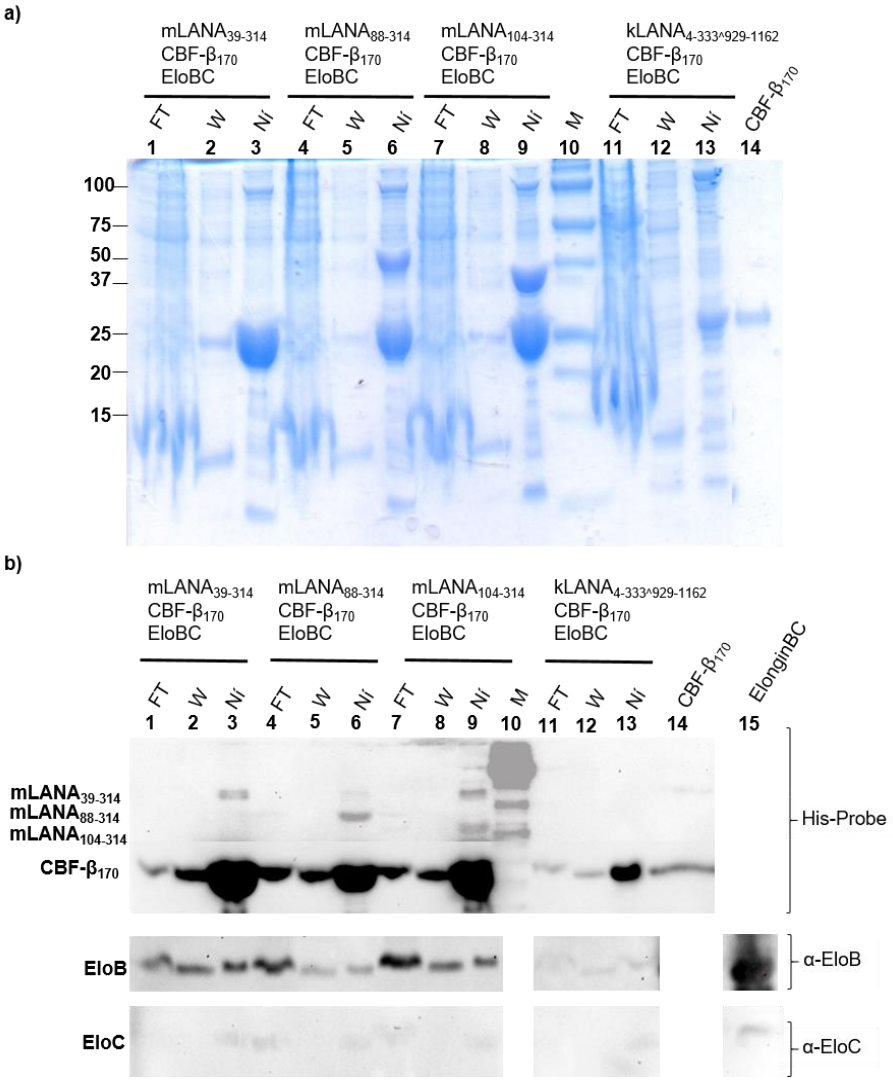

Figure 3.6- Elongin BC proteins interact directly with mLANA truncations. *His-mLANA truncations were co-expressed with CBF-β1-170 (21 kDa), Elongin B (13 kDa) and Elongin C (11 kDa). The cells were lysed using a cell lysis solution, and Ni-NTA-agarose beads were added to the clarified lysates. (a) His-mLANA39-314 (30 kDa), His- mLANA88-314 (25 kDa), His- mLANA104-314 (23 kDa), GST-His-kLANA4-333^929-1162 (87 kDa) fractions from the beads purification were run on a 12 % Tricine SDS-PAGE gel, and (b) western blot representing the immunoblotting of the same samples. The LANA truncations and CBF-β are His-tagged and Elongin BC is not tagged. Elongin B was pull down with mLANA truncations. FT- Flow through from beads purification, W-Wash of unbound of proteins, Ni- Ni-NTA agarose beads purification.*

On the other hand, the latency-associated protein ORF73 (mLANA) has an unconventional SOCS-box-like motif, where the Elongin B and C box and Cullin box are spatially together. It was also shown that mLANA assembles in an ElonginC/Cullin5/SOCS (suppressors of cytokine signalling)-like complex, [44] to inhibit the transcriptional activity of the host nuclear factor-kappa B (NF-kB).

Therefore, it will be interesting to compare the interaction of Elongin BC between kLANA and mLANA proteins. It was also decided to investigate the potential contribution of the C-terminal domain of kLANA in the interaction with CBF-β and Elongin BC, using kLANA$_{4\text{-}333\text{^}929\text{-}1162}$ truncation (From residue 1 to 333 with a deletion between N-terminal and C-terminal domain and from 929 to 1162 residue) that contains a deletion between the N- and C- terminal domains, but contains the Elongin BC and Cullin5 boxes.

Pull downs assays were performed using mLANA$_{39\text{-}314}$, mLANA$_{88\text{-}314}$, mLANA$_{104\text{-}314}$ and kLANA$_{1\text{-}333\text{^}929\text{-}1162}$ truncations, all containing the SOCS-box-like motif. mLANA truncations are cloned into pET 47 b (+), including a His-tag and kLANA$_{4\text{-}333\text{^}929\text{-}1162}$ cloned into pET 49 b (+) , which includes a GST and His-tag.

The pull-down assays were repeated as before. mLANA truncations, and kLANA$_{4\text{-}333\text{^}929\text{-}1162}$ were co-transformed with CBF-β and Elongin BC. Since mLANA truncations are His-tagged, the pull down were performed using Ni-NTA beads, which means that His-CBF-β will also be pull down. Samples were run on a 12% SDS-gel (Figure 3.6a), and proteins were further confirmed by immunoblotting (Figure 3.6b).

The experiment successfully purified the three mLANA constructs, and Elongin B successfully interacted in the pull-down assay (Figure 3.6b, lanes 3, 6 and 9); Although, Elongin C has been shown to directly bind to BC-box, it does not appear in the blot, or in the Elongin C control experiment. It is likely that Elongin C antibody has lost its activity and we will assume that Elongin C was also pulled down since Elongin B was present.

The same did not happen for kLANA$_{4\text{-}333\text{^}929\text{-}1162}$ truncation (Figure 3.6b, lane 13). The protein was not bound to the beads, which indicates that probably the protein was not expressed.

3.1.6. Vif interacts with Elongin BC

Characterization of the interaction of full-length HIV-1 Vif protein with CBF-β and Elongin BC proteins is already well studied [49,51–54]. A similar approach was used in this study to define the interaction between Vif and CBF-β [55]. In order to validate the experimental protocol used in LANA pull downs, the same assay was performed with Vif protein (Figure 3.7). The four proteins were expressed and as expected, Vif was able pull down Elongin BC (Figure 3.7a), validating the protocol being used.

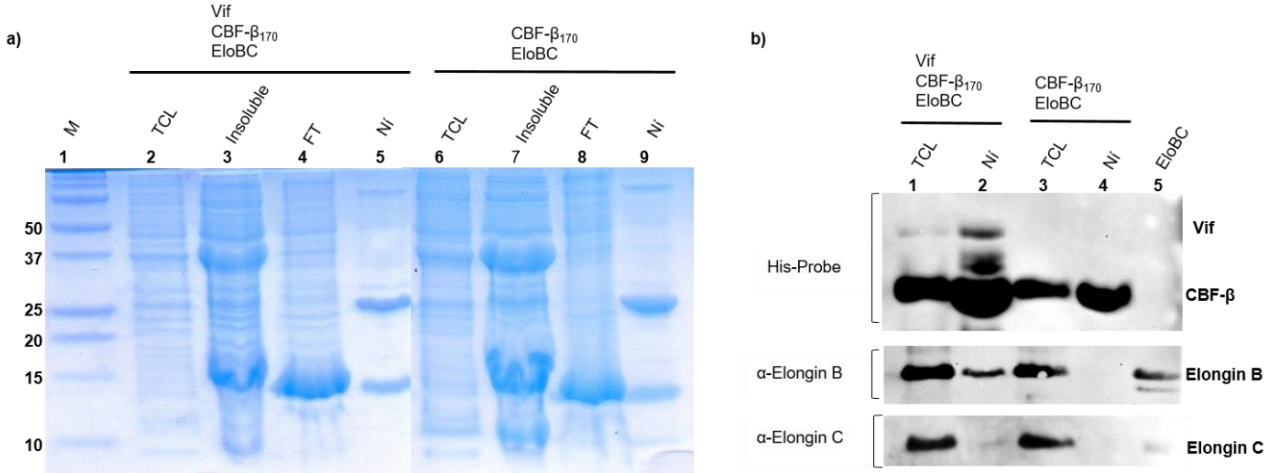

Figure 3.7- Vif interacts directly with Elongin BC. *Control experiment, with Vif protein for validation of LANA pull down assays. His-tagged Vif was co-expressed with CBF-β and Elongin BC. The cells were lysed using a cell lysis solution, and Ni-NTA-agarose beads were added to clarified lysates. (a) the fractions from the beads purification were run on a 12 % Tricine SDS-PAGE gel, and on (b) is represented the immunoblotting of the same samples. Vif (23 kDa) and CBF-β$_{1\text{-}170}$ (21 kDa) are His-tagged while Elongin B (13 kDa) and Elongin C (11 kDa) are not tagged. All proteins were expressed and Elongin BC was pull down along with Vif. TCL- Total cell lysate, FT- Flow through from beads purification, Ni-NTA agarose beads purification.*

3.2. Binding activity of LANA to USP7

3.2.1. Cloning, Expression and Purification

To facilitate structural studies, a new kLANA construct was designed in the C-terminal domain, and designated kLANA$_{971-1150}$ (from 971 to 1150 residue). This construct contains the USP7 binding region and was designed for the characterization of LANA interaction with USP7. For these studies the USP7$_{54-205}$ truncation was used and two more USP7 constructs were designed and cloned. The first construct (USP7$_{54-198}$) was a shorter version of USP7$_{54-205}$ truncation for use in crystallizing the USP7-kLANA complex and the second truncation (USP7$_{chimera}$) was designed to create a fusion version of USP7-kLANA complex, i.e. USP7 gene followed by insertion of kLANA gene (8 aa residues), where both proteins are encoded from a single continuous DNA gene sequence. This construct was also created for crystallization purposes.

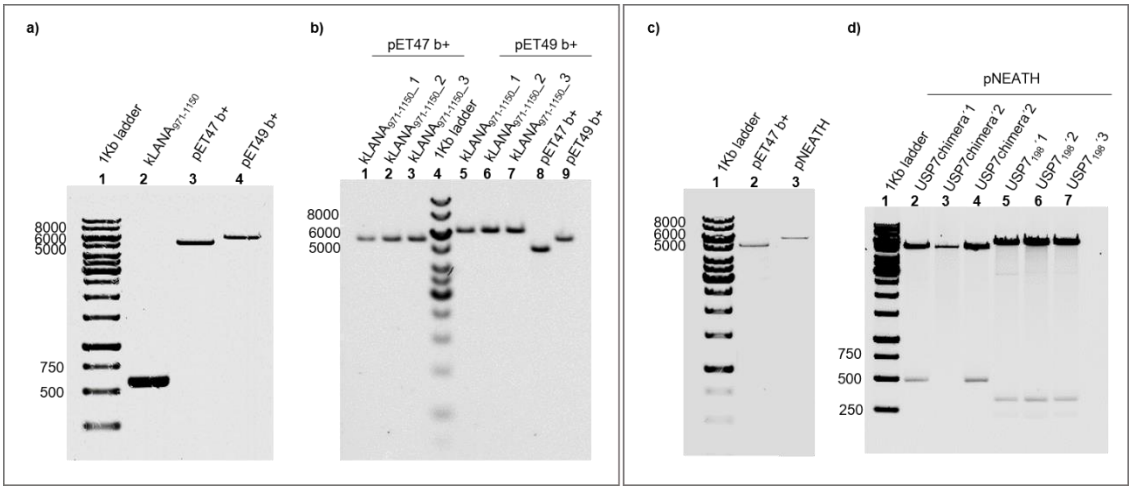

Figure 3.8- Cloning steps of kLANA$_{971-1150}$ and USP7 constructs. *(a) Vectors linearization with BamHI and NotI restriction enzymes and PCR amplification of kLANA$_{971-1150}$; lane2: PCR amplification of kLANA$_{971-1150}$ (540 bp); lane 3: restriction digestion of pET47 b(+) (5203 bp); lane 4: restriction digestion of pET49 b(+) (5926 bp). (b) kLANA$_{971-1150}$ clone validation by single restriction digestion. (c) Vectors linearization for USP7 cloning: lane 2: pET47 b(+) digestion with BamHI and XhoI restriction enzymes; lane 3: pNEATH (5763 bp) digestion with NdeI and BamHI restriction enzymes. (d) USP7 clone validation by restriction digestion.*

kLANA$_{971-1150}$ and USP7 truncations, were cloned in a similar way to kLANA. kLANA$_{971-1150}$ was also cloned into pET 47 b(+) and pET 49 b(+), using BamHI and NotI restriction enzymes (Figure 3.8a). Clone validation was performed by single restriction digestion (Figure 3.8b). USP7$_{Chimera}$ was cloned into pET 47 b(+), between BamHI and XhoI restriction sites and USP7$_{54-198}$ was cloned into pNEATH, between BamHI and NdeI restriction sites (Figure 3.8c). For both cases, inserts were confirmed by digestion, using the respective restriction enzymes (Figure 3.8d). Both kLANA$_{971-1150}$ and USP7 truncations were expressed successfully. However, kLANA$_{971-1150}$ turned out to be a very unstable protein, even in higher salt concentrations and it was only possible to purify the GST version without cleaving the tag.

