Novel Fluorescence and Fluorine Labelling Methods for Viruses and Virus-like Particles

A thesis submitted in partial fulfilment for the degree of Doctor of Philosophy at the University of Oxford

by **Lok Chun Rogen, Leung**

Department of Chemistry, University of Oxford

October, 2016

St Hugh's College

Declaration

I declare the work presented in this thesis was carried out at the Chemistry Research Laboratory and Department of Biochemistry, University of Oxford, UK. All the work is my own, except where otherwise stated, and has not been submitted for any other degree at this or any other university.

The transverse relaxation time study of Qβ particles and the ^{19}F diffusion experiment of intact Qβ particles were performed in collaboration with Prof. Andrew Baldwin in the Chemistry Department, University of Oxford.

MS/MS analysis was performed in collaboration with Dr. Benedikt M. Kessler in Target Discovery Institute, University of Oxford.

Lok Chun Rogen, Leung

October 2016

Abstract

Molecular imaging involves the development of probes which can specifically label a certain object in the body at cellular or subcellular level. This thesis consists of three parts, each involving the development of novel labelling methods for viruses or virus-like particles with specific applications.

Virus-like particles (VLP) derived from the *E. coli* bacteriophage Qβ are widely employed as a nano-carrier for drugs and vaccines, but a powerful method for tracing its circulation without affecting its structure is yet to be developed. In the first part of the thesis, the electrophilic fluorine source ^{19}F-Selectfluor^{TM} was employed for introducing single fluorine atoms on Qβ VLPs. For the 'tag-and-modify' approach, site-selective electrophilic C-F bond formation was achieved on the dehydroalanine (Dha) amino acid tag of VLPs under aqueous conditions. Chemoselective electrophilic aromatic fluorination on tyrosine residues were also achieved using the same reagent by manipulating the amino acid sequence. Similar results were observed in conditions required for ^{18}F-Selectfluor^{TM} reaction, indicating the potential of this technique for positron emission tomography (PET) imaging.

In addition, there is a lack of *in situ* technique for tracking the functional status of Qβ VLPs and hence the release of cargos. In the second part of the thesis, a simple way to monitor the disassembly of ^{19}F-labelled Qβ VLPs by ^{19}F NMR spectrosocpy is reported. Analysis of resonances, using experiments under a range of conditions, allowed determination not only of the intact particle but also the disassembled multimeric species and even smaller peptides upon digestion by cells. This in turn allowed mutational redesign of disassembly and testing in both bacterial and mammalian systems as a strategy for the creation of putative, targeted-VLP delivery systems.

In the third part of the thesis, a new type of rhodamine B fluorescent dye functionalised with a 2-imino-2-methoxyethyl (IME) group is reported. The amidine linkage formed between the IME group and lysine residue retains the pKaH of the original side chain, which cannot be achieved using commercially available conjugating dyes. This in turn minimises the change in net charge hence virus infectivity following virus labelling. By employing adenovirus (AV) as an example, the IME dye was shown to be a better choice in retaining virus infectivity compared to dyes linked with other coupling groups. In addition, preliminary experiments on dengue virus with the synthesised dyes were also performed.

Acknowledgement

First of all, I would like to show my immense gratitude to my parents for guiding and supporting me in all these years. I would not have a chance to study a PhD in Oxford without them. I am really happy that I can pursue my dream as a scientist. As I did not get any scholarship funding, it would be impossible for me to study abroad without their financial support.

I would also like to thank my PhD supervisor, Prof. Ben Davis, for his guidance during my PhD. I am so lucky to be in his group in which I can explore both chemistry and biology, which are my two favourite subjects. It is also an invaluable experience for me to involve in different interdisciplinary projects. Furthermore, he gave me a lot of inspiring ideas which made my projects more interesting.

Moreover, I cannot finish my projects without my helpful collaborators. Prof. Andrew Baldwin and Prof. Tim Claridge have given me a lot of advices regarding to all the NMR experiments. Prof. Christian Eggelings and Dr. Jakub Chojnacki have provided great support for microscopy experiments. Dr. Kerstin Lühn, Dr. Anwen Howell and Dr. Joanna Miller also taught me a lot of things related to dengue virus, a dangerous but exciting creature.

Also, I want to thank my dear lab members for proof-reading my thesis, teaching me a lot of techniques and making me laugh (!). Special thanks to Bala (for all the chemistry and TLC!), Leanne (for all the virus stuff), Seb (for the jokes), Kim, Anuchit, Billy, Yiqun, Cristina, Jens, Mylene and also my best friend Ritu Raj. I would also like to thank all the

members from LG8 and LG10, the two wonderful laboratories that I stayed. I cannot survive in these four years without them.

I would also like to show my great appreciation to Prof. Hagan Bayley and Prof. Jeroen Cornelissen for their guidance during my viva examination and their excellent advices on my thesis. This thesis would not be completed without them.

Last but not least, I want to thank my beloved girlfriend, Angel Chan for all her support throughout the years. Thanks for being so patient and kind to me. I am now finally submitting my thesis! This marks the end of my PhD and I truly enjoyed the time with no regrets.

「不悔夢歸處

只恨太匆匆」

Abbreviations

ADE	antibody-dependent enhancement
AF	Alexa Fluor®
Aha	azidohomoalanine
AV	adenovirus
CAR	coxsackie-adenovirus receptor
CNM	cyanomethyl
DAPI	4',6-diamidino-2-phenylindole
DC-SIGN	dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin
DBHDA	α,α′-di-bromoadipyl(bis)amide
DENV	dengue virus
DF	dengue fever
Dha	dehydroalanine
Dfa	(*N*,α)-dehydro-β-fluoro-alanine
Dfm	difluoromethionine
DiD	1,10-dioctadecyl-3,3,30,30-tetramethylindo- dicarbocyanine, 4-chlorobenzenesulfonate salt
DLS	dynamic light scattering
DTT	dithiothreitol
EDC	1-ethyl-3-(3-dimethylamino-propyl)-carbodiimide
ER	endoplasmic reticulum
FBEM	*N*-[2-(4-fluorobenzamido)-ethyl]maleimide
FcγR	Fcγ receptor
FDG	2-fluoro-2-deoxy-glucose
FITC	fluorescein isothiocyanate
Fluo-NHS	5/6-carboxyfluorescein succinimidyl ester
FPSC	3-fluoro-propanesulfonyl chloride
FWHH	full width at the half-height
GFP	green fluorescence protein

GlcNAc-SH	thiol-functionalised *N*-acetylglucosamine
Hfa	α-hydroxy-β-fluoro-alanine
HBV	hepatitis B virus
HEL	hen egg white lysozyme
IME	2-imino-2-methoxyethyl
ITC	isothiocyanate
LB	Luria-Bertani
LC-MS	liquid chromatography mass spectrometry
MR	mannose receptor
OI	optical imaging
Qβ-F	Qβ K16Tfm
Qβ-CS	Qβ K16Tfm C74S C80S
MRI	magnetic resonance imaging
MRS	magnetic resonance spectroscopy
NHS	*N*-hydroxysuccinimidyl
NMR	nuclear magnetic resonance
ON	oligonucleotide
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PET	positron emission tomography
SDS	sodium dodecyl sulfate
SET	single-electron transfer
SFB	*N*-(hydroxysuccimidyl)-4-fluorobenzoate
SPECT	photon emission computed tomography
TAPS	N-[Tris(hydroxymethyl)methyl]-3-aminopropane-sulfonic acid
TEM	transmission electron microscopy
TFA	trifluoroacetic acid
TFAcetone	1,1,1-trifluoroacetone
Tfm	trifluoromethionine
TLC	thin layer chromatography
VLP	virus-like particle
WT	wild type

Table of Contents

Introduction 1

CHAPTER

1 Introduction

1.1 What is Molecular Imaging?

Molecular imaging is an emerging field which involves the development of probes that can specifically label a certain object in the body at cellular or subcellular level.[1-3] It usually allows a non-invasive, repetitive and uniform study of a pathological or physiological process in a living subject. The analysis typically comes with both spatial and temporal data, which is vital for studying a dynamic biological process.[4-6] These characteristics render molecular imaging an invaluable tool for detecting and characterising early disease,[7, 8] understanding pharmacokinetics[9, 10] and studying other molecular pathways[11-13] taking place in an intact organism.

