Multi-level characterization of bacteriophage T5 capsids and capsid-like particles using mass spectrometry.

Kavya Clément

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : Biologie Structurale et Nanobiologie

Arrêté ministériel : 25 mai 2016

Présentée par

KAVYA CLEMENT

Thèse dirigée par **Christophe MASSELON**

préparée au sein du **Laboratoire Biologie et Biotechnologies pour la Santé**
dans **l'École Doctorale Chimie et Sciences du Vivant**

Caractérisation multi-échelle de capsides et de pseudo-capsides de bactériophages T5 par spectrométrie de masse.

Multi-level characterization of bacteriophage T5 capsids and capsid-like particles using mass spectrometry.

Thèse soutenue publiquement le **24 septembre 2021**,
devant le jury composé de :

Monsieur Christophe MASSELON
CADRE SCIENTIFIQUE DES EPIC, CEA Grenoble, Directeur de thèse
Madame Cécile BREYTON
DIRECTEUR DE RECHERCHE, CNRS, Présidente
Madame Joelle VINH
DIRECTEUR DE RECHERCHE, cnrs, Examinatrice
Monsieur Mark VAN RAAIJ
CHERCHEUR, Centro Nacional de Biotecnologia, Rapporteur
Monsieur Richard COLE
PROFESSEUR DES UNIVERSITES, Sorbonne Université, Rapporteur

ACKNOWLEDGEMENTS

Everything, I believe takes place as per God's plan, and hence I would like to begin by thanking The Almighty for showering all the blessings on me throughout this wonderful journey.

In addition to my efforts, the support of several individuals and teams equipped me with the guidance and strength to make this thesis possible. I would like to express my sincere gratitude to all who helped me during my Ph.D. study at the Commissariat à l'Énergie Atomique / French Alternative Energies and Atomic Energy Commission (CEA), Grenoble.

My supervisor Mr. Christophe MASSELON, Senior Research Scientist at CEA, was welcoming right from the moment I was picked for his laboratory. He has offered constant encouragement and fruitful discussions day in and day out. Heartfelt thanks to him for entrusting me with the honor to present my dissertation before this jury.

Thanks to the members of my thesis committee — Mme. Cécile BREYTON (Research Director at CNRS), Mme. Joelle VINH (Research Director at CNRS), Mr. Mark VAN RAAIJ (Researcher at CNB), and Mr. Richard COLE (Professor at Sorbonne University) — for taking time out of their busy schedules to provide advice and suggestions for my work. Special thanks to Mr. Mark VAN RAAIJ and Mr. Richard COLE, for having accepted to judge my work and to be the reviewers.

I would also like to acknowledge with wholehearted thanks the efforts taken by Mr. Sébastien HENTZ to guide me through my first steps in the Nano-Electro-Mechanical-Sensor based Mass spectrometry (NEMS-MS). Special thanks to Mme. Pascale BOULABGER (Research Director at Institute Of Biology Intégrative De La Cellule) for sharing her knowledge that benefited the successful completion of this study. I am also indebted to Mme. Virginie BRUN and Mr. Yohann COUTE, for their scientific and technical help that allowed me to carry out this work on my own. Thanks to Mme. Elisabetta BOERI ERBA for her valuable advice and suggestions with the experiments.

The NEMS team has also been great in accompanying me, especially during the times that helped me begin with my thesis project. Sergio DOMINGUEZ-MEDINA and Halim MOHAMMAD ABDUL shared their experimental knowledge while Szu-Hsueh LAI and Bogdan VYSOTSKI offered me expertise on data. Thomas FORTIN, Adrien REYNAUD, and Vaitson ÇUMAKU have been reliable colleagues throughout. Thanks to all.

I would also like to thank Julia NOVION DUCASSOU for her technical assistance round the clock. Special thanks to Marie COURÇON for being a happy friend at the office.

Several members from the Exploring the Dynamics of Proteomes (EDyP) lab have contributed in their own way to make these three years a sacred good time for me, especially Lucid BELMUDES, Sandrine MIESCH-FREMY, Michel JAQUINOD, Delphine PELIGER, Myriam FERRO, Celine FLEURY, Veronique DUPIERRIS, and Christophe BRUELEY among others. Even as I face difficulty to recollect the names of others, their contributions — both small and big — have assisted me a lot. Thanks to them from the bottom of my heart.

Not to forget, my family has stayed rock solid with me with all support to help me achieve my aspirations. My parents Clement Silvester and Lissy Clement gave me the opportunity to study abroad, encouraged me to explore, and to always aim higher shaping me into who I am today. My sister, Divya CLEMENT, has always tried to cheer me up on days when things were not well. My better half Ralph ALEX ARAKAL has ensured my good mood, with this indulgence and patience. Thanks to them for being there wholeheartedly. Special thanks to my dearest friends Nipia SUGATHAN and Santhini V M for all the motivation and support as well.

TABLE OF CONTENT

GENERAL INTRODUCTION

Vaccination is a method of prevention and control over infectious disease. With the help of vaccine about 2-3 million of human race has been saved from death [1]. It works by priming the immune system to recognize and destroy a given pathogen upon future exposure. To achieve that, they must prompt the presentation of pathogen-specific antigens on the surface of specialized antigen presenting cells (APCs). A central challenge to vaccination lies in the efficient delivery of antigens to APCs without causing adverse reactions. This is possible by two routes: either by direct introduction of antigenic proteins, or by the incorporation of genetic material (RNA or DNA) coding for the desired antigen.

The current vaccines are either having active ingredients or adjuvants, they suffer from stability issues, which questions its side effects and inefficiency [2]. A drive to develop cheaper and safer immunization led to the emergence of second-generation vaccines composed of purified (recombinant) antigenic components of pathogens [3]. These subunit vaccines reduce the risk of side effects, but often necessitate multiple injections, and require adjuvants to enhance the immune response. Adjuvants can unintentionally alter antigens, and can cause safety issues [4]. More recently, mRNA and DNA vaccines have emerged, which induce the in situ production of the target antigen to produce protective immunity against diseases. Unfortunately, these new vaccines suffer from specific challenges associated with possible toxicity, oligonucleotide degradation, and concerns over integration into the host genome, and requirements to prime the immune response with other immunogenic agents [5].

Nanoparticle-based vaccines attract growing attention as a means for targeted antigen delivery [6]. Such nano-carriers can enhance cellular uptake by APC, and may even display self-adjuvant properties. Among them, engineered “Virus Like Particles” (VLP) offer some distinct advantages over synthetic particles. They self-assemble into biocompatible repetitive structures, analogous to pathogen-associated molecular patterns, triggering potent immune responses, which offers ways to circumvent the need for adjuvants [7]. While most VLP are derived from human pathogens (HBV, HPV, Adenovirus), viruses infecting bacteria, or bacteriophages (e.g. Qβ, T4), have recently been proposed as safer, more stable, and versatile antigen nano-carriers [8] [9].

Preliminary evidence suggest that capsids of bacteriophage T5, a bacterial virus devoid of toxicity to humans, could be engineered into safe, efficacious, stable, and versatile antigen presentation nano-carriers [10]. These VLP have the same structure as native T5 capsids [11] [12], but are devoid of the viral genome and of their tails which makes them innocuous. Initial evidence indicates that they are very likely as stable as the bacteriophages themselves, which can tolerate harsh environments [13]. They bind at their external surface 120 copies of a “decoration” protein that can be fused to a variety of antigenic proteins [10]. This antigen grafting process based on molecular recognition far surpasses cross-linking in efficiency and flexibility. Finally, due to their natural immunogenicity, VLP may circumvent the need for adjuvants to boost the immune response. An easy-to-produce, well-characterized, and modular antigen presentation nano-carrier would significantly lower the cost and duration of vaccine development.

Structural integrity and immunogenicity are essential factors for nano-vaccine development. In contrast to inorganic nanoparticles, VLP result from tightly regulated auto-assembly processes, which yield strictly identical particles of definite molecular composition. Therefore, their molecular mass could be used as an alternative to study their structural integrity, which could

be assessed by mass spectrometry. Compared to size and morphological characterization methods like Dynamic Light Scattering (DLS) or Cryogenic electron microscopy (cryo-EM), Mass spectrometry (MS) provides a convenient metric, while avoiding selection biases.

MS is a key analytical technique to collect essential information from minute amounts of samples present in the gas, solid or liquid phase. MS has been the fastest growing analytical technique over the last two decades [14]. Conventional MS requires particle ionization and observes how ion trajectories respond to various electromagnetic fields, in order to determine the mass to charge ratio (m/z) distribution of the species of interest. The molecular weight of each compound in the analyzed mixture is deduced using the m/z ratio and the mass spectrum displays the abundance with respect to the m/z. This technique is used to analyze both organic and inorganic molecules. The introduction of electrospray ionization (ESI) in the field of MS enabled the transfer of fragile biomolecules into the gas phase [15]. This benefited to the development of the field of proteomics, an emerging technique for the analysis of the protein complement of biological systems. MS has since become a routine analytical tool in modern biological research, and in recent years, it has gained a foothold in the realm of clinical diagnostic and screening. For instance, MS is routinely used to screen newborns for metabolic defects, or to phenotype bacterial infections [16] [17].

The present project focuses on studying bacteriophage T5 capsids from different perspectives using multiple MS approaches. One of the aspect of this project focus towards the contribute to an improved understanding of T5 capsid assembly process by determining the stoichiometry of the T5 capsid protease using nanoscale liquid chromatography-tandem mass spectrometry (nano LC-MS/MS) based proteomics. Another aspect of this thesis is to assess the integrity and stability of the phage T5 capsids by analyzing their intact mass using Nano-electro-mechanical Sensor based Mass Spectrometry (NEMS-MS).

Chapter III of this thesis describes an original experimental study method to determine the presence and the copy number of an important structural protein of bacteriophage T5 capsid. The maturation of the capsid begins with the activation of a protease called pb11 [18]. In spite of its importance for T5 capsid maturation and assembly, the exact quantity of this protein at each stage of the capsid assembly is still subject to debate.

With the help of shotgun proteomic discovery, MS based approaches is possible to quantify proteins by measuring the relative abundance of their tryptic peptide by data dependent acquisition, counting either the MS/MS spectra or from extracted ion chromatography (XIC) [19]. Moreover, when shotgun proteomic join hands with isotopically labelled strategies it is possible to investigate the quantity of proteins in different biological samples [20]. Herein, I developed a novel targeted proteomic assay for the quantification of pb11 in mature T5 capsid using heavy isotopically labelled QconCAT dipeptides.

Currently there is no commercially available mass-spectrometric technique, which is able to measure the mass in the mega- to giga-Dalton mass range required for VLP analysis (One Dalton = 1.66×10^{-27} kg). This limitation has prevented the study of large scale biomolecular assemblies and bio-particles of high biomedical importance, such as virus, bacteria or cellular organelles, NEMS-MS constitute a promising technology for this challenging mass range. This is a new method for mass sensing at the single particle level [21]. It allows investigating biological species that have so far escaped mass characterization and thus expands the applications of MS. A proof of concept of NEMS based mass spectrometry was demonstrated by CEA, Grenoble [22].

NEMS are nanoscale devices, usually planar in structure like cantilevers or suspended beams. When a mass lands onto the surface of a NEMS, the device's resonance frequency shifts downwards. Mass sensing can be performed by monitoring this frequency in real time. This emerging technology demonstrated the ability to characterize the mass of empty and filled T5 capsids. In this previously published study, the uncertainties due to physical phenomenon involved in NEMS mass measurements were quantified [23]. However, there has been so far no integrated perspective on the various factors influencing the measured mass using nano-resonators from an analytical chemistry point of view.

Chapter IV of the present work proposes to investigate the various factors influencing mass measurements using nano-resonator. This study focused on how properties independent from particle mass have an influence on the measurement and how to deal with these issues such as filtering out particles landing events occurring at sub-optimal positions on the sensor. In addition, device fabrication discrepancies was also taken in account to avoid the device-to-device variability. An evaluation of the magnitude of these effects on doubly clamped beams using bacteriophage T5 capsids was performed. Other than the physical property, a study on how particle desolvation affects mass measurements is also discussed here.

In this thesis, we will introduce bacteriophages, in particularly T5 capsid as a potential VLP (Chapter I). We will then present the technologies on which we will rely for our experiments (Chapter II). We then propose a contribution to understand T5 capsid maturation by studying T5 protease using a novel proteomic approach (Chapter III). Finally, we will describe an instrument that is able to measure mass on a routine analysis in the MDa-GDa range and discuss about various factors and parameters that affects the mass measurement (Chapter IV).

REFERENCES

1. Andre F, Booy R, Bock H, Clemens J, Datta S, John T, Lee B, Lolekha S, Peltola H, Ruff T, Santosham M, Schmitt H (2008) Vaccination greatly reduces disease, disability, death and inequity worldwide. Bull World Health Org 86:140–146 . https://doi.org/10.2471/BLT.07.040089

2. Bloom DE, Cadarette D (2019) Infectious Disease Threats in the Twenty-First Century: Strengthening the Global Response. Front Immunol 0: . https://doi.org/10.3389/fimmu.2019.00549

3. Moyle PM, Toth I (2013) Modern Subunit Vaccines: Development, Components, and Research Opportunities. ChemMedChem 8:360–376 . https://doi.org/10.1002/cmdc.201200487

4. Larson HJ, de Figueiredo A, Xiahong Z, Schulz WS, Verger P, Johnston IG, Cook AR, Jones NS (2016) The State of Vaccine Confidence 2016: Global Insights Through a 67-Country Survey. EBioMedicine 12:295–301 . https://doi.org/10.1016/j.ebiom.2016.08.042

5. Liu MA (2019) A Comparison of Plasmid DNA and mRNA as Vaccine Technologies. Vaccines 7:37 . https://doi.org/10.3390/vaccines7020037

6. Pati R, Shevtsov M, Sonawane A (2018) Nanoparticle Vaccines Against Infectious Diseases. Front Immunol 0: . https://doi.org/10.3389/fimmu.2018.02224

7. Mohsen MO, Gomes AC, Vogel M, Bachmann MF (2018) Interaction of Viral Capsid-Derived Virus-Like Particles (VLPs) with the Innate Immune System. Vaccines (Basel) 6:37 . https://doi.org/10.3390/vaccines6030037

8. Sharma J, Shepardson K, Johns LL, Wellham J, Avera J, Schwarz B, Rynda-Apple A, Douglas T (2020) A Self-Adjuvanted, Modular, Antigenic VLP for Rapid Response to Influenza Virus Variability. ACS Appl Mater Interfaces 12:18211–18224 . https://doi.org/10.1021/acsami.9b21776

9. Tao P, Zhu J, Mahalingam M, Batra H, Rao VB (2019) Bacteriophage T4 nanoparticles for vaccine delivery against infectious diseases. Advanced Drug Delivery Reviews 145:57–72 . https://doi.org/10.1016/j.addr.2018.06.025

10. Vernhes E, Renouard M, Gilquin B, Cuniasse P, Durand D, England P, Hoos S, Huet A, Conway JF, Glukhov A, Ksenzenko V, Jacquet E, Nhiri N, Zinn-Justin S, Boulanger P (2017) High affinity anchoring of the decoration protein pb10 onto the bacteriophage T5 capsid. Scientific Reports 7:41662 . https://doi.org/10.1038/srep41662

11. Preux O, Durand D, Huet A, Conway JF, Bertin A, Boulogne C, Drouin-Wahbi J, Trévarin D, Pérez J, Vachette P, Boulanger P (2013) A two-state cooperative expansion converts the procapsid shell of bacteriophage T5 into a highly stable capsid isomorphous to the final virion head. J Mol Biol 425:1999–2014 . https://doi.org/10.1016/j.jmb.2013.03.002

12. Huet A, Duda RL, Boulanger P, Conway JF (2019) Capsid expansion of bacteriophage T5 revealed by high resolution cryoelectron microscopy. PNAS 116:21037–21046 . https://doi.org/10.1073/pnas.1909645116

13. Jończyk E, Kłak M, Międzybrodzki R, Górski A (2011) The influence of external factors on bacteriophages—review. Folia Microbiol 56:191–200 . https://doi.org/10.1007/s12223-011-0039-8

14. Yates III JR (2011) A century of mass spectrometry: from atoms to proteomes. Nat Methods 8:633–637 . https://doi.org/10.1038/nmeth.1659

15. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1989) Electrospray ionization for mass spectrometry of large biomolecules. Science 246:64–71 . https://doi.org/10.1126/science.2675315

16. Using Tandem Mass Spectrometry for Metabolic Disease Screening Among Newborns </P>. https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5003a1.htm. Accessed 24 Jul 2021

17. Bardin EE, Cameron SJS, Perdones-Montero A, Hardiman K, Bolt F, Alton EWFW, Bush A, Davies JC, Takáts Z (2018) Metabolic Phenotyping and Strain Characterisation of Pseudomonas aeruginosa Isolates from Cystic Fibrosis Patients Using Rapid Evaporative Ionisation Mass Spectrometry. Sci Rep 8:10952 . https://doi.org/10.1038/s41598-018-28665-7

18. Zivanovic Y, Confalonieri F, Ponchon L, Lurz R, Chami M, Flayhan A, Renouard M, Huet A, Decottignies P, Davidson AR, Breyton C, Boulanger P (2014) Insights into Bacteriophage T5 Structure from Analysis of Its Morphogenesis Genes and Protein Components. Journal of Virology 88:1162–1174 . https://doi.org/10.1128/JVI.02262-13

19. Bauer M, Ahrné E, Baron AP, Glatter T, Fava LL, Santamaria A, Nigg EA, Schmidt A (2014) Evaluation of Data-Dependent and -Independent Mass Spectrometric Workflows for Sensitive Quantification of Proteins and Phosphorylation Sites. J Proteome Res 13:5973–5988 . https://doi.org/10.1021/pr500860c

20. Zhang Y, Fonslow BR, Shan B, Baek M-C, Yates JR (2013) Protein Analysis by Shotgun/Bottom-up Proteomics. Chem Rev 113:2343–2394 . https://doi.org/10.1021/cr3003533

21. E S, M S, S F, M D, M G, Ak N, R M, L D, Ml R, T A, G J, E C, C M, A B, S H (2018) Single-particle mass spectrometry with arrays of frequency-addressed nanomechanical resonators. Nat Commun 9:3283–3283 . https://doi.org/10.1038/s41467-018-05783-4

22. Hentz S, Masselon C (2014) Device for determining the mass of a particle in suspension or in solution in a fluid

23. Dominguez-Medina S, Fostner S, Defoort M, Sansa M, Stark A-K, Halim MA, Vernhes E, Gély M, Jourdan G, Alava T, Boulanger P, Masselon CD, Hentz S (2018) Neutral mass spectrometry of virus capsids above 100 megadaltons with nanomechanical resonators. Science 362:918–922 . https://doi.org/10.1126/science.aat6457

CHAPTER I: INTRODUCTION TO BACTERIOPHAGES

I.1 GENERALITIES

Bacteriophages are the biological entities present in greatest numbers on our planet [1]. Their total number is estimated to be about 10^{30} -10^{31} and the ratio of the number of phages to that of bacteria is about 10. It is fascinating to imagine that all these phages, each measuring about 100 nm long, put end to end would constitute a chain of approximately 10 to 100 million light years, corresponding to 2 to 20 round trips to the Andromeda galaxy from the earth, or 100 to 1000 times the diameter of our galaxy. There are enormous amount of bacteriophages whose genes have not yet been sequenced. Phages are of great importance in the regulation, dynamics and biodiversity of bacterial populations. The influence of phages on bacterial populations occurs in several ways such as by regulating the population of fastest growing bacteria, acting as a driving force for microbial evolution; which makes it possible to preserve large number of species and a great diversity of bacterial genomes [2].

I.2 HISTORY

In 1896, Ernest Hanbury Hankin, an English biologist described the first observation of bacteriophages in the Annales de l'Institut Pasteur [3]. His study described the antibacterial properties of a previously unknown agent against cholera microbes present in the water of Jumna and Ganges in India. In 1914, Frederick Twort, an English microbiologist observed glassy areas in his bacterial colonies using a microscope. After realizing that these areas were the result of the destruction of bacterial cells, he transferred them from one colony to another. This study published in 1915 in The Lancet, thus describes an agent endowed with antibacterial properties, but does not specify its nature [4]. Two years later, Félix d´Hérelle, a French-Canadian microbiologist finally described the isolation and properties of what he considered first to be an antagonistic microbe, and to which he gave the name "bacteriophage" (from the Greek *baktêria*: stick and *phagein*: to eat) [5].

I.3 PHAGE THERAPY

In the article published in 1917, Félix d´Hérelle, considered the use of bacteriophages against bacterial infection. In 1919, he used phages to treat avian typhoid in chicken, hemorrhagic septicemia and human plague *(Yersinia pestis)* [6]. This approach, termed phage therapy, was quickly recognized by a large number of scientists as a possible remedy in the fight against bacterial infections. A decade from 1920, several experiments were carried out using phages to curb epidemics, especially cholera. Despite very promising results, some scientists and the general public remained skeptical about the effectiveness of this mode of treatment. As a matter of fact, the experiments carried out often lacked adequate controls (such as untreated control groups or groups treated with conventional methods) to prove that the observed effects were indeed due to phage therapy. In some cases, the use of bacteriophages failed to cure patients, and sometimes it even worsened their state. These negative effects are now believed to be probably due to bacterial toxins present in the administered phage preparations [7]. In addition, as the exact nature of phages was not known, the conditions of preparation and storage were often unsuitable, leading to inefficient preparations [8] [9]. Following the discovery and development of antibiotics during the second half of the 20th century, interest in phages declined sharply. However, the former USSR still continued to develop phage therapy because

of their limited access to antibiotics. To this day, the Eliava Institute named after its founder George Eliava located in Georgia has remained a world reference in phage therapy [8]. Nowadays, there is a renewed interest in the use of phages in medicine due to the emergence of pathogenic bacterial strain that are resistant to antibiotics. Contrary to antibiotics, phages have the ability to co-evolve with their hosts: when bacteria develop resistance to phages, the phage can evolve to bypass these resistances. In addition, phage therapy is targeting only a specific pathogen without affecting other bacteria involved in the proper functioning of the body human. Thus, more and more research groups are now working on phage therapy as a promising alternative to the use of antibiotics in the treatment of certain infections. However, legislative obstacles remain in many countries, including France: bacteriophages have a very different mode of action than other therapies currently used and authorized. Regulations must therefore be adapted so that phage therapy can be properly developed [10] [11].

I.4 BIODIVERSITY OF BACTERIOPHAGES

Bacteriophages are the most abundant biodiversity species of life on Earth. Indeed, it is estimated that there are more than 10 species of phage infecting each species of bacteria [12]. The number of bacterial species is estimated to be 10^7, which makes the phage species count to be more than 10^8 [13]. On the other hand, due to strong structural constraints the enormous genomic diversity of phages corresponds to very limited structural diversity [14]. There is no single universal and systematic method of classifying phages [15]. Their classifications are mainly based on:

- Morphology (capsid symmetry: icosahedral, helical), presence of a tail, lipid envelope surrounding the capsid (subunit capsomere).
- Nature of their genome: double or single stranded, linear or circular, DNA or RNA.
- Three-dimensional structural analogies between proteins that constitute the viral particle
- Other properties, such as the method of assembly, or the attachment site on the bacterial host.

Conversely, the nature of their host is little or not taken into account [16]. Contrary to structural analysis, the criterion of sequence homology is only used for classification into genus, this homology being too weak to classify the phages in higher taxa. Of the 100 million putative phage species, only 5,568 have been observed by electron microscopy [17]. Of these, 96% feature a tail and an icosahedral capsid. These phages belong to the order Caudovirales, and are commonly called caudate phages. The other known phages (208 species) are polyhedral, filamentous or pleomorphic [Table I.1].

Shape	Nucleic Acid	Family	Genera in Species	Charecteristics	Examples
Tailed	dsDNA [L]	*Myoviridae*	6 in 1,320	Contractile tail	T4
		Siphoviridae	7 in 3,229	Long non contractile tail	λ, T5, HK97, SPP1
		Podoviridae	4 in 771	Short tail	T7
Polyhedral	ssDNA [C]	*Microviridae*	40 in 40	Conspicuous capsomers	Phi X 174
	dsDNA[C,S]	*Corticoviridae*	1 in 3	Complex capsid, lipids	PM2
	dsDNA[L]	*Tectivirdae*	1 in 19	Double capsid, lipids, *psuedo* tail	PRD1
	ssRNA[L]	*Leviviridae*	2 in 39	Poliovirus-like	MS2
	dsRNA[L,M]	*Cystoviridae*	1 in 3	Envelope, lipids	Phi 6, Phi 12
Filamentous	ssDNA[C]	*Inoviridae*	2 in 67	Long filaments, short rods	M13, Fd
	dsDNA[L]	*Lipothrixviridae*	4 in 7	Envelope, lipids	TTV1
	dsDNA[L]	*Rudiviridae*	10 in 3	Stiff rods, TMV like	TMV-like
Pleomorphic	dsDNA [C,S]	*Plasmaviridae*	1 in 5	Envelope, no capsid, lipids	L2
		Fuselloviridae	1 in 11	Lemon-shaped, envelop	SSV1

Table I.1. Classification and properties of bacteriophages and archaea viruses. The number of species indicated corresponds to the number of species characterized by electron microscopy. L=Linear, C=Circular, M=Multiparti, S=Supercoiled. (Adapted from Ackermann et al.) [18

I.5 MORPHOLOGY OF TAILED PHAGES

Tailed or caudate phages are by far the most common bacteriophages. They possess a capsid of icosahedral or pseudo-icosahedral geometry and a tail [see *Figure* I-1]. Their capsid is devoid of lipid envelope and it contains a linear double stranded DNA (size between 18,000-500,000 base pairs). Their tail has a helical symmetry. The host-binding site is at the end of this tail. The order of *Caudovirales* comprises three families:

- *Siphoviridae* (~3200 observations, i.e. 61% of known species). They have a long, flexible and non-contractile tail. The family of Siphoviridae comprises in particular phages λ, HK 97, T5, SIO-2 and SPP1.
- *Myoviridae* (~1300 observations, or 25% of known species). They feature a long contractile tail. The Myoviridae family includes in particular the phages T4 and P2.
- *Podoviridae* (~750 observations, or 14% of known species). They possess a short, non-contractile tail. This family includes, among others, the phages T7, P22 and Φ29.

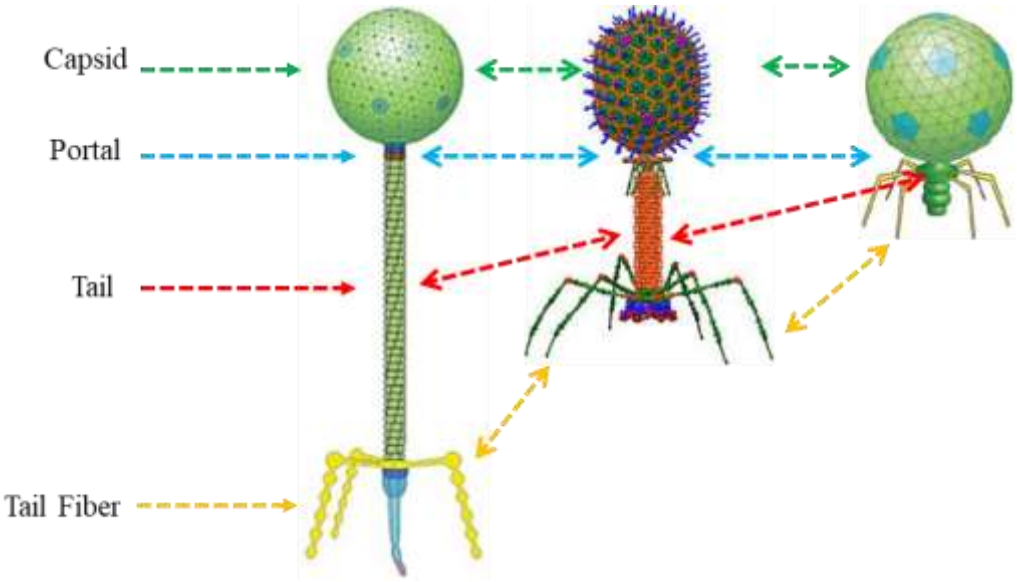


Figure I-1. Diagram illustrating the morphologies of different phage families of the order Caudovirales. From left to right: T5 (Siphoviridae, long non-contractile tail), T4 (Myoviridae, contractile tail), and T7 (Podoviridae, short tail). [18]

I.6 TAILED PHAGES CAPSIDS

I.6.1 ICOSAHEDRAL SYMMETRY

The capsids of caudate phages are made up of multiple copies of one [or sometimes two, e.g. phage T4] major head protein, which co-assemble into an icosahedral architecture. A homo-dodecamer of a so-called portal protein having the shape of a ring forming a gate for the DNA entry and exit is assembled at a unique vortex of the icosahedron and sometimes additional proteins, called accessory proteins are located at the outer surface of the capsid. The icosahedral organization of the capsid is shared by many others viruses infecting Eukaryotes and Archaea: this type of structure based on a single type of protein allows to maximize the volume to surface ratio, therefore providing as much space as possible for the genome. An icosahedron [Figure I-2] is a geometric figure whose shape is close to a sphere. It is made up of 20 identical equilateral triangles and 12 vertices. It has several sets of axes of symmetry by rotation passing through its center: 6 axes of order 5, 10 axes of order 3 and 15 axes of order 2. An axis of order 5 (respectively 3) means that a rotation by an angle $\frac{2\pi}{5}$ (respectively $\frac{2\pi}{3}$) leaves the figure unchanged. Axes of order 5 are located at each vertex, axes of order 3 at the center of each of the 20 triangular faces and axes of order 2 in the middle of each of the 30 edges. An icosahedron is made up of 60 asymmetric units. From an asymmetric unit and by applying the symmetries mentioned above, we build the icosahedron.

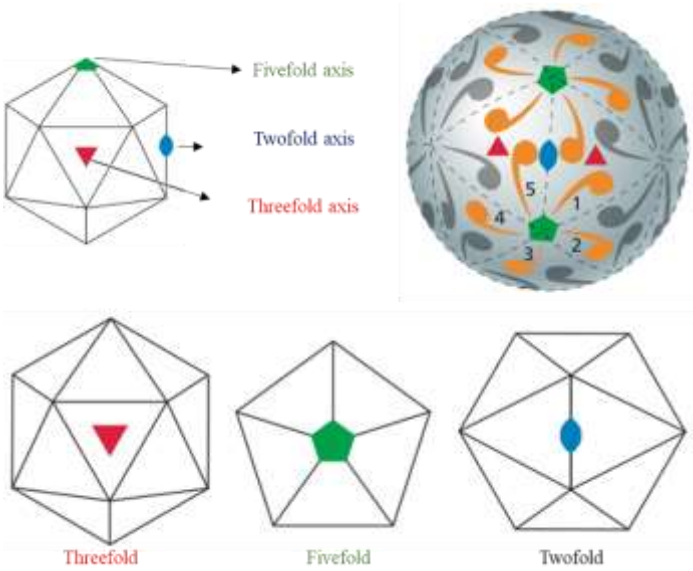


Figure I-2. Simplified representation of an icosahedral capsid with 20 faces, each constituted by an equilateral triangle. The pattern is in T = 1 triangulation. Axes of order 5 are represented in green, of order 3 in red, and those of order 2 in blue

I.6.2 TRIANGULATION NUMBER

The simplest icosahedral symmetry capsid is obtained with 60 copies of the major head protein. These protein subunits are organized into 12 pentamers, located at the vertices of the icosahedron. These pentamers are all equivalent, both chemically and with respect to their interactions. The major capsid proteins of all phages and icosahedral capsid viruses are made up of a structurally conserved protein fold including comparable number of amino acids, about 300 [19]. Larger size capsids are therefore composed of additional copies of the major head protein. In such capsids, a unit is composed of a number of protein greater than 1. This number is called the triangulation number "T" [20]. This number, characterizes the size of the capsids. It also indicates the number of hexamers between two pentamers using the formula:

$$T = h^2 + k^2 + h \times k \qquad (1)$$

where, h and k are two positive and non-zero integers, as defined in the Table I.2. Indeed, in addition to pentamers forming the top angles, capsids with triangulation number greater than one feature hexamer planes, thus forming the faces of the icosahedron. In that case, the major head proteins of an asymmetric unit are chemically identical but do not form exactly the same interactions with the other major head proteins. For example, a major head protein located in a pentamer interacts with six other major proteins, and a major protein located in a hexamer interacts with seven other major proteins (See T=13 of Table I.2). This slight variation in the physico-chemical environment of the major proteins of capsid is called 'quasi-equivalence'.

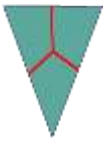
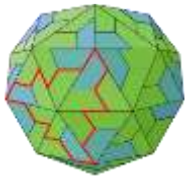
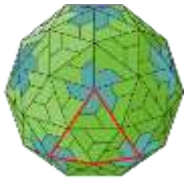
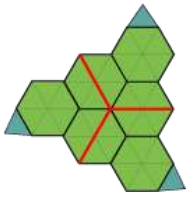
Structural Unit	Organisation at 20-fold axis	Capsid	Total number of subunits[60T]	[h,k]
			60	[1,0]
X60 T=1 icosahedral assymetric unit	X20 Traingulation facet	T=1		
			180	[1,1]
X60 T=3 icosahedral assymmetric unit	X20 Triangular facets	T=3		
			240	[2,0]
X60 T=4 icosahedral assymmetric unit	X20 Triangular facets	T=4		
			780	[3,1]
X60 T=13 Icosahedral assymmetric unit	X20 Triangular facets	T=13		

Table I.2. Triangulation number "T". Examples of icosahedron with different triangulation numbers T= 1,3,4,13. Reprinted from [18]

I.7 ASSEMBLY AND MATURATION OF THE PHAGE PARTICLE

I.7.1 PHAGE LIFE CYCLE

The infectious cycle of bacteriophages begins with the attachment of the virus to its host's membrane receptor, followed by ejection of the viral DNA from the capsid and its transfer into the bacterial cytoplasm. At this stage, two types of bacteriophages can be distinguished: lytic

and temperate phages. Lytic, (coming from Greek, meaning "to loosen, to cut apart, to divide") or virulent phages only perform lytic cycles, whereas temperate or non-virulent phages, depending on the infection conditions, can perform lytic or lysogenic cycles [21] [see Figure I-3]. During a lytic cycle, the biosynthetic machinery of the bacteria is hijacked by the viral genome, allowing the viral genome replication, viral protein synthesis and leading to the assembly of new infectious particles. The release of mature virions is ultimately achieved through bacterial lysis. During a lysogenic cycle, the viral genome is integrated into the bacterial genome, becoming a so-called prophage. This prophage is unable to replicate autonomously, but is copied at the same time as the bacterial genome during cell division. A protein called a repressor prevents transcription of viral genes. However, under certain conditions of stress, the viral genome becomes active and "comes out" of the bacterial genome, which triggers the lytic cycle[22].

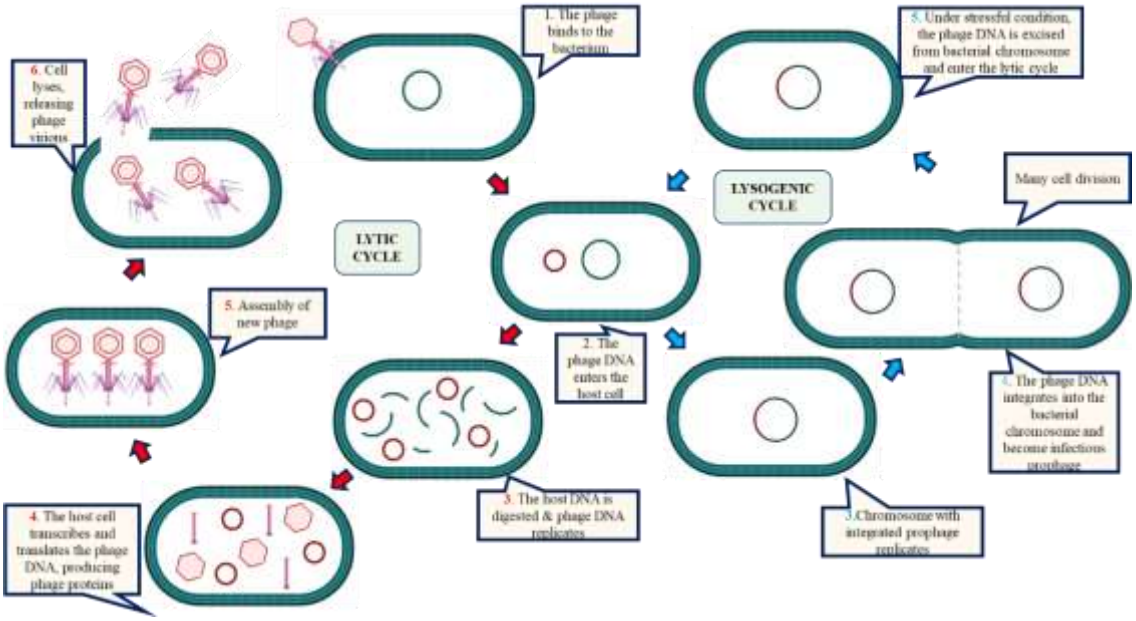


Figure I-3. Schematic diagram of phage life cycle. Red arrow indicates the Lytic cycle and blue arrow indicates the Lysogenic cycle[23] [24]

I.8 VIRION ASSEMBLY

The assembly process for an infectious virion is largely the same for all phages of the order *Caudovirales* [25] [Figure I-4]. It points towards several processes:

- Formation of a DNA-free protein shell, called a procapsid.
- Removal of the "scaffolding" protein or domain that support the formation of this procapsid
- Viral DNA packaging by a molecular motor into the capsid. This process triggers major conformational rearrangement of the capsid proteins, resulting in the capsid expansion and its inner volume increase.
- A possible strengthening of the stability of the capsid, by adding accessory cementing. (bacteriophage *lambda*) or by autocatalytic covalent cross-linking between capsomers (bacteriophage *HK97*). In some phages, there is also binding of proteins called decoration proteins. The role of these non -essential proteins is still poorly understood [26].

- The closure of the capsid is done by one or more proteins called head-completion proteins.
- Assembly of the tail: In *Podoviridae*, tail proteins are attached to the capsid sequentially [27]. For *Siphoviridae* and *Myoviridae*, the tail is assembled independently and then attached to the capsid, at the portal.

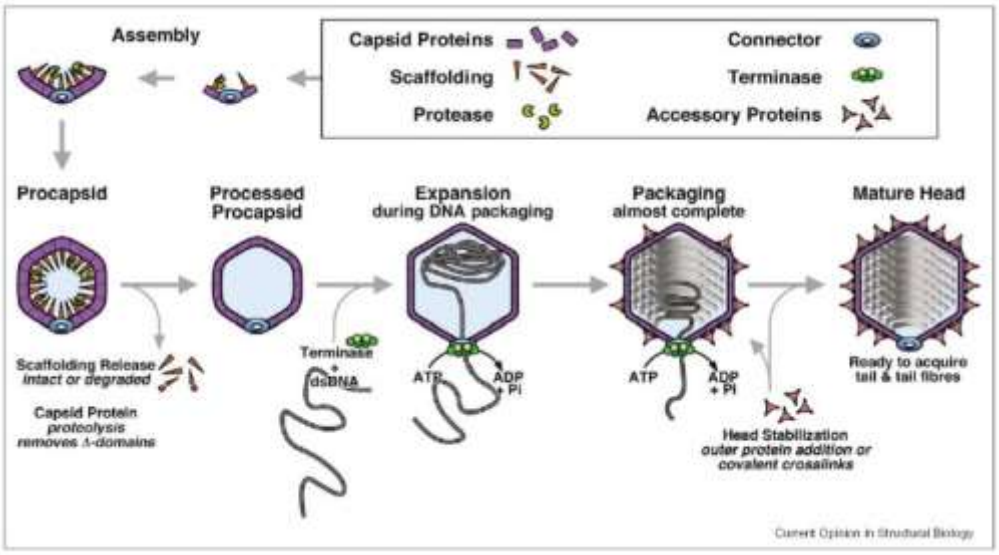


Figure I-4. . Schematic diagram representing the generic assembly pattern of the capsid of tailed phages. Reprinted from Ref [28]

I.8.1 PROCAPSIDS

The procapsid is made up of the portal complex, multiple copies of the major capsid protein as well as multiple copies of a protein or domain called scaffolding protein or domain. In most phages, the major capsid protein subunits alone cannot self-assemble alone into a capsid of correct geometry. There occurs the necessity of scaffolding proteins or domains to "mount" this assembly [29]. Indeed, it was shown for phage φ29 with a mutation of the scaffolding protein leads to the formation of isometric (not elongated) capsids, which are too small to incorporate the entire phage genome [30]. In phage T4, the absence of this protein leads to the formation of large cylindrical structures, not viable for further assembly of the virion [31]. The structure of the phage φ29 scaffolding protein was determined by crystallography [32]. It indicates that it is a homodimer structured in coiled coil conformation, which interacts with the major leader protein via ionic interactions. In phages T5 and HK97, the scaffolding protein is not encoded by a separated gene, but this function is performed by an N-terminal extension of the major capsid protein called the scaffolding domain [33] [34]. The mechanisms by which the scaffolding protein or domain acts on the shape of the procapsid are still poorly understood. Yet, it is likely that procapsid assembly is initiated from the portal [35]. This structure, by which DNA enters and leaves the capsid, is formed by twelve copies of the portal protein [36]. For T4, the interaction with the bacterial membrane happens via an additional scaffolding protein, and viral assembly takes place from this nucleation point [37]. However, this membrane anchoring of procapsids during assembly has not been demonstrated in any other phage. In addition, some phages, such as P22, T7, P4 [38] and HK97 [39] have the ability to form procapsids even in the absence of portal. Nevertheless, these procapsids are not able to encapsulate DNA and thus form infectious virions.

I.8.2 CAPSID MATURATION

Following the formation of the procapsid, the scaffolding domains or proteins are removed. In the case of bacteriophages φ29 [40] and T7 [41], these proteins leave the capsid intact. In phage φ29, the output of this protein could be induced by interactions with the viral DNA, and would therefore accompany its encapsulation [42]. In phages HK97, P2, λ, T4 and T5, the scaffolding proteins or domains are cleaved by a phage protease and the resulting peptides removed prior to packaging of the genome and capsid expansion. An enzymatic complex called terminase transfers the DNA into the capsid. It consists of two subunits: a subunit called the small terminase, which recognizes the DNA to be encapsulated, and a catalytic subunit called the large terminase whose ATPase activity provides the energy to transport the DNA through the portal gate. The driving force produced by this powerful molecular motor was measured to be greater than 60pN by using optical tweezers [43]. The speed of encapsidation in vitro varies according to the quantity of genome to be packaged, from 180 (φ29, genome size: 19.3 kbp) to 1800 (T4, size of the genome: 168.9 kbp) base pairs per second, allowing packaging of the entire genome in 2-3 minutes [44]. The large terminase subunit also has a DNAse activity that cleaves DNA concatamers produced by the viral DNA replication, allowing a unit of genome to enter each capsid [45].