On the other hand, all USP7 truncation behave very stably and they were very soluble proteins. During purification of these proteins, we saw immediately different behaviour between the three USP7 truncations. While USP7$_{54-205}$ and USP7$_{54-198}$ truncations behave as monomer, USP7 chimera behaves as dimer (Figure 3.9) giving already one indication of the interaction between USP7 and the LANA peptide in which it is interacting in “trans”, that is USP7 fused with C-terminal LANA peptide is interacting with another USP7, hence resulting in dimer

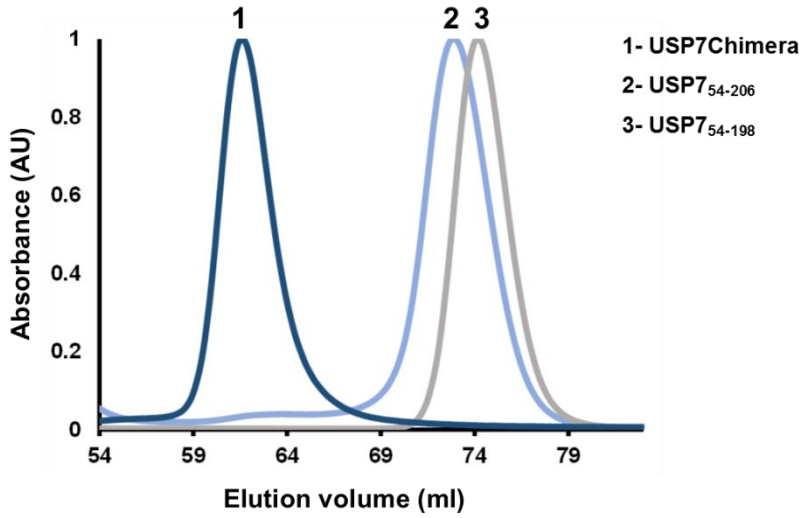

Figure 3.9- Size exclusion analysis of USP7 truncations. *The USP7 truncations were loaded into a* HiLoad 16/600 Superdex 75 pg column and the elution profiles were analysed. *$USP7_{54-205}$ and $USP7_{54-198}$ behave as monomer while $USP7_{Chimera}$ behaves as a dimer.*

3.2.2. Effect of USP7 on DNA binding by LANA

USP7 was previously reported to interact with herpes viral proteins such as, HSV-1 ICP0 [59] and EBV EBNA-1 [94], and to be important for the life cycle of both viruses. USP7 was also shown to have stimulated the DNA binding activity of EBNA-1.

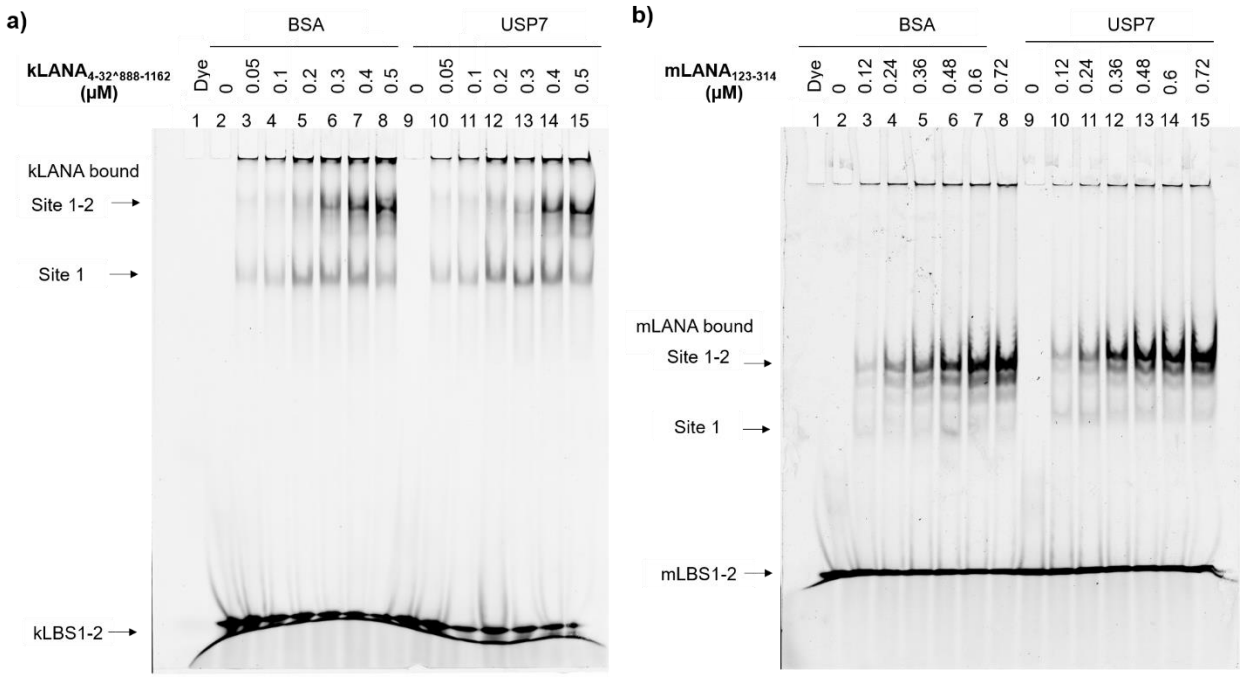

Figure 3.10- Analysis of USP7 effect on LANA's DNA binding activity*. EMSAs showing titrations of LANA with 0.2 μM of DNA recognition site, in the presence of 40 μM of USP7 or in the presence of 40 μM of BSA as a negative control. a) Complexes of $kLANA_{4-32^{\wedge}888-1162}$ with kLBS1-2 and b) Complexes of $mLANA_{124-314}$ with mLBS1-2. The 5% acrylamide gels were run for 3h at 150 V in 0.5x TBE buffer.*

Recently, USP7 was shown to interact with both kLANA and mLANA. Similar to EBNA-1, a motif region in kLANA preceding the DNA binding site has been mapped for USP7 binding site. However, a similar motif was absent in mLANA but has still been shown to interact with USP7 [69]. Since LANA is a functional homologue of EBNA-1, we investigated whether USP7 could also have an effect on DNA binding activity of both kLANA and mLANA [71].

To assess the effect of USP7 on the DNA binding activity of LANA, an electrophoretic mobility shift assay (EMSA) was performed with a kLANA truncation that contains both N-terminal and C-terminal domains ($kLANA_{4\text{-}32\text{^}888\text{-}1162}$) and includes the USP7 binding site. For mLANA, a C-terminal domain ($mLANA_{124\text{-}314}$) was used for the EMSA. Purified kLANA and mLANA was incubated with Fluorescein labelled-DNA (Flc-DNA) containing two LANA binding sites (LBS 1-2 (high and low affinity site) from the TR element of the viral episome DNA); kLBS1-2 or mLBS1-2, respectively, in the presence of excess USP7 or in the presence of the same quantity of BSA (as a negative control).

Repeated experiments consistently showed no stimulation for both kLANA and mLANA's DNA binding activity by USP7, different from the homologous EBNA1 protein (Figure 3.10a and b). The fact that USP7 enhances the DNA binding in the case of EBNA-1, but not in kLANA and mLANA could indicate a different binding mechanism of USP7 or DNA, between LANA and EBNA1proteins.

3.2.3. Thermofluor

For thermofluor experiments, $USP7_{54\text{-}205}$ truncation (USP7-TRAF) was initially screened in 24 different buffers, among which six were chosen and incubated with two kLANA peptides (k10 and k16) and p53 peptide. The experiment was performed in six different buffers, keeping the buffer concentration (100 mM) constant. For data analyses, the melting temperature of each sample in each buffer was calculated, Figure 3.1 and the corresponding melting curves are represented in supplementary information, figure S2.

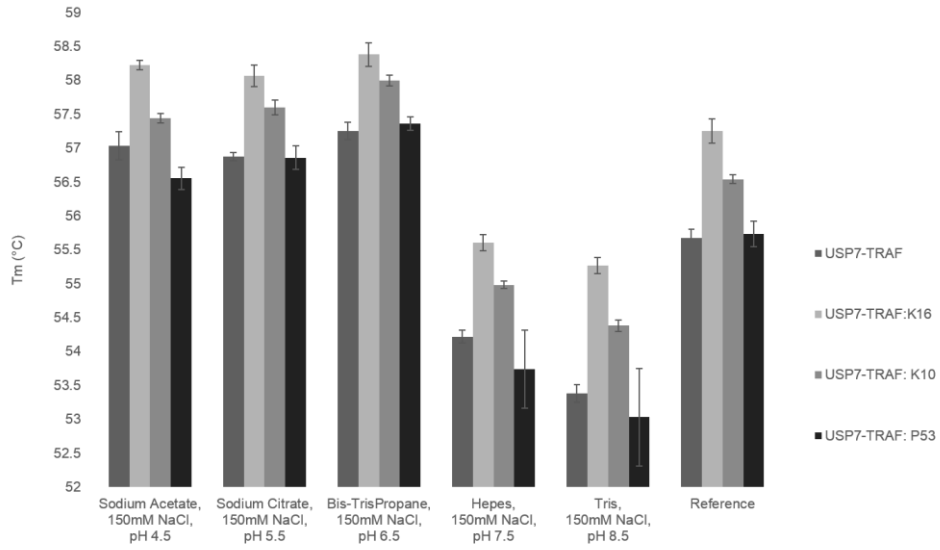

Figure 3.11- Melting temperatures of USP7-TRAF alone and incubated with three peptides. *Melting temperatures of USP7 incubated with three peptides. Midpoint temperatures of the protein-unfolding transition (T_m) were calculated for US7-TRAF and incubated with three peptides, k10, k16 and p53 in the presence of six buffers.*

To identify a ligand or a buffer condition that stabilizes a protein, the T_m value of the protein under each condition of the screen needs to be compared with the reference T_m.
The T_m values measured in buffer were subtracted from the T_m values for the control experiments, and the change in unfolding temperature, ΔT_m, was calculated (Figure 3.12).

A positive ΔT_m can be associated with an increase in structural order and a reduced conformational flexibility, whereas a decrease in stability, negative ΔT_m, indicates either the buffer induces protein structural changes to a disordered conformation, or it can be a sign of misfolding.

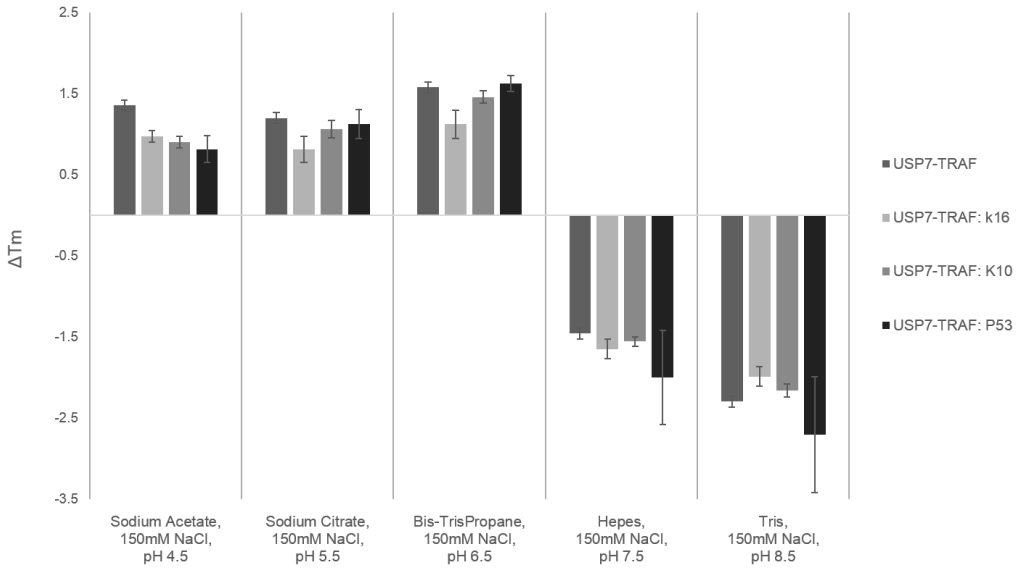

Figure 3.12- Assessment of buffer stabilizing effect through thermofluor analysis. *The T_m value of each condition were compared to reference T_m. Changes in the unfolding temperature (ΔT_m) were calculated for measurements in five buffers for four different protein samples. The histogram represents the median ΔT_m values and error bars the standard deviation. A negative median ΔT_m value signifies that the buffer destabilizes the protein and a positive ΔT_m value indicates that the buffer has a stabilizing effect. Buffers with lower pH (4.5, 5.5 and 6.5) have a more stabilizing effect compare to buffers with a higher pH (7.5 and 8.5).*

Obviously for most proteins, the goal of thermofluor experiment is to identify an increased melting temperature in order to obtain a protein sample with higher stability and hopefully homogeneity. However, the goal of this particular experiment was to look for the opposite effect: identifying a buffer system which will destabilize the protein. In such a buffer system, protein solubility would be decreased and consequently the chance of crystallization may be enhanced. Of note: USP7 solubility is very high, up to 100 mg/ml (reference buffer: 25mM HEPES, pH 7.5), hence we needed a higher quantity of peptide.