The main modalities encompassed within molecular imaging include radionuclide imaging, magnetic resonance imaging (MRI), optical imaging (OI), X-ray computed tomography imaging and ultrasound imaging. Each technology has its own advantages and disadvantages (**Figure 1.1**).[14, 15] Recently, imaging agents with multi-modality, such as dual-function PET/near-infrared fluorescence probes were developed,[16] which further expand the potential of molecular imaging. In this thesis, only radionuclide imaging, MRI and OI are discussed.

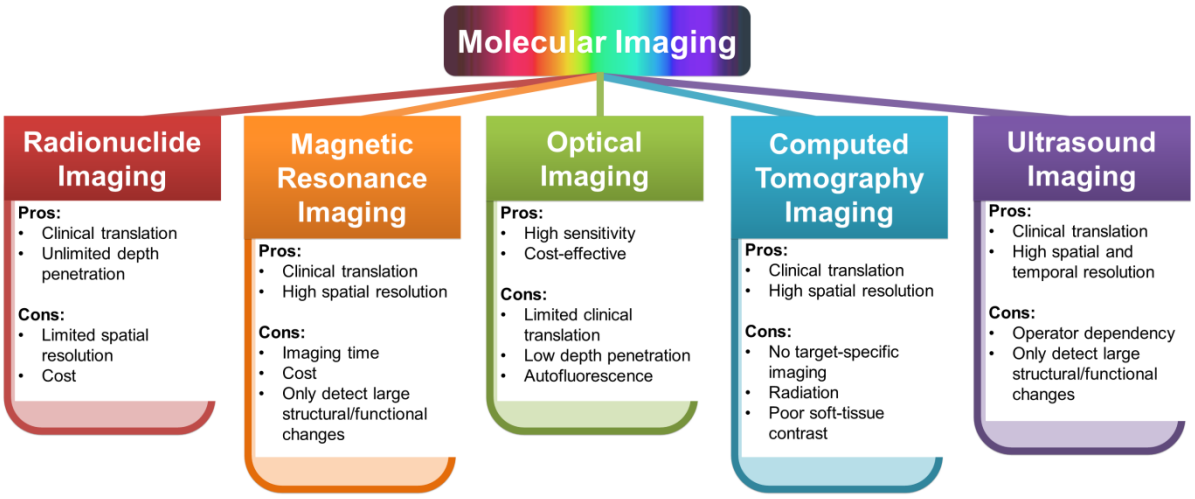


Figure 1.1: The five main modalities of molecular imaging, with pros and cons of each listed.[14, 15]

1.1.1 Radionuclide imaging

Radionuclide imaging includes both positron emission tomography (PET) and single photon emission computed tomography (SPECT). Their high sensitivity renders them popular in clinical and preclinical studies.[17-19] For PET imaging, the imaging agent usually possesses a trace amount of a positron-emitting radioisotope. The emitted positron loses its kinetic energy rapidly by colliding with an electron on the same atom to form two 511 keV γ-rays travelling at directions 180^{o} apart (**Figure 1.2A**).[20] The emitted γ-rays are then detected and converted into an electrical signal that can be processed into tomographic images. The commonly used radioisotopes includes ^{11}C ($t_{1/2}$ = 20 min), ^{13}N ($t_{1/2}$ = 10 min), ^{15}O ($t_{1/2}$ = 2 min), ^{18}F ($t_{1/2}$ = 110 min), ^{64}Cu ($t_{1/2}$ = 13 h) and ^{76}Br ($t_{1/2}$ = 16 h).[21] PET has exceptionally high sensitivity (10^{-11} to 10^{-12} M) and unlimited penetration depth, which make it very powerful in the detection of cancer.[22, 23] The ^{18}F-labelled imaging agent ^{18}F-2-fluoro-2-deoxy-glucose (^{18}F-FDG) developed in 1978 is one of the most effective imaging agents in this regard since it acts as a glucose analogue, thus enabling easy detection of the abnormally high glucose uptake rate in cancer cells.[24, 25]

On the other hand, SPECT detects the γ-rays emitted by the radionuclide directly during its decay (**Figure 1.2B**). The γ-rays of each type of radionuclide have their own characteristic energy, thus allowing the use of different radioisotopes in the same study.[26] Commonly used radionuclides include ^{99m}Tc (140 keV; $t_{1/2}$ = 6 h), ^{111}In (171 keV and 245 keV; $t_{1/2}$ = 3 d) and ^{123}I (27 keV and 159 keV; $t_{1/2}$ = 8 d). SPECT has a relatively low sensitivity compared to PET as it requires the use of physical collimators in which a lower percentage of emitted photon is detected.[27] In here, only applications towards PET will be discussed.

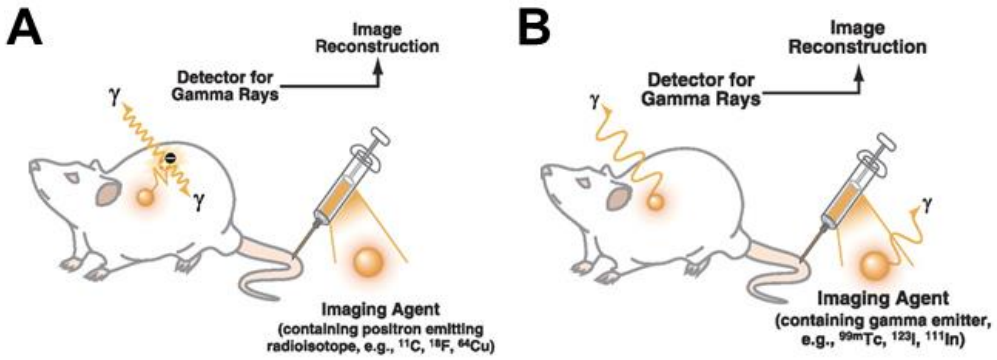


Figure 1.2: Basic principles of (A) PET and; (B) SPECT imaging.[14]

1.1.2 Magnetic resonance imaging

MRI is a technique extended from nuclear magnetic resonance (NMR) (also called magnetic resonance spectroscopy (MRS) in the clinical field) methodology commonly used in chemistry. Only atoms with nuclei possessing a net angular momentum (i.e. spin) can be detected by NMR spectroscopy. In the presence of a high static magnetic field, nuclei such as ^{1}H and ^{19}F (with a 1/2 spin) will split into two orientations with different energy levels, with one pointing in the same direction of the magnetic field having a lower energy. The energy difference depends on the type of nuclei used and also the strength of the magnetic field. A range of radio frequency is then applied at a perpendicular direction to that of the magnetic field. Once the nuclei with lower energy state absorb the energy at a specific frequency, they become excited and a change of angular momentum occurs. The nuclei will then return to the

lower energy state again after some time by emitting radio waves (i.e. relaxation), which are detected by the machine (**Figure 1.3**). The biggest difference between MRI and NMR spectroscopy is that the former has a gradient of magnetic field present in the sample. In other words, the magnetic field strength varies through the sample, allowing analysis of certain points at a time using a specific radio frequency. In principle, if a system works in NMR spectroscopy, it can be applied for MRI.[28]

The main drawback of MRI is its low sensitivity (minimum concentration required: 10^{-3} to 10^{-4} M) compared to other methods. At 37 ^{o}C, only around 1 in 10^{6} nuclei aligns with the applied magnetic field. The amount of signal received is thus extremely low, resulting in poor sensitivity.[29] Clinical imaging is greatly restricted by this sensitivity problem, even for detecting the most abundant nuclei, such as ^{1}H in water. This is commonly solved by methods such as addition of contrast agents, in which the T1 relaxation time of nearby water protons are usually shortened.[30, 31] On the other hand, NMR spectroscopy (or MRS) has been used extensively for *in vivo* studies, providing useful information about tissue energetic and cellular metabolism.[32-35] For instance, the introduction of ^{19}F into drugs such as fluoxetine allows the investigation of the drug distribution in different tissues.[36]