I.9 FINAL STAGE: VIRION ASSEMBLY

The last step in the assembly process is sealing of the capsid by specific proteins. The best-described case is that of phage SPP1, in which two homo-dodecamers called connector and plug attach to the portal to form a complex called gatekeeper, preventing the premature release of DNA from the capsid [46]. Two connector proteins (gp4 and gp10) and a plug protein (gp26) also close the phage P22 capsid [47]. These proteins constitute the tail attachment site, yielding infectious virions. Interestingly, some phage like T5 use a single protein to play this gate keeper function[48].

I.10 CONCLUSION

In this chapter, after having a brief introduction to bacteriophages, we presented structural information regarding the Caudovirales capsid assembly mechanism. Bacteriophages assembly has been mostly studied using biochemical and structural biology approaches, but only a few mass spectrometry (MS)-based investigations have been reported. In the present work, we used mass spectrometric-based proteomics approaches to study a specific protein involved in bacteriophage T5 assembly, and nanomechanical mass spectrometry as a tool to assess capsid structural integrity.

REFERENCES

1. Suttle C (2007) Marine viruses—Major players in the global ecosystem. Nature reviews Microbiology 5:801–12 . https://doi.org/10.1038/nrmicro1750

2. Weinbauer M, Rassoulzadegan F (2004) Weinbauer MG, Rassoulzadegan F.. Are viruses driving microbial diversification and diversity? Environ Microbiol 6: 1. Environmental microbiology 6:1–11 . https://doi.org/10.1046/j.1462-2920.2003.00539.x

3. Pasteur PI, scientifique (France) C national de la recherche, Institut Pasteur (Paris F (1887) Annales de l'Institut Pasteur. Paris :Masson,

4. Twort FW (1915) An Investigation on the Nature of Ultra-Microscope Viruses. The Lancet 186:1241–1243 . https://doi.org/10.1016/S0140-6736(01)20383-3

5. Publications service (2007) On an invisible microbe antagonistic toward dysenteric bacilli: brief note by Mr. F. D'Herelle, presented by Mr. Roux. Research in Microbiology 158:553–554 . https://doi.org/10.1016/j.resmic.2007.07.005

6. Summers WC (2012) The strange history of phage therapy. Bacteriophage 2:130–133 . https://doi.org/10.4161/bact.20757

7. Pellett S (2012) Learning from the past: historical aspects of bacterial toxins as pharmaceuticals. Current Opinion in Microbiology 15:292–299 . https://doi.org/10.1016/j.mib.2012.05.005

8. Sulakvelidze A, Alavidze Z, Morris JG (2001) Bacteriophage Therapy. Antimicrob Agents Chemother 45:649–659 . https://doi.org/10.1128/AAC.45.3.649-659.2001

9. Calendar R (2015) Bacteriophages: Biology and Applications. Edited by Elizabeth Kutter and, Alexander Sulakvelidze. Boca Raton (Florida): CRC Press. $129.95. xvi + 510 p; ill.; index. ISBN: 0-8493-1336-8. 2005. | The Quarterly Review of Biology: Vol 81, No 3. The Quarterly Review of Biology. https://doi.org/10.1086/509419

10. Verbeken G, Pirnay J-P, Lavigne R, Jennes S, De Vos D, Casteels M, Huys I (2014) Call for a Dedicated European Legal Framework for Bacteriophage Therapy. Arch Immunol Ther Exp 62:117–129 . https://doi.org/10.1007/s00005-014-0269-y

11. Debarbieux L, Pirnay J-P, Verbeken G, De Vos D, Merabishvili M, Huys I, Patey O, Schoonjans D, Vaneechoutte M, Zizi M, Rohde C (2016) A bacteriophage journey at the European Medicines Agency. FEMS Microbiol Lett 363:fnv225 . https://doi.org/10.1093/femsle/fnv225

12. Rohwer F (2003) Global Phage Diversity. Cell 113:141 . https://doi.org/10.1016/S0092-8674(03)00276-9

13. Curtis TP, Sloan WT, Scannell JW (2002) Estimating prokaryotic diversity and its limits. Proc Natl Acad Sci U S A 99:10494–10499 . https://doi.org/10.1073/pnas.142680199

14. Abrescia NGA, Bamford DH, Grimes JM, Stuart DI (2012) Structure Unifies the Viral Universe. Annual Review of Biochemistry 81:795–822 . https://doi.org/10.1146/annurev-biochem-060910-095130

15. Ackermann H-W (2009) Phage classification and characterization. Methods Mol Biol 501:127–140 . https://doi.org/10.1007/978-1-60327-164-6_13

16. Claude Fauquet M.A. Mayo J. Maniloff U. Desselberger L.A. Ball (2005) Virus Taxonomy - 1st Edition. https://www.elsevier.com/books/virus-taxonomy/fauquet/978-0-08-057548-3

17. Ackermann H-W (2007) 5500 Phages examined in the electron microscope. Arch Virol 152:227–243 . https://doi.org/10.1007/s00705-006-0849-1

18. Tequintavirus ~ ViralZone. https://viralzone.expasy.org/511

19. Suhanovsky MM, Teschke CM (2015) Nature's favorite building block: Deciphering folding and capsid assembly of proteins with the HK97-fold. Virology 0:487–497 . https://doi.org/10.1016/j.virol.2015.02.055

20. Caspar DL, Klug A (1962) Physical principles in the construction of regular viruses. Cold Spring Harb Symp Quant Biol 27:1–24 . https://doi.org/10.1101/sqb.1962.027.001.005

21. Dodd IB, Shearwin KE, Egan JB (2005) Revisited gene regulation in bacteriophage lambda. Curr Opin Genet Dev 15:145–152 . https://doi.org/10.1016/j.gde.2005.02.001

22. Oppenheim AB, Kobiler O, Stavans J, Court DL, Adhya S (2005) Switches in Bacteriophage Lambda Development. Annual Review of Genetics 39:409–429 . https://doi.org/10.1146/annurev.genet.39.073003.113656

23. Aman B, Ciobanu G Mutual Mobile Membranes Systems with Surface Objects. 12

24. The Viral Life Cycle | Microbiology. https://courses.lumenlearning.com/microbiology/chapter/the-viral-life-cycle/. Accessed 24 Jul 2021

25. Aksyuk AA, Rossmann MG (2011) Bacteriophage Assembly. Viruses 3:172–203 . https://doi.org/10.3390/v3030172

26. Dedeo CL, Teschke CM, Alexandrescu AT (2020) Keeping It Together: Structures, Functions, and Applications of Viral Decoration Proteins. Viruses 12:1163 . https://doi.org/10.3390/v12101163

27. Lander GC, Khayat R, Li R, Prevelige PE, Potter CS, Carragher B, Johnson JE (2009) The P22 Tail Machine at Subnanometer Resolution Reveals the Architecture of an Infection Conduit. Structure 17:789–799 . https://doi.org/10.1016/j.str.2009.04.006

28. Steven AC, Heymann JB, Cheng N, Trus BL, Conway JF (2005) Virus maturation: dynamics and mechanism of a stabilizing structural transition that leads to infectivity. Curr Opin Struct Biol 15:227–236 . https://doi.org/10.1016/j.sbi.2005.03.008

29. Fane B, Prevelige P (2003) Mechanism Of Scaffolding-Assisted Viral Assembly. Advances in protein chemistry 64:259–99 . https://doi.org/10.1016/S0065-3233(03)01007-6

30. Choi KH, Morais MC, Anderson DL, Rossmann MG (2006) Determinants of bacteriophage phi29 head morphology. Structure 14:1723–1727 . https://doi.org/10.1016/j.str.2006.09.007

31. Kellenberger E (1990) Form determination of the heads of bacteriophages. European Journal of Biochemistry 190:233–248 . https://doi.org/10.1111/j.1432-1033.1990.tb15568.x

32. Morais MC, Kanamaru S, Badasso MO, Koti JS, Owen BAL, McMurray CT, Anderson DL, Rossmann MG (2003) Bacteriophage phi29 scaffolding protein gp7 before and after prohead assembly. Nat Struct Biol 10:572–576 . https://doi.org/10.1038/nsb939

33. Duda RL, Hempel J, Michel H, Shabanowitz J, Hunt D, Hendrix RW (1995) Structural transitions during bacteriophage HK97 head assembly. J Mol Biol 247:618–635 . https://doi.org/10.1006/jmbi.1995.0168

34. Huet A, Conway JF, Letellier L, Boulanger P (2010) In Vitro Assembly of the T=13 Procapsid of Bacteriophage T5 with Its Scaffolding Domain. J Virol 84:9350–9358 . https://doi.org/10.1128/JVI.00942-10

35. Valpuesta JM, Carrascosa JL (1994) Structure of viral connectors and their function in bacteriophage assembly and DNA packaging. Q Rev Biophys 27:107–155 . https://doi.org/10.1017/s0033583500004510

36. Johnson JE, Chiu W (2007) DNA packaging and delivery machines in tailed bacteriophages. Curr Opin Struct Biol 17:237–243 . https://doi.org/10.1016/j.sbi.2007.03.011

37. Michaud G, Zachary A, Rao VB, Black LW (1989) Membrane-associated assembly of a phage T4 DNA entrance vertex structure studied with expression vectors. Journal of Molecular Biology 209:667–681 . https://doi.org/10.1016/0022-2836(89)90599-8

38. Dokland T, Wang S, Lindqvist BH (2002) The structure of P4 procapsids produced by coexpression of capsid and external scaffolding proteins. Virology 298:224–231 . https://doi.org/10.1006/viro.2002.1485

39. Conway JF, Duda RL, Cheng N, Hendrix RW, Steven AC (1995) Proteolytic and conformational control of virus capsid maturation: the bacteriophage HK97 system. J Mol Biol 253:86–99 . https://doi.org/10.1006/jmbi.1995.0538

40. Morphogenesis of bacteriophage phi 29 of Bacillus subtilis: oriented and quantized in vitro packaging of DNA protein gp3. https://journals.asm.org/doi/epdf/10.1128/jvi.45.1.383-396.1983

41. Cerritelli ME, Conway JF, Cheng N, Trus BL, Steven AC (2003) Molecular Mechanisms In Bacteriophage T7 Procapsid Assembly, Maturation, And Dna Containment. In: Virus Structure. Academic Press, pp 301–323

42. Morais MC, Kanamaru S, Badasso MO, Koti JS, Owen BAL, McMurray CT, Anderson DL, Rossmann MG (2003) Bacteriophage φ29 scaffolding protein gp7 before and after prohead assembly. Nat Struct Mol Biol 10:572–576 . https://doi.org/10.1038/nsb939

43. Sun S, Gao S, Kondabagil K, Xiang Y, Rossmann MG, Rao VB (2012) Structure and function of the small terminase component of the DNA packaging machine in T4-like bacteriophages. PNAS 109:817–822 . https://doi.org/10.1073/pnas.1110224109

44. Single-molecule studies of viral DNA packaging. - Abstract - Europe PMC. https://europepmc.org/article/med/22297530

45. Rao VB, Feiss M (2008) The bacteriophage DNA packaging motor. Annu Rev Genet 42:647–681 . https://doi.org/10.1146/annurev.genet.42.110807.091545

46. Structure of a viral DNA gatekeeper at 10 Å resolution by cryo-electron microscopy. https://www.embopress.org/doi/epdf/10.1093/emboj/cdg123. Accessed 24 Jul 2021

47. Olia A, Bhardwaj A, Joss L, Casjens S, Cingolani G (2007) Role of Gene 10 Protein in the Hierarchical Assembly of the Bacteriophage P22 Portal Vertex Structure †. Biochemistry 46:8776–84 . https://doi.org/10.1021/bi700186e

48. Zivanovic Y, Confalonieri F, Ponchon L, Lurz R, Chami M, Flayhan A, Renouard M, Huet A, Decottignies P, Davidson AR, Breyton C, Boulanger P (2014) Insights into Bacteriophage T5 Structure from Analysis of Its Morphogenesis Genes and Protein Components. Journal of Virology 88:1162–1174 . https://doi.org/10.1128/JVI.02262-13

CHAPTER II: **ANALYTICAL METHODS**

II.1 **MASS SPECTROMETRY**

Mass spectrometry (MS) is an analytical chemistry technique to characterize chemical structures based on their mass to charge ratio (m/z). In 1897, Sir Joseph John Thomson made an apparatus to measure the charge to mass ratio (denoted as e/m) of electrons [1]. After a couple of years, with his assistant E. Everett, he built an instrument able to measure e/m and m simultaneously. This work was rewarded by the Noble Price in 1906 for the discovery of the electron [2]. Thomson's earlier research work on cathode rays forms the basis of the field of MS. Sir Thomson together with Francis Aston conducted an experiment to determine the characteristics of positively charged particles. Their instrument was later recognized as the first mass spectrometer. Later on, Aston redesigned the system (see *Figure* II-1) to improve its resolving power and proved the existence of elemental isotopes [3].

Figure II-1. Aston's third mass spectrograph [3]

Until the1940's, physicists were using MS to study the fundamentals of atoms. Around 1950, mass spectrometers became commercially available and the technique was used by physicists and chemists alike [4]. By 1980's, many organic molecules were routinely analyzed by MS [5]. However, proteins and other fragile bio-macromolecules remained a challenge because there was no way to ionize these large particles without disintegration. In 1988, Matrix-assisted laser desorption/ionization (MALDI) and Electrospray ionization (ESI) appeared on the scene. Macromolecules that had so far escaped mass measurements became amenable to MS analysis using these ionization techniques. In 2002, John Fenn and Koichi Tanaka shared a Noble Price in Chemistry for the development of these methods. On one hand, ESI produces intact gaseous ions of biomolecules using either transfer of nonvolatile solutes from highly charged droplets of volatile solvents [6] [7], and on the other hand MALDI performs laser desorption of fragile biomolecules from a solid matrix absorbing at the laser wavelength [8]. By 1989, Fenn and his team were able to ionize macromolecules having a molecular weight of 130,000 Da using ESI – MS [9]. Almost at the same time, Franz Hillenkamp and Michael Karas were developing MALDI to overcome the same issues [8].

Since then, MS has become one of the fastest growing analytical technique. This technique has become so significant that it is now used in a wide range of fields : to track contaminants in the environment[10], study biochemical processes taking place in nature [11], or even to support different programs in outer space [12]. MS is unique among spectrometric and spectroscopic techniques because it is able to characterize a broad diversity of biological processes at multiple

levels (nucleic acids, proteins, lipids, sugars, metabolites). This method is thus used as a routine analytical tool in the field of modern biology and analytical chemistry, ranging from fundamental research to clinical studies, as well as in routine disease screening. Recently MS has become a contributor in the field of structural virology especially with the development of high-resolution native MS and single particle mass sensing [13].

In MS, ionized particles such as atoms, molecules, or clusters, are separated using differences in their mass to charge ratios (m/z), which can be used to derive their molecular weight [14]. A mass spectrometer consists of three major components (see *Figure* II-2):

- Ion source – in which analytes are converted into their gaseous ionized state.
- Analyzer – where ions are separated according to their m/z using electromagnetic fields.
- Detector – which responds to ions in a quantitative manner, and provides estimates of ion abundances as a function m/z.

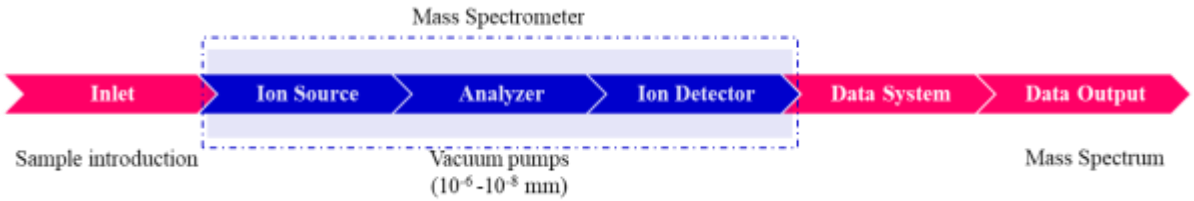


Figure II-2. Components of a Mass Spectrometer (in blue)

In the present work, two MS-related techniques have been used: 1. Nano-scale Liquid Chromatography-tandem Mass Spectrometry (nano LC-MS/MS) based proteomics, to reveal the protein composition of virus capsids and the stoichiometry of a viral protein, and 2. Nano Electro Mechanical System based Mass Spectrometry (NEMS MS) a method to achieve mass measurement of intact viral capsid in the Mega- to Giga-Dalton range.

II.2 LIQUID CHROMATOGRAPHY-TANDEM MASS SPECTROMETRY

Modern biology requires comprehensive investigations of the all the constituents of living organisms, including nucleic acids, lipids, carbohydrates, metabolites, and proteins. Powerful tools like Polymerase chain reaction (PCR) combined with Next Generation Sequencing (NGS) allow routine oligonucleotides sequencing and characterization [15]. However, oligonucleotides sequencing is not enough to account for many biological phenomena such as gene transcription or post-translational modifications, which determine the assembly of proteins and their function. Knowledge of protein dynamics is thus vital for understanding these molecular effectors of life.

The field of proteomics studies the entire protein complement of an biological system (e.g. cell, tissue, organism, sub-cellular compartment), with the aim of identifying and quantifying specific proteins, and obtain information about their roles as a part of a larger network of proteins [16]. Proteomics heavily relies on MS, which is a proven characterization technique for proteins in complex biological samples [17]. There are several proteomics based MS methods such as: Matrix Assisted Laser Desorption Ionization coupled with Time-of-Flight (MALDI-TOF), Liquid Chromatography tandem MS (LC-MS/MS) or Surface Enhanced Laser Desorption Ionization (SELDI) [18]. They can be categorized into bottom-p and top-down proteomics [19]. Herein we will focus on LC-MS/MS based Proteomics, the most commonly used “bottom-up” method.

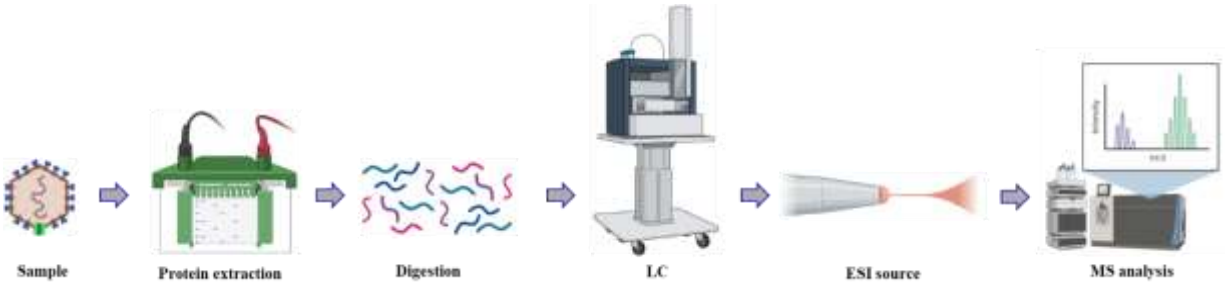


Figure II-3. LC-MS/MS based proteomics overview. Following protein extraction and purification using e.g. SDS-Page (Sodium Dodecyl Sulphate–Polyacrylamide Gel Electrophoresis), the proteins are digested by trypsin, and the resulting peptides separated by reverse phase liquid chromatography, ionized by electrospray, and analyzed by MS.

In this method, the proteins extract of interest undergoes enzymatic digesting, yielding proteolytic peptides. The resulting peptides are analyzed by liquid chromatography coupled to mass spectrometry [20]. A schematic representation of how the experiment proceeds is shown in *Figure* II-3. First, the sample is prepared and proteins are separated from contaminants (e.g. nucleic acids, lipids) using e.g. gel electrophoresis. The proteins of interested are subjected to digestion by trypsin, and the resulting peptides passed through a reverse phase liquid chromatography (RP-LC) column. The eluted peptides are ionized by ESI and then introduced in the mass spectrometer. Each scan provides a mass spectrum consisting of peak intensities versus *m/z*. Tandem mass spectrometry (MS/MS) consists in the dissociation of selected ions *m/z* present in the acquired parent ion spectra (MS1) and the collection of the corresponding fragment ion spectra (MS2).

The peptide's measured *m/z* and the list of *m/z* detected in the associated tandem mass spectrum are compared with entries in a database of known protein sequences using a software called Mascot™ [21]. Protein abundances can be evaluated in different ways from the list of identified peptides. A simple way is to compute the number of times peptides corresponding to the given protein are detected and aggregate those peptide counts [22]. Another way is by calculating the peak area of the extracted ion chromatogram (XIC) for each peptide of the protein.

Quantification of peptides abundances is possible using label-free quantification strategies or stable-isotope label strategies [23] [24]. In label-free, relative quantification such as spectral counting of the identified peptide is used to have an estimation on the protein abundance. But this technique is less reliable for low abundant proteins having few spectral counts [25]. Another available relative quantification method is Accurate Mass Retention (AMRT), which is based on intensity of the peptide observed in the MS spectrum. But the results can be biased based on reproducibility in sample preparation and LC-MS operation [26]. Conversly, labelled quantification relies on stable isotopes. It was in 1999 that quantification using isotopic labeling of proteins or peptides was introduced [27]. In this technique, peptides having similar physical property as the peptide of interest are coupled with various isotopes. These peptides can be identified in the same spectrum but with a mass shift, which is based on the labeled isotope. Quantification of these peptides can be made by the MS by comparing the ratio of the XIC value for the native peptide of interest and its heavy-isotope labelled analog added to the sample. Integration of isotopes can be done in three different way: (1) *In vitro or In vivo* labelling where the corresponding atoms or labelled amino acids are added into the growth medium during the cell culture; (2) Chemical labelling, where the isotope is introduced chemically; at the peptide level or at the full length protein level. Stable isotopes such as ^{13}C, ^{2}H, ^{18}O or ^{15}N are used for this purpose.

For the quantification of protein, enzyme-linked immunosorbent assay (ELISA) has been used as a reference method as a targeted quantification method [28]. However, this method had limitations with respect to the accessibility and cross reactivity of antibodies and cost of production. Label free proteomic techniques has the ability to quantify many proteins in the samples of interest. But then again interpretation of the data using this method had to be done very cautiously because of the uncertainties in abundance measurements. On the other hand, MS quantification using internal standards concerns one or few proteins but is not suitable for global proteomics. This can be performed using an absolute quantification of proteins (AQUA) strategy, where an isotopically labelled analog of the peptide of interest is used as an internal standard [29]. In this technique, a peptide is selected from the protein of interest; this peptide is synthesized *de novo* bearing isotopes at its amino acids. AQUA peptide thus have same physical properties as the corresponding native peptide (such as charge, hydrophobicity, size and ionization). The sample to be analyzed is added with a known quantity of AQUA standard peptide and analyzed by LC-MS. MS is able to distinguish the mass of the standard and the native peptide. By comparing the ratio of native and AQUA peptide, peptide quantity is calculated. The same technique when performed with full length protein standard, is called as protein standard for absolute quantification (PSAQ) [30]. When compared with AQUA, PSAQ is more expensive because it requires the production of isotopic labeled full length proteins.

To simultaneously quantify peptides multiple proteins a quantification using concatenated peptides (QconCAT) were proposed, where concatemers of signature peptides are chemically synthesized. Here, two or more a unique peptide is selected for each of the targeted protein. The QconCAT is added to the sample and digested, which generates peptides containing labeled peptides standards as internal standards [31].

II.3 NANO ELECTRO MECHANICAL SENSOR BASED MASS SPECTROMETRY

Like all pharmaceutical products, Virus Like Particles (VLP) used in vaccine applications require tight quality controls to validate their structural integrity and immunogenicity. VLP structure assessment mostly relies on particle sizing or morphological characterization methods (e.g. Light Scattering, electron microscopy). However, these methods either suffer from imperfect metrics (1 or 2-D rendering of 3-D structures), selection biases, and/or limited statistics. Interestingly, VLP result from tightly auto-regulated assembly processes, yielding strictly identical particles of definite molecular composition, making them amenable to MS analysis.

Unfortunately, no commercial MS system operates in the mega- to giga-Dalton mass range which is required for VLP analysis. In a review article in 2009, Wang discussed the lack of technology to deal with masses of bio-particles like virus, bacteria, cellular organelles, that lie in the MDa to GDa range, although their study would have significant impact on the biomedical field [32]. In another article published in 2012, Heck and coworkers explained that even though there has been a lot of improvements with native MS, the demand of MS capable of measuring MDa complexes is vital especially in the field of bio-therapeutics [33]. Although native-MS could indeed measure the mass of pure bio-nanoparticles up to few MDa, it requires charging, and involves ensemble averaging of mass to charge ratios (m/z), yielding complex spectra for high mass and heterogeneous samples [27]. Charge detection MS (CDMS) avoids some of these issues by measuring individual particles charge and m/z ratio, but also requires charging, and

long acquisition times [34]. Finally, gas-phase electrophoresis can derive the mass of such particles, but relies on a compound-specific size to mass calibration [35].

Nano-electro mechanical systems (NEMS)-based MS constitutes a promising technology for this challenging mass range. NEMS sensors are capable of detecting the mass of particles accreting on their surface through changes in their resonance frequency [36] [37] [38]. In the 1990's, micro-cantilevers as mass sensors were proposed by Chen et al. and Thundat et al. [39]. At that time, MEMS were able to detect adsorbed material with picogram mass resolution. Recently with the development of nano-mechanical resonators, mass measurement is possible in the femtogram to attogram (Giga- to Mega Dalton) mass range. Such mass sensitivity enables a wide range of technological applications including the analysis of biological or artificial nanoparticles such as viruses, or soot particles for environmental purposes.

Early mass sensing experiments typically showed the effect of a single mass addition on the device's resonance frequency, but none had the capability to perform actual MS analyses by measuring the mass of many single particles in real time and aggregating the results in an actual 'mass spectrum'. In 2009, the first nanoelectromechanical system based mass spectrometer (NEMS MS) was developed by Roukes et al. [40]. In 2012, in collaboration with CEA Grenoble, this team achieved the analysis of IgM antibody aggregates on a similar system using hexapole ion guides to transfer analytes to a single nano-resonator, albeit with very low capture efficiency, and prohibitively large sample consumption and analysis time. [41].

In 2013, researchers at CEA Grenoble patented a new concept for NEMS-based mass spectrometry taking into account remarkable characteristics of these nano-sensors [42]. NEMS can act both as detector and analyzer, and unlike conventional mass analyzers, they do not require charged particles, thus circumventing the need for electromagnetic fields. This capability, which had never been fully taken advantage of, was demonstrated experimentally in 2015 by Sage et al. [43], paving the way to a new type of high efficiency NEMS based architecture independent on particle charge. As their name implies, NEMS are very small devices, therefore requiring tightly focused particle beams to efficiently capture analytes. To mitigate this issue, arrays of frequency addressed nano-resonators were proposed and implemented [44]. In addition, using highly efficient particle nebulization methods and an aerodynamic lens was proposed to focus the particle beam onto a NEMS array. All these advances were implemented in a novel NEMS-MS prototype that ultimately allowed the measurement of the mass of bacteriophage capsids with and without their genome at 27 and 108 MDa respectively. [45]

Figure II-4 displays the system architecture used in these measurements. Briefly, it consists in 3 consecutive sections with differential pressures: 1. An atmospheric nebulization source in which the particles in solution are transferred into the gas phase using Surface Acoustic Wave Nebulization or nano-Electrospray. 2. A focusing stage consisting in an inlet capillary, an aerodynamic lens and a skimmer to collimate particles into a tight beam. 3. A detector array of 20 nanoresonators situated in high vacuum (10^{-5} Torr). We will describe these three stages in more details in the following sections.

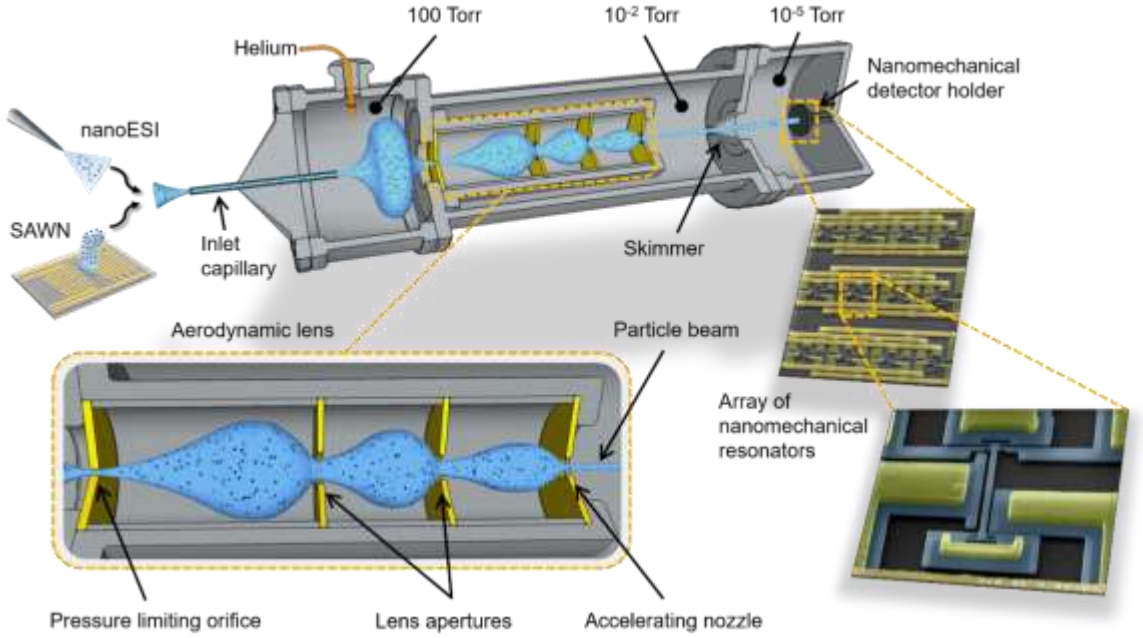


Figure II-4 Novel architecture or nano-resonator mass spectrometer (reprinted from [35]) The setup consists of three differentially pumped chambers. Analytes in solution are nebulized at atmospheric pressure and aspirated through a heated metal capillary inlet. An aerodynamic lens focuses the particle stream in its carrier gas (blue shade), and transfers them onto an array of frequency-addressed nanomechanical resonators. Bottom left: Simplified scheme of the aerodynamic lens. Bottom right: False color SEM image of a portion of a 20 NEMS array and magnified SEM image of a single resonator. Metallic layer in yellow and silicon in gray-blue.

II.3.1 NEBULIZATION STAGE

Nebulization can be carried out by surface acoustic wave nebulization (SAWN) or electrospray ionization (ESI). In this work, nanoelectrospray ionization (nano-ESI) was used for analytes nebulization, enabling lower sample consumption and flow rates.

ESI is a commonly used soft ionization technique, which enables the analysis of intact biomolecules using MS. This technique was a revolution because it was able to convert fragile analytes from liquid phase to their electrically charged gas phase state without fragmentation. Thanks to ESI-MS, scientist have been able to measure the mass of macromolecules up to the MDa range since 1998 [46].

In ESI, a spray of droplets is produced using a high voltage applied to the liquid at the tip of a capillary (*Figure* II-5). At the outlet of the capillary, the liquid cross section decreases due to charge repulsion, creating a so-called Taylor cone. Then, the flow becomes a jet, which is atomized into residue particles. The generated primary charged droplets experience Coulombic repulsion, which induce expansion of the spray. The solvent of the droplets is quickly evaporated, increasing the charge density. When the droplets charge overcomes the Rayleigh limit, progeny droplets experience jet fission. Evaporation and fission are the two mechanisms leading to the formation of nanoparticle aerosols. This atomization process depends on many parameters such as the electric field, the diameter of the capillary tip, the solution flow-rate or the properties of the solvent (for example: viscosity, volatility) [47].

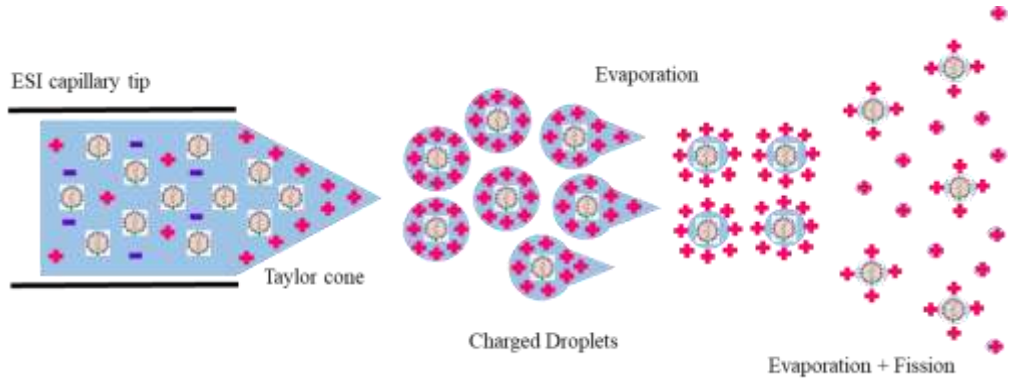


Figure II-5. Sketch of the electrospray ionization mechanism. Adapted from [48]

II.3.2 AERODYNAMIC FOCUSING LENS

To achieve efficient particle detection using NEMS resonators, the ratio of particle beam to detector size should be low, because of the small detection area of the resonator. For this purpose, an aerodynamic focusing lens is used in the NEMS-MS system. This method is generally used for aerosol sampling and transfer, or to achieve precise deposition of micro patterns on a substrate [49] [50]. Since this method can focus nanoparticle over a wide range of size, it is broadly applied in aerosol characterization, including by MS, and material synthesis [51] [52].

An aerodynamic focusing lens consists in a series of pressure reducing orifices placed between multiple relaxation chambers (*Figure* II-6) [53] . The device features a pressure limiting orifice where the particle flux from ESI is received and an accelerating nozzle facing the NEMS resonator array. Narrowing of the nanoparticle beam happens when the carrier gas undergoes several consecutive contractions and expansions. At each constriction, the carrier gas pushes particles closer to the axis of the device. When the gas expands in the subsequent relaxation chamber, particles remain closer to the axis due to their much larger inertia. By forcing the flow to pass through multiple lenses, it is possible to confine particles into a gradually narrower beam. This arrangement therefore provides the transfer of particles regardless of their charge state with maximum efficiency. Lighter particles diffuse and escape, whereas the heavy particles (like biomolecules) pick-up momentum and turn into a very fine jet.

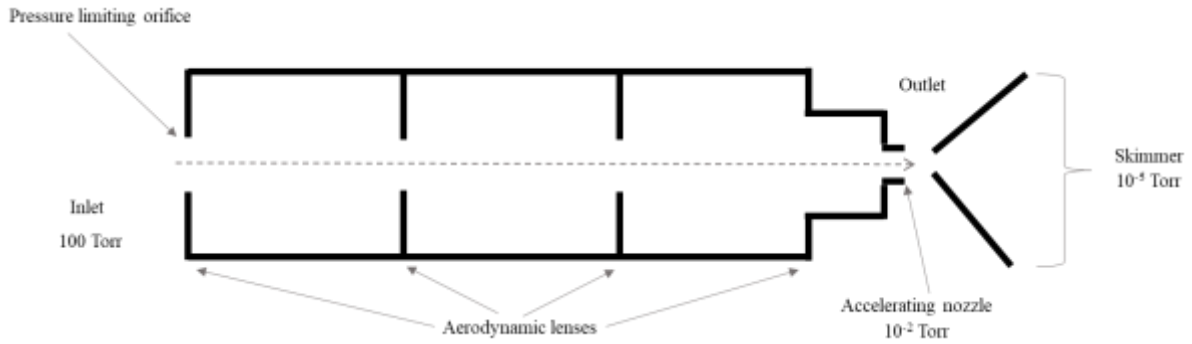


Figure II-6. Sketch of the aerodynamic lens system.

II.3.3 NEMS RESONATOR

A nanoresonator vibrates upon actuation by piezoelectric, thermal, electrostatic or electromagnetic actuators [54] [55] [56]. The frequency readout may be performed using piezoelectric, piezoresistive, capacitive or optical techniques as reported in a review by Boisen et al. on cantilever-like sensors [57]. The nanoresonators used in our NEMS MS architecture are arrays of silicon-based doubly pinned beams, whose vibration at their resonance frequencies upon electrical actuation are detected using piezoelectric gauges [58]. The NEMS-array detector is composed of 20-nanomechanical resonator. The array consists of five rows of four devices covering a total area of 50 µm × 230 µm = 11500 µm^2. Considering the NEMS to be approximately 20 µm long and 300 nm wide, the active sensing surface is 20*3 µm^2 = 60 µm^2 or ~0.5% of the total array surface. NEMS arrays are fabricated from Complementary metal–oxide–semiconductor (CMOS) materials using Very Large Scale Integration (VLSI) processes. NEMS are manufactured on 200 mm Silicon On Insulator (SOI) wafers having a 160 nm thick silicon layer. To achieve distinct resonance frequencies for each of the 20 resonators, they are fabricated with slightly different lengths.

When a particle lands onto the vibrating part of the resonator, its resonance frequency shift downwards. This shift in frequency is proportional to the mass of the particle landed onto the surface. This approach enable the measurement of the mass of individual particles falling onto the surface of the beams in real time and their precise location on the NEMS resonator. This is made possible by real-time tracking of the resonance frequencies of two modes for each resonator [43] [44].

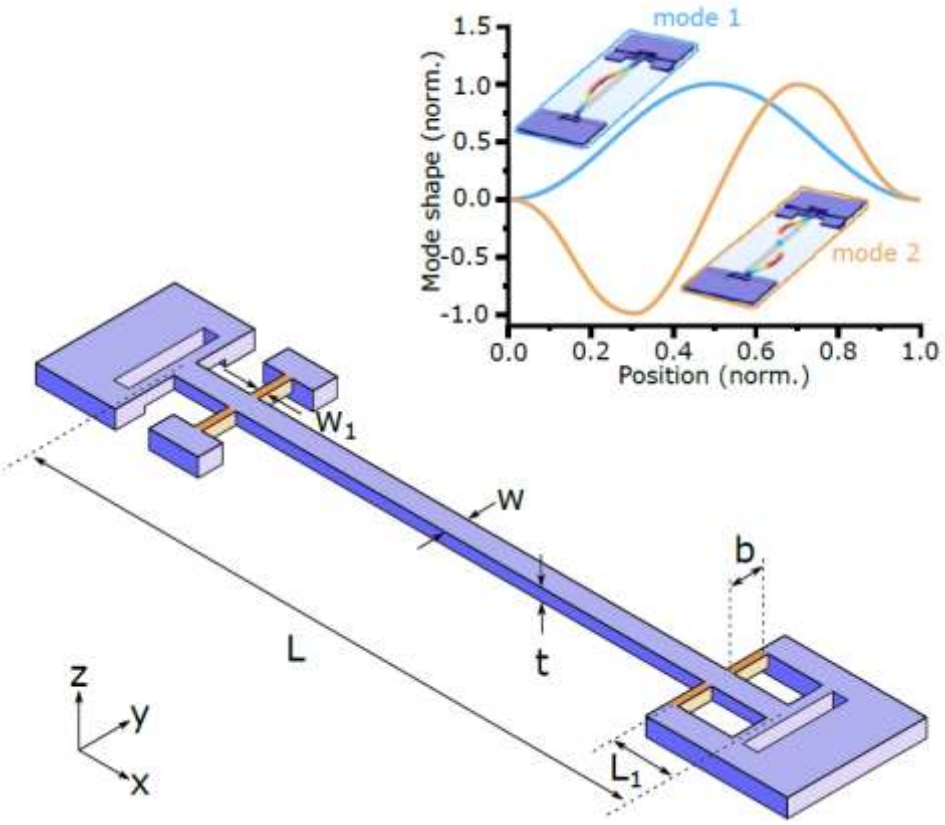


Figure II-7. Nano-resonator device : Elongated doubly clamped beam oscillating laterally (in the x/y plane); two piezoresistive nanogauges (in orange, top) placed in a bridge configuration capture these oscillations. The other end of the beam is not electrically contacted but the mechanical design preserves mode symmetry. Inset displays

normalized mode shapes used to actuate the beam and derive the particle mass and the position on the sensor [59].

The resonators in the array are designed to achieve a resonance frequencies of ~30 MHz and ~80 MHz for mode 1 and mode 2 respectively. The devices geometry and dimensions are summarized in Figure II-7 and in Table II.1. Piezoresistive nanogauges are used to perform in plane transduction motion. The actuation of mode 1 and mode 2 are made possible by drive electrodes (not shown). The inset in Figure II-7 illustrates the beam shape upon maximum displacement under large amplitudes for both the modes. The highlighted zones are the most sensitive areas of the sensor and these regions are distinct for each mode.

Beam length (*L*)	Beam width (*w*)	Beam thickness (t)	Anchor/gauge distance (L_1)	Gauge length (*b*)	Gauge width ($w1$)
7.61-10 µm	300 nm	160 nm	305-400 nm	400 nm	80 nm

Table II.1. Typical dimensions of the nano-resonator devices used in the present study. Ranges pertain to the variable dimensions of individual devices in an array of 20 frequency-addressed resonators [59].

II.3.4 NEMS RESONATOR DEVICE AND MASS MEASUREMENT

To measure the mass of successive particles landing onto the surface of the beam, the resonance frequencies as a function of time are recorded. An example of frequency traces over time obtained in an actual experiment is presented in the Figure II-8. These data were obtained during an analysis of empty bacteriophage T5 capsids. Simultaneous frequency downshifts in both modes correspond to single particle landing events.

