Gonçalo Pedro da Silva Nunes

Bachelor of Science in Biochemistry

Master's Student in Biochemistry for Health

The role of Ubiquitin Specific Protease 7 on Latency Associated Nuclear Antigen DNA binding.

Dissertation submitted for the degree of Master of Science in Biochemistry for Health

Supervisor: Doctor Colin McVey, Principal Investigator, ITQB NOVA UNL

September 2018

Gonçalo Pedro da Silva Nunes
Bachelor of Science in Biochemistry
Master's Student in Biochemistry for Health

The role of Ubiquitin Specific Protease 7 on Latency Associated Nuclear Antigen DNA binding.

Dissertation submitted for the degree of Master of Science in Biochemistry for Health

Supervisor: Doctor Colin McVey, Principal Investigator, ITQB NOVA UNL

Júri:

President: Dr. Pedro Manuel Henrique Marques Matias, Principal Investigator, ITQB NOVA
Examiner: Dr. Elin Moe, Investigator, ITQB NOVA
Jury Members: Dr. Margarida Archer Baltazar Pereira da Silva Franco Frazão, Principal Investigator, ITQB NOVA
Dr. Colin Edward McVey, Principal Investigator, ITQB NOVA

ITQB NOVA - Instituto de Tecnologia Química e Biológica António Xavier

September 2018

Acknowledgements

It is with immense gratitude that I acknowledge the incredible support of my adviser Doctor Colin McVey, it was a great honour to work with him and to learn not only about structural virology, but also about new ways to push myself to go above and beyond my limits. I also owe my deepest gratitude to Rute Chitas, who stood by my side day after day, coaching me on several subjects, techniques and methods, with an impressive persistence and patience. This thesis would have not been possible without them, holding my hand and sparing no efforts, in such a kind and motivating manner. I cannot find words to express my gratitude for all your help.

I also would like to thank the whole team of the macromolecular crystallography unit of ITQB NOVA for the assistance, support and friendship, everyone welcomed me in and made me feel comfortable during this whole year.

I thank my wife, for her everlasting understanding and support, and above all for her love. I also wish to thank my big family of pets, my 3 dogs and my 3 cats, the best friends I ever had.

“O'er ladies lips, who straight on kisses dream,

Which oft the angry Mab with blisters plagues,

Because their breaths with sweetmeats tainted are:

Sometime she gallops o’er a courtier’s nose,

And then dreams he of smelling out a suit;

And sometime comes she with a tithe-pig’s tail

Tickling a parson’s nose as a’ lies asleep...”

Mercutio (to Romeo) in Romeo and Juliet by Shakespeare.

Abstract

The implications of USP7 (Ubiquitin Specific Protease 7, deubiquitinating enzyme involved in several crucial molecular pathways) interactions with LANA (Latency Associated Nuclear Antigen protein, that facilitates the tethering of viral γ-Herpesvirus episomes into the host's DNA and promotes its replication) has multiple health consequences, since USP7 is a protein that has an essential role it the regulation of many cellular functions, like the p53-mdm2 pathway, that regulates cellular apoptosis and prevents cancer.

We aim to study the structural, biochemical and biophysical relations between USP7 and LANA, to understand the role of USP7 in the USP7-LANA DNA binding function. Several constructs of both USP7 and LANA (from different species: Murine gamma-herpesvirus 68 LANA, or mLANA; Kaposi's sarcoma-associated herpesvirus, or kLANA, from humans) were designed to be cloned, expressed, purified and complexed, for structural determination through X-ray crystallography, and for DNA binding affinity determination, through EMSA and ITC techniques.

So far, the role of increasing affinity for the tethering of the viral DNA into the host's chromatin, while inhibiting and thus regulating the replication of the viral DNA episome seems to be more likely. The results indicate that USP7 aids LANA in its tethering functions and also regulates its activity in the replication process, thus maintaining a lower quantity of episomes in the cell, during the latent cycle. The "stealth" mode of the virus depends on preventing its detection, so this might contribute greatly to its own survival and immortalization.

Keywords: Kaposi's sarcoma herpesvirus; Latency associated nuclear antigen; Ubiquitin Specific Protease 7; LANA-USP7 complex; Lana Binding Site DNA.

Resumo

As implicações das interações de USP7 (Protease Específica da Ubiquitina 7, enzima desubiquitinante envolvida em várias vias moleculares cruciais) com o LANA (Antigenio Nuclear Associado à Latência proteína, que facilita a fixação de episssomas virais γ-Herpesvirus no DNA do hospedeiro e promove sua replicação) têm múltiplas consequências para a saúde, já que a USP7 é uma proteína que tem um papel essencial na regulação de muitas funções celulares, como a via p53-mdm2, que regula a apoptose celular e previne o cancro.

O nosso objetivo é estudar as relações estruturais, bioquímicas e biofísicas entre USP7 e LANA, para entender o papel da USP7 na função de ligação do DNA ao complexo USP7-LANA. Diversas construções tanto do USP7 quanto do LANA (de diferentes espécies: gama-herpes vírus 68 LANA murino, ou mLANA; vírus do sarcoma de Kaposi associado, ou kLANA, de humanos) foram projetadas para serem clonadas, expressas, purificadas e complexadas, para determinação estrutural através de cristalografia de raios X, e para determinação de afinidade de ligação ao DNA, através de técnicas de EMSA e ITC.

Até agora, o papel do aumento da afinidade para a fixação do DNA viral na cromatina do hospedeiro, enquanto inibe e, portanto, regula a replicação do episssoma do DNA viral parece ser mais provável. Os resultados indicam que o USP7 auxilia o LANA nas suas funções de ancoragem e também regula a sua atividade no processo de replicação, mantendo assim uma menor quantidade de episssomas na célula, durante o ciclo latente. O modo “furtivo” do vírus depende da prevenção da sua detecção, portanto isso pode contribuir muito para sua própria sobrevivência e imortalização.

Palavras-chave: herpesvírus do sarcoma de Kaposi; Antigenio nuclear associado à latência; Protease Específica da Ubiquitina 7; Complexo LANA-USP7; Local de DNA de ligação ao LANA.

Contents

Figures Content

Tables Content

Abbreviations

USP7	Ubiquitin-Specific Protease 7.
HAUSP	Herpesvirus-Associated Ubiquitin-specific Protease 7.
DUB	Deubiquitinating enzyme.
TNF	Tumor Nercosis Factor.
TRAF	TNF-receptor associated Factors.
LANA	Latency Associated Nuclear antigen.
KS	Kaposi's Sarcoma.
KSHV	Kaposi's Sarcoma Associated Herpesvirus.
HIV-1	Human Immunodeficiency Virus.
HHV8	Human Herpesvirus 8.
MHV-68	Murine gamma-herpesvirus 68.
ORF73	Open Reading Frame 73.
kLANA	KSHV LANA.
mLANA	MHV-68 LANA.
LBS	LANA Binding Site.
DNA	Deoxyribonucleic Acid.
DBD	DNA Binding Site.
TR	Terminal Repeats.
PCR	Polymerase Chain Reaction.
EMSA	Electrophoretic Mobility Shift Assay.
MATH	Meprin and TRAF Homology.
NTD	N-Terminal Domain.
EBV	Epstein-Barr virus.
Tm	Melting Temperature.
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis.
His	Histidine.
SAP	Shrimp Alkaline Phosphatase.

1 Introduction

1.1 Herpes Viridae

Herpesviridae are one of the most leading causes of human viral disease, right after influenza and cold viruses. They can either unleash disease or remain silent for many years, only to be reactivated when a more suitable environment manifestates itself (48). The name herpes comes from the Latin *herpes* which, in turn, comes from the Greek *herpein*, which means to creep. It is a direct reflection of the creeping or spreading nature of the skin lesions caused by many herpes virus types (48). There are at least 25 viruses in the family Herpesviridae (currently divided into three sub-families). Eight or more herpes virus types are known to infect man frequently: Herpes simplex virus Type 1 (HSV-1); Herpes simplex virus Type 2 (HSV-2); Epstein Barr virus (EBV); Cytomegalovirus (CMV); Varicella Zoster Virus (VZV); Human herpes virus 6 (exanthum subitum or roseola infantum); Human herpes virus 8 (Kaposi's sarcoma-associate herpes virus) (48).

Once a patient is infected, it remains for life. The initial infection may be followed by latency with subsequent reactivation. Usually, after middle age, patients have antibodies to most herpes viruses, except for HHV-8 (or KSHV, Kaposi's sarcoma Herpesvirus) (48). Like Epstein Barr Virus, KSHV is located in the B lymphocytes during the latent state of infection (81), while its target cell type are endothelial cells during lytic infection. Its transmission vector is suggested to be the exchange of body fluids (48).

Herpes viruses have a unique four-layered structure: genome, capsid, tegument and envelope. A core containing the large, double-stranded DNA genome is enclosed by an icosapentahedral capsid, which is composed of capsomers. The capsid is surrounded by an amorphous protein coat called the tegument. it is encased in a glycoprotein-bearing lipid bilayer envelope (**Figure 1.1**) (48, 49).

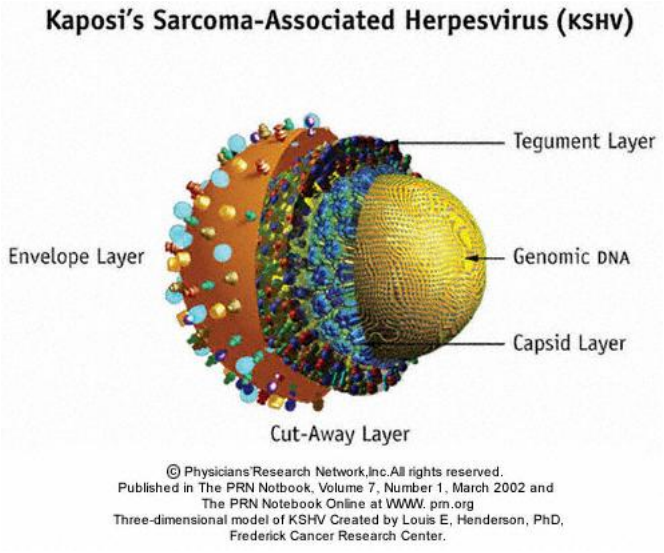


Figure 1. 1: Kaposi's sarcoma Herpes Virus 3D model. Adapted from Physicians Research Network.

According to its classification, herpes viruses are divided into three groups: the alpha-herpesviruses, comprised of herpes simplex virus types 1 and 2 and varicella-zoster virus, have a short replicative cycle, induce cytopathology in monolayer cell cultures and have a broad host range; beta-herpesviruses, include cytomegalovirus and human herpesviruses 6 and 7, with a long replicative cycle and restricted host range; and finally gamma-herpesvirus, to which the Epstein-Barr virus and the human herpesvirus 8 (KSHV) belong, and which have a very restricted host range (49).

1.2 Kaposi's sarcoma-associated Herpesvirus (KSHV)

Kaposi's sarcoma-associated herpesvirus (KSHV), also referred to as human herpesvirus 8 (HHV-8), is a recently identified human gamma-herpesvirus which is closely associated with several malignancies, such as Kaposi's sarcoma (KS), primary effusion lymphoma (PEL) and multicentric Castleman's disease (MCD) (50-57). KSHV is a double-stranded DNA virus with a genome that contains a 140 kb unique coding region, with multiple GC-rich terminal repeats (55). The most interesting feature of the herpesvirus is an initial lytic infection in hosts, followed by the establishment of a lifelong latent infection, maintaining the viral genome in the cell in the form of episomes, in different tissues. Occasionally, reactivation of the virus occurs, shifting from latent to lytic state (58). KSHV infection, like the infection by other herpesvirus, displays latent and lytic replication modes in its life cycle (59). While maintaining a tightly regulated latent infection in infected cells, KSHV is also capable of keeping a lytic infection undergoing in a percentage of cells in the population. Since only a small number

of the viral genes are expressed during latency, the virus can escape host immune surveillance, establishing a persistent latent infection (55, 60).

1.3 Latency Associated Nuclear Antigen (LANA)

Among the very few genes expressed during KSHV (Kaposi Sarcoma Herpesvirus) latent infection, LANA (encoded by ORF73) is the dominant one (18). LANA is a multifunctional protein essential for latency, required for the maintenance of the viral episomes. It directs the replication and tethers the viral episome to mitotic chromosomes (**Figure 1.2**), to ensure efficient segregation of the newly replicated molecules during mitosis (19). It also activates the expression of other KSHV genes. All these functions are accomplished by binding to specific terminal repeat (TR) DNA sequences in the KSHV genome.

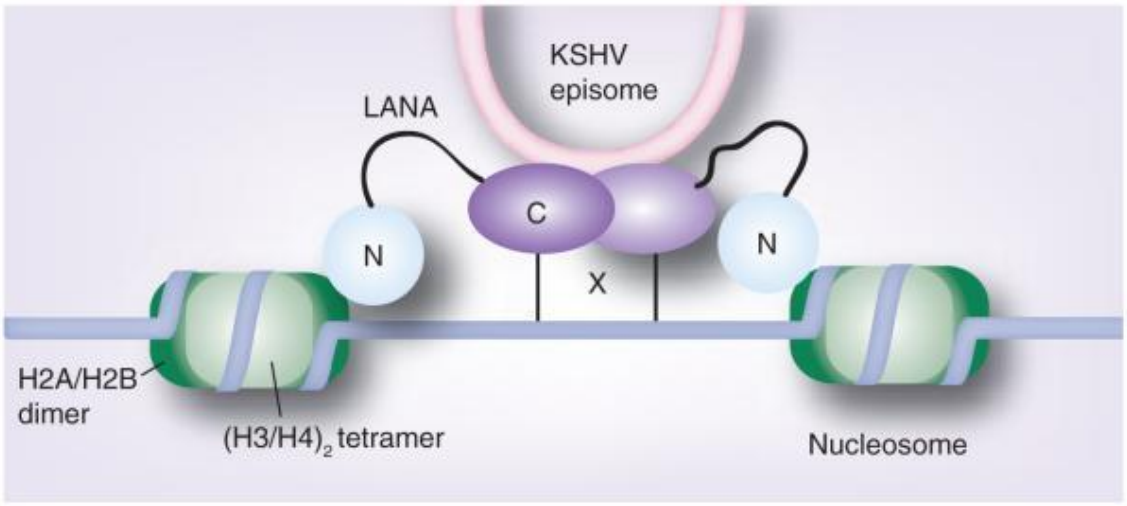


Figure 1. 2: Model of latency-associated nuclear antigen binding to Kaposi's sarcoma-associated herpesvirus episome and cellular chromosomes. Adapted from (46).

kLANA (Kaposi sarcoma herpesvirus LANA, of human origin) consists of 1162 amino acid residues, divided into four unique regions: a proline-rich region, a central acidic repeat region, a putative leucine zipper and a dimerization and viral DNA binding domain (DBD) (20). Its N-terminal region is flexible and intrinsically disordered, while the C-terminal region is structurally ordered and stable (19). The first 33 N-terminal residues interact with histones H2A/H2B on the nucleosome surface (21-24). The C-terminus DBD operates both dimerization and DNA-binding. The DBD binds to specific sequences present within the terminal repeat (TR) of the viral episomal DNA, known as LANA binding site (LBS). Three LBS (LBS-1, -2 and -3) have been identified so far, each of them having different binding affinities to LANA protein. Each LBS region binds one LANA dimer, containing two recognition half-sites. The LBS sequences are present in tandem within TR DNA, so they require LANA to oligomerize and interact with LBS in a cooperative manner (**Figure 1.3**) (25-29). The central region of kLANA contains several repeat sequences which vary in length among KSHV isolates (30).

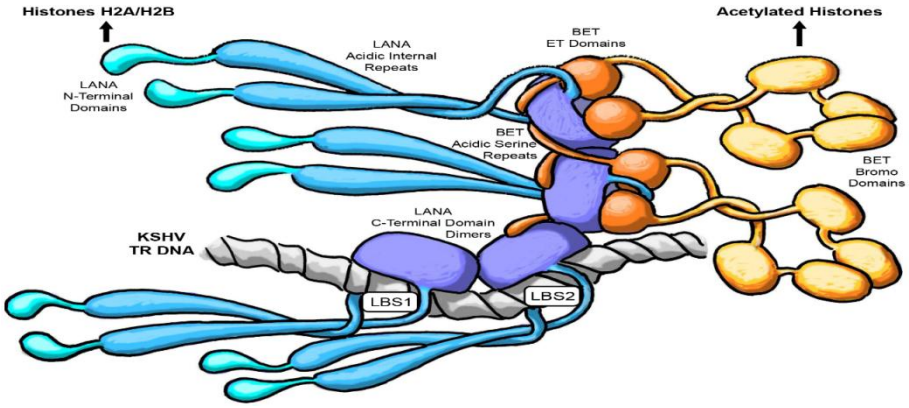


Figure 1. 3: Hypothetical model of KSHV LANA oligomers. Dark blue: kLANA CTD dimers. Cyan: kLANA acidic internal repeat regions and N-terminal domains. Grey: DNA. Orange: BET proteins. LBS3 is not considered on this model, which should bind the KSHV TR DNA right after LBS1, using a biphasic mode of binding. Adapted from (47).

The MHV-68 LANA (Murine Herpesvirus-68, or mLANA) protein, homologue to kLANA, consists of only 314 residues, lacking the acidic repeat and the glutamine-rich sequence, but is also able to efficiently tether the viral DNA and maintain viral episomes (31). mLANA binding sites (mLBS) have also been identified in MHV-68 TR DNA (32). Both kLANA and mLANA, other than having the tethering function, are also required to initiate episome replication. For this purpose, they recruit various cellular replication proteins. They are also involved in activation and repression of transcription (**Figure 1.4**) (33-35). In order to perform its multiple functions, LANA requires intrinsic flexibility in its structure and employs strategies such as oligomerization and post-translational modification.

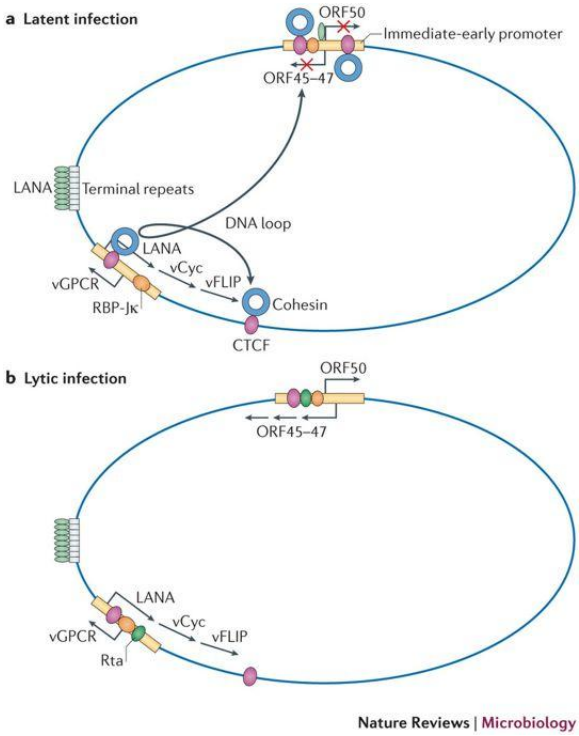


Figure 1. 4: Establishment of the KSHV epigenome. Primary infection of KSHV results in combination of both lytic and latent transcription. Adapted from (45).