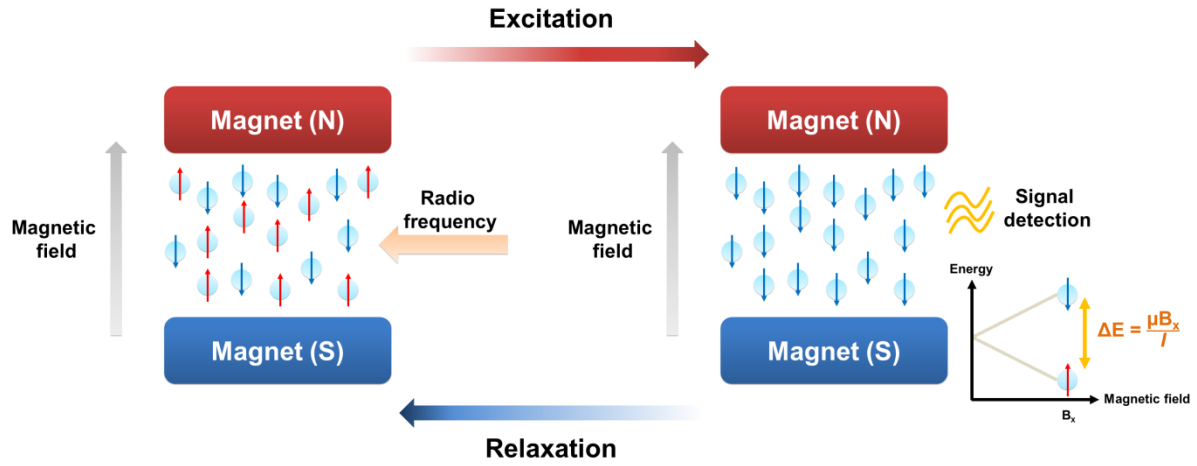


Figure 1.3: Basic principle of NMR spectroscopy or MRI. When the nuclei are excited by radio frequency under an external magnetic field, they will emit energy (ΔE) during relaxation. **μ**: magnetic moment of the nucleus; $\mathbf{B_x}$: external magnetic field; **I**: spin of the nucleus.

1.1.3 Optical imaging

The most widely used OI methods can be divided into fluorescence and bioluminescence imaging. These modalities are usually cost effective, easily performed and have high sensitivity (up to 10^{-15} M). However, they cannot penetrate into deep tissues, thus restricting them to cellular level or to preclinical study in small animals. As a result, OI plays a crucial role in studying tumour and relevant drug development.[37] Only fluorescence imaging is discussed here.

Fluorescence is a phenomenon in which a fluorophore absorbs and then emits light (**Figure 1.4A**).[38] An electron in the fluorophore molecular system can be excited from the ground state (S_0) after absorbing energy from an external light source having high energy (or short wavelength). The electron in the excited state (S_1) then undergoes non-radiative relaxation in which some energy is released as heat (i.e. vibrations) to the solvent. In most of the cases, the electron then returns to the ground state by emitting a photon with a lower energy (or longer wavelength). The absorption and emission of light of a specific fluorophore can be described by its characteristic absorption and emission spectra (**Figure 1.4B**).

The introduction of fluorescence microscopy in the past two decades has rendered fluorescence an invaluable tool for cell biology.[39] Fluorophores can be introduced to the target *via* genetic incorporation of fluorescent proteins[40] or bio-conjugation with fluorescent-tagged molecules.[41] The use of fluorescence microscopy allows the visualisation of intracellular protein-protein interactions[42] and other molecular events such as gene delivery in the nucleus.[43]

At the animal level, the living subjects are usually anaesthetised and placed in a dark chamber with a Charged Coupled Device (CCD) camera. The camera can convert the recorded fluorescence light to electronic signal and subsequently generate an image. Examples of fluorescence imaging in animals include the monitoring of cancer and

metastases with green fluorescence protein (GFP)-expressing tumour cells[44] and the detection of human epidermal growth factor receptor 2 using fluorescence-labelled antibodies.[45]

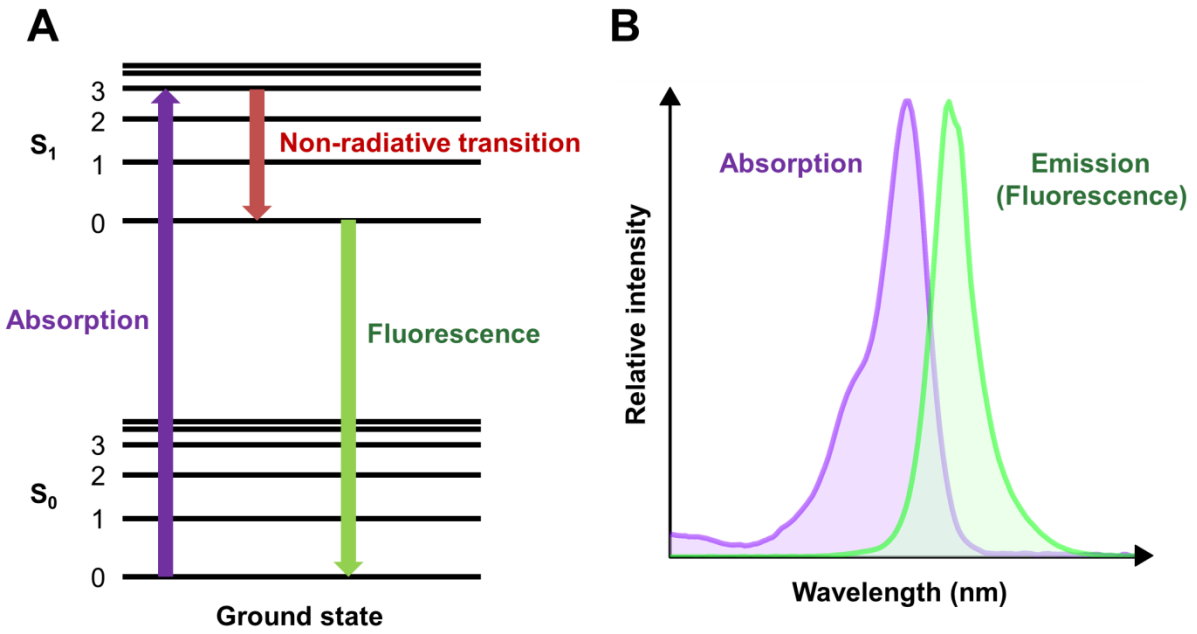


Figure 1.4: (A) Jablonski diagram showing the basic principle of fluorescence. In the presence of an external light, the electron in the molecular system is excited both electronically and vibrationally. The electron then relaxes vibrationally by releasing heat, and eventually fluoresces at a longer wavelength; (B) The absorption and emission spectra of a typical fluorophore.

1.2 Molecular Imaging for Drug Delivery

1.2.1 General background

For a drug to be effectively consumed by the target site in the body, it is vital to strictly localise its pharmacological activity to a specific organ or tissue. Targeted drug delivery can greatly reduce the dose of drug required and hence the drug toxicity. In addition, it is crucial to identify other pharmacological parameters such as dose schedule, biologically active drug

concentration, inhibiting effect of the target protein, etc.[5] Molecular imaging provides an excellent method for answering the above questions. By labelling the drug or its nano-carrier, a single multifunctional system with both therapeutic and imaging constituents is formed.[46] The drug delivery process can then be observed clearly in real time, providing useful information for drug development.