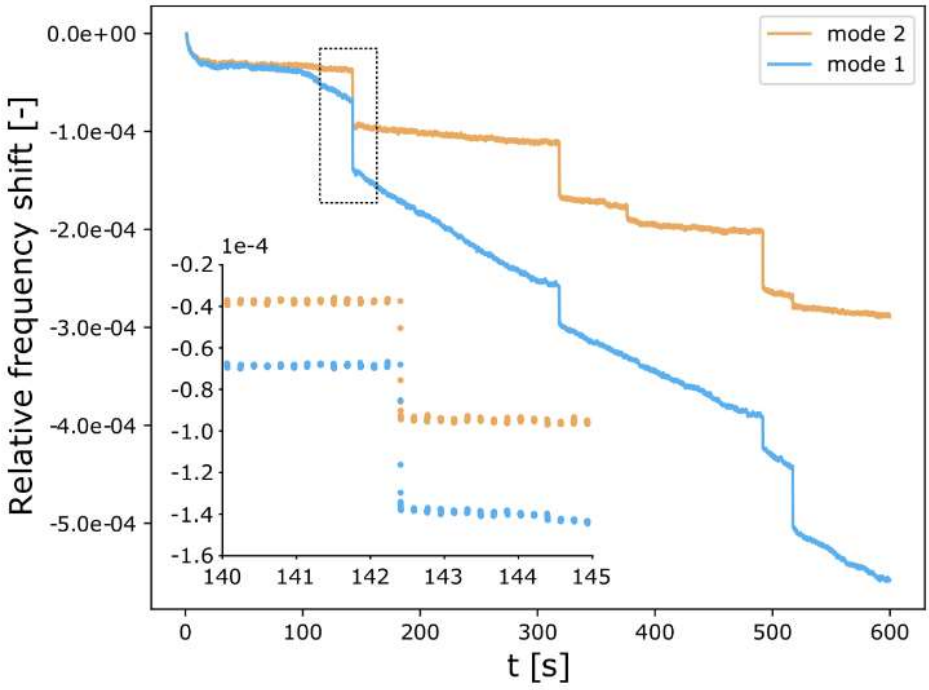


Figure II-8. Frequency trace for both oscillation modes during an actual particle deposition experiment. The inset represents a zoom of the dashed rectangle [59].

Following each particle landing, the resonance frequency suddenly changes as evidence by the frequency discontinuities in Figure II-8. By measuring the relative frequency shifts for a given

particle landing event, it is possible to derive the added mass Δm and its location x using the formula:

$$\Delta m = M \frac{\Delta f_n}{f_n} \frac{\alpha_n}{\phi_n^2(x)} \quad (1)$$

where, M is the total beam mass, $\Delta f_n / f_n$ the relative change in resonance frequency for the nth mode of vibration, $\phi_n(x)$ the mode shape function, and α_n a constant equal to $-2 \int_{x=0}^{x=1} \phi_n^2(x)\, dx$ The process by which the mass and position of the landed particle are derived and the corresponding uncertainties will be discussed in details in Chapter IV.

II.4 NANOPARTICLE TRACKING ANALYSIS

To characterize our samples prior to mass analysis, we needed a reference technique that could provide information on capsid integrity, concentration, as well as aggregation state. Ideally, the method should be easy to perform, require small amounts of sample, and would not alter the sample. Capsids can be viewed as some kind of biological nanoparticles and we thus turned to nanoparticle characterization techniques. Particle sizing using Dynamic Light Scattering (DLS), or imaging with Electron Microscopy (EM) and Atomic Force Microscopy (AFM) are commonly used to derive nanoparticle size, polydispersity and morphology [60] [61]. Each method has specific advantages and disadvantages. Although imaging techniques allow particle visualization at high resolution, they involve specialized skills and are time consuming. Besides, measurements are performed on limited number of particles, which may provide biased information [62]. Finally, it is problematic to retrieve the original sample after analysis to perform additional characterizations.

Light scattering techniques are particularly attractive for our application as they require low amounts of sample in solution and do not degrade the sample. DLS is frequently used for particle sizing and concentration estimation because of its simplicity. DLS determines the size distribution of particles in solution from variations in the scattered light intensity caused by the particle's Brownian motion. The scattered light intensity is proportional to the square of the volume of the particle. Due to this, DLS technique can be very sensitive to larger particles in a sample, and small amounts of large particles can preclude size determination even if the sample is composed mainly of particles with considerably smaller size [63] [64]. Therefore, DLS is ill suited for very poly-dispersed samples, and tends to overestimate particle aggregation when it occurs.

The Nanoparticle Tracking Analyzer (NTA) uses scattered light to visualize and track the Brownian motion of individual nanoparticles in suspension in a liquid, in order to determine their size. NTA can size individual particles from 30 nm to 1 µm. Contrary to DLS, this technique does not depends on the refractive index of nanoparticle [65]. Because this technique deals with individual nanoparticles, it is perfectly suited for poly-dispersed samples. [66] Table II.2 shows a comparison between DLS and NTA from various experimental points of views. Of high interest to us was the ability of NTA to provide information about the sample concentration for each modality in the observed particle size distribution [67].

	DLS	**NTA**
Size range	1-1000 nm	30-1000 nm
Concentration range	Wide : 10^8 - 10^{12} particle/ml	Narrow: 10^7 - 10^9 particle/ml
Peak resolution	Low (Log scale)	High (linear scale)
Particle visualization	No	Yes
Duration	~2-5 min/measurement	~5-10 min/measurement
Reproducibility	High	May depend on operator's skills and experimental parameters.
Sensitivity to polydispersity	Better suited for monodispersed sample	Precise for both mono and poly-dispersed samples
Sensitivity to contamination	Highly sensitive to low number of larger particles (dirt, aggregates)	Aggregates or dust particles are easily visualized. Moderate amounts have minor impact on the results.
Device handling	Calibration required User friendly Disposable cuvettes limit risk of cross-contamination	No calibration required Requires optimization by a skilled operator After each measurement, the flow cell must be cleaned. Risk of biofouling.
Information	Size distribution Intensity weighted size distribution Concentration Polydispersity index	Individual particle sizes Number weighted size distribution Concentration

Table II.2. Comparative characteristics of DLS and NTA

II.4.1 PRINCIPLE AND METHODOLOGY

NTA is a recent technique that allows nanoparticles analysis in real time (Figure II-9) [68] [69]. It relies on a laser illuminated video-microscopy technique in which the movement of nanoparticles under Brownian motion is recorded. NTA tracks the position of nanoparticles in the field of view using a computer program that detects light scattered from the nanoparticle in suspension as they are illuminated by a laser source. Figure II-9 A shows the Malvern NTA LM300 system used in our laboratory. It features a 20x magnification microscope equipped with a charge coupled device (CCD) camera, a sample chamber and laser.

In this instrument, a 405 nm laser beam is focused onto an optical prism and refracted into a liquid suspension of the nanoparticles of interest (Figure II-9 C). Refraction causes a strong illumination area in the liquid medium, and the light scattered by particles renders them easily visible using a 20X magnification microscope (Figure II-9 B). A CCD camera, functioning at 30 frames/second is used to record the field of view as particles move under Brownian motion. The camera's field of view is approximately 100 µm x 80 µm x 10 µm. A dedicated software is used to detect and track the moving particles over time within the camera's field of view (Figure II-9 D). The average distance travelled by each particle along the x and y-axis is computed over a large multiplicity of frames. Using this information, a diffusion coefficient can be derived for each particle. By knowing the samples temperature and viscosity it is then possible to determine the hydrodynamic diameter of each individual particle, corresponding to the diameter of the equivalent hard sphere that would diffuse at the same average speed as the particle under study. The hydrodynamic radius can be determined by using the relationship

between distance travelled by the nanoparticle during a given time and its hydrodynamic diameter, derived from the Stokes Einstein Equation, is as follows:

$$\overline{(x,y)}^2 = D = \frac{4TK_B t}{3d_h \pi \eta} \qquad (2)$$

where $\overline{(x,y)}^2$ is the particles mean squared displacement in the focal plane, T the temperature in K°, KB the Boltzmann's constant = 1.380649 × 10−23 (J/K), t the time period during which the displacement was observed, dh the hydrodynamic diameter and η the solvent's dynamic viscosity.

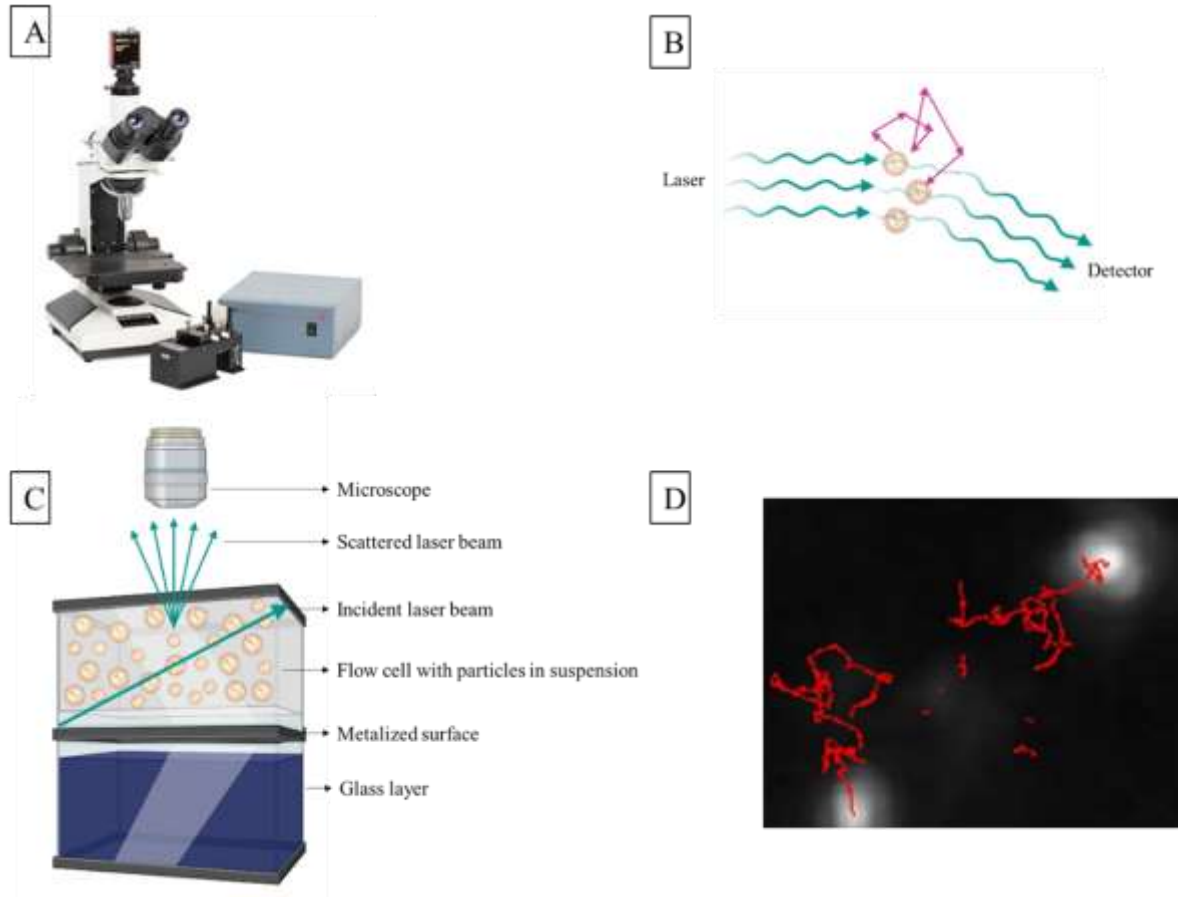


Figure II-9. (A) Nanoparticle Tracking Analysis-Nanosight NS 300 system used in this study, showing the 20x magnification microscope, the proprietary flow-cell compartment, and the 405 nm laser housing. This system is coupled with a desktop computer (not shown) to control the system, record and process videos. (B) Brownian motion of particles give rise to intensity fluctuation in the scattered light. In DLS, this fluctuations are analyzed using a detector to yield particle the size distribution. In NTA, the scattered light is merely used to visualize the particles using a light microscope (C) Schematic of the optical configuration used in the NTA flow cell. (D) Computer traces corresponding to the Brownian motion of individual particles within the field of view.

NTA records many consecutive frames to track particles within the field of view, thereby eliminating the chance of counting the same particle repeatedly. Data obtained from the acquired frames are taken into account to evaluate the mean displacement for each detected particle. Aggregating this information, one can compute the particle size distribution and the average concentration of the sample for each modality. NTA tracks particles over a timescale of 30 sec to 1 minute. NTA experiments require a minimum of 10^7 particles/ml and a maximum

of 10^9 particles/ml. Some key experimental factors to perform measurements by NTA are provided in the Table II.3. This underscores the requirement of trained operators to perform the measurements.

Parameter	Description	Degraded settings	Optimized settings
Shutter speed	Length of time when the camera shutter is open. This controls the sensitivity of the camera.	Too short shutter speed obscures small particles and too long causes blurring.	Nanoparticles look like clear moving dots, tracable by the software.
Capture duration	Length of the captured video.	Too long durations may cause multiple detections of the same particle; and too short durations increase the error in size.	Reproducible size distribution
Blur	Smoothening of the obtained video due to the motion (visual noise) of the particle.	Too much blur makes some particles disappear and too little causes false light scattering centers.	Eliminates the fake light scattering centers and allow the software to track genuine particles.
Detection threshold	Determines the ability to track a particle	Too high threshold causes loss of small particles and too low causes loss of larger particles.	All nanoparticles in the obtained video are tracked.
Minimum expected particle size (MEPS)	Determines the field around the particle where the software searches for it in the following frame.	High MEPS results in loss of particles and low MEPS results in particles overlap.	Nearly all particles in the obtained video are accounted for.
Laser beam position	Appropriate area to view the illuminated particles	Wrong position causes streaking in the image 	No streaks. Particles are clearly visible 
Centering the image	Adjusting the image center	Shifted field of view 	Well centered field of view 
Image focus	Focus of the microscope	Out of focus 	In focus 
Camera level adjustments	Camera settings should be well adjusted prior capturing of the video	Improper setting 	Optimal setting 
Concentration	Optimal particle concentration is 10^7 -10^9particles/ml (20-100 particles in the viewing area)	Too high concentration 	Acceptable concentration 

Table II.3. Important NTA experimental parameters that can affect the measurement. Settings are illustrated to achieve optimal performance

In practice, the first step in a measurement consists in proper alignment of the field of view of the microscope and its focal plane with respect to the area of the sample flow cell highlighted by the laser. The flow cell is then cleaned by introducing deionized water using a syringe connected to a Luer type fitting. The blank is then checked for absence of particles. If particles are visible, it means that either the flow cell is contaminated, or the water is not pure. Cleaning is then repeated until no particles are visible. Next, the sample, properly diluted to obtain a concentration range within specifications, is introduced into the flow cell using the syringe. It is very important to avoid introducing air bubbles in the flow cell at this stage, as they will scatter light and cause artifacts. Following sample introduction, the image is optimized iteratively by tuning the beam position, the focus, and the camera sensitivity level. At this stage, the sample can be further diluted or concentrated if required to obtain between 20 and 100

particles within the field of view. After stabilization of the temperature, and visual assessment of the particles within the field of view, a user specified number of videos are recorded following consecutive introduction of small amounts of sample solution into the flow cell. Each video is 30 s to 1 minute long depending on user setting. Typical number of repetitions are 3-10. Following automated image processing, the system outputs a .pdf file displaying the overlaid size distribution plot for each repetition as well as data on the settings used along with a results summary (see *Figure* II-10). It is also possible to obtain a CSV file to capture all details of the experiment.

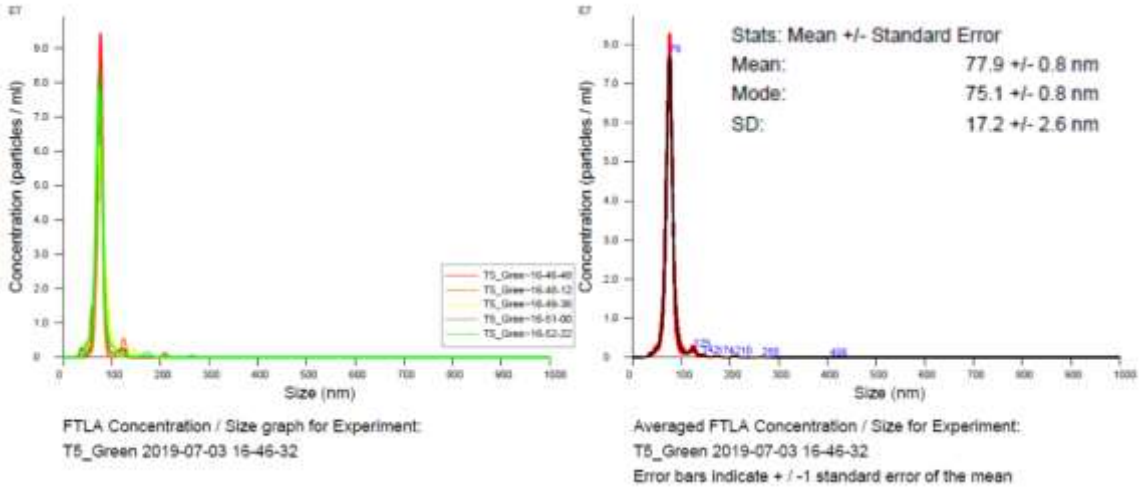


Figure II-10. Summary output from NTA measurement of bacteriophage T5 capsids with 5 replications overlaid (left), or averaged (right). The excellent agreement between replicate measurements supports the accuracy of this result.

II.5 VIRAL VACCINES AND VIRUS LIKE PARTICLES (VLP) ANALYSIS

Measuring the size and counting the titer of viral particles and their aggregation state plays a crucial role in developing vaccines and for phage therapy. NTA is able to provide such information in few minutes. Viruses ranging from 15 – 300 nm are amenable to NTA measurement. Studying virus batches using NTA is ideal because this system is able to deliver information about the concentration and size distribution on a particle-by-particle basis. This information is important because estimating the concentration of the sample and degree of agglomeration is of major factor in the production of vaccine. Since this method does not require infectivity or nucleic acid to perform a measurement, it is ideal for measuring the titer and quality of VLP structures [70]. In the present work, we relied on NTA to compare sample preparation methods by assessing particle integrity, aggregation status and potential sample losses prior to mass measurement using nano-resonators Mass Spectrometry (See Chapter IV).

II.6 CONCLUSION

In this chapter, we have presented two MS based techniques: Liquid Chromatography-Tandem Mass Spectrometry and Nano Electro Mechanical Sensor Based Mass Spectrometry, as an analytical tool to characterize bacteriophage T5 capsid particles. We have also mentioned a reference technique: Nanoparticle Tracking Analysis, which was used to access additional information on the structural integrity, size and concentration of the capsid sample. In the next two chapters, we will be presenting our results and conclusions based on these techniques.

REFERENCES

1. (2007) XL. Cathode Rays - by J.J. Thomson. Philosophical Magazine Letters 87:303–326 . https://doi.org/10.1080/09500830701306165

2. Strutt RJ (1941) Joseph John Thomson, 1856 - 1940. Obituary Notices of Fellows of the Royal Society 3:587–609 . https://doi.org/10.1098/rsbm.1941.0024

3. Squires G (1998) Francis Aston and the mass spectrograph. J Chem Soc, Dalton Trans 3893–3900 . https://doi.org/10.1039/a804629h

4. Berson JA (2007) Shifting and Rearranging. Physical Methods and the Transformation of Modern Chemistry. By Carsten Reinhardt. Angewandte Chemie International Edition 46:4818–4819 . https://doi.org/10.1002/anie.200785463

5. Posthumus MA, Kistemaker PG, Meuzelaar HLC, Ten Noever de Brauw MC (1978) Laser desorption-mass spectrometry of polar nonvolatile bio-organic molecules. Anal Chem 50:985–991 . https://doi.org/10.1021/ac50029a040

6. Fenn J (2002) Electrospray ionization mass spectrometry: How it all began. J Biomol Tech 13:101–118

7. Fenn JB (2003) Electrospray Wings for Molecular Elephants (Nobel Lecture). Angewandte Chemie International Edition 42:3871–3894 . https://doi.org/10.1002/anie.200300605

8. Kim J (2015) Sample Preparation for Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. Mass Spectrometry Letters 6:27–30 . https://doi.org/10.5478/MSL.2015.6.2.27

9. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1989) Electrospray ionization for mass spectrometry of large biomolecules. Science 246:64–71 . https://doi.org/10.1126/science.2675315

10. Heberer T (2002) Occurrence, fate, and removal of pharmaceutical residues in the aquatic environment: a review of recent research data. Toxicol Lett 131:5–17 . https://doi.org/10.1016/s0378-4274(02)00041-3

11. Robinson CV, Sali A, Baumeister W (2007) The molecular sociology of the cell. Nature 450:973–982 . https://doi.org/10.1038/nature06523

12. Palmer PT, Limero TF (2001) Mass spectrometry in the U.S. space program: past, present, and future. Journal of the American Society for Mass Spectrometry 12:656–675 . https://doi.org/10.1016/S1044-0305(01)00249-5

13. Wörner TP, Shamorkina TM, Snijder J, Heck AJR (2021) Mass Spectrometry-Based Structural Virology. Anal Chem 93:620–640 . https://doi.org/10.1021/acs.analchem.0c04339

14. Murayama C, Kimura Y, Setou M (2009) Imaging mass spectrometry: principle and application. Biophys Rev 1:131 . https://doi.org/10.1007/s12551-009-0015-6

15. Miller MB, Tang Y-W (2009) Basic concepts of microarrays and potential applications in clinical microbiology. Clin Microbiol Rev 22:611–633 . https://doi.org/10.1128/CMR.00019-09

16. Karpievitch YV, Polpitiya AD, Anderson GA, Smith RD, Dabney AR (2010) Liquid Chromatography Mass Spectrometry-Based Proteomics: Biological and Technological Aspects. Ann Appl Stat 4:1797–1823 . https://doi.org/10.1214/10-AOAS341

17. Nesvizhskii AI, Vitek O, Aebersold R (2007) Analysis and validation of proteomic data generated by tandem mass spectrometry. Nat Methods 4:787–797 . https://doi.org/10.1038/nmeth1088

18. Tang N, Tornatore P, Weinberger SR (2004) Current developments in SELDI affinity technology. Mass Spectrom Rev 23:34–44 . https://doi.org/10.1002/mas.10066

19. Top-Down versus Bottom-Up Approaches in Proteomics. In: Chromatography Online. https://www.chromatographyonline.com/view/top-down-versus-bottom-approaches-proteomics-0

20. Han X, Aslanian A, Yates JR (2008) Mass spectrometry for proteomics. Curr Opin Chem Biol 12:483–490 . https://doi.org/10.1016/j.cbpa.2008.07.024

21. Mascot overview | Protein identification software for mass spec data. https://proteomicsresource.washington.edu/mascot/search_intro.html

22. Liu H, Sadygov RG, Yates JR (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem 76:4193–4201 . https://doi.org/10.1021/ac0498563

23. Wang W, Zhou H, Lin H, Roy S, Shaler TA, Hill LR, Norton S, Kumar P, Anderle M, Becker CH (2003) Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. Anal Chem 75:4818–4826 . https://doi.org/10.1021/ac026468x

24. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17:994–999 . https://doi.org/10.1038/13690

25. Lundgren DH, Hwang S-I, Wu L, Han DK (2010) Role of spectral counting in quantitative proteomics. Expert Rev Proteomics 7:39–53 . https://doi.org/10.1586/epr.09.69

26. Silva JC, Gorenstein MV, Li G-Z, Vissers JPC, Geromanos SJ (2006) Absolute Quantification of Proteins by LCMSE: A Virtue of Parallel ms Acquisition. Molecular & Cellular Proteomics 5:144–156 . https://doi.org/10.1074/mcp.M500230-MCP200

27. Oda Y, Huang K, Cross FR, Cowburn D, Chait BT (1999) Accurate quantitation of protein expression and site-specific phosphorylation. Proceedings of the National Academy of Sciences 96:6591–6596 . https://doi.org/10.1073/pnas.96.12.6591

28. Engvall E, Perlmann P (1972) Enzyme-Linked Immunosorbent Assay, Elisa. J Immunol 109:129

29. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci U S A 100:6940–6945 . https://doi.org/10.1073/pnas.0832254100

30. Brun V, Dupuis A, Adrait A, Marcellin M, Thomas D, Court M, Vandenesch F, Garin J (2007) Isotope-labeled Protein Standards: Toward Absolute Quantitative Proteomics. Molecular & Cellular Proteomics 6:2139–2149 . https://doi.org/10.1074/mcp.M700163-MCP200

31. Beynon RJ, Doherty MK, Pratt JM, Gaskell SJ (2005) Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. Nat Methods 2:587–589 . https://doi.org/10.1038/nmeth774

32. Chang H-C (2009) Ultrahigh-mass mass spectrometry of single biomolecules and bioparticles. Annu Rev Anal Chem (Palo Alto Calif) 2:169–185 . https://doi.org/10.1146/annurev-anchem-060908-155245

33. Rose RJ, Damoc E, Denisov E, Makarov A, Heck AJR (2012) High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. Nat Methods 9:1084–1086 . https://doi.org/10.1038/nmeth.2208

34. Keifer DZ, Pierson EE, Jarrold MF (2017) Charge detection mass spectrometry: weighing heavier things. Analyst 142:1654–1671 . https://doi.org/10.1039/C7AN00277G

35. Weiss VU, Pogan R, Zoratto S, Bond KM, Boulanger P, Jarrold MF, Lyktey N, Pahl D, Puffler N, Schelhaas M, Selivanovitch E, Uetrecht C, Allmaier G (2019) Virus-like particle size and molecular weight/mass determination applying gas-phase electrophoresis (native nES GEMMA). Anal Bioanal Chem 411:5951–5962 . https://doi.org/10.1007/s00216-019-01998-6

36. Thundat T, Wachter EA, Sharp SL, Warmack RJ (1995) Detection of mercury vapor using resonating microcantilevers. Appl Phys Lett 66:1695–1697 . https://doi.org/10.1063/1.113896

37. Ilic B, Yang Y, Aubin K, Reichenbach R, Krylov S, Craighead HG (2005) Enumeration of DNA molecules bound to a nanomechanical oscillator. Nano Lett 5:925–929 . https://doi.org/10.1021/nl050456k

38. Gil-Santos E, Ramos D, Martínez J, Fernández-Regúlez M, García R, San Paulo Á, Calleja M, Tamayo J (2010) Nanomechanical mass sensing and stiffness spectrometry based on two-dimensional vibrations of resonant nanowires. Nature Nanotech 5:641–645 . https://doi.org/10.1038/nnano.2010.151

39. Chen GY, Thundat T, Wachter EA, Warmack RJ (1995) Adsorption-induced surface stress and its effects on resonance frequency of microcantilevers. Journal of Applied Physics 77:3618–3622 . https://doi.org/10.1063/1.359562

40. Naik AK, Hanay MS, Hiebert WK, Feng XL, Roukes ML (2009) Towards single-molecule nanomechanical mass spectrometry. Nat Nanotechnol 4:445–450 . https://doi.org/10.1038/nnano.2009.152

41. Hanay MS, Kelber S, Naik AK, Chi D, Hentz S, Bullard EC, Colinet E, Duraffourg L, Roukes ML (2012) Single-protein nanomechanical mass spectrometry in real time. Nature Nanotech 7:602–608 . https://doi.org/10.1038/nnano.2012.119

42. Hentz S, Masselon C (2016) Device for determining the mass of a particle in suspension or in solution in a fluid. US patent #9506852B2

43. Sage E, Brenac A, Alava T, Morel R, Dupré C, Hanay MS, Roukes ML, Duraffourg L, Masselon C, Hentz S (2015) Neutral particle mass spectrometry with nanomechanical systems. Nat Commun 6:6482 . https://doi.org/10.1038/ncomms7482

44. E S, M S, S F, M D, M G, Ak N, R M, L D, Ml R, T A, G J, E C, C M, A B, S H (2018) Single-particle mass spectrometry with arrays of frequency-addressed nanomechanical resonators. Nat Commun 9:3283–3283 . https://doi.org/10.1038/s41467-018-05783-4

45. Neutral mass spectrometry of virus capsids above 100 megadaltons with nanomechanical resonators | Science. https://science.sciencemag.org/content/362/6417/918?rss=1. Accessed 11 Dec 2020

46. Schultz JC, Hack CA, Benner WH (1998) Mass determination of megadalton-DNA electrospray ions using charge detection mass spectrometry. J Am Soc Mass Spectrom 9:305–313 . https://doi.org/10.1016/S1044-0305(97)00290-0

47. Boeri Erba E, Petosa C (2015) The emerging role of native mass spectrometry in characterizing the structure and dynamics of macromolecular complexes. Protein Sci 24:1176–1192 . https://doi.org/10.1002/pro.2661

48. Unraveling the Mechanism of Electrospray Ionization | Analytical Chemistry. https://pubs.acs.org/doi/10.1021/ac302789c. Accessed 25 Jul 2021

49. Liu P, Ziemann PJ, Kittelson DB, McMurry PH (1995) Generating Particle Beams of Controlled Dimensions and Divergence: I. Theory of Particle Motion in Aerodynamic Lenses and Nozzle Expansions. Aerosol Science and Technology 22:293–313 . https://doi.org/10.1080/02786829408959748

50. Liu P, Ziemann PJ, Kittelson DB, McMurry PH (1995) Generating Particle Beams of Controlled Dimensions and Divergence: II. Experimental Evaluation of Particle Motion in Aerodynamic Lenses and Nozzle Expansions. Aerosol Science and Technology 22:314–324 . https://doi.org/10.1080/02786829408959749

51. Lee K-S, Cho S-W, Lee D (2008) Development and experimental evaluation of aerodynamic lens as an aerosol inlet of single particle mass spectrometry. Journal of Aerosol Science 39:287–304 . https://doi.org/10.1016/j.jaerosci.2007.10.011

52. Di Fonzo F, Gidwani A, Fan MH, Neumann D, Iordanoglou DI, Heberlein JVR, McMurry PH, Girshick SL, Tymiak N, Gerberich WW, Rao NP (2000) Focused nanoparticle-beam deposition of patterned microstructures. Appl Phys Lett 77:910–912 . https://doi.org/10.1063/1.1306638

53. Hentz S, Masselon C (2014) Device for determining the mass of a particle in suspension or in solution in a fluid

54. Black JP, Elium A, White RM, Apte MG, Gundel LA, Cambie R (2007) 6D-2 MEMS-Enabled Miniaturized Particulate Matter Monitor Employing 1.6 GHz Aluminum Nitride Thin-Film Bulk Acoustic Wave Resonator (FBAR) and Thermophoretic Precipitator. In: 2007 IEEE Ultrasonics Symposium Proceedings. pp 476–479

55. Hajjam A, Rahafrooz A, Wilson JC, Pourkamali S (2009) Thermally actuated MEMS resonant sensors for mass measurement of micro/nanoscale aerosol particles. In: 2009 IEEE SENSORS. pp 707–710

56. Schmid S, Kurek M, Boisen A (2013) Towards airborne nanoparticle mass spectrometry with nanomechanical string resonators. In: Micro- and Nanotechnology Sensors, Systems, and Applications V. International Society for Optics and Photonics, p 872525

57. Boisen A, Dohn S, Keller SS, Schmid S, Tenje M (2011) Cantilever-like micromechanical sensors. Rep Prog Phys 74:036101 . https://doi.org/10.1088/0034-4885/74/3/036101

58. Mile E, Jourdan G, Bargatin I, Labarthe S, Marcoux C, Andreucci P, Hentz S, Kharrat C, Colinet E, Duraffourg L (2010) In-plane nanoelectromechanical resonators based on silicon nanowire piezoresistive detection. Nanotechnology 21:165504 . https://doi.org/10.1088/0957-4484/21/16/165504

59. Clement K, Reynaud A, Defoort M, Vysotskyi B, Fortin T, Lai S-H, Çumaku V, Dominguez-Medina S, Hentz S, Masselon C (2021) Requirements and attributes of nano-resonator mass spectrometry for the analysis of intact viral particles. Analytical and Bioanalytical Chemistry. https://doi.org/10.1007/s00216-021-03511-4

60. Sokolova V, Ludwig A-K, Hornung S, Rotan O, Horn PA, Epple M, Giebel B (2011) Characterisation of exosomes derived from human cells by nanoparticle tracking analysis and scanning electron microscopy. Colloids and Surfaces B: Biointerfaces 87:146–150 . https://doi.org/10.1016/j.colsurfb.2011.05.013

61. Anderson W, Kozak D, Coleman VA, Jämting ÅK, Trau M (2013) A comparative study of submicron particle sizing platforms: Accuracy, precision and resolution analysis of polydisperse particle size distributions. Journal of Colloid and Interface Science 405:322–330 . https://doi.org/10.1016/j.jcis.2013.02.030

62. Merkus HG (2009) Particle Size, Size Distributions and Shape. In: Merkus HG (ed) Particle Size Measurements: Fundamentals, Practice, Quality. Springer Netherlands, Dordrecht, pp 13–42

63. Carpenter DK (1977) Dynamic Light Scattering with Applications to Chemistry, Biology, and Physics (Berne, Bruce J.; Pecora, Robert). J Chem Educ 54:A430 . https://doi.org/10.1021/ed054pA430.1

64. Maulucci G, Spirito MD, Arcovito G, Boffi F, Castellano AC, Briganti G (2005) Particle Size Distribution in DMPC Vesicles Solutions Undergoing Different Sonication Times. Biophysical Journal 88:3545–3550 . https://doi.org/10.1529/biophysj.104.048876

65. Malloy A (2011) Count, size and visualize nanoparticles. Materials Today 14:170–173 . https://doi.org/10.1016/S1369-7021(11)70089-X

66. Filipe V, Hawe A, Jiskoot W (2010) Critical evaluation of Nanoparticle Tracking Analysis (NTA) by NanoSight for the measurement of nanoparticles and protein aggregates. Pharm Res 27:796–810 . https://doi.org/10.1007/s11095-010-0073-2

67. Carr B, Hole P, Malloy A, Nelson P, Wright M, Smith J (2009) Applications of nanoparticle tracking analysis in nanoparticle research - A mini review. Eur J Parenter Pharm Sci 14:45–50

68. Review of Nanoparticle Tracking Analysis.pdf

69. Malloy A (2011) Count, size and visualize nanoparticles. Materials Today 14:170–173 . https://doi.org/10.1016/S1369-7021(11)70089-X

70. WP140912VirusVaccines.pdf

CHAPTER III: **PROTEOMICS STUDY OF T5 PROTEASE**

In this chapter after a brief introduction on Bacteriophage T5 capsid, we discuss about the importance of the capsid protein: Protease. To have a better understanding on this protein we investigated the proteins present in T5 capsid using a global proteomic approach, which showed us the evidence of protease in the mature capsid. Based on these experiments we then designed a novel targeted proteomic method using Quantitative Concatamer, which determined the stoichiometry of protease protein in the mature T5.

III.1 **INTRODUCTION**

Bacteriophage T5, also called Escherichia virus T5, from the order of Caudovirales, is a lytic phage infecting E. coli. It has a long non-contractile tail, which classifies it in the Siphoviridae family [*Figure III-1*]. T5 has an icosahedral capsid of about 900 Å in diameter containing its genome, a large double stranded 121 kbp DNA genome, compacted into an almost crystalline structure. Its capsid has a triangulation of T = 13, and is composed of :

- 775 copies of a Major Capsid Protein [pb8] forming 11 pentamers and 120 hexamers. pb8 has two domains. Only the C-terminal mature domain (pb8m) forms the icosahedral shell, while the N-terminal scaffolding domain (Δ-domain) has been processed by the head maturation protease.
- An oligomer of 12 copies of the Portal Protein [pb7], assembled into a ring on a single vertex [12th vertex], and forming an entry and exit channel for the viral DNA
- Multiple copies, whose exact number remains still unknown, of a the head Protease [pb11], performing proteolytic cleavage and thereby regulating capsid maturation [1]
- 120 copies of a Decoration Protein [pb10], attached to the center of each hexamer and whose function is unknown
- The Head Completion Protein [p144], forming a plug that closes the portal gate and connecting the mature capsid the tail terminator protein p142. The copy number of p144 and of p142 in the phage particle has not been determined, however, by similarity with other phages of known structure, we can expect this number to be 12.

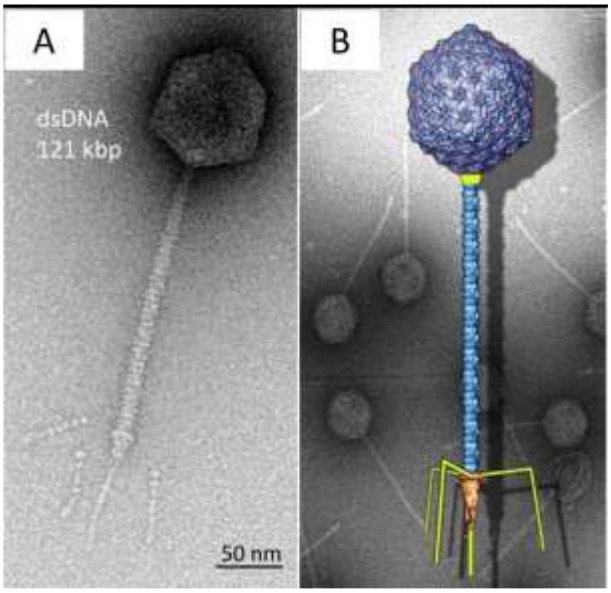


Figure III-1. Bacteriophage T5. [A] Negative stain electron microscopy image. [B] First reconstruction of the viral particle from cryo electron microscopy. The head and tail are at 19 Å and 30 Å of resolution, respectively [2] [2]. pb8 is shown in blue and pb10 in red. The connector as well as the tail fibers are shown in yellow. Background: Negative staining electron microscopy image of several T5 phages. Reprinted from [2]

The high resolution structure of the T5 capsid was recently solved by cryo-electro microscopy and reconstruction image with a resolution of 20 Å (*Figure III-2*).

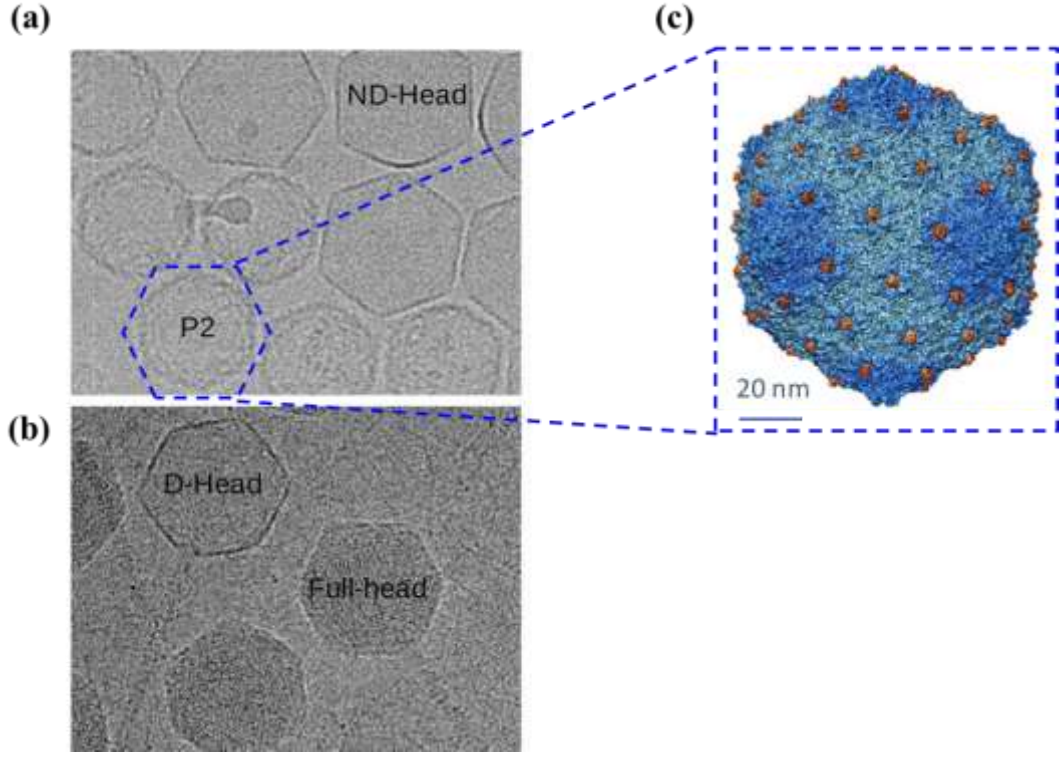


Figure III-2. The cryo-EM and image reconstruction of T5 capsid. (a) P=Prohead II and ND-Head= Empty capsid without decoration, (b) D-Head= Decorated empty capsid and (c) reconstruction image of T5 capsid. Reprinted from [3]

The major capsid protein pb8 adopts a fold similar to that of the first resolved capsid protein structure, the gp5 protein of phage HK97 [4]. The HK97-like fold is very prevalent for icosahedral capsid proteins. A model of the structure of pb8 based on the structure of gp5 is shown in Figure III-3[5]. Domain A (for "axial") is located at the center of capsomers and the P domain (for "peripheral") allows contacts between capsomers.

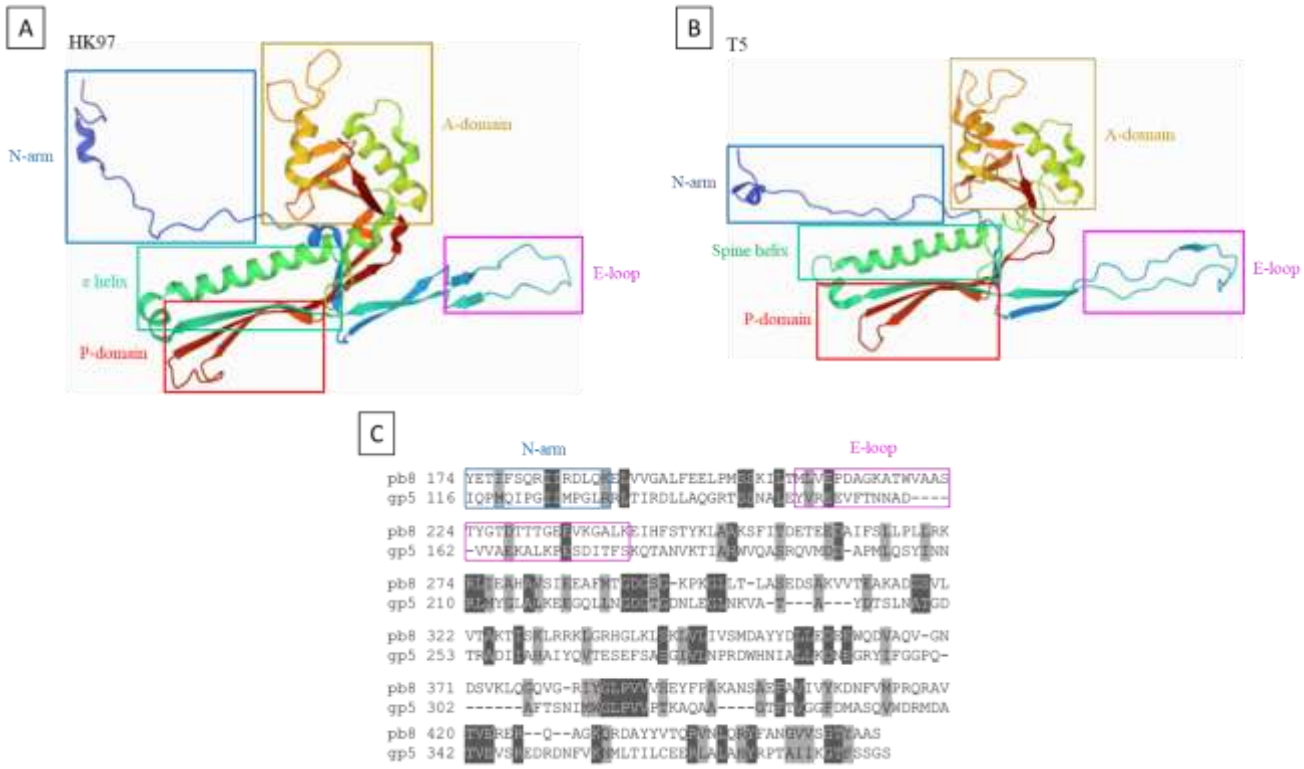


Figure III-3. pb8 of T5 has a HK97 type fold. [A] Structure of the gp5 capsid protein of phage HK97 [6]. [B] Structural model of the pb8 protein of T5 obtained by homology with the gp5 protein of HK97 and generated from the PHYRE server [7]. [C] Sequence alignment pb8 and gp5 proteins [8].