The efficient maintenance of viral episomes requires two minimal replicator elements (MRE) containing the LBS sequence to be separated by 801 bps, irrespective of the DNA (36, 37). This fact indicates that a spatial arrangement is critical for accommodating a super molecular structure that incorporates LANA bound to TR DNA, in addition to other host cell proteins. This type of assembly could be responsible for the characteristic LANA speckles observed in mitotic KSHV infected cells (19).

The C-terminal DBD of LANA from KSHV and MHV-68 have similar structural topologies, as the reported crystal structures revealed (38-40). The LANA DBD dimerizes by forming a central eight-stranded anti parallel β-barrel with four strands contributing from each monomer and flanked on either side by three α-helices (**Figure 1.5**). The LANA DBD structure is similar to EBV EBNA-1 (**Figure 1.6**) and HPV E2 latent proteins (40-42).

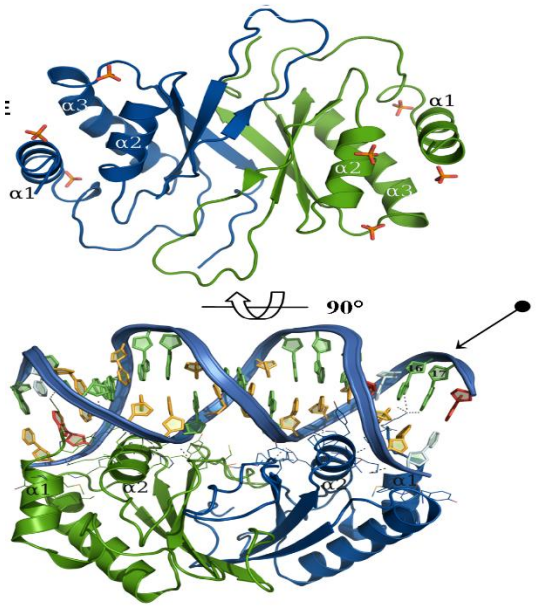


Figure 1. 5: Structural representation of mLANA DBD. (TOP) A ribbon representation of the dimer in the same orientation, showing the secondary structure arrangement. (BOTTOM) Structural model of a mLANA DBD DNA complex (rotated relative to top perspective) using mLBS1 DNA docked onto the 18 bp EBNA1 DNA recognition sequence followed by energy minimisation using YASARA. Adapted from (40).

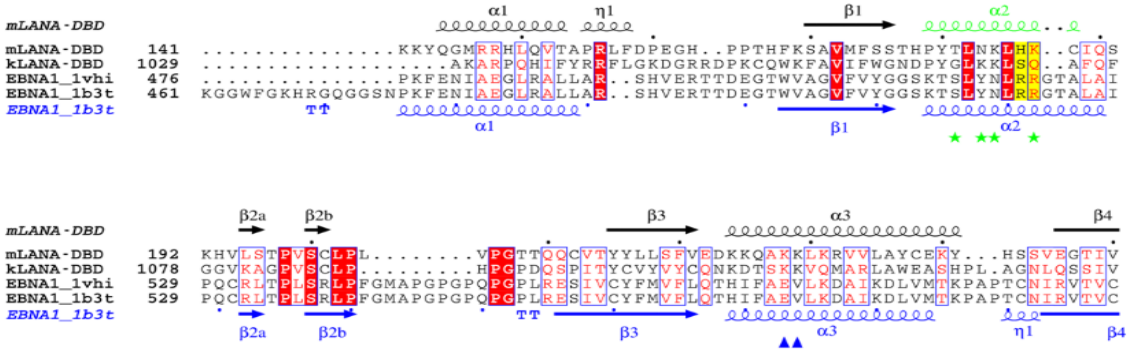


Figure 1. 6: Conserved sequence of mLANA DBD. Structure based sequence alignment of the MHV-68 mLANA DBD with the DBD domains of KSHV kLANA, EBV EBNA1 (PDB ID: 1VHI), and EBNA1 DNA complex (PDB ID: 1B3T). Adapted from (40).

A reported multiple-point mutant kLANA protein crystal structure bound with kLBS-1 fragment showed an asymmetric binding mode, different to the previously observed EBNA-1 and E2 DNA complexes (29). A total of nine mutations on the opposite side of the DNA binding interface made the structure similar to the EBNA-1 protein, to improve the crystallizability of the LANA protein. The DNA-free structures show higher oligomers formation for both kLANA and mLANA DBD. kLANA shows a closed ring with a bend occurring at the dimer-dimer interface, mLANA shows linear oligomerization in an end-to-end fashion (38-40). The disruption of this type of interface leads to impared plasmid replication and maintenance *in vivo* (38, 39).

The biological relevance of the bend vs. linear tetramer formation, displayed in the crystal structures of KSHV and MHV-68 LANA DBD respectively, still remains unclear due to the absence of structural data in the presence of LBS 1-2 TR DNA. Quite recently, a ground-breaking research reported that kLANA and mLANA are partially interchangeable in their function, as they can bind to the respective non-cognate LBS DNA *in vitro* (19). The same researchers demonstrated also that kLANA and mLANA bind DNA very differently, with opposite thermodynamic binding profiles: mLANA DBD DNA binding is exothermic, while kLANA DBD DNA binding is endothermic, due to the fact that kLANA absorbs energy, in order to bend. They also showed that the presence of tandem low and high affinity binding sites is obligatory for correct LANA binding (19).

1.4 Ubiquitin Specific Protease 7 (USP7)

The Ubiquitin-Specific Protease 7 (USP7), or HAUSP (Herpesvirus-associated USP), is an important deubiquitinating enzyme (**Figure 1.7**), involved in the regulation processes of several key proteins that have crucial functions, such as tumour suppressing (p53 protein), DNA repair, immune responses, viral activity, epigenetic modulation and transcription. USP7 is involved in cellular responses to various stress stimuli and in cellular homeostasis maintenance. Dysregulation of USP7 is associated with diseases such as cancer and neurodegenerative disorders (1).

USP7 is comprised by 1102 amino acid residues, with a total molecular weight of 128.3 kDa. Its regions can be divided into 3 domains: An N-terminal TRAF/MATH domain (residues 50-208); a catalytic domain (or USP domain, residues 214-521); a C-terminal region with a total of five Ub-like domains (**Figure 1.7**). The TRAF (tumour necrosis factor receptor-associated factors) corresponds to a region that allow proteins to interact and regulate different members of the TNF (tumor necrosis factor) receptors. Interestingly, USP7 is the only ubiquitin specific protease that holds a TRAF region. The USP7 TRAF domain adopts an eight-stranded anti-

parallel β-sandwich structure, with strands β1, β5, β6 and β8 in one sheet and strands β2, β3, β4 and β7 in the other (**Figure 1.8**) (4).

Originally, USP7 was identified as an interaction partner of ICP0 (or Vmw110), from the herpes simplex virus 1 (HSV-1). ICP0 plays a role in the initiation of the viral lytic life cycle (11, 61). ICP0 is a ubiquitin E3 ligase and induces auto-ubiquitination, this is counteracted by USP7 binding which leads to its stabilization (12, 61). ICP0 can also ubiquitinate USP7 and reduce cellular USP7 levels (13). However, when ICP0 is expressed in limiting amounts, no reduction in USP7 levels can be observed, suggesting that ICP0 deubiquitination by USP7 is dominant over ICP0-induced degradation of USP7 (13). The difference, in the case of ICP0, is that it binds within the C-terminal domain of USP7, in contrast with other USP7-binding proteins (14).

USP7 has been identified as a binder of many other substrates, enhancing or inhibiting their functionalities, which are related to the stress response pathway, epigenetic silencing, neurodegenerative disorders and progression of infection by DNA viruses (62). For example, USP7 participates in the p53-MDM2 pathway, deubiquitinating both proteins, but has a higher affinity for MDM2. USP7 stabilizes MDM2 and leads to MDM2-catalyzed degradation of p53 protein (1, 2, 3, 4, 5, 6, 7). Both proteins (MDM2 and p53) bind in a mutually exclusive manner to the USP7 N-terminal tumor necrosis factor (TNF) receptor-associated factor (TRAF)-like domain, recognizing the same shallow groove (**Figure 1.8**) on the USP7 surface (4, 7).

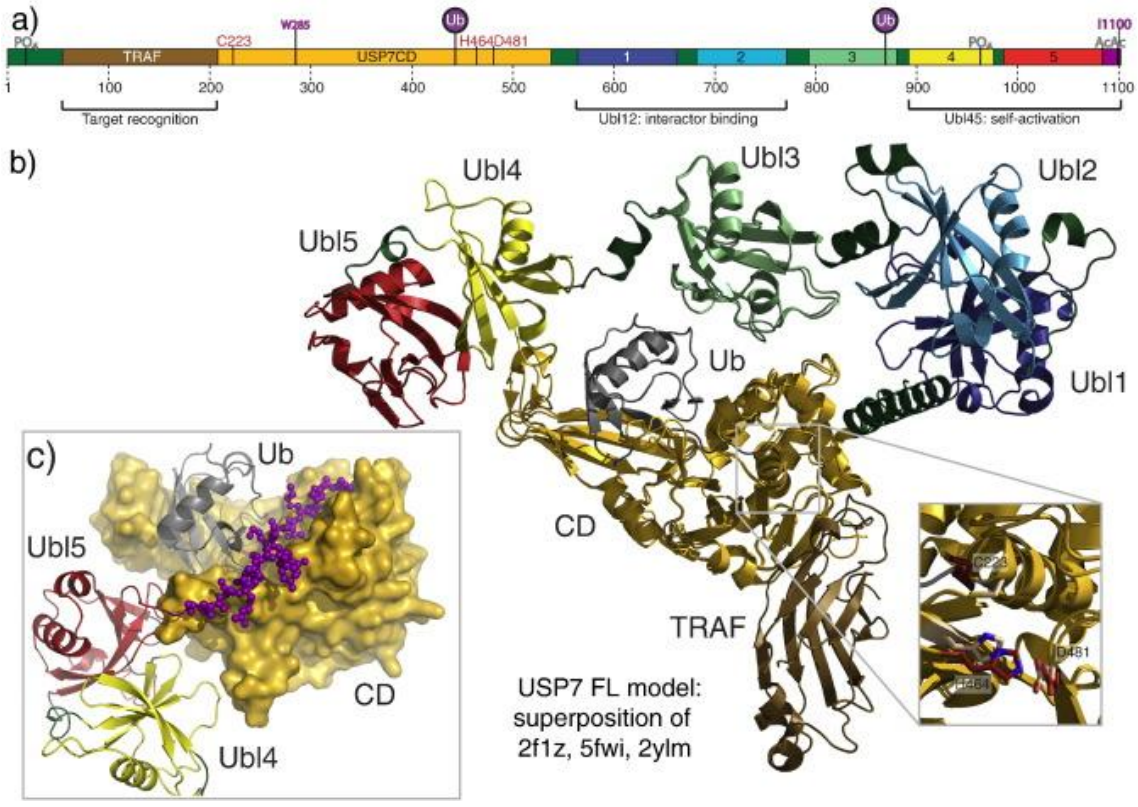


Figure 1. 7: Structural modelling of USP7. (a) Schematic figure of USP7 domain architecture. (b) A structural model of the full-length USP7. (c) The crystal structure of USP7 in activated state (5JTV) illustrating how Ubl45 and the activating peptide bind onto the catalytic domain (CD). Adapted from (44).

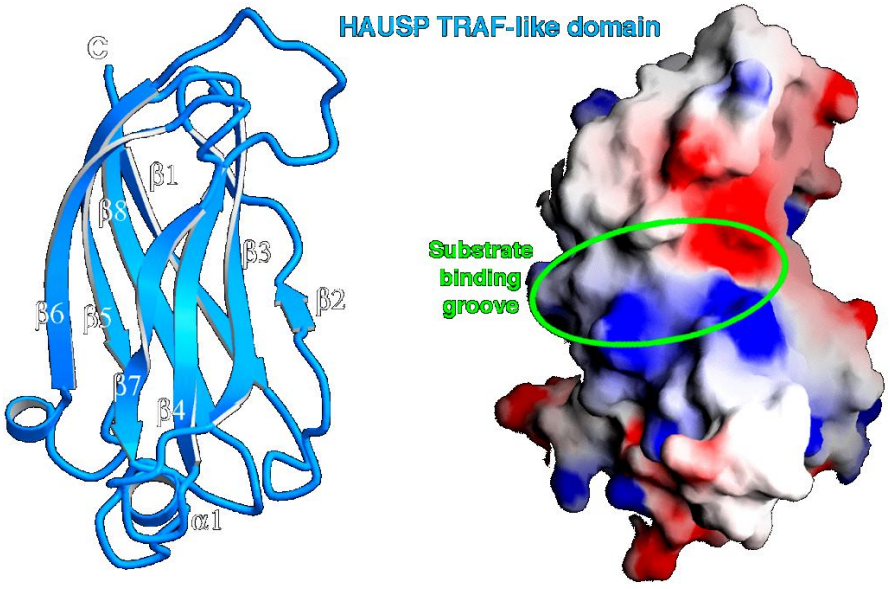


Figure 1. 8: Structure of the HAUSP TRAF-like domain in a ribbon diagram (left) and a surface representation (right). Secondary structural elements (left) and the putative substrate-binding groove (right) are labelled. Adapted from (4).

The higher affinity for MDM2 can be accounted for the more extensive contacts that it makes to USP7 (4, 7). Supporting this fact is that, in competition assays, a MDM2 peptide efficiently displaced a p53 peptide (4). Upon genotoxic stress or nucleotide deprivation, the E3 ubiquitin ligase TRIM21 releases GMPS (Guanosine 5' monophosphate synthetase), which transcends from the cytosol into the nucleus and there it binds USP7 and p53, preventing the binding of MDM2 and therefore leading to p53 stabilization and transactivation of its target genes (**Figure 1.9**).

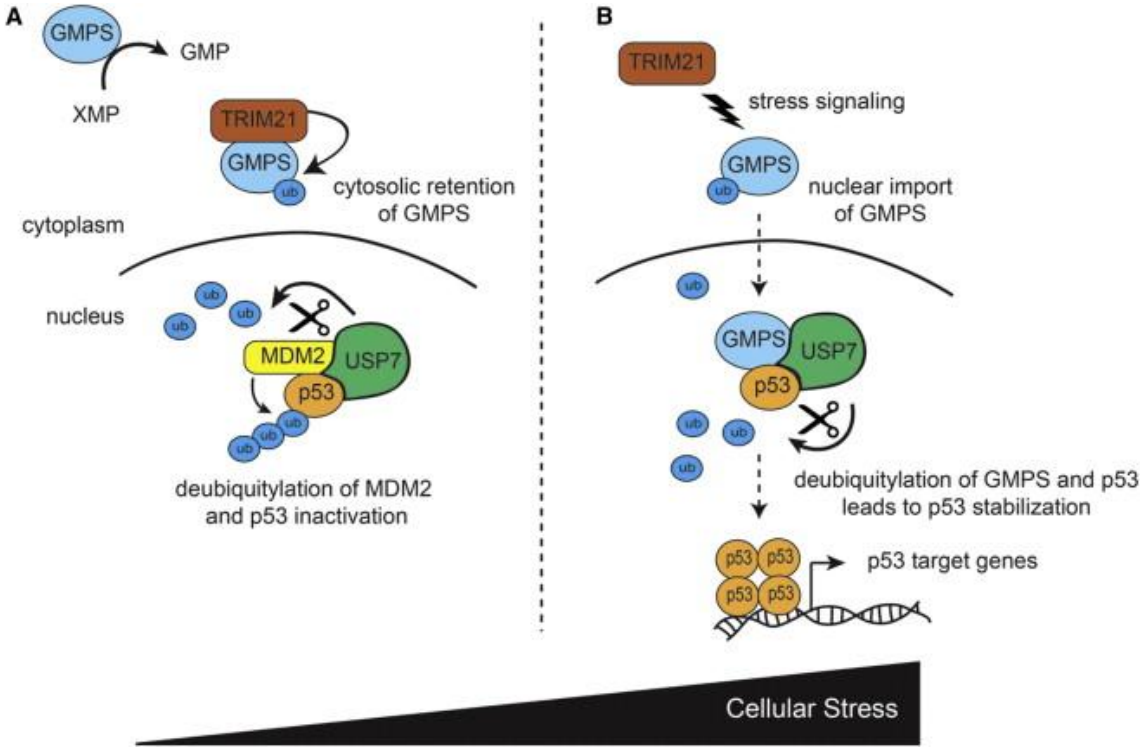


Figure 1. 9: The Stress-Induced TRIM21/GMPS/USP7 Cascade for p53 Stabilization. Adapted from (43).

The consensus peptide sequence for recognition by USP7 has been identified as P/A-X-X-S (with X as any residue), in which the Serine (S) residue plays an important role mediating the contact to USP7 (7, 8). USP7 is also involved in the regulation of DAXX (death-domain-associated protein) and HdmX, both regulatory participants in the p53-Mdm2 pathway (9, 10). DAXX fine tunes the interplay between MDM2 and USP7, which demonstrates the relevance of a tight regulation of USP7 in cell fate decisions (10).

Another protein that interacts with USP7 is Epstein-Barr virus EBNA-1, which is a functional homologue of LANA protein. It was found that an EBNA-1 mutant, unable to interact with USP7, performs decreased transcriptional activation and increased replication activity of the latent EBV origin, but maintains normal activity on the segregation function of EBNA-1 (15). USP7 also shows augmented recruitment to the FR (family of repeats) element of *oriP*, the origin of latent EBV replication (15), and can decrease H2B ubiquitination together with GMPS (guanosine 5' monophosphate synthetase), influencing gene expression (15, 16). For these reasons, there are evidences that USP7-EBNA-1 interaction regulates plasmid maintenance and/or latent replication during EBV infection (15). It was also shown that EBNA-1 binds with higher affinity to the N-terminal domain of USP7 than p53, so it competes with p53 for USP7 binding during infection, which is supported by structural studies (7, 17). LANA protein, as well as the LANA homologues of two other gamma-2 herpesviruses, also interact with USP7 (1). The residues critical for USP7 binding are in the C-terminal domain of LANA, in a short region that displays high sequence homology to EBNA-1, and like EBNA-1, LANA interacts with the N-terminal part of USP7, like the USP7 targets p53, MDMX and MDM2.