Image-guided drug delivery is an emerging field and a few examples have been reported. Harrington *et al.* employed ^{111}In-labelled PEGylated liposomes for monitoring drug delivery to cancer cells by SPECT imaging.[47] The anti-cancer protein tumour necrosis factor was also labelled with ^{64}Cu for studying its bio-distribution *via* PET imaging.[48] Fluorescent quantum dots have also been introduced to doxorubicin for treating cancer.[49]

1.2.2 Qβ virus-like particle as a drug nano-carrier

Virus-like particles (VLPs) are protein cages made up of hundreds of capsid protein subunits from different viruses. They have well-defined structures and can be strikingly stable under extremes of temperature,[50, 51] pH,[51] and in different solvents.[52] These render them potentially suitable for encapsulating materials such as proteins,[53, 54] synthetic polymers,[55, 56] oligonucleotides (ONs)[57, 58] and smaller molecules.[59, 60] Also, their surface can be used to append different functional groups, ligands or antigens for targeting,[61-63] imaging,[64] vaccination,[65] and other biomedical purposes.[66]

The bacteriophage Qβ belongs to the *Leviviridae* family. It has an icosahedral capsid with a T = 3 symmetry.[67] Qβ VLP with a diameter of 28 nm is formed from 180 copies of a 132 amino acid subunit (14.1 kDa) *via* strong protein-protein interactions (**Figure 1.5A**).[50, 68, 69] In the particle, the subunit has three different conformations (namely chain A, B and C) even though they have the same amino acid sequence. One unique advantage of using Qβ VLP is ascribed to its remarkable stability. Similarly to other members of the family, Qβ monomers

can form dimers *via* strong non-covalent interaction.[67, 70] Moreover, the dimers in Qβ are further stabilised by inter-subunit disulphide linkage at Cys-74 and Cys-80, which are present near the five- and three-fold axes respectively (**Figure 1.5B**).[71] These disulphide linkages are not necessary for particle formation and can be removed without damaging the particle structure.[72] Several different cargos have been successfully inserted into the Qβ cage of about 21 nm,[67] including proteins,[73] enzymes,[74] siRNA,[75] short DNA,[76-83] polyanionic macromolecules[84] and other small molecules.[75, 85, 86]

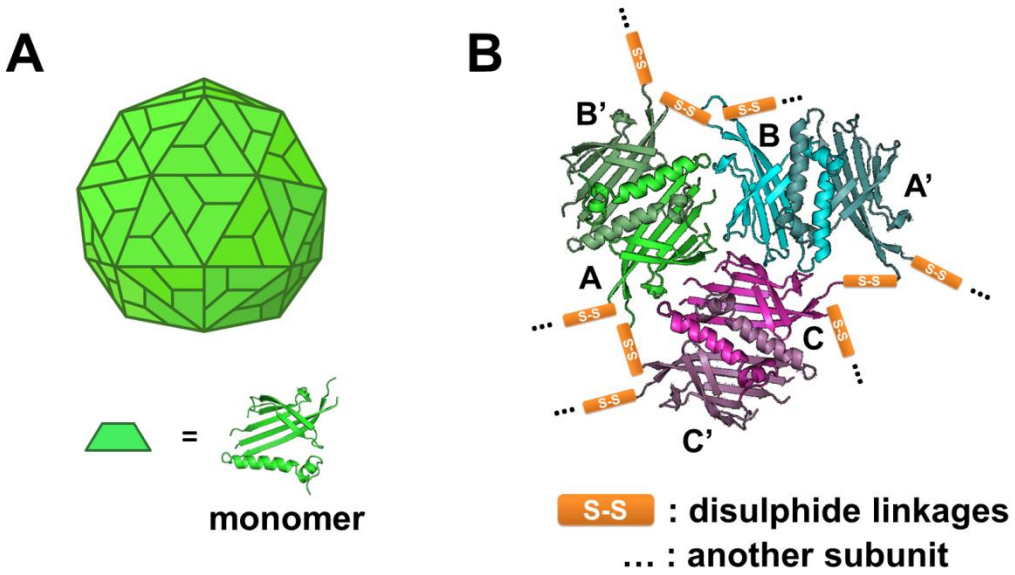


Figure 1.5: (A) Qβ VLP is formed by 180 copies of monomeric subunits; (B) The monomers interact with each other to form dimers (conformer A & B or conformer C & C). The particle is further stabilised by the presence of disulphide linkages between dimers.

In order to employ Qβ VLPs for drug delivery, the particle should first be labelled by an imaging agent. Previous reports have demonstrated the conjugation of quantum dots[75] and fluorescent proteins[73] to the interior of the particle. The surface of the particle has also been labelled with fluoro-mannose in a previous study.[87] However, these techniques require the introduction of bulky cargos or surface groups to the particle, which can greatly change the size and structure of the particle. It is therefore essential to develop a method for the

introduction of a small imaging group - or even a single atom - in order to minimise the effect of labelling. On the other hand, there is also a lack of techniques for detecting the release of drugs from the particle. In this work, fluorine was proposed as an imaging agent to address these problems.

1.2.3 Fluorine as a special element suitable for both NMR and PET

In the field of chemical biology, fluorine labelling is extensively used in protein structure determination,[88-90] protein design and protein engineering.[91, 92] Fluorine has two isotopes, ^{18}F and ^{19}F, with the later having nearly 100% natural abundance. The radioisotope ^{18}F releases positrons during decay, rendering it applicable to PET imaging.[20] On the other hand, the signal of the abundant ^{19}F isotope can be easily detected by NMR spectroscopy[93] and MRI[94] due to its nuclear spin property. In addition, unlike hydrogen, which is ubiquitous in the body, imaging involving fluorine does not give much background signal in biological samples. These advantages have made fluorine an attractive imaging agent especially in non-invasive studies of biological phenomena.[92, 95]

Strategies for fluorine labelling can be classified into three types: (i) direct protein labelling (**Figure 1.6A**); (ii) synthesis of labelled peptide (**Figure 1.6B**) and; (iii) biosynthetic labelling (**Figure 1.6C**). In this thesis, novel fluorine labelling methodologies and applications were developed based on the first and the third approach.

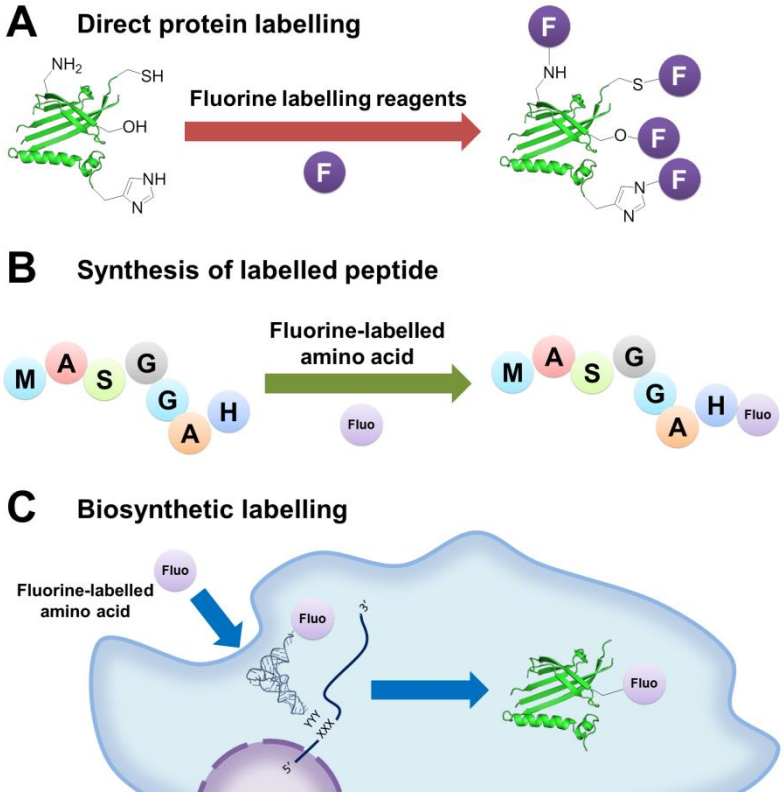


Figure 1.6: Schematic diagrams showing the three different ways for introducing fluorine into proteins: (A) Direct protein labelling; (B) synthesis of labelled peptide and; (C) biosynthetic labelling. XXX and YYY are the codons and anticodons used for unnatural amino acid incorporation respectively.

1.2.3.1 Direct protein labelling

Direct protein labelling can be achieved by mixing proteins and an excess amount of labelling reagent together, with purification afterwards. A cysteine residue is usually the labelling target due to its high nucleophilicity and low natural abundance.[96] Other amino acid residues like lysine,[97] serine,[98] threonine[98] and histidine[99] have also been reported as possible targeting sites. There are many commercially available ^{18}F or ^{19}F labelling reagents with different properties and amino acid specificities (**Figure 1.7**).[90] Thus, with appropriate mutagenesis, site-selective fluorination can be achieved by choosing suitable labelling reagents.[100] The direct labelling method is usually easy and straightforward, but sometimes problems with mutagenesis and purification after the reaction may arise. Also, additional chains or functional groups are usually added onto the target amino acid after the reaction,

which may affect the structure and property of the protein. Since the reagents are specific to particular amino acids but do not discriminate between different protein, it can only work *in vitro* and with pure proteins.