III.2 **HEAD ASSEMBLY**

The T5 head assembly process is similar to that of many other phages and to the eukaryotic herpes viruses. It consists in the assembly of an empty compact procapsid, which is then filled with viral DNA, with concomitant capsid expansion [6]. The T5 head assembly process was described in details by Huet et al. [7]and is illustrated in Figure III-4. The first step is the assembly of a compact structure called a procapsid, or "prohead I". This assembly involves three proteins: the portal protein pb7 [shown in blue in Figure III-4], the major capsid protein in its precursor form pb8p [purple] and the maturation protease pb11 [light green]. The pb8p protein self-assembles thanks to its scaffolding domain, also called "Δ-domain" [symbolized by a gray triangle]. The maturation of Prohead I results from on a mutltiple proteolyic processing by the protease pb11, which cleaves itself at both N and C-termini, cleaves the first 10 amino acids of pb7 and the scaffolding domain of pb8p, freeing interior space. This empty procapsid is called "prohead II". The exact copy number of protease involved in this process is still undetermined, but a few copies of pb11 cleaved at both ends are known to remain associated with Prohead II [8]. DNA is packaged in prohead II by an enzyme called terminase through the channel formed by the portal protein. This packaging process is commonly seen in tailed phages and herpesviruses. The packaging machinery is made up of two subunits of the terminase TerS and TerL, which assemble into oligomers at the portal channel located at one of the vertices of

the procapsids. The small TerS subunit forms a ring of 8 to 12 copies that recognizes phage DNA replicated as concatamers [8]. The large TerL subunit is pentameric and has dual ATPase and DNAse activity. It allows the translocation of DNA in the capsid, thanks to the hydrolysis energy of ATP. In the case of T5, when a genome unit has been packaged, TerL cuts DNA at a specific site to separate it from the concatamer, thanks to its endonuclease activity [9]. Thus, all the produced virions contain exactly the same DNA sequence. The capsid expands during packaging, under the pressure exerted by compacted DNA and/or electrostatic interactions between DNA and the pb8 protein. In this process, all major capsid protein subunits simultaneously undergo irreversible conformational rearrangement effectively doubling the volume of the capsid which can then accommodate the entire genome. The average diameter of the capsid increases from 70 nm (prohead II) to 90 nm (expanded capsid) [10].

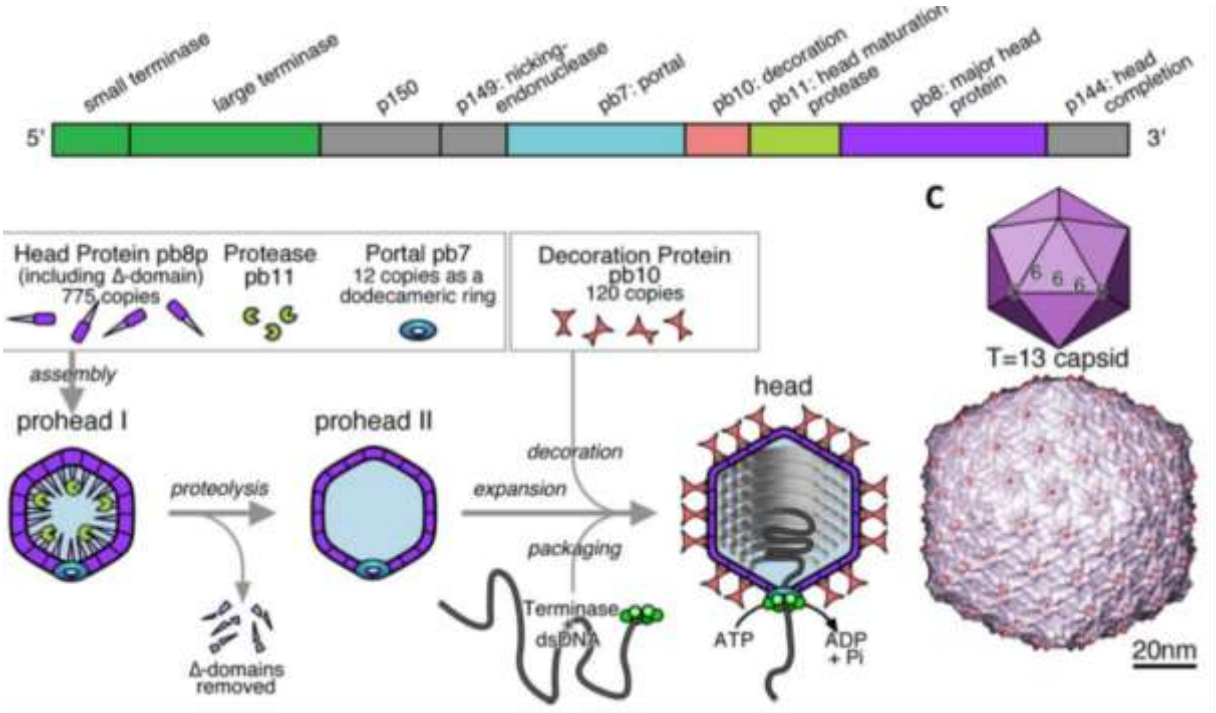


Figure III-4. Bacteriophage T5 head assembly. [A] Schematic of the genes encoding the head assembly. [B] Schematic of the head assembly process. [C] At the top, a diagram of a capsid having a triangulation number T = 13 on which the hexamers are symbolized by the number 6 and the pentamers by the number 5. Below, a model of the T5 capsid reconstructed from cryo-microscopy data electronic. Reprinted from [7]

III.3 INVOLVEMENT OF THE HEAD PROTEASE

Knowing the above-mentioned role of pb11, it is necessary for the capsid to incorporate a certain amount of pb11 for the capsid assembly process to take place. Even though pb11 is essential for T5 capsid assembly and maturation, the exact amount of this protein present at each stage of capsid assembly remains unknown. A previous study on the regulation of T5 capsid assembly pointed out that the stoichiometry of pb11 in prohead I might be as high as 200 copies [11]. In another article published in 2014, Zivanovic et al. confirmed the presence of pb11 in prohead II and in mature T5 capsids [8]. They estimated the existence of a low number (9±3) of pb11 in mature capsid by densitometry analysis of T5 structural protein resolved by SDS-PAGE.

Mass spectrometry (MS) based shotgun proteomic approaches allow the quantification of proteins by measuring the relative abundance of their tryptic peptides following data dependent acquisition, through MS/MS spectra counting or from extracted ion chromatography (XIC) [12] [13]. Moreover, when shotgun proteomic join hands with isotopically labelled strategies such as stable isotope-labeled amino acids (SILAC), isotope-coded affinity tag (ICAT), or isobaric tags for relative absolute quantitation, it is possible to investigate the absolute quantity of proteins in biological samples [14][15]. When it comes to targeted proteomic quantification, AQUA and QconCAT based isotopic labelling methods are ideal as internal standards that can be introduced into the sample before or after proteolysis [16] [17]. Based on the amount of these standards that is added to the sample, one could determine the intensity signal ratio between the native peptide and the corresponding standard, and calculate the amount of the native peptide.

In this chapter, we will describe a global proteomic approach using nanoscale liquid chromatography mass spectrometry (nano LC-MS/MS) to assess proteins present in bacteriophage T5 capsids, and a novel targeted proteomic assay for the quantification of pb11 in mature T5 capsids using a heavy isotopically labelled QconCAT AQUA dipeptide standard.

III.4 SHOTGUN PROTEOMIC ANALYSIS OF T5 CAPSIDS

III.4.1 CAPSID SAMPLES

The bacteriophage T5 capsids used in this present work were produced at the laboratory of Dr. Pascale Boulanger (Institute for Integrative Biology of the Cell/ Gif- sur- Yvette/Paris/France). Three types of capsids were used in this experiment: 1. DNA filled capsids having decoration protein at its exterior (DFC), 2. Undecorated DNA filled capsids (FC), 3. Capsids without DNA and undecoration (empty capsids-EC). (D)FC and EC samples were produced by infecting Escherichia coli F cells with T5D18am-Δdec, and T5AmN5-del3 respectively as described in [Vernhes 2017] [18]. The capsid particles were purified by using 0.5M Sodium chloride – Polyethelyne (NaCl-PEG) glycol precipitation and Caesium chloride (CsCl) density-gradient centrifugation. They were then concentrated by NaCl-PEG precipitation and then purified by sedimentation velocity through a glycerol gradient followed by anion exchange chromatography [19]. Later, purified capsids were expanded in 25 mM Hepes Buffer pH 7.0 without NaCl and dialyzed against Phosphate buffer. The final concentration of DFC and FC were 2×1013 capsids/ml and 1×1013 capsids/ml respectively. In this thesis, I have prepared three batches of EC and the detailed description on the production of these samples are detailed in the next session.

III.4.2 PHAGE MUTANTS REPLICATION: INFECTION WITH HIGH MULTIPLICITY OF INFECTION

Empty capsids were obtained by infection of bacteria (E.coli F Strain) with a bacteriophage T5 amber mutant for the terminase protein responsible for DNA packaging (T5AmN5-del3). Due to the defect in DNA packaging in this mutant, capsid assembly stops after cleavage of the major capsid protein pb8 by the viral protease pb11, at the prohead 2 stage. Importantly, these viral particles, devoid of DNA will not bind to their tails.

After an overnight pre-culture of E.coli F in lysogeny broth (LB), the bacteria were cultured in LB supplemented with 1 mM MgCl2 and 1 mM CaCl2 (these salts being important for the lysis

of infected bacteria) at 37 °C, 220 rpm. When OD600 reached 0.5, the bacteria were infected with the phage amber mutant at a multiplicity of infection (MOI) of 5, such that all bacteria were statistically infected by at least one phage. After lysis of the bacteria, the larger debris were removed by centrifugation for 20 min at 6000 g and 4°C.

III.4.3 PHAGE PARTICLES PURIFICATION

III.4.3.1 NACL-PEG PRECIPITATION

The phage particles obtained following bacterial lysis were concentrated by precipitation in 0.5 M NaCl- 7.5% PEG 6000 at 4°C and centrifuged overnight.

III.4.3.2 GLYCEROL GRADIENT

The pellet obtained after precipitation by NaCl-PEG was resuspended in 4 mL of phage buffer (10 mM tris pH 7.6, 100 mM NaCl, 1 mM CaCl2, 1 mM MgCl2). The concentrated procapsids were centrifuged at 21,000 rpm at 8°C for 2 hours on a gradient of 10% to 40% glycerol, allowing separation according to sedimentation rates. After recovering all fractions from top to bottom, they were analyzed by SDS-Page to determine the capsid-containing fractions (see Figure III-5). Capsid-containing fractions displaying the presence of pb8 [4-10] were pooled for subsequent purification by anion exchange chromatography.

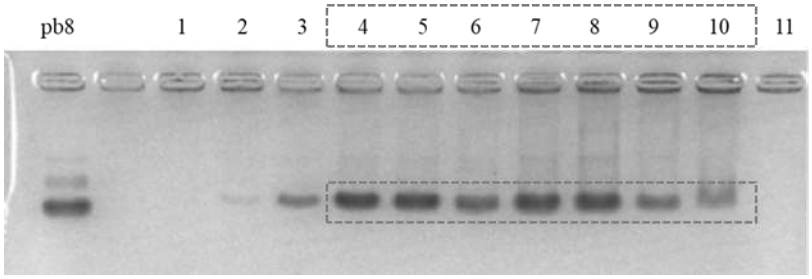


Figure III-5. SDS Page of fractions collected following centrifugation on glycerol gradient.

III.4.4 PURIFICATION BY ION EXCHANGE CHROMATOGRAPHY

The selected fractions from the glycerol gradient containing the protein of interest were loaded onto a HiTrapQ HP anion exchange column (GE Healthcare). (Figure III-6). The fraction of interest was eluted by a gradient linear from 0 to 100% of buffer B containing 1 M NaCl (Na + ions competing with protein). 260 nm and 280 nm wavelengths were monitored to assess the presence of DNA and proteins through the ratio between absorbance at the two wavelengths. The peak corresponding to fractions 36 to 40 shows the presence of pro-capsids devoid of DNA (same absorbance at 260 and 280 nm). After anion exchange purification pro-capsids were expanded in 25 mM Hepes Buffer pH 7.0 without NaCl and dialyzed against phosphate buffer (the final buffer).

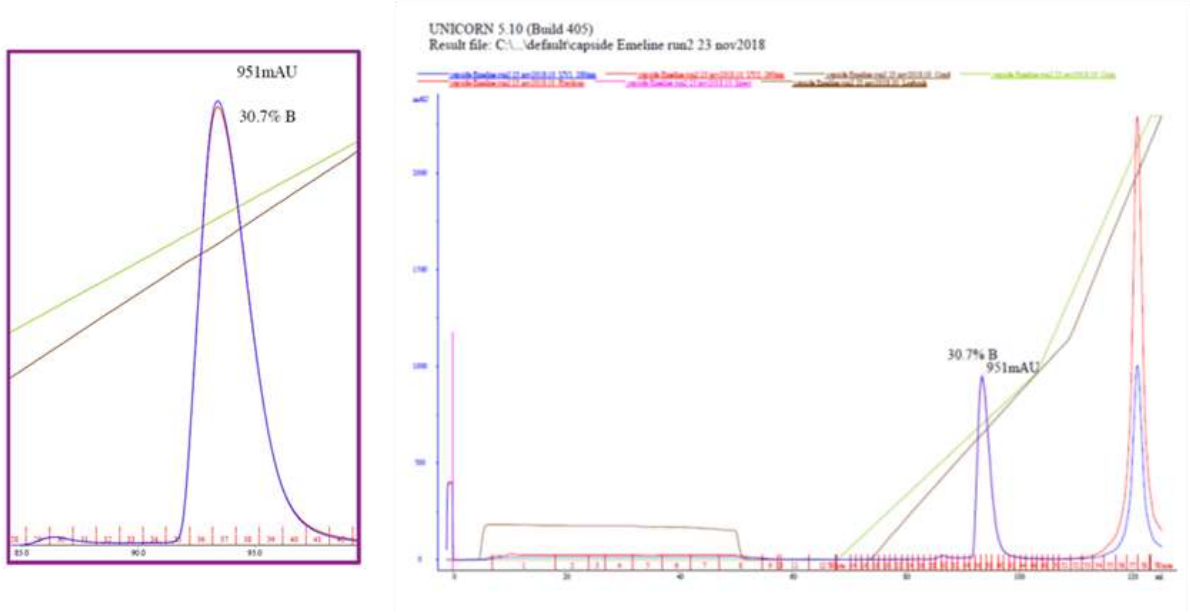


Figure III-6. Purification of empty capsids by anion exchange chromatography. (Equilibration Buffer A = 25mM Tris pH 8.0, 200 mM NaCl. Elution buffer B = 25 mM Tris pH 8.0, 1 M NaCl.) Fractions containing procapsids were loaded onto the column, and eluted from 0 to 42% of buffer B. The zoomed area shows the peak corresponding to the presence of phage procapsids.

III.4.5 SDS PAGE & IN-GEL TRYPSIN DIGESTION

The global proteomic assessment of T5 capsid started with 1D in-gel trypsin digestion of DFC samples. Capsids were solubilized in Laemmli buffer and heated at 95°C for 5 minutes. They were then stacked on top of a precast polyacrylamide Gel (NuPAGE™ 4%–12% bis-Tris protein gel, Invitrogen) and subsequently stained by Coomassie Blue. Gel bands containing capsid proteins were manually excised and chopped into small parts. The finely chopped pieces were washed several times for destaining by successive incubation in 25mM NH4HCO3 and 25mM NH4HCO3 in 50% acetonitrile (ACN) solution. To dehydrate the gel pieces, they were washed two times in 100% ACN. Denaturation of disulfide bonds was achieved by incubation in 10mM Dithiothreitol (DTT) in 25mM NH4HCO3 at 56°C for 45minutes. Cysteine alkylation was obtained by incubation with 55mM Iodoacetamide (IAM) in 25mM NH4HCO3 at room temperature for 35 minutes in the dark. To stop the alkylation reaction using 25mM NH4HCO3 for 10 minutes at 56°C. The next step was washing two times in 100% ACN. This step was followed by an overnight incubation at room temperature with Trypsin (Promega, Madison, WI) for digestion. Extraction of peptides from gel pieces were performed by in four washing steps [(1) 50% ACN/H2O, (2) 5% Formic acid, (3) 100% ACN and (4) 100% ACN], over 15 minutes long. The obtained peptide extracts were dried using a vacuum concentrator.

III.4.6 NANO LC/MS/MS BASED PROTEOMIC ANALYSIS

The trypsin digested peptides were resuspended in 5% acetonitrile, 0.1% Formic Acid acid and 95% of water. Solutions were analyzed by nano LC-MS/MS on an Ultimate 3000 & Q-exactive Plus Orbitrap mass spectrometer (Thermo Fisher Scientific). Peptides were separated on a C18 Reprosil (Dr. MAISH GmbH) column. The elution was performed using a 120-minute gradient (5% -37% of ACN in 0.1% Formic acid for 114 minutes and the last 6 minutes with 70% of ACN in 0.1% Formic acid) at a flowrate of 300 nl/min. Raw files were processed using Thermo Fisher Xcalibur Qual Browser Application. Mascot search engine version 2.6.0 was used to

analyze the MS/MS data by performing database searches against the UniProt Protein Databank Version 2021_01. The database search parameters are summarized in the *Table* III.1.

Enzyme	Trypsin
Missed cleavage	2
Fixed modification	Carbamidomethyl (C)
Variable modification	Acetyl (N-terminal), Oxidation
Fragment mass tolerance	25 mmu
Peptide charge state	2+ and 3+
Peptide mass error tolerance	10 ppm
Minimum peptide length	7
Specific peptide	1
Unique peptide	1
Minimum Score	25
Minimum false discovery rate	1% on score

Table III.1. Database search parameters, where mmu = Milli mass unit and ppm = Parts per million

III.4.7 PROTEIN IDENTIFICATION

Shotgun proteomic analysis approach detected the presence of T5 proteins, along with E.coli and other contaminants based on two unique peptides and 1% false discovery rate (FDR) (See Table III.2). Based on the data search parameters, all the five unique proteins of decorated filled T5 capsid including the protease were identified (See *Table* III.3). Identification of protein from protein sequence database using Mascot and quantification using Proline software [20] provided information on retention time, score, and abundance of protein, which helped us cross-reference the MS/MS data to the chromatographic peaks. In Table 8, we observe more peptides than the theoretically observable peptides because these are calculated without taking into account some of the miscleaved or modified peptides that we see in the "real" results.

	Serial No.	UniProt Accession #	Protein name	Description	Protein score	Peptide count
T5 proteins	1	Q6QGD8	CAPSD_BPT5	Major capsid protein	5131,699	60
	2	Q6QGD5	PORTL_BPT5	Portal protein	3198,011	37
	3	Q6QGD6	DECO_BPT5	Decoration protein	1514,51	22
	4	Q6QGD7	PRO_BPT5	Prohead protease	730,26	11
	5	Q6QGD9	HCP_BPT5	Head completion protein	200,54	3
E.coli Proteins	6	P0CE47	EFTU1_ECOLI	Elongation factor Tu 1	980,83	14
	7	P0A853	TNAA_ECOLI	Tryptophanase	667,3	11
	8	P0A799	PGK_ECOLI	Phosphoglycerate kinase	505,09	7
	9	P0A9B2	G3P1_ECOLI	Glyceraldehyde-3-phosphate dehydrogenase A	402,3	5
	10	P0A6M8	EFG_ECOLI	Elongation factor G	333,35	5
	11	P0ABB0	ATPA_ECOLI	ATP synthase subunit alpha	227,68	3
	12	P0A6Y8	DNAK_ECOLI	Chaperone protein DnaK	203,07	4
	13	P0A825	GLYA_ECOLI	Serine hydroxymethyltransferase	186,07	3
	14	P0A6F5	CH60_ECOLI	60 kDa chaperonin	147,38	2
	15	P06996	OMPC_ECOLI	Outer membrane protein C	133,86	2
	16	P0A6P9	ENO_ECOLI	Enolase	125,99	2
	17	P0AE08	AHPC_ECOLI	Alkyl hydroperoxide reductase C	118,36	3
	18	P0A836	SUCC_ECOLI	Succinate--CoA ligase [ADP-forming] subunit beta	87,58	2
Other contaminants	19	P13645	#C#P13645	Keratin, type I cytoskeletal 10	1383,04	20
	20	P04264	#C#P04264	Keratin, type II cytoskeletal 1	1373,45	19
	21	P35527	#C#P35527	Keratin, type I cytoskeletal 9	819,98	13
	22	P35908	#C#P35908	Keratin, type II cytoskeletal 2 epidermal	647,88	11
	23	P00761	#C#P00761	Trypsin	413,37	6
	24	P13646-1	#C#P13646-1	Keratin, type I cytoskeletal 13	332,75	6
	25	P02533	#C#P02533	Keratin, type I cytoskeletal 14	328,73	6
	26	P13647	#C#P13647	Keratin, type II cytoskeletal 5	275,43	6

Table III.2. Proteins identified by LC-MS/MS for the sample decorated filled capsid (DFC) with at least two peptide and false discover rate (FDR) of 1%.

Capsid Protein	Short name	Protein score	Peptides		MW	Abundance
			# Observable	# Observed		
Major capsid protein	pb8	5131.70	31	60	50885	2.44E+10
Decoration protein	pb10	1514.51	11	22	17247	3.40E+09
Portal Protein	pb7	3198.01	26	37	45087	6.99E+08
Protease protein	pb11	730.26	16	11	23379	5.46E+08
Head completion protein	pb144	200.54	11	3	19210	5.06E+06

Table III.3. The identified T5 structural proteins.

Figure below shows the sequence coverage of all detected T5 capsid proteins including the pb11 protease. Peptides of the T5 capsid that were detected by LC/MS are highlighted in violet. The pink highlighted portions shows the cleaved scaffolding domain of the major capsid protein and the cleaved portions of the capsid protease and portal protein. The fact that no peptides from the cleaved portion have been observed supports the processing mechanisms proposed earlier [7]. The DFC proteins with the highest sequence coverage was pb8 (~90%) and the least covered was p144 (19%). 70% of sequence coverage was obtained for both pb7 and pb10. And for pb11; it was about 66%, which provided unquestionable confirmation of the presence of pb11 in mature DFC.

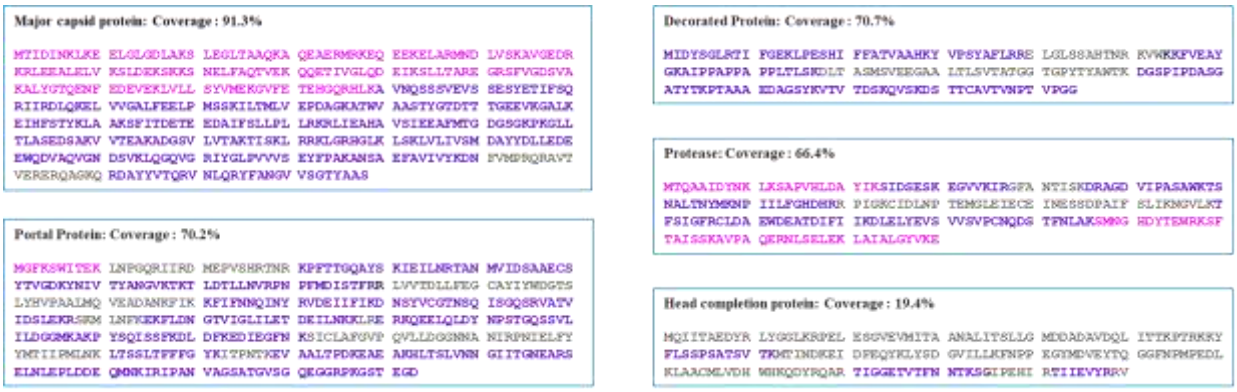


Figure III-7. Sequence coverage of bona fide T5 capsid proteins that were identified in the proteome characterization.

III.4.8 SEMI-QUANTITATIVE ANALYSIS

To gain insight into the quantity of pb11 in mature DFC, we initially relied on the so called "Proteomics Ruler Approach" [21]. This method was developed in 2015 by Jacek R.Wiśniewski et. al. to estimate the copy number of proteins in a sample using the sum of the intensities of MS signals corresponding to peptides from each protein. They reported that results obtained by this approach were in fair agreement with isotopic label quantification.

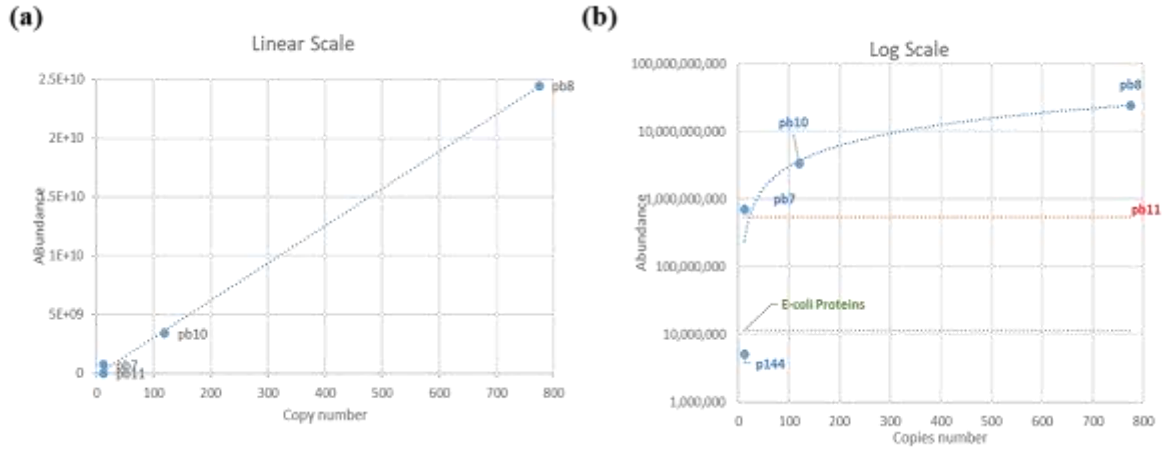


Figure III-8. Semi-quantitative analysis of capsid proteins using proteomics ruler. Sum of the abundance of peptides corresponding to each protein vs. copy number. The red line indicates the abundance of pb11. (b) The green line indicates the abundance of the most prevalent host contaminant protein.

Using this approach, we plotted the observed protein abundance vs. known copy number of the proteins of known stoichiometry. As postulated by the Proteomic Ruler Approach, we obtained a linear relationship between the sum of peptides abundances and the known protein copy numbers (*Figure III-8*a). When plotting the abundance of pb11 on this graph, it appeared similar to that of pb7, known to be present in 12 copies per T5 capsid. Plotting the same relationship with a log scale (*Figure III-8* b) confirmed that pb11 was present in higher abundance than contaminant proteins, and in amounts similar to pb7. The shotgun proteomics results thus suggested that the copy number of pb11 and pb7 could be similar. This was consistent with an estimation mentioned in a previously published article by Zivanoic et al. [8]. However, the proteomics ruler approach provides only a semi-quantitative view of the sample under investigation, and experimental biases may differ from one protein to another. To determine a more exact copy number, we turned to a novel method for absolute quantification using isotope-labeled QconCAT dipeptides.

III.5 ABSOLUTE QUANTIFICATION USING ISOTOPE-LABELED DIPEPTIDES

In 2018, Alexey Nesvizhskii et al. introduced a novel approach to determine Allele Specific Protein Expression (ASPE) using a QconCAT based proteomics approach to identify Cis-acting genetic variants [22] (*Figure* III-9). They developed this method based on high resolution multiple reaction monitoring (hrMRM) with a heavy stable isotope-labelled QconCAT internal protein standard. The standard was a concatemer of two peptides, corresponding to the mutant and wild type peptides to ensure equimolar amount upon digestion and thereby accurate quantification. In the present work, a similar approach was used to develop a novel targeted proteomic assay for the quantification of pb11 relative to pb7 in mature T5 capsid using heavy isotopically labelled QconCAT dipeptides.

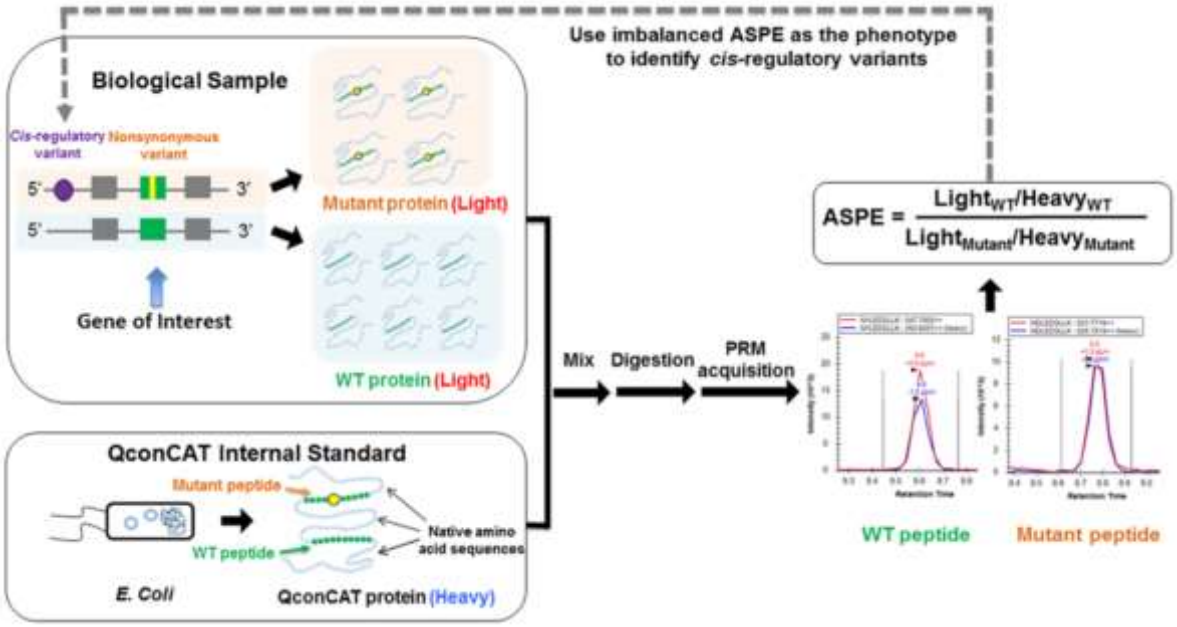


Figure III-9. Quantification of allele-specific protein expression (ASPE) using quantitative concatmer (QconCAT) internal protein standard. Reprinted from [22]

III.5.1 SELECTION OF STANDARD PEPTIDES

Selection of the two peptide to be concatenated in the QconCAT standard dipeptide was based on the shotgun proteomic analysis performed earlier. *Table* III.4 and *Table* III.5 list the identified peptides for pb7 and pb11 respectively. The criteria for choosing suitable quantification peptides are the following:

- Length of the peptide = 8 to 25 amino acid
- This peptide should be unique to that particular protein. This can be confirmed by performing Basic Local Alignment Search Tool (BLAST).
- The selected tryptic peptide should not contain any mis-cleavage area.
- There should not be any modified in the amino acid.
- 2+ and 3+ charge state are favorable precursor charge states, because of their m/z ranges.
- Symmetrical and narrow chromatographic peaks are preferred
- The retention time of the peptides should be different

Prev. AA	Peptide Sequence	Next AA	PTM (Carbamidomethyl& Oxidation)	Score	Charge	m/z	RT	Pep. match count	Abundance	Start	Stop	Missed Cl.
R	IPANVAGSATGVSGQEGGRPK	G		120.79	3	651.67756	27.0497	10	1.97E+08	377	397	1
R	VATVIDSLEKR	S		90.66	2	615.85663	40.9116	7	1.46E+08	177	187	1
R	VATVIDSLEK	R		73.86	2	537.80631	45.4086	4	4.64E+07	177	186	0
R	VDEIIFIK	D		43.97	2	488.78942	60.5164	3	4.28E+07	152	159	0
K	HLTSLVNNGIITGNEAR	S		135.57	2	904.98361	46.6794	3	3.63E+07	343	359	0
R	SELNLEPLDDEQMNK	I	Oxidation (M13)	65.67	2	895.9076	51.0255	6	3.12E+07	360	374	0
R	SELNLEPLDDEQMNK	I		74.77	2	887.91209	55.9963	3	2.46E+07	360	374	0
K	IRIPANVAGSATGVSGQEGGRPK	G		112.03	3	741.40285	36.1211	3	2.32E+07	375	397	2
R	KQEELQLDYNPSTGQSSVLILDGGMK	A	Oxidation (M25)	112.28	3	956.14351	63.4864	4	2.06E+07	222	247	1
K	DLDFKEDIEGFNK	S		94.43	2	785.37647	60.0932	4	1.96E+07	259	271	1
K	FIFNNQINYR	V		77.56	2	664.8414	57.2361	1	1.68E+07	142	151	0
K	KFIFNNQINYR	V		42.62	3	486.26174	46.9978	1	1.65E+07	141	151	1
K	DNSYVCGTNSQISGQSR	V	Carbamidomethyl (C6)	145.61	2	936.90819	29.1174	4	1.14E+07	160	176	0
K	AKPYSQISSFK	D		67.76	2	628.33662	30.4232	3	9.98E+06	248	258	1
K	TLDTLLNVRPNPFMDISTFR	R	Oxidation (M14)	64.05	3	789.41003	81.5616	9	8.89E+06	80	99	1
R	KQEELQLDYNPSTGQSSVLILDGGMK	A		117.89	3	950.81123	68.1137	4	8.67E+06	222	247	1
K	FLDNGTVIGLILETDEILNKK	L		89.89	3	782.43719	106.58	5	6.86E+06	197	217	1
R	SELNLEPLDDEQMNKIR	I	Oxidation (M13)	48.04	3	687.33666	49.4003	1	6.78E+06	360	376	1
K	LTSSLTFFFGYK	I		71.52	2	705.86862	83.2313	3	6.35E+06	311	322	0
R	SELNLEPLDDEQMNKIR	I		69.64	3	682.00562	56.8656	1	4.73E+06	360	376	1
K	TLDTLLNVRPNPFMDISTFR	R		84.46	3	784.0794	86.6664	2	3.51E+06	80	99	1
R	KPFTTGQAYSKIEILNR	T		114.29	3	656.03051	46.6636	1	3.30E+06	31	47	2
K	IRIPANVAGSATGVSGQEGGR	P		77.65	3	666.35619	40.3229	1	1.90E+06	375	395	1
R	ERKQEELQLDYNPSTGQSSVLILDGGMK	A		80.58	4	784.64605	63.7706	1	1.46E+06	220	247	2
R	TANMVIDSAAECSYTVGDK	Y	Carbamidomethyl (C12)	112.02	2	1016.4531	50.0965	1	1.19E+06	48	66	0
R	IPANVAGSATGVSGQEGGR	P		119.44	2	864.43695	30.507	2	8.85E+05	377	395	0
R	TANMVIDSAAECSYTVGDK	Y	Oxidation (M4); Carbamidomethyl (C12)	131.51	2	1024.451	44.5062	2	7.17E+05	48	66	0
R	TANMVIDSAAECSYTVGDKYNIVTYANGVK	T	Carbamidomethyl (C12)	107.29	3	1085.5131	67.3335	1	4.43E+05	48	77	1
R	TANMVIDSAAECSYTVGDKYNIVTYANGVK	T	Oxidation (M4); Carbamidomethyl (C12)	89.48	3	1090.8444	64.489	1	3.85E+05	48	77	1
K	QEELQLDYNPSTGQSSVLILDGGMK	A	Oxidation (M24)	90.75	3	913.44303	70.7088	1	3.06E+05	223	247	0
K	PYSQISSFK	D		59.06	2	528.77198	30.577	1	1.93E+05	250	258	0
K	EVAALTPDKEAEAK	H		56.44	3	491.25974	23.7334	3	1.68E+05	329	342	1
K	EKFLDNGTVIGLILETDEILNKK	L		103.58	3	868.14935	98.8717	1	1.63E+05	195	217	2
K	FLDNGTVIGLILETDEILNK	K		61.42	2	1109.1063	114.839	1	1.46E+05	197	216	0
K	QEELQLDYNPSTGQSSVLILDGGMK	A		89.3	3	908.11212	75.5696	1	1.41E+05	223	247	0
R	IPANVAGSATGVSGQEGGRPKGSTEGD	-		48.32	3	833.73733	28.2119	1	9.11E+04	377	403	2
K	EKFLDNGTVIGLILETDEILNK	K		53.81	3	825.45044	105.023	1	6.44E+04	195	216	1

Table III.4. List of pb7 peptides identified by shotgun proteomics. Carboxymethylated or oxidized residues shown in green and blue respectively. Missed-cleavage sites are highlighted in red. The yellow highlighted rows are suitable tryptic peptides. (AA= Amino Acid, PTM= Post-translational modification, RT= Retention Time, m/z= mass to charge ratio, Missed Cl= Missed cleavage)

Peptide Sequence	Next AA	PTM (Carbamidomethyl& Oxidation)	Score	Charge	m/z	RT	Pep. match count	Abundance	Start	Stop	Missed Cl.
AGDVIPASAWK	T		70.87	2	557.7981	46.81	6	2.87E+08	48	58	0
TFSIGFR	C		44.14	2	414.22435	53.41	7	1.54E+08	120	126	0
TSNALTNYMK	N		65.22	2	571.77942	40.39	4	6.44E+07	59	68	0
NPIILFGHDHR	R		83.24	2	659.85431	39.59	5	1.34E+07	69	79	0
DRAGDVIPASAWK	T		62.13	2	693.36286	44.31	2	1.30E+07	46	58	1
CLDAEWDEATDIFIIK	D	Carbamidomethyl (C1)	101.62	2	969.96116	88.87	4	4.62E+06	127	142	0
TSNALTNYMKNPIILFGHDHR	R	Oxidation (M9)	43.94	5	492.45201	54.28	1	3.85E+06	59	79	1
DLELYEVSVVSVPCNQDSTFNLAK	S	Carbamidomethyl (C14)	87.09	2	1364.1629	82.90	2	3.62E+06	143	166	0
TSNALTNYMKNPIILFGHDHR	R		43.37	4	611.31332	66.81	1	1.36E+06	59	79	1
TSNALTNYMK	N	Oxidation (M9)	72.18	2	579.77795	29.06	21	3.02E+05	59	68	0
SIDSESKEGVVKIR	G		56.46	3	516.28638	25.31	1	2.75E+05	24	37	2

Table III.5. List of pb11 peptides identified by shotgun proteomics Carboxymethylated or oxidized residues shown in green and blue respectively. Missed-cleavage sites are highlighted in red. The yellow highlighted rows are suitable tryptic peptides. (AA= Amino Acid, PTM= Post-translational modification, RT= Retention Time, m/z= mass to charge ratio, Missed Cl= Missed cleavage)

Accordingly, two tryptic peptides from proteins pb7 and pb11 were chosen as per their unique peptide sequence and for their detectability in LC-MS analysis. *Table* III.6 shows the two selected tryptic peptides and some of their properties.

Protein	Prev. AA	Peptide	Next AA	Score	z	m/z	Missed cleav.	RT (min)	Abundance	Start	Stop
Portal	R	VDEIIFIK	D	43,97	2	488,7894	0	60,5	4,28E+07	152	159
Protease	K	TFSIGFR	C	44,14	2	414,2244	0	53,4	1,54E+08	120	126

Table III.6. Peptides selected as quantification targets for pb7 and pb11. (AA= Amino Acid, PTM= Post-translational modification, z=Charge, m/z= mass to charge ratio, RT= Retention Time, Missed Cl= Missed cleavage)

III.5.2 PRODUCTION OF ISOTOPE LABELLED QCONCAT AQUA DIPEPTIDES

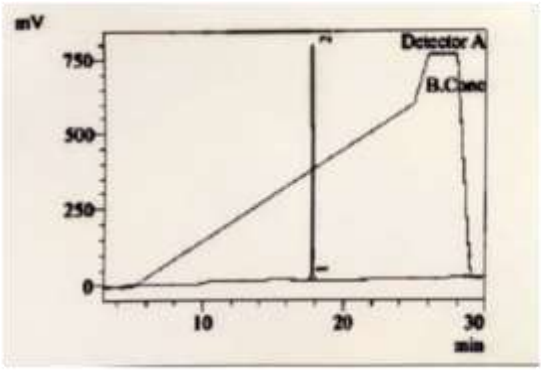
For this particular study, Lysine (K) and Arginine (R) labelled QconCAT peptides were purchased from Thermo Fisher Scientific Life Technologies SAS (Courtaboeuf, France). Figure III-10 shows the sequence of the dipeptide standards, which has been isotopically labelled [+10Da] on Arginine (R) at position 8 and [+8Da] on Lysine (K) at position 15. Quantitative information is provided in *Table* III.7.

Figure III-10. Sequence of the dipeptide concatemer standard. Boldface indicate isotopically labelled amino acids at position 8 and 15.