For all USP7 samples, sodium acetate pH 4.5, sodium citrate pH 5.5 and Bis-Tris Propane pH 6.5 buffers were stabilizing, whereas HEPES pH 7.5 and Tris-HCl pH 8.5 buffers significantly destabilize the protein. In other words, protein solubility decreased at higher pHs.

If the ligand binds, the protein-ligand complex denatures at a higher temperature than the protein alone. The difference in the T_m value between the presence and absence of ligand reflects the binding ability.

The T_m values measured in each buffer for USP7-ligand were subtracted to the T_m values for USP7-TRAF alone, for each buffer and the change in unfolding temperature, ΔT_m, was calculated (Figure 3.13). A positive ΔT_m can be associated to a more stabilizing ligand and a negative ΔT_m, indicates the opposite.

A positive thermal shift is associated with a stabilizing effect and also with a higher binding affinity, meaning that k16 peptide seems to be the most stabilizing ligand for the protein and also the one with the higher affinity between the three tested ligands. In summary, k16 peptide has a higher affinity than the k10 peptide and in turn, the k10 peptide has a higher affinity than p53. These results are in agreement with complementary ITC analysis (Unpublished data of Dr. Rajesh Ponnusamy, SVL, ITQB).

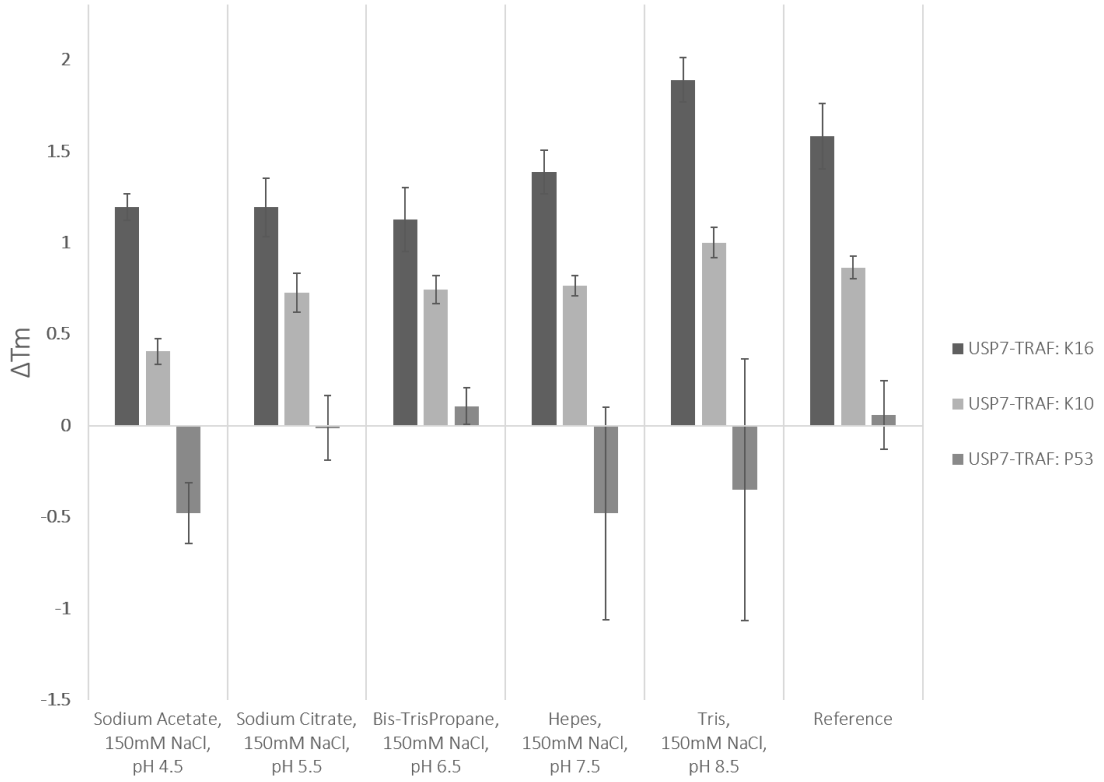

Figure 3.13- Characterization of the effect of peptides on the protein stability. *Each histogram indicates the thermal shift of a peptide from the melting temperature of USP7$_{54-205}$ (without ligand), in each buffer. All experiments were done in triplicate and standard deviation represented with error bars. A positive median ΔT_m value signifies that the peptide stabilizes the protein and a negative ΔT_m value indicates that the peptide has a less stabilizing effect. The k16 peptide is the more stabilizing peptide, while the p53 peptide is the less one.*

3.2.4. Crystallization and X-ray structure

3.2.4.1. USP7 co-crystallization with kLANA

Initial attempts to co-crystallized USP7 with LANA were performed by using the USP7$_{54-205}$ truncation incubated with k10 and k16 peptide. Several screening conditions were tested: with different protein concentrations, different molar ratios between USP7 and the peptide and different temperatures. Despite all these efforts, diffraction quality crystals were not able to be obtained.

At this point a new strategy was designed. By analyzing previous conditions of crystal growth of USP7 N-terminal TRAF-like domain and also its structure and crystal packing from previous studies from Saridakis group, we observed that the region from residue 198 to residue 206 was involved in the crystal packing (Figure 3.14) in all of their crystal structures. Interestingly, seeds of USP7-bound with p53 crystal, had to be used for all their USP7 complex crystals. [95].

Since USP7$_{54-205}$ construct failed to form quality crystals, USP7$_{54-198}$ truncation construct was created to obtain USP7LANA complex with different packing from Saridakis's crystals.

New screens were performed with USP7$_{54-198}$ incubated with both kLANA peptides (k10 and k16). Micro-crystals of USP7$_{54-198}$-k16 were grown in 28 % of PEG 3350, 100mM Tris-HCl pH 8.5 and 0.2M Lithium sulfate. Further optimization by streak seeding was performed but was unsuccessful. This USP7$_{54-198}$ construct also failed in getting quality crystals, so another completely different strategy was designed to obtain the USP7-LANA complex crystals.

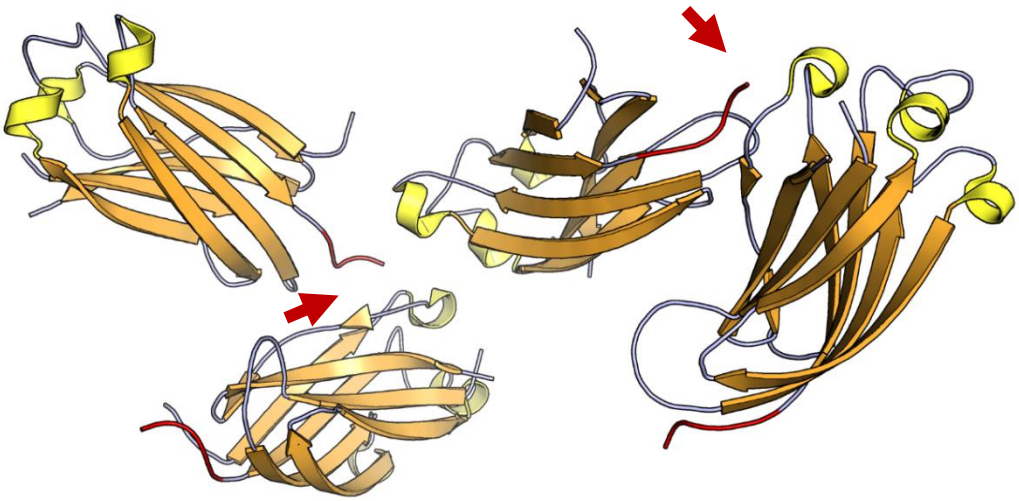

***Figure 3.14*- Crystal packing of the USP7 N-teminal TRAF-like Domain.** The residues 199 to 205 are coloured in red, showing the deleted region on *USP7*$_{54-198}$ truncation. Adapted from [95] , PDB code: 1YY6.

3.2.4.2. Crystallization of the USP7-k10 chimeric protein

There are several published X-ray structures of USP7 N-terminal TRAF-like domain bound to ligands, but most of them are from Saridakis's group, who have used the same seed to obtain crystals of different USP7 complexes [64]. However, crystals of USP7 with p53 and MDM2 ligands were obtained using chimeric proteins; that is, p53 and MDM2 peptide sequences were fused to the C-terminal region of USP7. Based on this strategy, the ^{975}PQPGPSRE^{982} aa residues from LANA sequence were fused to the C-terminal of USP7 to create USP7 chimeric protein.

The chimeric protein was purified in 20mM Tris-HCl pH 8.0, 100mM NaCl buffer. This buffer choice was based on the previous thermofluor assays (Figure 3.12), where it was shown as a buffer with lower T_m to reduce protein solubility. In order to maintain protein stability, this buffer condition was only used in the last step of purification.

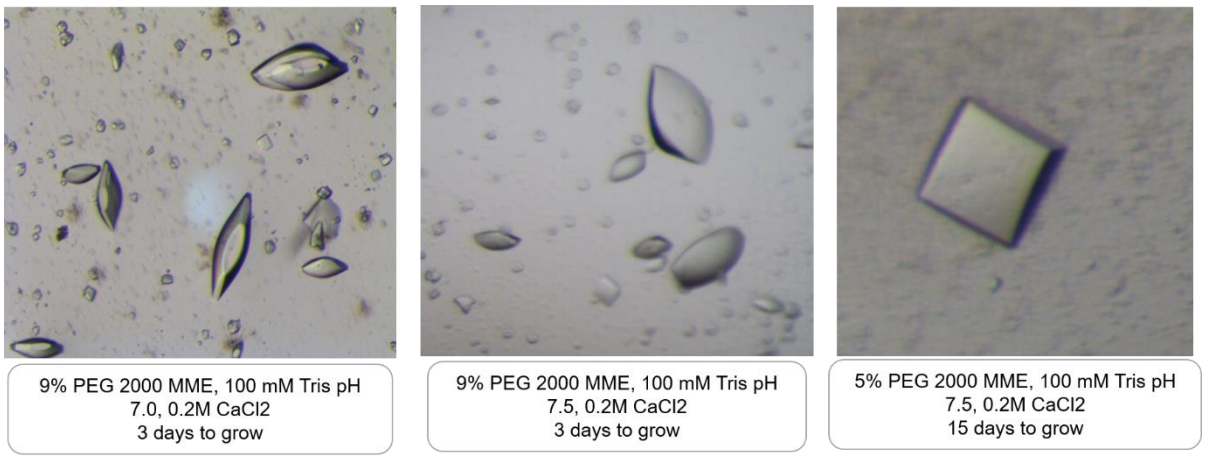

***Figure 3.15- USP7-k10 chimeric protein crystals**. Crystals were grown in different conditions and was observed different shapes.*

In agreement with thermofluor data (Figure 3.12), the crystals were grown in higher pH buffers, which are destabilizing buffers, i.e. buffers were the protein solubility decreased (HEPES pH7.5 and Tris pH 8.5). The crystals were grown in PEG 2000 MME, 100 mM Tris-HCl and 0.2M $CaCl_2$. The crystals had different shapes and growth times depending on both the PEG concentration and pH used (Figure 3.16).

Small crystals were also obtained in PEG 1500, 100 mM Tris-HCl pH 9.0 and 0.2M $MgCl_2$. To improve crystal growth in this condition, serial streak-seeding was performed in clear drops, using small crystals as seed in increasing precipitant concentrations (Figure 3.16).

