Přírodovědecká fakulta
Faculty of Science

Jihočeská univerzita v Českých Budějovicích
University of South Bohemia in České Budějovice

Mass spectrometry in structural proteomics: study of viral protein complexes

Bachelor Thesis

Anirudhan Ravi

Supervisor: Dyčka Filip, Mgr. Ph.D.

Co-supervisor: Tůma Roman, doc. Ph.D.

České Budějovice 2021

Ravi A., 2021: Mass spectrometry in structural proteomics: study of viral protein complexes, Bc. Thesis, in English 46 p. Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

ANNOTATION

This thesis investigated the avian reovirus σNS protein and its interaction with RNA using mass spectrometric methods. The σNS structure is still not known; thus, hydrogen-deuterium exchange and cross-linking coupled to mass spectrometry were employed to understand protein conformation. This research provides insights into protein-RNA interactions and the obtained data will be used in future work to predict protein structure.

DECLARATION

I declare that I am the author of this qualification thesis and that in writing it I have used the sources and literature displayed in the list of used sources only.

České Budějovice, 15.12.2021

……..

Anirudhan Ravi

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Dyčka Filip, Mgr. Ph.D for giving me the chance to work on this bachelor thesis and also for the patience in explaining whenever I had questions regarding the project. Next, my thanks are dedicated to whole Makrokomplex group which welcomed from the first day.

Furthermore, my acknowledgement goes to my family who supported and encouraged me in studying biological chemistry Bachelor's degree program.

ABSTRACT

Avian Reovirus ARV belongs to the *Orthoreovirus* genus, part of the *Reoviridae* family. They are pathogens that infect poultry fowl causing many diseases. The non-structural σNS protein in ARV is important for viral assembly nucleation and the creation of viral inclusions. In this work, Rosetta and Phyre2 softwares were used in an attempt to predict the structure of the σNS protein. Various mass spectrometric methods were employed to study protein structure and σNS-RNA interactions. Matrix assisted laser/desorption ionization mass spectrometry was used to control the quality of protein purification before the analysis. Cross-linking mass spectrometry revealed an inter-molecular cross-linker at the Lys-193 position showing a location of subunit interactions during the formation of the σNS oligomer. In total 30 cross-linked peptides were identified in the MeroX software, but only seven peptides were unique. Although, three cross-links were detected in the His-tag and SUMO sequence of the protein. The cross-linked peptides were mapped onto the structural models of σNS. According to obtained data, predicted protein structures should have been modified as distances between several linked amino acids were over the expected threshold. Hydrogen-deuterium exchange coupled to mass spectrometry was employed to study the σNS-RNA interaction. Obtained results showed many regions of σNS to be less protected in protein alone than in presence of RNA due to local stabilization upon RNA binding. In contrast, peptides at the C-terminal part showed higher local stability in σNS compared to the σNS-RNA complex.

CONTENTS

1 INTRODUCTION 1
1.1 Structure and types of Viruses 1
1.2 Avian Reovirus ARV 2
1.3 Protein σNS 4
1.4 Mass Spectrometry 5
1.5 Cross-linking mass spectrometry 7
1.6 Hydrogen-deuterium exchange mass spectrometry 9
2 AIMS OF THESIS 11
3 MATERIALS AND METHODS 12
3.1 Chemicals 12
3.2 Preparation of σNS protein 12
3.3 Protein purification using ZipTip 13
3.4 Intact mass MALDI MS measurement 13
3.5 Cross-linking reaction of protein σNS 13
3.6 HDX MS analysis 15
3.7 Softwares for prediction and visualization of protein structure 16
4 RESULTS 17
4.1 Prediction of σNS protein structure 17
4.2 Estimation of intact protein mass by MALDI MS 18
4.3 Study of σNS protein by cross-linking mass spectrometry 20
4.4 Protein-RNA complex analysis using hydrogen deuterium exchange mass spectrometry 22
5 DISCUSSION 27
6 CONCLUSIONS 30
7 REFERENCES 31
8 SUPPLEMENTARY 38

LIST OF ABBREVIATIONS

ARV – avian reovirus

BS3 – bis(sulfosuccinimidyl)suberate

CDI – carbonyldiimidazole

CLMS – cross-linking mass spectrometry

DSBU – disuccinimidyl dibutyric urea

dsDNA – double stranded DNA

dsRNA – double stranded RNA

DSS – disuccinimidyl suberate

DTT – dithiothreitol

ESI – electrospray ionization

FDR – false discovery rate

GHCl – guanidine hydrochloride

HDX – hydrogen-deuterium exchange

HPLC – high-performance liquid chromatography

IMS – ion mobility mass spectrometry

IPTG – isopropyl β-D-1-thiogalactopyranoside

LC – liquid chromatography

MALDI – matrix assisted laser desorption ionization

MD – molecular dynamics

MeOH – methanol

MRV – mammalian reovirus

NHS – N-hydroxysuccinimide

PLGS – ProteinLynX Global SERVER

Q – quadrupole

RT – reverse transcription

ssDNA – single stranded DNA

ssRNA – single stranded RNA

SUMO – small ubiquitin-like modifier

TFA – trifluoroacetic acid

TOF – time-of-flight

1 INTRODUCTION

1.1 Structure and types of Viruses

Viruses represent a current issue in the world since they can cause many diseases such as influenza, Ebola, and pneumonia. These tiny parasites, with a diameter between 30-300 nm are surrounded by protein shells known as capsids that protect the genome, which can be either DNA or RNA (Ryu, 2017). To replicate themselves, they have to rely on a cell that can accommodate them since their purpose is to transmit their genetic information to be transcribed and translated by the host cell (Rampersad & Tennant, 2018). In nature, viruses can be either enveloped or non-enveloped pathogens. In the first case, the capsid is covered by a lipid bilayer composed of glycoproteins. On the other hand, non-enveloped viruses contain only the genetic material and protein shell (Sakudo et al., 2010). The viruses without an envelope such as norovirus, poliovirus, human hepatitis A virus (HAV), rotavirus, and coxsackie viruses can be very infectious for a long time, resistant to alcohols and heat (Sakudo et al., 2010; Wörner et al., 2020).

From a structure's perspective, viruses can be classified based on their capsid arrangements to icosahedral, helical, and complex types. Typically, an icosahedral virus has a polyhedral structure in which the faces are in an equilateral triangular orientation (Louten J., 2016). When the genome twists in a helical shape, with protein shells turning around it, they belong to helical viruses (Louten J., 2016). A complex structure with a genome surrounded by a phage protein capsid is important for bacteriophages to be able to attack and replicate in prokaryotic cells (E. White & V. Orlova, 2020).

What is important to know is whether viruses are encoded by RNA or DNA, which is classified according to the Baltimore system (Louten J., 2016). The system classifies viruses into seven categories. Group I and II are double and single-stranded DNA viruses (**Fig. 1A**). Group III, IV, and V include double-stranded RNA, positive single-stranded RNA, and negative single-stranded RNA viruses (**Fig. 1B**). Furthermore, Baltimore's classification consists of positive single-stranded RNA viruses which can do reverse transcription from RNA to DNA (group VI) and double-stranded DNA with RNA intermediates (group VII) (**Fig. 1 A&B**) (Mahmoudabadi & Phillips, 2018).

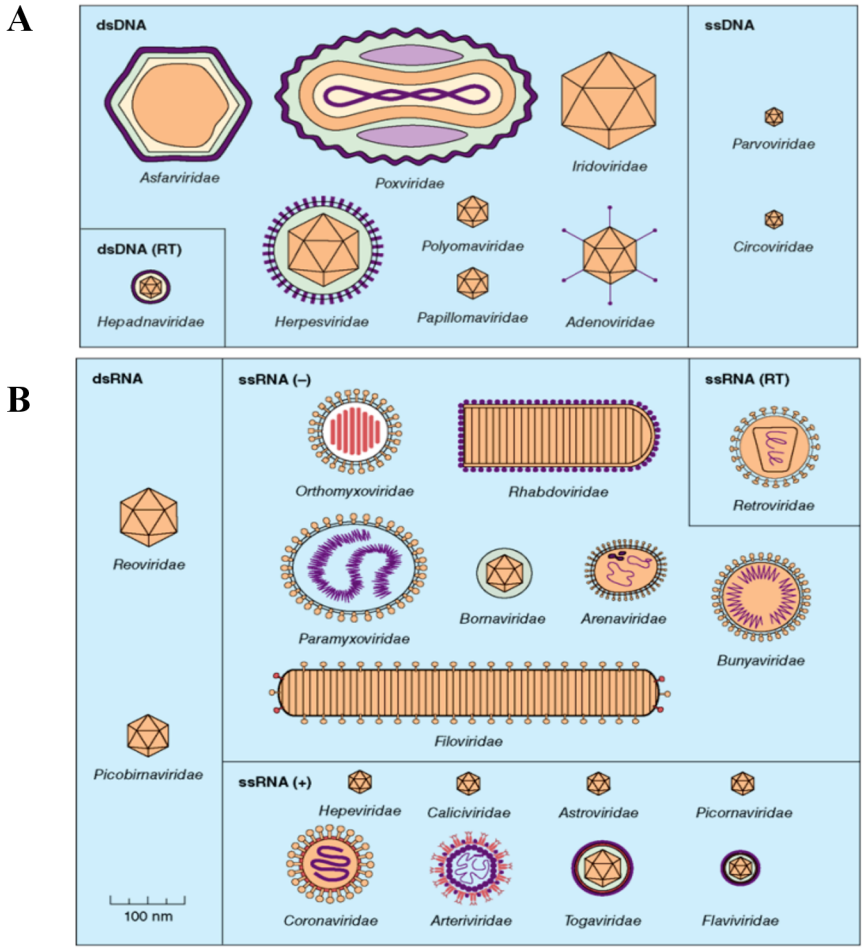

Fig. 1: (**A**) double-stranded (ds) and single-stranded (ss) DNA viruses (**B**) double-stranded (ds) and single-stranded (ss) RNA viruses. RT means reverse transcribed virus (Pierce, 2020).

1.2 Avian Reovirus ARV

Avian Reoviruses ARV are non-enveloped pathogens that infect birds causing different diseases such as hepatitis, myocarditis as well as respiratory illness. These agents, with an icosahedral structure, are classified in the *Reoviridae* family, part of *the Orthoreovirus* genus (Benavente& Martinez-Costas, 2007; Sahin et al., 2013). With a genome composition structured into 10 segments of double-stranded RNA (dsRNA), they are capable of coding four nonstructural and eight structural proteins (Bodelòn et al., 2001; Schnitzer, 1985; Varela & Benavente, 1994). Apart from ARV, other four species are members of *the Orthoreovirus* genus which can be classified into two major groups, ARV and the mammalian reovirus MRV. Although they share similar characteristics, they are very different regarding pathogenicity and biological properties (Benavente& Martinez-Costas, 2007). MRV as opposed to ARV doesn't fuse infected cells but causes clumps of red blood cells (Nibert & Schiff, 2001).

ARV viral particles (**Fig. 2**) with an icosahedral shape have an external diameter of 85 nm approximately, and consist of an inner core that contains segments of the genome that are protected by inner and outer protein shells (Spandidos & Graham, 1976; Zhang et al., 2005;). The inner core is built by λA and σA proteins which conform to the stability of the shell. Pentamers of λC protein form turrets from the inner to outer capsid and the external part is where trimers of σC attach to the cell. The outer protein shell which covers the virion particle is composed of μB and σB proteins (Benavente& Martinez-Costas, 2007).

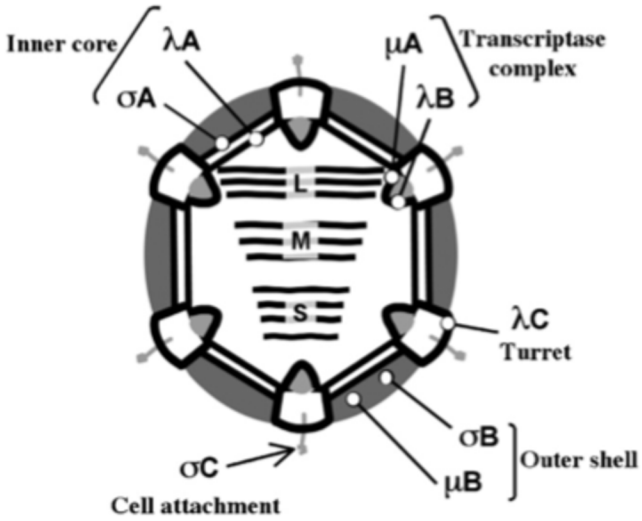

Fig. 2: Schematic representation of ARV structure (Benavente& Martinez-Costas, 2007).

The ARV genome is composed of 10 double-stranded RNA, which is divided into three components according to their size. Four segments are small (S1, S2, S3, and S4), three are medium (M1, M2, and M3), and three large (L1, L2, and L3) segments. Apart from the S1 segment that encodes 3 products, all of them are monocistronic which can only encode one protein (Mirbagheri et al., 2019; Bodelòn et., 2001). Each segment with a negative-strand consists of a pyrophosphate group at the 5' end whereas in the positive strand it can be found a type-1 cap at this position (Martinez-Costas et al., 1995). The ARV genome expresses 12 primary translational products since nine segments are monocistronic and only one protein codes for three products. Four are non-structural proteins that are located on infected cells, whereas the remaining eight are the structural ones that belong to the virion. Two non-structural proteins known as σNS and μNS are encoded by S4 and M3 genes, which are present in the cytoplasm of infected cells. Furthermore, the p10 and p17 segments are encoded by the S1 segment (Benavente& Martinez-Costas, 2007; Bodelòn et al., 2001).

1.3 Protein σNS

The non-structural protein interacting with RNA in ARV is σNS encoded by the S4 gene, with a mass of around 40 kDa and a length of 367 amino acids (Chiu & Lee, 1997; Schnitzer, 1985; Varela & Benavente, 1994). It can bind single-stranded RNA (ssRNA), however, a minimal RNA length between 10-20 nucleotides is required for σNS binding (Tourís-Otero et al., 2005). The structure of σNS has not been identified yet despite many trials using different techniques such as X-ray crystallography (Bravo, 2019).

According to some studies, it has been proposed that σNS plays an important role in viral genome packaging (Benavente J & Martinez-Costas J., 2007). Different functions of ARV σNS include segment assortment and RNA polymerase activity (Stamatos & Gomatos, 1982). Together with μNS, it was found that σNS protein is present as a large ribonucleoprotein complex located in the cytoplasm of infected cells (Touris-Otero et al., 2004). The ARV σNS protein acts as an RNA chaperone and it is important in the replication process (Rajkowitsch et al., 2007; Borodavka et al., 2015). When σNS binds to RNA, the single strand is opened and the disulfide bridges are disrupted between them causing the removal of the RNA secondary structure (Borodavka et al., 2015; Bravo et al., 2018). In nature, σNS forms hexamer from homodimer and homotrimer when it's not bound to RNA (Tourís-Otero et al., 2005).

Native electrospray ionization mass spectrometry (ESI-MS) and ion mobility mass spectrometry (IMS-MS) showed the evidence of hexamer by spectra analysis (Borodovka et al., 2015). However, monomers, dimers, and tetramers were also present in small quantities due to dissociation from electrospray ionization and it was found that dimer is the basis for hexamer formation (**Fig. 3**) (Borodovka et al., 2015). The σNS protein undergoes from hexamer to octamer transition upon coupling multiple RNA molecules (Bravo, 2019). Therefore, a stable octameric ribonucleoprotein (RNP) complex is formed when σNS is bound to multiple RNAs (Bravo et al., 2018). However, the interaction of σNS protein with RNAs is poorly understood (Borodovka et al., 2015).