Standard dipeptide	Weight (Da)	Quantity per aliquot (nmol)	Weight (µg)	Volume (µL)	Concentration (g/L)
VDEIIFIKTFSIGFR	1801.99	1	1.8	200	0.009

Table III.7. Quantitative information of the dipeptide standard

Each vial of the standard contained 1 nmol of the lyophilized dipeptide standard. It was provided in precipitated form and subsequently resolubilized in 5% ACN, 0.1% Formic Acid, 94.9% H_2O. The manufacturer used analytical High Performance Liquid Chromatography (0.1% TFA in 50% ACN/H2O) method to check the purity of the sample (see *Figure* III-11)



Coloumn	150 X 4.6mm 3.6 C18 120 A	
Solvent	A: 0.05% TFA in water, B: 0.05% TFA in ACN	
Flowrate	1 ml/min	
Peak Table		
Peak #	RT	Area
1	17.553	1.551
2	17.776	98.449
Total		100

Figure III-11. Quality control HPLC of dipeptide standard showing >98% purity.

III.5.3 DIGESTION EFFICIENCY OF DIPEPTIDE STANDARDS

We evaluated the trypsin digestion efficiency of the standard by nano LC-MS/MS. Suitable digestion conditions for the standard dipeptide were obtained after trying different digestion protocols. A vial containing 1.8µg of the dipeptide standard was resuspended in 200 µL of 25mM ammonium bicarbonate. The solution was then digested by adding Trypsin/Lys-C at three different ratios: 1/50, 1/10 and 1/200 (See *Table* III.8). The solution was incubated at 37°C for 3 hours. Digestion was stopped by adding Formic acid to a pH 2. Desalting was performed by using Harvard C18 Reverse Phase Spin Columns following the manufacturer's protocol. Assessments of the digestion of the standard dipeptide was done by nano LC-MS/MS analysis.

VDEIIFIKTFSIGFR	1 nmol	200 µL
ABC 50mM	50 mM	200 µL
Digestion 1/10, 1/50 and 1/200		
Trypsin/Lys-C	1/10	11,25 µL
	1/50	2,25 µL
	1/200	0,55 µL
3 hours, 37°C		
Acidification		
TFA	10%	until pH 2
Desalting		

Table III.8 : Optimization of tryptic digestion conditions of the dipeptide standard.

Figure III-12 shows the base peak chromatograms obtained from the nano LC-MS/MS experiment for different quantities of trypsin and Lys-C. As expected the elution times of both peptides were not affected by the quantity of trypsin digestion ratio. Trypsin/Lys-C at a ratio of 1/10 (*Figure* III-12a) provided much better signal to noise when compared with 1/50 and 1/200 (*Figure* III-12b & c). Large background signals at low quantity of trypsin suggested incomplete digestion of the standard. We selected a Trypsin/Lys-C ratio of 1/10 for further experiments.

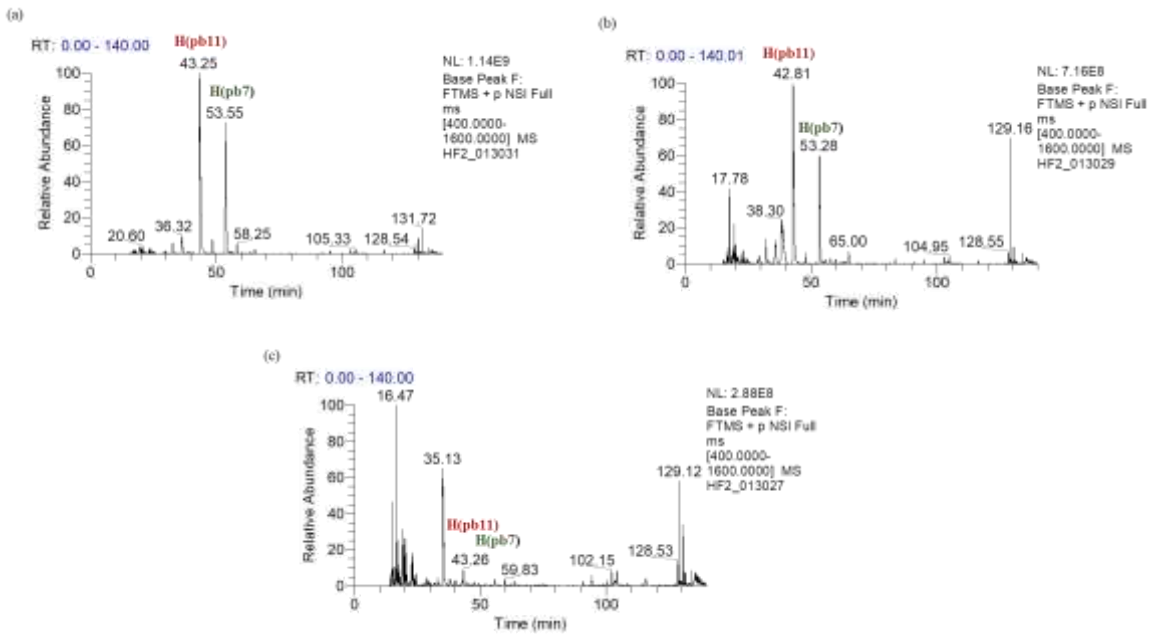


Figure III-12. Base peak chromatograms of LC-MS acquisitions of dipeptide standard digested with different LysC/Trypsin to standard ratios. a) 1/10; b) 1/50; c) 1/200.

III.5.4 pb11 QUANTIFICATION USING QCONCAT DIPEPTIDE STANDARD

DFC were subjected to in-gel trypsin digestion. The in-solution predigested dipeptide standard was subsequently added to the digested sample. This sample was subjected for LC-MS/MS Proteomic analysis. We performed two duplicate experiments, in which the injected volume was changed. *Figure* III-13 shows the observed chromatograms and mass spectra of the native tryptic peptides and standard peptides for DFC. To determine the best range of signal for quantification, the injection volume was set to 1µl for the first set of experiments (*Figure* III-13) and 5µl for the second set (Appendix I).

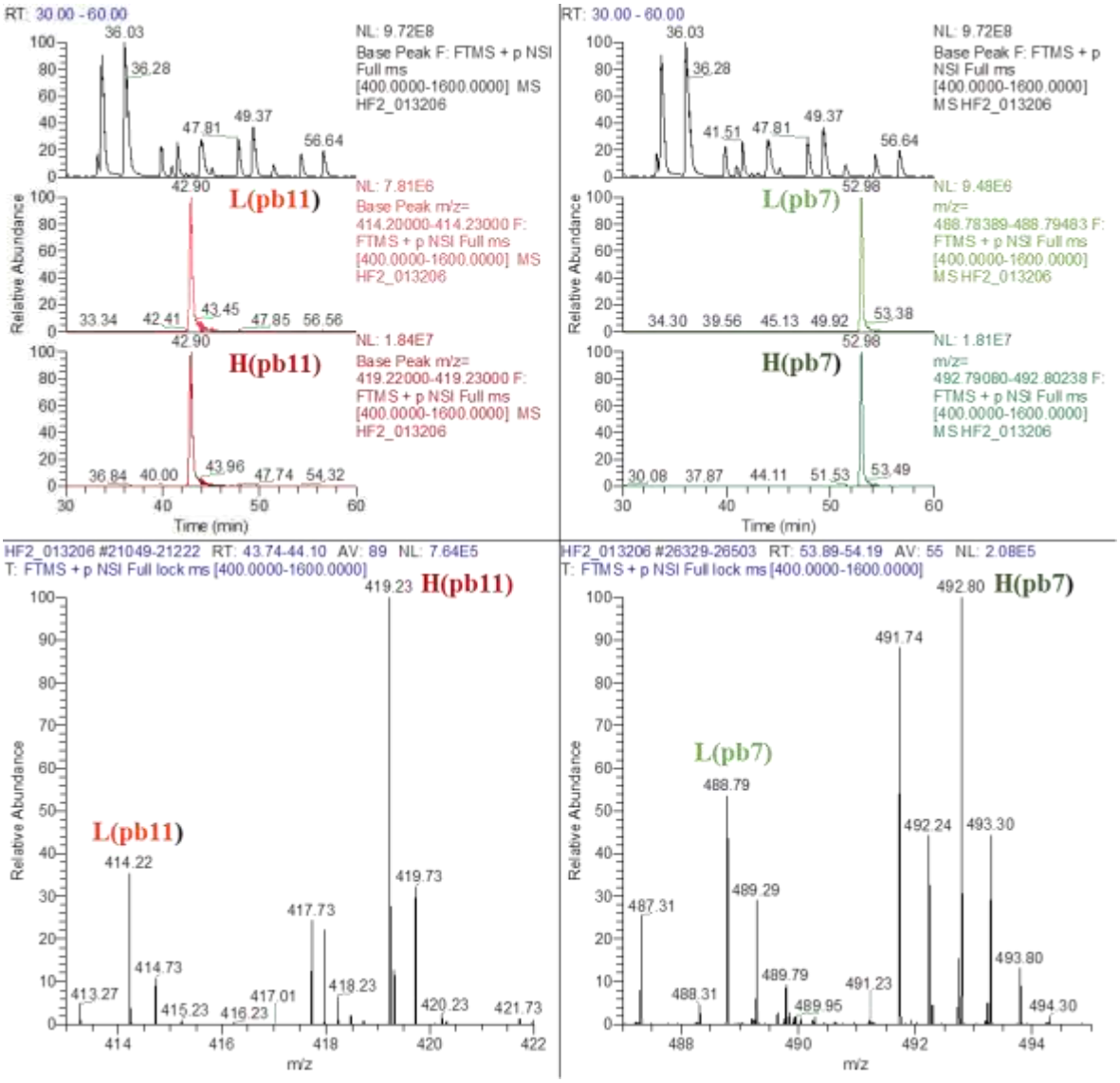


Figure III-13. Observed base peak chromatograms, and extracted chromatograms for the native tryptic peptides [Lpb7 and Lpb11] and the corresponding quantitation standards of pb11 (a) and pb7 (b) and the corresponding quantitation standards (Hpb7 and Hpb11) for the DFC sample. (c) and (d) represents their corresponding mass spectra when 1 µl was injected.

Table III.9 contains the information from the extracted ion chromatogram (XIC) of the DFC native peptide and dipeptide standard. We calculated the copy number of pb11 [N(pb11)] in each sample by comparing the ratio between native (Lpb7, Lpb11) and standard peptide (Hpb7, Hpb11) present in the analyzed sample, knowing the copy number of the portal protein (12

copies per capsid) and the 1:1 ratio of both standard peptide, due to the design of the standard. The calculation relied on the following equation:

$$N(pb11) = N(pb7) \times \frac{H(pb7) \times L(pb11)}{L(pb7) \times H(pb11)} \quad (3)$$

Sample	Inj. vol. (µL)	Signal intensity				Ratio L/H pb11	Ratio L/H pb7	pb11 Copy Number
		L(pb11) TFSIGFR	H(pb11) TFSIGF**R**	L(pb7) VDEIIFIK	H(pb7) VDEIIFI**K**			
DFC	1	7,81E+06	1,84E+07	9,48E+06	1,81E+07	0,42	0,52	10
DFC repl	1	1,03E+07	2,39E+07	1,15E+07	2,15E+07	0,43	0,53	10
DFC	5	4,53E+07	1,12E+08	4,65E+07	8,67E+07	0,40	0,54	9
DFC repl	5	5,22E+07	1,24E+08	4,12E+07	7,96E+07	0,42	0,52	10

Table III.9. Quantitative analysis of DFC with Dipeptide standards. L(pb11) and L(pb7) represents the native peptides of pb11 and pb7 whereas, H(pb11) and H(pb7) represents the isotopically labelled standard peptides of pb11 and pb7. While the ratios were quite consistent between the two injected quantities, the analysis with 1 µl injected displayed larger coefficients of variation.

From this experiment, we determined a copy number of pb11 in DFC of 10±1 copies per capsid in DFC sample. The copy number estimated from these experiments was consistent with the previous semi quantitative estimates by Zivanovic et al. [8]. However, since preliminary experiments were conducted with in-gel digestion, and because dipeptide standards were digested separately and mixed with the digested capsid proteins, questions regarding the relative digestion efficiency between in gel and in solution could be raised. Therefore, we decided to perform experiments where the capsid sample and dipeptide standard were mixed together and codigested in-solution.

Digestion in solution was performed using Preomics iST kit 8x [23]. This kit included chemicals to perform denaturation, reduction, alkylation of proteins and tryptic digestion. In solution digestion, using this kit was performed in three steps: 1. Lysis, 2. Digestion and 3. Purification. The total time that was required to perform this experiment was less than three hours. Importantly, as capsid protein and standard were co-digested, it circumvented issues of differential digestion of native and standard versions of the peptides. Different digestion (in gel opposed to in solution) but also extraction/purification. In-gel, extracted native peptides could have different retaining capabilities in contrast to the Harvard purification columns. Using Preomics iST kit 8x we do an in-solution digestion in the same vial, using the same enzyme. In addition, the heavy and native peptides are eluted from the purification column in the exact same conditions. Whatever happens to one the native peptide happens to the heavy peptide.

In-solution digestion experiment were performed in triplicate on three capsid samples: decorated filled capsid (DFC), filled capsid (FC) and empty capsid (EC). Nano LC/MS-MS based Proteomic Analysis was performed as described previously with identification using Mascot and quantification using Proline. Like before, the amount of protease present was determined by comparing the ratio between native (Lpb7, Lpb11) and standard peptide (Hpb7, Hpb11) present in the analyzed sample, knowing the copy number of the portal protein (12 copies per capsid). The observed ratios and the calculated copy number of pb11 for each capsid sample is summarized in *Table* III.10 for triplicate experiments. In these experiments, we observed that the signal intensity obtained for H(pb11) and H(pb7) are for each of the triplicate experiments. But for in-gel quantification, this was not the case and we observed similar signal intensity between the multiple experiments. In average DFC displayed a number of copies of pb11 of 17±1, FC 18±2 and EC 30±2.

Sample	Replicate experiment	Signal intensity				Ratio L/H pb11	Ratio L/H pb7	Copy Number pb11
		L(pb11) TFSIGFR	H(pb11) TFSIGF**R**	L(pb7) VDEIIFIK	H(pb7) VDEIIFI**K**			
DFC	1	6,670,000	789,000	27,100,000	4,430,000	8.45	6.12	17
	2	15,600,000	981,000	43,100,000	3,970,000	15.90	10.86	18
	3	10,800,000	1,270,000	39,200,000	5,990,000	8.50	6.54	16
FC 1	1	12000000	849,000	41800000	4120000	14.13	10.15	17
	2	11,100,000	568,000	35,700,000	2,990,000	19.54	11.94	20
	3	13,700,000	928,000	42,300,000	4,470,000	14.76	9.46	19
EC	1	2,410,000	627,000	6,640,000	4,280,000	3.84	1.55	30
	2	5,450,000	1,410,000	13,500,000	8,340,000	3.87	1.62	29
	3	6,520,000	1,660,000	15,100,000	10,100,000	3.93	1.50	32

Table III.10. Quantification results using in-solution digestion protocol.

These results confirm the presence of similar quantities of pb11 in mature capsid after DNA packaging, regardless of decoration. In addition, EC contained approximately 10 copies of pb 11 (~30%) more than FC and DFC. This difference could be due to partial ejection of pb11 from the capsid upon DNA packaging. Importantly, we could always detect peptides from pb11 in T5 capsids, but never any trace of the scaffolding domain of pb8 or the cleaved portions of pb7. Given the highly sophisticated and efficient processes at play in viral replication, the systematic presence of pb11 in mature capsids could hint at a role of this protease beyond its involvement in capsid assembly. Could this role be structural or functional? Could pb11 also be involved during infection? We can only speculate at this stage.

III.6 CONCLUSION

Even though pb11protein plays an important role in capsid maturation process of T5, the exact number of copies involved in this process is still undetermined. We have developed an isotope labelled standard based targeted proteomic assay to study the stoichiometry of pb11 in T5 capsids. We started with a global proteomic approach using nano LC-MS/MS to assess the proteins present in the bacteriophage T5 capsid. This semi qualitative approach showed the presence of all five capsid proteins including the protease, along with E.coli and contaminants due to the host and the sample handling respectively. Using the proteomics ruler method we estimated the copy number of individual capsid proteins using the abundance of the corresponding MS signals. Using this method, a linear relationship was observed proteins between copy number and their abundance, confirming that pb11 was present in higher abundance than contaminant proteins, and in amounts similar to the pb7. A targeted proteomic assay was proposed for the quantification of pb11 relative to the portal protein pb7 in mature T5 capsid using a heavy isotope labelled quantification concatemer (QconCAT) of two peptides. Upon trypsin digestion, this dipeptide releases two quantitation standards in a 1:1 ratio, allowing direct deduction of pb11 copy numbers knowing pb7 stoichiometry. We thus observed 18±2 copies of pb11 in capsids containing DNA and 30 ±2 copies in empty procapsids (prohead II) that were isolated before DNA packaging. The lower number of copies in FC and DFC suggests that few copies of pb11 might have left the capsid in order to accommodate the DNA.

Confirmation on the presence of pb11 in mature T5 capsid even after the incorporation of DNA supports the fact mentioned by Zivanovic et. al that pb11 is not only a morphogenetic enzyme but also has a potential role in the structural composition of the mature capsid [8]. Presence of protease in the mature capsid may also suggest a role in DNA replication or perhaps ingenome stability. However, the fact that copies of pb11 seem to exit the capsid upon DNA packaging

do not support this hypothesis. It is our aspiration that the knowledge provided in the present study, regarding the copy number of the pb11 protease in mature T5 capsids, may help future research to reveal its exact situation in the capsid structure.

REFERENCES

1. Capsid expansion of bacteriophage T5 revealed by high resolution cryoelectron microscopy | PNAS. https://www.pnas.org/content/116/42/21037

2. Effantin G, Boulanger P, Neumann E, Letellier L, Conway JF (2006) Bacteriophage T5 structure reveals similarities with HK97 and T4 suggesting evolutionary relationships. J Mol Biol 361:993–1002 . https://doi.org/10.1016/j.jmb.2006.06.081

3. Huet A, Duda RL, Boulanger P, Conway JF (2019) Capsid expansion of bacteriophage T5 revealed by high resolution cryoelectron microscopy. PNAS 116:21037–21046 . https://doi.org/10.1073/pnas.1909645116

4. Wikoff WR, Liljas L, Duda RL, Tsuruta H, Hendrix RW, Johnson JE (2000) Topologically linked protein rings in the bacteriophage HK97 capsid. Science 289:2129–2133 . https://doi.org/10.1126/science.289.5487.2129

5. Garcia-Doval C, Luque D, Castón JR, Boulanger P, van Raaij MJ (2013) Crystallization of the C-terminal domain of the bacteriophage T5 L-shaped fibre. Acta Crystallogr Sect F Struct Biol Cryst Commun 69:1363–1367 . https://doi.org/10.1107/S1744309113028959

6. Rossmann MG, Rao VB (2012) Rossmann MG, Rao VB(2012) Viral Molecular Machines.Springer-Verlag, New York. DOI: 10.1007/978-1-4614-0980-9

7. Huet A, Conway JF, Letellier L, Boulanger P (2010) In Vitro Assembly of the T=13 Procapsid of Bacteriophage T5 with Its Scaffolding Domain. J Virol 84:9350–9358 . https://doi.org/10.1128/JVI.00942-10

8. Zivanovic Y, Confalonieri F, Ponchon L, Lurz R, Chami M, Flayhan A, Renouard M, Huet A, Decottignies P, Davidson AR, Breyton C, Boulanger P (2014) Insights into Bacteriophage T5 Structure from Analysis of Its Morphogenesis Genes and Protein Components. Journal of Virology 88:1162–1174 . https://doi.org/10.1128/JVI.02262-13

9. Ponchon L, Boulanger P, Labesse G, Letellier L (2006) The Endonuclease Domain of Bacteriophage Terminases Belongs to the Resolvase/Integrase/Ribonuclease H Superfamily: A BIOINFORMATICS ANALYSIS VALIDATED BY A FUNCTIONAL STUDY ON BACTERIOPHAGE T5*. Journal of Biological Chemistry 281:5829–5836 . https://doi.org/10.1074/jbc.M511817200

10. Preux O, Durand D, Huet A, Conway JF, Bertin A, Boulogne C, Drouin-Wahbi J, Trévarin D, Pérez J, Vachette P, Boulanger P (2013) A two-state cooperative expansion converts the procapsid shell of bacteriophage T5 into a highly stable capsid isomorphous to the final virion head. J Mol Biol 425:1999–2014 . https://doi.org/10.1016/j.jmb.2013.03.002

11. Huet A, Duda RL, Hendrix RW, Boulanger P, Conway JF (2016) Correct Assembly of the Bacteriophage T5 Procapsid Requires Both the Maturation Protease and the Portal Complex. J Mol Biol 428:165–181 . https://doi.org/10.1016/j.jmb.2015.11.019

12. Liu H, Sadygov RG, Yates JR (2004) A Model for Random Sampling and Estimation of Relative Protein Abundance in Shotgun Proteomics. Anal Chem 76:4193–4201 . https://doi.org/10.1021/ac0498563

13. Nahnsen S, Bielow C, Reinert K, Kohlbacher O (2013) Tools for Label-free Peptide Quantification. Molecular & Cellular Proteomics 12:549–556 . https://doi.org/10.1074/mcp.R112.025163

14. Ong S-E, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1:376–386 . https://doi.org/10.1074/mcp.m200025-mcp200

15. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17:994–999 . https://doi.org/10.1038/13690

16. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci U S A 100:6940–6945 . https://doi.org/10.1073/pnas.0832254100

17. Beynon RJ, Doherty MK, Pratt JM, Gaskell SJ (2005) Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. Nat Methods 2:587–589 . https://doi.org/10.1038/nmeth774

18. Vernhes E Maturation de la capside du bactériophage T5: étude structurale et fonctionnelle de la protéine de décoration pb10. 235

19. Preux O, Durand D, Huet A, Conway JF, Bertin A, Boulogne C, Drouin-Wahbi J, Trévarin D, Pérez J, Vachette P, Boulanger P (2013) A Two-State Cooperative Expansion Converts the Procapsid Shell of Bacteriophage T5 into a Highly Stable Capsid Isomorphous to the Final Virion Head. Journal of Molecular Biology 425:1999–2014 . https://doi.org/10.1016/j.jmb.2013.03.002

20. (2020) Proline: an efficient and user-friendly software suite for large-scale proteomics. In: ProFI Proteomics. https://www.profiproteomics.fr/proline-an-efficient-and-user-friendly-software-suite-for-large-scale-proteomics/. Accessed 25 Jul 2021

21. Wiśniewski JR, Hein MY, Cox J, Mann M (2014) A "proteomic ruler" for protein copy number and concentration estimation without spike-in standards. Mol Cell Proteomics 13:3497–3506 . https://doi.org/10.1074/mcp.M113.037309

22. Shi J, Wang X, Zhu H, Jiang H, Wang D, Nesvizhskii A, Zhu H-J (2018) Determining Allele-Specific Protein Expression (ASPE) Using a Novel Quantitative Concatamer Based Proteomics Method. J Proteome Res 17:3606–3612 . https://doi.org/10.1021/acs.jproteome.8b00620

23. 5f8567b72a11f1ff9e759725_PreOmics_iST_method_8rxn_v5.pdf

CHAPTER IV: NANORESONATOR-BASED MS OF T5 CAPSIDS

This chapter studies the factors influencing the NEMS-MS mass estimates. First, we demonstrate the mass measurements that were obtained for Empty T5 capsids, which lead us to investigate the multiple factors that concerns the performance of NEMS-MS measurements. After discussing the particles extraneous physical properties, we proceed to the influence of frequency noise and device fabrication defects. We then present experimental results investigating the effects of particles desolvation. This work has been published in the journal of Analytical and Bioanalytical Chemistry in 2021[1].

IV.1 STATE OF THE ART

The potential of NEMS-MS for mass measurement of individual capsids of bacteriophage T5 has been demonstrated recently [2]. In the published study, solutions containing DNA filled capsids (FC) and empty capsids (EC, i.e. devoid of the viral DNA), produced by infecting Escherichia coli (E.coli) F cells using T5stAmN5 and T5D18am-Δdec T5 mutants respectively, were nebulized using nano-ESI prior to mass analysis. To do so, the salt concentration of capsid samples had to be controlled in order to maintain a stable spray. For this purpose, dialysis against 25mM ammonium acetate (AmAc) was performed, followed by addition of a 10% methanol. The final concentrations of EC and FC used in these experiments were 5.9×1011 and 8.8×1011 capsid/ml respectively.

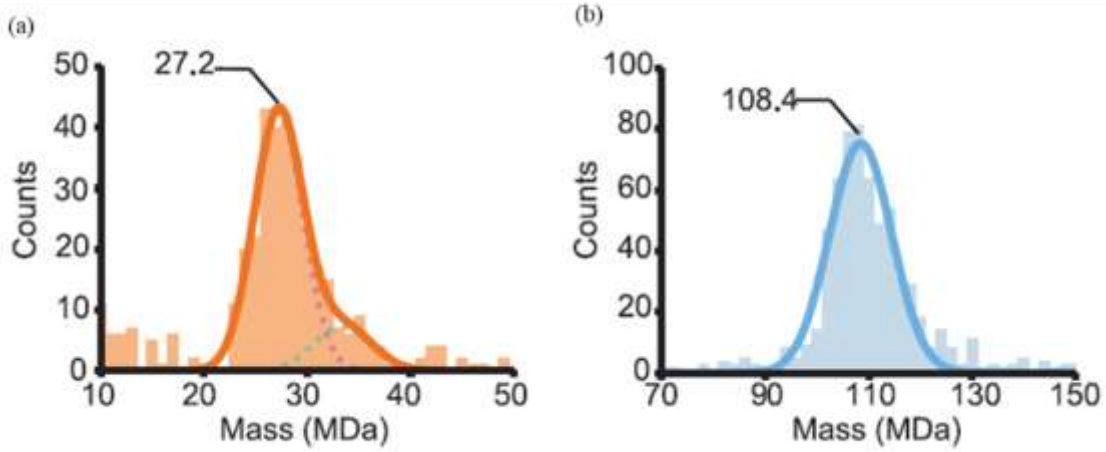


Figure IV-1. Accumulated Mass histogram and Gaussian fits for Empty capsids (a) and Filled capsids (b), reprinted from [2].

The mass distribution obtained for the EC sample is shown in *Figure* IV-1 a. For this spectrum, 363 individual particle-landing events, in the mass range of interest were collected over a period of 302 minutes. The data exhibited a bimodal distribution and was fitted using two Gaussian peaks. The central mass of the main peak was 27.2MDa and the second peak 33.4 MDa, to be compared to a theoretical mass of ~26.0 MDa for EC. The justification provided for the second peak was the partial incorporation of host DNA fragments in the capsid. The mass distribution observed for the FC sample is shown in *Figure* IV-1 b. For this spectrum, 648 individual particle landing events were recorded in 655 minutes. The center mass obtained for FC was 108.4 MDa, which was ~2.5% higher than the theoretical mass of 105.4 MDa. The explanation provided for

the observed mass discrepancies were the presence of residual solvent or salt adducts to the capsid particles, during the ESI process. SEM (Scanning Electron Microscopy) images provided in the study, that capsids reaching the NEMS resonator apparently preserved their structural integrity.

In the present work, I intended to extend this proof of concept experiment and develop a routine methodology to analyze the mass of T5 capsids as a substitute to size or morphological characterization. For this purpose, I produced and purified several lots of empty bacteriophage T5 capsids, and attempted to replicate the previous analytical protocol. However, issues with spray instabilities and resonators unlocking precluded routine analysis of the capsids. I therefore investigated sample desalting conditions and their effects on EC samples using Nanoparticle-Tracking Analysis (NTA). This finally yielded ESI compatible solutions and facilitated NEMS-MS mass measurements. However, the masses were always in excess of the theoretical estimates. To make sense of the observed mass discrepancies, I made an inventory of the various factors that affect mass determination by NEMS-MS and evaluated the uncertainty of the measurement. I finally investigated how spray parameters would affect the mass measurement.

IV.2 PRODUCTION AND CHARECTERIZATION OF EMPTY T5 CAPSIDS

This work started with the production and purification of bacteriophage T5 empty capsids following a protocol described previously by Huet et al. [3]. These experiments were performed at the laboratory and under the expert guidance of Pascale Boulanger, at the Institute for Integrative Biology of the Cell, University Paris Saclay. The protocol used to prepare the empty capsids is described in Chapter III (section: Phage Mutants Replication: Infection With High Multiplicity Of Infection). We prepared three batches of empty capsids that we will refer to EC2a, EC2b and EC2c. EC1 corresponds to the empty capsids analyzed previously by Dominguez-Medina et al. [1] and that serve as our benchmark. EC3 corresponds to an additional batch of capsids kindly provided by Pascale Boulanger's team at the end of our project.

IV.2.1 NANOPARTICLE-TRACKING ANALYSIS

The size and concentration of the produced capsid samples were estimated using Nanoparticle Tracking Analysis (see *Table* IV.1and *Figure* IV-2). The NTA results show fairly monodisperse samples with minor contributions from larger and smaller particles. Concentrations were in excess of 1×10^3 particles /ml for EC2a and EC2b, but the yield was lower for sample EC2c. The larger mean radius obtained for EC2c, along with the lower concentration suggests that capsid assembly may not have progressed optimally for this sample, and it was not considered for further experiments.

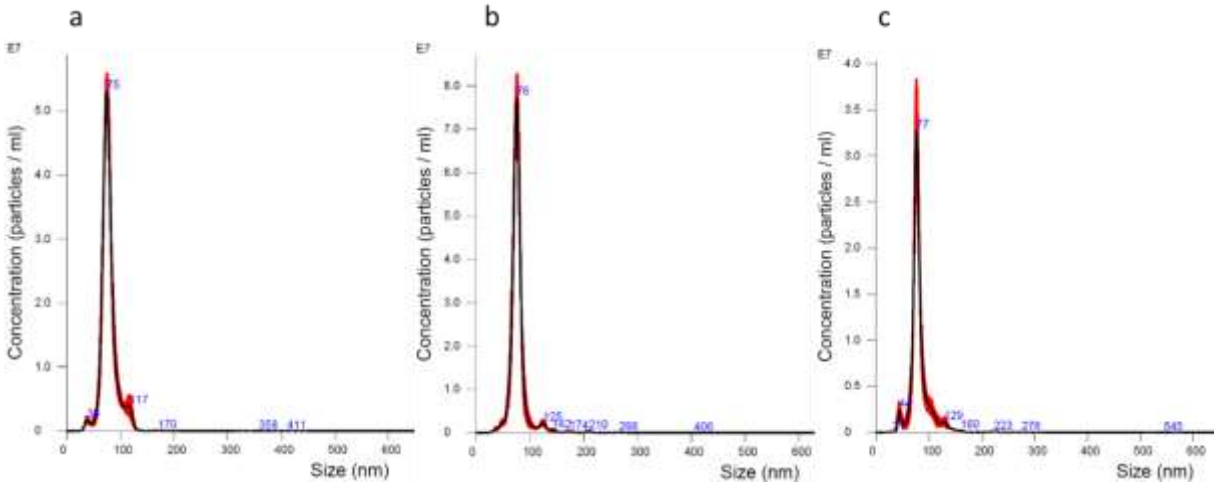


Figure IV-2. NTA measurement of hydrodynamic radius and concentration for samples EC2a, b and c.

Abbreviation	Mean	Mode	SD	Concentration NTA	Concentration estimate
EC1*	78.0	69.8	10.0	$1.66x10^{13}$	$1.00x10^{13}$
EC2a	78.3	74.6	16.0	$1.10x10^{13}$	$4.97x10^{13}$
EC2b	77.9	75.8	17.8	$1.40x10^{13}$	$2.32x10^{13}$
EC2c	81.0	76.6	19.7	$5.70x10^{12}$	$4.00x10^{12}$
EC3	79.1	79.8	20.2	$1.63x10^{13}$	$1.00x10^{13}$

*Table IV.1 Nanoparticle Tracking Analyzer characterization of empty T5 capsids batches. * Benchmark data from Ref [1]*

IV.3 SAMPLE PREPARATION STUDY

Following the Dominguez-Medina *et. al.* method of sample preparation [1], empty capsids from batch EC2a were dialyzed against 25 mM Ammionum Acetate in deionized water for 24 hours at 0°C, and 10% Methanol was added to a final concentration of ~5×10^{11} capsid/ml. The sample was sprayed at a flow rate of 30 µl/hr, and voltage ranging from 2-3 kV. The inlet capillary temperature was set to 150°C. It should be noted that under these conditions, the spray was difficult to stabilize due to clogging of the ESI tip, and spray disruptions were frequent (see *Figure* IV-3). In addition, NEMS resonators were frequently unlocking (i.e. the phase lock loop was unable to track the resonator's frequency) when performing these experiments. Thus, in my hands, the original protocol did not yield exploitable results.

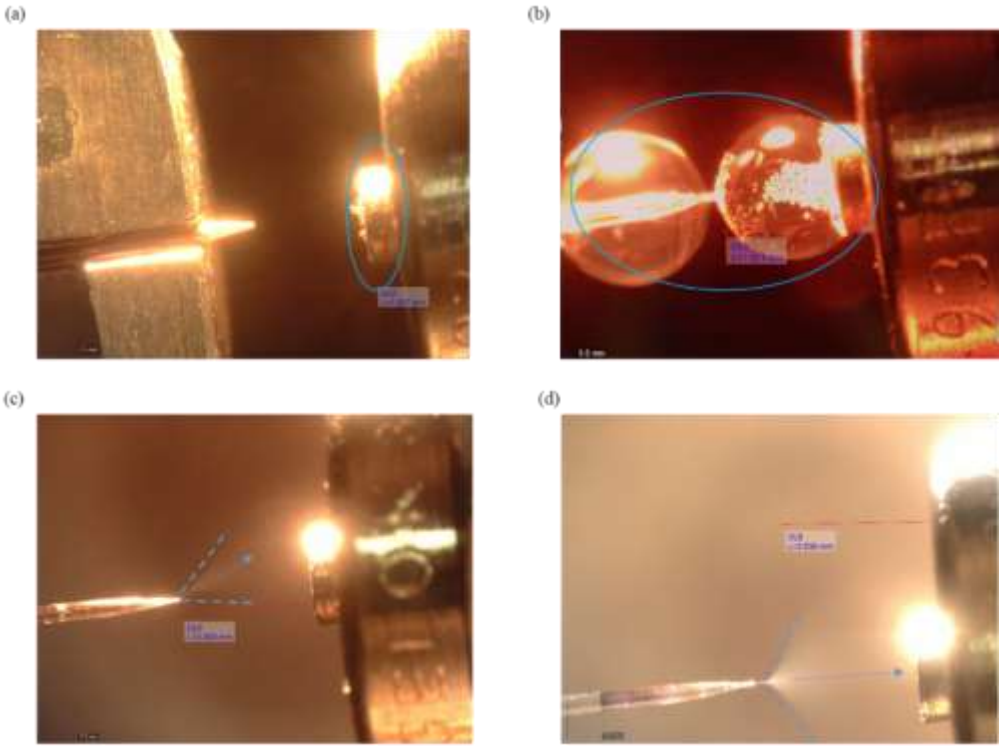


Figure IV-3. Images of the ESI source of NEMS MS system during spraying of T5 capsids. (a) Salt/impurity crystals formed on the inlet capillary (b) loss of electrical contact within the ESI tip, leading to the formation of large drops blocking the inlet capillary (c) Unstable ESI spraying condition caused by partial clogging of the tip (d) Suitable ESI spraying condition

Figure IV-4 shows a typical mass spectrum obtained using the original sample preparation methodology. The x-axis shows the mass in MDa and y-axis shows the number of particles measured. This spectrum clearly shows deposition of particles with variable masses all over the mass range. Only 118 events were recorded over a period of 378 minutes. The difficulty in spraying, the deposit of large masses and the unlocking of NEMS devices could be attributed to the presence of residual salts and of possible particle aggregates in solution. These results question this sample preparation protocol for routine capsid mass analysis.

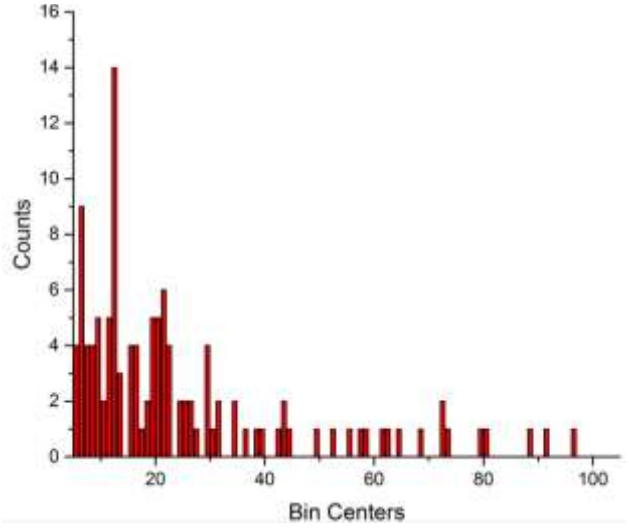


Figure IV-4. Mass measurement of an empty capsid sample after 24hr dialysis, showing no clearly distinguishable peak around the expected mass (26 MDa). Events were scattered all over the 1-100 MDa mass range (average event rate: ~1 event/100sec).

IV.4 BUFFER EXCHANGE

High amounts of salt in the capsid storage buffer constitute a challenge for ESI at the front end of our NEMS-MS system. Therefore, it was necessary to exchange the buffer to a suitable solvent. Deionized water and 25 mM ammonium acetate in water were tested. Two methods were available to us for buffer exchange: ultra-centrifugal filtration (AmiconTM, 30kDa cutoff) or membrane dialysis (Slide-A-LyzerTM dialysis cassettes, 20kDa cutoff) [4].

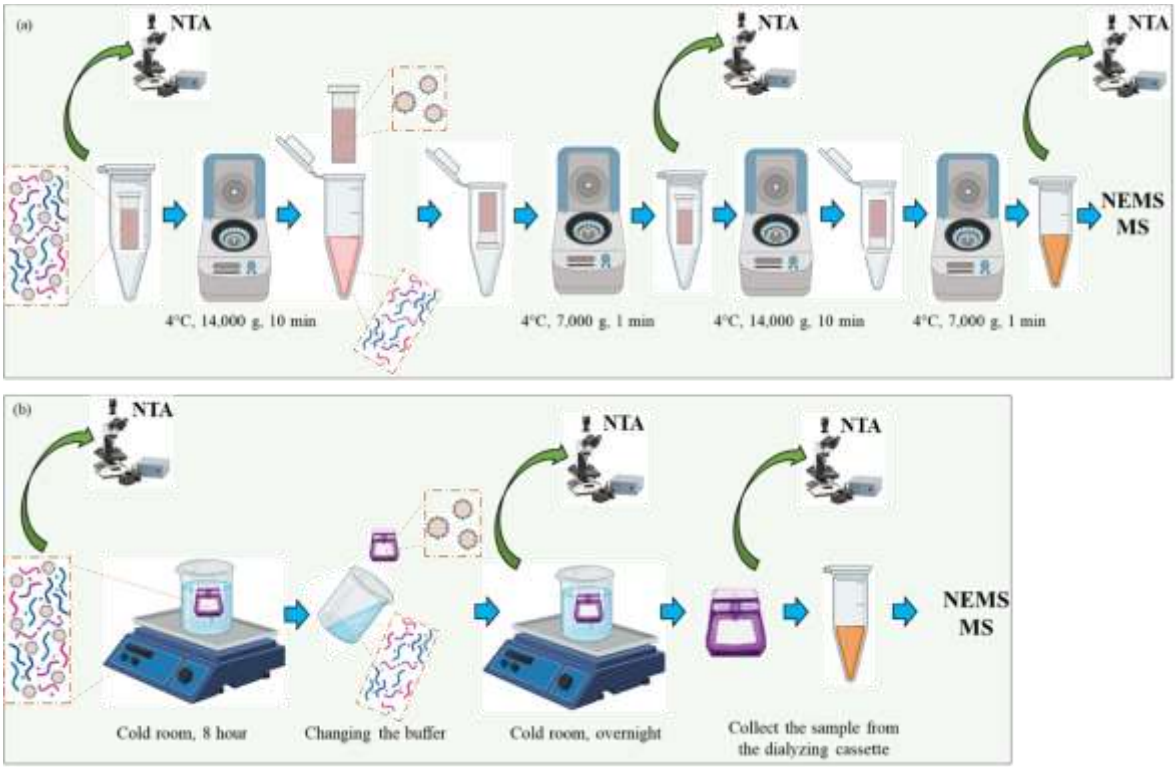


Figure IV-5. Buffer exchange protocols using (a) ultra-centrifugal filtration and (b) membrane dialysis. Aliquots were collected at each stage of the two protocols to derive capsid sizes and concentration by NTA.

We designed an experiment to study the effect of buffer exchange protocols on capsid samples using NTA (see *Figure* IV-5). For each sample preparation method, two rounds of buffer exchange were performed (several experiments based on the solvents and number of filtration were performed, the data is shown in Appendix II). For ultra-filtration, the sample was deposited in the cartridge and subjected to two cycles of dilution and centrifugation. Aliquots were taken from the original sample, and after each resuspension and then analyzed by NTA. For dialysis, aliquots were collected at t=0, t=8hr and after overnight treatment, and later analyzed by NTA.