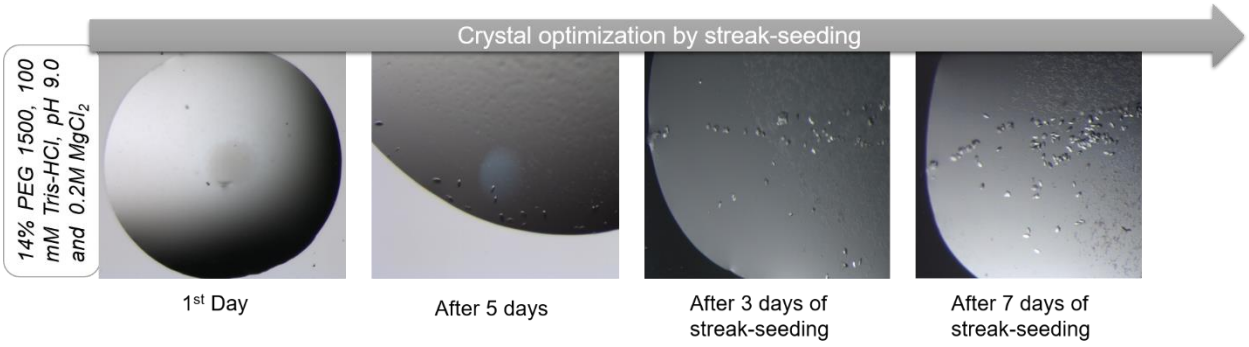

***Figure 3.16- Crystal optimization by streak-seeding**. Small crystals of USP7-k10 chimeric protein grown in 14% PEG 1500, 100mM Tris-HCl pH 9.0 and 0.2M of magnesium chloride were optimized by streak seeding. In seven days diffraction-quality crystals were obtained.*

In crystallization, another important parameter is to determine the best cryosolution and the best technique to cryocool down crystals. There are two relevant considerations to take into account: vitrification of the drop and the interaction between the cryo-solute and the protein inside the crystal. Crystals immersed in various cryoprotectants often develop cracks, bends, etc. indicative of stress by cryo solution to the crystal packing. Even small variations can be detrimental to crystal health and diffraction quality. This unwanted stress can be related to the different density/osmolarity between the crystal and the solution in which it is placed. Sucrose, glycerol and ethylene glycol are normally the most common cryo-agents that crystals like (i.e. survives in)[96].

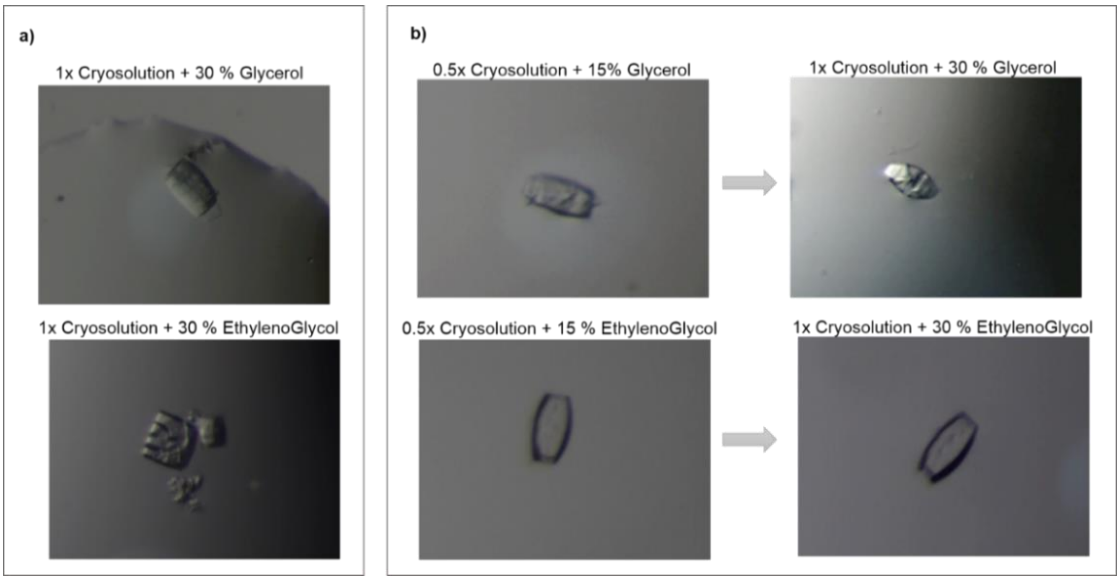

***Figure 3.17- Cryo cool down test of $USP_{7chimera}$ crystals.** Two different cryo-agents were tested in two different approaches. Cryo soluition glycerol in the top panel and ethylene glycol in the bottom panel were tested in (a) one-step approach and (b) two-step approach.*

These three cryo-agents were tested, and the best cryo-agent based on the external appearance was ethylene glycol. In order to maintain crystal quality we had to cryo-soak in a two-step approach. When transferring the crystals to the cryo-solution they instantly cracked (Figure 3.17a); however, if we first soak them in a solution with half of the final concentration, where they stayed stable, and then transfer them to the final cryo-solution the crystals survived (Figure 3.17b).

3.2.4.3. X-ray structure of USP7-kLANA chimera

To gain insight into the molecular basis of LANA interaction with USP7, the $USP7_{chimera}$ crystallized in 9% PEG 2000 MME, 100 mM Tris-HCl pH 8.0, 0.2M $CaCl_2$ turn out to be the highest resolution data and the structure was determined using molecular replacement (Figure 3.19a).

The model was refined to 1.9 Å of resolution (see Table 3.1 for structure solution and refinement statistics).The kLANA peptide was located in the F_0-F_c difference electron density map. From residues 975 to 980 the map showed excellent density, which allowed clear building of the kLANA peptide (Figure 3.18). For residues 981 and 982 the density map was less well defined, but adequate to allow theses residues to be built.

Table 3.1- X-ray data collection and refinement statistics.

X-ray diffraction data		Values in parentheses are for the highest resolution shell. [1] $R_{merge} = \sum_{hkl}\sum_i \left\| I(hkl)_i - \langle I(hkl) \rangle \right\| / \sum_{hkl}\sum_i I(hkl)_i$, [2] $R_{pim} = \sum_{hkl} \sqrt{\frac{1}{n-1}} \sum_i \left\| I(hkl)_i - \langle I(hkl) \rangle \right\| / \sum_{hkl}\sum_i I(hkl)_i$, where I(hkl) is the intensity of reflection hkl and ⟨I(hkl)⟩ is the average intensity over all equivalent reflections. [3] $R_{cryst} = \sum hkl \left\| F_o(hkl) - F_c(hkl) \right\| / \sum_{hkl} F_o(hkl)$. R_{free} was calculated for a test set of reflections (5%) omitted from the refinement.
Wavelength (Å)	0.98	
Space group	P 2_1 2_1 2_1	
Cell dimensions: a, b, c (Å)	49.3, 61.1, 61.8	
α, β, γ (°)	90.0, 90.0, 90.0	
Resolution (Å)	61.83-1-79 (1.83-1.79)	
R_{merge} (%) [1]	5.9 (8.3)	
R_{pim} (%) [2]	3.3 (29.7)	
I/σI	11.3 (2.2)	
Completeness (%)	99.6 (95.6)	
Multiplicity	4.0 (3.3)	
Total measured reflections	72556 (3586)	
Unique reflections	18173 (1072)	
Wilson B-factor (Å^2)	25.15	
Refinement		
Resolution (Å)	43.50-1.79	
R_{cryst}/R_{free} [3]	20.9 /25.0	
Protein atoms	1308	
B factor (Å^2)	38.47	
Rmsd bond length (Å)	0.021	
Rmsd cond angles (°)	1.970	
Ramachandran analysis: favoured/allowed (%)	98.6/1.4	

Like all TRAF domains, the overall structure of USP7-TRAF forms an eight-stranded antiparallel β-sandwich, with strands β1, β5, β6 and β8 in one sheet and strands β2, β3, β4 and β7 in the other (Figure 3.19a). The interactions between LANA and USP7 mainly occur in the β7 strand (Figure 3.19b), similar to other peptides interactions of USP7 [64,75,95].

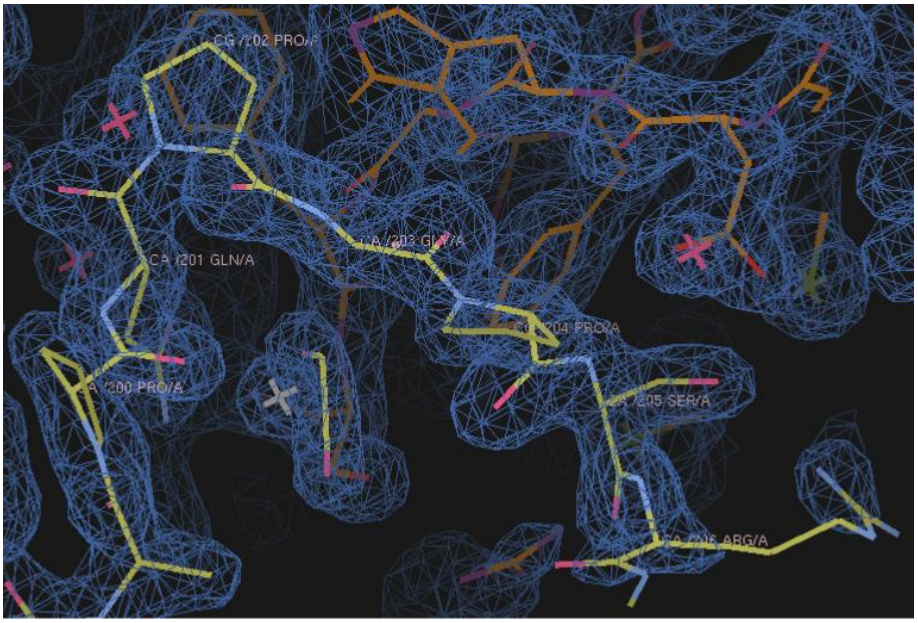

***Figure 3.18*- The $2F_0$-F_c density map of kLANA peptide region.** The map is in blue contoured at 1.2 σ.

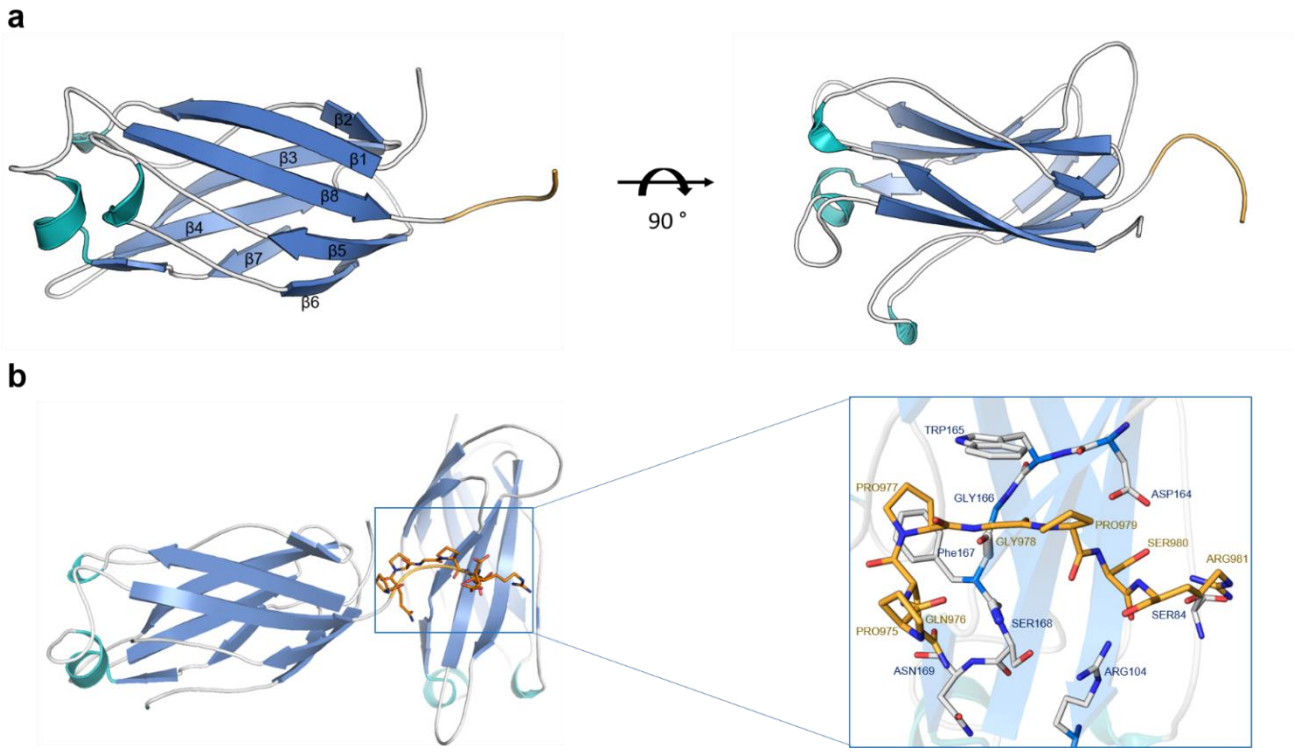

Figure 3.19- Structure of USP7-kLANA chimera protein. *(a) Cartoon representation of USP7-kLANA chimera protein, with the respectively structure with 90 degrees rotation on the right. The kLANA peptide is coloured in orange and USP7 strands are labelled. (b) kLANA peptide interaction region with USP7. On the right, detailed representation of interaction between USP7 (silver) and kLANA (orange) and the residues involved in the interaction are represented as a stick.*

USP7-TRAF preferentially shown to interact with P/AxxS motifs [95]. In kLANA this motif is formed by ^{977}PGPS^{980}, which we observed in the solved structure to interact with residues from the β7 strand (Figure 3.19b).

kLANA interaction with USP7 is chiefly hydrophobic, similar to other USP7-TRAF domain partners. The binding region of USP7 contains a hydrophobic core centred on residues Trp165-Gly166-Phe167. The residue Trp and Phe have a sandwich conformation and make π interaction with the hydrophobic region of the kLANA peptide Pro977-Gly978- Pro979 (Figure 3.19). The main chain atoms make two hydrogen bonds at the centre by residue 166 (USP7) and residue 978 (kLANA).