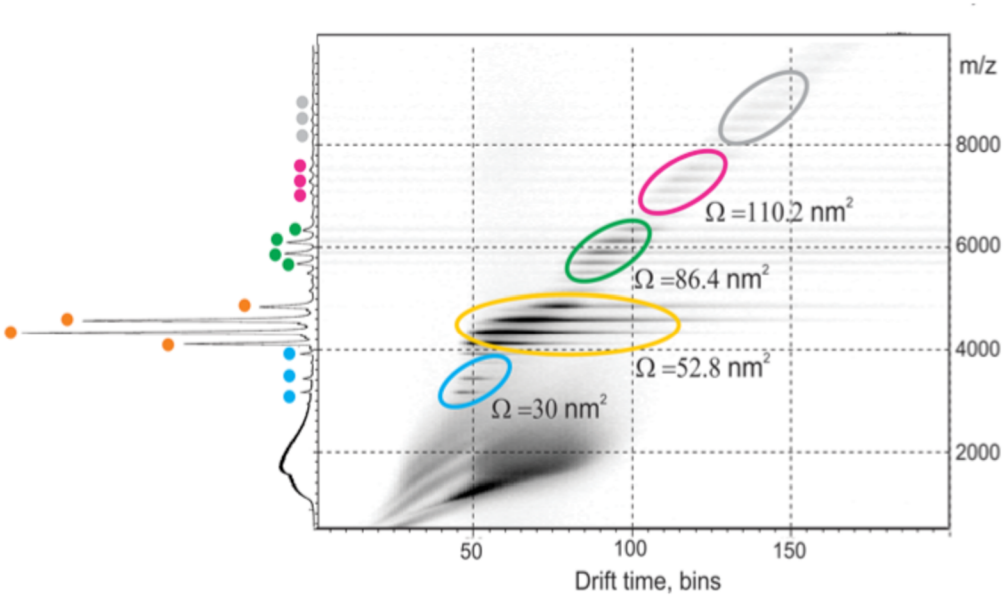

Fig. 3: Native ESI-IMS-MS analysis of ARV σNS protein. Blue color indicates the presence of monomer, orange dimers, green tetramers, magenta hexamers, whereas grey color explains the presence of other types of oligomers (Borodovaka et al., 2015).

1.4 Mass Spectrometry

Mass spectrometry is a method that is widely used in the present day to identify proteins. It is possible to determine the exact mass of a particular molecule using this technique. The main idea behind MS consists of three steps (**Fig. 4**). The first step is ionization, by means of changing particles from liquid or solid form to vacuum and providing the charge to them (Olshina & Sharon, 2016). The next step consists of separating the particles that are present in a sample according to the mass to charge ratio (*m/z*). In the end, the ions are detected using a detector. A final result is a mass spectrum which represents the mass to charge (*m/z*) of ions on the horizontal axis and relative abundance of ions on the vertical axis (Urban, 2016).

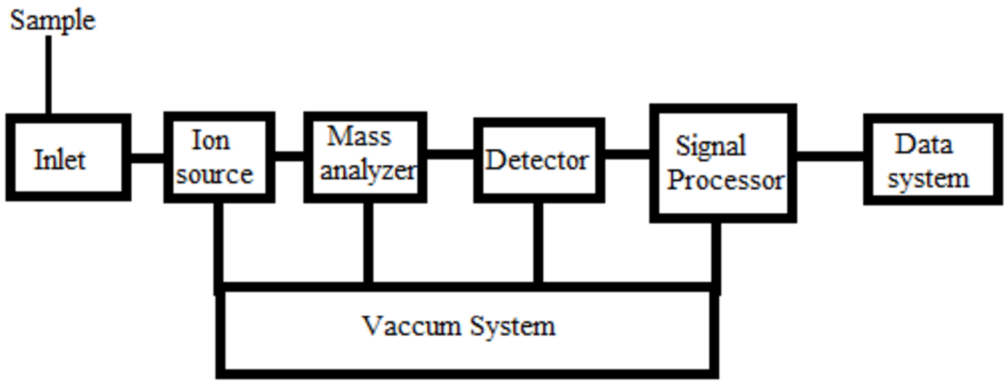

Fig. 4: Scheme of a mass spectrometer components.

Molecule of interest is transformed to the gas phase, ionized, and loaded into the inlet. However, generating stable ions of high mass molecules, like proteins or peptides, is not easy, and to solve this problem two types of ionization are commonly used, electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI) (Kicman et al., 2007).

Electrospray ionization is used to transform samples of interest into the gas phase. During this process, the solution is poured into a capillary or needle, then attracted by an electric force that induces a charge in the analytes (Olshina & Sharon, 2016). The fluid of charged particles at the end of the needle has a conical shape known as the Taylor cone. When the droplets travel from the capillary to the inlet, their radius will become smaller due to the evaporation of the solvent. The droplets with ions continue to reduce their size until they reach the Rayleigh limit (Banerjee & Mazumdar, 2012). Therefore, a Coulomb explosion will be reached since the Coulomb forces prevail over the surface tension. Tiny droplets then travel to the mass analyzer (Kicman et al., 2007).

Molecules with a large molecular weight up to 500 kDa can be easily analyzed with matrix-assisted laser desorption ionization (MALDI) which is a soft ionization technique (Chong et al., 1997). This method is used to determine the intact mass of proteins, peptides, lipids, and other organic macromolecules.

An analyte is spotted onto the MALDI plate together with the matrix (**Fig. 5**). During sample preparation, the analyte has to co-crystallize with a matrix. The matrix is commonly an organic acid molecule that can absorb energy from a laser beam, by bringing the matrix to the excited state of energy (Cobo, 2013). Since molecules are excited, the matrix will transfer a proton to the analyte and the energy will modify its phase from solid to gas form (Clark et al., 2013). As a consequence, the ionized particles are subjected to an electric field by speeding up the velocity and entering into a mass analyzer (Clark et al., 2013).

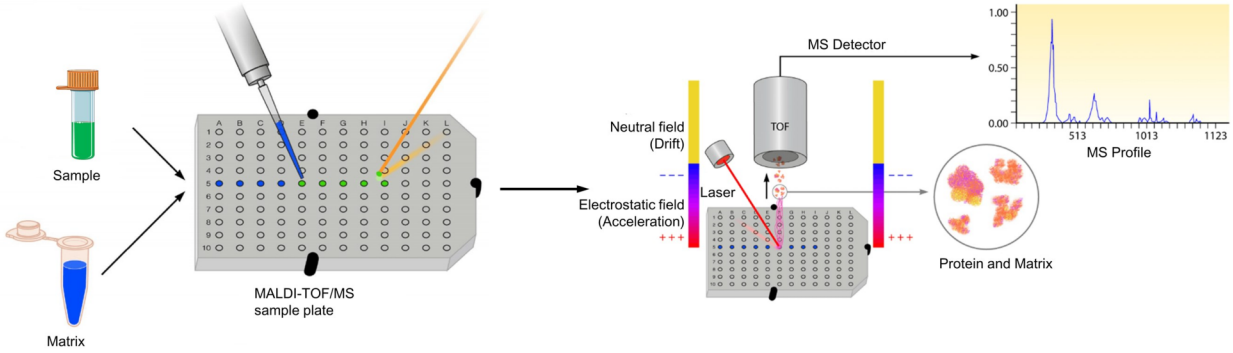

Fig. 5: Schematic principle of MALDI-TOF mass spectrometry (Clark et al., 2013).

Different types of mass spectrometers can be used in mass spectrometry. One of the most often used mass analyzers in proteomics is quadrupole (Q) and time of flight (TOF).

The quadrupole instrument is composed of four parallel rods that are displayed as a hyperbolic or circular cross-section, where the opposing rods are connected electrically (Chong et al., 2018). The instrument produces an electric field that allows it to pass through only ions with a defined mass-to-charge ratio. This is obtained by applying radiofrequency (RF) and direct current (DC) signals to each opposing rod pair which creates a low-pass and high pass filter in the two orthogonal planes (Chong et al., 2018). Therefore, only ions with a defined *m/z* range can go through quadrupole, whereas other ions hit with the rods and don't reach the detector (Thomas, 2019). This type of instrument is very robust and inexpensive and it can be coupled to gas chromatography (GC-MS) and liquid chromatography (LC-MS) (Hübschmann, 2015).

In the TOF mass analyzer, accelerated ions by the electric field travel through a flight tube, and get separated according to the velocity. Ions with a large mass need more time to travel the flight path whereas lighter mass reaches the detector faster (Canas, 2006). In the analysis of mass lower than 20 kDa, TOF can use reflectron mode to correct kinetic energy dispersion and spatial spread in initial ion energy (Murphy, 2016).

1.5 Cross-linking mass spectrometry

Cross-linking mass spectrometry (CLMS) is an useful tool in protein structure studies. The method provides information regarding the spatial position of amino acid residues and complex protein folding. In addition, it is a very sensitive method and protein analysis could be done even at femtomole concentrations (Sinz, 2018). Cross-linking reactions involve the addition of covalent bonds between a cross-linking agent and a pair of protein residues. These covalent linkages are important to derive distance constraints regarding proteins of interest, which are then used to achieve computational modeling of protein structure (Iacobucci et al., 2018).

The selection of cross-linking reagents is fundamental because the cross-linkers target specific functional groups. Nowadays, N-hydroxysuccinimide (NHS) esters (**Fig. 6A**) are the most commonly used cross-linkers since they can target primary amines, especially lysine residues. However, the disadvantage is the production of big peptides if trypsin is involved in an experiment (Holding, 2015). The reagents disuccinimidyl dibutyric urea (DSBU), 1,1-carbonyldiimidazole (CDI), and bis(sulfosuccinimidyl)suberate (BS3) (**Fig. 6 B, C, D**) are also popular as MS-cleavable cross-linkers. The advantage of MS-cleavable cross-linkers is their ability to exhibit specific fragmentation during MS/MS experiments. The characteristic fragment ion patterns make cross-linked peptide assignment easier (Iacobucci et al., 2018; Lu et al., 2018). Cross-linkers consist of

spacer and reactive end groups. Spacer length is important to obtain information about the distance between peptides in 3D structures (O'Reilly & Rappsilber, 2018).

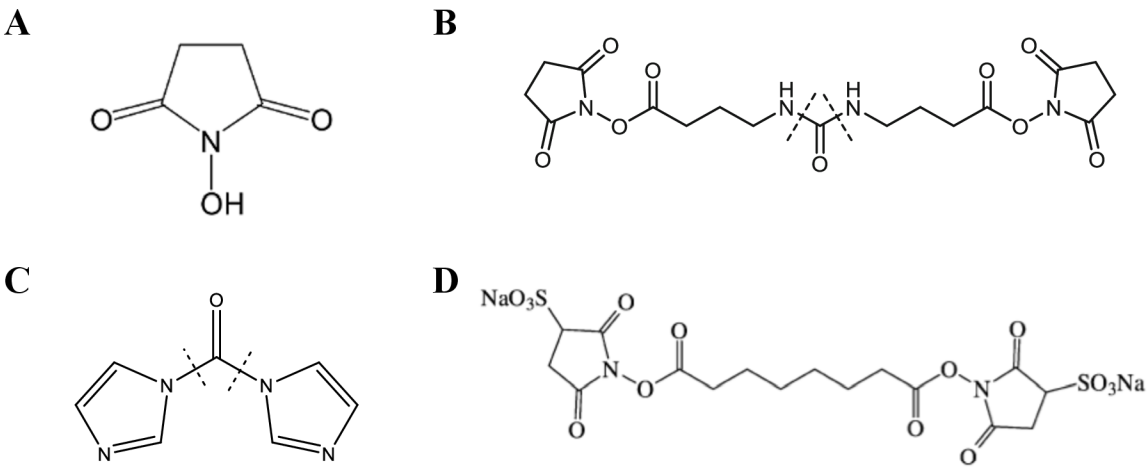

Fig. 6: (**A**) NHS structure (Klykov & Weller, 2015) (**B**) DSBU structure (Piersimoni & Sinz, 2020) (**C**) structure of CDI (Matzinger & Mechtler, 2020) (**D**) BS3 structure (Dihazi & Sinz, 2003).

Generally, a typical cross-linking MS experiment (**Fig. 7**) consists of selecting cross-linker properties (length, functional groups, and solubility) for the required reaction with a protein complex. Once the cross-linker is selected, the protein complex of interest is incubated with reagents in solution followed by quenching the cross-linking reaction and digestion of the protein to obtain peptides. Optionally, cross-linked peptides can be enriched using size exclusion chromatography Afterwards, it is followed by LC-MS analysis. It is possible to identify the pair match of residues in protein complexes according to identified cross-linked peptides from data analysis (O'Reilly & Rappsilber, 2018).

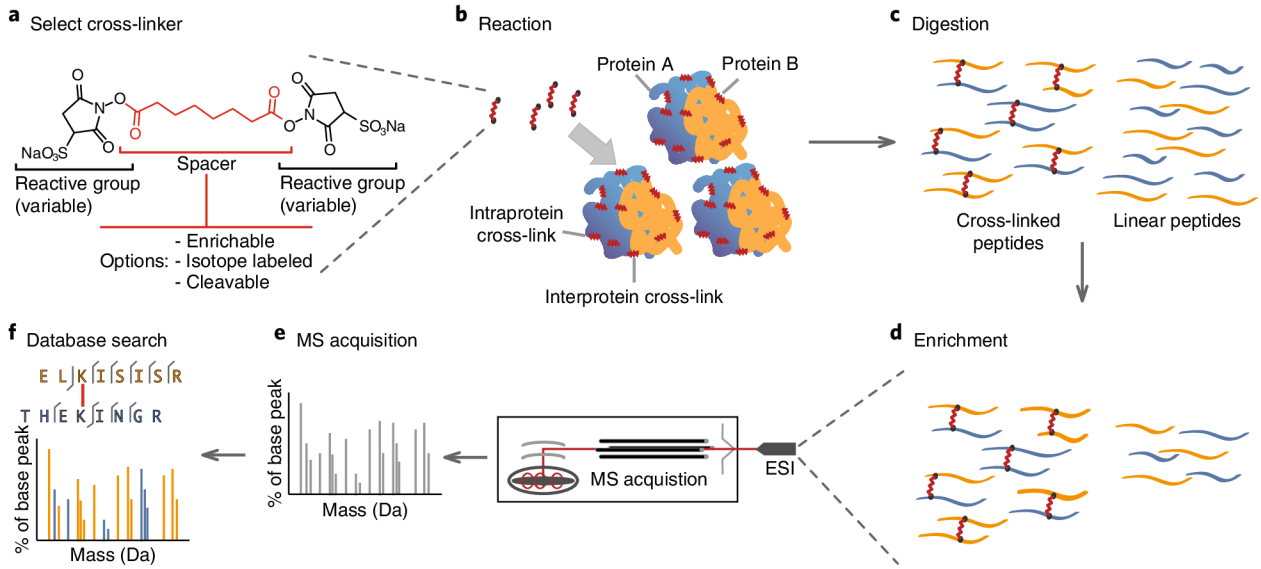

Fig. 7: General workflow of cross-linking MS experiment (O'Reilly & Rappsilber, 2018).

1.6 Hydrogen-deuterium exchange mass spectrometry

Hydrogen-deuterium exchange coupled to MS has become a key technique for the investigation of protein structure dynamics and protein-ligand interaction (Burke, 2019; Konermann et al., 2011; Kostyukevich et al., 2018). In this method, the amide functional group will undergo a hydrogen-deuterium exchange reaction (Hamuro et al., 2003). The deuterium is incorporated into the backbone amide group during protein labeling time (**Fig. 8**). The centroid mass is then shifted according to the deuterium content in the protein (Guttman et al., 2016). Factors affecting the rate of exchange depend on pH, temperature as well as protein structure and dynamics (Narang et al., 2020; Radou et al., 2014). This means that hydrogen bonds of protected parts in protein tertiary structures undergo slower exchange because they are not accessible to the solvent (Narang et al., 2020). Regions of protein occluded by a ligand are slowly exchanged and hidden compared to parts that are fully exposed to the solvent (Malito et al., 2013). This is due to the bad accessibility of amide hydrogen to the water or structural rigidity of the backbone protein which causes low deuterium incorporation in the presence of a ligand (Eisinger et al., 2017). The rate of exchange can be calculated from MS data analysis at different labeling times after exposure to deuterium.