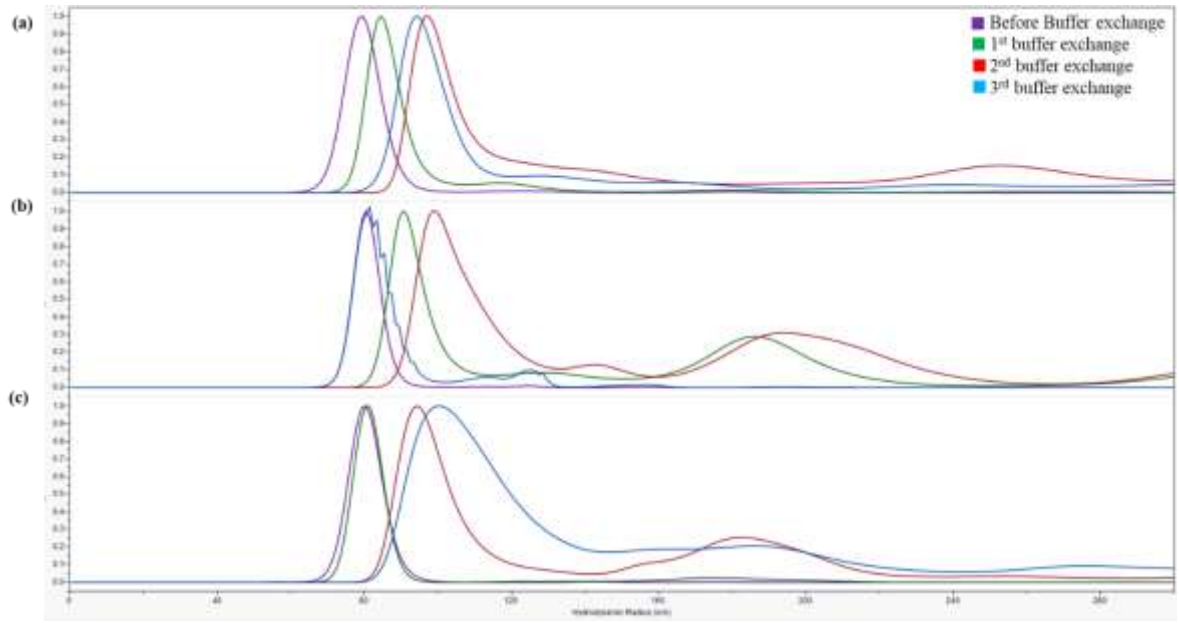


Figure IV-6. NTA size measurements of filled capsids before and during the buffer exchange process. a) Dialysis against ammonium acetate led to an increase in capsid size, followed by aggregation. b) Dialysis against water caused immediate aggregation after the first round of buffer exchange. c) The first round of ultrafiltration against pure water did not alter the capsid profile, although further treatment yielded capsid swelling and aggregation.

Figure IV-6 shows the hydrodynamic diameter distributions measured by NTA for FC particles at each stage of sample preparation for the two proposed methods. Dialysis against ammonium acetate resulted in a slight increased in particle size most likely due to osmosis, followed by particle aggregation as shown in *Figure* IV-6 a. When dialysis was performed against pure water, this behavior was even more pronounced and aggregates dominated the NTA profile even after a single round of buffer exchange (*Figure* IV-6 b). Conversely, for ultra-centrifugal filtration, the particle size was not affected by a single round of buffer exchange, even with pure water (*Figure* IV-6 c). Therefore, for the further experiments with NEMS MS system we opted for a single ultra-centrifugal filtration process with water.

For these experiments, we used filled capsids (FC) instead of EC as they were readily available in larger quantities. Since the surface proteins are identical for EC and FC, we believed that the results obtained could be extended to empty capsids. Indeed, limited tests on EC samples confirmed the presented findings (*Figure* IV-7).

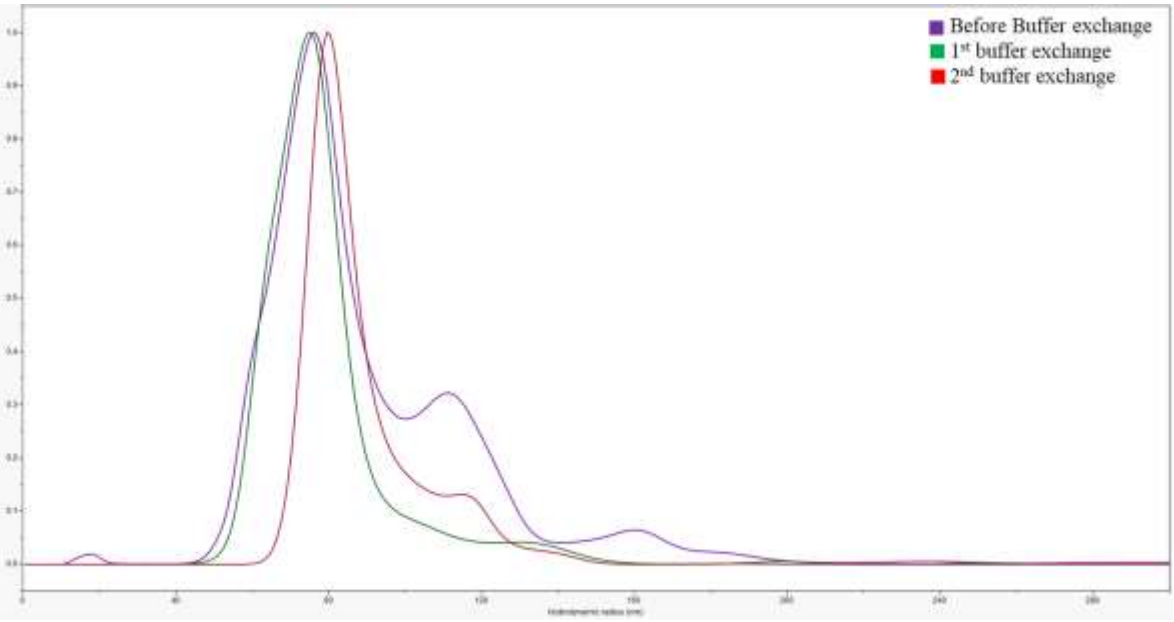


Figure IV-7. NTA size measurements of empty capsids using ultracentrifugation. One round of ultrafilteration did not change the size of the capsid.

IV.5 NEMS-MS MEASUREMENTS

Before the analysis NEMS system were set into the desired vacuum conditions. Once the pressure stabilized, the inlet capillary was heated to a temperature of 175°C. A slightly higher temperature than in previous experiments was selected in order to provide better desolvation in an attempt to mitigate the observed mass discrepancy [5]. Following a single round buffer exchange against water using Amicon filter, EC samples were diluted to a final concentration of 5×10^{11} capsid/ml and nebulized using nano-ESI with 30 µm ID New Objective (Littleton, MA, USA) picotip ESI emitter, at a flow rate of 30µl/hour. While performing this experiment the ESI spray was more stable, the needle didn't clog and the NEMS devices didn't unlock as frequently, supporting our sample preparation protocol.

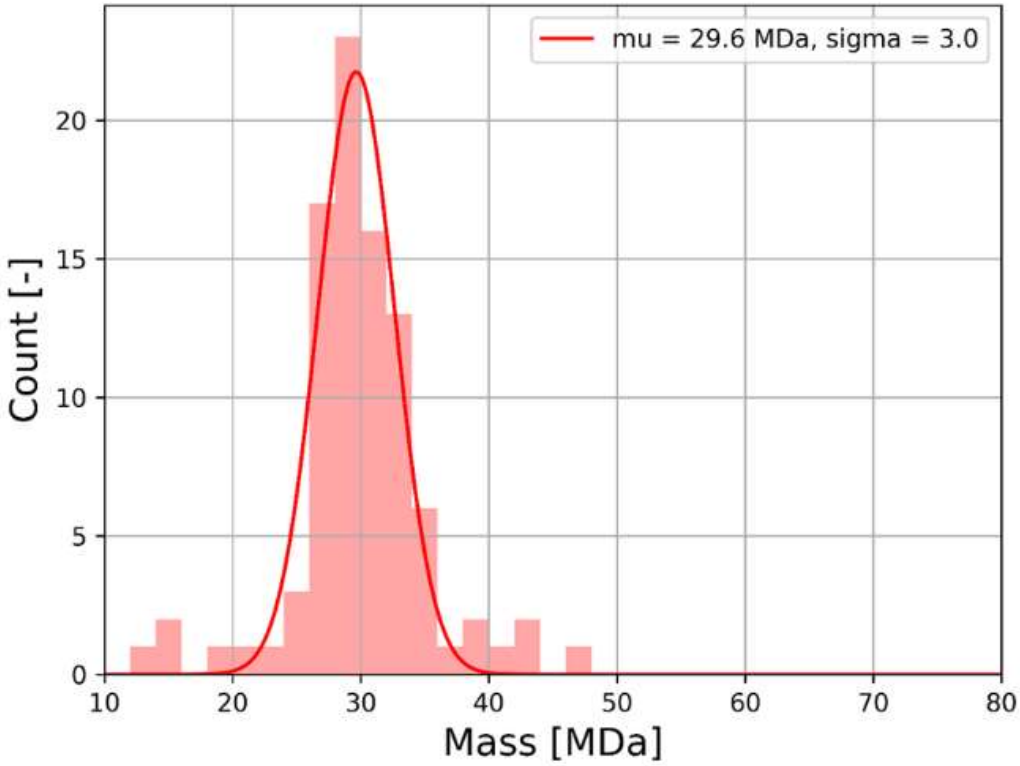


Figure IV-8. Measured mass histogram for EC with Gaussian fit.

Figure IV-8 shows the mass distribution obtained for EC. It was fitted using a Gaussian peak, yielding a central mass of 29.6 MDa with a standard deviation of 3 MDa. 225 particle-landing events were detected, over the course of 350 minutes, with an event rate ranging from 1.5-2 events/minute.

When comparing to the dialyzed sample experiment, performing experiment with ultra-filtered sample was easier. Mostly because the preparation of the sample saved a lot of time: dialyzing sample took 24 hours, while amicon filtration took only 10min. We believe that the frequent unlocking issues experienced in our previous attempts could be due to the large aggregates observed by NTA after multiple rounds of buffer exchange.

The theoretical mass of T5 capsids can be calculated based on the elementary composition of its constitutive proteins. The atomic element that compose T5 proteins are reported in Table IV.2. The reference values for the lower and upper bound of their Standard Atomic Weight of elements were taken from the official site of NIST [6] [7].

Elements	Standard atomic weight	
	Lower bound	Upper bound
C	12.00960	12.01160
H	1.00784	1.00811
N	14.00643	14.00728
O	15.99903	15.99977
S	32.05900	32.07600

Table IV.2. Standard atomic weights of selected elements, taking into account the exact masses (Da) of their isotopes and the uncertainties in their natural distribution [1]

According to Huet et al., the copy number for the proteins pb7 (processed portal protein) and pb8 (processed head protein) are 775 and 12 respectively. In addition, we detected 30 copies of the protease pb11 in empty capsids in our proteomic experiments (see Chapter III) [8]. Using the upper and lower bounds of atomic weights of the atoms contained in each protein, we can make the estimates reported in Figure IV-9 .Finally, we conclude that the theoretical mass of bacteriophage T5 empty capsid lies in the range 26,213,576 to 26,217,020 Daltons, with an average of 26,215,298 Da. Even though experiments with ultra-filtered capsids were easier to perform, the measured masses nevertheless showed discrepancies with the theoretical ones. Whether this was due to the sample itself or to uncertainties in NEMS mass measurement remained undetermined. In previously published studies, uncertainties due to the physical phenomena involved in NEMS mass measurements were evaluated and reported [2]. However, there has been so far no integrated perspective on the various factors influencing the measured mass using nano-resonators from an analytical chemistry point of view. To establish the origin of the observed mass discrepancies, we decided to examine this topic from both fundamental and practical standpoints.

Components	Molecular Formula	Mass per Protein		Subunits	Total Mass	
		Lower Bound	Upper Bound		Lower Bound	Upper Bound
pb8 (Processed major head protein)	C1466H2349N391O455S5	32,889.86	32,894.18	775	25,489,640	25,492,988
pb7 (Processed portal protein)	C1948H3099N523O601S13	43,875.54	43,881.39	12	526,507	526,577
pb11 (Processed head protease)	C290H466N86O87S1	6,580.97	6,581.83	30	197,429	197,455
				Total	26,213,576	26,217,020

Figure IV-9. Calculated masses (Da) for proteins of bacteriophage T5 capsid based on their elemental composition and total mass based on numbers of subunits

IV.6 FACTORS INFLUENCING NANORESONATOR MASS MEASUREMENTS

The factors influencing the mass measurements using NEMS resonators can be sorted into three broad categories:

- Physical principles of the measurement
- Mechanical and geometrical characteristics of the landing particle

- Mechanical properties and geometry of the nano-resonator

While these factors manifest in all nano-resonators, their effects may be more or less pronounced depending on the sensor's design and the system architecture. We will thus focus on our particular implementation of nano-resonators.

IV.6.1 MASS AND POSITION COMPUTATION

As described in Chapter II, nanoresonators used in our NEMS MS are doubly clamped beams oscillating in plane. This design was selected because its symmetry allows the determination of the mass Δm of the landed particles by tracking only two resonance frequencies. The frequency shift that occur due to Δm at position x for the nth mode of vibration can be derived from the following equation:

$$\frac{\Delta f_n}{f} = \frac{\Delta m}{M}\frac{\phi_n^2(x)}{\alpha_n} \qquad (4)$$

Hence, the resonance frequency of NEMS nano-resonator depends not only on its total mass but also on the mass density. As the landed mass alters the mass distribution, we need extra information regarding the frequency shift for more than one mode. Thus, we track the resonance frequency of two modes and solve the following system of equation for the two independent variables Δm and x :

$$\begin{cases} \Delta f_1 \propto \Delta m\, \phi_1^2(x) \\ \Delta f_2 \propto \Delta m\, \phi_2^2(x) \end{cases} \qquad (5)$$

where, Δm is the mass of the landed particle, x its position along the beam, Δf_n = shift in frequency for the i-th mode of oscillation (i={1,2}), ϕ_n the mode shape for the i-th mode with $x\epsilon[0; 0.5]$ due to the symmetry of the beam.

Figure IV-10, represents the resonance frequency shifts as a function of position for both modes when a mass of 1 to 3 MDa lands onto its surface. The graph shows that the frequency response to the same mass of a landed particle strongly depends on position. The dotted line represents the frequency noise level, assumed to be identical for both modes. Below this value, the shift in frequency is not detectable anymore.

Depending on the mode considered, different portions of the beam have low sensitivity to the added mass, precluding the detection of particles in these regions. Precise mass determination can thus be achieved only on portions of the beam where both mode 1 and mode 2 responses are higher than the frequency noise. Taking into account landing positions effects, we can define a reliable sensing interval. In practice, only the regions situated in the interval x=[0.25-0.47] and by symmetry x=[0.53-0.75] along the beam should be taken into account for particle mass determination [9].

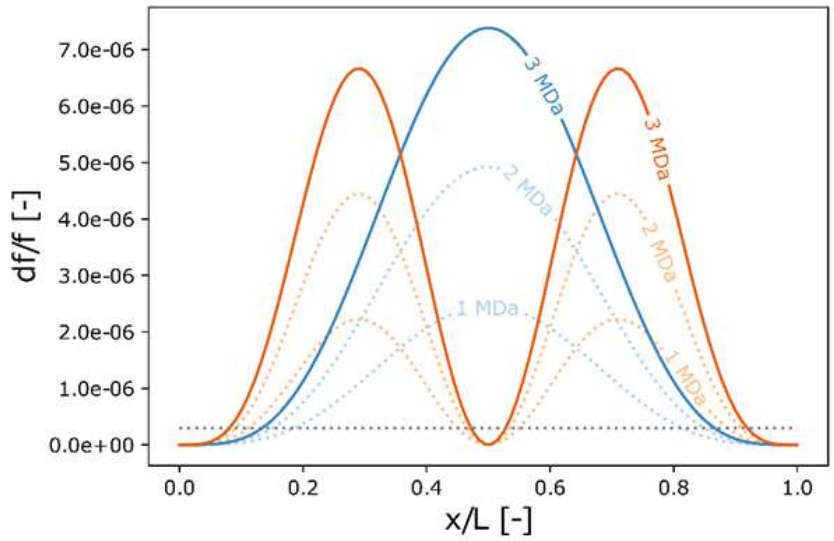


Figure IV-10. Relative frequency shift due to different added masses as a function of the relative position along the resonator beam. Mode 1 is represented in blue and the mode 2 in orange. The grey dashed line represents the level of noise (the typical noise recorded is ~ 3×10^{-7} depending on beam length) [1].

The thickness and width of our doubly clamped beam are 160 nm and 300 nm respectively. Such a beam resonates at frequencies of ~24.6MHz (mode 1) and ~66.8MHz (mode 2). Deleterious effects such as electrode short-circuit, shifts in frequencies and mass errors can occur when a particle larger than the beam width lands on its surface. To avoid this, we choose to analyze spherical or quasi-spherical nanoparticles such as virus particles ranging between 20-200 nm in size [10] [11].

IV.6.2 **PARTICLE STIFFNESS**

Tamayo and coworkers have systematically studied how to modulate the frequency shift caused by molecular adsorption such that either the stiffness effect or the landed particle mass effect can control the frequency shift [12]. This effect appear to be very significant for very thin flexural beams [13] [14]. This effect is complicated to derive experimentally for all device geometries, as it requires conducting similar experiments with particles having the same mass but different Young's modulus. However, it is difficult to obtain nanoparticle with defined stiffness and narrow mass distributions in the same range. Fortunately, this phenomenon has been previously studied in the article by Dominguez et al. [1] using numerical simulations.

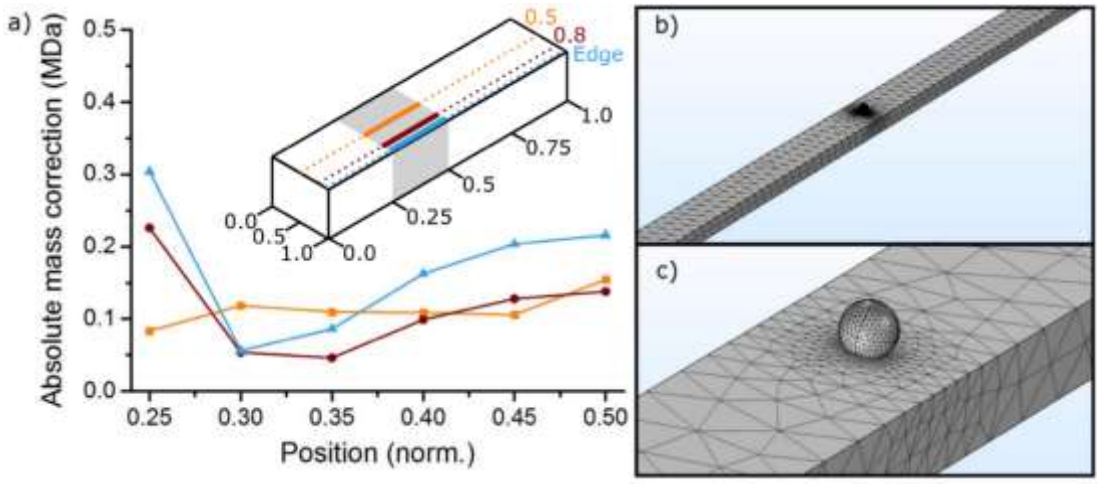


Figure IV-11. Influence of the stiffness of the particle on mass estimation. The mass correction shown in (a) corresponds to the difference between the mass of the particle with and without taking into account the capsid's stiffness. It was computed for different positions along and across the beam. Figures (b) and (c) show the computational domain and its meshing by zooming in the particle area. Adapted from [1]

Figure IV-11 shows the results of a COMSOL Multiphysics finite-element simulation of the device used in our experiment [1]. To simulate the deposition of a bacteriophage T5 capsid on the beam surface, the particle was modelled as a sphere with a 92nm diameter weighing of 106 MDa, and with a disk contact area of 67nm. The stiffness was set to a Young's modulus to 1GPa based on the properties of bacteriophage HK97 capsids, which has been determined experimentally [15]. The mass of the particle was calculated from the resulting frequency shifts with and without taking into account the particle stiffness. Mass discrepancies were computed for the particle landing on favorable positions of the beam, i.e. between 0.25 and 0.5.

The curves shown in Figure IV-11 show the impact of the particle's stiffness on mass measurement as a function of the landing position along and across the axis of the beam. Depending on position, the mass difference averaged 0.1 MDa and could go up to 0.3MDa for particles landing at the extreme of the range considered. These values are well below the absolute mass discrepancy and the full width half-maximum (FWHM) obtained in NEMS MS measurement of T5 capsids. Therefore, while it may contribute by ~0.1 MDa in average, we could safely eliminate the effect of particle stiffness causing the observed mass differences.

IV.6.3 **FREQUENCY NOISE**

Noises affecting NEMS resonance frequencies traces have an important impact on particles mass estimates: by randomly increasing or decreasing frequency shifts on both oscillation modes, they induce a loss of mass accuracy and mass resolution. [16] [17]

The precision attainable for the measurement of the device's resonant frequency changes can be expressed with a statistical tool called the Allan deviation σ_A, which relates to the frequency stability of the device and its electronics [18]. Allan deviation is used in many scientific domains for example atomic clocks, crystal oscillators [19] [20]. Its value depends on the readout scheme and the fluid (viscosity, pressure, temperature) surrounding the resonator. Typically, our devices, operating in high vacuum, display sub-ppm resolution for the frequency shift after particle landing ($10^{-7} < \sigma_A < 10^{-6}$).

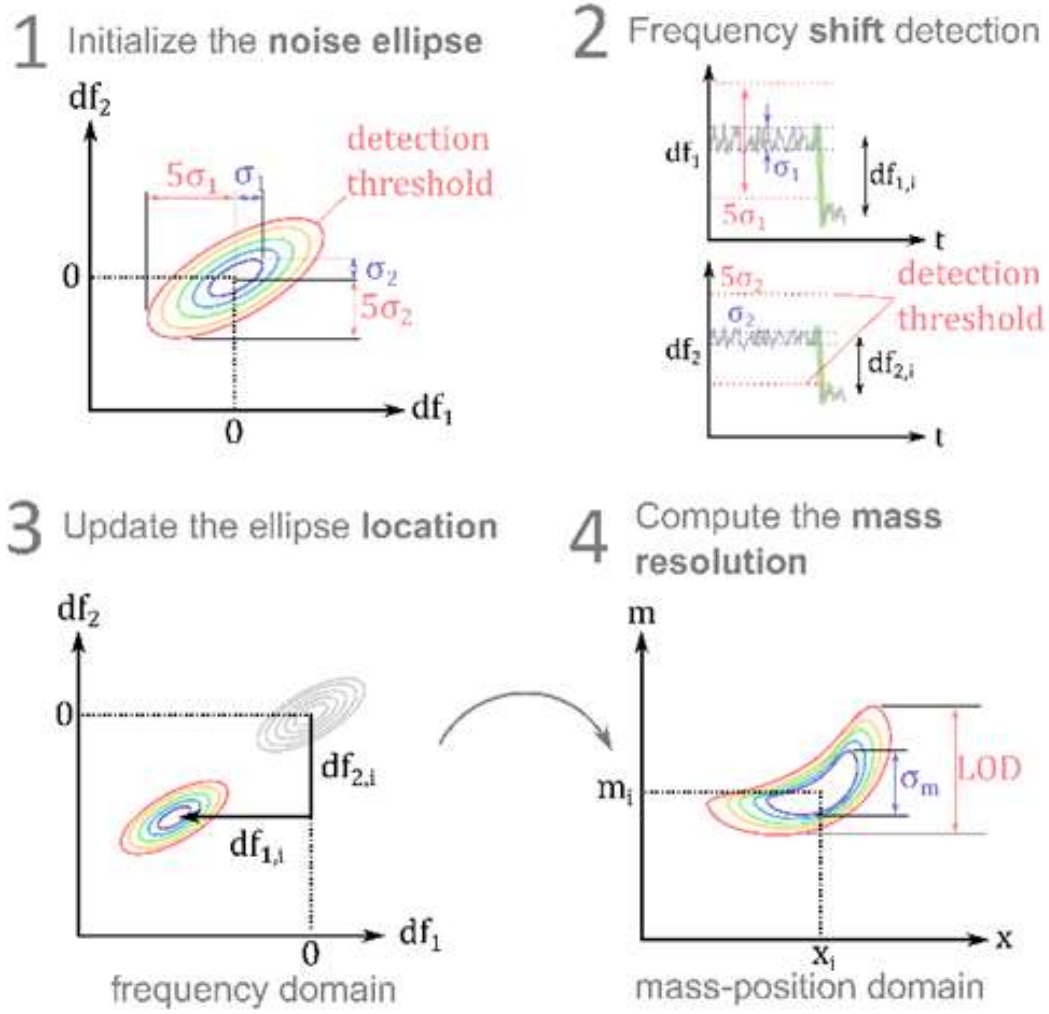


Figure IV-12. Simplified data process workflow to extract the mass resolution and the limit of detection (LOD) from one frequency jump. (1) The noise is characterized by evaluating σ_1 and σ_2, leading to the noise ellipse. (2) The whole frequency trace is swept to detect the shifts corresponding to landing events and the mass resolution is computed by (3) updating the resonance frequency and (4) extracting the extrema values of the noise ellipsoid in the mass-position domain [1].

An added complication stems from the requirement to monitor two vibration modes. In order to discriminate noise from real particle-induced frequency shifts, fluctuations in both resonance frequencies are recorded at the beginning of every experiment, before any particle deposition. Then, a scatter plot of the shifts between consecutive frequency samples for the second oscillation mode vs. the same shifts in the first oscillation mode is drawn: the resulting surface is fitted with a bi-variate Gaussian distribution with standard deviations σ_1 and σ_2 – the standard deviation of both modes 1 and 2 frequency shifts in the absence of particle landing events. We thus encapsulate the points that are considered to be due to noise within an ellipse of parameters $N \times \sigma_n$ (see *Figure* IV-12). N is typically chosen between 3 and 5. This ellipse is called the noise ellipse. At last, we plot the pairwise frequency shifts in both modes obtained during a particle deposition experiment:

- Points that lie inside the previously defined 'noise ellipse' are attributed to noise and are discarded
- Points outside the 'noise ellipse' are attributed to particle landing events and flagged for subsequent estimation of particles masses.

The ellipse defined by $N \times \sigma_n$ can be considered as a limit of detection (LOD), as frequency shifts that lay within the ellipse are not detected as particle landing events. This limit is typically in the range 0.5 – 1 MDa depending on the quality factor of the resonator and the electronic noise. In order to compute it, it is required to transfer the noise ellipse from the frequency shift domain into the mass-position domain. Details about this process can be found in Ref [21].

The next data processing step consists in extracting the magnitude of the simultaneous relative frequency jumps corresponding to each landed particle in both oscillations modes. These values are used to resolve the system of equations for position and mass. Because of the frequency noise, this last step also yields an uncertainty in mass determination. This uncertainty is called the mass resolution σ_m and is defined as:

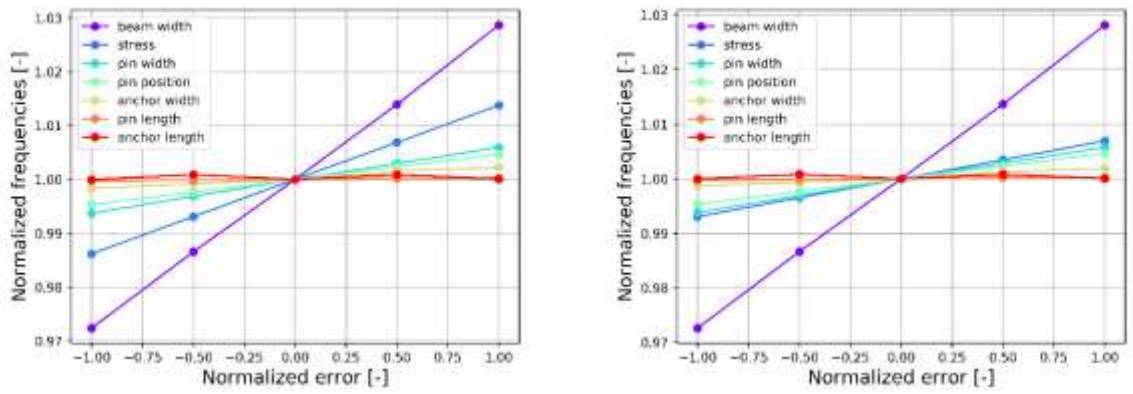
$$\sigma_m = 2M_{eff}\sigma_n \qquad (6)$$

Where σ_n is the standard deviation of the resonator relative frequency (dimensionless number) and M_{eff} is the effective mass of the resonator, defined by $M_{tot}\int_{x=0}^{x=1}\phi(x)^2 dx$. Importantly, Equation (7) is valid for one mode only, and thus does not take into account the landing position, which plays a significant role (see section "*MASS AND POSITION COMPUTATION"*). Therefore, the mass resolution has to be computed for every deposition event. The method used to compute it is similar to the one used for LOD, but the ellipse is defined by $1 \times \sigma_n$. In the active detection area, the typical mass resolution of the devices used in this study is ~0.1 MDa in average.

IV.6.4 RESONATOR MASS AND RESIDUAL STRESS

After many NEMS design iterations, a strong experience have been gathered along the years. This experience can be used in a very tangible way when it comes to the comparison between the theoretical resonance frequencies and the ones observed. Indeed, differences measured over several device generations have lead to the legitimate questions: "what is the origin of such resonance frequency deviation and how does it affect the mass measurement?"

It is well known that the processes used to make nano-sized objects may eventually yield slight variations between the expected and the actual dimensions. In addition, the same processes may induce residual stress over the device. These two phenomena may induce errors in resonance frequency estimates and even larger errors when estimating the total mass of the beam. Such mass beam discrepancies directly leads to an error on the particle mass Δm, as shown by equation (5).



(a) *Mode 1*

(b) *Mode 2*

Figure IV-13. Frequency shifts associated with several NEMS parameters and residual stress. The normalized error represents the deviation from the design, ±1 being the extrema of the studied range. These plots are valid for a 10 µm-long beam, similar results were observed for all beams in the array.

To understand how defects caused during fabrication affect the device's resonance frequency, mechanical simulations were undertaken using COMSOL Multiphysics (see *Figure* IV-13 a). The computational domain was made to fit with the actual dimensions of the devices (*Figure* IV-13 b). The boundary conditions used to simulate the pined-pined beam were defined as fixed constraints for the bottom walls of the anchors. The other walls were defined as free. Finally, the stress was initialized to simulate the residual stress in the structure due to the fabrication processes. Multiple defects were taken into consideration, such as beam width (w), anchor dimensions (L_1, w_1), pins dimensions, shifts in pins positions and residual isotropic stress. However, the device thickness was available through quality control measurements during fabrication and directly taken in consideration in the model. The discrepancies of the dimensions were studied in the range $\pm 10\ nm$. The residual stress was studied in the range $\sigma_{res} \in [-100; 100]\ MPa$ based on feedback from fabrication. The variable that influence the resonance frequency the most were found to be the discrepancy in beam width w and the residual stress σ_{res} as shown in *Figure* IV-13. In the range studied, these effects could be considered linearly dependent of the stress and width parameters.

IV.6.5 **BEAM MASS CALIBRATION**

Based on these simulations, we were able compute a better estimation of the beam mass and thereby correct the particle mass computed from Equation 5. For this, let us assume that the $i-th$ resonance frequency of a given resonator is a function of a global residual stress (σ_global), which is assumed to be isotropic across the whole array, and the individual beam width (w_i) for this resonator. A model for the resonance frequency is shown in Equation 8:

$$f_{model}^{i} = g(\sigma_{global}, w_i) \qquad (7)$$

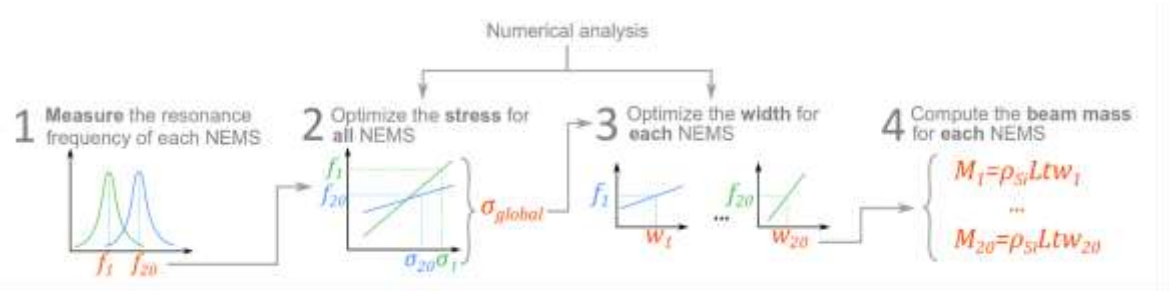


Figure IV-14. Algorithm of the calibration method. The indexes of the variables f, σ, w and M refer to the different beams of the array (only resonators 1 and 20 are displayed for clarity) [1].

To determine the width w_i, a new approach was developed by Bogdan Vysotskyi using the measured resonance frequencies and the results of the numerical analysis as inputs (see *Figure* IV-14). This approach consists in four steps. First, the resonance frequencies of all resonators in the array are measured. Next, the residual stress across all devices is computed through a optimization algorithm. Then, the computed global stress is used in the model to calculate the devices widths. In both optimizations, the error between the observed frequencies and those provided by the model are minimized. Finally, the corrected mass corresponding to each beam are obtained and used in Equation 5 to calculate the mass added.

Thanks to this method, the defects in the fabrication influences the beam width and the residual stress can be corrected prior to particle deposition. To assess the accuracy of these corrections, the beam width of resonators within an array were measured using scanning electron microscope (SEM), yielding a value of approximately 315 nm (*Figure* IV-15). The estimations provided by the beam calibration procedure ranged between 4.3% and 5.6%, with an average of 5 %, corresponding to a width of 315 nm. These values were in good agreement with the SEM measurements.

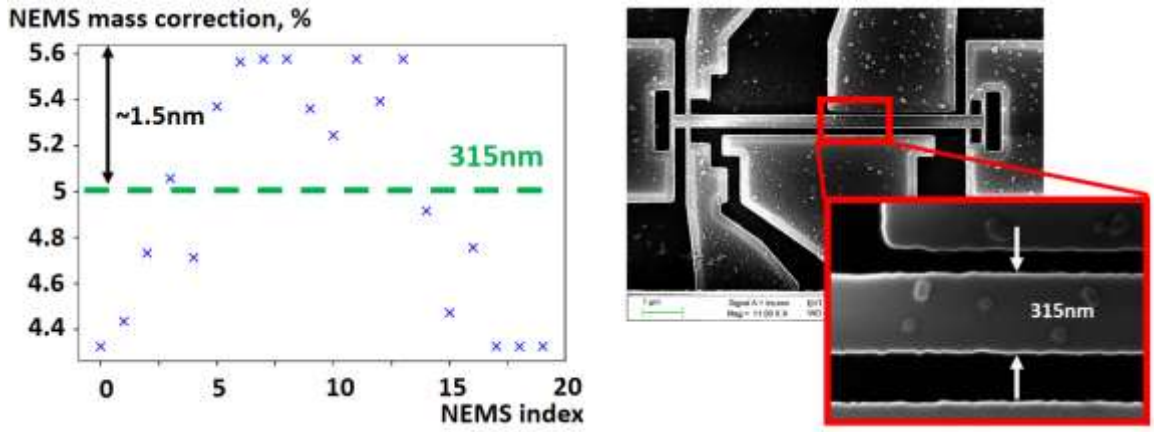


Figure IV-15. Calibration of the beam width through the two processes. Left panel: Comparison between the experimental value (green dashed line) and the calibration results (blue crosses). Right panel : SEM micrograph of a representative beam.

The effect of beam mass calibration on the mass measurement of bacteriophage T5 capsids is shown in *Figure* IV-16. The calibration of the beam mass affected the mass distribution, which increased in average by ~0.9MDa. This correction is larger than the NEMS resolution (typically 0.1 MDa) and should be taken into account. However, its effect is to further increase the discrepancy between the capsid's measured and theoretical mass.

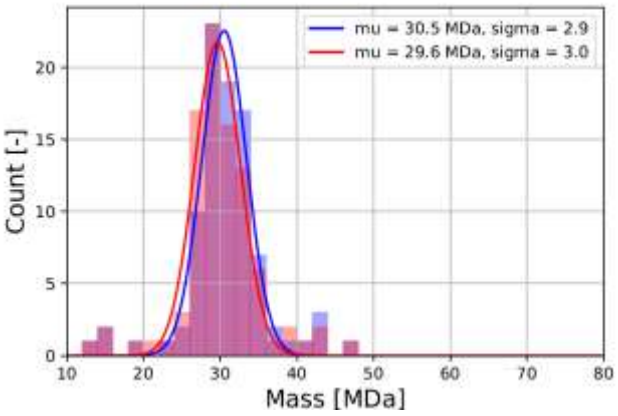


Figure IV-16. Spectrum of the bacteriophage T5 sample using a 20 NEMS array. The red and the blue histogram represents the mass before and after devices stress and width calibration, respectively. The parameters of the Gaussian fit (average mass μ and standard deviation σ) were computed for each distribution.

IV.6.6 PARTICLE SOLVATION EFFECTS

Having examined the parameters affecting mass precision allowed us to provide a conservative estimate of our experiment's overall mass uncertainty: ~1 MDa. Part of this uncertainty is corrected with the proposed beam mass calibration procedure. As a matter of fact, mass accuracy is an essential parameter required to determine if observed mass differences are due to changes in experimental conditions or not. As an example, we will show an investigation into the effect of the nanoESI flow rate on the mass measurement of a viral capsid.

In previously published work [1] Dominguez et al. observed that NEMS mass measurement of empty T5 capsids displayed a bimodal distribution, with a main peak at 27.2MDa, which was in close agreement with the capsid calculated mass, and a second peak at 33.4MDa. They attributed these mass discrepancies to salt aggregation or incomplete particle desolvation. In the field of native mass spectrometry, incomplete desolvation of particle in ESI is a well-known concern [22]. Solvation of a particle can be influenced by multiple factors such as solvent composition, solution feeding flow rate, diameter of the ESI tip, ESI voltage, ambient humidity level, inlet capillary temperature, and pressure in the interface [23].

To investigate this, we studied the effect of temperature, ambient humidity and flow rate on mass measurements. For the study of temperature, we performed experiments at three different temperatures: 150, 175 and 185° C. *Figure* IV-17 shows the effect of temperature on Empty capsid mass measurements. The results shows that at high temperature (185° C), mass is measured all over the axis, which could be due to breakage of capsid particles because of the high temperature [24]. On the contrary, the experiments at low temperature (150° C) had less events at lower mass and showed a clear peak at 30.8MDa the expected mass range. At 175° C, we measured a mass of 30.9MDa and obtained an event rate of ~203 events/acqusitions including low mass.

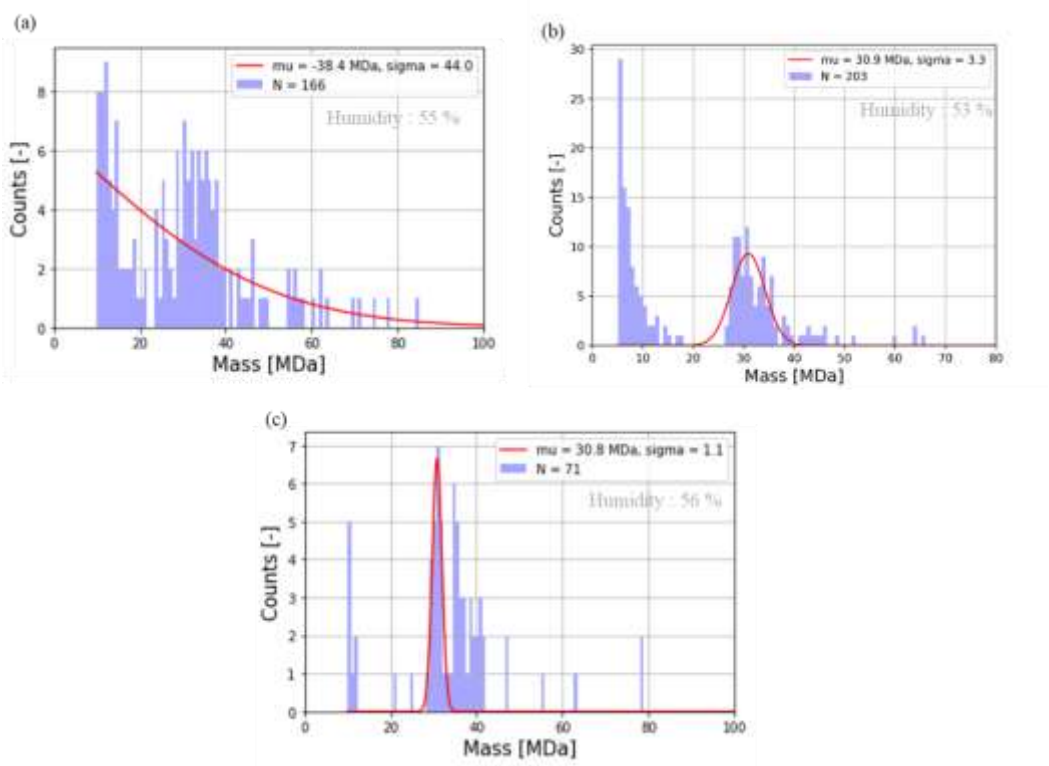


Figure IV-17. Shows the effect of temperature on NEMS-MS mass measurements. (a), (b) and (c) shows one complete run of experiments conducted at 185, 175 and 150 °C respectively.

To investigate the effect of humidity, the humidity in the laboratory was recorded each day before the measurement, using a digital hygrometer. The humidity was arbitrarily labelled as "low" or "high" when the relative humidity was below or above 45% respectively. *Figure* IV-18 shows the average vs. the standard deviation of the measured mass for the two humidity levels. Higher humidity level corresponded to a higher average mass albeit with a lower standard deviation. Importantly, each point may correspond to different temperature and flowrate conditions, and the humidity was not controlled, so these results are barely qualitative. One hypothesis was that higher humidity decreased the efficiency of the particles' desolvation in the electrospray interface, affecting the accuracy of the NEMS-MS measurement [25].

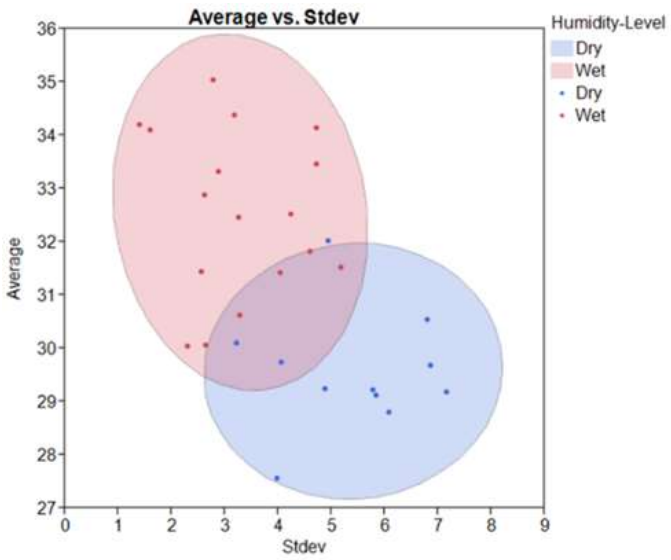


Figure IV-18. Effect of relative humidity on mass measurements : Average mass v/s standard deviation at low & high humidity level.

To have a clearer understanding on the effects of particle desolvation on NEMS-MS mass measurements, we needed to have one entire experiment with all parameter, under control, which was complicated due to the changing environmental conditions (humidity), since the entry of the system is in open air. Thus, we chose to investigate the effect of flow rate on the mass measurement under fixed temperature and humidity conditions. Contrary to the other influential parameters identified, the flow rate can be readily controlled and does not entail long stabilization times between experiments.