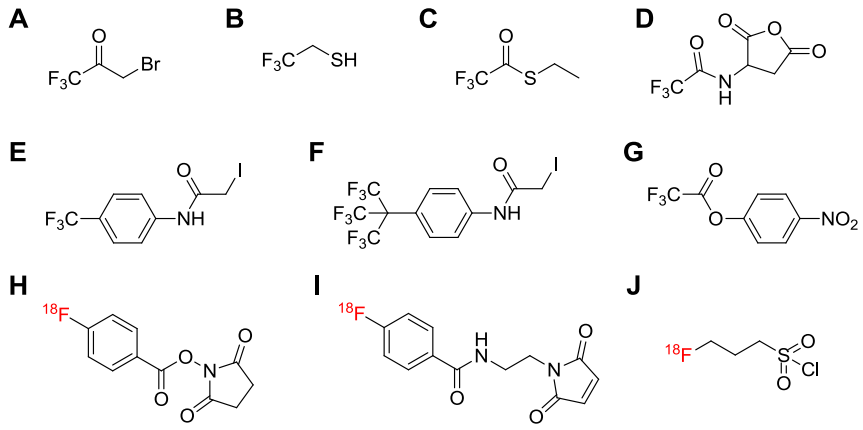


Figure 1.7: Commercially available reagents for fluorine labelling: (A) 3-Bromo-1,1,1-trifluoro-acetone; (B) 2,2,2-trifluoroethanthiol; (C) *S*-ethyl-trifluorothioacetate; (D) trifluoroacetamido-succinic anhydride; (E) *N*-[4-(trifluoromethyl)phenyl]iodoacetamide; (F) 4-(perfluorotertbutyl)- phenyliodoacetamide; (G) *p*-nitrophenyl trifluoroacetate; (H) *N*-(hydroxysuccimidyl)-^{18}F-fluoro-benzoate (SFB); (I) *N*-[2-(4-^{18}F-fluorobenzamido)-ethyl]maleimide (FBEM) and; (J) 3-^{18}F-fluoro- propanesulfonyl chloride (FPSC).[90, 101]

1.2.3.2 Biosynthetic labelling

Biosynthetic labelling of proteins can be achieved through two methods. The first of these requires the use of an orthogonal suppressor-tRNA/aminoacyl-tRNA synthetase pair. An amber suppressor codon is usually introduced to the desired site by mutagenesis. With the expression of the orthogonal pair, unnatural fluorinated amino acids can be incorporated in specific sites. This strategy is widely used for incorporation of fluorinated derivatives of aromatic amino acids such as 4-fluorophenylalanine,[102] 4-trifluoromethylphenylalanine[103] and *O*-nitrobenzyl-2-fluorotyrosine.[104] Although this technique was limited by low expression yield and incorporation efficiency in the past,[102] recent techniques such as auto-induction have greatly solved the problem in which protein expression is regulated under metabolic control of the host cells.[105] On the other hand, bacterial strains that are auxotrophic to a certain amino acid can be employed for biosynthetic labelling. As the

culture medium is the only source for that amino acid, a fluorinated derivative of that amino acid can be added to the culture medium for incorporation into proteins. This method is mainly applied to aliphatic amino acids like trifluoromethionine (Tfm),[106] 5-fluoroleucine[107] and 4,4,4-trifluorovaline.[108] A simpler strategy using inhibitors such as glyphosate for suppressing enzymes responsible for aromatic amino acid synthesis is also available.[109-111]

1.3 Molecular Imaging of Virus Infection

1.3.1 General background

The infection pathway of viruses is one of the most studied areas in virology as understanding how viruses replicate in cells can shed light on potential disease treatments. It is also particularly useful when viruses are utilised as drug or vaccine delivery vehicles.[112, 113] In order to track the internalisation and subsequent translocation during virus infection, it is essential to label the virus using different techniques.

Optical imaging is the most common strategy for tracking viruses as it can be easily performed on a cellular level with high sensitivity.[114-116] There are two main strategies for labelling viruses: fusing the virus particle with fluorescent proteins[117] or direct labelling of the virus with small fluorescent molecules.[115] The major problem of fluorescent proteins like GFP comes from their bulky size, which can affect the viral structure and hence infectivity. On the other hand, labelling the virus capsid proteins[118] or lipid membrane[119] with small dye molecules not only avoids this problem, but can also omit the troublesome gene engineering step.

1.3.2 Fluorescent dyes for virus labelling

There are many commercially available dyes that are ready to use for virus labelling, such as fluorescein, rhodamine, Oregon green, 1,1'-dioctadecyl-3,3,3',3'-tetramethylindo-dicarbocyanine 4-chlorobenzenesulfonate salt (DiD), the Alexa Fluor® series, etc. To decide whether to use a specific dye for labelling, fluorescence properties (**Table 1.1**) of the dye are always the first thing to consider.[120] For virus labelling, however, one should also consider the effect of dye labelling on virus infectivity.[115] Indeed, this is a rarely studied field and not much effort has been spent on improving the fluorescent dyes for minimising their effect in this regard. One of the aims in this thesis is thus to develop a dye which can retain the infectivity of a virus after labelling.

Table 1.1: Fluorescence properties of dyes[120]

Property	Significance
Fluorescence absorption and emission spectra	The absorption and emission maxima of a dye can be revealed by the plots. The choice of dyes is limited by the wavelength of the laser sources available.
Stokes shift	It is the difference between the absorption and emission maxima of a dye. A large Stokes shift can eliminate spectral overlapping and hence reduce the interference in measurement.
Extinction coefficient (ε)	It is the capacity of a dye for absorbing light at a specific wavelength. It is defined by the Beer-Lamber law (A = εcl) where A is absorbance, c is molarity and I is optical path length. Fluorescence intensity of a dye is directly proportional to its extinction coefficient.
Quantum yield (Φ)	It is defined as the number of fluorescence photons released per excitation photon absorbed. Fluorescence intensity of a dye is directly proportional to its quantum yield.
Photobleaching	An excited fluorophore can be destructed by non-specific reactions with surrounding molecules such as reactive oxygen species, resulting in an irreversible loss of fluorescence signal. It depends on the intensity and duration of the excitation light, and also varies in different dyes.

1.3.3 Virus models proposed for testing the newly developed dyes

1.3.3.1 Human adenovirus (AV)

Human AV is a medium-sized (120 nm), non-enveloped, icosahedral dsDNA-virus possessing a nucleocapsid.[121, 122] The viral capsid is formed by 240 hexon trimeric proteins (108 kDa) and 12 penton base proteins (63 kDa) each linked with a fiber protein (62 kDa) (**Figure 1.8**). The hexon protein is responsible for supporting the viral structure, while the penton and fiber proteins are vital for virus infection.[121] 57 serotypes have been discovered in humans where serotype 5 (AV-5) is the most studied. [121]

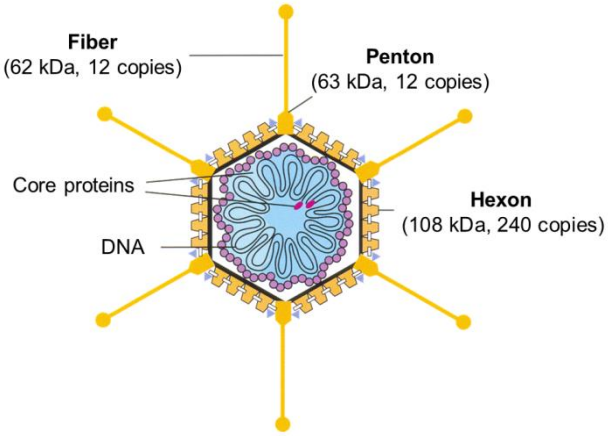


Figure 1.8: The structure of human adenovirus.[121]

AV infects the target cell by binding to the coxsackie-adenovirus receptor (CAR) with the fiber knob region, which then promotes the interactions between the penton-base Arg-Gly-Asp (RGD) motif and the cellular $\alpha_v\beta$ integrins, eventually leading to clathrin-mediated endocytosis of virus.[123] The acidic pH in the endosome causes the nucleocapsid to disassemble, releasing the viral DNA to the nucleus.

Although the virus can cause respiratory and gastrointestinal infections in human,[121] it is more well-known as a gene therapy vector for cancer cells.[124-126] AV vectors can hold a large

segment of DNA (~7.5 kb) and it can be easily manipulated and produced in high titers.[127] Also, it can efficiently transduce both dividing and non-dividing cells, which is crucial for a viral vector.[127] Since AV has been used extensively for gene therapy, its infection pathway has been well studied with fluorescence microscopy.[128, 129] It can thus act as a perfect model for testing novel fluorescent dyes.