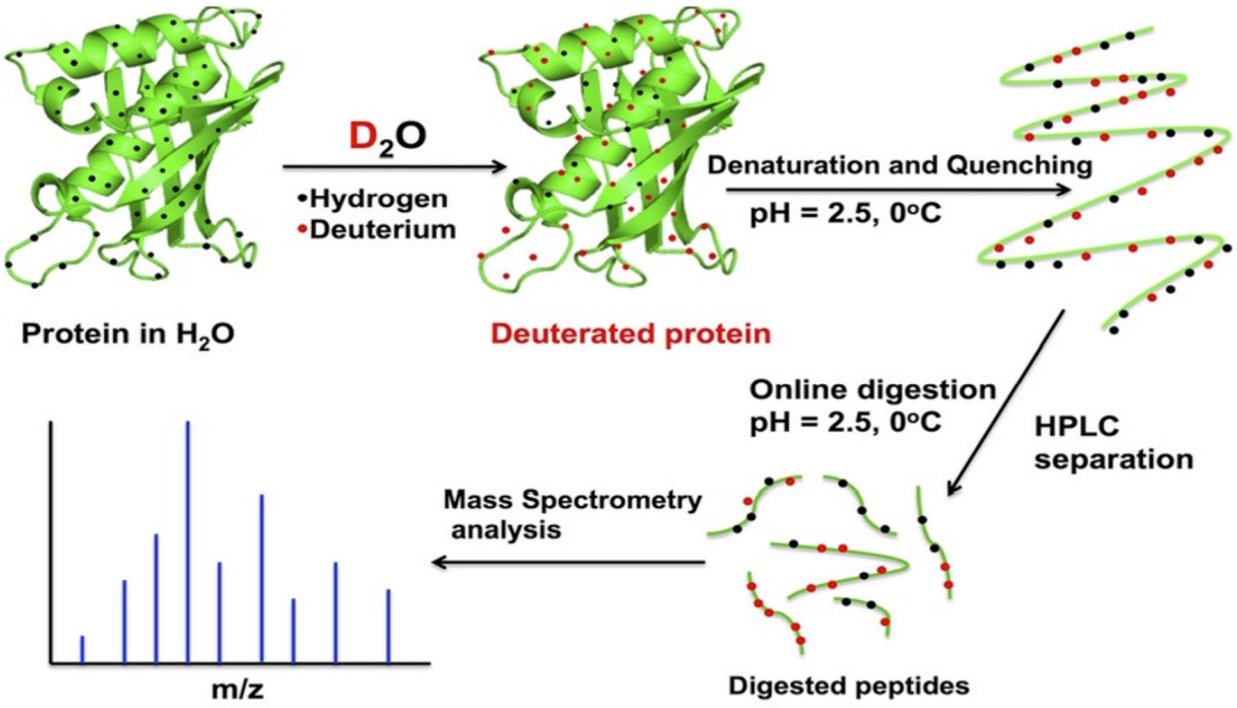

Fig. 8: Hydrogen deuterium exchange working principle (Cao et al., 2013).

In the HDX method, proteins can be subjected to the pepsin column before MS analysis. Pepsin is used to digest the protein and to obtain peptides, which subsequently can be separated by High-performance liquid chromatography (HPLC) column and examined by a mass spectrometer tool

(**Fig. 8**). Pepsin is commonly used because it works better in acidic conditions with a low temperature such as 0 °C (Wei et al., 2012). The deuterium exchange rate is determined at the peptide level to increase spatial resolution. The advantages of HDX coupled to MS are the simple sample preparation, low protein concentration, and the possibility to use a wide range of buffers for deuterium exchange. However, it is important to control parameters that influence the deuterium exchange rate such as pH and temperature. The protein labeling can be performed in a wide pH range but then the reaction has to be quenched at a low pH and low temperature. Short analysis time, low pH, and temperature have to be used to avoid back exchange of deuterium for hydrogen in a protein.

2 AIMS OF THESIS

This thesis aimed to study the structure of viral σNS protein. The purpose of this work was to understand the formation of σNS oligomer and provide a better understanding of σNS and RNA interactions.

In summary, the goal of this project was to investigate σNS protein using different mass spectrometry methods:

1. Employing the MALDI MS to control the quality of purification and stability of σNS protein.
2. Prediction of σNS protein structure using available softwares.
3. Characterization of σNS 3D structure and subunit interactions in oligomer using CLMS method.
4. Determine σNS structure and its changes upon RNA binding by HDX MS technique.

3 MATERIALS AND METHODS

3.1 Chemicals

Acetonitrile (C_2H_3N; for LC-MS, ≥ 99.9%) was purchased from Honeywell Research Chemicals (ND, USA). Ammonium bicarbonate ($(NH_4)HCO_3$; P.A., ≥ 99%) and sodium chloride (NaCl; P.A., ≥ 99.5%) were obtained from Lach-Ner, Czech Republic. Deuterium oxide (D_2O; 99.9%), dimethylated trypsin from porcine pancreas (proteomic grade), dipotassium phosphate (K_2HPO_4; ACS reagent, ≥ 99%), DL-dithiothreitol (DTT; ≥ 99%), formic acid (CH_2O_2; ACS reagent, ≥ 98%), guanidine hydrochloride (GHCl; ≥ 98%), HEPES (≥ 99.9%), iodoacetamide (C_2H_4INO; ≥ 99%), magnesium dichloride ($MgCl_2$; ≥ 98%), potassium dihydrogen phosphate (KH_2PO_4; ACS reagent, ≥ 98%), trifluoroacetic acid (TFA; ≥ 99%), and Tris (≥ 99.9%) were purchased from Sigma Aldrich (MO, USA). Methanol (CH_3OH; for HPLC, ≥ 99.8%) was purchased from VWR Chemicals (PA, USA). All HDX experiments were performed with RNA (polyU 10 bases) obtained from Integrated DNA Technologies (IA, USA).

3.2 Preparation of σNS protein

The protein σNS with SUMO and His-tag has been prepared by a colleague Mgr. Barbora Kaščáková from Department of chemistry, Faculty of Science, the University of South Bohemia in České Budějovice. The main steps involved in the experiment were protein expression, cell lysis, affinity chromatography, sucrose gradient, and size exclusion chromatography to purify the protein.

Briefly, protein expression was done in *E. coli* strain BL21 (New England Biolabs, MA, USA) in LB medium (Merck, Germany) with Kanamycin at 37°C overnight. Protein over-expression was induced by 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG; Carl Roth, Germany) after reaching OD600 of 0.6 and harvesting cells 4h after induction. Next, extraction of protein from cells was performed using a lysis buffer (50 mM Tris, 500 mM NaCl, 5mM MgCl2, pH 7.6) containing protease, RNAsel, and DNAseA inhibitors (Merck, Germany). For cell disruption, a french press (Fisherbrand) was used and debris was removed by ultracentrifugation at 40 000 g at 4 °C for 60 minutes. Affinity chromatography method, using His-tag purification column (I.D. 16 mm, length 25 mm, avg. part. size 34 µm; Merck, Germany) on AKTA PURE M2 systems (GE Healthcare, NY, USA) was required for the protein purification step, followed by sucrose gradient fractionation (5-60%) with ultra-centrifugation (40 000 g at 4 °C for 24 h). The last step in protein purification was size-exclusion LC on AKTA PURE M2 systems using Superdex 200 Prep Grade column (avg. part. size 34 µm; Merck, Germany).

3.3 Protein purification using ZipTip

During the MS analysis of the samples, an important part is the purification of proteins to remove salts and other impurities. The ZipTip C4 (4 carbon chain; Merck Millipore, MA, USA) would be used for this purpose. A first step during the preparation of ZipTip C4 is to wet resin in 50% acetonitrile with 0.05% trifluoroacetic acid (TFA) and then wash it with 0.1% TFA three times. Afterward, the protein is diluted in 2 M guanidine hydrochloride (GHCl) with 0.1% TFA is loaded onto the ZipTip C4. The use of GHCl was critical because of protein precipitation in 0.1% TFA. Resin with loaded protein is then washed in 5% MeOH with 0.05% TFA. It is followed by the elution of protein from C4 by 6 µl of 50% acetonitrile with 0.05% TFA solution, which is the first fraction subjected to MS analysis. The second step of elution was performed in 70% acetonitrile with 0.03% TFA and the obtained second fraction was also analyzed using MALDI-TOF MS.

3.4 Intact mass MALDI MS measurement

Purified proteins (see section 3.3) were spotted onto an MSP AnchorChip^{TM} 384 target plate (Bruker Daltonik, Bremen, Germany) and mixed with sinapinic acid (Bruker Daltonik, Bremen, Germany) (30 mg/ml in 70% acetonitrile with 0.03% TFA). The dried droplet method was used for sample preparation, followed by mass spectrometric measurements on an Autoflex Speed MALDI-TOF/TOF (Bruker Daltonik, Bremen, Germany). Mass spectra were acquired in the positive linear ion mode in the mass range 20-100 kDa using a pulsed extraction laser with an acceleration voltage of 19.42 kV, an extraction voltage of 17.67 kV, a lens voltage of 7.49 kV, and a delayed extraction time of 450 ns. The measurement in the mass range 5-20 kDa was also performed with a similar condition but an extraction voltage was set up at 18.12 kV, a lens voltage at 7 kV and a delayed extraction time was 290 ns. The resulting spectra were accumulated from up to 500 laser shots.

3.5 Cross-linking reaction of protein σNS

Protein was cross-linked using bis(sulfosuccinimidyl)suberate (BS3, Thermo Fisher Scientific, MA, USA) agent. The protein was stored in Tris buffer which reacts with BS. Therefore, Tris had to be removed before the experiment. Protein in Tris buffer was loaded onto Microcon 10 centrifugal filter (Merck, Germany, Darmstadt). Afterward, a centrifugation step was performed (10 000 *g* at 4 °C) to decrease buffer volume to a minimum, followed by adding 500 µl of HEPES buffer to the protein and repeating centrifugation another time with the same conditions. In total, the washing step using the HEPES buffer was repeated twice. The addition of a cross-linking agent to 10 µM protein in 100 molar excesses followed. Cross-linking reaction was performed in 20 mM

HEPES buffer containing 150 mM NaCl at room temperature and the reaction was stopped after 45 min by addition of the same volume of 100 mM ammonium bicarbonate.

After cross-linking experiments, the protein was reduced with 10 mM dithiothreitol (DTT) at 56 °C for 40 min and alkylated with 55 mM iodoacetamide in the dark for 20 min. Alkylation was stopped in 50 mM DTT for 20 min. Finally, trypsin was added to the protein at the ratio of 1:50 and the sample was incubated at 37 °C overnight. The digestion step was terminated by the addition of formic acid to a final concentration of 5%. The obtained peptide mixtures were purified using C18 Empore™ disks (3M, USA, St. Paul) (Rappsilber et al., 2007).

Peptides were dissolved in 30 µl 3% acetonitrile/0.1% formic acid. The peptide separation was carried out on an UltiMate™ 3000 RSLCnano system (Thermo Fisher Scientific, MA, USA) online coupled to mass spectrometer timsTOF Pro (Bruker Daltonics, Bremen, Germany). The injection volume was 2 µl with a flow rate of 2.5 µl/min of 2% acetonitrile/0.1% formic acid. The sample was loaded onto an Acclaim™ PepMap™ 100 C18 trapping column (300 µm i.d., 5 mm length, particle size 5 µm, pore size 100 Å; Thermo Fisher Scientific) for 2 min followed by elution step of peptides from the trapping column onto an Acclaim™ PepMap™ 100 C18 analytical column (75 µm i.d., 150 mm length, particle size 2 µm, pore size 100 Å; Thermo Fisher Scientific) and separated by a 48 min long linear gradient of 5-35% acetonitrile/0.1% formic acid at a constant flow rate of 0.3 µl/min. The temperature of column oven temperature was set at 35 °C and the data were acquired in PASEF scan mode with positive polarity. Electrospray ionization was performed using a CaptiveSpray (Bruker Daltonics, Bremen, Germany) with capillary voltage at 1500 V, dry gas at 3 l/min, and dry temperature at 180 °C. Ions were accumulated for 100 ms and 10 PASEF MS/MS scans were acquired per topN acquisition cycle with an ion mobility range (1/K0) set at 0.6-1.6 Vs/cm^2. Mass spectra were collected over an *m/z* range of 100 to 1700. A polygon filtering was applied to exclude the low *m/z* of singly charged ions. The collision energies were ramped from 20 to 59 eV in 5 steps of equal width between 0.6 and 1.6 Vs/cm^2 of 1/K0 values.

Mascot generic files (MGF) were obtained from timsTOF raw data in Compass Data Analysis software version 5.3 (Bruker Daltonics, Bremen, Germany). These, together with fasta files with protein sequence were loaded to the MeroX software version 2.0.1.4 (Iacobucci et al., 2018). The search was performed with default settings for BS3/DSS cross-linker but the minimum peptide score was changed to 1, and deisotoping was turned off. Precursor precision was set at 5 ppm and fragment ion precision at 10 ppm. Lower and upper mass limits were set at 1000 and 6000 Da, respectively. The minimal signal-to-noise ratio to accept the signal was 2 and the minimum number of fragments per peptide was 3. All ions with charges lower than 2 were ignored. False discovery rate (FDR) cut off was set up at 5% and a decoy database was created as a shuffled sequence but

keeping protease sites. Cysteine carbamidomethylation and methionine oxidation were used as static and variable modifications, respectively.

3.6 HDX MS analysis

For the HDX experiment 20 mM Tris, 50 mM NaCl, 5 mM $MgCl_2$, (pH/pD 7.6) was used as the equilibration and labeling buffer. All pH values were corrected ($pD = pH_{read} + 0.4$) from the actual pH reading (Li et al., 2014). The HDX reaction was initiated by mixing 4 µl of protein solution with 56 µl of D_2O labeling buffer and the mixed solution was incubated at 20 °C for a predetermined set of times. The exchange reaction was then stopped by transferring of 50 µl labeled protein solution to 50 µl quench buffer (50 mM phosphate buffer, pH 2.6) kept at 0 °C. It was important to keep protein solution at low pH and low temperature from this step to avoid back exchange to hydrogen.

The sample was then injected into the HDX manager system (Waters, MA, USA) and online digestion of protein was performed by Enzymate™ BEH Pepsin Column (Waters) connected to Acquity UPLC M-Class system (Waters) with a constant flow rate of 100 µl/min 0.2% formic acid. Obtained peptides were trapped onto an Acquity BEH C18 Vanguard Column (1.7 µm; Waters) for 2 min followed by elution of peptides from the trapped column onto an Acquity BEH C18 analytical column (1.7 µm, 1 mm i.d., 100 mm length; Waters) and separated by a 7 min long linear gradient of 8-35% acetonitrile/0.2% formic acid with flow rate 40 µl/min. Automation of protein deuterium labeling and injection into UPLC system is provided by PAL RTC system (LEAP technologies, USA). Through the separation, the columns were kept at low temperature in HDX manager chambers, and finally, the peptides were analyzed by Synapt G2-Si mass spectrometer (Waters). Within the ESI source, a capillary voltage of 3 kV was used, the sampling cone voltage was 40 V, and the source offset was 80 V. The source temperature was set up at 100 °C and desolvation temperature at 250 °C. The cone gas flow was 50 l/h, the flow of desolvation gas was 600 l/h and nebulizer gas were set up at 6.5 bar. Measurement was performed in MS^E or MS mode in positive polarity and MS^E mass data were acquired in low- and high-energy modes. The trap collision energy was ramping from 20 up to 40 V in high energy mode.