Another hydrogen bond, presumably determines the specificity, is made between the hydroxyl group of Ser980 (kLANA) and carboxyl group of Asp164 from USP7. Residue 164 from USP7 also makes another hydrogen bond with the main chain amide group of residue 980 from kLANA (Figure 3.20).

Final interaction between the kLANA peptide and USP7 are between main chain carbonyl atom of residue 980 and side chain atom NH1 of Arg104, respectively. In addition, a unique interaction is made by the residue Gln976 at the start of kLANA peptide; two hydrogen bonds are made by the side chain atoms carboxyamide and carboxyl groups of Gln976 with carboxyl group and main chain amide group of Asn169 (USP7), respectively. Furthermore hydrophobic residues, Phe118 and Ile154 from USP7 are in the vicinity of kLANA peptide making the USP7-kLANA interaction more hydrophobic (Figure 3.20).

.

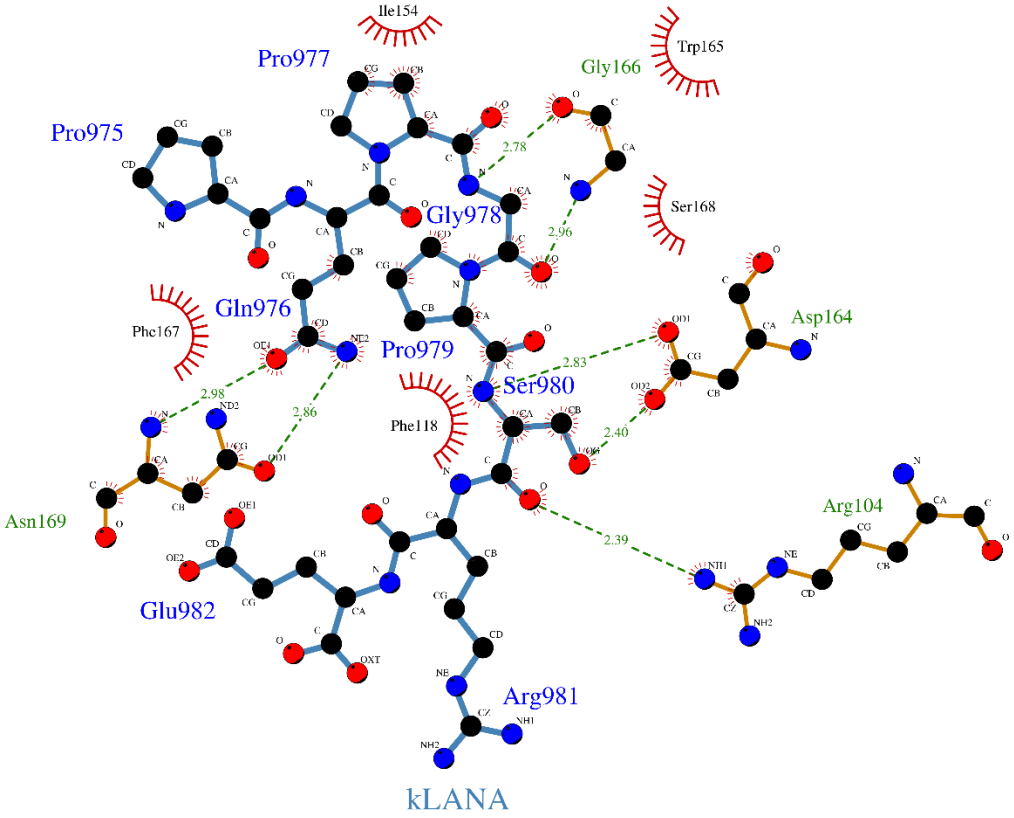

Figure 3.20- Schematic diagram illustrating the USP7-kLANA interactions, generated by LIGPLOT*. The kLANA residues are displayed in blue and USP7 interacting residues are in orange. Hydrogen bonds are indicated by green dashed lines, while hydrophobic contacts are represented by a red arc with spokes radiating towards kLANA atoms they contact. The contacted atoms are shown with spoke radiating back*.

4

DISCUSSION AND CONCLUSION

The ubiquitin pathway is one of the major systems of protein regulation and has been implicated in the immune response, development, and programmed cell death [34,35,97]. Viruses have a remarkable capacity to evolve and adapt in such a way that they can take advantage of the ubiquitin system for its own purposes, e.g. to modulate host proteins for viruses to enter and replicate [33]. KSHV LANA protein is one of more examples which has the capacity to recruit a ubiquitin ligase complex for degradation of VHL and p53 tumour suppressors [43].

LANA was already shown to interact with Elongin BC complex, through binding with Elongin C via its BC-box motif located within its N-terminal domain. Furthermore, the LANA-Elongin BC complex assembles with a Cullin5/Rbx1 module to reconstitute a ubiquitin ligase complex (Figure 4.1) [43].

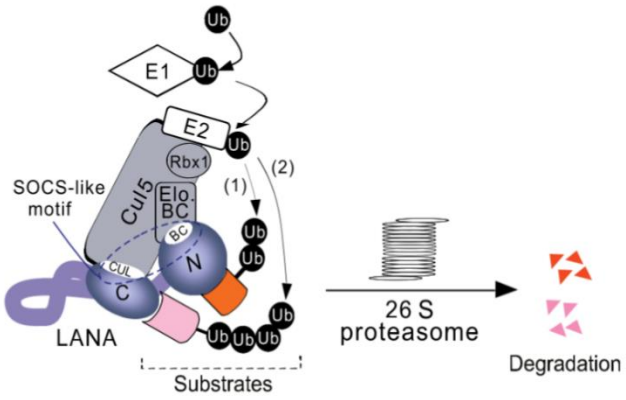

Figure 4.1- Model for KSHV LANA assembles EC$_5$S Ubiquitin complex to target restriction factors for degradation. *Adapted from [43].*

The ubiquitin E3 ligase is the key enzyme in ubiquitination pathway, because it recognizes a specific protein substrate and catalyses the transfer of activated Ub to it. In the case of kLANA, the E3 ligase from the Cullin-Ring family, Cul5 was preferentially recruited [43]. The mechanism by which Cul5 functions in this process remains unclear. To uncover other interaction partners will be important to gain insights into LANA recruitment of the ubiquitin assembly process.

One potential partner is CBF-β protein, that is crucial for HIV-1 Vif recruitment to the ubiquitin-ligase complex [42,51,52], being functional similar to kLANA on targeted protein degradation. Vif recruitment process is well studied and the tertiary structure of Vif ubiquitin complex were recently resolved [49]. Recent studies also indicated that, CBF-β binds with the N-terminal domain of Vif protein, which induces structural changes, to enable Cul5 to bind Vif and form a functional ubiquitin ligase complex [42].

One of the goals of this work was to understand the mechanism behind kLANA recruitment of the EC$_5$S ubiquitin complex. The starting point was to establish whether or not CBF-β interacts with kLANA. The pull down assays performed suggested no binding between CBF-β and kLANA, but neither with Elongin BC although it was previously identified to interact with the N-terminal domain of kLANA. Due to this, a control experiment was performed with Vif protein, which was able to pull down both CBF-β and Elongin C, which has validated the protocol being used for the study. Taking this into consideration, kLANA N-terminal truncation (207 to 321) may not be sufficiently long enough for the establishment of a direct interaction.

One of the issues was the fact that the kLANA N-terminal domain is predicted to be intrinsically unfolded [74]. Due to the presence of an unstructured region in the N-terminus of LANA could impair formation of a complex, as well as contributing to poor expression, solubility and stability. Unlike kLANA, the homologous MHV-68 LANA (mLANA) was able to pull down Elongin BC, consistent with the previous study [44]. It is worth mentioning the differences between kLANA and mLANA; while kLANA SOCS-box is spatially separated and present in the N-terminal domain (Elongin BC-box) and C-terminal domain (Cul5-box), mLANA has an unconventional SOCS-box which was present together at the C-terminal, which may make Elongin BC binding easier. Additional interaction partners from the ubiquitin complexes, like CBF-β and Cul5, were not able to be studied for mLANA truncations in the present studies. Further analyses are required to dissect the LANA-ubiquitin complexes. The results of this study suggested that kLANA does not interact directly with CBF-β and neither with Elongin BC. In contrast, mLANA is able to directly interact with Elongin BC.

Apart from recruiting ubiquitin ligase complex, LANA also regulate the reversal of ubiquitination, through interaction with a deubiquitinating enzyme, USP7. LANA and p53 (tumour-suppressor protein) competes for the same pocket to bind in the USP7-TRAF domain. EBNA1 was also shown to bind to this USP7-TRAF region and showed to interfere with the binding and stabilization of p53 by USP7, resulting in p53-mediated apoptosis in response to DNA damage [66]. Other interesting feature was, USP7 enhancement on EBNA1 DNA binding activity.

EBNA1 is the closest homolog protein from the gammaherpesvirus family of LANA and was previously shown that USP7 modulates the replication of KSHV latent episomal DNA [69]. For this reason, the possibility of USP7 enhance LANA DNA binding was investigated. EMSAs results for both kLANA and mLANA, indicated that USP7 does not affect LANA's DNA binding activity.

A recent publication reported the X-ray crystal structure of kLANA bound to DNA, which shows an asymmetric binding, different from the EBNA1 DNA complex [24,98]. Since DNA binding mechanisms are different compared to kLANA and EBNA1, might explain why USP7 enhances DNA binding in EBNA1 but not in kLANA.

Previously, residues 971-986 of kLANA were shown to be interacting with USP7 [69]. However by further analysing the sequence, a conserved USP7 binding motif ^{978}PGPS^{981} was present within amino acid residues 971-986 in kLANA. Another such motif ^{258}PPTS^{262} is found in the N-terminal domain of kLANA, which is also likely to interact with USP7 [95]. Similarly, homologous mLANA also contains USP7 binding motif ^{272}PSTS^{275}, but at the end of C-terminus, spatially differently located compared to kLANA and EBNA1. The crystal structure of USP7-kLANA complex revealed a novel finding. Until now it has been observed that USP7-TRAF preferentially interacts with P/AxxS motifs [64,65,75]. Here our findings show that USP7 recognizes a ^{976}QPGPS^{980} motif in kLANA. Interesting, mLANA also possesses a glutamine within the USP7 recognition motif, ^{271}QPSTS^{275}.

Therefore, additional to the standard motif, LANA proteins include a glutamine (Gln977) to establish hydrogen bonds with Asn169 of USP7, making Gln the new member of the recognition site in USP7 motif. To understand why kLANA utilizes this additional residue to establish extra hydrogen bonds with USP7, the previously mapped USP7-p53 interactions [64] was compared to USP7-kLANA interactions.

The p53 peptide ($^{358`}$EPGGSR$^{´363}$) with an identical peptide-binding motif (PGGS), has an identical conformation to kLANA PGPS motif (Figure 4.2).