1.5 USP7-LANA interaction

Since the interaction between USP7 and ORF73 proteins seems to be conserved among different gamma-herpesviruses, it has been suggested that it may be significant for the viral latent life cycle (1). As such, it is expected to have an impact on LANA's episomal tethering and viral DNA latent replication functions.

The C-terminal domain of LANA alone can mediate the interaction with USP7 (1), so truncations of the protein, including residues from that region, are suitable for the study of the interaction. A small region of 16 amino acid residues (localized between residues 971 to 986) has been shown to still interact with USP7, and the deletion of this region from the C-terminal domain abolishes the interaction (1). This USP7 binding region was also found to be conserved in EBNA-1, which is a functional homologue of LANA, but displays a limited sequential homology to it. In relation to this region, several residues are identical, which again shows conservation of the USP7 binding site in LANA and EBNA-1 (1). This USP7 binding region is centered on a PGPS motif in LANA, corresponding to published consensus sequences (71, 72). it was also shown that substitution of this motif with alanines eliminated the binding of USP7 to the C-terminal domain of LANA (1). So, the residues 971 to 986 in the C-terminal domain of LANA demonstrate a potential USP7 binding site, with its central PGPS motif mediating contact to USP7. Typically, USP7 interacts with its partners through the N-terminal TRAF domain (15), except for a few proteins, such as ICP0. LANA was also observed binding to the N-terminal TRAF domain of USP7 (1), similarly to EBNA-1, p53, MdmX and Mdm2. The possibility of further USP7 regions being involved in the interaction with LANA was not excluded, though.

1.6 Possible roles of USP7 in LANA DNA binding activity

Regarding the role of USP7 in EBNA-1 activity, it has been demonstrated that EBNA-1 competes for the same binding pocket as p53 and Mdm2 in the N-terminal TRAF domain of USP7, and it binds EBNA-1 with higher affinity than p53 or Mdm2 (63, 64, 65, 66), so since USP7 is necessary to stabilize p53, removing from it the polyubiquitin chains that normally signal degradation (67, 68, 69), EBNA-1 USP7 binding reduces the levels of available free USP7 in the cell, causing decreased p53 accumulation in response to DNA damage (63, 70), so it dampens p53 induced cell cycle arrest and induction of apoptosis, therefore allowing the Epstein Barr virus infected cells to become immortalized.

As for the role of USP7 in LANA activity, this thesis was focused on LANA's DNA binding activity, specifically on the viral episome tethering (which is related to LBS 1-2) and on the viral DNA latent replication (which is related to LBS 1-3) mechanisms. Other interactions of

USP7-LANA complexes with other proteins and with other expression promoters has been disregarded. For example, the influence that USP7 may have on LANA's inhibition of Rta expression has not been considered. Whether USP7 has a simple affinity increasing role in LANA DNA binding, or it participates in a more elaborated intrinsic mechanism upon the binding is what remains to be determined, since the affinity increase has already been demonstrated by previous studies (1), even though it is still not clear if this increase is specific to a particular DNA binding site or if it delivers a more efficient structural capacity to bind any sequence of DNA. A LANA mutant, lacking the USP7 binding site, has consistently showed increased replication activity, in a transient replication assay (1). The same has been observed with EBNA-1 (15). These observations may indicate that recruitment of USP7 to LANA can modulate latent viral DNA replication (1). EMSA assays have shown that the increased replication activity of LANA mutant (with deleted USP7 binding site) cannot be attributed to an enhanced DNA binding ability (1).

Since it is widely accepted that the LANA region required for DNA binding and dimerization is comprised of residues 1025 to 1162 (1, 73, 74), the USP7 binding site is therefore outside of this domain. Since USP7 was shown to deubiquitinate histone H2B (75), and mono-ubiquitination of H2B was reported to influence transcription, both positively and negatively (65, 66, 76, 77), effects on processes such as DNA replication cannot be excluded (1). Therefore, a complete lack or reduced recruitment of USP7 to LANA could influence H2B mono-ubiquitination, enhancing or inhibiting deubiquitination of H2B, causing an impact on viral DNA accessibility for transcription or replication (1). The KSHV episome also forms complexes with histones (55, 78), so USP7 could influence accessibility of the viral DNA for latent replication or for viral persistence mediated by LANA, since LANA is able to bind to histone H2B directly (1). In fact, it has been suggested that a USP7's role in the replication of latent viral DNA, among gamma-1 and gamma-2 herpesvirus, seems to be a conserved feature (1).

The work in this thesis is set to determine the role of USP7 in MHV and KSHV LANA's DNA binding activity and specify USP7's impact on LANA activity during the latent viral DNA episome tethering and replication, through the analysis of ITC (isothermal titration calorimetry) assays and EMSA (electrophoretic mobility shift assays), and also the analysis of 3D structural data of the USP7-LANA and USP7-LANA/DNA complexes, using X-ray crystallography.

2 Materials and methods

2.1 Constructs Design

The experimental design focused on constructs mostly involving $kLANA_{971-978}$ region (USP7 binding site) and mLANA C-Terminal region. Initially, two constructs of MHV-68 LANA (Murine Herpesvirus-68 LANA, or mLANA) were prepared, namely $mLANA_{140-314}$ (this construct contains less residues in the DBD, so it is expected to have less binding affinity, since it will have less interaction sites) and in $mLANA_{124-314}$ (containing more N-terminal residues in the DBD, including a predicted structural arm conformation) (**Figure 2.1**). The $mLANA_{140-314}$ construct has a molecular weight of 18.89 kDa, and corresponds to a total of 175 amino acid (aa) residues with an isoelectric point (pI) of 9.18 (so, it will show a basic behaviour), with an extinction coefficient of 10.68 M^{-1} cm^{-1} (280 nm) and a dimer molecular weight of 37.78 kDa. The first samples of these two mLANA constructs were previously expressed, purified and characterized.

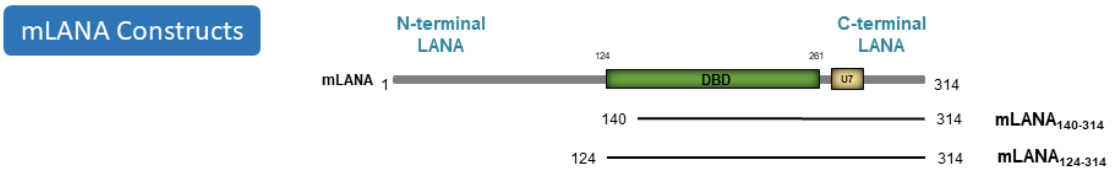


Figure 2. 1: Schematic diagram of mLANA and the initial constructs used. The segment starting at residue 140 includes less residues in the region necessary for DNA binding.

The kLANA constructs were designed according to different functions. The $kLANA_{974-983}$ was designed to have only the 10 residues in which the LANA's binding motif is included, while minimizing the number of residues that could hinder the structural interactions. It was designed for the structural determination of the binding region of the USP7-LANA complex. The $kLANA_{971-986}$ was designed for the same purpose, but with a bit longer sequence, to verify if it would favour the binding. The $kLANA_{971-1150}$ includes the USP7 binding region and the whole terminal repeats (TR) DNA binding and dimerization regions. It does not include all the DBD sequence, but the referred sites of interest are represented. The $kLANA_{930-1162}$ construct includes the whole DBD sequence, so it is expected to have higher affinity for the DNA than the previous constructs. The $kLANA_{4-32\text{^}888-1162}$ has a small segment of the beginning of the sequence at the N-terminal region together with the C-terminal region, including a portion of the glutamine-glutamate segment, before the DBD (Figure 2.2).

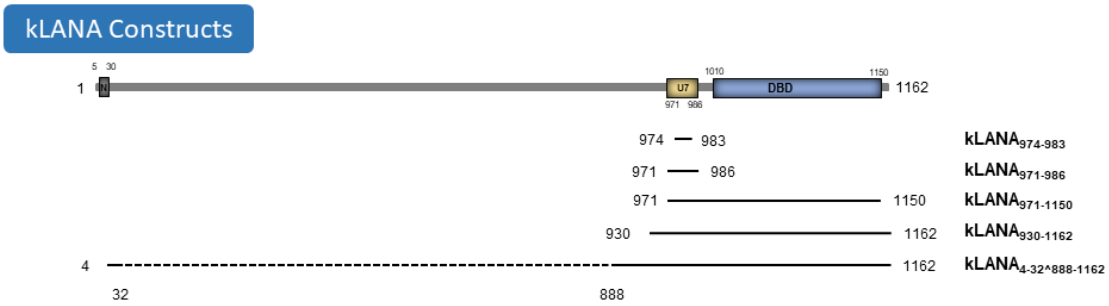


Figure 2. 2: Schematic diagram of kLANA and the respective constructs designed. The region between residues 5 and 30 (N) includes the nucleosome H2A/H2B binding region (residues 5 to 13) and the nuclear localization signal (residues 24 to 30). The USP7 binding region (U7, from 971 to 986) is presented as a segment apart from the DBD that is comprised between residues 1010 and 1150.

As for the USP7, two constructs were prepared: $USP7_{54-205}$ (it represents the N-terminal domain of the protein), and $USP7_{54-198}$ (missing the last 7 residues of the N-terminal domain) (**Figure 2.3**). The $USP7_{54-198}$ construct has a molecular weight of 17.11 kDa, corresponds to a total of 145 residues, its pI is 4.85 (so, it will have an acidic behaviour, interacting with the basic nature of the LANA's constructs) and it holds an extinction coefficient of 30.6 M^{-1} cm^{-1} (280 nm). The first samples of this two USP7 constructs were previously expressed and purified.

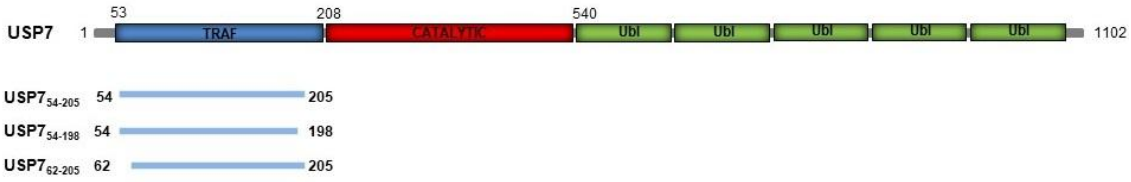


Figure 2. 3: Schematic diagram of USP7. The TRAF, catalytic and ubiquitin-like domains are indicated. The constructs are indicated as fragments of the TRAF domain. The fragments were designed to have less residues at the edges of the TRAF domain, to allow a better understanding of the type of interference that the missing residues might have on the overall binding interactions.

2.2 Cloning and Transformation

The Sequence and Ligation Independent Cloning (SLIC) technique (**Figure 2.4**) was used to clone all the previously described constructs. This technique holds the advantage of requiring only the use of the T4 DNA Polymerase, with 3'-5' exonuclease activity. The use of restriction endonucleases tends to create small extensions on the overlapping ends, which results in a very small number of hydrogen bonds between the few exposed nucleotides, so the vector and the insert don't hold together permanently, unless a DNA ligase enzyme is also used to join the nucleotides covalently to the phosphodiester backbone. With the SLIC Cloning technique, since the primers incorporate the necessary extensions, the complementarity of the extended (10 to 15 nucleotides, instead of the usual 4 to 6 from the endonuclease activity)

overlapping overhangs, resulting from the polymerase activity during the PCR amplification, ensures that the annealing of the insert to the vector occurs with a 90% rate of success by simple mixing of both DNA fragments (79).

The linearization of the vectors was made using the appropriate restriction enzymes, according to the inserts that they were meant to anneal with at the respective overlapping ends. The linearized vectors were confirmed by electrophoresis in a 1% agarose gel stained with GelGreen Nucleic Acid Stain (Biotium). The DNA inserts, specific for each construct, were obtained with the PCR (Polymerase Chain Reaction) technique. Primer sequences, containing recognition sites for the polymerase, were annealed with the template DNA and the required sequence (the DNA insert, which encodes for the protein to be expressed) is polymerized in several cycles of denaturation, annealing and extension, amplifying the sequence polymers in the solution (**Table 1.1**). The overlapping ends of the insert must be complementary to the overlapping ends of the linearized vector, to anneal and thus clone the insert into the vector. The primers were synthesized by SIGMA-ALDRICH.

Table 2. 1: **PCR cycling conditions.** Tm refers to the melting Temperature at which half of the DNA strands are in the single stranded (ssDNA) state.

Step	Temperature (°C)	Time	Cycles
Initial Denaturation	98°C	30 sec	1
Denaturation	98°C	10 sec	25-30
Annealing	Tm - 5°C	30 sec	
Extension	72°C	30 sec	
Final Extension	72°C	10 min	1

The gel purification of the inserts was carried out using the QIAquick PCR purification kit, from Qiagen. The linearized vectors were treated with Shrimp Alkaline Phosphatase (SAP) enzyme to dephosphorylate its phosphate from the phosphodiester backbone, which prevents the religation of the vector.

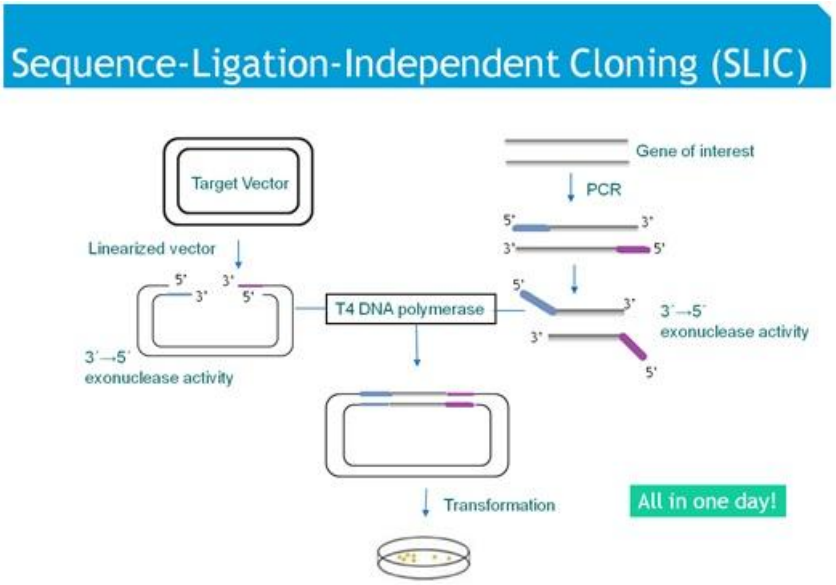


Figure 2. 4: Schematic illustration of the Sequence-Ligation-Independent Cloning (SLIC) technique.

The PCR reaction was performed in sample solutions (total volume of 50 µL) with the following components: 0.5 µL of Phusion DNA polymerase; 10 µL of 5×Phusion HF buffer; 200 µM of dNTPs; 20 mM of magnesium cloride; 20 ng of DNA template and 10 pmol of each primer. During the SLIC reaction, 40 ng of vector along with 80 ng of the DNA insert are incubated at room temperature for 10 minutes with T4 DNA polymerase and 1× NEB buffer 2.1 in a 10 µL solution (all components used were from New England Biolabs). The reaction product is placed on ice for 10 minutes to allow temperature decrease and consequent single-strand annealing. Table 1.2 describes the cloning details for the USP7 and kLANA constructs:

Table 2. 2: **Cloning details of USP7 and kLANA constructs.**

Protein Construct	**Vector [resistance]**	**RE site**	**Tag – protease site**	**SLIC Primers (Tm)**
$USP7_{62-205}$ (USP7 62)	pET 47b (+) [Kanamycin]	SacII/ XhoI	His – 3C Protease	FOR: 5'-GTCCTCTTTCAGGGACCCGGGgacaccagttggcgctcc -3' (Tm=89.8°C) REV: 5'-AGAGCGAGCTCTGCGGCCgcttaccacgcaactccatggggagc -3' (Tm=92.6°C)
$kLANA_{930-1162}$ (kLANA F)	pET 49b (+) [Kanamycin]	SacII/ XhoI	GST – 3C Protease	FOR: 5'-GTCCTCTTTCAGGGACCCGGGcaggagacggtggaagagcc -3' (Tm=89.5°C) REV: 5'-AGAGCGAGCTCTGCGGCCTTATGCGGCCGcttatgtcatttcc -3' (Tm=88.6.°C)

The transformation is done directly into DH5α *Escherichia coli* competent cells. 10 µL of the SLIC reaction product (vector with insert) were added to 100 µL of DH5α competent cells and the solution was incubated on ice for 30 minutes. The thermal shock, to allow the cells to take up the vectors, was delivered by keeping the solution for 2 minutes at 42°C, followed by 1 hour in an incubator shaker at 37°C and 180 rpm. Afterwards, 150 µL of the cell suspension was plated onto previously prepared LB agar plates, containing the appropriate antibiotics (50 µg/ml kanamycin, 100 µg/ml Ampicillin and/or 34 µg/ml of chloramphenicol). The remaining volume of the cell suspension was plated onto LB agar plates as well. The plates were incubated overnight at 37°C. On the next day, 3 colonies from each construct were chosen to be used for colony PCR confirmation, using GoTag DNA polymerase (Promega) and for inoculation in LB medium, with the appropriate antibiotics for overnight culture of each clone for mini-prep, which was prepared with the use of the Minipreps DNA Purification Systems kit (Wizards plus – Promega). To confirm the success of the cloning procedure, restriction enzyme double digestion of 200 ng of DNA, in a 10 µL reaction at 37°C for 30 minutes was performed, followed by an agarose gel (1%) electrophoresis.