The acquired data were submitted for processing and database searching using ProteinLynx Global Server software (PLGS) version 3.0.3 (Waters) against an in-house prepared database containing σNS protein sequence supplemented with common contaminants (Max Planck Institute of Biochemistry, Martinsried, Germany). The low and high energy threshold values were set to 200 and 40 counts, respectively. The parameters used for database search were: enzyme specificity: non-specific, variable modification: methionine oxidation, false discovery rate: 4, min. fragment

ion matches per peptide: 1, min. fragment ion matches per protein: 7, min. peptide matches per protein: 3.

For HDX mass data analysis was used DynamX software version 3.0 (Waters) in which the exported peptide tables from PLGS and raw data were loaded. Only peptides of σNS with a minimum intensity of 500, the minimum sequence length of 5, and the minimum number of products of 2 (number of fragments in MS^E) were processed. DynamX assigns the ions from raw data to peptides but it was required a manual correction of all data in software. The deuterium uptake plots were visualized in DynamX software to show the deuteration level of identified peptides. Relative fractional uptake in all labeling times was exported from DynamX and the data were loaded to protein structure in pyMOL (2.5.2 version).

3.7 Softwares for prediction and visualization of protein structure

Rosetta software was used to obtain the 3D structure of protein σNS (σNS model) (Baek et al., 2021). Phyre2 software was used to predict a structure of σNS containing His-tag and SUMO (σNS SUMO model) (Kelley et al., 2015). Both structural models of protein were visualized in pyMOL software (2.5.2 version). To draw cross-links in pyMOL a plugin PyXlinkViewer was used (Schiffrin et al., 2020). The software Xinet was employed for a schematic overview of crosslinked peptides (Combe et al., 2015).

4 RESULTS

4.1 Prediction of σNS protein structure

Protein σNS containing His-tag and Small-Ubiquitin Modifier (SUMO) at N-terminal site was used in experiments (**Fig. 9**). The theoretical average mass and isoelectric point (pI) of the protein was calculated in Compute pI/MW tool of the expasy.org web page (**Fig. 9, Fig. S1**). The sequence of σNS SUMO construct was compared to proteins in the non-redundant database using Basic local alignment tool (BLAST) and it was found an identity of 99.18% with σNS protein of avian *orthoreovirus* ARV. His-tag and SUMO are located in the protein on the N-terminal site till amino acid position 119.

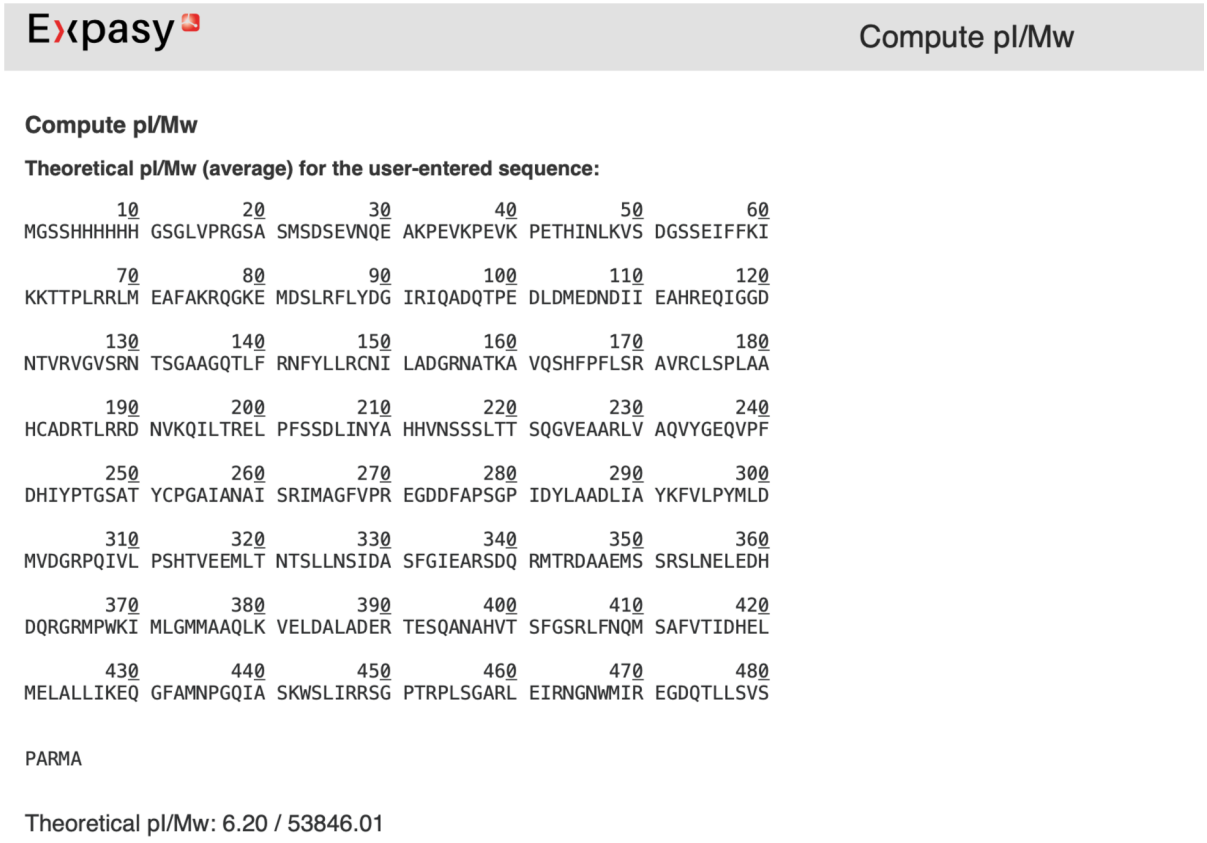

Fig. 9: The average mass and pI of protein σNS SUMO computed in expasy.org.

Although σNS structure is not known, it is possible to predict its interaction and construction using models of 3D protein structure which can be obtained by combining the 1D and 2D approaches (Baek et al., 2021). In our study, we used two different software to predict the structure of σNS. Protein sequence missing His-tag and SUMO part was loaded into Rosetta software (**Fig. 10A**) (Baek et al., 2021). The prediction of σNS His-tag SUMO protein was performed in Phyre2 (**Fig. 10B**) (Kelley et al., 2015). The sequence analysis in Rosetta Fold software was done by the colleague Mgr. Barbora Kaščáková. We investigated whether obtained σNS and σNS SUMO

models are similar at secondary structures. What has been found is that both models (**Fig. 10 A&B**) are very different even at the secondary level.

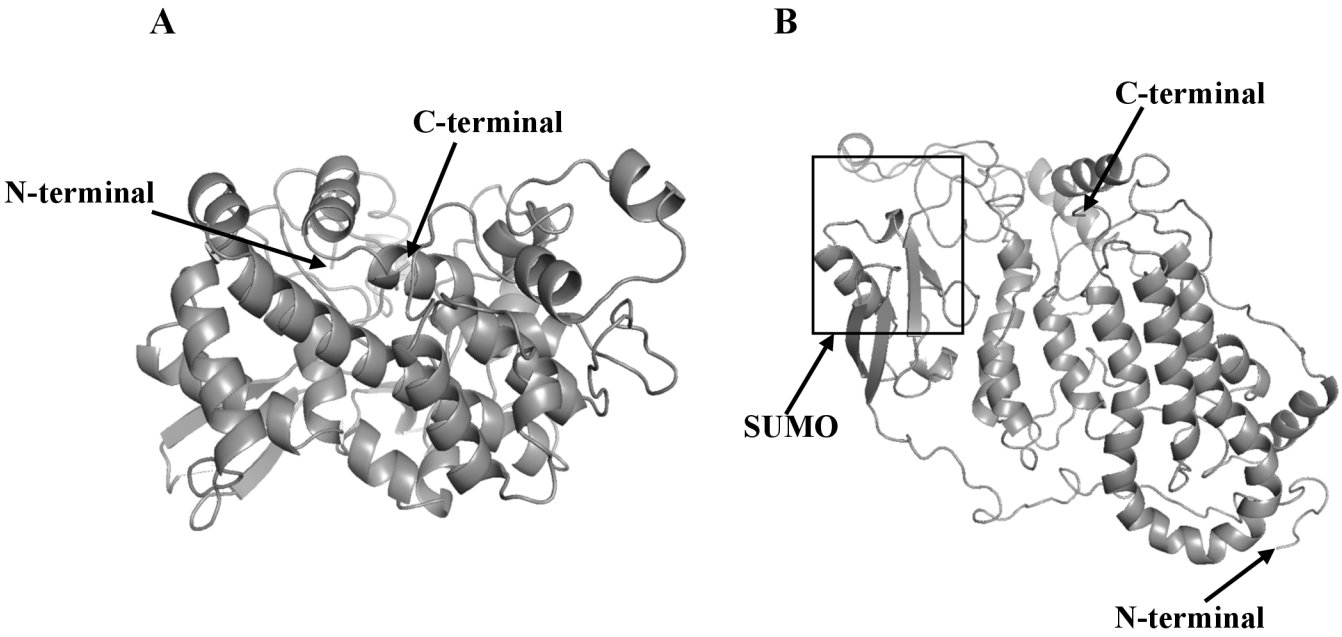

Fig. 10: Prediction of (**A**) σNS and (**B**) σNS SUMO protein structures obtained by Rosetta and Phyre2 software, respectively.

4.2 Estimation of intact protein mass by MALDI MS

Determination of intact mass by MALDI MS is a fast method to inspect the composition of proteins in a sample. The great advantage of MALDI MS is the possibility of analyzing high mass proteins even more than 100 kDa with good accuracy and a detection limit (Gatlin-Bunai et al., 2007). Therefore, the mass spectrometer tool can confirm the presence of a protein of our interest, its purity, and its properties such as protein modification and artificial fragmentation.

Protein σNS has been expressed and purified by liquid chromatography, particularly, the metal-chelating system specific to His-tag bound to our protein and the size exclusion method where the molecules are being separated according to their size. Protein construct contained His-tag and SUMO. The presence of His-tag in the protein can help in the purification step. The role of the SUMO is the capacity to attach covalently within the substrates and modify the function of a protein, such as solubility. In the experiment two fractions of σNS were prepared, a protein sample containing the SUMO and σNS after cleavage of His-tag and SUMO by protease. The MALDI MS technique was used to examine the samples, especially to reveal if the SUMO part was cleaved. In addition, the stability of the protein was controlled by MALDI MS before cross-linking and HDX

experiments. The protein expression, purification, and SUMO cleavage were performed by our colleagues Dr. Zdeněk Franta and Mgr. Barbora Kaščáková.

The proteins were purified using ZipTip C4 before MS analysis. Two fractions using different content of acetonitrile in the elution step (50% or 70% acetonitrile in 0.1% TFA) were collected and spotted on a MALDI plate with sinapinic acid as a matrix. The less hydrophobic proteins should be eluted by 50% acetonitrile whereas the more hydrophobic species would stay at resin until 70% acetonitrile is used. We observed no significant differences between the two fractions (**Fig. 11**). Therefore, only 70% acetonitrile should be preferred in the purification step using ZipTip C4.

The measurement of spectra was performed by Autoflex Speed MALDI-TOF/TOF mass spectrometer in a linear model. Furthermore, we measured in two mass ranges (5-20 kDa and 20-100 kDa) but ions were detected in the 20-100 kDa range. Three ion signals were observed in the mass spectra of the σNS containing SUMO (**Fig. 11 A, B**). The first was around 27 000 *m/z*, the second at 41 000 *m/z*, and the last at 54 000 *m/z*. The calculated mass of σNS with and without SUMO is 53.8 kDa and 40.4 kDa, respectively. The mass of the third ion signal corresponds to the theoretical mass of σNS with SUMO. For higher mass molecules the doubly and triply charged ions are commonly visible in the MALDI mass spectrum. Therefore, we supposed that the ion signal at 27 000 *m/z* is doubly charged σNS with SUMO which corresponds to the mass of 54 kDa.

The ion signal of the lowest intensity corresponds to the mass of the σNS without SUMO, over 40 kDa Therefore, we can expect that σNS without SUMO is present in a small amount in the sample. This impurity might depend on incorrect purification or degradation of the protein SUMO. On the other hand, only two ion signals had been observed in the mass spectra of the σNS without SUMO (**Fig. 11 C, D**). The ion signal at 20 300 *m/z* and *m/z* around 40 600 corresponds to σNS molecule after cleavage of SUMO in the charge state 2+ and 1+, respectively.

We confirmed that the analyzed fractions contained purified σNS as the only dominant molecule in the mass spectra (**Fig. 11**). The sample of σNS SUMO protein contains predominantly protein with His-tag and SUMO parts. But we also observed that the σNS SUMO protein was partially cleaved or other impurities were present in the sample in a low amount. In contrast, only σNS was detected in the sample after cleavage of His-tag and SUMO.

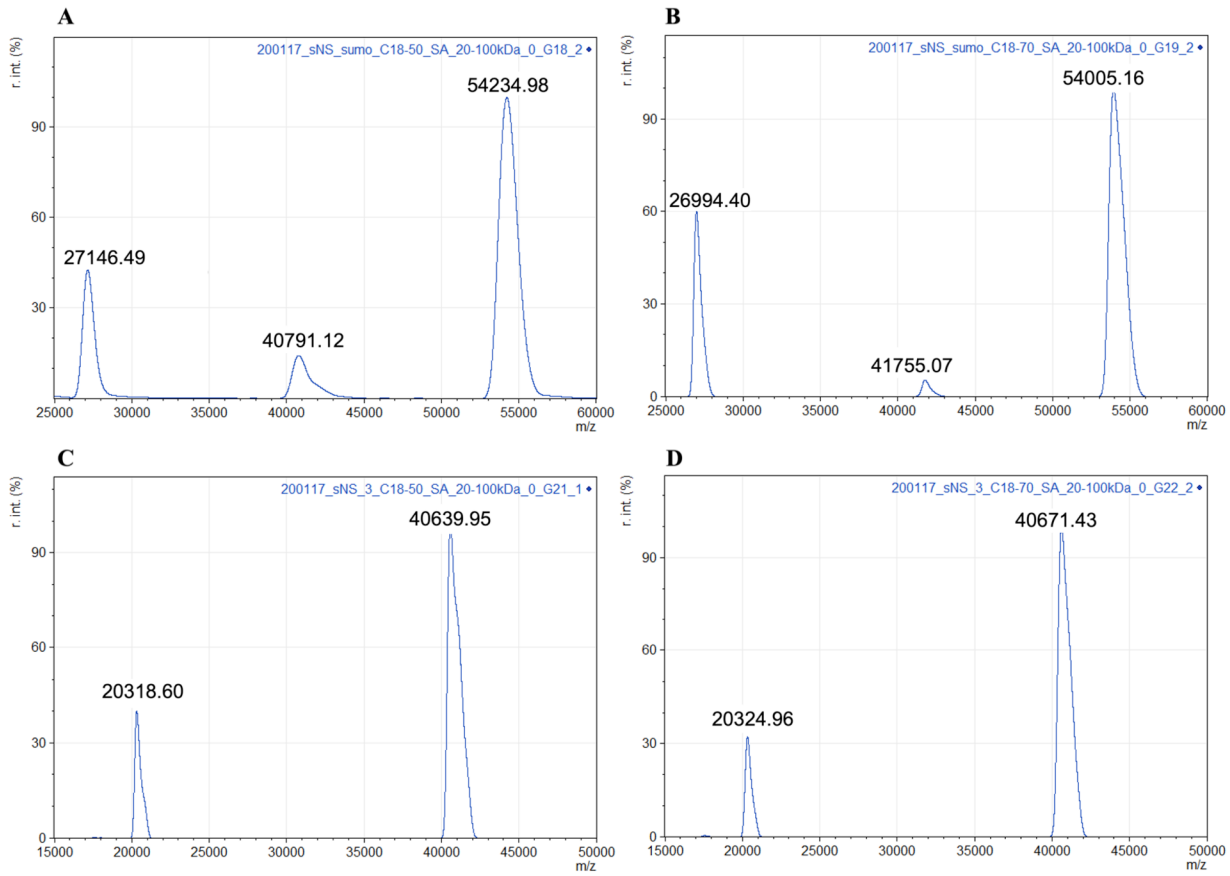

Fig. 11: MALDI-TOF mass spectra of σNS protein (**A, B**) and protein containing His-tag and SUMO (**C, D**). The spectra were obtained after ZipTip C4 purification and protein elution in 50% (**A, C**) and 70% (**B, D**) acetonitrile, respectively.