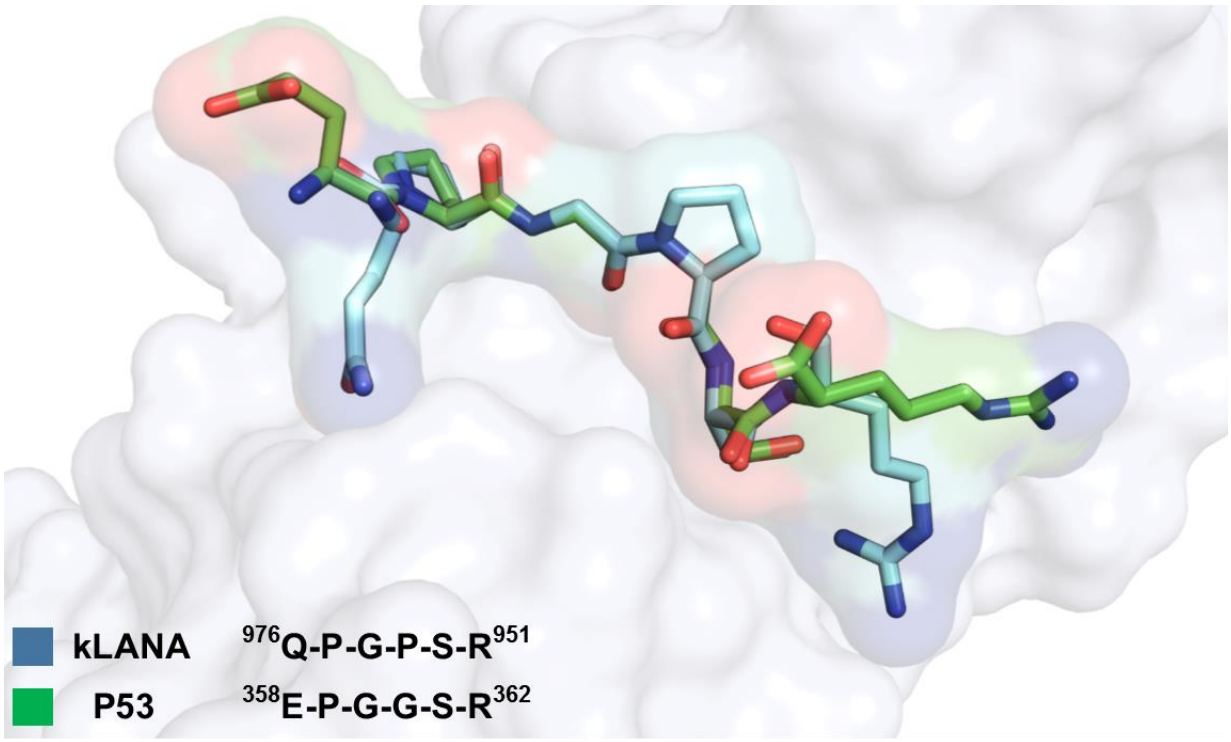

***Figure 4.2- Comparison between kLANA and p53 peptides**. kLANA is displayed in cyan and p53 in green. PDB code of P53 structure: 2FOO.*

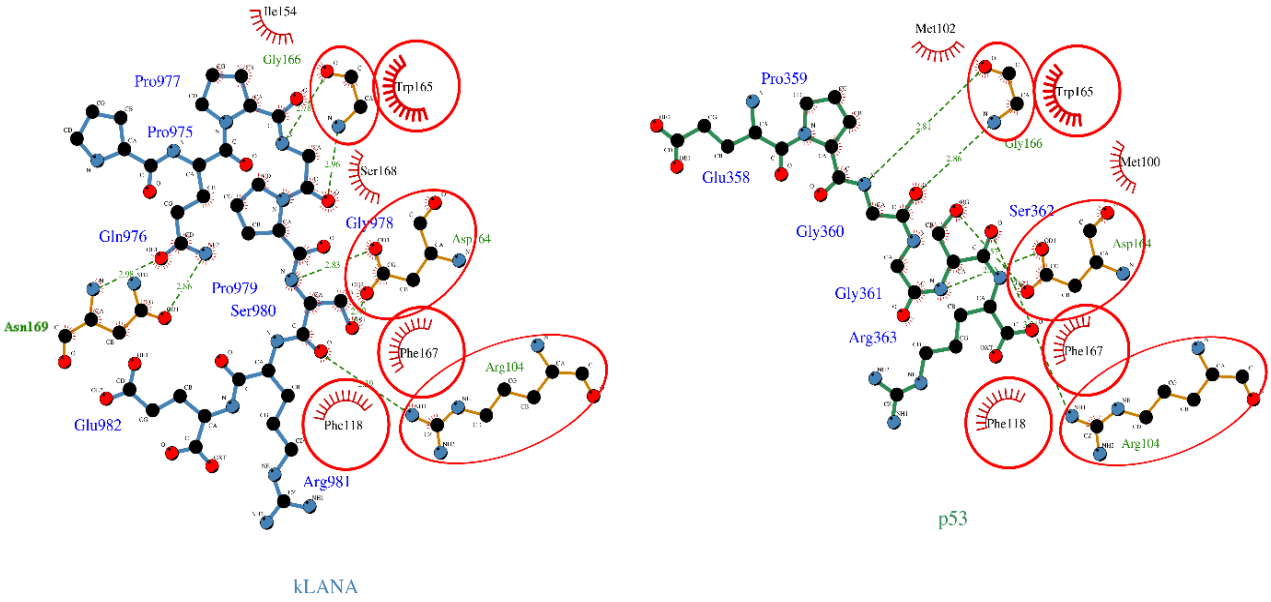

***Figure 4.3- Schematic diagram illustrating the USP7-kLANA interactions on the left and USP7-p53 interactions on the right.** kLANA residues are displayed in blue, p53 residues in green and USP7 interacting residues in orange. Hydrogen bonds are indicated by green broken lines, while hydrophobic contacts are represented by a red arc with spokes radiating towards kLANA or p53 atoms to which they contact. The contacted atoms are showed with spoke radiating back.*

For a more detailed analysis, LIGPLOTplus was again used to generate a schematic diagram of USP7-kLANA and USP7-p53 interactions, highlighting with a red circle similar interactions between the structures (Figure 4.3). Overall, the residues and the type of contacts made by USP7 with the peptides are similar. The main differences between kLANA and p53 on USP7 binding are the additional hydrogen bonds established between the carboxyamide group of kLANA Gln976 and the side chain carbonyl and main chain amino group of USP7 Asn169. The common features between kLANA and p53 peptides interaction with USP7 is the hydrophobic contacts centred on ^{165}WGF^{167} and additional two hydrogen bonds made by residues Asp164 and Arg104.

In summary, the USP7-TRAF domain has a structure that resembles the TRAF/MATH domain forming an eight anti-parallel β-sheet. Extensive structural and biochemical studies have shown that all of the USP7-peptide complex structures are interacting through the P/AxxS standard motif and mainly through the β7 strand of USP7. Most of the solved structures of USP7 complex showed that the peptides do not form any direct contacts with USP7-NTD outside of P/AxxS region. For the first time, glutamine residue from kLANA is shown to be part of the USP7 binding motif. Interestingly, also mLANA has glutamine residue as a part of the USP7 binding motif. A similar motif was found in two other herpesviruses, *Saimiriine herpesvirus 2 and Bovine herpesvirus 4,* indicating a common feature within the herpesvirus family. LANA may use the additional interaction residues to compete with USP7's natural substrates like p53. In concurrence, ITC analysis revealed a kLANA peptide has 3.6-fold higher affinity than p53 for USP7 (Unpublished data of Dr. Rajesh Ponnusamy, SVL, ITQB).

Since USP7 plays a vital role in p53 regulation, disturbances by kLANA interaction with USP7 may lead to a positive regulation of USP7, which in turns stabilizes the tumour-suppressor protein p53. The availability of high resolution structures will help to search for potential drug-like molecule screening and for new structure-based drug design.

REFERENCES

1. Whitley RJ. Herpesviruses. 1996. Available at: http://www.ncbi.nlm.nih.gov/books/NBK8157/. Accessed August 25, 2015.
2. Dreyfus DH. Herpesviruses and the microbiome. *J Allergy Clin Immunol*. 2013;132(6):1278-1286. doi:10.1016/j.jaci.2013.02.039.
3. Sternbach G, Varon J. Moritz Kaposi: Idiopathic pigmented sarcoma of the skin. *J Emerg Med*. 1995;13(5):671-674. doi:10.1016/0736-4679(95)00077-N.
4. Kaposi M, European E. Aposi ' s. 2000:1027-1038.
5. Boshoff C, Weiss RA. Epidemiology and pathogenesis of Kaposi's sarcoma-associated herpesvirus. *Philos Trans R Soc Lond B Biol Sci*. 2001;356(1408):517-34. doi:10.1098/rstb.2000.0778.
6. Chang Y, Cesarman E, Pessin MS, et al. Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science*. 1994;266(5192):1865-1869. doi:10.1126/science.7997879.
7. Schulz TF. Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8): epidemiology and pathogenesis. *J Antimicrob Chemother*. 2000;45(T3):15-27. doi:10.1093/jac/45.suppl_4.15.
8. New THE, Journal E, Medicine OF. Kaposi's sarcoma–associated herpesvirus-like dna sequences in aids-related body-cavity–based lymphomas e. 1995:1186-1191.
9. Soulier J, Grollet L, Oksenhendler E, et al. Kaposi's sarcoma-associated herpesvirus-like DNA sequences in multicentric Castleman's disease. *Blood*. 1995;86(4):1276-80. Available at: http://www.bloodjournal.org/content/86/4/1276.abstract. Accessed August 3, 2015.
10. Nash AA, Dutia BM, Stewart JP, Davison AJ. Natural history of murine gamma-herpesvirus infection. *Philos Trans R Soc Lond B Biol Sci*. 2001;356(1408):569-79. doi:10.1098/rstb.2000.0779.
11. Pedro Simas J, Efstathiou S. Murine gammaherpesvirus 68: a model for the study of gammaherpesvirus pathogenesis. *Trends Microbiol*. 1998;6(7):276-282. doi:10.1016/S0966-842X(98)01306-7.
12. Speck SH, Ganem D. Viral latency and its regulation: Lessons from the ??-Herpesviruses. *Cell Host Microbe*. 2010;8(1):100-115. doi:10.1016/j.chom.2010.06.014.
13. Gradoville L, Gerlach J, Grogan E, et al. Kaposi ' s Sarcoma-Associated Herpesvirus Open Reading Frame 50 / Rta Protein Activates the Entire Viral Lytic Cycle in the HH-B2 Primary Effusion Lymphoma Cell Line †. 2000;74(13):6207-6212.
14. Lan K, Kuppers DA, Verma SC, Sharma N, Murakami M, Robertson ES. Induction of Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen by the lytic transactivator RTA: a novel mechanism for establishment of latency. *J Virol*. 2005;79(12):7453-65. doi:10.1128/JVI.79.12.7453-7465.2005.
15. Yang Z, Yan Z, Wood C. Kaposi's sarcoma-associated herpesvirus transactivator RTA promotes degradation of the repressors to regulate viral lytic replication. *J Virol*. 2008;82(7):3590-603. doi:10.1128/JVI.02229-07.
16. Chang P-C, Kung H-J. SUMO and KSHV Replication. *Cancers (Basel)*. 2014;6(4):1905-24. doi:10.3390/cancers6041905.
17. Gould F, Harrison SM, Hewitt EW, Whitehouse A. Kaposi's sarcoma-associated herpesvirus RTA promotes degradation of the Hey1 repressor protein through the ubiquitin proteasome pathway. *J Virol*. 2009;83(13):6727-38. doi:10.1128/JVI.00351-09.
18. Lan K, Kuppers DA, Verma SC, Robertson ES. Kaposi's sarcoma-associated herpesvirus-encoded latency-associated nuclear antigen inhibits lytic replication by targeting Rta: a potential mechanism

for virus-mediated control of latency. *J Virol*. 2004;78(12):6585-94. doi:10.1128/JVI.78.12.6585-6594.2004.

19. Kelley-Clarke B, Ballestas ME, Srinivasan V, et al. Determination of Kaposi's sarcoma-associated herpesvirus C-terminal latency-associated nuclear antigen residues mediating chromosome association and DNA binding. *J Virol*. 2007;81(8):4348-56. doi:10.1128/JVI.01289-06.

20. Rainbow L, Platt GM, Simpson GR, et al. The 222- to 234-kilodalton latent nuclear protein (LNA) of Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) is encoded by orf73 and is a component of the latency-associated nuclear antigen. *J Virol*. 1997;71(8):5915-21. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=191847&tool=pmcentrez&rendertype=abstract. Accessed August 24, 2015.

21. Kedes DH, Lagunoff M, Renne R, Ganem D. Identification of the gene encoding the major latency-associated nuclear antigen of the Kaposi's sarcoma-associated herpesvirus. *J Clin Invest*. 1997;100(10):2606-10. doi:10.1172/JCI119804.

22. Verma SC, Lan K, Robertson E. Structure and function of latency-associated nuclear antigen. *Curr Top Microbiol Immunol*. 2007;312:101-36. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3142369&tool=pmcentrez&rendertype=abstract. Accessed August 3, 2015.

23. Campbell M, Izumiya Y. Post-Translational Modifications of Kaposi's Sarcoma-Associated Herpesvirus Regulatory Proteins - SUMO and KSHV. *Front Microbiol*. 2012;3:31. doi:10.3389/fmicb.2012.00031.

24. Hellert J, Weidner-Glunde M, Krausze J, et al. The 3D structure of Kaposi sarcoma herpesvirus LANA C-terminal domain bound to DNA. *Proc Natl Acad Sci U S A*. 2015;112(21):6694-9. doi:10.1073/pnas.1421804112.

25. Garber AC, Hu J, Renne R. Latency-associated nuclear antigen (LANA) cooperatively binds to two sites within the terminal repeat, and both sites contribute to the ability of LANA to suppress transcription and to facilitate DNA replication. *J Biol Chem*. 2002;277:27401-27411. doi:10.1074/jbc.M203489200.