2.3 Protein Expression

Once positive results arise, the vectors were transformed into an expression strain (BL21 Star pRARE 2 (DE3)) cell line, using 0.2 µL of the miniprep and 100 µL of competent cells. Same procedure was used as for the previous transformation into the cloning cell lines (30 minutes on ice, 2 minutes at 42°C, 1 hour at 37°C with shaking). To speed up the process and start expressing on the next day, a pre-inoculum was prepared: freshly grown colonies were used to inoculate 50 mL of ZY-0.8G medium (for growth to high densities with little or no induction of expression), supplemented with the appropriate antibiotics. The cells were grown in an incubator shaker at 37°C and 180 rpm. In the next day, 1 L of a preparation of ZYP-5052 auto-induction medium (auto-induction allows an activation signal to be triggered when a certain concentration threshold of the inducer is reached, leading to stimulation of altered gene expression; in this case, the lactose auto-induction system was combined with enzymatic glucose release to provide a constant feed of glucose, instead of using glycerol as a carbon source), supplemented with the appropriate antibiotics, was inoculated with 10 mL of the pre-inoculum (with a 0.07 starting Optical Density (OD)), and 500 mL of culture was grown in 2 L Flasks at 37°C and 180 rpm, for 2 hours, followed by an 18°C overnight incubation in a shaker with cooling system. The cells were harvested using ultra-centrifugation at 9000 rpm for 20 minutes at 4°C and the supernatant was discarded. The cell pellets were frozen and stored at -20°C.

2.4 Protein purification

Pellets were resuspended in lysis buffer (100 mL, containing 25 mM Na/K Phosphate, 500 mM NaCl, 10% (v/v) glycerol, 1 M NDSB 201, 10 mM DTT, pH 7.5) by vigorous stirring overnight. On the next day, the cell suspension was supplemented with lysozyme (0.25 mg/mL), 10 mM $MgCl_2$, 5 unit/mL OmniCleave (Epicentre) and one EDTA-free protease inhibitor cocktail tablet (Roche). The cells were left in incubation for 40 minutes at 4°C, and then the cell suspensions were lysed with the use of a Branson 450D sonicator (10% amplitude, cycles of 30 seconds of pulse on followed by 30 seconds off, for 2 minutes and 30 seconds. The cell lysates were then clarified by ultra-centrifugation at 19000 rpm, for 40 minutes at 4°C (82).

The cell lysis of USP7 pellets were done the same way as for LANA, except that NDSB 201 was not necessary. The adition of 20 mM imidazole to the buffer was used to minimize non-specific protein binding during the His-Trap purification.

2.4.1 mLANA and kLANA purifications

Protein purification was carried using an AKTA Explorer 10 FPLC System at room temperature (GE Healthcare). The supernatants were passed through 2×5 mL affinity columns (depending on the protein construct tag, for example GSTrap columns for the constructs with glutathione-S-transferase tag), connected in tandem, using a peristaltic P-1 pump (GE Healthcare), at 4°C. Unbound proteins and general compounds were washed with buffer A and the LANA proteins were eluted in buffer B (25 mM Na/K Phosphate, 500 mM NaCl, 10% (v/v) glycerol, 20 mM of compound depending on the column (for example, Glutathione for the GSTrap columns), pH 7.5), using a two steps gradient. The main fractions were pooled and 3C protease was added in a 1:100 molar ratio to the tagged LANA proteins with an addition of 2 mM DTT and 1 mM EDTA. The cleaved tags and non-specific DNA were removed from the LANA constructs passing through HiTrap Heparin HP columns (GE Healthcare), since heparin mimics DNA and the LANA constructs contain a DNA Binding Domain (DBD), so the LANA protein will be retained in such columns. The column was washed in washing buffer (25 mM Na/K Phosphate, 200 mM NaCl, pH 7.5). The protein eluted using a linear salt gradient (25 mM Na/K Phosphate, 0.2 - 2 M NaCl, pH 7.5). The fractions of the protein peak were pooled, concentrated and injected into Superdex 200 10/300 GL size exclusion column (GE Healthcare, which was pre-equilibrated with washing buffer (25 mM Na/K Phosphate, 250 mM NaCl, 10% (v/v) glycerol, pH 7.5 at 25°C). The purity of the proteins was monitored at all stages of the

purification process, running samples from the relevant fractions in 12.5% SDS-PAGE gel electrophoresis, at 200 V for 1 hour.

2.4.2 Purification of USP7-TRAF

Purification of USP7-TRAF was carried out by His-tag affinity purification using HisTrap HP columns (GE Healthcare) using the same protocol as for the LANA constructs, except for the Heparin affinity purification step, this was followed by the Hiload Superdex 16/600 S75 PG (124ml) size exclusion column using buffer containing 25 mM Na/K phosphate pH 7.5, 250 mM NaCl, 10% glycerol, 0.1 mM TCEP.

2.5 Preparation of LANA USP7 complexes

Complexes were prepared with a minimum molar ratio of 1:2.5 LANA/USP7. Since LANA forms dimers, the calculated molar amount is adjusted fto account for bindng to USP7, some excess was used to guarantee saturation. The complexation buffer contained 25 mM Na/K phosphate pH 7.5, 250 mM NaCl, 10% glycerol, 0.1 mM TCEP. On some occasions, desalting processes was needed, so the sample buffer would be altered to reduce the salt composition. In such cases, the buffer would be the same, but with half of the salt concentration (125 mM NaCl).

2.6 Crystallization and preliminary crystallographic analysis

Crystallization screenings were performed for $USP7_{54-198}$ $mLANA_{140-314}$ (32.14 mg/mL) and $USP7_{54-205}$ kLANA K16 (14.6 mg/mL) complexes (both complexes prepared with an excess of 2.5× of the LANA ligand, in 25 mM Na/K phosphate pH 7.5 and 500 mM NaCl protein solution), using the BCS HT-96 MD1-105 and the LMB MD1-98 Screening sets (Molecular Dimensions) on the Crystallization robot Mosquito LCP (TTP Labtech), with the sitting-drop vapor-diffusion method. Further optimizations, mainly deriving from the results of the screenings, were made using both the sitting-drop and the hanging-drop vapor-diffusion method. The crystallization drops were imaged regularly, using a Minstrel DT UV imager (Rigaku).

2.7 Electrophoretic Mobility Shift Assay

In order to analyse the way USP7 enhances the binding of LANA to DNA, several Electrophoretic Mobility Shift Assays (EMSA) were performed. Several LANA Binding Sites (LBS) DNA constructs were used, combining 3 different binding sites (site 1 and 2 are considered necessary for LANA's DNA tethering, site 3 is considered necessary for LANA replication activity). The DNA constructs used are incubated with the protein complexes and the shift of electrophoretic mobility is evaluated, to determine if binding occured. For this reason, the DNA constructs are usually named DNA probes. For visualization, the DNA probes are labelled with fluorescein (Flc) at the 5'-end (Flc-DNA synthesized by SIGMA-ALDRICH). The probes sequences were:

kLBS1-2 (5'-TTTGACGCCGCCGGGGCCTGCGGCGCCTCCCGCCCGGGCATGGGGCCGC-3');

kLBS1-3 (5'-TTTTCCCGCCCGGGCATGGGGCCGCGCGCCGCCTCAGGGCCCGGCGCGG-3')

corresponding to the LBS1-2 and LBS1-3 sites of the terminal repeat (TR) DNA sequences of KSHV. For control verification, LANA incubations with both USP7 and BSA (negative control) were used. Protein samples were incubated with 0.2 µM of DNA probe at room temperature, for 30 minutes. Each incubation was made in 4 µL of 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). The gel was run in TBE buffer for 1 hour at 150 V at 4°C. All results were scanned with a Fuji TLA-5100 scanner.

2.8 Isothermal Titration Calorimetry (ITC)

Titrations were performed with with a MicroCal^{TM} iTC200 Isothermal Titration Calorimeter (Malvern) at 25°C. Protein absorbance (280 nm) was measured by NanoDrop (NanoDrop Technologies) and their concentrations were determined using their respective extinction coefficients. Twenty-three 1.5 µL injections of titrant ($USP7_{TRAF}$) were titrated into $LANA_{DBD}$ protein solution. Data were corrected for nonspecific heat and analysed using MicroCal Origin® 7.0 software, using a one-site binding model. Nitpic (MBR software, integrates peaks and normalizes baseline), Sedphat (NIH software, allows graph analysis) were used for fitting and calculating thermodynamic parameters, while Gussi (MBR software, illustrates the output of Sedphat) was used to generate the graphical data (80). Proteins were freshly prepared for the ITC assays.

3 Results

3.1 Cloning

USP7 N-terminal construct $USP7_{62-205}$ was cloned into pET-47b (+) vector, which includes a N-terminal 6xHis-Tag and 3C protease cleavage site after the tag and is kanamycin resistant. LANA DNA binding domain construct $kLANA_{930-1162}$ was cloned into pET-49b (+) vector, which has a N-terminal GST-6xHis-Tag with a 3C protease cleavage site after the tag. pET-47b (+) and pET-49b (+) vector maps are illustrated in Figure S1A and S1D, respectively. The inserts were generated by PCR and the vectors were linearized by digestion with SacII and XhoI restriction enzymes and treated with SAP, to prevent re-ligation. The PCR products and the linearized vectors were ligated with SLIC cloning. To confirm if the reaction of ligation was efficient, the same restriction enzymes were utilized with the product of the SLIC cloning. The linear vectors were visualized using a 1% agarose gel electrophoresis, to confirm the insert presence (**Figure 3.1**).

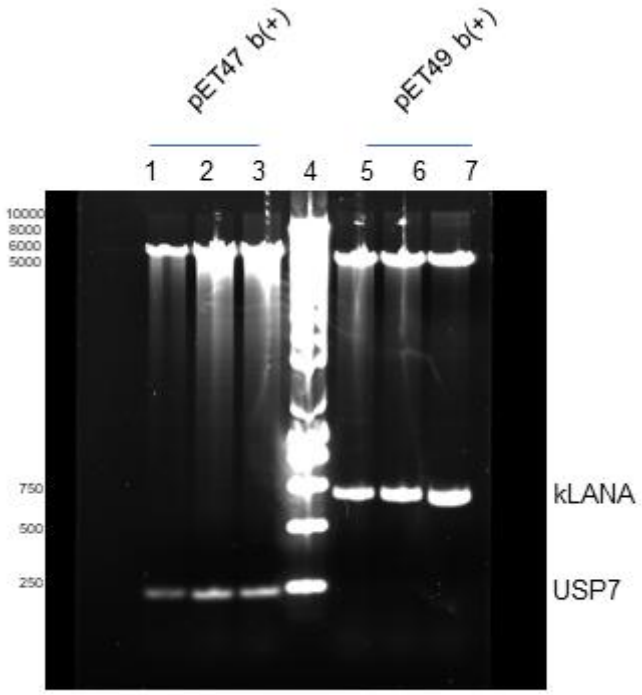


Figure 3. 1: Restriction digestion of the cloning of $USP7_{62-205}$ and $kLANA_{930-1162}$. Lanes 1-3 show clone validation by SacII/XhoI digest of different isolated $USP7_{62-205}$ clones. Lane 4 is a 1kb ladder marker, Lanes 5-7 show clone validation by SacII/XhoI digest of $kLANA_{930-1162}$ isolated clones.

The gel demonstrates that the ligation was efficient, since the presence of the insert is visible in all experimental procedures.

3.2 Expression and purification

$USP7_{62-205}$ and $kLANA_{930-1162}$ were expressed successfully, as were the other protein truncations. Results were confirmed during purification. $USP7_{NTD}$ $mLANA_{DBD}$ complex was successfully purified (**Figure 3.2**), to be used in crystallization, EMSA, ITC and limited proteolysis assays.

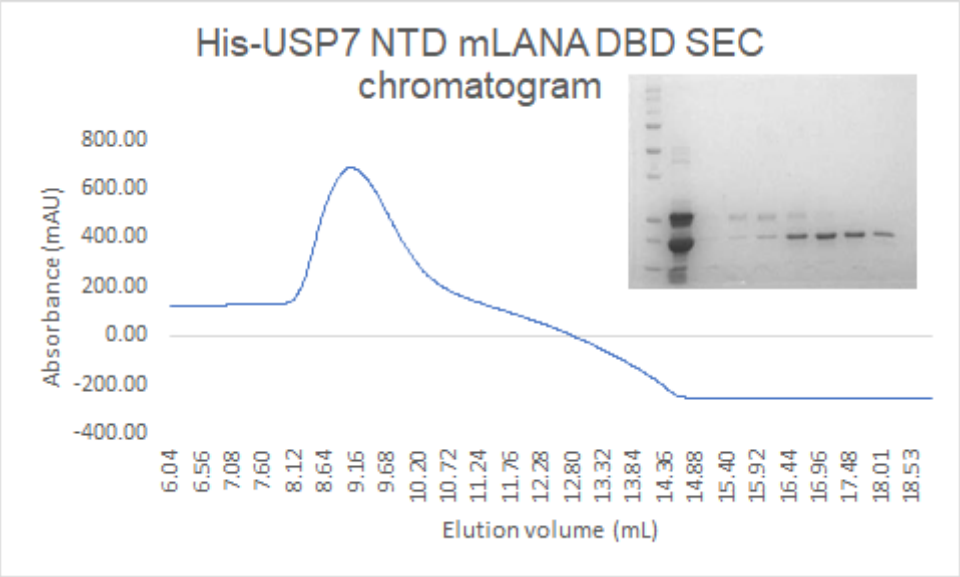


Figure 3. 2: Size exclusion analysis of $USP7_{NTD}$ $mLANA_{DBD}$ complex protein. The complex sample was loaded into a Superdex 75 10/300 GL column and the elution profile was analysed. A 12.5% SDS-PAGE gel was run (u).

3.3 Crystallization of LANA-USP7

The only precipitant condition, in the initial screening made, that produced crystals was the condition 0.2 M Ammonium sulfate, 0.05 M Magnesium sulfate heptahydrate, 0.1 M Bicine pH 9.0, 20% v/v PEG Smear solution, only for the $USP7_{54-205}$ kLANA K16 complex. Subsequent optimizations were tested in several crystallization assays. Crystals from that condition and from similar optimized conditions were tested for X-ray diffraction data collection at 100k, beamline ID30A, at the European Synchrotron Radiation Facility, Grenoble France, proving to have no diffractable quality (**Figure 3.3 (left)**). A crystal from a modified precipitant condition (0.1 Tris pH 8.5, 30% PEG4000, 0.6 M Lithium Sulfate) produced data sets, showing 4 of the 16 residues (corresponding to the known motif, mentioned in the introduction) from $kLANA_{971-986}$ ligand involved in direct interactions with $USP7_{198}$ (**Figures 3.3 (right) to 3.5**).

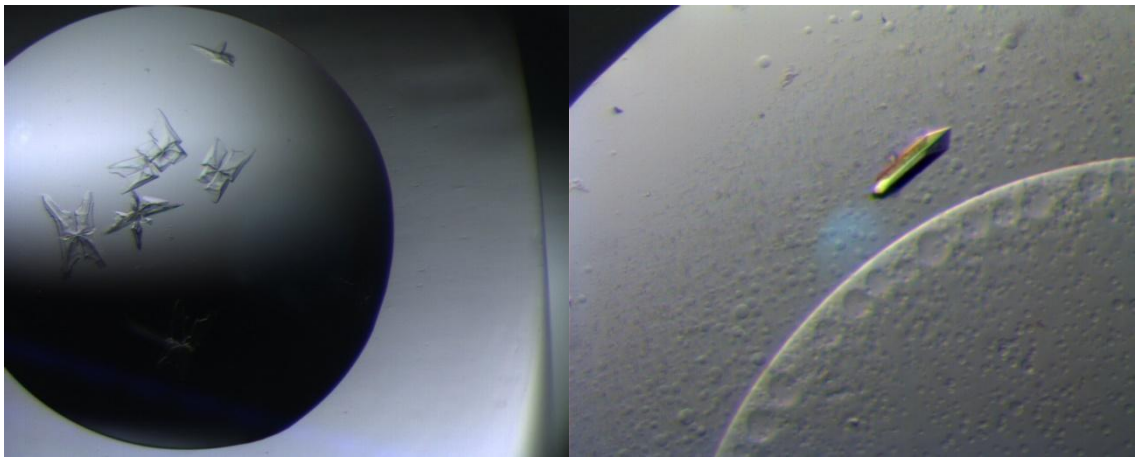


Figure 3. 3: (**Left)** USP7$_{54-198}$ kLANA$_{971-986}$ crystals from the BCS screen, with non diffractable quality. The precipitant solution consisted of 0.2 M Ammonium sulfate, 0.05 M Magnesium sulfate heptahydrate, 0.1 M Bicine pH 9.0, 20% v/v PEG Smear solution. **(Right)** USP7$_{54-205}$ kLANA$_{971-986}$ crystal, diffracted at 1.6 Å resolution. The precipitant solution consisted of 0.1 M Tris-HCl pH 8.5, 30% PEG4000, 0.6 M Lithium Sulfate.

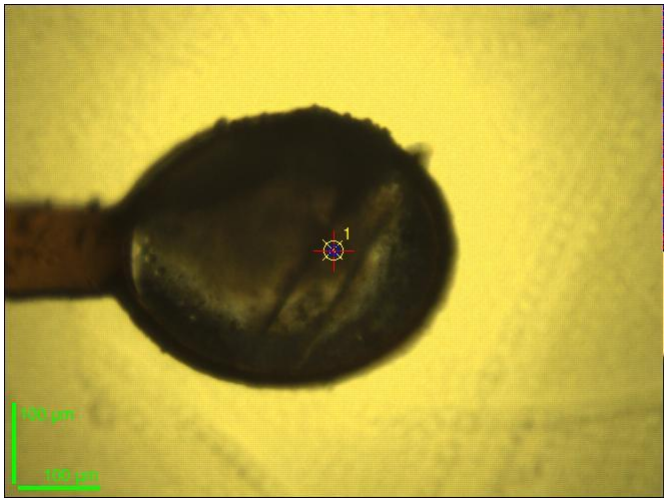


Figure 3. 4: Crystal from figure 3.3 (right) mounted on a cryo-loop and centred automatically as detected by MX-cube software, at beamline ID30A, at the European Synchrotron Radiation Facility, Grenoble France.

The ESRF MX beamline control graphical user interface (GUI) MX-cube analysed and centred the crystal. Clear sharp single lattice spots indicate the presence of a single diffractable crystal, with no evidence of salt diffraction which give rise to very strong spots. There are, however, indications of ice, because some rings are visible in the diffraction patterns (**Figure 3.5**). The crystal was automatically indexed, and diffraction data collection strategy was determined by EDNA software.