4.3 Study of σNS protein by cross-linking mass spectrometry

To determine the protein complex folding and position of amino acid residues, CLMS can be used for the analysis which involves the selection of a cross-linker to incubate with proteins. The proteins can be digested and obtained peptides are analyzed by MS. The cross-linked peptides will be examined by software which then detects a pair of amino acid residues position.

During our experiment, we employed MeroX as a tool for the identification of cross-linked peptides. The MeroX software analyzes the MS/MS signal and calculates the score which is determined by the intensity of detected ions and length of peptides. From this, the crosslinked residues position of two peptides can be obtained (Iacobucci, 2018). We identified in total 30 peptides (see **Table S1**) but only seven unique cross-linked peptides (**Table 1**) in MeroX software using a 5% FDR cut-off. All cross-linked residues were located in range 79-193 of position in protein σNS SUMO (**Fig. 12A**). According to this, six were assigned as intra-protein cross-links (**Fig. 12A**) and one homotypic cross-linker at position 193. However, it is not possible to link

residue with itself by BS3. Therefore, we suggest that it is an inter-molecular cross-linker between two proteins σNS SUMO, which provides information regarding protein-protein interaction.

Table 1: List of unique cross-linked peptides identified in MeroX software

n.	Peptide sequence 1	Residue 1	Peptide sequence 2	Residue 2	Type	Score	Distance σNS model [Å]	Limit [Å] S/V	Distance σNS SUMO model [Å]	Limit [Å] S/V
1	NATKAVQSHFPFLSR	159	QGKEMDSLR	79	Intraprotein	112	-	-	22	S
2	NATKAVQSHFPFLSR	159	BLSPLAAHBADR	176	Intraprotein	101	17.2	S	22.3	S
3	BLSPLAAHBADR	176	DNVKQILTR	193	Intraprotein	98	25.1	**V**	17	S
4	NTSGAAGQTLFR	131	QGKEMDSLR	79	Intraprotein	83	-	-	31.4	**V**
5	RDNVKQILTR	193	DNVKQILTR	193	**Homotypic**	75	0	S	0	S
6	RDNVKQILTR	193	NTSGAAGQTLFR	138	Intraprotein	63	47.1	**V**	46	**V**
7	AVQSHFPFLSR	169	QGKEMDSLR	79	Intraprotein	44	-	-	24.9	S

The data obtained from MeroX were loaded to the PyXlink Viewer plugin (Iacobucci et al., 2018; Schiffrin et al., 2020) in pyMOL to get the insight of cross-linked residues in the 3D structure of σNS (**Fig. 12 B&C**). Although, the protein structure of σNS is still not known. Therefore, σNS and σNS SUMO protein structures were only predicted using Rosetta and Phyre2, respectively (Baek et al., 2021; Kelley et al., 2015). Distance between protein residues in 3D models was calculated by PyXlink Viewer (see **Table 1**). We set a threshold of 25 Å, as an expected maximum length of BS3 cross-linker and lysine residues are 24 Å (Merkley et al., 2014).

In the σNS model, three crosslinks were not displayed in the 3D structure because they link σNS and SUMO part which was not subject to prediction (**Table 1, Fig. 12B**). The distance between cross-linked residues of two peptide pairs, (linked 193-193 and 159-176 residues) were below the threshold limit (**Table 1**, **Fig. 12B**). However, the other two cross-linkers linking residues 176-193 and 193-138 had a distance above 25 Å (**Table 1, Fig. 12B**). In the σNS SUMO model, the distance between cross-linked residues of five out of seven cross-linked peptides was within the threshold (**Table 1, Fig. 12C**). Although two cross-links, linking residues 193-138 and 131-79 had a distance of 46 and 31.4 Å, respectively (**Table 1, Fig. 12C**).

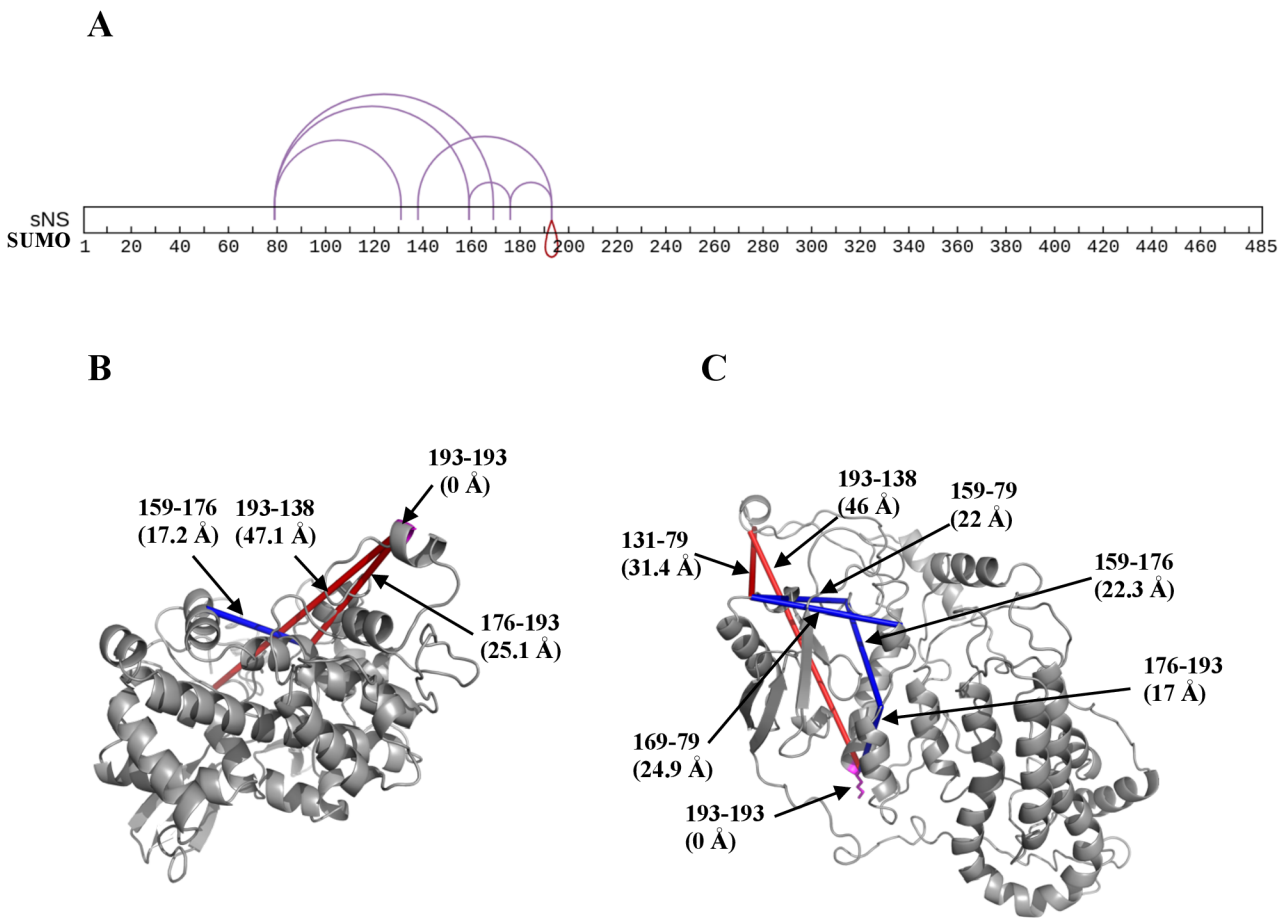

Fig. 12: (**A**) Schematic overview of cross-linked peptides of σNS SUMO in Xinet software (**B**) 3D structure of protein σNS and (**C**) σNS SUMO model, depicts all crosslinked peptides. Blue sticks indicate that the distance is below 25 Å, whereas red means a length greater than the expected threshold. The self-linker Lys-193 is depicted with magenta color.

4.4 Protein-RNA complex analysis using hydrogen deuterium exchange mass spectrometry

The protein σNS and its interaction with RNA had been analyzed by the HDX MS technique which can give important information about protein conformation since the solvent accessibility of backbone amide and hydrogen bonds affect the rate of exchange (Li et al, 2014).

During the optimization step, several protein concentrations were used for the injection of protein into the HDX system for MS analysis. The highest protein coverage was observed for 5 µM or higher protein concentrations. Therefore, we used a 5 µM protein concentration and three times the molar excess of RNA (PolyU, 10 bases) in stock solution. The stock solution was mixed with D_2O buffer in ratio 4:56 with a final concentration of protein to be 333 nM during the HDX experiment. Labeling was performed at various times from 10 seconds up to 8 hours, which was used for the deuterium incorporation comparison in apoprotein and σNS-RNA complex. To identify proteins

and peptides in a sample, the raw data were analyzed using PLGS software and the obtained peptide list was uploaded then to DynamX. The total number of identified peptides of σNS protein was 220 with a protein coverage of 94.6%. However, only 64 peptides were successfully assigned in DynamX software and protein coverage was 81% for the σNS-RNA complex (**Fig. 13**). The first 119 amino acid residues are part of His-tag sequence SUMO in the σNS protein construct.

Fig. 13: Protein coverage for state-comparison σNS-RNA complex. Only peptides assigned in DynamX software are displayed.

Deuterium uptake plots were generated in DynamX software to obtain information regarding HDX kinetics for all identified peptides calculated from the mass shift of isotope centroids at different labeling times (**Fig. 14**). We investigated deuterium incorporation differences between σNS protein alone and after adding RNA and also mapped the changes of relative fractional deuterium uptake at various labeling times on models of 3D protein structure in pyMOL software (**Fig. 15 & S2**).

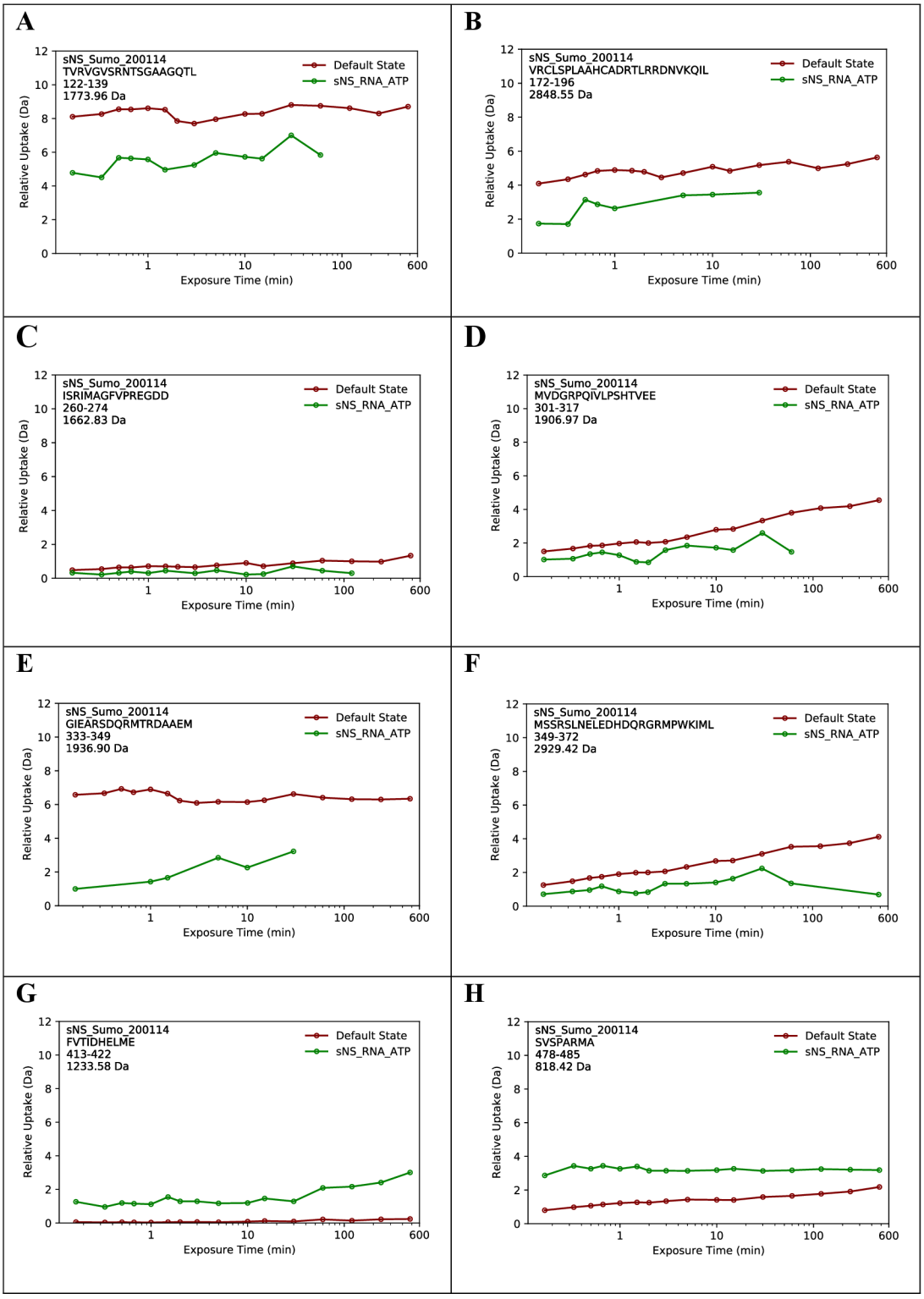

Fig. 14: HDX kinetics plots as a function of time for 8 different peptides: (**A**) 122-139, (**B**) 172-196, (**C**) 260-274 (**D**) 301-317, (**E**) 333-349, (**F**) 349-372, (**G**) 413-422, (**H**) 478-485.

According to HDX kinetics, we observed towards the N-terminal part of the peptide (residues 122-139) a very fast exchange and a high deuterium uptake for σNS and σNS-RNA complex reaching maximum level already at 10 seconds (**Fig. 14 A**). Although, in presence of RNA exhibited a reduction in deuterium incorporation compared to σNS protein (**Fig. 14 A; Fig. 15 A&B**).

The peptide (residues 172-196) showed a very fast exchange rate and moderate deuterium uptake for the protein alone (**Fig. 14 B**). The α-helix spans from residues 177-185 in the N-terminal part. This region exhibited high relative fractional uptake in the absence of RNA (**Fig. 15 B**). However, the C-terminal part of α-helix (residues 182-185) and the following loop (residues 186-192) demonstrated similar deuterium incorporation in σNS and σNS-RNA complex (**Fig. 15 B**).

Very slow HD exchange was obtained for peptide at residues 260-274 throughout the labeling periods (**Fig. 14 C**). In the σNS and σNS SUMO models, the residues 260-264 of α-helix that was solvent inaccessible in the absence of RNA also showed no change in deuteration levels in the presence of RNA (**Fig. 15 A&B**).

A slow and gradual increase in deuterium uptake was noticed on peptide fragment 301-317 for σNS (**Fig. 14 D**). In the σNS model, the β-strand between residues 298-306 and α-helix from 315-319 demonstrated no significant difference in deuterium incorporation for σNS with and without RNA (**Fig. 15 A**).