26. Correia B, Cerqueira S a, Beauchemin C, et al. Crystal structure of the gamma-2 herpesvirus LANA DNA binding domain identifies charged surface residues which impact viral latency. *PLoS Pathog*. 2013;9(10):e1003673. doi:10.1371/journal.ppat.1003673.

27. Ballestas ME, Kaye KM. Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen 1 mediates episome persistence through cis-acting terminal repeat (TR) sequence and specifically binds TR DNA. *J Virol*. 2001;75(7):3250-8. doi:10.1128/JVI.75.7.3250-3258.2001.

28. Domsic JF, Chen H-S, Lu F, Marmorstein R, Lieberman PM. Molecular basis for oligomeric-DNA binding and episome maintenance by KSHV LANA. *PLoS Pathog*. 2013;9(10):e1003672. doi:10.1371/journal.ppat.1003672.

29. Bochkarev A, Barwell JA, Pfuetzner RA, Bochkareva E, Frappier L, Edwards AM. Crystal Structure of the DNA-Binding Domain of the Epstein–Barr Virus Origin-Binding Protein, EBNA1, Bound to DNA. *Cell*. 1996;84(5):791-800. doi:10.1016/S0092-8674(00)81056-9.

30. Skiadopoulos MH, McBride AA. The bovine papillomavirus type 1 E2 transactivator and repressor proteins use different nuclear localization signals. *J Virol*. 1996;70(2):1117-24. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=189919&tool=pmcentrez&rendertype=abstract. Accessed August 30, 2015.

31. Ponnusamy R, Petoukhov M V, Correia B, et al. KSHV but not MHV-68 LANA induces a strong bend upon binding to terminal repeat viral DNA. *Nucleic Acids Res*. 2015. doi:10.1093/nar/gkv987.

32. Jana NR. Protein homeostasis and aging: role of ubiquitin protein ligases. *Neurochem Int*. 2012;60(5):443-7. doi:10.1016/j.neuint.2012.02.009.

33. Calistri A, Munegato D, Carli I, Parolin C, Palù G. The ubiquitin-conjugating system: multiple roles in viral replication and infection. *Cells*. 2014;3(2):386-417. doi:10.3390/cells3020386.

34. Lecker SH, Goldberg AL, Mitch WE. Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. *J Am Soc Nephrol*. 2006;17(11):1807-1819. doi:10.1681/ASN.2006010083.

35. Komander D. The emerging complexity of protein ubiquitination. *Biochem Soc Trans*. 2009;37:937-953. doi:10.1042/BST0370937.

36. Pickart CM, Fushman D. Polyubiquitin chains: polymeric protein signals. *Curr Opin Chem Biol*. 2004;8(6):610-6. doi:10.1016/j.cbpa.2004.09.009.

37. Herrmann J, Lerman LO, Lerman A. Ubiquitin and ubiquitin-like proteins in protein regulation. *Circ Res*. 2007;100:1276-1291. doi:10.1161/01.RES.0000264500.11888.f0.

38. Herrmann J, Lerman LO, Lerman A. Ubiquitin and ubiquitin-like proteins in protein regulation. *Circ Res*. 2007;100(9):1276-91. doi:10.1161/01.RES.0000264500.11888.f0.

39. Pickart CM. Ubiquitin in chains. *Trends Biochem Sci*. 2000;25(11):544-548. doi:10.1016/S0968-0004(00)01681-9.

40. Pickart CM. Back to the Future with Ubiquitin. *Cell*. 2004;116(2):181-190. doi:10.1016/S0092-8674(03)01074-2.

41. Scheffner M, Werness BA, Huibregtse JM, Levine AJ, Howley PM. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell*. 1990;63(6):1129-1136. doi:10.1016/0092-8674(90)90409-8.

42. Evans SL, Schön A, Gao Q, et al. HIV-1 Vif N-terminal motif is required for recruitment of Cul5 to suppress APOBEC3. *Retrovirology*. 2014;11:4. doi:10.1186/1742-4690-11-4.

43. Cai Q-L, Knight JS, Verma SC, Zald P, Robertson ES. EC5S ubiquitin complex is recruited by KSHV latent antigen LANA for degradation of the VHL and p53 tumor suppressors. *PLoS Pathog*. 2006;2(10):e116. doi:10.1371/journal.ppat.0020116.

44. Rodrigues L, Filipe J, Seldon MP, et al. Termination of NF-kappaB activity through a gammaherpesvirus protein that assembles an EC5S ubiquitin-ligase. *EMBO J*. 2009;28(9):1283-95. doi:10.1038/emboj.2009.74.

45. Widjaja I, de Vries E, Tscherne DM, García-Sastre A, Rottier PJM, de Haan CAM. Inhibition of the ubiquitin-proteasome system affects influenza A virus infection at a postfusion step. *J Virol*. 2010;84(18):9625-31. doi:10.1128/JVI.01048-10.

46. Delboy MG, Roller DG, Nicola A V. Cellular proteasome activity facilitates herpes simplex virus entry at a postpenetration step. *J Virol*. 2008;82(7):3381-90. doi:10.1128/JVI.02296-07.

47. Duda DM, Scott DC, Calabrese MF, Zimmerman ES, Zheng N, Schulman BA. Structural regulation of cullin-RING ubiquitin ligase complexes. *Curr Opin Struct Biol*. 2011;21(2):257-64. doi:10.1016/j.sbi.2011.01.003.

48. Okumura F, Matsuzaki M, Nakatsukasa K, Kamura T. The Role of Elongin BC-Containing Ubiquitin Ligases. *Front Oncol*. 2012;2(February):10. doi:10.3389/fonc.2012.00010.

49. Guo Y, Dong L, Qiu X, et al. Structural basis for hijacking CBF-β and CUL5 E3 ligase complex by HIV-1 Vif. *Nature*. 2014;505:229-33. doi:10.1038/nature12884.

50. Bergeron JRC, Huthoff H, Veselkov D a., et al. The SOCS-Box of HIV-1 vif interacts with elonginBC by induced-folding to recruit its Cul5-containing ubiquitin ligase complex. *PLoS Pathog*. 2010;6(6). doi:10.1371/journal.ppat.1000925.

51. Fribourgh JL, Nguyen HC, Wolfe LS, et al. Core binding factor beta plays a critical role by facilitating the assembly of the Vif-cullin 5 E3 ubiquitin ligase. *J Virol*. 2014;88:3309-19. doi:10.1128/JVI.03824-13.

52. Jäger S, Kim DY, Hultquist JF, et al. Vif hijacks CBF-β to degrade APOBEC3G and promote HIV-1 infection. *Nature*. 2011:1-16. doi:10.1038/nature10693.

53. Lu Z, Bergeron JRC, Atkinson RA, et al. Insight into the HIV-1 Vif SOCS-box-ElonginBC interaction. *Open Biol*. 2013;3:130100. doi:10.1098/rsob.130100.

54. Wang X, Wang X, Zhang H, et al. Interactions between HIV-1 Vif and human ElonginB-ElonginC are important for CBF-β binding to Vif. *Retrovirology*. 2013;10:94. doi:10.1186/1742-4690-10-94.

55. Zhou X, Evans SL, Han X, Liu Y, Yu XF. Characterization of the interaction of full-length HIV-1 Vif protein with its key regulator CBFβ and CRL5 E3 ubiquitin ligase components. *PLoS One*. 2012;7(3):1-10. doi:10.1371/journal.pone.0033495.

56. Nijman SMB, Luna-Vargas MPA, Velds A, et al. A genomic and functional inventory of deubiquitinating enzymes. *Cell*. 2005;123(5):773-86. doi:10.1016/j.cell.2005.11.007.

57. Holowaty MN, Sheng Y, Nguyen T, Arrowsmith C, Frappier L. Protein interaction domains of the ubiquitin-specific protease, USP7/HAUSP. *J Biol Chem*. 2003;278(48):47753-61. doi:10.1074/jbc.M307200200.

58. Jagannathan M, Nguyen T, Gallo D, et al. A role for USP7 in DNA replication. *Mol Cell Biol*. 2014;34(1):132-45. doi:10.1128/MCB.00639-13.

59. Everett RD, Meredith M, Orr A, Cross A, Kathoria M, Parkinson J. A novel ubiquitin-specific protease is dynamically associated with the PML nuclear domain and binds to a herpesvirus regulatory protein. *EMBO J*. 1997;16(7):1519-30. doi:10.1093/emboj/16.7.1519.

60. Lee JT, Gu W. The multiple levels of regulation by p53 ubiquitination. *Cell Death Differ*. 2010;17(1):86-92. doi:10.1038/cdd.2009.77.

61. Hu M, Li P, Li M, et al. Crystal Structure of a UBP-Family Deubiquitinating Enzyme in Isolation and in Complex with Ubiquitin Aldehyde. *Cell*. 2002;111(7):1041-1054. doi:10.1016/S0092-8674(02)01199-6.

62. Zapata JM, Martínez-García V, Lefebvre S. Phylogeny of the TRAF/MATH domain. *Adv Exp Med Biol*. 2007;597:1-24. doi:10.1007/978-0-387-70630-6_1.

63. Faesen AC, Dirac AMG, Shanmugham A, Ovaa H, Perrakis A, Sixma TK. Mechanism of USP7/HAUSP activation by its C-Terminal ubiquitin-like domain and allosteric regulation by GMP-synthetase. *Mol Cell*. 2011;44(1):147-159. doi:10.1016/j.molcel.2011.06.034.

64. Hu M, Gu L, Li M, Jeffrey PD, Gu W, Shi Y. Structural basis of competitive recognition of p53 and MDM2 by HAUSP/USP7: implications for the regulation of the p53-MDM2 pathway. *PLoS Biol*. 2006;4(2):e27. doi:10.1371/journal.pbio.0040027.

65. Hu M, Li P, Li M, et al. Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell*. 2002;111:1041-1054. doi:10.1016/S0092-8674(02)01199-6.

66. Saridakis V, Sheng Y, Sarkari F, et al. Structure of the p53 binding domain of HAUSP/USP7 bound to epstein-barr nuclear antigen 1: Implications for EBV-mediated immortalization. *Mol Cell*. 2005;18:25-36. doi:10.1016/j.molcel.2005.02.029.

67. Lee H-R, Choi W-C, Lee S, et al. Bilateral inhibition of HAUSP deubiquitinase by a viral interferon regulatory factor protein. *Nat Struct Mol Biol*. 2011;18(12):1336-44. doi:10.1038/nsmb.2142.

68. Sarkari F, Wheaton K, La Delfa A, et al. Ubiquitin-specific protease 7 is a regulator of ubiquitin-conjugating enzyme UbE2E1. *J Biol Chem*. 2013;288:16975-16985. doi:10.1074/jbc.M113.469262.

69. Jager W, Santag S, Weidner-Glunde M, et al. The Ubiquitin-Specific Protease USP7 Modulates the Replication of Kaposi's Sarcoma-Associated Herpesvirus Latent Episomal DNA. *J Virol*. 2012;86(12):6745-6757. doi:10.1128/JVI.06840-11.

70. Marchenko ND, Wolff S, Erster S, Becker K, Moll UM. Monoubiquitylation promotes mitochondrial p53 translocation. *EMBO J*. 2007;26(4):923-934. doi:10.1038/sj.emboj.7601560.

71. Sarkari F, Sanchez-Alcaraz T, Wang S, Holowaty MN, Sheng Y, Frappier L. EBNA1-mediated recruitment of a histone H2B deubiquitylating complex to the Epstein-Barr virus latent origin of DNA replication. *PLoS Pathog*. 2009;5(10). doi:10.1371/journal.ppat.1000624.

72. Linding R, Russell RB, Neduva V, Gibson TJ. GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Res*. 2003;31(13):3701-3708. doi:10.1093/nar/gkg519.

73. Dosztányi Z, Mészáros B, Simon I. ANCHOR: Web server for predicting protein binding regions in disordered proteins. *Bioinformatics*. 2009;25(20):2745-2746. doi:10.1093/bioinformatics/btp518.

74. Ward JJ, McGuffin LJ, Bryson K, Buxton BF, Jones DT. The DISOPRED server for the prediction of protein disorder. *Bioinformatics*. 2004;20(13):2138-2139. doi:10.1093/bioinformatics/bth195.

75. Saridakis V, Sheng Y, Sarkari F, et al. Structure of the p53 binding domain of HAUSP/USP7 bound to Epstein-Barr nuclear antigen 1 implications for EBV-mediated immortalization. *Mol Cell*. 2005;18(1):25-36. doi:10.1016/j.molcel.2005.02.029.

76. Jeong JY, Yim HS, Ryu JY, et al. One-step sequence-and ligation-independent cloning as a rapid and versatile cloning method for functional genomics Studies. *Appl Environ Microbiol*. 2012;78(15):5440-5443. doi:10.1128/AEM.00844-12.

77. Stevenson J, Krycer JR, Phan L, Brown AJ. A practical comparison of ligation-independent cloning techniques. *PLoS One*. 2013;8(12):8-14. doi:10.1371/journal.pone.0083888.