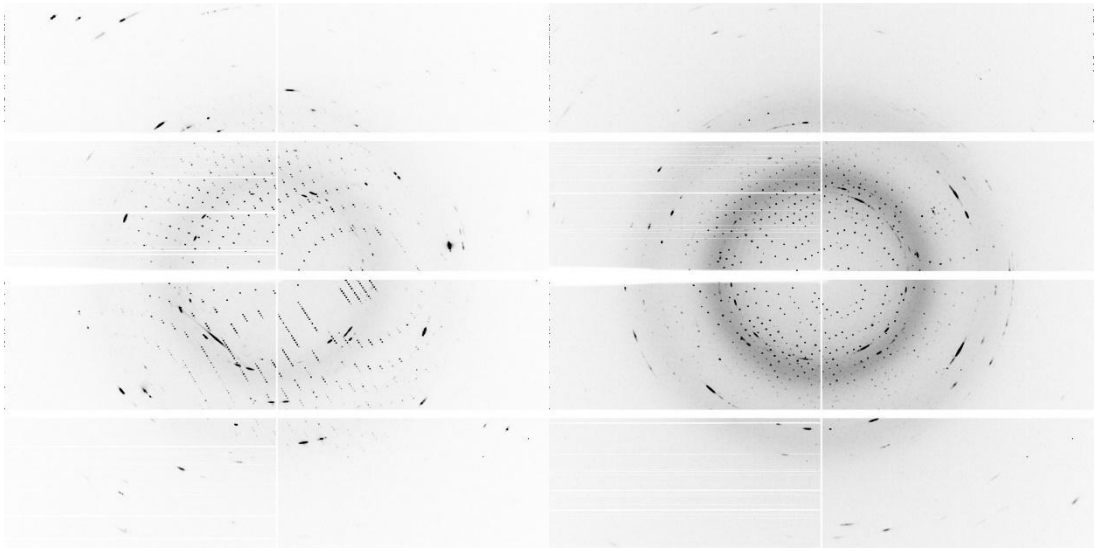


Figure 3. 5: Diffraction pattern from the diffraction experiment of the crystal from figure 3.3 (right). **Left:** X-ray screening of crystal diffraction. **Right:** Second diffraction image 45^0 apart from diffraction pattern on the left.

The unit cell dimensions of the crystal were a= 50.7628, b=50.7628, c=130.347 (Å); α=90, β=90, γ=90 (°). Space group was P 4_1 2_1 2 (Number 92, of the tetragonal crystal system). Resolution range was 33 - 1.48 Å. Work is currently ongoing to complete data refinement and analysis of the crystal structure of the $USP7_{TRAF}$-LANA complex.

Additional crystallization attempts where made, using several strategies, such as streak seeding, macro seeding and micro seeding, salt adding after drop equilibration and maintaining the drops at different temperatures (8^0 C, 14^0 C and 20^0 C). The most significant results from those trials are presented next (**Figures 3.6 - 3.11**).

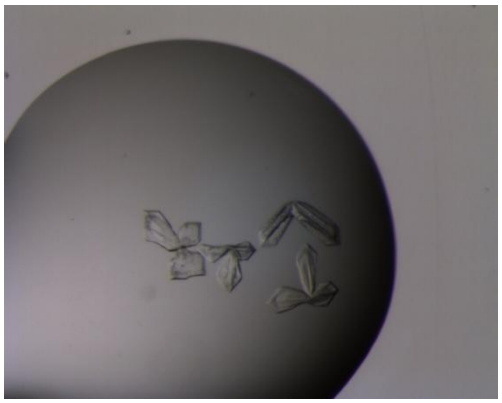


Figure 3. 6: Optimization from PEG Smear screening condition.

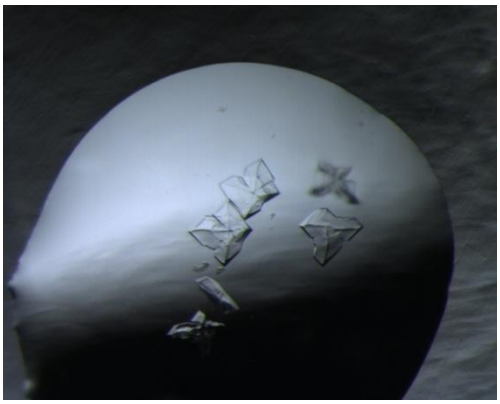


Figure 3. 7: Further PEG combinations optimized from PEG Smear.

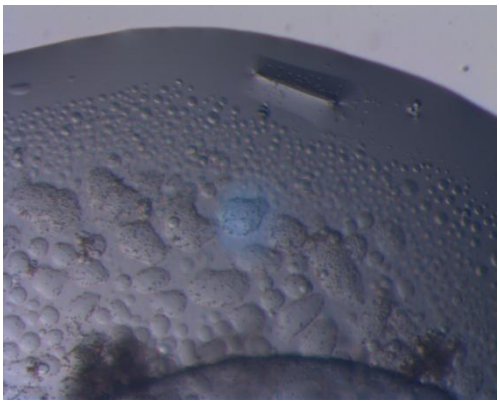


Figure 3. 8: Crystal obtained in the 8^0 C crystallization room.

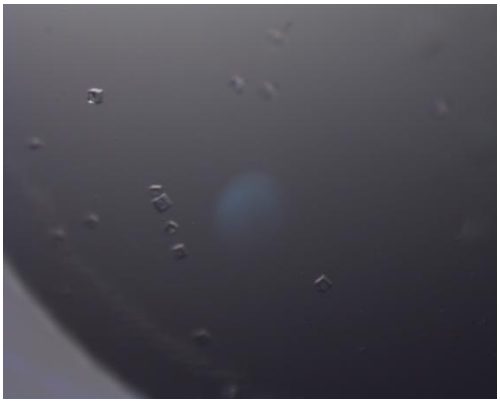


Figure 3. 9: Altered salt concentrations.

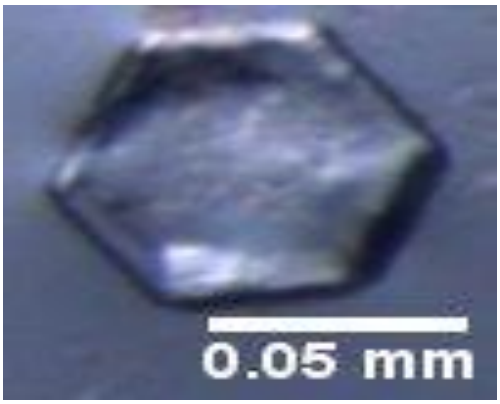


Figure 3. 10: Salt adding after drop equilibration.

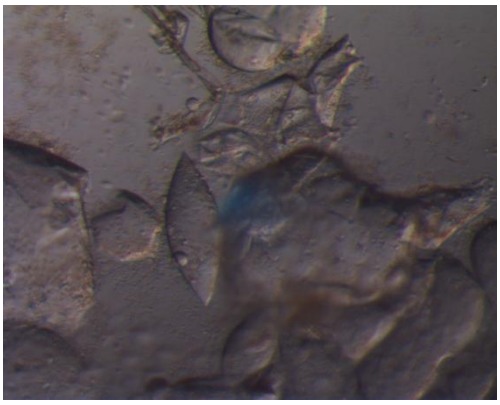


Figure 3. 11: Extreme precipitation condition, 8^0 C room.

Unfortunately, these crystals did not produce the desired results. Crystals from figures 3.6 and 3.7 did not diffracted and crystals from figures 3.8 and 3.11 couldn't even be collected because their condition deteriorated during the procedure. The crystals from figures 3.9 and 3.10 reveal to be just salt crystals.

3.4 Electrophoretic Mobility Shift Assays

3.4.1 Preliminary assays

Several herpes viral proteins (HSV-1 ICP0, EBV EBNA-1) have been demonstrated to interact with USP7 and such interactions have shown some degree of enhancement of DNA binding affinity. Therefore, to verify USP7's influence on DNA binding affinity by LANA, preliminary assays were performed, using $USP7_{54-205}$ complexed with $kLANA_{4-32\text{^}888-1162}$ (**Figure 3.12**).

kLBS 1-2					kLBS 1-3				
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
0.2 µM kLBS 1-2	0.1 µM USP7 NTD 0.2 µM kLBS 1-2	0.1 µM BSA 0.1 µM kLANA E 0.2 µM kLBS 1-2	0.1 µM Complex (kLANA + USP7) 0.2 µM kLBS 1-2	0.4 µM Complex (kLANA + USP7) 0.2 µM kLBS 1-2	0.2 µM kLBS 1-3	0.1 µM USP7 NTD 0.2 µM kLBS 1-3	0.1 µM BSA 0.1 µM kLANA E 0.2 µM kLBS 1-3	0.1 µM Complex (kLANA + USP7) 0.2 µM kLBS 1-3	0.4 µM Complex (kLANA + USP7) 0.2 µM kLBS 1-3

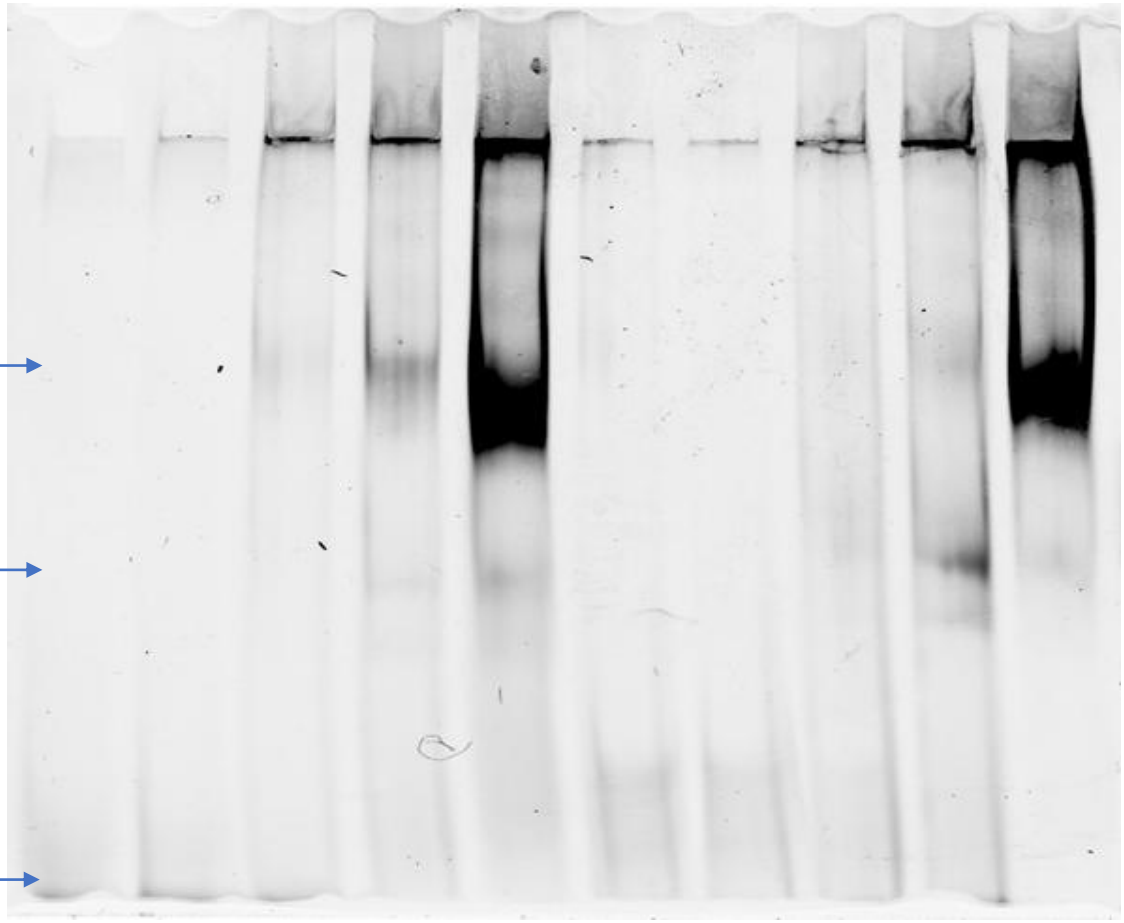


Figure 3. 12: EMSA of titrations of 0.1 µM and 0.4 µM of kLANA USP7 complex with 0.2 µM of DNA recognition sites, also having equimolar presence of BSA with kLANA (0.1 µM) as a negative control and 0.1 µM USP7 alone titrations.

This preliminary assay shows a positive influence of USP7 in the complex for enhancing LANA's LBS 1-2 DNA binding affinity, whereas in the case of kLBS 1-3, USP7 seems to inhibit LANA's DNA binding, at least at lower concentrations of the USP7-LANA complex. These results are in line with what was previously suggested in studies with EBNA-1 (15) and USP7-LANA (1).

3.4.2 mLANA model assays

To determine the USP7-mLANA-DNA ternary complex binding affinities, $USP7_{62-205}$ complexed with $mLANA_{124-314}$ and mLBS 1-2 DNA were used, in two separate assays: firstly, USP7 and mLANA were incubated overnight prior to binding it to the DNA and alternatively, mLANA and DNA were incubated prior to binding to USP7 (**Figure 3.13**).

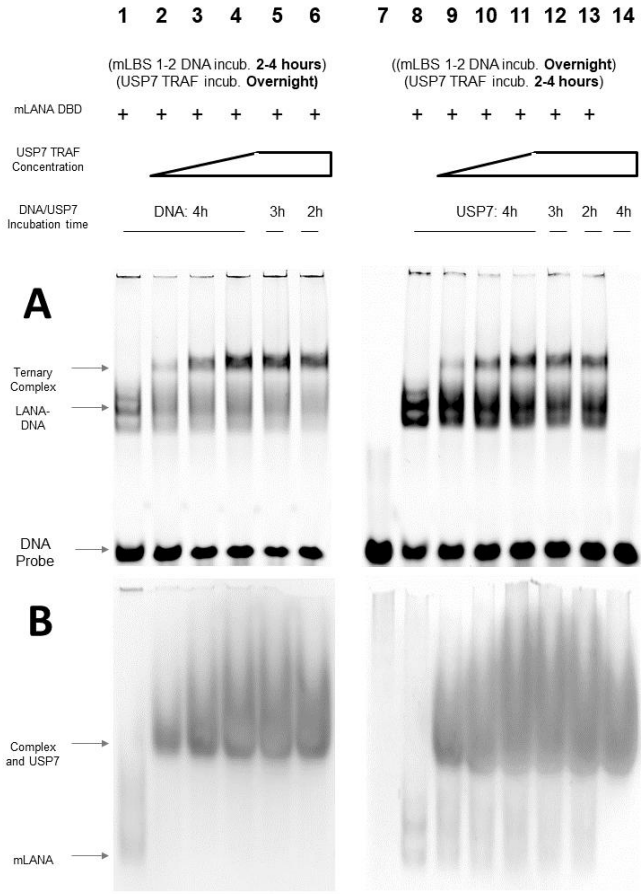


Figure 3. 13: Analysis of DNA binding mLANA DBD – USP7 TRAF ternary complex. A clear shift between the ternary LANA/USP7 DNA complex and the LANA-DNA complex is visible. (**A):** Electrophoretic mobility shift assays (EMSA). Lane 1 corresponds to a mLANA control; Lanes 2 to 6, DNA binds to previously formed LANA-USP7. Lane 7 shows a DNA control; Lane 8 shows a mLANA control; Lanes 9-13, USP7 TRAF binds to previously formed LANA-DNA, and lane 14 shows an USP7 control. In all the samples DNA was labelled with fluorescein. **(B):** Native gels from the EMSA assays stained with coomassie blue.

The ternary complex from the EMSA assay clearly shows that the mLANA-USP7 assembly prior to DNA binding results in a more stable complex. This is demonstrated by the difference of intensity between the complex bands and the corresponding DNA probe bands. In the first assay (lanes 1 to 6), USP7 was incubated with mLANA overnight, while the DNA was added 2 to 4 hours before the run (as indicated). A clear shift can be observed between the band of DNA binding mLANA (lane 1) and the band binding mLANA-USP7 complex (lanes 2-6). Also, its clear that the intensity of the DNA binding complex is proportional to the concentration of USP7, revealing a higher affinity of DNA to mLANA complexed with USP7. On the next assay (lanes 7 to 14), DNA was incubated overnight with mLANA and USP7 was added a few hours before the run. The difference of intensities between the lanes 1-6 and the lanes 7-14 is due to a much higher incubation period with fluorescein labelled DNA in the case of the lanes 1-6, allowing a higher ratio of binding. The coomassie blue stained native gels shows the mLANA band is quite distinct from the complex band. It is interesting to notice that the USP7 band coincides with the complex band, revealing that in the native state, the LANA-USP7 complex electrophoretic mobility is mainly determined by the USP7 characteristics.

3.5 Isothermal Titration Calorimetry (ITC) analysis

3.5.1 $USP7_{TRAF}$ titration on $LANA_{DBD}$

ITC assays were performed to determine the thermodynamic parameters of the LANA-USP7 binding, measuring the energetics of the binding process as it occurs throughout the titration. 153 µM of His-$USP7_{62-205}$ were titrated into the cell, containing 10 µM of $mLANA_{124-314}$ at a cell temperature of 25°C. Both proteins were kept in the same buffer (25 mM Na/K phosphate pH 7.0, 250 mM NaCl, 5% glycerol). In the first assay, a complete saturation was not achieved, because the concentration in the syringe wasn't enough. Furthermore, a bubble most have entered the cell, causing a disturbance in the titration, so the data was unable to be fitted for a one-site binding model and was not used for further analysis. (**Figure 3.14**).

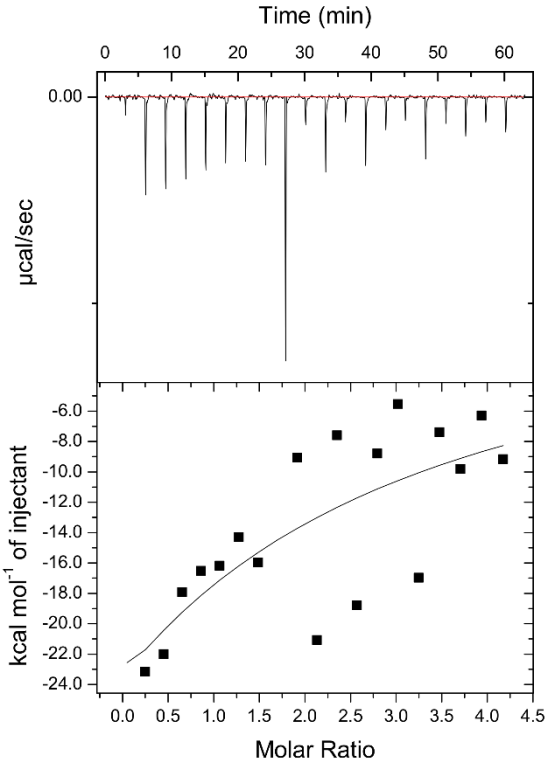


Figure 3. 14: First ITC assay, a bubble on the cell prevented the titration correct reading. Also, the concentrations did not lead to saturation.