Towards the C-terminal part, a very rapid exchange rate (below 10 s), and very high deuterium uptake were observed for peptide (residues 333-349) in σNS protein. In contrast, the peptide had a low deuterium content in the σNS-RNA complex (**Fig. 14 E**). The loop (residues 333-338) and adjacent part of α-helix (residues 339-342) demonstrated a very high level of deuterium accumulation in the σNS state. A significant reduction in deuterium incorporation to α-helix (residues 339-342) was detected after adding RNA to the protein (**Fig. 15 B**). A similar but lower difference between relative fractional uptake of σNS and σNS-RNA states was observed at the following part of α-helix (residues 343-354) (**Fig. 15 B**). No significant difference in the deuteration level of σNS and σNS-RNA complex was noticed at the adjacent loop (residues 355-364) and α-helix (residues 365-372) (**Fig. 15 B**). Moreover, the peptide covering this region (residues 349-372) showed a gradual increase in deuterium uptake in the presence and absence of RNA (**Fig. 14 F**). However, deuterium accumulation declined in the peptide of the σNS-RNA complex after a long time (more than 30 min) (**Fig. 14 F**).

A very low deuterium uptake was observed in the peptide at the C-terminal part (residues 413-422) in the σNS state, but a slight increase in deuterium content was detected after adding RNA (**Fig. 14 G**). The α-helix (residues 413-422) demonstrated high deuterium incorporation for σNS-RNA compared to σNS protein alone (**Fig. 15 A&B**). We observed a very rapid deuterium exchange rate of C-terminal peptide (residues 478-485) for the σNS-RNA complex. In contrast, deuterium content gradually increased over time in the peptide in the σNS state. The faster kinetics in the σNS-RNA complex described the higher stability of the C-terminal peptide in the σNS protein (**Fig. 14 H**). In both structural models, the C-terminal peptide is represented as a loop (**Fig. 15 A&B**). In

the σNS model, the β-strand was predicted as a structure of the previous C-terminal part showing a high difference in deuterium content between the two states (**Fig. 15 A**). In contrast that part was represented as a non-structural loop in the σNS SUMO model (**Fig. 15 B**).

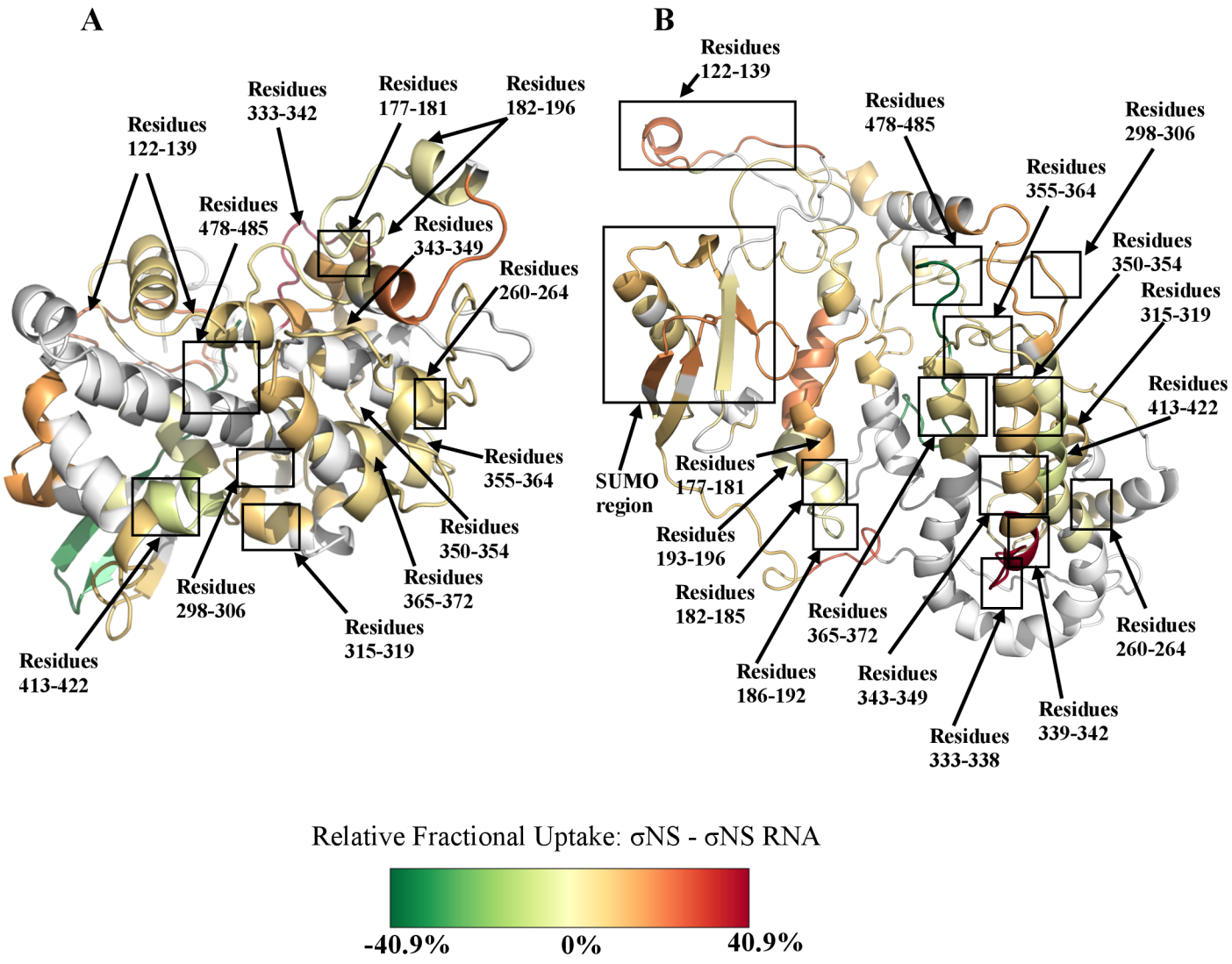

Fig. 15: HDX MS analysis of σNS protein and σNS-RNA complex state comparison in 3D for (**A**) σNS model and (**B**) σNS SUMO model at 60 s HDX labeling time. The color code represents the relative changes in deuterium uptake between protein and σNS-RNA complex states (see legend). The regions in the protein structure colored in white were not detected by HDX.

5 DISCUSSION

Protein σNS was studied using CLMS and HDX MS techniques. Whereas the 3D structure of σNS is unknown, prediction of protein structure can be helpful for further biological investigation. Determination of protein conformation using an experimental method such as X-ray crystallography requires a lot of time and making good crystals for protein is not an easy task (Zheng et al., 2015). To overcome this problem, proteins can be predicted as a structural model using computational programs (Baek et al., 2021; Kelley et al., 2015). In the thesis, models of σNS and σNS His-tag SUMO protein structures were obtained in the Rosetta and Phyre2 software (**Fig. 10 A&B**). These two models were used in the data interpretation of our MS experiment. It was found by native-MS that σNS forms a hexamer as a trimer of dimers (Borodovka et al., 2015). Protein binding to ssRNA leads to an octameric transition from hexamer when σNS is coupled to RNA molecules (Bravo et al., 2018; Borodovka et al., 2015).

To investigate the structure of σNS and oligomeric subunit contacts CLMS method was used. In total 30 cross-linked peptides were identified in the MeroX software but only seven unique peptides were observed in the mass spectra (**Table 1, Table S1**). However, three unique peptides were linked to the SUMO part of the σNS protein (**Fig. 12 A**). Therefore, only four cross-links can show biological relevance in further analysis.

The CLMS analysis was applied to study σNS protein and RNA interactions in the thesis of Bravo (2019). It was found six cross-links and two of them were self-linkers. Although, it is not possible to link amino acids to itself by BS3 cross-linker. Therefore, it was suggested that self-linked residues participate in inter-monomer interactions. This means that between these sites the protein is linked to another subunit-forming oligomer (Bravo, 2019). Similar way, we investigated in our experiment a residue participating in the interaction during the formation of protein oligomers. According to the results, the self-linker Lys-193 is homotypic and we suggested that σNS is linked to another subunit at Lys-193 position forming an oligomer (**Table 1, Fig. 12**). Although protein used in our experiments contained His-tag and SUMO. Therefore, the actual amino acid residue position in the sequence of σNS without the modifier is Lys-75 (**Fig. S1**). Nevertheless, the number of identified unique cross-links was low (**Table 1**); thus, it was a limitation in our study.

We mapped the data obtained in MeroX software into the PyXlink Viewer (pyMOL plugin) to observe cross-linked residues in 3D models (**Fig. 12 B&C**). In the σNS model, only one cross-link is below the distance threshold which was set to 25 Å as the expected maximum length of the BS3 cross-linker (Merkley et al., 2014). Whereas the other two cross-linkers have a distance greater than the expected length (**Fig. 12 B**). To be cross-linked both residue pairs should be accessible

and remain as close as possible in the protein structure (Chen & Rappsilber, 2018). In the σNS SUMO model, four cross-linked peptides (two of them are linked to the SUMO part) showed the distance between residues lower than the threshold. The distance between the two cross-linkers (one is linked to the SUMO part) is above the threshold (**Fig. 12 C**). According to our data, predicted models do not display a final structure of the σNS protein (**Fig. 12 B&C**). Although, the higher distance between some residues can also be explained by subunit interactions in the oligomer. Molecular dynamics simulations can be used for further investigation to predict the protein 3D structure using our results.

The σNS and σNS-RNA complex structures were investigated by the HDX MS method. Protein structure changes upon different conditions can be monitored by the rate of HD exchange and deuterium uptake at different labeling times for the identified peptides. Factors affecting the rate of backbone amide hydrogen exchange are pH, temperature, and ionic strength of buffers (Lento et al., 2015). The HDX kinetics of amino acid residues located in different parts of the protein shows us the local stability of that protein part. Based on the protection pattern, we investigated structural changes in σNS upon RNA binding.

Peptide located in the N-terminal part at position 122-139 demonstrated a very fast exchange and high deuterium uptake for σNS and σNS-RNA complex (**Fig. 14 A**). This indicates that this protein part is flexible (Shkriabai et al., 2014). Deuterium incorporation was slightly lower in the presence of RNA compared to protein alone (**Fig. 14 A**) which means that peptide sequence 122-139 is partially protected in the σNS-RNA complex and fully exposed to solvent in the absence of RNA. The peptide 333-349 exhibited also very high uptake and rapid exchange in the σNS protein indicating that it is flexible and less protected. However, the significant reduction in deuterium incorporation suggests higher local stability and higher protection in the σNS-RNA complex (**Fig. 14 E; Fig. 15 B**).

The α-helix at position 177-181 is less protected in σNS indicating lower local stability compared to protein when bound to RNA (**Fig. 15 B**). The Lys-193 residue which is located on the α-helix demonstrated a slight difference in deuteration level for σNS and σNS-RNA state (**Fig. 15 A&B**). Based on our cross-linking data it is possible to assume that on Lys-193 residue the protein interacts with another subunit forming an oligomer (**Table 1, Fig. 12 A**). We suggest that the presence of RNA influences the Lys-193 subunit binding site as a slightly higher local stability was observed after adding RNA (**Fig. 14 B, Fig. 15**).

The slow and gradual increase of deuterium uptake was noticed for peptide sequences 301-317 and 349-372 (**Fig. 14 D, F**) in σNS and σNS-RNA state. In both situations, the rate of exchange was

slightly lower in presence of RNA compared to protein alone. The decrease of deuterium incorporation for long incubation times (more than 30 min) shows the problem of protein stability if incubated in deuterium buffer with RNA (**Fig. 14 D, F**).

The peptide, residues 260-274 have a very low deuterium uptake for σNS and σNS-RNA complex (**Fig. 14 C**). Therefore, we suggest that the α-helix part and following loop at 260-274 are buried into the protein structure causing very low accessibility of solvent in the absence and presence of RNA (**Fig. 15**).

The σNS in presence of RNA is less protected than protein alone for peptides 413-422 and 478-485 in the C-terminal part. The α-helix (413-422) shows very low accessibility of solvent in σNS (**Fig. 14 G**). Therefore, we hypothesize that it is occluded to the solvent and more protected in absence of RNA. The loop (peptide 478-485) displayed low HD exchange in σNS (**Fig. 14 H; Fig. 15A&B**) which indicates that it undergoes intramolecular interactions if the protein is not bound to RNA. Lower deuterium uptake for protein-RNA complex compared to σNS was observed only in the C-terminal segment. No other part of protein showed similar results.

According to our results, many peptide regions are less protected in σNS protein and more protected in σNS-RNA complex. The research group of Lísal et al. (2006) investigated the P4 protein and suggested that it becomes protected when bound to procapsid (PC) leading to the decreased accessibility of solvent in the P4-PC complex. Our outcome demonstrated that many parts of the protein are less accessible to the solvent when σNS is bound to RNA due to high local stabilization compared to σNS.

In the HDX MS experiment, many peptides were unassigned and the problem mainly occurs after adding RNA. We suggest the presence of RNA negatively influences the digestion of protein and therefore, a lower peptide coverage can be observed. All our experiments were performed with 10 nucleotides long ssRNA. The ideal length of ssRNA is between 10-20 nucleotides for σNS binding (Touris-Otero et al., 2005). Therefore, longer base RNA would be used in future experiments to control the proper protein-RNA binding. In contrast, longer RNA will negatively affect the HDX MS experiment. Furthermore, we observed the lower deuterium uptake for long incubation times of σNS-RNA complex because σNS degraded or precipitated after a long time of incubation in presence of RNA at 20 °C. However, a major factor is data reproducibility since the HDX MS experiment generates a large number of data that has to be analyzed accurately and requires a lot of time (Lento et al., 2015). This affected our analysis since some time points are missing due to non-assigned peptides.

6 CONCLUSIONS

The presented work focused on studying σNS protein using different mass spectrometric methods such as CLMS and HDX. Several different techniques such as X-ray crystallography and nuclear magnetic resonance spectroscopy (NMR) can be used to determine protein structure (Kaufmann et al., 2010). However, the use of these methods is expensive, time-consuming and to fulfill requirements would be hard to achieve. Therefore, an alternative way is to employ sophisticated software to obtain 3D structural models of the molecule. We predicted the 3D structure of the protein in Phyre2 and Rosetta software (Kelley et al., 2015; Baek et al., 2021) for the data interpretation, since the structure of the protein is not known.

In this thesis, we employed MALDI MS to estimate the mass of the protein σNS and consequently control the purity and stability of the protein before further analysis. We investigated the σNS protein structure using CLMS. Seven unique cross-linked peptides were obtained, but three of them targeted the His-tag SUMO part of the protein. The experiment revealed a cross-linker at Lys-193 residue which was linked to itself. This indicates a location where σNS subunit is linked to another σNS subunit forming an oligomer. The protein-RNA interaction was studied using the HDX MS method. Comparative analysis of the protection pattern revealed many regions of higher HDX rate in σNS than in σNS-RNA complex. This suggests the protein or oligomer stabilization upon RNA binding.

Further experiments need to be conducted to explain better the σNS protein and its conformation. In the thesis, we used CLMS only for σNS protein, but, in future experiments, it will be interesting to employ this method for analysis of the σNS-RNA complex to show if similar interactions will be identified in the presence of RNA. However, optimization of the experiment will be required to obtain more cross-linked peptides. Our data set can be used as a support in future experiments employing molecular dynamics simulation to predict the protein structure.