1.3.3.2 Dengue virus (DENV)

DENV is a 50 nm enveloped (+)ssRNA virus belonging to the *Flaviviridae* family. It has four serotypes (DENV-1, DENV-2, DENV-3 and DENV-4) which share 70-80% sequence homology.[130] DENV is enveloped with a lipid membrane which contains 180 copies of both membrane protein (M) and envelope protein (E), forming a T = 3 symmetry (**Figure 1.9**).[69, 128] The E protein is essential for the virus to attach the cell surface and fuse with the membrane for escaping.[131, 132] The M protein exists as membrane precursor protein (prM) in immature virus particle and it forms a heterodimer with the E protein.[133] The capsid protein (C), on the other hand, forms the nucleocapsid with unknown structure which surrounds the viral RNA.[133]

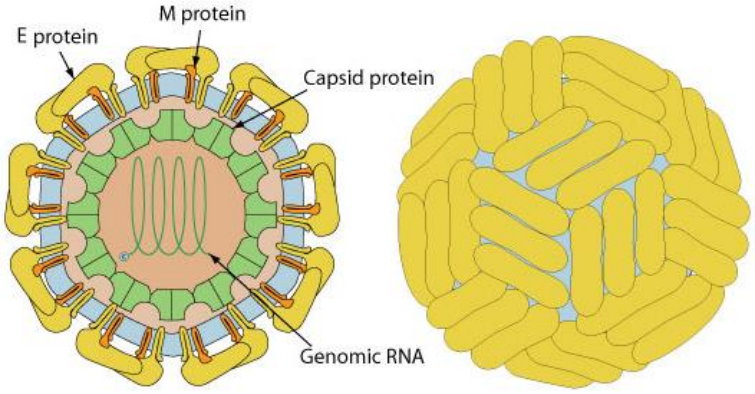


Figure 1.9: The structure of DENV.[134]

DENV is known to cause dengue fever (DF) which can be classified into mild and severe form. Mild DF has syndromes similar to normal fever,[135] while severe DF is characterised by thrombocytopenia, haemorrhagic manifestations and leakage of intravascular fluid into interstitial spaces, which can cause death.[136] It has been a challenge for scientists to understand why some of the patients develop severe DF instead of mild DF when they are infected by DENV. Recently, some research has shown that pre-existing antibodies can enhance the binding and infection of viruses to the target cells, which led to the hypothesis of the antibody-dependent enhancement (ADE) effect.[137, 138]

Although the binding factor responsible for DENV infection is not fully understood,[139] the infection pathway of DENV in the absence of antibodies has been well-studied. DENV enters the cell *via* clathrin-mediated endocytosis, and escapes from the endosomes by membrane fusion.[134, 135] The nucleocapsid released from the fusion process is then dissociated to release the viral RNA, which is subsequently translated and processed by proteases in the endoplasmic reticulum (ER) (**Figure 1.10A**).[140]

When the body is infected by a DENV serotype for the first time, the immune system will generate antibodies specifically targeting the E and prM proteins on the surface of that serotype.[141-144] Some of these antibodies, however, may cross-react with other serotypes without neutralising them, in turn providing a shortcut for them to enter the cell (i.e. ADE).[145] Previous studies have shown that antibody-bound virus particles can enter the immune cells *via* the Fcγ receptor (FcγR) on the cell surface (**Figure 1.10B**).[146-148] Nevertheless, the exact entry route and fate of the antibody-bound particles are not fully understood.[130] An imaging reagent for the virus with little effect on its infectivity is thus urgently needed for understanding more about ADE and the causes of severe DF.

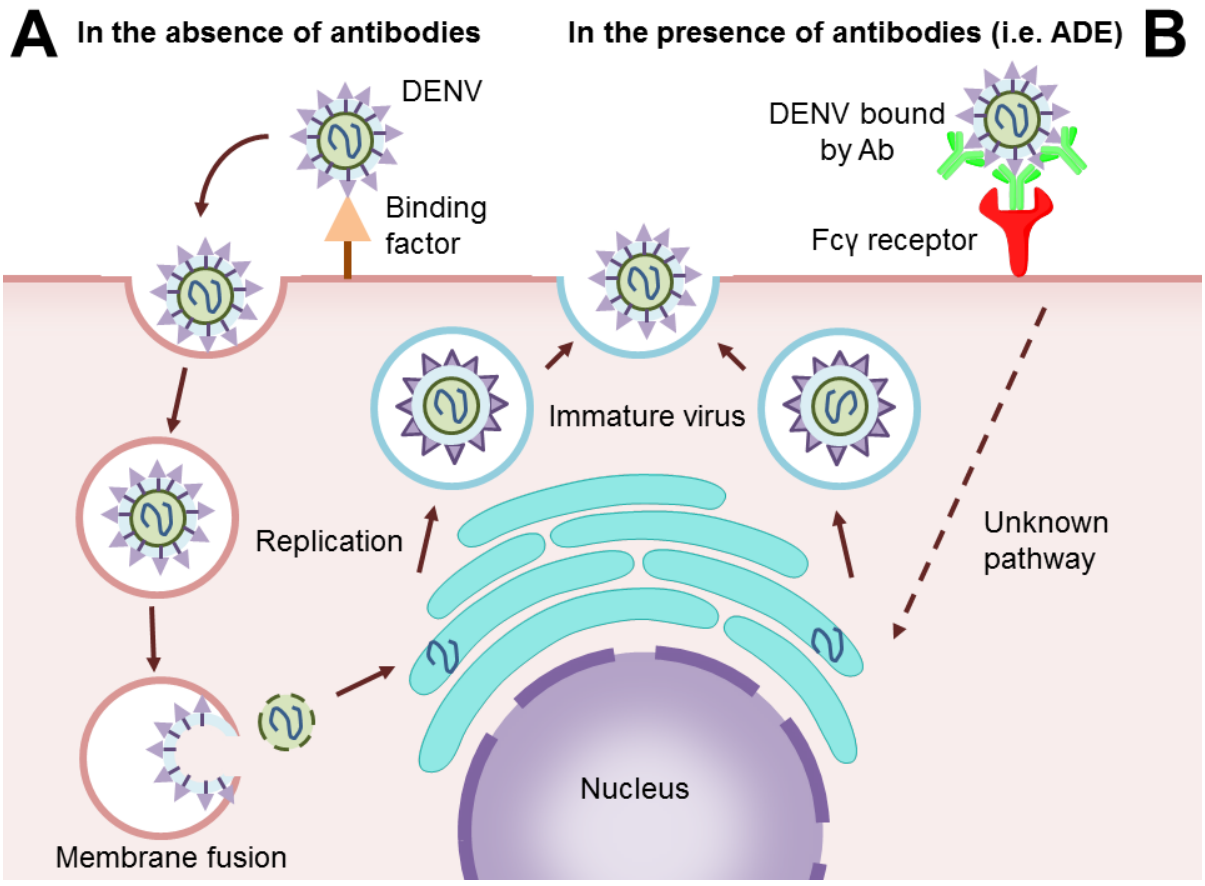


Figure 1.10: Life cycle of DENV in a host cell. (A) In the absence of antibodies, DENV enters the cell by interacting with a binding factor. The virus then escapes *via* membrane fusion and replicates in endoplasmic reticulum-associated membrane compartments; (B) During ADE, the antibody-bound DENV interacts with the Fcγ receptor on the cell membrane. The entry mechanism and cellular localisation of DENV is however unknown.

1.4 Project Aims

In summary, there are three goals in this thesis:

(i) To develop new fluorine-labelling methods for VLPs that can avoid the concomitant introduction of bulky prosthetic group

An electrophilic fluorine source Selectfluor™ was proposed for introducing single fluorine atoms on Qβ VLPs in aqueous conditions. Preliminary studies in our group have revealed that fluorination can be achieved on both dehydroalanine (i.e. 'tag-and-modify' approach) and tyrosine residues (i.e. tyrosine labelling approach) on proteins.[149] With the availability of ^{18}F-Selectfluor™,[150] it is believed that these methods can provide a novel way for labelling Qβ VLPs without affecting its structure.