A second ITC assay was prepared with optimized concentrations. this time, 361 µM of His-USP7$_{62-205}$ were titrated into the cell, containing 20 µM of mLANA$_{124-314}$ at a cell temperature of 25°C. Both proteins were kept in the same buffer (25 mM Na/K phosphate pH 7.0, 250 mM NaCl, 5% glycerol), the buffer was freshly made. This time, a saturation curve was achieved, and the titration allowed correct readings all across (**Figure 3.15**). The curve was fitted for a one-site binding model.

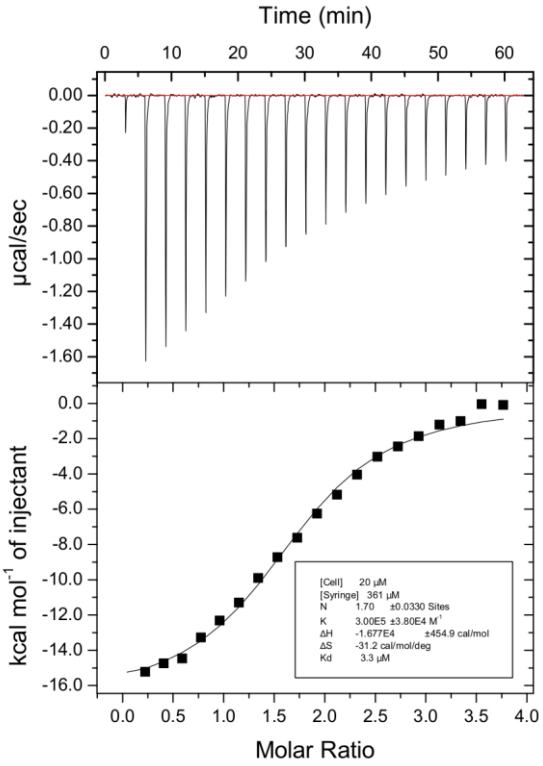


Figure 3. 15: Analysis of His-USP7$_{62-205}$ binding by mLANA$_{124-314}$. The upper panel isotherms indicate the USP7 binding raw data and the heat released versus time. The lower curve is obtained after integration of the peak intensities plotted against the protein/protein molar ratio in the calorimeter cell. The solid lines correspond to the best fit using a one-site binding model. The thermodynamic parameters describing the fit for the USP7-mLANA titration are presented on the table on the lower right corner. Titrations were performed with a series of 20 injections, each with 2.0 µL of USP7, to the LANA sample.

The titration shows an exothermic reaction, since heat was being released upon the binding. The USP7-mLANA binding signature has ΔH = -70.2 kJ/mol and $-T\Delta S$ = 38.9 kJ/mol, with a K_D of 3.3 µM. The favourable ΔH proves that USP7 binding to mLANA is driven by hydrogen-bonding and van der Waals interactions. The gained Gibbs free energy of binding is $\Delta G = \Delta H - T\Delta S$ = -31.2 kJ/mol.

3.6 Limited proteolysis

3.6.1 Preliminary small-scale assay

To obtain a more crystallizable USP7-LANA complex, eliminating disordered regions that are not relevant for the binding, limited proteolysis assays were conducted. This technique uses protease enzymes to trim the complex, removing mostly disordered loops that are not stabilized by the binding, thus not interfering with the interacting region, while creating a much more stable complex. A preliminary small-scale assay, to determine the optimal time of digestion, was the first step (**Figure 3.16**).

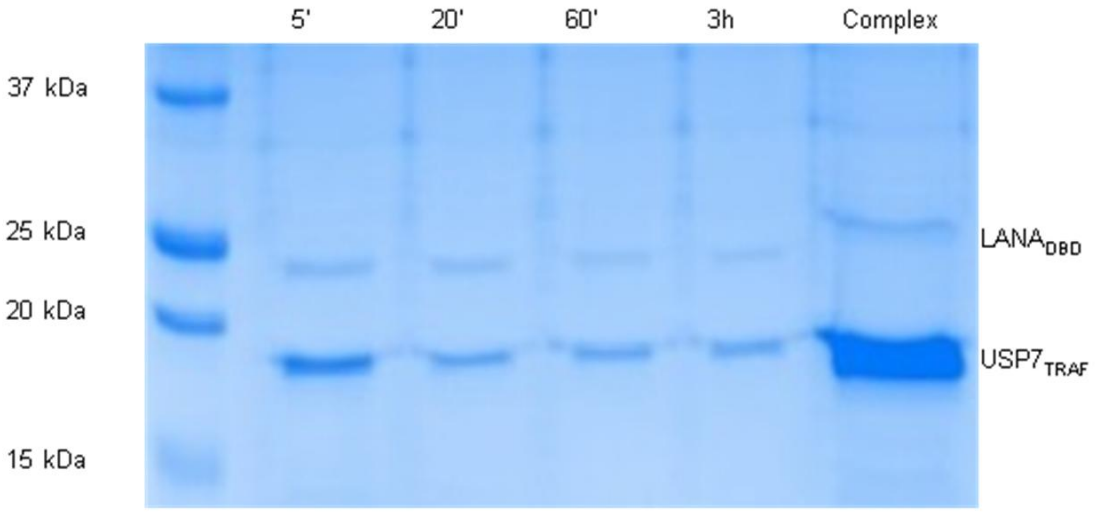


Figure 3. 16: Limited proteolysis of USP7-mLANA peptide complex, using 1/50 dilution of 0.1 mg/mL of trypsin. The last well contains undigested complex sample.

A clear cut on LANA can be observed, reducing its size by approximately 1.6 kDa, without cutting USP7, so the complex has remained stable. All the digestion periods, from 5 minutes to 3 hours, seem to have been efficient, so a 10-minute digestion period was decided for further assays.

3.6.2 Scale-up of the limited proteolysis assays

Since crystallization assays require a considerable amount of the proteins, a scale-up of the limited proteolysis was initiated. The mLANA$_{124-314}$ His-USP7$_{NTD}$ complex was purified in the Superdex S75 10/300 (24ml) size exclusion column (**Figure 3.17**).

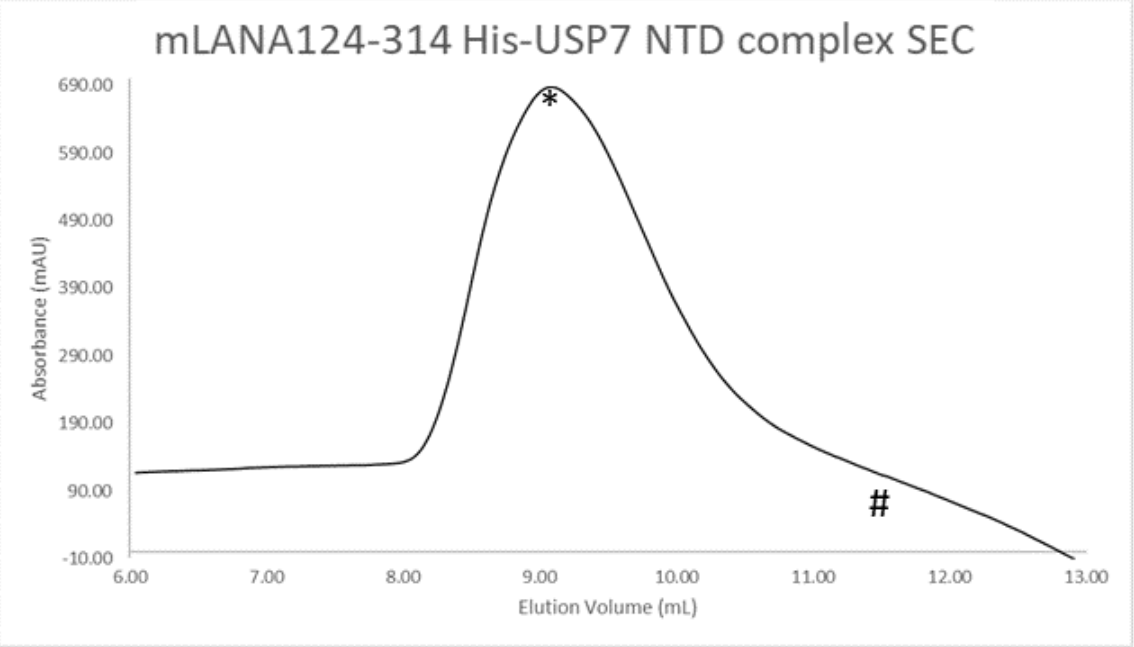


Figure 3. 17: mLANA$_{124-314}$ His-USP7$_{NTD}$ complex purification chromatogram.

The symbols on the graph signal the relative position of elution of the complex (*) and a mixture of incomplete complex and LANA dimer (#) (**Figure 3.18**).

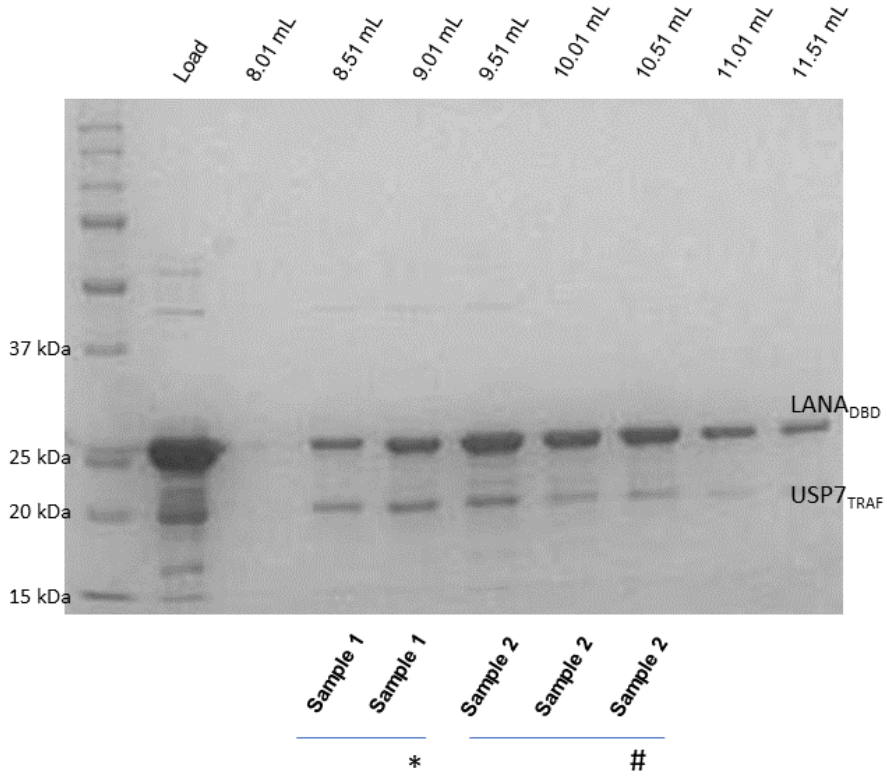


Figure 3. 18: 15% SDS-PAGE gel from the complex elution peak.

As it is shown, 2 separate samples were extracted, since sample 1 presents equimolar bands (which means that it corresponds to the elution of the complex) and sample 2 does not. Sample 2 was later treated with a desalination process, to allow the proteins to properly bind and thus form the complex. After some optimizing steps, to determine the needed quantities of protease enzyme in relation to the total complex protein amount, the final assay was performed using 0.1 µg of Trypsin in a total of 340 µg of complex, for a period of 20 minutes, followed by the injection of the sample into the SEC column (**Figure 3.19**). The resulting peaks were analysed in 15% SDS-PAGE gel (**Figure 3.20**).

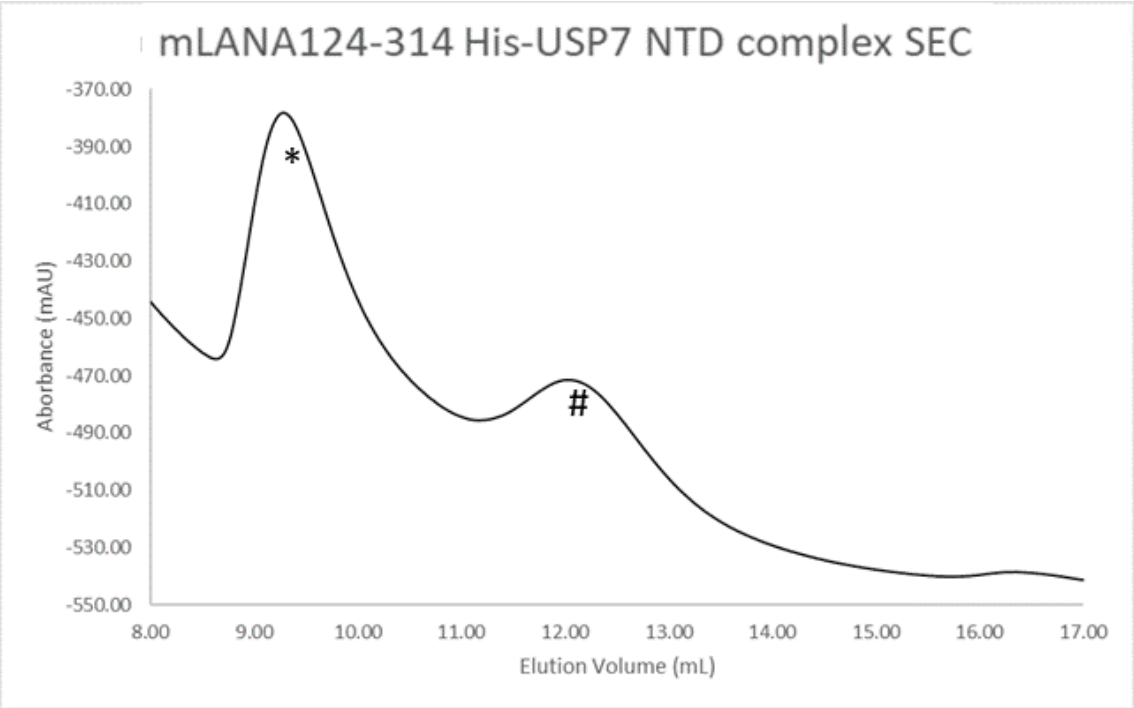


Figure 3. 19: Size exclusion chromatography of the mLANA$_{123-314}$ His-USP7$_{NTD}$ complex following the optimized limited proteolysis.

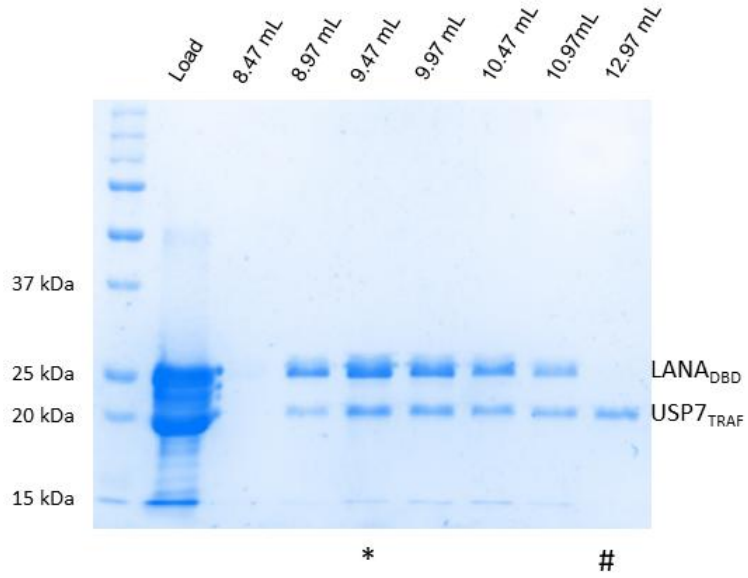


Figure 3. 20: 15% SDS-PAGE gel analysis of the limited proteolysis complex elution peaks.

The results show that it is still partial, so to obtain a complex sample fully digested, a higher quantity of trypsin should be used, so 0.4 µg of Trypsin in a total of 340 µg of complex, for a period of 20 minutes, should constitute a more optimized experimental condition.

4 Discussion and conclusions

USP7 is targeted by LANA, that uses USP7 N-terminal TRAF (tumour necrosis factor [TNF] receptor-associated factor) domain. A short sequence of residues (amino acids [aa] 971 to 986), contained in the C-terminal domain of LANA, interacts directly with USP7 N-terminal TRAF. This sequence has strong similarities with the USP7 binding site of the Epstein-Barr virus (EBV) EBNA-1 protein (1). The location of kLANA's USP7 binding motif is at the N-terminus of its DBD, while in the case of mLANA, the USP7 binding motif ^{272}PSTS^{275} is located at the C-terminus of the DBD, differently than kLANA and EBNA-1.

In this work, i aimed to try to understand the mechanism by which LANA recruits the de-ubiquitinase (DUB) USP7, for its own trDNA binding purposes. A previous study revealed that USP7 modulates the replication of KSHV latent episomal DNA (1), so i tried to evaluate the impact of USP7 in LANA DNA binding activity. I have concluded that the EMSAs results for both kLANA and mLANA indicate a clear influence of USP7 over LANA's DNA binding. I was also able to perform an ITC analysis of $USP7_{TRAF}$ binding to the more stable mLANA DBD dimer.

4.1 Biochemical characterization of USP7-LANA complex

Purified $USP7_{TRAF}$ monomer conformation is 18.9 kDa in size, while $LANA_{DBD}$ dimer is sized 43 kDa. Since the complex, generated both by co-expression and complexation of its individual components, eluted at 9 mL in the size-exclusion chromatography (SEC) column (**Figures 3.2 and 3.22**), the calculated value was determined to be around 80.9 kDa size for the complex. This interesting observation suggests that each LANA dimer recruits two $USP7_{TRAF}$ monomers on opposing sides.

4.2 Effect of USP7 on DNA binding by LANA *in vitro*

Earlier studies have shown that USP7 or $USP7_{NTD}$ binding to EBV EBNA-1 increases its DNA binding activity by 20 or 15-fold, respectively (75). In the case of LANA, deleting USP7 interaction region (aa residues 971-986) caused no change in its DNA binding activity (1). To clarify the influence that USP7 has on LANA's DNA binding activity *in vitro*, KSHV an MHV-68 LANA (kLANA and mLANA, respectively) proteins with both USP7 and DNA binding regions

included in their structures were incubated with Fluorescein labelled LBS trDNA in the presence of an excess of $USP7_{NTD}$ or BSA (negative control). EMSAs results for both kLANA and mLANA indicated a clear influence of USP7 on LANA's DNA binding activity, although it remains to be seen what, if any, consequence this may have on virus infection. Studies are underway to explore this interaction with an *in vivo* model of infection.