7 REFERENCES

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., Wang, J., Cong, Q., Kinch, L. N., Schaeffer, R. D., Millán, C., Park, H., Adams, C., Glassman, C. R., DeGiovanni, A., Pereira, J. H., Rodrigues, A. V., van Dijk, A. A., Ebrecht, A. C., . . . Baker, D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. Science, 373(6557), 871–876. https://doi.org/10.1126/science.abj8754

Banerjee, S., & Mazumdar, S. (2012). Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *International Journal of Analytical Chemistry*, *2012*, 1–40. https://doi.org/10.1155/2012/282574

Benavente, J., & Martínez-Costas, J. (2007). Avian reovirus: Structure and biology. *Virus Research*, *123*(2), 105–119. https://doi.org/10.1016/j.virusres.2006.09.005

Bodelòn, G., Labrada, L., Martínez-Costas, J., & Benavente, J. (2001). The Avian Reovirus Genome Segment S1 Is a Functionally Tricistronic Gene That Expresses One Structural and Two Nonstructural Proteins in Infected Cells. *Virology*, *290*(2), 181–191. https://doi.org/10.1006/viro.2001.1159

Borodavka, A., Ault, J., Stockley, P. G., & Tuma, R. (2015). Evidence that avian reovirus σNS is an RNA chaperone: implications for genome segment assortment. *Nucleic Acids Research*, *43*(14), 7044–7057. https://doi.org/10.1093/nar/gkv639

Bravo, J. P. K. (2019). *Mechanisms of Genome Segment Assortment in Reovirus* (PhD Thesis). University of Leeds, UK. https://etheses.whiterose.ac.uk/25399/

Bravo, J. P. K., Borodavka, A., Barth, A., Calabrese, A. N., Mojzes, P., Cockburn, J. J. B., Lamb, D. C., & Tuma, R. (2018). Stability of local secondary structure determines selectivity of viral RNA chaperones. *Nucleic Acids Research*, *46*(15), 7924–7937. https://doi.org/10.1093/nar/gky394

Burke, J. E. (2019). Dynamic structural biology at the protein membrane interface. Journal of Biological Chemistry, 294(11), 3872–3880. https://doi.org/10.1074/jbc.aw118.003236

Canas, B. (2006). Mass spectrometry technologies for proteomics. *Briefings in Functional Genomics and Proteomics*, *4*(4), 295–320. https://doi.org/10.1093/bfgp/eli002

Cao, J., Burke, J. E., & Dennis, E. A. (2013). Using Hydrogen/Deuterium Exchange Mass Spectrometry to Define the Specific Interactions of the Phospholipase A2 Superfamily with Lipid Substrates, Inhibitors, and Membranes. *Journal of Biological Chemistry*, *288*(3), 1806–1813. https://doi.org/10.1074/jbc.r112.421909

Chen, Z. A., & Rappsilber, J. (2018). Protein Dynamics in Solution by Quantitative Crosslinking/Mass Spectrometry. *Trends in Biochemical Sciences*, *43*(11), 908–920. https://doi.org/10.1016/j.tibs.2018.09.003

Chiu, C. J., & Lee, L. H. (1997). Cloning and nucleotide sequencing of the S4 genome segment of avian reovirus S1133. *Archives of Virology*, *142*(12), 2515–2520. https://doi.org/10.1007/s007050050258

Chong, B. E., Wall, D. B., Lubman, D. M., & Flynn, S. J. (1997). Rapid profiling ofE. coli proteins up to 500 kDa from whole cell lysates using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, *11*(17), 1900–1908.

Chong, Y. K., Ho, C. C., Leung, S. Y., Lau, S. K., & Woo, P. C. (2018). Clinical Mass Spectrometry in the Bioinformatics Era: A Hitchhiker's Guide. *Computational and Structural Biotechnology Journal*, *16*, 316–334. https://doi.org/10.1016/j.csbj.2018.08.003

Clark, A. E., Kaleta, E. J., Arora, A., & Wolk, D. M. (2013). Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry: a Fundamental Shift in the Routine Practice of Clinical Microbiology. *Clinical Microbiology Reviews*, *26*(3), 547–603. https://doi.org/10.1128/cmr.00072-12

Cobo, F. (2013). Application of MALDI-TOF Mass Spectrometry in Clinical Virology: A Review. The Open Virology Journal, 7(1), 84–90. https://doi.org/10.2174/1874357920130927003

Combe, C. W., Fischer, L., & Rappsilber, J. (2015). xiNET: Cross-link Network Maps With Residue Resolution. Molecular & Cellular Proteomics, 14(4), 1137–1147. https://doi.org/10.1074/mcp.o114.042259

Dihazi, G. H., & Sinz, A. (2003). Mapping low-resolution three-dimensional protein structures using chemical cross-linking and Fourier transform ion-cyclotron resonance mass spectrometry. *Rapid Communications in Mass Spectrometry*, *17*(17), 2005–2014. https://doi.org/10.1002/rcm.1144

E. White, H., & V. Orlova, E. (2020). Bacteriophages: Their Structural Organisation and Function. *Bacteriophages - Perspectives and Future*. Published. https://doi.org/10.5772/intechopen.85484

Eisinger, M. L., Dörrbaum, A. R., Michel, H., Padan, E., & Langer, J. D. (2017). Ligand-induced conformational dynamics of the Escherichia coli Na+/H+ antiporter NhaA revealed by hydrogen/deuterium exchange mass spectrometry. *Proceedings of the National Academy of Sciences*, *114*(44), 11691–11696. https://doi.org/10.1073/pnas.1703422114

Gatlin-Bunai, C. L., Cazares, L. H., Cooke, W. E., Semmes, O. J., & Malyarenko, D. I. (2007). Optimization of MALDI-TOF MS Detection for Enhanced Sensitivity of Affinity-Captured Proteins Spanning a 100 kDa Mass Range. Journal of Proteome Research, 6(11), 4517–4524. https://doi.org/10.1021/pr0703526

Guttman, M., Wales, T. E., Whittington, D., Engen, J. R., Brown, J. M., & Lee, K. K. (2016). Tuning a High Transmission Ion Guide to Prevent Gas-Phase Proton Exchange During H/D

Exchange MS Analysis. Journal of the American Society for Mass Spectrometry, 27(4), 662–668. https://doi.org/10.1007/s13361-015-1330-8

Hamuro, Y., Coales, S. J., Southern, M. R., Nemeth-Cawley, J. F., Stranz, D. D., & Griffin, P. R. (2003). Rapid analysis of protein structure and dynamics by hydrogen/deuterium exchange mass spectrometry. Journal of biomolecular techniques: JBT, 14(3), 171–182.

Holding, A. N. (2015). XL-MS: Protein cross-linking coupled with mass spectrometry. *Methods*, *89*, 54–63. https://doi.org/10.1016/j.ymeth.2015.06.010

Hübschmann, H. (2015). Handbook of GC-MS: Fundamentals and Applications (3rd ed.). Wiley-VCH.

Iacobucci, C., Götze, M., Ihling, C. H., Piotrowski, C., Arlt, C., Schäfer, M., Hage, C., Schmidt, R., & Sinz, A. (2018). A cross-linking/mass spectrometry workflow based on MS-cleavable cross-linkers and the MeroX software for studying protein structures and protein–protein interactions. *Nature Protocols*, *13*(12), 2864–2889. https://doi.org/10.1038/s41596-018-0068-8

Kaufmann, K. W., Lemmon, G. H., DeLuca, S. L., Sheehan, J. H., & Meiler, J. (2010). Practically Useful: What the Rosetta Protein Modeling Suite Can Do for You. Biochemistry, 49(14), 2987–2998. https://doi.org/10.1021/bi902153g

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols, 10(6), 845–858. https://doi.org/10.1038/nprot.2015.053

Kicman, A. T., Parkin, M. C., & Iles, R. K. (2007). An introduction to mass spectrometry based proteomics—Detection and characterization of gonadotropins and related molecules. *Molecular and Cellular Endocrinology*, *260–262*, 212–227. https://doi.org/10.1016/j.mce.2006.02.022

Klykov, O., & Weller, M. G. (2015). Quantification of N-hydroxysuccinimide and N-hydroxysulfosuccinimide by hydrophilic interaction chromatography (HILIC). *Analytical Methods*, *7*(15), 6443–6448. https://doi.org/10.1039/c5ay00042d

Konermann, L., Pan, J., & Liu, Y. H. (2011). Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem. Soc. Rev.*, *40*(3), 1224–1234. https://doi.org/10.1039/c0cs00113a

Kostyukevich, Y., Acter, T., Zherebker, A., Ahmed, A., Kim, S., & Nikolaev, E. (2018). Hydrogen/deuterium exchange in mass spectrometry. Mass Spectrometry Reviews, 37(6), 811–853. https://doi.org/10.1002/mas.21565

Lento, C., Wilson, D. J., & Audette, G. F. (2015). Dimerization of the type IV pilin from Pseudomonas aeruginosa strain K122-4 results in increased helix stability as measured by

time-resolved hydrogen-deuterium exchange. Structural Dynamics, 3(1), 012001. https://doi.org/10.1063/1.4929597

Li, J., Rodnin, M. V., Ladokhin, A. S., & Gross, M. L. (2014). Hydrogen–Deuterium Exchange and Mass Spectrometry Reveal the pH-Dependent Conformational Changes of Diphtheria Toxin T Domain. *Biochemistry*, *53*(43), 6849–6856. https://doi.org/10.1021/bi500893y

Lísal, J., Kainov, D. E., Lam, T. T., Emmett, M. R., Wei, H., Gottlieb, P., Marshall, A. G., & Tuma, R. (2006). Interaction of packaging motor with the polymerase complex of dsRNA bacteriophage. *Virology*, *351*(1), 73–79. https://doi.org/10.1016/j.virol.2006.03.025

Louten, J. (2016). Virus Structure and Classification. *Essential Human Virology*, 19–29. https://doi.org/10.1016/b978-0-12-800947-5.00002-8

Lu, L., Millikin, R. J., Solntsev, S. K., Rolfs, Z., Scalf, M., Shortreed, M. R., & Smith, L. M. (2018). Identification of MS-Cleavable and Noncleavable Chemically Cross-Linked Peptides with MetaMorpheus. *Journal of Proteome Research*, *17*(7), 2370–2376. https://doi.org/10.1021/acs.jproteome.8b00141

Mahmoudabadi, G., & Phillips, R. (2018). A comprehensive and quantitative exploration of thousands of viral genomes. *ELife*, *7*. https://doi.org/10.7554/elife.31955

Malito, E., Faleri, A., lo Surdo, P., Veggi, D., Maruggi, G., Grassi, E., Cartocci, E., Bertoldi, I., Genovese, A., Santini, L., Romagnoli, G., Borgogni, E., Brier, S., lo Passo, C., Domina, M., Castellino, F., Felici, F., van der Veen, S., Johnson, S., . . . Masignani, V. (2013). Defining a protective epitope on factor H binding protein, a key meningococcal virulence factor and vaccine antigen. Proceedings of the National Academy of Sciences, 110(9), 3304–3309. https://doi.org/10.1073/pnas.1222845110

Martinez-Costas, J., Varela, R., & Benavente, J. (1995). Endogenous Enzymatic Activities of the Avian Reovirus S1133: Identification of the Viral Capping Enzyme. *Virology*, *206*, 1017–1026. https://doi.org/10.1006/viro.1995.1024

Matzinger, M., & Mechtler, K. (2020). Cleavable Cross-Linkers and Mass Spectrometry for the Ultimate Task of Profiling Protein–Protein Interaction Networks in Vivo. *Journal of Proteome Research*, *20*(1), 78–93. https://doi.org/10.1021/acs.jproteome.0c00583

Merkley, E. D., Rysavy, S., Kahraman, A., Hafen, R. P., Daggett, V., & Adkins, J. N. (2014). Distance restraints from crosslinking mass spectrometry: Mining a molecular dynamics simulation database to evaluate lysine-lysine distances. Protein Science, 23(6), 747–759. https://doi.org/10.1002/pro.2458

Mirbagheri, S. A., Hosseini, H., & Ghalyanchilangeroudi, A. (2019). Molecular characterization of avian reovirus causing tenosynovitis outbreaks in broiler flocks, Iran. Avian Pathology, 49(1), 15–20. https://doi.org/10.1080/03079457.2019.1654086

Murphy, D. M. (2016). The sTOF, a Favorable Geometry for a Time-of-Flight Analyzer. *Journal of the American Society for Mass Spectrometry*, *28*(2), 242–246. https://doi.org/10.1007/s13361-016-1518-6

Narang, D., Lento, C., & J. Wilson, D. (2020). HDX-MS: An Analytical Tool to Capture Protein Motion in Action. Biomedicines, 8(7), 224. https://doi.org/10.3390/biomedicines8070224

Nibert, M.L., Schiff, L.A., 2001. Reoviruses and their replication. In: Knipe,D.M., Hooley, P.M. (Eds.), Fields Virology, fourth ed. Lippincott Willams& Wilkins, Philadelphia, PA, pp. 1679–1728

O'Reilly, F. J., & Rappsilber, J. (2018). Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. *Nature Structural & Molecular Biology*, *25*(11), 1000–1008. https://doi.org/10.1038/s41594-018-0147-0

Olshina, M. A., & Sharon, M. (2016). Mass spectrometry: a technique of many faces. Quarterly Reviews of Biophysics, 49. https://doi.org/10.1017/s0033583516000160

Pierce, M. M. (2020). Virus classification. *AccessScience*. Published. https://doi.org/10.1036/1097-8542.733650

Piersimoni, L., & Sinz, A. (2020). Cross-linking/mass spectrometry at the crossroads. *Analytical and Bioanalytical Chemistry*, *412*(24), 5981–5987. https://doi.org/10.1007/s00216-020-02700-x

Radou, G., Dreyer, F., Tuma, R., & Paci, E. (2014). Functional Dynamics of Hexameric Helicase Probed by Hydrogen Exchange and Simulation. *Biophysical Journal*, *107*(4), 983–990. https://doi.org/10.1016/j.bpj.2014.06.039

Rajkowitsch, L., Chen, D., Stampfl, S., Semrad, K., Waldsich, C., Mayer, O., Jantsch, M. F., Konrat, R., Bläsi, U., & Schroeder, R. (2007). RNA Chaperones, RNA Annealers and RNA Helicases. *RNA Biology*, *4*(3), 118–130. https://doi.org/10.4161/rna.4.3.5445

Rampersad, S., & Tennant, P. (2018). Replication and Expression Strategies of Viruses. *Viruses*, 55–82. https://doi.org/10.1016/b978-0-12-811257-1.00003-6

Rappsilber, J., Mann, M., & Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nature Protocols, 2(8), 1896–1906. https://doi.org/10.1038/nprot.2007.261

Ryu, W. S. (2017). Discovery and Classification. *Molecular Virology of Human Pathogenic Viruses*, 3–20. https://doi.org/10.1016/b978-0-12-800838-6.00001-1

Sahin, E., Egger, M. E., McMasters, K. M., & Zhou, H. S. (2013). Development of Oncolytic Reovirus for Cancer Therapy. *Journal of Cancer Therapy*, *04*(06), 1100–1115. https://doi.org/10.4236/jct.2013.46127

Sakudo, A., Onodera, T., & Tanaka, Y. (2010). Inactivation of Viruses. In A. Sakudo & H. Shintani (Eds.), *Sterilization and Disinfection by Plasma: Sterilization Mechanisms, Biological and Medical Applications* (1st ed., pp. 49–60). Nova Science Publisher, USA.