(*Part I: Introduction of single F atoms on virus-like particles with Selectfluor™*)

(ii) To establish a detection system for monitoring the disassembly of fluorine-labelled VLPs

^{19}F is NMR-active with a wide chemical shift range, and it has a low background when biological samples are analysed.[90] It was proposed that ^{19}F NMR spectroscopy can be used as a tool for monitoring the structural changes of ^{19}F-labelled Qβ VLPs. Introduction of magnetically equivalent fluorine atoms onto the particles can be achieved by incorporation of the unnatural amino acid trifluoromethionine.[106, 151, 152] It is believed that the particle, which consists of 180 monomers, will have a different spectrum from the monomeric or oligomeric species, in turn indicating its disassembly and hence the release of cargos.

(*Part II: Monitoring the disassembly of virus-like particles by ^{19}F NMR spectroscopy*)

(iii) To develop novel fluorescent dyes which have a minimal effect on virus infectivity after virus labelling

The surface net charge of different viruses has been revealed to be a crucial component for guiding viral infection.[153-155] Introduction of a 2-imino-2-methoxyethyl (IME) coupling group on fluorescent dyes was proposed as it is known to retain the pKaH of the side chain of lysine residue after conjugation, hence the surface net charge of the protein.[156] The synthesised dyes will be tested with both adenovirus and dengue virus.

(*Part III: Development of 2-imino-2-methoxyethyl (IME) fluorescent dyes for retaining the infectivity of labelled viruses*)

1.5 References

1. Massoud, T.F. & Gambhir, S.S. Molecular imaging in living subjects: seeing fundamental biological processes in a new light. *Genes Dev* **17**, 545-80 (2003).
2. Weissleder, R. & Pittet, M.J. Imaging in the era of molecular oncology. *Nature* **452**, 580-9 (2008).
3. Weissleder, R. Molecular imaging in cancer. *Science* **312**, 1168-71 (2006).
4. Chen, Z.-Y. et al. Advance of Molecular Imaging Technology and Targeted Imaging Agent in Imaging and Therapy. *BioMed Research International* **2014**, 12 (2014).
5. Niu, G. & Chen, X. The Role of Molecular Imaging in Drug Delivery. *Drug delivery (London, England. 2007)* **3**, 109-113 (2009).
6. Couillard-Despres, S. et al. In vivo optical imaging of neurogenesis: watching new neurons in the intact brain. *Mol Imaging* **7**, 28-34 (2008).
7. New, S.E.P. & Aikawa, E. Molecular Imaging Insights Into Early Inflammatory Stages of Arterial and Aortic Valve Calcification. *Circulation Research* **108**, 1381-1391 (2011).
8. Atreya, R. & Goetz, M. Molecular imaging in gastroenterology. *Nat Rev Gastroenterol Hepatol* **10**, 704-712 (2013).
9. Gomes, C.M., Abrunhosa, A.J., Ramos, P. & Pauwels, E.K.J. Molecular imaging with

SPECT as a tool for drug development. *Advanced Drug Delivery Reviews* **63**, 547-554 (2011).

10. Byrne, A.T. et al. Vascular-targeted photodynamic therapy with BF2-chelated Tetraaryl-Azadipyrromethene agents: a multi-modality molecular imaging approach to therapeutic assessment. *Br J Cancer* **101**, 1565-1573 (2009).
11. Chudakov, D.M., Lukyanov, S. & Lukyanov, K.A. Fluorescent proteins as a toolkit for in vivo imaging. *Trends Biotechnol* **23**, 605-13 (2005).
12. Contag, C.H. et al. Visualizing gene expression in living mammals using a bioluminescent reporter. *Photochem Photobiol* **66**, 523-31 (1997).
13. Bhaskar, S. et al. Multifunctional Nanocarriers for diagnostics, drug delivery and targeted treatment across blood-brain barrier: perspectives on tracking and neuroimaging. *Part Fibre Toxicol* **7**, 3 (2010).
14. James, M.L. & Gambhir, S.S. A Molecular Imaging Primer: Modalities, Imaging Agents, and Applications. *Physiological Reviews* **92**, 897-965 (2012).
15. Willmann, J.K., van Bruggen, N., Dinkelborg, L.M. & Gambhir, S.S. Molecular imaging in drug development. *Nat Rev Drug Discov* **7**, 591-607 (2008).
16. Xu, H. et al. Design, synthesis, and characterization of a dual modality positron emission tomography and fluorescence imaging agent for monoclonal antibody tumor-targeted imaging. *Journal of medicinal chemistry* **50**, 4759-4765 (2007).
17. Anderson, C.J. & Ferdani, R. Copper-64 Radiopharmaceuticals for PET Imaging of Cancer: Advances in Preclinical and Clinical Research. *Cancer Biotherapy and Radiopharmaceuticals* **24**, 379-393 (2009).
18. Xavier, C. et al. Synthesis, preclinical validation, dosimetry, and toxicity of 68Ga-NOTA-anti-HER2 nanobodies for iPET imaging of HER2 receptor expression in cancer. *Journal of Nuclear Medicine* **54**, 776-784 (2013).
19. M Catafau, A. & Bullich, S. Molecular imaging PET and SPECT approaches for improving productivity of antipsychotic drug discovery and development. *Current medicinal chemistry* **20**, 378-388 (2013).
20. Phelps, M.E. Positron emission tomography provides molecular imaging of biological processes. *Proceedings of the National Academy of Sciences* **97**, 9226-9233 (2000).
21. Halldin, C., Gulyas, B., Langer, O. & Farde, L. Brain radioligands--state of the art and new trends. *Q J Nucl Med* **45**, 139-52 (2001).
22. Afshar-Oromieh, A., Haberkorn, U., Eder, M., Eisenhut, M. & Zechmann, C. [68Ga]Gallium-labelled PSMA ligand as superior PET tracer for the diagnosis of prostate cancer: comparison with 18F-FECH. *European Journal of Nuclear Medicine and Molecular Imaging* **39**, 1085-1086 (2012).
23. Kessler, H. PET Imaging of Integrin α3βv Expression. *Theranosttiics* **1**, 48-57 (2011).

24. Ido, T. et al. Labeled 2-deoxy-D-glucose analogs. 18F-labeled 2-deoxy-2-fluoro-D-glucose, 2-deoxy-2-fluoro-D-mannose and 14C-2-deoxy-2-fluoro-D-glucose. *Journal of Labelled Compounds and Radiopharmaceuticals* **14**, 175-183 (1978).
25. Lin, T.-J. et al. Validation of an animal FDG PET imaging system for study of glioblastoma xenografted mouse and rat models. *Ann Nucl Med Sci* **23**, 77-83 (2010).
26. Hijnen, N.M., de Vries, A., Nicolay, K. & Grüll, H. Dual-isotope 111In/177Lu SPECT imaging as a tool in molecular imaging tracer design. *Contrast Media & Molecular Imaging* **7**, 214-222 (2012).
27. Blake, P., Johnson, B. & VanMeter, J.W. Positron emission tomography (PET) and single photon emission computed tomography (SPECT): clinical applications. *Journal of neuro-ophthalmology* **23**, 34-41 (2003).
28. Faulkner W. Basic Principles of MRI. OutSource (1996).
29. Medarova, Z. & Moore, A. Imaging of pancreatic cancer: a promise for early diagnosis through targeted strategies. *Cancer Ther* **2**, 329-44 (2004).
30. Padhani, A.R. & Husband, J.E. Dynamic contrast-enhanced MRI studies in oncology with an emphasis on quantification, validation and human studies. *Clinical radiology* **56**, 607-620 (2001).
31. Thorek, D.L., Chen, A.K., Czupryna, J. & Tsourkas, A. Superparamagnetic iron oxide nanoparticle probes for molecular imaging. *Annals of biomedical engineering* **34**, 23-38 (2006).
32. Schweinhardt, P., Bountra, C. & Tracey, I. Pharmacological FMRI in the development of new analgesic compounds. *NMR Biomed* **19**, 702-11 (2006).
33. Ballon, D. & Koutcher, J. Nuclear magnetic resonance: principles and applications to pathology. *Clinics in laboratory medicine* **8**, 737-751 (1988).
34. Glunde, K., Artemov, D., Penet, M.F., Jacobs, M.A. & Bhujwalla, Z.M. Magnetic resonance spectroscopy in metabolic and molecular imaging and diagnosis of cancer. *Chem Rev* **110**, 3043-59 (2010).
35. Mullins, M.E. MR spectroscopy: truly molecular imaging; past, present and future. *Neuroimaging Clinics of North America* **16**, 605-618 (2006).
36. Salmeron, B.J. & Stein, E.A. Pharmacological applications of magnetic resonance imaging. *Psychopharmacol Bull* **36**, 102-29 (2002).
37. Edgington, L.E. et al. Noninvasive optical imaging of apoptosis by caspase-targeted activity-based probes. *Nature medicine* **15**, 967-973 (2009).
38. Lakowicz, J.R. Principles of Fluorescence Spectroscopy (Springer US, 2013).
39. Stephens, D.J. & Allan, V.J. Light Microscopy Techniques for Live Cell Imaging. *Science* **300**, 82-86 (2003).
40. Chudakov, D.M., Matz, M.V., Lukyanov, S. & Lukyanov, K.A. Fluorescent proteins

and their applications in imaging living cells and tissues. *Physiological reviews* **90**, 1103-1163 (2010).