As the EMSA results suggest, USP7 increases LANA's affinity to bind specifically LBS 1-2 DNA, while it seems to decrease somewhat its affinity for LBS 1-3 DNA. Since LBS 1-2 DNA is directly involved in the process of viral DNA tethering into the host cell's chromatin and LBS 1-3 DNA relates to the LANA's episome replication functions, the results indicate that USP7 may play multiple roles, while complexed with LANA, on its DNA binding activities. The USP7-LANA complex will offer an optimal structure for the tethering function, while it may regulate LANA's role in the replication of the viral DNA episome, during the latent or lytic cycle of infection. From a survival and immortalization of the infected cells point of view, aside from the fact that LANA will compete with p53 for the same binding location on USP7 and thus contribute to destabilization of p53, it also seems that USP7 will favour LANA's capacity to hide the presence of the virus in the cell, during the latent cycle.

4.3 LANA interacts directly with $USP7_{TRAF}$

The titration of $mLANA_{DBD}$ with $USP7_{TRAF}$ was analysed by ITC. Their binding is governed by exothermic heat. The dissociation constant (K_D) for LANA binding to $USP7_{TRAF}$ is 3.3 µM and a n-values of 0.99 for LANA monomer indicated a 1:1 binding stoichiometry. The binding is stronger than host binding proteins p53 (8 µM). Interestingly, the interaction of LANA with $USP7_{NTD}$ through the conserved motif (P/AxxS) is predominantly hydrophobic, and the same is true for other USP7 binders, such as p53, HDM2 and EBNA-1, because the $USP7_{NTD}$ binding region contains a hydrophobic core, centred in residues Trp165-Gly166-Phe167 from strand β7. The residues Trp and Phe form an aromatic clamp for the ^{271}QPSTS^{275} motif region of mLANA. However, the ITC results in this work clearly demonstrate a spontaneous enthalpy driven binding between $USP7_{TRAF}$ and $LANA_{DBD}$, promoted by hydrogen bond and Van der Waals interactions. This might be a strong indicator of further interactions outside of the known motif.

The preliminary crystallographic results confirm the previously suggested binding motif P/AxxS and no other residues have been identified as relevant for the binding of the peptide used. This may be a consequence of the reduced size of the peptide motifs used, rather then the complete LANA DBD, since it can offer a more stable binding interface for LANA USP7 interaction.

The nanomolar affinity of LANA C-terminal DBD for LBS 1-2 DNA is lower, but further addition of another low-affinity site (LBS 3) increases the affinity to picomolar for LBA 2-1-3

DNA (19). LANA's affinity for the trDNA increases with its cooperativity in binding. In the case of EBNA-1, it has also shown to oligomerize and its nanomolar affinity for the dyad symmetry sequence at the origin of replication is lower (83). The location of the USP7 binding motif in EBNA-1 might be the reason why it may stimulate its binding affinity, since it is adjacent to its DNA binding site. However, in the case of MHV-68 LANA homologue, the USP7 binding motif is located further away from the DNA binding site (**Figure 2.1**). Therefore, it is still unclear whether USP7 is really necessary for these latency proteins DNA binding stimulation.

Alternatively, LANA might adopt a simple strategy of interference in the normal regulation of USP7 over the p53-HDM2 pathway, by competitive binding of the P/AxxS motif. ITC results suggest the LANA may in fact be able to disrupt both USP7-p53 and USP7-HDM2 interactions, thus altering p53 levels, which could lead to increased cell survival and proliferation, as it has been observed in most forms of KSHV-latent infection. On the other side, if LANA sequesters USP7, levels of HDM2 may decrease, increasing p53 levels in turn, unleashing its deadly downstream apoptotic effects on the cell. The subtle interplay between LANA and USP7 may be determinant to either cell-death, lytic infection or cellular immortalization in long-lived memory B cells during latent infection.

The USP7 binding motif closeness to the DNA binding interface of LANA can be a cause of interference in the assembly of the origin of replication, imparing latent replication. Moreover, the proliferating cell nuclear antigen (PCNA) is a DNA clamp that plays an essential role in KSHV replication (84). PCNA ubiquitination is crucial for the maintenance of genomic stability during DNA replication, and it is also deubiquitinated by USP7, regulating DNA-damage-induced mutagenesis. LANA recruits PCNA to the viral replication centre, so the LANA-USP7 complex may eventually also have some function in the event of the PCNA post-translational modifications being affected during KSHV infection.

4.4 Work ahead

Further studies may reveal more data on the structure of the ternary complexes (USP7-LANA/DNA), as well as the thermodynamic parameters of the binary complex (USP7-LANA) binding to DNA, allowing to understand the energetics of the different LBS DNA and correlate this data to obtain a clearer view of their interactions.

References

1. Jäger, W., Santag, S., Weidner-Glunde, M., Gellermann, E., Kati, S., Pietrek, M., Viejo-Borbolla, A., F. Schulz, T. The Ubiquitin-Specific Protease USP7 Modulates the Replication of Kaposi's Sarcoma-Associated Herpesvirus Latent Episomal DNA *Journal of Virology,* 2012; *Volume 86, Number 12, p. 6745-6757.* **DOI:** 10.1128/JVI.06840-11.

2. Cummins JM, et al. Tumour suppression: disruption of HAUSP gene stabilizes p53. *Nature,* 2004; 428:1 p following 486. **DOI:** 10.1038/nature02501.

3. Cummins JM, Vogelstein B. HAUSP is required for p53 destabilization. *Cell Cycle*, 2004; 3:689–692. **DOI**: 10.4161/cc.3.6.924.

4. Hu M, et al. Structural basis of competitive recognition of p53 and MDM2 by HAUSP/USP7: implications for the regulation of the p53MDM2 pathway. *PLoS Biol*. 2006; 4: e27. **DOI:** 10.1371/journal.pbio.0040027.

5. Li M, Brooks CL, Kon N, Gu W. A dynamic role of HAUSP in the p53-Mdm2 pathway. *Mol. Cell, 2004;* 13:879–886.

6. Li M, et al. Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization. *Nature*, 2002; 416:648–653.

7. Sheng Y, et al. Molecular recognition of p53 and MDM2 by USP7/ HAUSP. *Nat. Struct. Mol. Biol.,* 2006; 13:285–291.

8. Sarkari F, et al. Further insight into substrate recognition by USP7: structural and biochemical analysis of the HdmX and Hdm2 interactions with USP7. *J. Mol. Biol.*, 2010; 402:825–837.

9. Meulmeester E., et al. Loss of HAUSP-mediated deubiquitination contributes to DNA damage-induced destabilization of Hdmx and Hdm2. *Mol. Cell,* 2005; 18:565–576.

10. Tang J, et al. Critical role for Daxx in regulating Mdm2. *Nat. Cell Biol.*, 2006; 8:855–862.

11. Everett RD, et al. A novel ubiquitin-specific protease is dynamically associated with the PML nuclear domain and binds to a herpesvirus regulatory protein. *EMBO J.*, 1997; 16:1519–1530.

12. Canning M, Boutell C, Parkinson J, Everett RD. A RING finger ubiquitin ligase is protected from autocatalyzed ubiquitination and degradation by binding to ubiquitin-specific protease USP7. *J. Biol. Chem.*, 2004; 279:38160–38168.

13. Boutell C, Canning M, Orr A, Everett RD. Reciprocal activities between herpes simplex virus type 1 regulatory protein ICP0, a ubiquitin E3 ligase, and ubiquitin-specific protease USP7. *J. Virol.*, 2005; 79:12342–12354.

14. Holowaty MN, Sheng Y, Nguyen T, Arrowsmith C, Frappier L. Protein interaction domains of the ubiquitin-specific protease, USP7/ HAUSP. *J. Biol. Chem., 2003;* 278:47753–47761.

15. Holowaty MN, et al. Protein profiling with Epstein-Barr nuclear antigen-1 reveals an interaction with the herpesvirus-associated ubiquitin-specific protease HAUSP/USP7. *J. Biol.*, 2003; Chem. 278:29987–29994.

16. van der Knaap JA, et al. GMP synthetase stimulates histone H2B deubiquitylation by the epigenetic silencer USP7. *Mol. Cell*, 2005; 17:695–707.

17. Saridakis V, 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.

18. Gao, S.J., Kingsley, L., Hoover, D.R., Spira, T.J., Rinaldo, C.R., Saah, A., Phair, J., Detels, R., Parry, P., Chang, Y. et al. Seroconversion to antibodies against Kaposi's sarcoma-

associated herpesvirus-related latent nuclear antigens before the development of Kaposi's sarcoma. *N. Engl. J. Med., 1996;* 335, 233–241.

19. Ponnusamy R, Petoukhov MV, Correia B, Custodio TF, Juillard F, Tan M, Pires de Miranda M, Carrondo MA, Simas JP, Kaye KM, Svergun DI, McVey CE. KSHV but not MHV-68 LANA induces a strong bend upon binding to terminal repeat viral DNA. *Nucleic Acids Res*. 2015; Nov 16;43(20):10039-54. doi: 10.1093/nar/gkv987. Epub 2015 Sep 30.

20. Verma, S.C., Lan, K. and Robertson, E. Structure and function of latency-associated nuclear antigen. *Curr. Top. Microbiol. Immunol.*, 2007; 312, 101–136.

21. Piolot, T., Tramier, M. and Coppey, M. Close but distinct regions of human herpesvirus 8 latency-associated nuclear antigen 1 are responsible for nuclear targeting and binding to human mitotic chromosomes*. J. Virol.*, 2001; 75, 3948–3959.

22. Szekely, L., Kiss, C., Mattsson, K., Kashuba, E., Pokrovskaja, K., Juhasz, A., Holmvall, P. and Klein, G. Human herpesvirus-8-encoded LNA-1 accumulates in heterochromatin-associated nuclear bodies. *J. Gen. Virol.*, 1999; 80, 2889–2900.

23. Barbera, A.J., Chodaparambil, J.V., Kelley-Clarke, B., Joukov, V., Walter, J.C., Luger, K. and Kaye, K.M. The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science*, 2006; 311, 856–861.

24. Barbera, A.J., Ballestas, M.E. and Kaye, K.M. The Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen 1 N terminus is essential for chromosome association, DNA replication, and episome persistence. *J. Virol.*, 2004; 78, 294–301.

25. Ballestas, M. and Kaye, K.M. 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, 3250–3258.

26. Garber, A.C., Hu, J. and 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.

27. Cotter, M.A., Subramanian, C. and Robertson, E.S. The Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen binds to specific sequences at the left end of the viral genome through its carboxy-terminus. *Virology*, 2001; 291, 241–259.

28. Komatsu, T., Ballestas, M.E., Barbera, A.J., Kelley-Clarke, B. and Kaye, K.M. KSHV LANA1 binds DNA as an oligomer and residues N-terminal to the oligomerization domain are essential for DNA binding, replication, and episome persistence*. Virology*, 2004; 319, 225–236.

29. Hellert, J., Weidner-Glunde, M., Krausze, J., Lunsdorf, H., Ritter, C., Schulz, T.F. and Luhrs, T. The 3D structure of Kaposi sarcoma herpesvirus LANA C-terminal domain bound to DNA. *Proc. Natl. Acad. Sci.*, 2015; 112, 6694–6699.

30. Gao, S.J., Zhang, Y.J., Deng, J.H., Rabkin, C.S., Flore, O. and Jenson, H.B. Molecular polymorphism of Kaposi's sarcoma-associated herpesvirus (Human herpesvirus 8) latent nuclear antigen: evidence for a large repertoire of viral genotypes and dual infection with different viral genotypes. *J. Infect. Dis.*, 1999; 180, 1466–1476.

31. Habison, A.C., Beauchemin, C., Simas, J.P., Usherwood, E.J. and Kaye, K.M. Murine gammaherpesvirus 68 LANA acts on terminal repeat DNA to mediate episome persistence. *J. Virol.*, 2012; 86, 11863–11876.

32. Paden, C.R., Forrest, J.C., Tibbetts, S.A. and Speck, S.H. Unbiased mutagenesis of MHV68 LANA reveals a DNA-binding domain required for LANA function in vitro and in vivo. *PLoS Pathog.*, 2012; 8, e1002906.

33. Lan, K., Kuppers, D.A., Verma, S.C., Erle, S. and Robertson, E.S. Kaposi's sarcoma-associated herpevirus-encoded latency-associated nuclear antigen inhibits lytic replication by

targeting Rta: a potential mechanism for virus-mediated control of latency. *J. Virol.*, 2004; 78, 6585–6594.

34. Stedman, W., Deng, Z., Lu, F. and Lieberman, P.M. ORC, MCM, and histone hyperacetylation at the Kaposi's sarcoma-associated herpesvirus latent replication origin. *J. Virol.*, 2004; 78, 12566–12575.

35. Sun, Q., Tsurimoto, T., Juillard, F., Li, L., Li, S., De Léon Vázquez, E., Chen, S. and Kaye, K. Kaposi's sarcoma-associated herpesvirus LANA recruits the DNA polymerase clamp loader to mediate efficient replication and virus persistence. *Proc. Natl. Acad. Sci. U.S.A.*, 2014; 111, 11816–11821.

36. Shrestha, P. and Sugden, B. Identification of properties of KSHV's latent origin of replication essential for the efficient establishment and maintenance of intact plasmids. *J. Virol.*, 2014; 88, 8490–8503.

37. Skalsky, R.L., Hu, J. and Renne, R. Analysis of viral cis elements conferring Kaposi's sarcoma-associated herpesvirus episome partitioning and maintenance. *J. Virol.*, 2007; 81, 9825–9837.

38. Hellert, J., Weidner-Glunde, M., Krausze, J., Richter, U., Adler, H., Fedorov, R., Pietrek, M., Ruckert, J., Ritter, C., Schulz, T.F. et al. A structural basis for BRD2/4-mediated host chromatin interaction and oligomer assembly of Kaposi sarcoma-associated herpesvirus and murine gammaherpesvirus LANA proteins. *PLoS Pathog.*, 2013; 9, e1003640.

39. Domsic, J.F., Chen, H.-S., Lu, F., Marmorstein, R. and Lieberman, P.M. Molecular basis for oligomeric-DNA binding and episome maintenance by KSHV LANA. *PLoS Pathog.*, 2013; 9, e100367.

40. Correia, B., Cerqueira, S.A., Beauchemin, C., Pires de Miranda, M., Li, S., Ponnusamy, R., Rodrigues, L., Schneider, T.R., Carrondo, M.A., Kaye, K.M. 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, e1003673.

41. Bochkarev, A., Barwell, J.A., Pfuetzner, R.A., Furey, W., Edwards, A.M. and Frappier, L. Crystal structure of the DNA-binding domain of the Epstein-Barr virus origin-binding protein EBNA 1. *Cell*, 1995; 83, 39–46.

42. Liang, H., Petros, A.M., Meadows, R.P., Yoon, H.S., Egan, D.A., Walter, K., Holzman, T.F., Robins, T. and Fesik, S.W. Solution structure of the DNA-binding domain of a human papillomavirus E2 protein: evidence for flexible DNA-binding regions. *Biochemistry*, 1996; 35, 2095–103.

43. Guarnieri, Anna L. et al*.* Back to Bases: How a Nucleotide Biosynthetic Enzyme Controls p53 Activation. *Molecular Cell*, 2014; Volume 53 , Issue 3 , 365 – 367.

44. Robbert Q.KimTitia K.Sixma. Regulation of USP7: A High Incidence of E3 Complexes. *J.Mol.Biol.*, 2017; Volume 429, Issue 22, Pages 3395-3408.

45. Lieberman PM. Keeping it quiet: chromatin control of gammaherpesvirus latency. *Nat Rev Microbiol.*, 2013;11(12):863-75. doi: 10.1038/nrmicro3135. Epub 2013 Nov 6.

46. Kelley-Clarke B, DE Leon-Vazquez E, Slain K, Barbera AJ, Kaye KM. Role of Kaposi's sarcoma associated herpesvirus C-terminal LANA chromosome binding in episome persistence. *J. Virol.*, 2009; 83(9):4326–4337. [PubMed: 19225000].

47. Jan Hellert, Magdalena Weidner-Glunde, Joern Krausze, Ulrike Richter, Heiko Adler, Roman Fedorov, Marcel Pietrek, Jessica Rückert, Christiane Ritter, Thomas F. Schulz, Thorsten Lührs. A Structural Basis for BRD2/4-Mediated Host Chromatin Interaction and Oligomer Assembly of Kaposi Sarcoma-Associated Herpesvirus and Murine Gammaherpesvirus LANA Proteins. *PLOS Pathogens*, 2013; 9(10): e1003640. doi: 10.1371/journal.ppat.1003640.

48. Hunt R., Herpes Viruses, Virology-Chapter eleven, Microbiology and Immunology On-line. http://www.microbiologybook.org/virol/herpes.htm. October 29th, 2016. Accessed Jully 3rd, 2018.

49. Whitley R.J. *Medical Microbiology*. 4th edition. Chapter 68 – Herpesviruses. Galveston (TX): Baron S, editor; 1996.

50. Boshoff, C., and Wiess R.A. Epidemiology and pathogenesis of Kaposi's sarcoma-associated herpesvirus*. Philos. Trans. R. Soc. Lond.* 2001; B356**:**517-534.

51. Cesarman, E., Y. Chang, P. S. Moore, J. W. Said, and D. M. Knowles. Kaposi's sarcoma-associated herpesvirus-like DNA sequences in AIDS-related body-cavity-based lymphomas*. N. Engl. J. Med*., 1995; 332**:**1186-1191.

52. Cesarman, E., R. G. Nador, K. Aozasa, G. Delsol, J. W. Said, and D. M. Knowles. Kaposi's sarcoma-associated herpesvirus in non-AIDS-related lymphomas occurring in body cavities. *Am. J. Pathol.*, 1996; 149**:**53-57.

53. Cesarman, E., R. G. Nador, F. Bai, R. A. Bohenzky, J. J. Russo, P. S. Moore, Y. Chang, and D. M. Knowles. Kaposi's sarcoma-associated herpesvirus contains G protein-coupled receptor and cyclin D homologs which are expressed in Kaposi's sarcoma and malignant lymphoma*. J. Virol.*, 1996; 70**:**8218-8223.

54. Chang, Y., E. Cesarman, M. S. Pessin, F. Lee, J. Culpepper, D. M. Knowles, and P. S. Moore. Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science*, 1994; 266**:**1865-1869.

55. Russo, J. J., R. A. Bohenzky, M. C. Chien, J. Chen, M. Yan, D. Maddalena, J. P. Parry, D. Peruzzi, I. S. Edelman, Y. Chang, and P. S. Moore. Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8). *Proc. Natl. Acad. Sci. USA*, 1996; 93:14862-14867.