78. Schägger H. Tricine-SDS-PAGE. *Nat Protoc*. 2006;1(1):16-22. doi:10.1038/nprot.2006.4.

79. Mahmood T, Yang PC. Western blot: Technique, theory, and trouble shooting. *N Am J Med Sci*. 2012;4(9):429-434. doi:10.4103/1947-2714.100998.

80. Boivin S, Kozak S, Meijers R. Optimization of protein purification and characterization using Thermofluor screens. *Protein Expr Purif*. 2013;91(2):192-206. doi:10.1016/j.pep.2013.08.002.

81. Ericsson UB, Hallberg BM, DeTitta GT, Dekker N, Nordlund P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem*. 2006;357:289-298. doi:10.1016/j.ab.2006.07.027.

82. Reinhard L, Mayerhofer H, Geerlof A, Mueller-Dieckmann J, Weiss MS. Optimization of protein buffer cocktails using Thermofluor. *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 2013;69(October 2012):209-214. doi:10.1107/S1744309112051858.

83. Hellman LM, Fried MG. NIH Public Access. 2009;2(8):1849-1861. doi:10.1038/nprot.2007.249.Electrophoretic.

84. Holden NS, Tacon CE. Principles and problems of the electrophoretic mobility shift assay. *J Pharmacol Toxicol Methods*. 2011;63(1):7-14. doi:10.1016/j.vascn.2010.03.002.

85. Krauss IR, Merlino A, Vergara A, Sica F. An overview of biological macromolecule crystallization. *Int J Mol Sci*. 2013;14:11643-11691. doi:10.3390/ijms140611643.

86. Rhodes G. *Crystallography Made Crystal Clear*. Elsevier; 2006. doi:10.1016/B978-012587073-3/50015-5.

87. Wlodawer A, Minor W, Dauter Z, Jaskolski M. Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *FEBS J*. 2008;275:1-21. doi:10.1111/j.1742-4658.2007.06178.x.

88. Powell HR, Johnson O, Leslie AGW. Autoindexing diffraction images with iMosflm. *Acta Crystallogr Sect D Biol Crystallogr*. 2013;69:1195-1203. doi:10.1107/S0907444912048524.

89. Evans P. Scaling and assessment of data quality. *Acta Crystallogr Sect D Biol Crystallogr*. 2006;62:72-82. doi:10.1107/S0907444905036693.

90. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr*. 2007;40:658-674. doi:10.1107/S0021889807021206.

91. Emsley P, Cowtan K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr Sect D Biol Crystallogr*. 2004;60:2126-2132. doi:10.1107/S0907444904019158.

92. Murshudov GN, Skubák P, Lebedev A a., et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr Sect D Biol Crystallogr*. 2011;67:355-367. doi:10.1107/S0907444911001314.

93. Romier C, Ben Jelloul M, Albeck S, et al. Co-expression of protein complexes in prokaryotic and eukaryotic hosts: Experimental procedures, database tracking and case studies. *Acta Crystallogr Sect D Biol Crystallogr*. 2006;62:1232-1242. doi:10.1107/S0907444906031003.

94. Holowaty MN, Zeghouf M, Wu H, et al. Protein profiling with Epstein-Barr nuclear antigen-1 reveals an interaction with the herpesvirus-associated ubiquitin-specific protease HAUSP/USP7. *J Biol Chem*. 2003;278(32):29987-94. doi:10.1074/jbc.M303977200.

95. Sarkari F, La Delfa A, Arrowsmith CH, Frappier L, Sheng Y, Saridakis V. Further Insight into Substrate Recognition by USP7: Structural and Biochemical Analysis of the HdmX and Hdm2 Interactions with USP7. *J Mol Biol*. 2010;402(5):825-837. doi:10.1016/j.jmb.2010.08.017.

96. Pflugrath JW. Macromolecular cryocrystallography - Methods for cooling and mounting protein crystals at cryogenic temperatures. *Methods*. 2004;34:415-423. doi:10.1016/j.ymeth.2004.03.032.

97. Woelk T, Sigismund S, Penengo L, Polo S. The ubiquitination code: a signalling problem. *Cell Div*. 2007;2:11. doi:10.1186/1747-1028-2-11.

98. Bochkarev A, Bochkareva E, Frappier L, Edwards AM. The 2.2 A structure of a permanganate-sensitive DNA site bound by the Epstein-Barr virus origin binding protein, EBNA1. *J Mol Biol*. 1998;284(5):1273-8. doi:10.1006/jmbi.1998.2247.

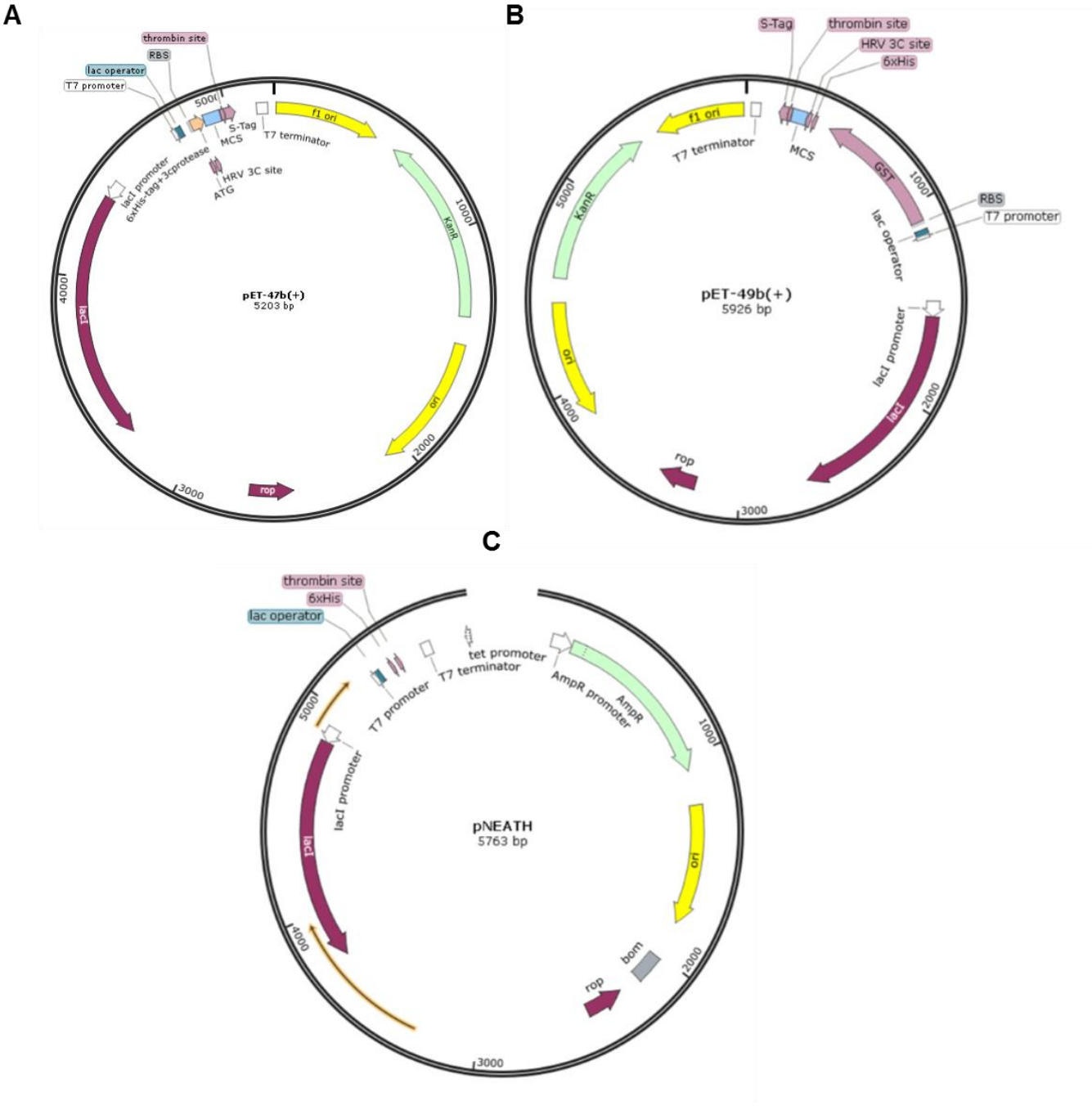

Figure S1 – Map illustration of the vectors used for cloning in this study. *Design of the reporter vector a) pET47 b(+), encoding resistance for kanamycin with a 6xHis-Tag; b) pET49 b(+) encoding resistance for kanamycin with a GST-6xHis-Tag and c) pNEATH encoding resistance for ampicillin with a 6xHis-Tag. The desired genes were cloned in the Multiple Cloning Site (MCS) region.*

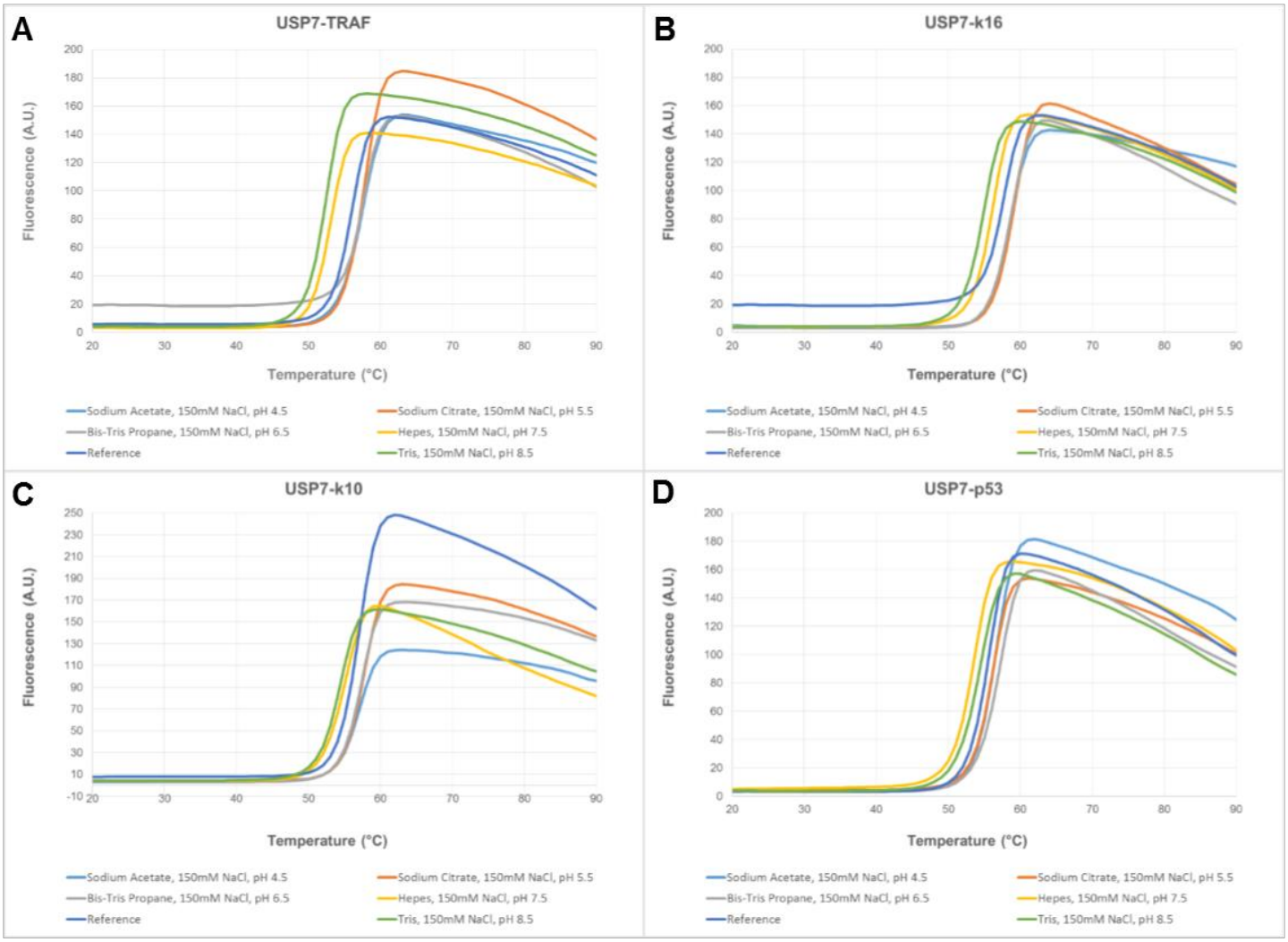

Figure S2 – Fluorescence- based thermal shift analysis of USP7-TRAF truncation alone and in complex with three peptides. *Thermofluor-based protein-unfolding curves of a) USP7-TRAF, b) USP7-k16, c) USP7-k10 and d) USP7-p53 for accessing the thermal stability effect of buffer and pH. The thermofluor experiment was prepared as described in 2.8, using 2.5 µg of protein and a 5-fold (final concentration) of SYPRO Orange. The experiment was performed from 20 °C to 90 °C in increments of 1 °C.*