Schiffrin, B., Radford, S. E., Brockwell, D. J., & Calabrese, A. N. (2020). PyXlinkViewer: A flexible tool for visualization of protein chemical crosslinking data within thePyMOLmolecular graphics system. Protein Science, 29(8), 1851–1857. https://doi.org/10.1002/pro.3902

Schnitzer, T. J. (1985). Protein coding assignment of the S genes of the avian reovirus S1133. *Virology*, *141*(1), 167–170. https://doi.org/10.1016/0042-6822(85)90194-1

Shkriabai, N., Dharmarajan, V., Slaughter, A., Kessl, J. J., Larue, R. C., Feng, L., Fuchs, J. R., Griffin, P. R., & Kvaratskhelia, M. (2014). A Critical Role of the C-terminal Segment for Allosteric Inhibitor-induced Aberrant Multimerization of HIV-1 Integrase. *Journal of Biological Chemistry*, *289*(38), 26430–26440. https://doi.org/10.1074/jbc.m114.589572

Sinz, A. (2018). Cross-Linking/Mass Spectrometry for Studying Protein Structures and Protein–Protein Interactions: Where Are We Now and Where Should We Go from Here? *Angewandte Chemie International Edition*, *57*(22), 6390–6396. https://doi.org/10.1002/anie.201709559

Spandidos, D. A., & Graham, A. F. (1976). Physical and chemical characterization of an avian reovirus. *Journal of Virology*, *19*(3), 968–976. https://doi.org/10.1128/jvi.19.3.968-976.1976

Stamatos, N. M., & Gomatos, P. J. (1982). Binding to selected regions of reovirus mRNAs by a nonstructural reovirus protein. Proceedings of the National Academy of Sciences, 79(11), 3457–3461. https://doi.org/10.1073/pnas.79.11.3457

Thomas, S. N. (2019). Mass spectrometry. *Contemporary Practice in Clinical Chemistry*, 171–185. https://doi.org/10.1016/b978-0-12-815499-1.00010-7

Tourís-Otero, F., Martínez-Costas, J., Vakharia, V. N., & Benavente, J. (2005b). Characterization of the nucleic acid-binding activity of the avian reovirus non-structural protein σNS. *Journal of General Virology*, *86*(4), 1159–1169. https://doi.org/10.1099/vir.0.80491-0

Touris-Otero, F., Martínez-Costas, J., Vakharia, V. N., & Benavente, J. (2004). Avian reovirus nonstructural protein μNS forms viroplasm-like inclusions and recruits protein σNS to these structures. *Virology*, *319*(1), 94–106. https://doi.org/10.1016/j.virol.2003.10.034

Urban, P. L. (2016). Quantitative mass spectrometry: an overview. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2079), 20150382. https://doi.org/10.1098/rsta.2015.0382

Varela, R., & Benavente, J. (1994). Protein coding assignment of avian reovirus strain S1133. *Journal of Virology*, *68*(10), 6775–6777. https://doi.org/10.1128/jvi.68.10.6775-6777.1994

Wei, H., Ahn, J., Yu, Y. Q., Tymiak, A., Engen, J. R., & Chen, G. (2012). Using Hydrogen/Deuterium Exchange Mass Spectrometry to Study Conformational Changes in Granulocyte Colony Stimulating Factor upon PEGylation. *Journal of the American Society for Mass Spectrometry*, *23*(3), 498–504. https://doi.org/10.1007/s13361-011-0310-x

Wörner, T. P., Shamorkina, T. M., Snijder, J., & Heck, A. J. R. (2020). Mass Spectrometry-Based Structural Virology. *Analytical Chemistry*, *93*(1), 620–640. https://doi.org/10.1021/acs.analchem.0c04339

Zhang, X., Tang, J., Walker, S. B., O'Hara, D., Nibert, M. L., Duncan, R., & Baker, T. S. (2005). Structure of avian orthoreovirus virion by electron cryomicroscopy and image reconstruction. *Virology*, *343*(1), 25–35. https://doi.org/10.1016/j.virol.2005.08.002

Zheng, H., Handing, K. B., Zimmerman, M. D., Shabalin, I. G., Almo, S. C., & Minor, W. (2015). X-ray crystallography over the past decade for novel drug discovery – where are we heading next? Expert Opinion on Drug Discovery, 10(9), 975–989. https://doi.org/10.1517/17460441.2015.1061991

8 SUPPLEMENTARY

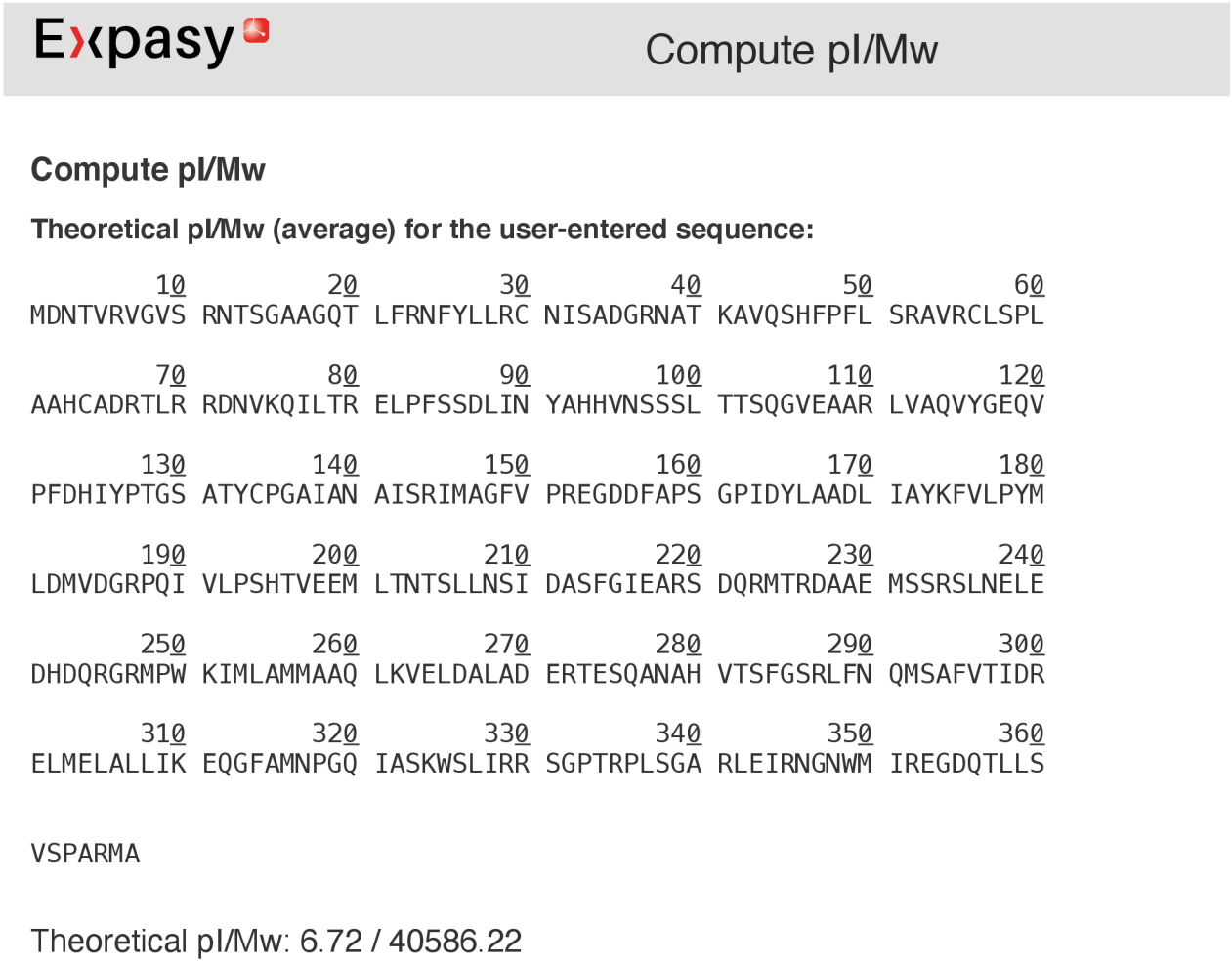

Fig. S1: The average mass and pI of protein σNS computed in expasy.org

Table S1: List of all identified peptides in MeroX software

N.	Score	m/z	Experimental mass (Da)	Calculated Mass (Da)	Deviation in ppm	Peptide 1	From	To	Peptide 2	From	To	Best linkage position peptide 1	Best linkage position peptide 2
1	43	606.55	2423.19	2423.20	-4.13	[NTSGAAGQTLFR]	130	141	[QGKEMDSLR]	77	85	T2	K3
2	60	606.55	2423.20	2423.20	-1.06	[NTSGAAGQTLFR]	130	141	[QGKEMDSLR]	77	85	T2	K3
3	83	606.56	2423.20	2423.20	0.75	[NTSGAAGQTLFR]	130	141	[QGKEMDSLR]	77	85	T2	K3
4	44	623.07	2489.26	2489.26	-1.34	[AVQSHFPFLSR]	160	170	[QGKEMDSLR]	77	85	S10	K3
5	43	649.34	2594.33	2594.32	3.14	[BLSPLAAHBADR]	174	185	[DNVKQILTR]	190	198	S3	K4
6	98	649.33	2594.32	2594.32	-0.89	[BLSPLAAHBADR]	174	185	[DNVKQILTR]	190	198	S3	K4
7	101	803.4	3210.60	3210.59	1.13	[NATKAVQSHFPFLSR]	156	170	[BLSPLAAHBADR]	174	185	K4	S3
8	96	649.33	2594.31	2594.32	-3.54	[BLSPLAAHBADR]	174	185	[DNVKQILTR]	190	198	S3	K4
9	38	642.92	3210.59	3210.59	-0.31	[NATKAVQSHFPFLSR]	156	170	[BLSPLAAHBADR]	174	185	K4	S3
10	85	649.33	2594.31	2594.32	-2.46	[BLSPLAAHBADR]	174	185	[DNVKQILTR]	190	198	S3	K4
11	72	726.62	2903.48	2903.48	-2.63	[NATKAVQSHFPFLSR]	156	170	[QGKEMDSLR]	77	85	T3	K3
12	94	581.5	2903.47	2903.48	-3.68	[NATKAVQSHFPFLSR]	156	170	[QGKEMDSLR]	77	85	T3	K3
13	69	726.63	2903.48	2903.48	-0.34	[NATKAVQSHFPFLSR]	156	170	[QGKEMDSLR]	77	85	K4	K3
14	81	803.41	3210.60	3210.59	1.79	[NATKAVQSHFPFLSR]	156	170	[BLSPLAAHBADR]	174	185	K4	S3
15	79	581.5	2903.47	2903.48	-3.53	[NATKAVQSHFPFLSR]	156	170	[QGKEMDSLR]	77	85	K4	K3
16	89	803.4	3210.58	3210.59	-4.04	[NATKAVQSHFPFLSR]	156	170	[BLSPLAAHBADR]	174	185	K4	S3
17	35	726.63	2903.48	2903.48	0.1	[NATKAVQSHFPFLSR]	156	170	[QGKEMDSLR]	77	85	K4	K3
18	97	803.4	3210.59	3210.59	-2.53	[NATKAVQSHFPFLSR]	156	170	[BLSPLAAHBADR]	174	185	K4	S3
19	72	726.63	2903.48	2903.48	-1.07	[NATKAVQSHFPFLSR]	156	170	[QGKEMDSLR]	77	85	S8	K3
20	99	726.62	2903.47	2903.48	-2.9	[NATKAVQSHFPFLSR]	156	170	[QGKEMDSLR]	77	85	K4	S7
21	57	642.92	3210.58	3210.59	-2.8	[NATKAVQSHFPFLSR]	156	170	[BLSPLAAHBADR]	174	185	K4	S3
22	112	726.63	2903.48	2903.48	-1.46	[NATKAVQSHFPFLSR]	156	170	[QGKEMDSLR]	77	85	K4	K3
23	112	726.63	2903.48	2903.48	-0.81	[NATKAVQSHFPFLSR]	156	170	[QGKEMDSLR]	77	85	K4	S7
24	48	651.36	2602.40	2602.41	-2.01	[RDNVKQILTR]	189	198	[NTSGAAGQTLFR]	130	141	K5	T9
25	109	726.62	2903.48	2903.48	-2.59	[NATKAVQSHFPFLSR]	156	170	[QGKEMDSLR]	77	85	K4	K3
26	64	726.62	2903.47	2903.48	-3.84	[NATKAVQSHFPFLSR]	156	170	[QGKEMDSLR]	77	85	S8	K3
27	32	651.36	2602.40	2602.41	-2.12	[RDNVKQILTR]	189	198	[NTSGAAGQTLFR]	130	141	K5	T9
28	63	651.36	2602.41	2602.41	1.09	[RDNVKQILTR]	189	198	[NTSGAAGQTLFR]	130	141	K5	T9
29	75	617.36	2466.41	2466.42	-1.21	[RDNVKQILTR]	189	198	[DNVKQILTR]	190	198	K5	K4
30	61	617.36	2466.41	2466.42	-1.97	[RDNVKQILTR]	189	198	[DNVKQILTR]	190	198	K5	K4

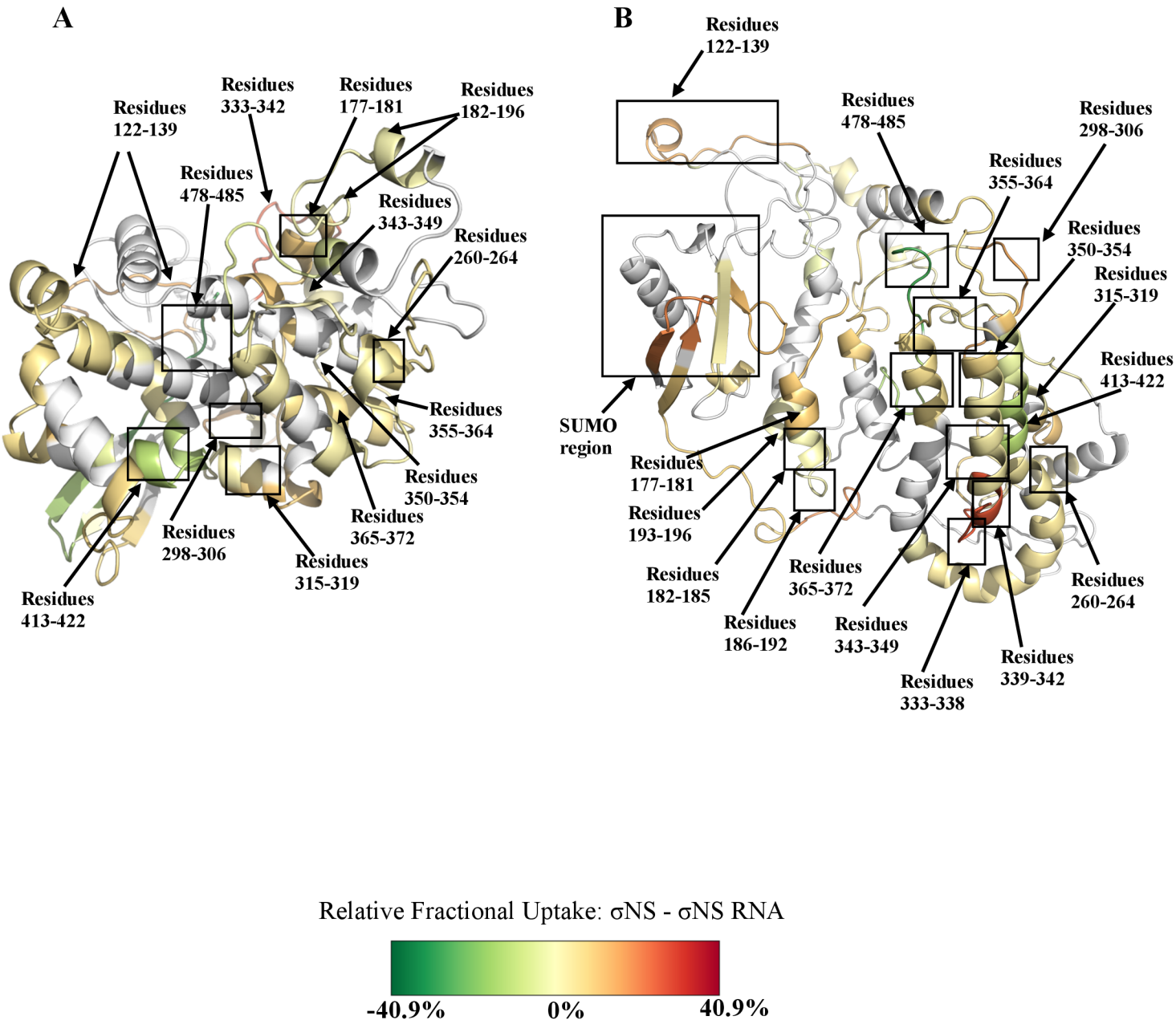

Figure S2: HDX MS analysis of σNS protein and σNS-RNA complex state comparison in 3D for (**A**) σNS model and (**B**) σNS SUMO model at 30 min HDX labeling time. The color code represents the relative changes in deuterium uptake between protein and σNS-RNA complex states (see legend). The regions in the protein structure colored in white were not detected by HDX.

.