Time interval (sec)	**Flow rate (µl/hr)**	**ESI voltage (kv)**	**Number of events**	**Average mass (MDa)**
0-840	10	2.2	30	30.98
840-1680	20	2.5	22	30.87
1680-2520	30	3	19	36.29
2520-3360	40	3.2	28	35.32
3360-4200	50	3.4	10	38.73

Table IV.3. Experimental parameters and data statistics for varying flowrates and ESI voltage.

A controlled experiment was designed in which the sample was electrosprayed at increasing flow rates over the course of a single acquisition under constant ambient humidity (53%) and inlet temperature conditions (175°C). The flow rate was varied from 10 µl/hr to 50 µl/hr by 10 µl/hr increments in every 14 minutes and the ESI voltage was adjusted on the fly to compensate for spray instabilities. *Table* IV.3 shows the time range, ESI flowrate, the number of particles landed and its average mass for the corresponding time interval. Importantly, the spray voltage was adjusted to yield stable conditions for each flow rate value. It is therefore impossible to disentangle the voltage from the flow rate and we will refer to these two variables in the next session as “spray parameters”.

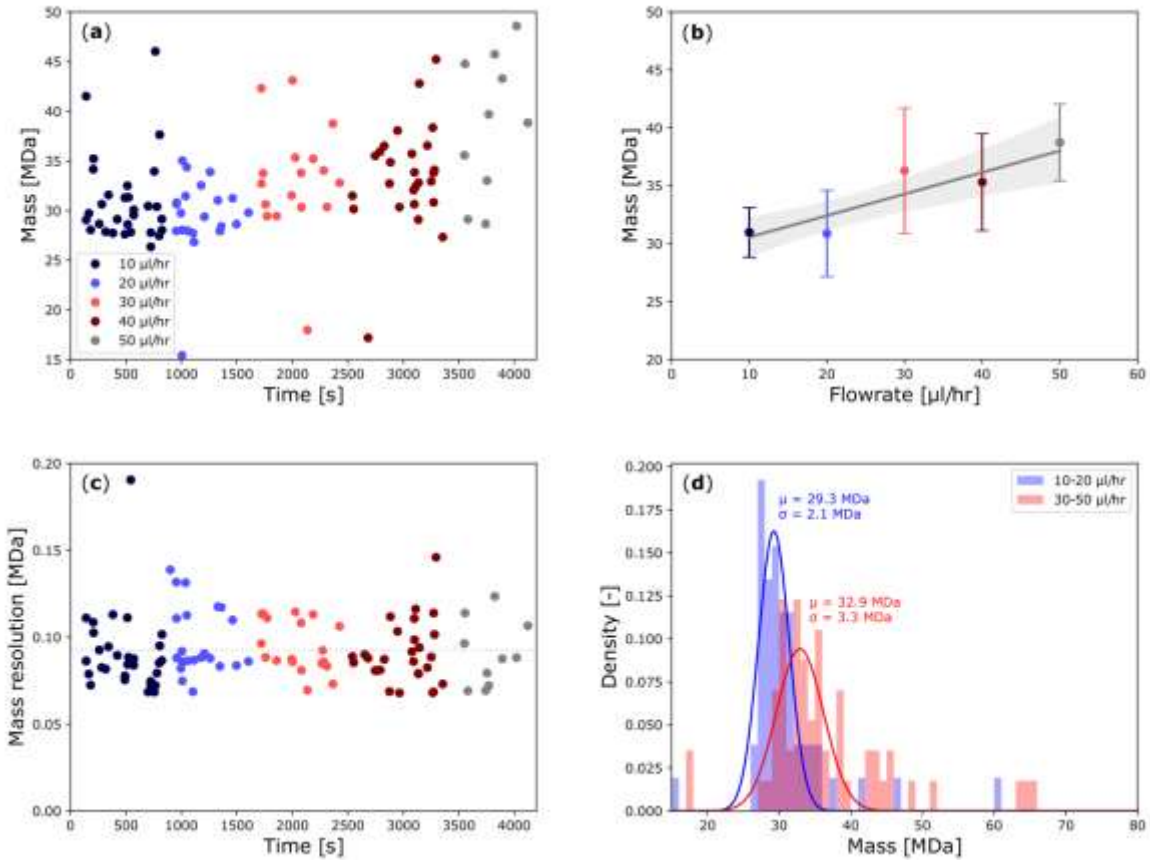


Figure IV-19. Investigation of the nanoESI flow rate and voltage effects. (a) Mass of individual particles landing on the resonator over time as the flowrate was increased. (b) Average mass vs. flowrate with linear trend. The data are fitted by the equation $0.18 \times x + 28.5$. *(c) Mass resolution for individual particle landing events over time. The grey dashed line corresponds to the average mass resolution of 93 kDa. (d) Composite spectrum at higher (30-40 µl/hr) and lower (10-20 µl/hr) flow rates showing a mass shift of several MDa [1].*

From *Figure* IV-19 a displays the mass of the individual particles detected in the mass range of interest over time, color-coded for spray parameters. We could observe a gradual increase in the measured mass as the flow rate and voltage were increased. *Figure* IV-19 b illustrates this trend more clearly by plotting the mass statistics vs. flow rate. The linear fit in this plot is mainly to guide the eye; actually, the points could be grouped into two conditions with respect to the flow rate. Considering the flow rates ≤ 20 µl/hr as low flow rate and > 20 µl/hr as high flow rate, we could say that, the mass at high flow rate was 3.4MDa heavier compared to that at low flow rate (*Figure* IV-19 d). This increase was in large excess to the resolution of each individual mass measurement (93 kDa) which appeared completely unaffected by changes in flow rate (*Figure* IV-19 c). Moreover, it was larger than our estimated uncertainties and therefore could be directly linked to changes in the flow rate parameter.

It seems to be reasonable to infer that the observed mass shifts were linked to incomplete desolvation, leading to residual water molecules inside or around the capsids. We attempted to make calculations to estimate the thickness of an equivalent mass of water on the outer or inner capsid surface. These calculations rely on the simplified geometric consideration, where the excess water is shaped into a spherical shell (*Figure* IV-20a and b).

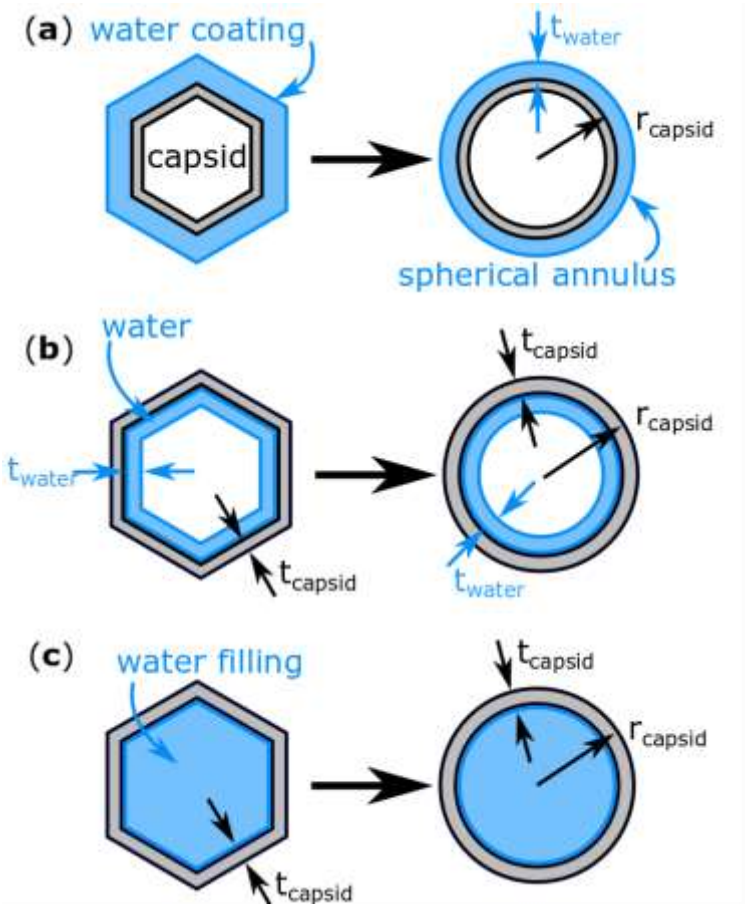


Figure IV-20. Sketch describing the geometric models used to compute the water layer thickness [1].

Based on the considerations outlined in *Figure* IV-20a, the thickness of the water coating t_{water} on the outside of the capsids can be computed as follows:

$$t_{water} = \left(\frac{3m_{H2O}}{4\pi\rho_{H2O}} + r_{capsid}^3\right)^{\frac{1}{3}} - r_{capsid} \qquad (8)$$

Where r_{capsid} is the outer diameter of the capsid, m_{H2O} is the total mass of water – estimated as the difference between the theoretical mass m_{th} and the measured average mass m_{exp} of capsids– and ρ_{H2O} is the water density.

If we consider that the water layer lies inside the capsid (*Figure* IV-20 b), one can compute its thickness t_{water} by using the formula:

$$t_{water} = r_{capsid} - t_{capsid} - \left((r_{capsid} - t_{capsid})^3 - \frac{3}{4\pi}\frac{m_{H2O}}{\rho_{H2O}}\right)^{\frac{1}{3}} \qquad (9)$$

Where t_{capsid} is the thickness of T5 capsids. Results of these calculations are summarized in Table IV.4.

		Low flowrate/voltage		High flowrate/voltage		Unit
mexp		29.3		32.9		MDa
tcapsid		1.5	4.5	1.5	4.5	nm
twater	in	0.22	0.253	0.464	0.535	nm
	out	0.204		0.426		nm

Table IV.4. Calculated water thickness

The parameters used in these calculation are summarized in *Table* IV.5.

Symbol	**Value**	**Unit**
r_{in}	46	nm
m_{th}	26.018	MDa
m_{exp}	29.3 or 32.9	MDa
ρ_{H20}	1000	kg/m^3
t_{capsid}	1.5 or 4.5	nm

Table IV.5. Parameters used to compute the water layer thickness.

It is important to place the observed mass differences in context by comparing them with the mass of water that would occupy the whole inner capsid volume. To estimate the mass of water needed to entirely fill the capsid, we computed the inner volume, taking into account the thickness of the capsid t_{capsid} (*Figure* IV-20c). The following equation was used :

$$m_{H2O,in} = \frac{4}{3}\pi\left(r_{capsid} - t_{capsid}\right)^3 \rho_{H2O} \qquad (10)$$

Thus, the mass of water that would fill the entire capsid could be estimated to lie somewhere between 222 MDa and 180 MDa depending on the capsid thickness [26].

Therefore, considering a 92 nm diameter capsid having a theoretical mass of 26.018 MDa, one could estimate the thickness of the water layer present on the inner or outer surface of capsids detected at low flow rate (29.3 MDa) and high flow rate (32.9 MDa) to be approximatively ~0.2 nm and ~0.5 nm respectively. While other explanations for the added mass could be invoked, such as residual salts or proteins inside the capsid, it is interesting to note that these values are on the same order of magnitude as the thickness of one or two water monolayer (0.25 ± 0.05 nm) [27] suggesting solvation layers on both capsid surfaces. Importantly, a capsid full of solvent should have a mass far in excess of the measured one, thus, almost all this water is removed during the electrospray process or the transfer through the interface. This in turn raises questions related to the exact capsid content in solution, and to the mechanism by which the water is expelled or evaporated during the electrospray process. Unfortunately, at this stage, we do not have enough information to answer these questions, but the question of removal of solvent is a critical issue to consider for future experiments.

4.7 CONCLUSION

A new approach to analyze large virus like particles basis using NEMS MS system was introduced in this chapter. Herein, the capsid sample was desalted using ultra centrifugation and nebulized using ESI technique resulting in efficient transfer into the gas phase. The use of aerodynamic focusing lens enabled the capsid particles to reach the resonator efficiently. The NEMS resonator was able to directly detect the mass of EC individually. Upon analysis of the results we observed a ~3 MDa discrepancy with the capsid calculated mass.

As we pointed out, the mass determination by NEMS can be affected by multiple parameters. While some of them are inherent to the particles themselves (landing position, size, stiffness, etc.), other relate to the resonator (frequency noise, fabrication defects). In this chapter, these two categories of phenomena have been investigated. This work was approved and validated in the journal Analytical and Bioanalytical Chemistry [1].

We clarified how the landing position dramatically degrades the mass resolution close to the anchors. Therefore, an increase of the range of landing position (i.e. an increase in terms of number of detected events) is at the cost of a degraded accuracy. From the experimentalist point of view, in specific conditions it might useful to tune these parameters according to the needs of the study [28].

Moreover, the impact of the T5 capsids stiffness was simulated assuming that it was close to HK97 capsids one. In this framework, it seems that this effect cannot be entirely neglected because its influence can scale up to 0.3 MDa in some cases, which is in the same order of magnitude than the mass resolution of the NEMS (0.1 MDa). This effect was first mentioned by Tamayo et al. for micro-cantilevers and the present study quantify its impact on mass measurement accuracy in the case of capsid-like particle adsorbed on the surface of pinned-pinned beams [14].

The beam and the anchors of the resonators are 300 nm and 80 nm wide respectively. Because nano-systems require complex fabrication processes, it may induce defects leading to errors when performing the frequency-to-mass conversion. The influence of these discrepancies was evaluated by means of a simulation and the results were used to propose an algorithm to correct it. This method suggests a correction of 0.9 MDa for bacteriophage T5, which is higher than the mass resolution of the NEMS. Moreover, it is important to mention that this error scales with the mass of the adsorbed particle, according to Equation (1).

Considering the bias due to the noise (0.1 MDa) and the influence of the particle stiffness (0.3 MDa), it is reasonable to assume that the uncertainty of the measurement is below 1 MDa. Quantifying this limit allowed us to investigate phenomena affecting the mass measurement within this order of magnitude, such as the influence of the ESI feeding flow rate. Mass measurements performed in the range from 10 to 50 µl/hr show the influence of this parameter over the average mass, which is increased by more than 3 MDa for highest flow rates (30 – 40 µl/hr).

This work shows what are the main parameters affecting mass measurement with NEMS-MS technique and proposes solutions to improve its accuracy. Moreover, the knowledge of the strength and limitations of the NEMS-MS technique is essential for further development of the current set-up. For instance, knowing the negative impact of the particle stiffness on the measurement accuracy, it would be relevant to use different type of sensors geometries.

Optomechanical devices have been recently described, that are neither sensitive to the particle size and aspect ratio, nor to the particle landing position [29]. Using such devices would definitely improve the mass accuracy of NEMS-MS measurements by circumventing the requirement to monitor multiple modes and the need to discard events occurring on unfavorable portions of the device.

REFERENCES

1. Clement K, Reynaud A, Defoort M, Vysotskyi B, Fortin T, Lai S-H, Çumaku V, Dominguez-Medina S, Hentz S, Masselon C (2021) Requirements and attributes of nano-resonator mass spectrometry for the analysis of intact viral particles. Analytical and Bioanalytical Chemistry. https://doi.org/10.1007/s00216-021-03511-4

2. Dominguez-Medina S, Fostner S, Defoort M, Sansa M, Stark A-K, Halim MA, Vernhes E, Gély M, Jourdan G, Alava T, Boulanger P, Masselon CD, Hentz S (2018) Neutral mass spectrometry of virus capsids above 100 megadaltons with nanomechanical resonators. Science 362:918–922 . https://doi.org/10.1126/science.aat6457

3. Huet A, Conway JF, Letellier L, Boulanger P (2010) In Vitro Assembly of the T=13 Procapsid of Bacteriophage T5 with Its Scaffolding Domain. J Virol 84:9350–9358 . https://doi.org/10.1128/JVI.00942-10

4. Phillips AT, Signs MW (2004) Desalting, Concentration, and Buffer Exchange by Dialysis and Ultrafiltration. Current Protocols in Protein Science 38:4.4.1-4.4.15 . https://doi.org/10.1002/0471140864.ps0404s38

5. Chen LC, Naito T, Ninomiya S, Hiraoka K (2018) Hyphenation of high-temperature liquid chromatography with high-pressure electrospray ionization for subcritical water LC-ESI-MS. Analyst 143:5552–5558 . https://doi.org/10.1039/C8AN01113C

6. Atomic Weights and Isotopic Compositions with Relative Atomic Masses. In: NIST. https://www.nist.gov/pml/atomic-weights-and-isotopic-compositions-relative-atomic-masses

7. Berglund M, Wieser ME (2011) Isotopic compositions of the elements 2009 (IUPAC Technical Report). Pure and Applied Chemistry 83:397–410 . https://doi.org/10.1351/PAC-REP-10-06-02

8. Huet A, Conway JF, Letellier L, Boulanger P (2010) In Vitro Assembly of the T=13 Procapsid of Bacteriophage T5 with Its Scaffolding Domain. J Virol 84:9350–9358 . https://doi.org/10.1128/JVI.00942-10

9. Hanay MS, Kelber S, Naik AK, Chi D, Hentz S, Bullard EC, Colinet E, Duraffourg L, Roukes ML (2012) Single-protein nanomechanical mass spectrometry in real time. Nature Nanotech 7:602–608 . https://doi.org/10.1038/nnano.2012.119

10. Louten J (2016) Virus Structure and Classification. Essential Human Virology 19–29 . https://doi.org/10.1016/B978-0-12-800947-5.00002-8

11. Baker TS, Olson NH, Fuller SD (1999) Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. Microbiol Mol Biol Rev 63:862–922, table of contents

12. Ruz JJ, Tamayo J, Pini V, Kosaka PM, Calleja M (2014) Physics of Nanomechanical Spectrometry of Viruses. Scientific Reports 4:6051 . https://doi.org/10.1038/srep06051

13. Malvar O, Ruz JJ, Kosaka PM, Domínguez CM, Gil-Santos E, Calleja M, Tamayo J (2016) Mass and stiffness spectrometry of nanoparticles and whole intact bacteria by multimode

nanomechanical resonators. Nature Communications 7:13452 . https://doi.org/10.1038/ncomms13452

14. Tamayo J, Ramos D, Mertens J, Calleja M (2006) Effect of the adsorbate stiffness on the resonance response of microcantilever sensors. Appl Phys Lett 89:224104 . https://doi.org/10.1063/1.2388925

15. Roos WH, Gertsman I, May ER, Brooks CL, Johnson JE, Wuite GJL (2012) Mechanics of bacteriophage maturation. PNAS 109:2342–2347

16. Sansa M, Sage E, Bullard EC, Gély M, Alava T, Colinet E, Naik AK, Villanueva LG, Duraffourg L, Roukes ML, Jourdan G, Hentz S (2016) Frequency fluctuations in silicon nanoresonators. Nat Nanotechnol 11:552–558 . https://doi.org/10.1038/nnano.2016.19

17. Cleland AN, Roukes ML (2002) Noise processes in nanomechanical resonators. Journal of Applied Physics 92:2758–2769 . https://doi.org/10.1063/1.1499745

18. Statistics of atomic frequency standards | IEEE Journals & Magazine | IEEE Xplore. https://ieeexplore.ieee.org/document/1446564

19. Chabuda K, Leroux ID, Demkowicz-Dobrzański R (2016) The quantum Allan variance. New J Phys 18:083035 . https://doi.org/10.1088/1367-2630/18/8/083035

20. Fsup S, Fsu S, Fsq S, Fsp S, Fsv S, Fsw S Time Domain Oscillator Stability Measurement Allan variance. 16

21. Hanay MS, Kelber S, Naik AK, Chi D, Hentz S, Bullard EC, Colinet E, Duraffourg L, Roukes ML (2012) Single-protein nanomechanical mass spectrometry in real time. Nature Nanotechnology 7:602–608 . https://doi.org/10.1038/nnano.2012.119

22. Kuprowski MC, Boys BL, Konermann L (2007) Analysis of protein mixtures by electrospray mass spectrometry: Effects of conformation and desolvation behavior on the signal intensities of hemoglobin subunits. J Am Soc Mass Spectrom 18:1279–1285 . https://doi.org/10.1016/j.jasms.2007.04.002

23. Bernier L, Pinfold H, Pauly M, Rauschenbach S, Reiss J (2018) Gas Flow and Ion Transfer in Heated ESI Capillary Interfaces. J Am Soc Mass Spectrom 29:761–773 . https://doi.org/10.1007/s13361-018-1895-0

24. Soleilhac A, Dagany X, Dugourd P, Girod M, Antoine R (2015) Correlating droplet size with temperature changes in electrospray source by optical methods. Analytical Chemistry 87:8210–8217 . https://doi.org/10.1021/acs.analchem.5b00976

25. Newsome GA, Ackerman LK, Johnson KJ (2016) Humidity Effects on Fragmentation in Plasma-Based Ambient Ionization Sources. J Am Soc Mass Spectrom 27:135–143 . https://doi.org/10.1007/s13361-015-1259-y

26. Lošdorfer Božič A, Šiber A, Podgornik R (2013) Statistical analysis of sizes and shapes of virus capsids and their resulting elastic properties. J Biol Phys 39:215–228 . https://doi.org/10.1007/s10867-013-9302-3

27. Opitz A, Scherge M, Ahmed SI-U, Schaefer JA (2007) A comparative investigation of thickness measurements of ultra-thin water films by scanning probe techniques. Journal of Applied Physics 101:064310 . https://doi.org/10.1063/1.2712155

28. Ilic B, Yang Y, Craighead HG (2004) Virus detection using nanoelectromechanical devices. Appl Phys Lett 85:2604–2606 . https://doi.org/10.1063/1.1794378

29. Sansa M, Defoort M, Brenac A, Hermouet M, Banniard L, Fafin A, Gely M, Masselon C, Favero I, Jourdan G, Hentz S (2020) Optomechanical mass spectrometry. Nat Commun 11:3781 . https://doi.org/10.1038/s41467-020-17592-9

GENERAL CONCLUSION

Since the development of the first vaccine for smallpox by Edward Jenner in 1796, huge advances have accumulated in the domain of pandemic disease vaccinology. Even though there is many achievements in this field, there is still great deal to be done. The field has modified in the traditional approach of vaccine development i.e. culturing the organism, attenuation or killing it to the production of target and safer vaccines using modern bioengineering [1]. However, these methods are less immunogenic and less defensive than the killed or attenuated vaccines.

With the development of recombinant protein vaccine technology, efficient and safer vaccines that are less reactogenic were produced [2]. Recombinant shingles is a good example for this trend . However this method require a potential adjuvant or efficient delivery mechanism. Even though lipid-based adjuvants [3] [4] is enormously great, Virus-like-particles (VLP) act as a desired platform for all recombinant antigen, providing safer and better technology with less cost of production [5]. In spite of the fact that virus, bacteria and other synthetic nanoparticles can because its large size, structural stability, biodegradability, surface proteins and low cost production.

In the present work, we studied bacteriophage T5 capsids a model for VLP. Phage T5 is a best option to use as a VLP platform for reasons such as their unique self-assembly process, its ability to bind with 120 copies of so called decoration protein that can be engineered at will, its stability to harsh environmental conditions and its large size. Our collaborator Dr. Pascale Boulanger from I2BC, observed that purified T5 capsids expanded *in-vitro* share similar geometry, stability and affinity to decoration proteins as capsids of native virus. These structures are thus ideal antigen presenting platform for vaccine development because they are high stability, DNA-free and bio-compactable nanoparticle capable of exposing 120 copies of customizable protein.

The study represented in this thesis, was aimed at providing an improved understanding on the structural knowledge of bacteriophage T5 capsids. The study was performed using two Mass Spectrometry (MS) tools, which provided inform on two different aspect. From the collective knowledge obtained from these data, a clearer understanding on the T5 capsid structural protein and a possible characterization method to have routine analysis was derived.

On this framework of the study we started by focusing on bacteriophages as an alternative for vaccine therapy. The study gave us information regarding the assembly of *Caudovirales* capsids. From which we understood that there is only a few MS based studies to investigate the bacteriophage assembly. This motivated us to have an analytical study on T5 phage capsid using MS. The tools used in this framework are; Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) and Nano Electro Mechanical Sensor based Mass Spectrometry (NEMS-MS).

Using LC-MS/MS based proteomic approaches we investigated the presence and copy number of protease protein, which is one of the five structural proteins of phage T5. For this purpose, we have developed a novel isotopically labelled QconCAT dipeptide based targeted proteomic approach. Using this dipeptide standard, we estimated the amount of pb11 in mature T5 capsid. We concluded that the protease protein present in phage T5 capsid does not only contribute to capsid morphogenesis unlike other phages like phage T4 but also have a structural function [6].

Information about the presence and stoichiometry of protease protein stands as a critical point in the development of VLP structures.

Before such VLPs reach the pharmaceutical industry a thorough check on their quality control of different batches to check the purity in preparation, homogeneity, their size and molecular weight is of high importance. Unfortunately, no commercial MS system operates in the mega- to giga-Dalton mass range required for VLP analysis.

Nano-electro-mechanical sensor (NEMS) mass spectrometry (MS) is emerging as a practical option to analyze species such as nanoparticles or viruses, in a mass range beyond the reach of commercial MS instruments [7]. It represents a new paradigm in the field, being the only technique that does not depend on mass to charge ratio. In addition, it features mass-independent resolution, which makes it even more valuable as mass increases. Previous study using NEMS-MS system demonstrated the analysis of bacteriophage T5 capsids.

In this thesis we have focused on empty T5 capsids (VLP form) to develop a routine analysis method for the determining the mass of virus and which will be helpful to ensure the quality of different batches of virus production. Herein we also discuss about the various phenomena influencing NEMS-MS mass estimates including particle's extraneous physical properties (size, aspect ratio, stiffness), and the influence of frequency noise and device fabrication defects. These factors being accounted for, we could begin to notice subtler effects linked with particle desolvation prior to mass measurements.

Overall, this thesis contributed to advancing out knowledge of the capsid assembly mechanism and demonstrated the use of NEMS-MS for capsid structure assessment.

REFERENCES

1. Lisa A. Raedler P (2018) Shingrix (Zoster Vaccine Recombinant) a New Vaccine Approved for Herpes Zoster Prevention in Older Adults

2. Nascimento IP, Leite LCC (2012) Recombinant vaccines and the development of new vaccine strategies. Braz J Med Biol Res 45:1102–1111 . https://doi.org/10.1590/S0100-879X2012007500142

3. Skwarczynski M, Toth I (2011) Lipid-Core-Peptide System for Self-Adjuvanting Synthetic Vaccine Delivery. In: Mark SS (ed) Bioconjugation Protocols: Strategies and Methods. Humana Press, Totowa, NJ, pp 297–308

4. Brown LE, Jackson DC (2005) Lipid-based self-adjuvanting vaccines. Curr Drug Deliv 2:383–393 . https://doi.org/10.2174/156720105774370258

5. Chroboczek J, Szurgot I, Szolajska E (2014) Virus-like particles as vaccine. Acta Biochim Pol 61: . https://doi.org/10.18388/abp.2014_1875

6. Yap ML, Rossmann MG (2014) Structure and function of bacteriophage T4. Future Microbiol 9:1319–1327 . https://doi.org/10.2217/fmb.14.91

7. Gupta A, Akin D, Bashir R (2004) Single virus particle mass detection using microresonators with nanoscale thickness. Appl Phys Lett 84:1976–1978 . https://doi.org/10.1063/1.1667011

APPENDIX I: PROTEOMICS

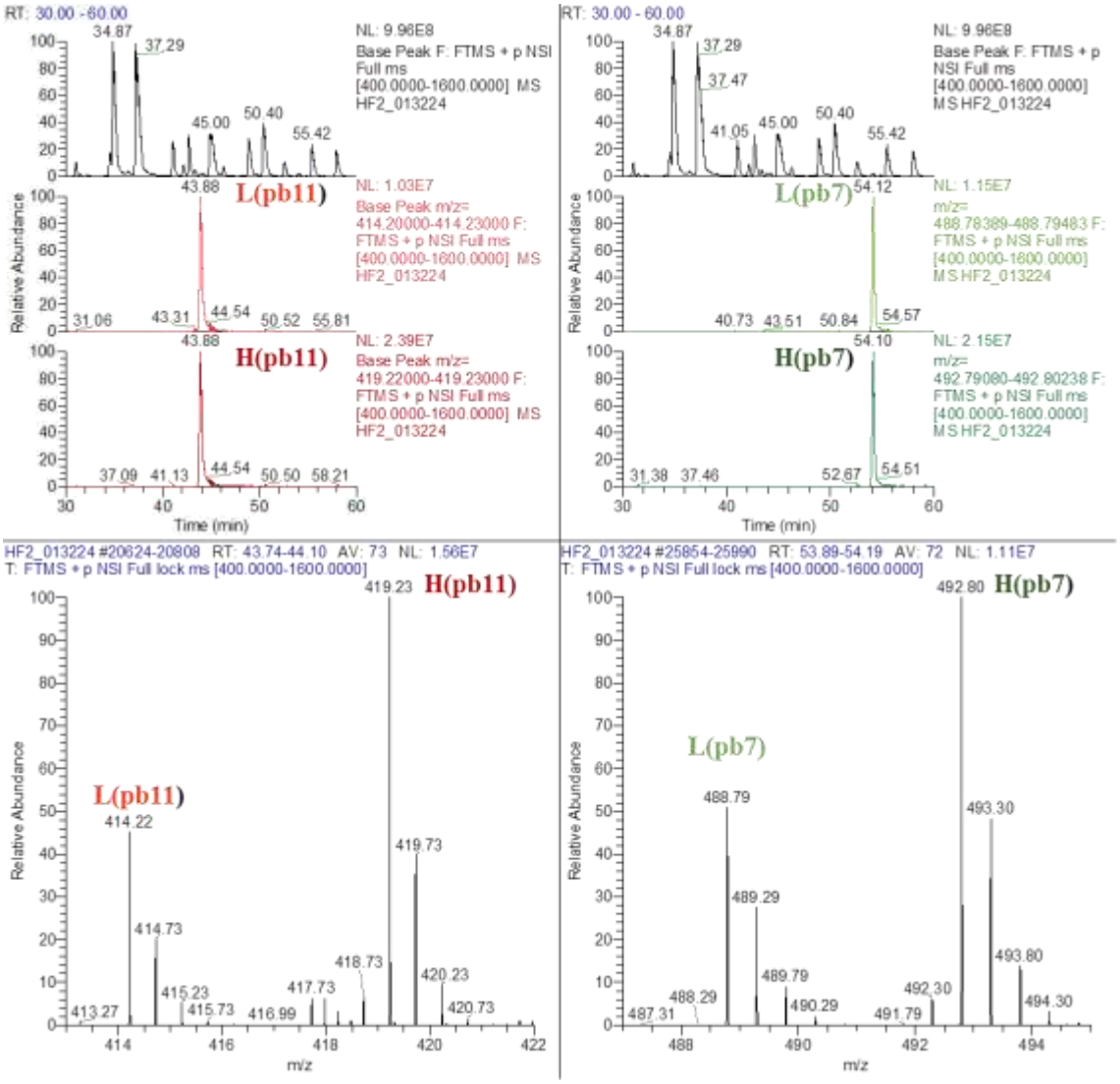


Figure 0-1. Observed base peak chromatograms, and extracted chromatograms of DFC repl with an injection volume 1µl in Table III 9.

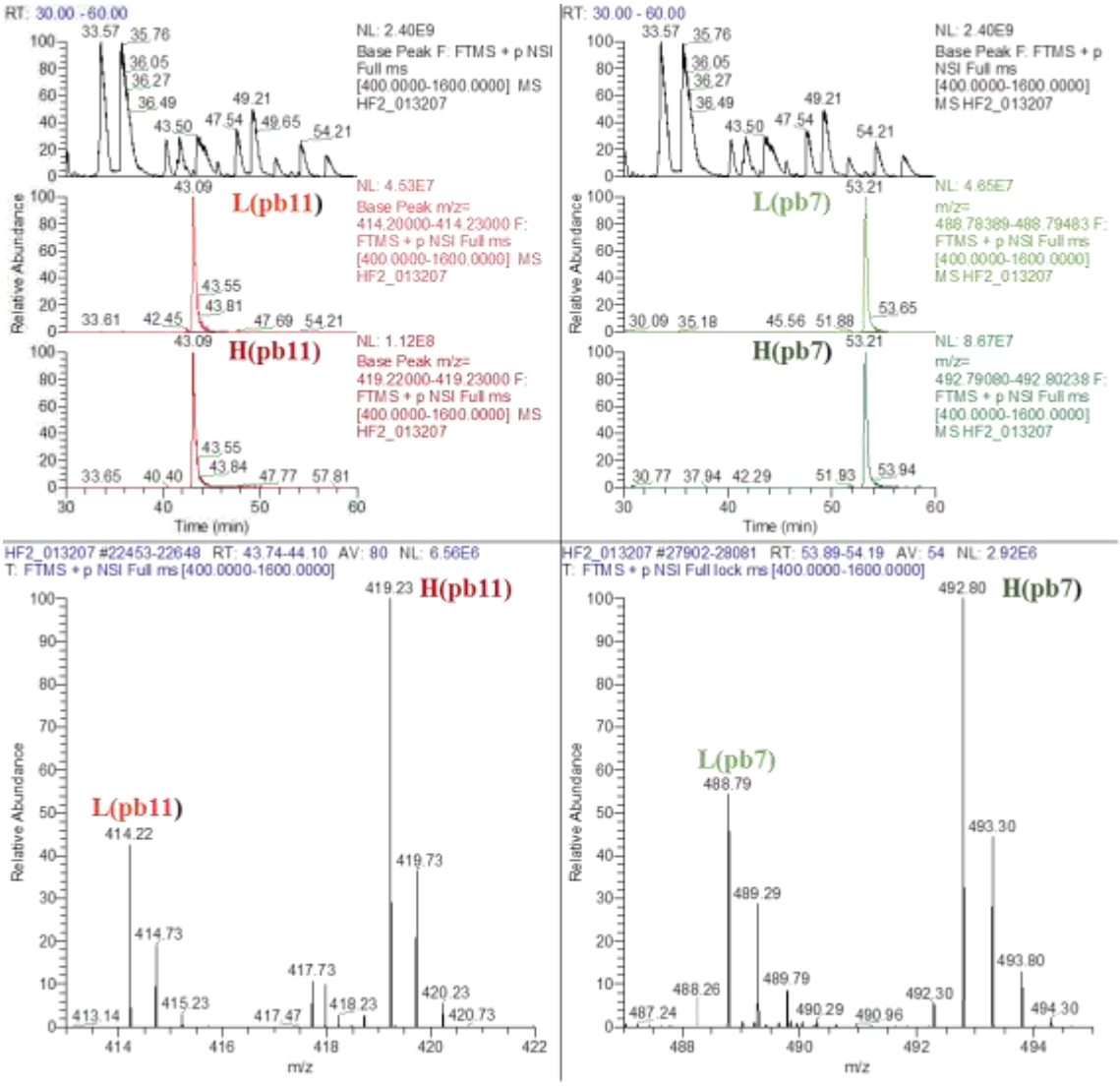


Figure 0-2. Observed base peak chromatograms, and extracted chromatograms of DFC sample with an injection volume 5µl in Table III.9

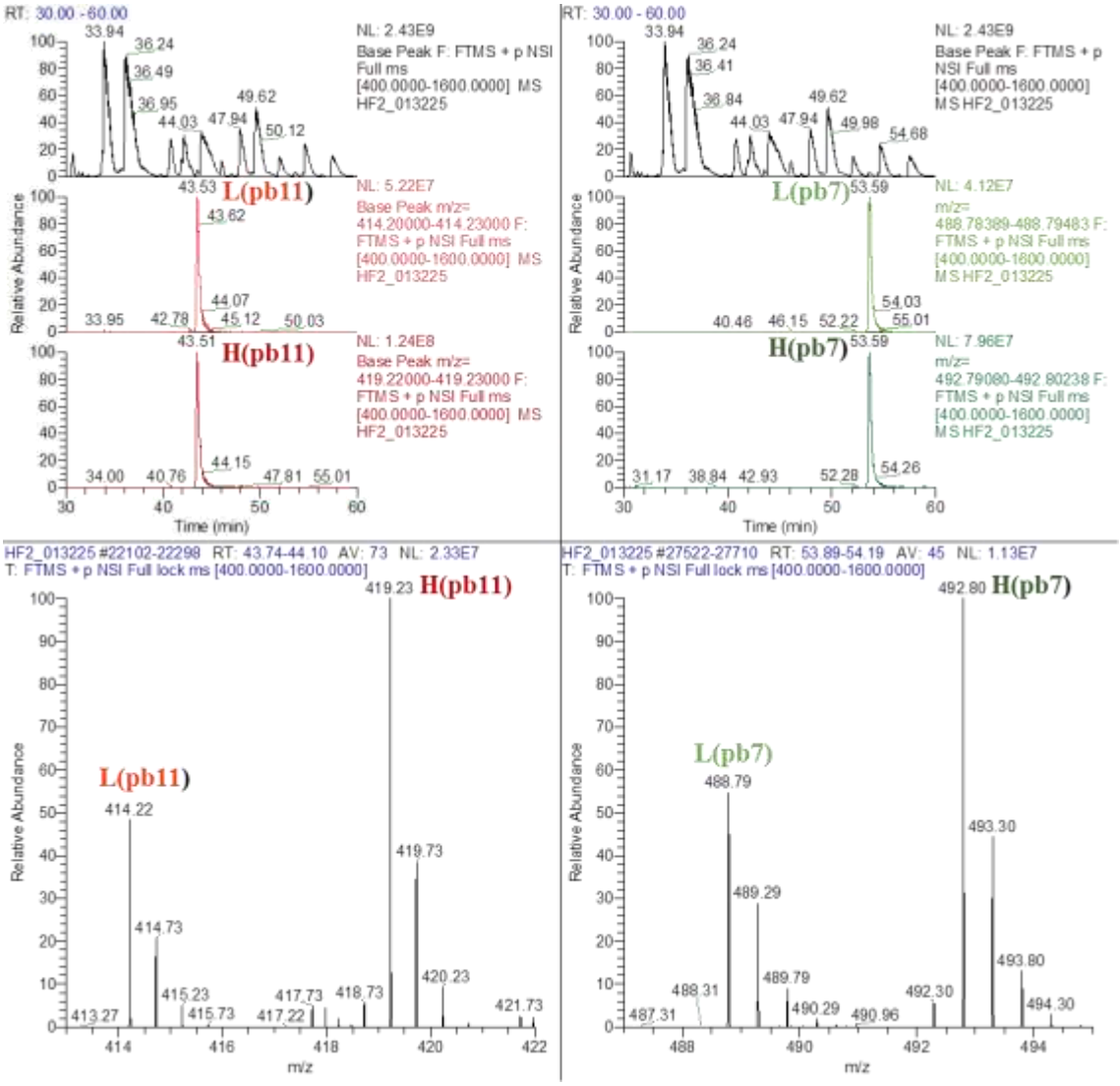


Figure 0-3. Observed base peak chromatograms, and extracted chromatograms of DFC repl sample with an injection volume 5µl in Table III. 9.

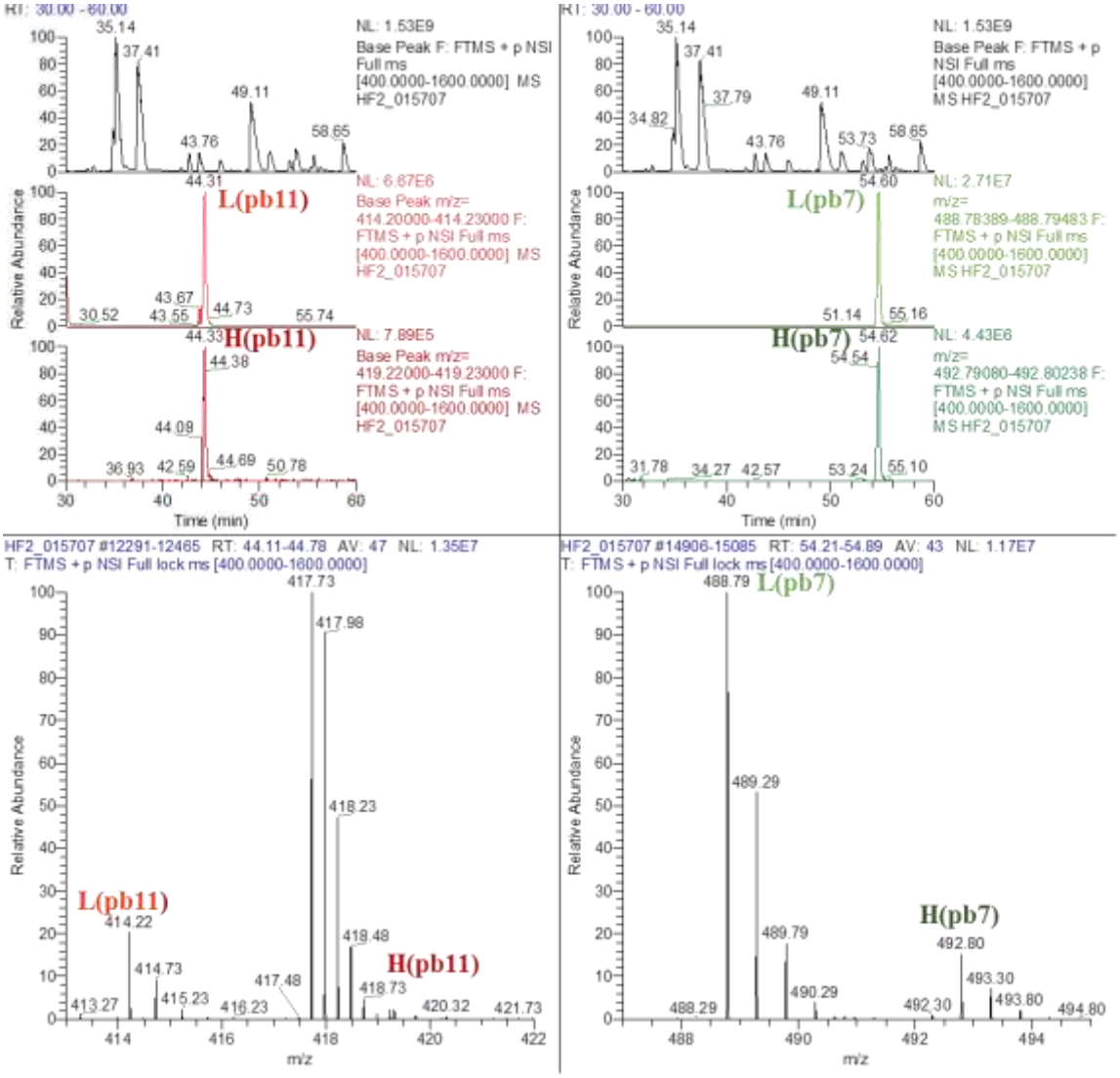


Figure 0-4. Observed base peak chromatograms, and extracted chromatograms of DFC 1, in Table III 10..