56. Sarid, R., A. Klepfish, and A. Schattner. Virology, pathogenetic mechanisms, and associated diseases of Kaposi sarcoma-associated herpesvirus (human herpesvirus 8). *Mayo Clin. Proc.*, 2002; 77:941-949.

57. Sarid, R., S. J. Olsen, and P. S. Moore. Kaposi's sarcoma-associated herpesvirus: epidemiology, virology, and molecular biology. *Adv. Virus Res.*, 1999; 52**:**139-232.

58. West, J. T., and C. Wood. The role of Kaposi's sarcoma-associated herpesvirus/human herpesvirus-8 regulator of transcription activation (RTA) in control of gene expression. *Oncogene, 2003;* 22**:**5150-5163.

59. Miller, G., L. Heston, E. Grogan, L. Gradoville, M. Rigsby, R. Sun, D. Shedd, V. M. Kushnaryov, S. Grossberg, and Y. Chang. Selective switch between latency and lytic replication of Kaposi's sarcoma herpesvirus and Epstein-Barr virus in dually infected body cavity lymphoma cells*. J. Virol.*, 1997; 71**:**314-324.

60. Sarid, R., O. Flore, R. A. Bohenzky, Y. Chang, and P. S. Moore. Transcription mapping of the Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) genome in a body cavity-based lymphoma cell line (BC-1). *J. Virol.*, 1998; 72**:**1005-1012.

61. 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.

62. 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.

63. Lori Frappier, Contributions of Epstein–Barr Nuclear Antigen 1 (EBNA1) to Cell Immortalization and Survival, *Viruses*. 2012 Sep; 4(9): 1537–1547. Published online 2012 Sep 13. doi: 10.3390/v4091537.

64. Holowaty M.N., Sheng Y., Nguyen T., Arrowsmith C., Frappier L. Protein interaction domains of the ubiqutin specific protease, USP7/HAUSP. *J. Biol. Chem.* 2003; 278:47753–47761.

65. Saridakis V., Sheng Y., Sarkari F., Holowaty M.N., Shire K., Nguyen T., Zhang R.G., Liao J., Lee W., Edwards A.M., 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.

66. Sheng Y., Saridakis V., Sarkari F., Duan S., Wu T., Arrowsmith C.H., Frappier L. Molecular recognition of p53 and mdm2 by USP7/HAUSP. *Nat. Struct. Mol. Biol.* 2006; 13:285–291. doi: 10.1038/nsmb1067.

67. Li M., Chen D., Shiloh A., Luo J., Nikolaev A.Y., Qin J., Gu W. Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization. *Nature.* 2002; 416:648–653.

68. Li M., Brooks C.L., Kon N., Gu W. A dynamic role of HAUSP in the p53-mdm2 pathway. *Mol. Cell.* 2004; 13:879–886.

69. Cummins J.M., Rago C., Kohli M., Kinzler K.W., Lengauer C., Vogelstein B. Tumour suppression: Disruption of HAUSP gene stabilizes p53. *Nature.* 2004; 428:486–487.

70. Sivachandran N., Sarkari F., Frappier L. Epstein-Barr nuclear antigen 1 contributes to nasopharyngeal carcinoma through disruption of PML nuclear bodies. *PLoS Pathog.* 2008; 4: e1000170. doi: 10.1371/journal.ppat.1000170.

71. Hu M, et al. Structural basis of competitive recognition of p53 and MDM2 by HAUSP/USP7: implications for the regulation of the p53MDM2 pathway. *PLoS Biol.* 2006; 4: e27.

72. Sheng Y, et al. Molecular recognition of p53 and MDM2 by USP7/ HAUSP. *Nat. Struct. Mol. Biol.* 2006; 13:285–291.

73. Kelley-Clarke B, et al. Determination of Kaposi's sarcomaassociated herpesvirus C-terminal latency-associated nuclear antigen residues mediating chromosome association and DNA binding. *J. Virol.* 2007; 81: 4348–4356.

74. Schwam DR, Luciano RL, Mahajan SS, Wong L, Wilson AC. Carboxy terminus of human herpesvirus 8 latency-associated nuclear antigen mediates dimerization, transcriptional repression, and targeting to nuclear bodies. *J. Virol.* 2000; 74:8532–8540.

75. Sarkari F, et al. EBNA1-mediated recruitment of a histone H2B deubiquitylating complex to the Epstein-Barr virus latent origin of DNA replication. *PLoS Pathog.* 2009; 5: e1000624.

76. Fleming AB, Kao CF, Hillyer C, Pikaart M, Osley MA. H2B ubiquitylation plays a role in nucleosome dynamics during transcription elongation. *Mol. Cell,* 2008*;* 31: 57–66.

77. Zhao Y, et al. A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counter acts heterochromatin silencing. *Mol. Cell*, 2008; 29:92–101.

78. Wong LY, Wilson AC. 2005. Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen induces a strong bend on binding to terminal repeat DNA. *J. Virol. 2005;* 79:13829–13836.

79. *Bitesize Bio.* "Get Your Clone 90% Of The Time with Ligation Independent Cloning". (Retrieved 2017-12-10).

80. Le, V.H., Buscaglia, R., Chaires, J.B. and Lewis, E.A. Modeling complex equilibria in isothermal titration calorimetry experiments: thermodynamic parameters estimation for a three-binding-site model. *Anal. Biochem*., 2013; 434, 233–241.

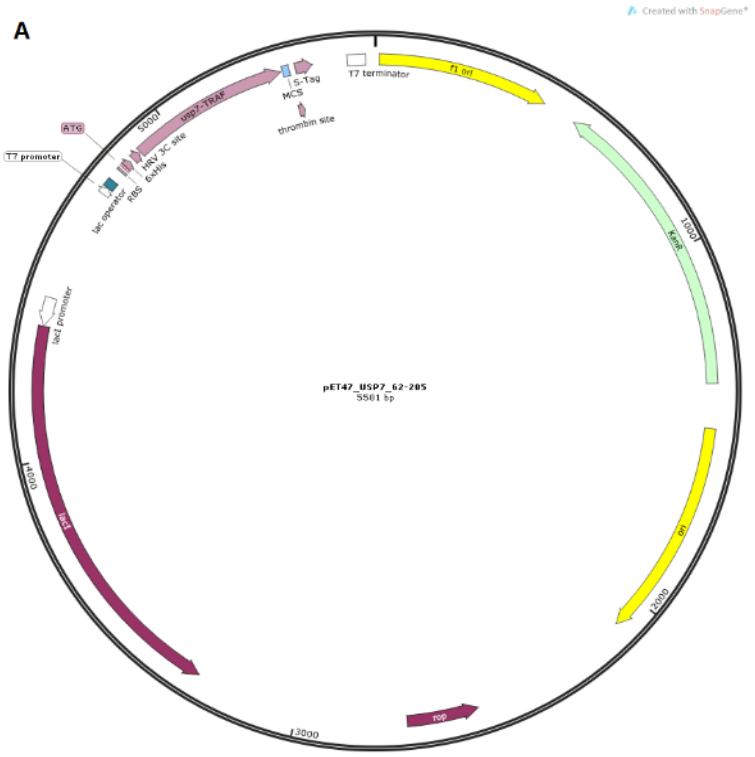
81. Nicol SM, Sabbah S, Brulois KF, Jung JU, Bell AI, Hislop AD. Primary B Lymphocytes Infected with Kaposi's Sarcoma-Associated Herpesvirus Can Be Expanded In Vitro and Are Recognized by LANA-Specific CD4+ T Cells. *J Virol.* 2016 Mar 28; 90(8):3849-3859. DOI: 10.1128/JVI.02377-15. Print 2016 Apr.

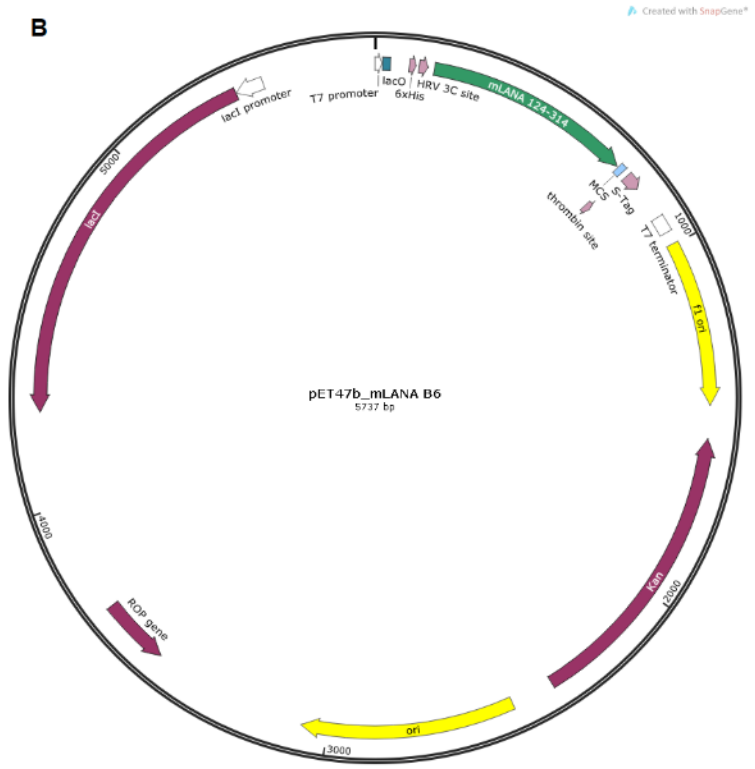
82. Custodio T, McVey CE, Ponnusamy R*, "Staying under the radar" - The Multifunctional LANA Protein*, [Master Thesis], 2015; p.14-15.

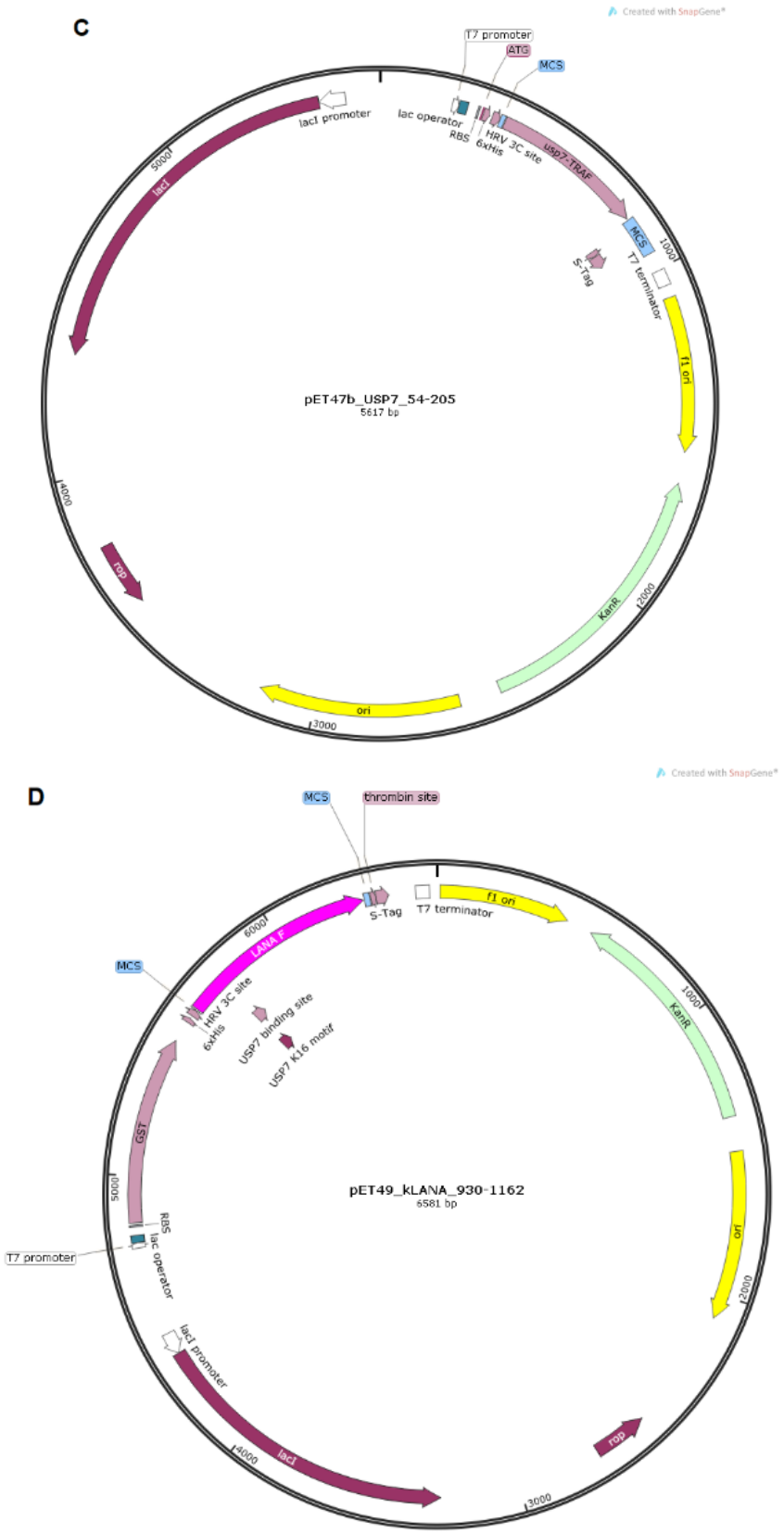
83. Summers H, Barwell JA, Pfuetzner RA, Edwards AM & Frappier L. Cooperative assembly of EBNA1 on the Epstein-Barr virus latent origin of replication. *J Virol,* 1996; 70, 1228–1231.

84. Sun Z, Jha HC, Robertson ES. Bub1 in Complex with LANA Recruits PCNA To Regulate Kaposi's Sarcoma-Associated Herpesvirus Latent Replication and DNA Translesion Synthesis, *J Virol.* 2015 Oct;89(20):10206-18. doi: 10.1128/JVI.01524-15. Epub 2015 Jul 29.

Supplementary information






C
Created with SnapGene®
T7 promoter
ATG
MCS
lac operator
RBS
6xHis
HRV 3C site
usp7-TRAF
MCS
T7 terminator
S-Tag
f1 ori
lacI promoter
lacI
5000
1000
pET47b_USP7_54-205
5617 bp
4000
rop
KanR
2000
ori
3000
D
Created with SnapGene®
MCS
thrombin site
S-Tag
T7 terminator
f1 ori
6000
LANA F
MCS
HRV 3C site
6xHis
USP7 binding site
USP7 K16 motif
1000
KanR
GST
5000
RBS
lac operator
T7 promoter
pET49_kLANA_930-1162
6581 bp
ori
2000
lacI promoter
lacI
4000
rop
3000

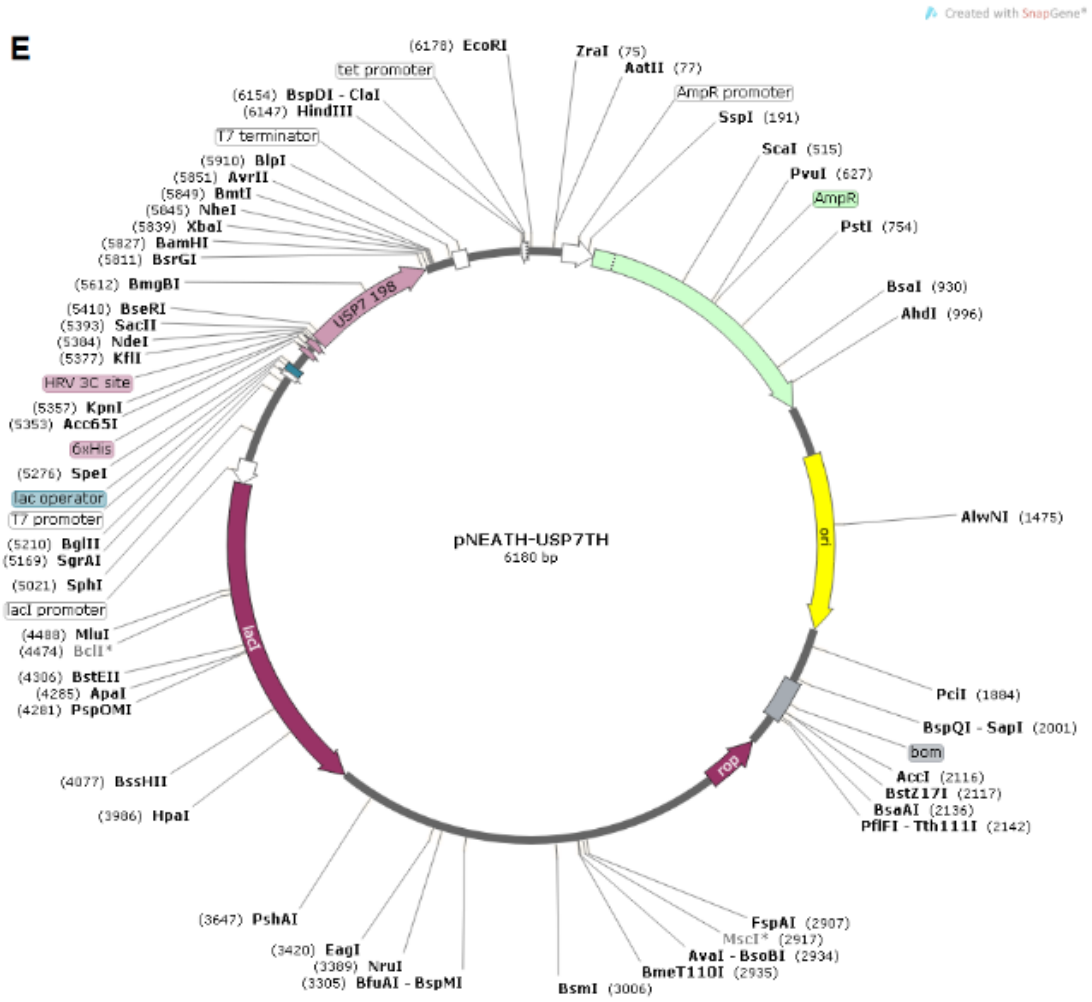


Figure S1: Map illustration of the vectors used for cloning in this thesis. Design of the reported vector **A)** pET 47 b(+), encoding resistance for kanamycin with a 6xHis-Tag and a $USP7_{62-205}$ insert; **B)** same vector, but with a $mLANA_{124-314}$ insert; **C)** same vector, but with an $USP7_{54-205}$ insert; **D)** pET49 b(+) encoding resistance for kanamycin with a GST-6xHis-Tag and a $kLANA_{930-1162}$ insert; **E)** pNEATH encoding resistance for ampicillin with a 6xHis-Tag and a $USP7_{54-198}$ insert. The desired genes were cloned in the Multiple Cloning Site (MCS) region.