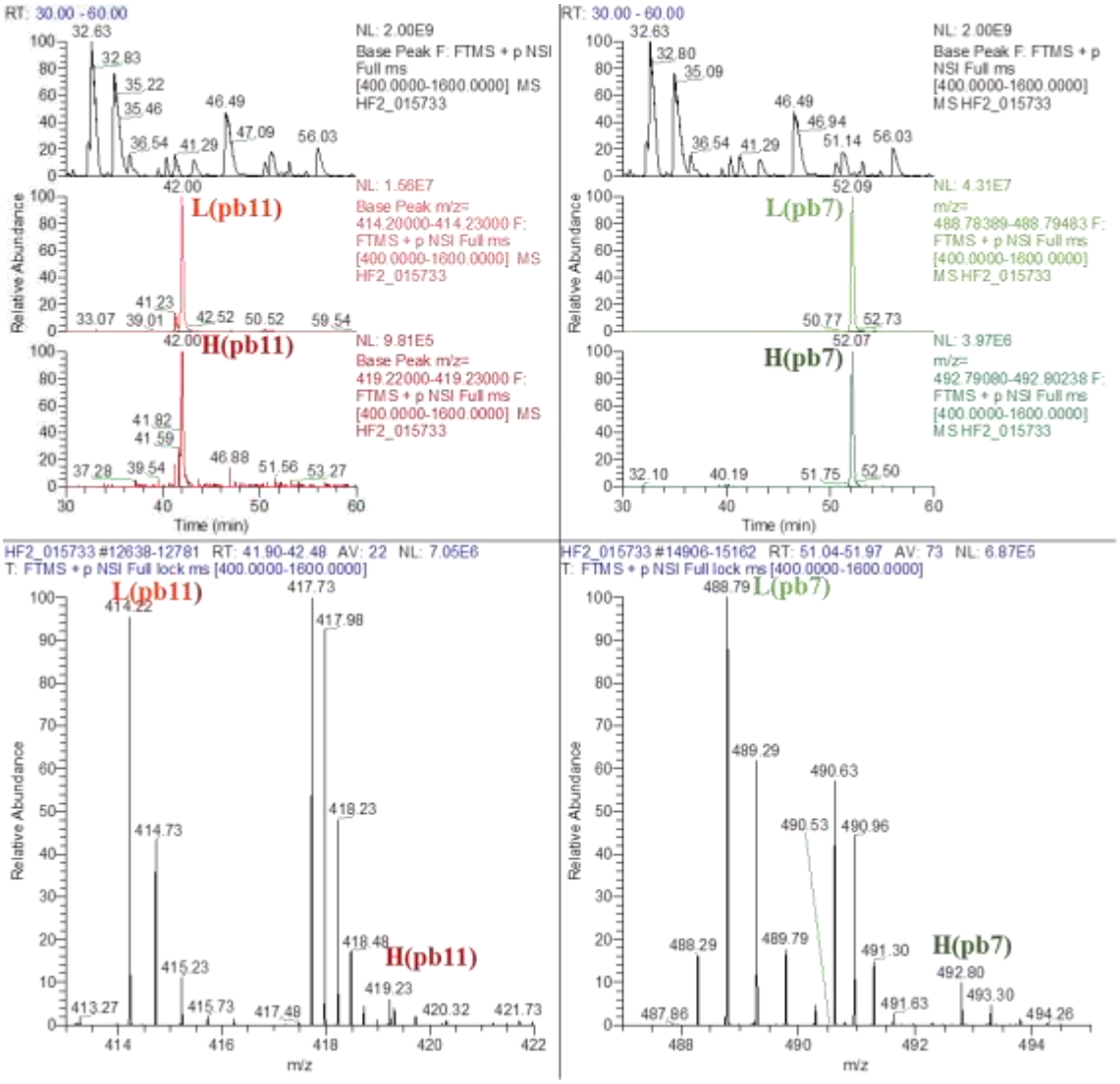


Figure 0-5. Observed base peak chromatograms, and extracted chromatograms of DFC 2, in Table III 10.

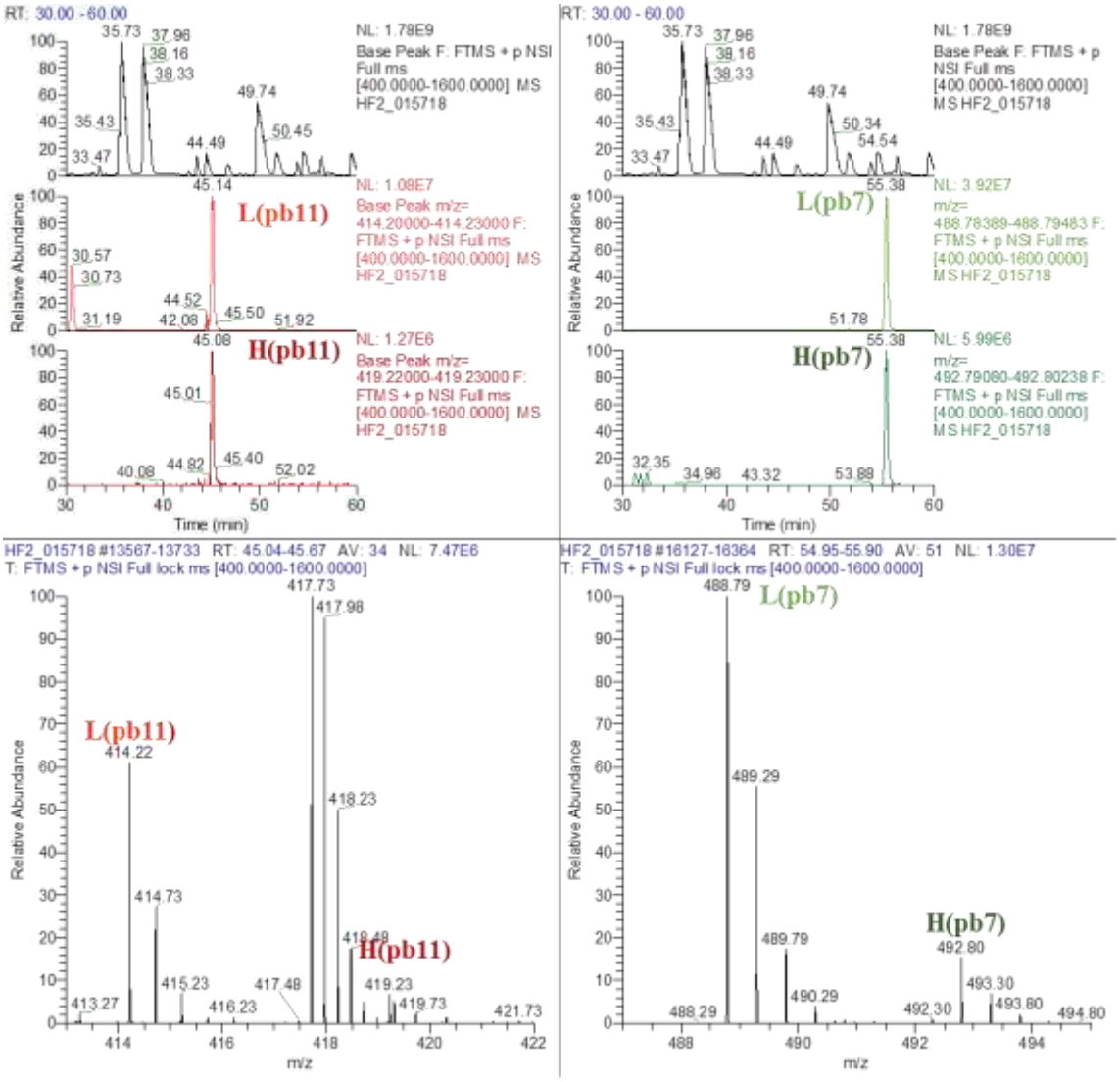


Figure 0-6. Observed base peak chromatograms, and extracted chromatograms of DFC 3, in Table III 10.

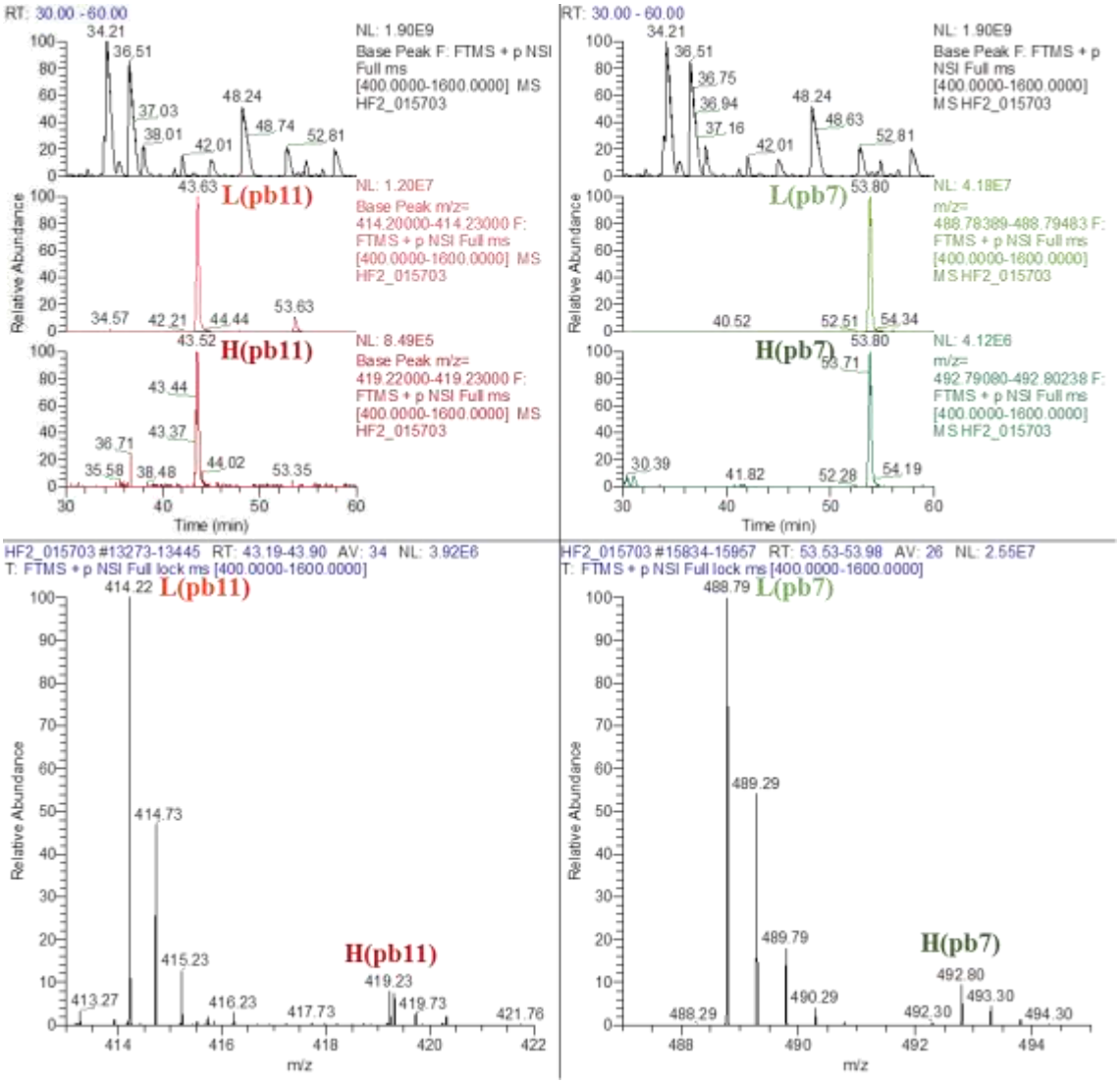


Figure 0-7. Observed base peak chromatograms, and extracted chromatograms of FC 1, in Table III 10.

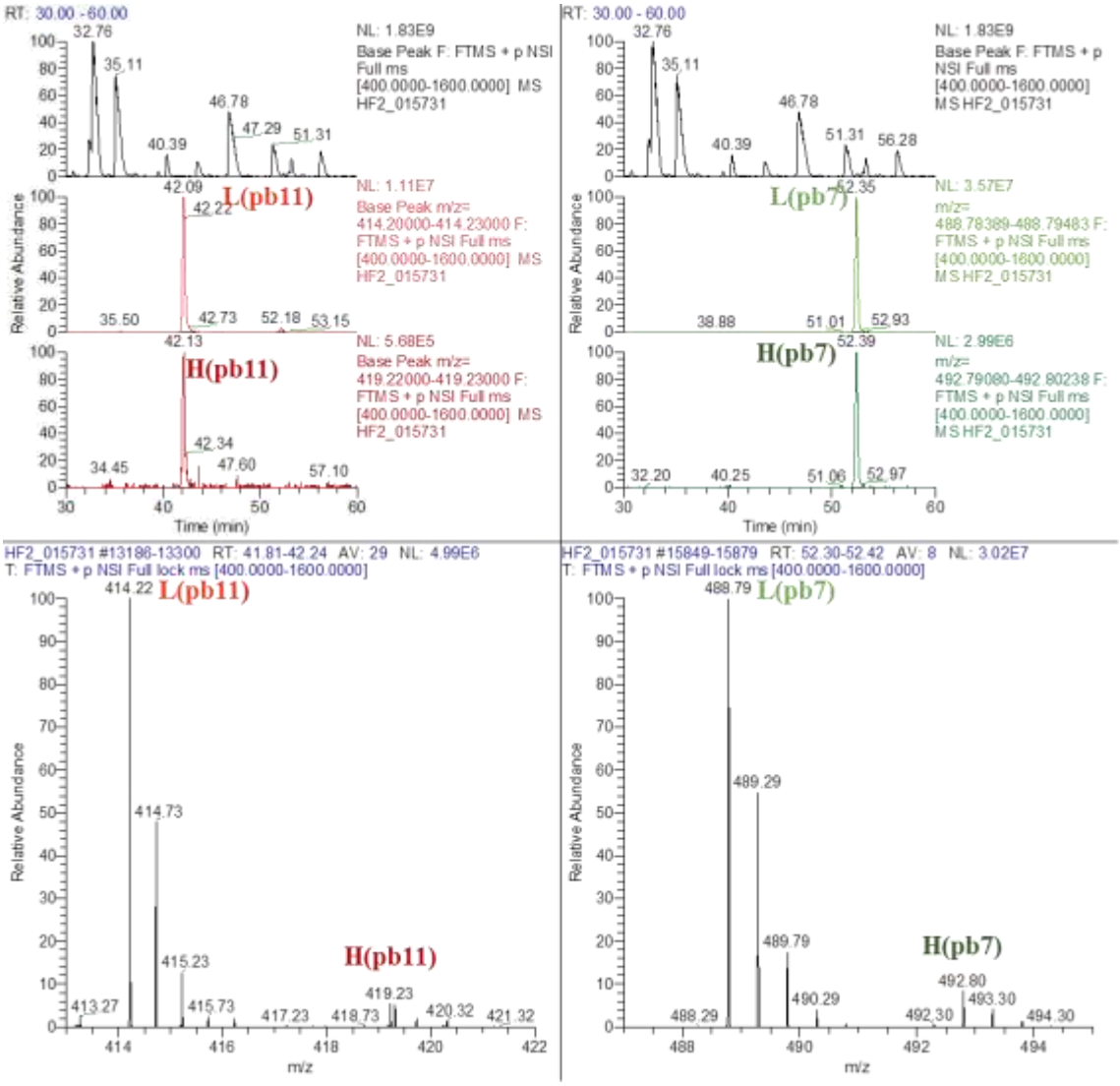


Figure 0-8. Observed base peak chromatograms, and extracted chromatograms of FC 2, in Table III 10.

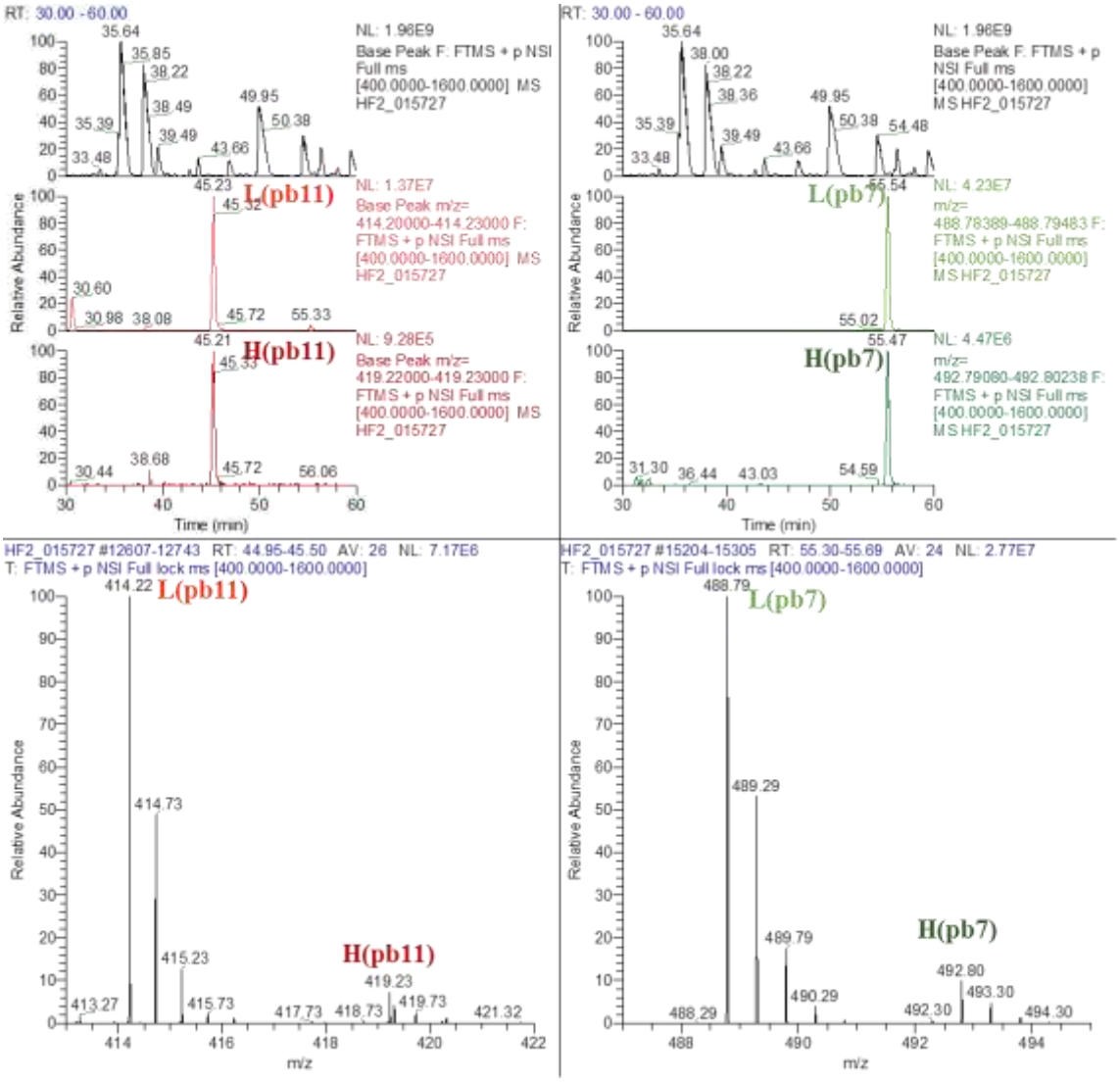


Figure 0-9. Observed base peak chromatograms, and extracted chromatograms of FC 3, in Table III 10.

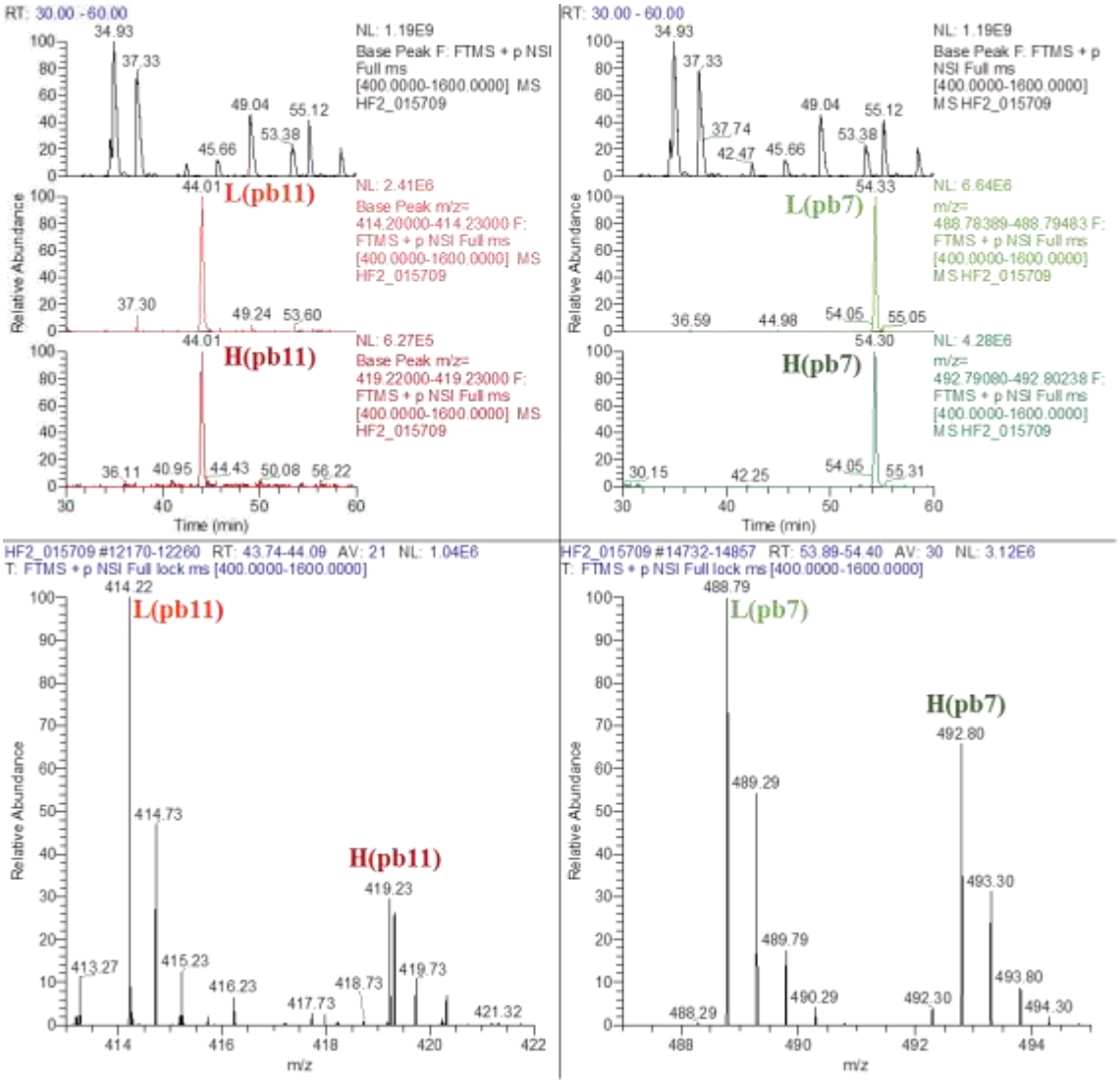


Figure 0-10. Observed base peak chromatograms, and extracted chromatograms of EC 1, in Table III 10.

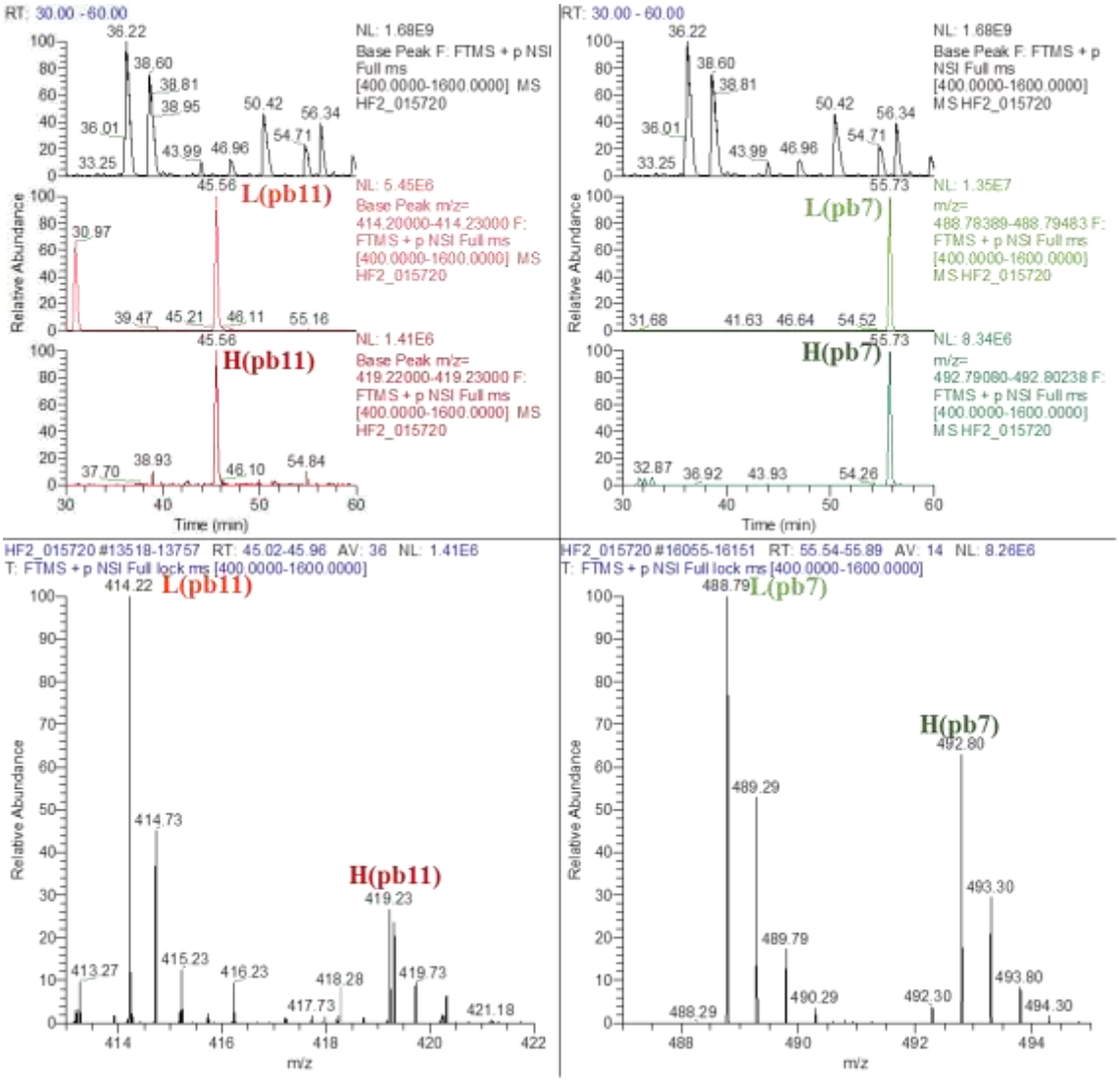


Figure 0-11. Observed base peak chromatograms, and extracted chromatograms of EC 2, in Table III 10.

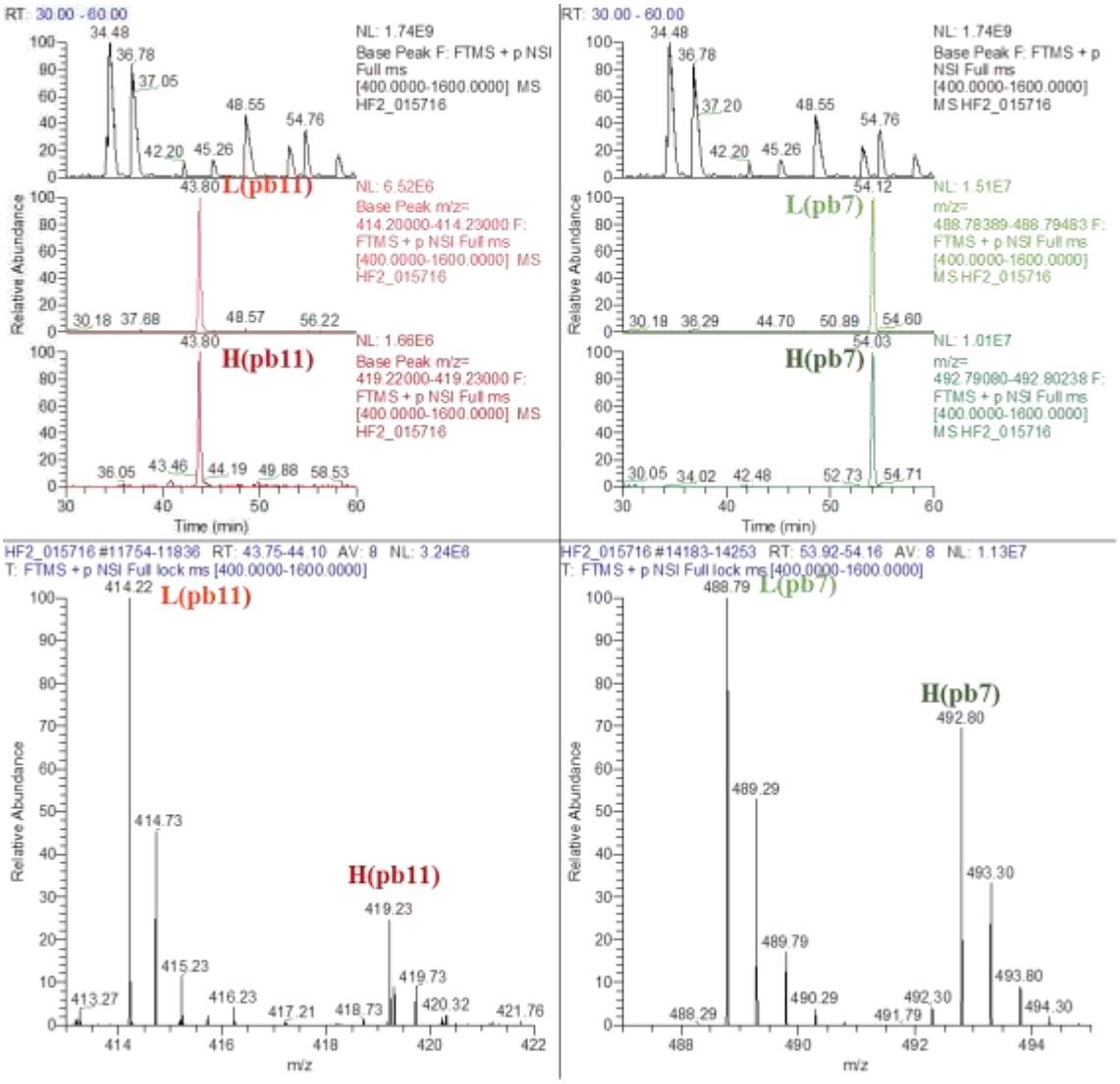


Figure 0-12. Observed base peak chromatograms, and extracted chromatograms of EC 3, in Table III 10.

APPENDIX II: NTA

NTA Date	*Sample*	*Buffer exchange*	*Solvent*	*Stage*	*NTA size*	*NTA conc*	*Aggregate*
24/01/2020	**FC2**	Dialysis	In water against Am.Ac	Before DY	84,0	1,47E+12	no
				1st DY	83,9	1,15E+12	no
				2nd DY	83,8	9,50E+11	no
				3rd DY	83,8	1,22E+12	no
06/02/2020	**FC2**	Dialysis	In water against Am.Ac	Before DY	78,8	3,32E+12	no
				1st DY	87,2	1,41E+12	no
				2nd DY	81,8	1,32E+12	no
				3rd DY	81,2	1,37E+12	no
Repeat on 7/02/2020	**FC2**	Dialysis	In water against Am.Ac	Before DY	79,3	2,56E+12	no
				1st DY	84,7	1,39E+12	no
				2nd DY	98,7	1,09E+12	yes
				3rd DY	94,8	9,60E+11	yes
11/02/2020	**FC2**	Dialysis	In water against Am.Ac	Before DY	78,1	3,58E+12	no
				1st DY	82,9	3,02E+11	no
				2nd DY	81,4	4,99E+11	no
				3rd DY	81,5	4,67E+11	no
24/01/2020	**FC2**	Dialysis	In and against Am.Ac	Before DY	83,5	1,79E+12	no
				1st DY	84,3	1,21E+12	no
				2nd DY	83,7	1,18E+12	no
				3rd DY	87,4	1,08E+12	no
06/02/2020	**FC2**	Dialysis	In and against Am.Ac	Before DY	78,8	3,32E+12	no
				1st DY	87,2	1,41E+12	no
				2nd DY	81,8	1,32E+12	yes
				3rd DY	81,2	1,37E+12	yes
Repeat on 7/02/2020	**FC2**	Dialysis	In and against Am.Ac	Before DY	79,3	2,56E+12	no
				1st DY	84,7	1,39E+12	no
				2nd DY	98,7	1,09E+12	yes
				3rd DY	94,8	9,60E+11	yes
11/02/2020	**FC2**	Dialysis	In and against Am.Ac	Before DY	80,9	4,31E+12	no
				1st DY	81,7	3,46E+11	no
				2nd DY	80,7	4,69E+11	no
				3rd DY	81,5	3,41E+11	no
23/01/2020	**FC2**	Ultra-filtration	Am.Ac	Before AF	79,0	2,79E+12	no
				1st AF	79,8	1,79E+12	no
				2nd AF	78,3	1,99E+12	no
				3rd AF	78,9	1,58E+12	no
				4th AF	79,9	8,25E+11	no
14/01/2020	**FC3**	Ultra-filtration	water	Before AF	80,1	1,91E+12	no
				1st AF	81,1	1,45E+12	no
				2nd AF	95,1	1,07E+12	yes
				3rd AF	102,2	1,45E+12	yes

NTA Date	*Sample*	*Buffer exchange*	*Solvent*	*Stage*	*NTA size*	*NTA conc*	*Aggregate*
14/01/2020	**FC2**	Ultra-filtration	water	Before AF	80,5	9,00E+11	no
				1st AF	80,5	8,05E+11	no
				2nd AF	85,0	1,70E+12	yes
				3rd AF	98,3	5,15E+11	yes
Repeat on 18/01/2020	**FC2**	Ultra-filtration	water	Before AF	80,2	3,95E+11	no
				1st AF	81,0	3,80E+11	no
				2nd AF	87,7	1,54E+12	yes
				3rd AF	90,9	5,89E+11	yes
Repeat on 04/02/2020	**FC2**	Ultra-filtration	water	Before AF	80,0	2,84E+11	no
14/01/2020	**FC2**	Ultra-filtration	water	Before AF	80,7	1,13E+12	no
				1st AF	81,7	1,21E+12	no
				2nd AF	84,2	1,71E+12	yes
				3rd AF	92,9	6,60E+11	yes
Repeat on 18/01/2020	**FC2**	Ultra-filtration	water	1st AF	80,9	1,36E+12	no
				2nd AF	86,5	1,91E+12	yes
				3rd AF	92,0	6,43E+11	yes
15-16/01/2020	**FC3**	Dialysis	water	Before DY	79,2	3,19E+12	no
				1st DY	94,5	1,01E+12	yes
				2nd DY	159,5	8,40E+11	yes
				3rd DY	584,0	8,85E+11	yes
Repeat on 18/01/2020	**FC3**	Dialysis	water	Before DY	80,6	1,85E+12	no
				1st DY	91,2	8,41E+11	yes
				2nd DY	101,6	7,05E+11	yes
				3rd DY	356,6	8,92E+08	yes
Repeat on 04/02/2020	**FC3**	Dialysis	water	Before Dy	81,0	1,01E+12	no
15-16/01/2020	**FC2**	Dialysis	water	Before DY	80,8	1,59E+12	no
				1st DY	87,7	7,60E+11	yes
				2nd DY	514,4	5,60E+11	yes
				3rd DY	353,4	7,10E+11	yes
Repeat on 04/02/2020	**FC2**	Dialysis	water	Before Dy	80,9	1,86E+12	no
15-16/01/2020	**FC2**	Dialysis	water	Before DY	80,6	1,67E+12	no
				1st DY	87,0	1,03E+12	yes
				2nd DY	333,7	7,30E+11	yes
				3rd DY	377,8	7,50E+11	yes
Repeat on 04/02/2020	**FC2**	Dialysis	water	Before Dy	80,2	2,04E+12	no

ABSTRACT

Vaccination is the most effective method of infectious disease prevention and control. Current vaccine development strategies do not provide broad access to vaccines in the developing world largely due to high costs and cold chain requirements. Virus Like Particles (VLP) are recombinant viral structures that exhibit immune-protective traits of native virus but are non-infectious. This developing technology has applications in the field of vaccines, gene therapy and antigen display. While most VLPs are derived from human pathogens, viruses infecting bacteria have recently been proposed as promising antigen display platforms. Bacteriophage T5 capsids exhibit 120 copies of a modifiable decoration protein at their surface, and could thus be used as a model for antigen-presenting VLP structures.

The present thesis focuses on studying bacteriophage T5 capsids from different perspectives using multiple mass spectrometry (MS) approaches. One of the aspect is to contribute to an improved understanding of T5 capsid assembly process by determining the stoichiometry of the T5 capsid protease using MS based Proteomics and the other aspect is to assess the integrity and stability of the phage T5 capsids by analyzing their intact mass using Nano-Electro-Mechanical-Sensors Mass Spectrometry (NEMS-MS).

The maturation of the T5 capsid begins with the activation of a protease called pb11, which in turn processes other structural proteins. In spite of its importance for T5 capsid maturation, the exact quantity of pb11 at each stage of the assembly is still subject to debate. The present work describes an original targeted proteomic strategy to determine the presence and copy number of this important protein in mature T5 capsids using a heavy isotope labelled quantification concatemer (QconCAT).

Characterization of massive supramolecular assemblies composed of millions of atoms such as VLPs is very challenging, yet it is an important requirement for vaccine production. NEMS-MS, a novel method for single particle mass sensing, is a promising technology to investigate large biological species that have so far escaped mass characterization. NEMS are nanoscale devices, such as cantilevers or suspended beams, vibrating at their resonance frequency. When a particle lands onto the surface of a NEMS, the device’s resonance frequency shifts downwards in proportion to the added mass. Mass sensing can thus be performed by monitoring this frequency in real time.Our study investigates properties independent from particle mass that may influence the uncertainty in mass measurement using NEMS-MS, and how to deal with these issues. An evaluation of the magnitude of all these effects on bacteriophage T5 capsid mass measurements using doubly clamped beams is proposed. Finally, we determine that particle desolvation affects mass measurement more than the physical parameters of the capsid or uncertainties in device geometry.

In conclusion, this work addresses the various attributes and parameters affecting capsid mass determination using nano-resonator-based MS, and reveals the actual copy number of protease protein in mature T5 capsid using nanoLC-MS proteomics.

Keywords: Bacteriophage T5, Mass spectrometry, NEMS, QconCAT

RESUME

La vaccination est la méthode la plus efficace de prévention et de contrôle des maladies infectieuses. Les stratégies actuelles de développement des vaccins ne permettent pas un large accès à la vaccination dans les pays en développement en raison de coûts élevés et de la nécessité de maintenir la chaine du froid. Les pseudo-particules virales (VLP) sont des structures virales recombinantes et qui possèdent les même propriétés immuno-protectrices que les virus natifs, mais ne sont pas infectieuses. Elles ont des applications en vaccination, mais également en thérapie génique et en présentation d'antigènes. Si la plupart des VLPs sont dérivées de pathogènes affectant les humains, les virus de bactéries ont été proposés comme plateformes de présentation d'antigènes. Les capsides de bactériophage T5 exposent à leur surface 120 copies d'une protéine modifiable, et constituent un modèle intéressant de structure VLP pour la présentation d'antigènes.

Cette thèse propose une étude des capsides de bactériophage T5 utilisant une multiplicité d'approches basées sur la spectrométrie de masse (MS). Un aspect est de contribuer à la compréhension fine des processus d'assemblage des capsides en déterminant la stœchiométrie de la protéase de T5 par une approche de protéomique basée sur la MS et l'autre aspect est de caractériser l'intégrité et la stabilité des capsides de phage T5 en analysant leur masse totale par NEMS-MS.

La maturation de la capside de T5 débute avec l'activation d'une protéase appelée pb11, qui vient ensuite cliver d'autres protéines de structure. Malgré son importance pour la maturation de la capside, la quantité exacte de pb11 à chaque étape de l'assemblage est encore sujette à discussions. Le présent travail décrit une stratégie expérimentale originale de protéomique ciblée pour déterminer la présence et le nombre de copies de cette importante protéine dans les capsides matures, en utilisant un concatémère (QconCAT) de deux peptides isotopiquement alourdis.

La caractérisation d'assemblages supramoléculaires massifs composés de millions d'atomes, tels que les VLPs est difficile, mais requise pour la production de vaccins. La spectrométrie de masse par capteurs nano-electro-mécaniques (NEMS-MS) est une technologie prometteuse pour étudier ces entités biologiques imposantes qui ont pour l'instant échappé à la mesure de masse. Les NEMS sont des dispositifs nanométriques, (leviers ou poutres suspendues), qui vibrent à leur fréquence de résonance. Lorsqu'une particule se pose à leur surface, la fréquence de résonance décroit proportionnellement à la masse déposée. Dans ce présent travail, nous étudions propriétés indépendantes de la masse des particules analysées qui peuvent affecter l'incertitude de mesure de masse par nano-résonateurs. Nous présentons une évaluation de la magnitude de ces effets sur la mesure de masse de capsides de bacteriophage T5 à l'aide de nano-poutres suspendues. Finalement, nous avons déterminé que la désolvatation incomplète des particules analysées affecte la mesure de masse bien plus largement que les paramètres physiques de la capside ou les imperfections de fabrication des dispositifs.

En conclusion, ce travail révèle le véritable nombre de copies de la protéase de T5 dans les capsides matures grâce à la protéomique par MS, etaborde les divers attributs et paramètres qui peuvent affecter la détermination de la masse de capsides intactes par MS sur nano-résonateurs.

Mots-clés : Bactériophage T5, Spectrométre de masse, NEMS, QconCAT