41. Yano, Y. & Matsuzaki, K. Tag-probe labeling methods for live-cell imaging of membrane proteins. *Biochim Biophys Acta* **1788**, 2124-31 (2009).
42. Siegel, R., Chan, F., Zacharias, D., Swofford, R. & Holmes, K. RY 18 Tsien, and MJ Lenardo. 2000. Measurement of molecular interactions in 19 living cells by fluorescence resonance energy transfer between variants of the 20 green fluorescent protein. *Sci STKE* (2000).
43. Payne, C.K. Imaging gene delivery with fluorescence microscopy. (2007).
44. Koyama, Y. et al. Spectral fluorescence molecular imaging of lung metastases targeting HER2/neu. *Clinical Cancer Research* **13**, 2936-2945 (2007).
45. Hoffman, R.M. Visualization of GFP-expressing tumors and metastasis in vivo. *Biotechniques* **30**, 1016-1027 (2001).
46. Liong, M. et al. Multifunctional inorganic nanoparticles for imaging, targeting, and drug delivery. *ACS nano* **2**, 889-896 (2008).
47. Harrington, K.J. et al. Effective targeting of solid tumors in patients with locally advanced cancers by radiolabeled pegylated liposomes. *Clin Cancer Res* **7**, 243-54 (2001).
48. Wang, H. et al. Integrin-targeted imaging and therapy with RGD4C-TNF fusion protein. *Molecular cancer therapeutics* **7**, 1044-1053 (2008).
49. Bagalkot, V. et al. Quantum dot-aptamer conjugates for synchronous cancer imaging, therapy, and sensing of drug delivery based on bi-fluorescence resonance energy transfer. *Nano letters* **7**, 3065-3070 (2007).
50. Ashcroft, A.E. et al. Engineering thermal stability in RNA phage capsids via disulphide bonds. *J Nanosci Nanotechnol* **5**, 2034-41 (2005).
51. Ausar, S.F., Foubert, T.R., Hudson, M.H., Vedvick, T.S. & Middaugh, C.R. Conformational stability and disassembly of Norwalk virus-like particles. Effect of pH and temperature. *J Biol Chem* **281**, 19478-88 (2006).
52. Johnson, H.R., Hooker, J.M., Francis, M.B. & Clark, D.S. Solubilization and stabilization of bacteriophage MS2 in organic solvents. *Biotechnology and Bioengineering* **97**, 224-234 (2007).
53. Wu, Y., Yang, H. & Shin, H.-J. Expression and Self Assembly of Cowpea Chlorotic Mottle Virus Capsid Proteins in Pichia pastoris and Encapsulation of Fluorescent Myoglobin. *MRS Online Proc. Libr.* **1317**, No pp. given (2011).
54. Lee, K.W. & Tan, W.S. Recombinant hepatitis B virus core particles: Association, dissociation and encapsidation of green fluorescent protein. *Journal of Virological Methods* **151**, 172-180 (2008).

55. Hu, Y., Zandi, R., Anavitarte, A., Knobler, C.M. & Gelbart, W.M. Packaging of a Polymer by a Viral Capsid: The Interplay between Polymer Length and Capsid Size. *Biophysical Journal* **94**, 1428-1436 (2008).
56. Minten, I.J. et al. CCMV capsid formation induced by a functional negatively charged polymer. *Organic & Biomolecular Chemistry* **7**, 4685-4688 (2009).
57. Cooper, A. & Shaul, Y. Recombinant viral capsids as an efficient vehicle of oligonucleotide delivery into cells. *Biochemical and Biophysical Research Communications* **327**, 1094-1099 (2005).
58. Galaway, F.A. & Stockley, P.G. MS2 Viruslike Particles: A Robust, Semisynthetic Targeted Drug Delivery Platform. *Molecular Pharmaceutics* **10**, 59-68 (2013).
59. Brasch, M. et al. Encapsulation of Phthalocyanine Supramolecular Stacks into Virus-like Particles. *J. Am. Chem. Soc.* **133**, 6878-6881 (2011).
60. Wu, W., Hsiao, S.C., Carrico, Z.M. & Francis, M.B. Genome-Free Viral Capsids as Multivalent Carriers for Taxol Delivery. *Angewandte Chemie International Edition* **48**, 9493-9497 (2009).
61. Banerjee, D., Liu, A.P., Voss, N.R., Schmid, S.L. & Finn, M.G. Multivalent Display and Receptor-Mediated Endocytosis of Transferrin on Virus-Like Particles. *Chembiochem* **11**, 1273-1279 (2010).
62. Lee, K.W., Tey, B.T., Ho, K.L. & Tan, W.S. Delivery of chimeric hepatitis B core particles into liver cells. *Journal of Applied Microbiology* **112**, 119-131 (2012).
63. Pan, Y. et al. Development of a microRNA delivery system based on bacteriophage MS2 virus-like particles. *FEBS Journal* **279**, 1198-1208 (2012).
64. Stephanopoulos, N., Carrico, Z.M. & Francis, M.B. Nanoscale integration of sensitizing chromophores and porphyrins with bacteriophage MS2. *Angew Chem Int Ed Engl* **48**, 9498-502 (2009).
65. Tissot, A.C. et al. A VLP-based vaccine against interleukin-1α protects mice from atherosclerosis. *European Journal of Immunology* **43**, 716-722 (2013).
66. Steinmetz, N.F. et al. Buckyballs Meet Viral Nanoparticles: Candidates for Biomedicine. *Journal of the American Chemical Society* **131**, 17093-17095 (2009).
67. Golmohammadi, R., Fridborg, K., Bundule, M., Valegard, K. & Liljas, L. The crystal structure of bacteriophage Q beta at 3.5 angstrom resolution. *Structure* **4**, 543-554 (1996).
68. Brown, S.D., Fiedler, J.D. & Finn, M.G. Assembly of Hybrid Bacteriophage Q beta Virus-like Particles. *Biochemistry* **48**, 11155-11157 (2009).
69. Shepherd, C.M. et al. VIPERdb: a relational database for structural virology. *Nucleic Acids Res* **34**, D386-9 (2006).
70. Strable, E. et al. Unnatural amino acid incorporation into virus-like particles.

Bioconjugate Chemistry **19**, 866-875 (2008).

71. Valegard, K., Fridborg, K. & Liljas, L. Crystallization and preliminary X-ray diffraction studies of the bacteriophage Qbeta. *Acta Crystallogr D Biol Crystallogr* **50**, 105-9 (1994).
72. Fiedler, J.D. et al. Engineered Mutations Change the Structure and Stability of a Virus-Like Particle. *Biomacromolecules* **13**, 2339-2348 (2012).
73. Rhee, J.K. et al. Colorful Virus-like Particles: Fluorescent Protein Packaging by the Q beta Capsid. *Biomacromolecules* **12**, 3977-3981 (2011).
74. Fiedler, J.D., Brown, S.D., Lau, J.L. & Finn, M.G. RNA-Directed Packaging of Enzymes within Virus-like Particles. *Angewandte Chemie-International Edition* **49**, 9648-9651 (2010).
75. Ashley, C.E. et al. Cell-specific delivery of diverse cargos by bacteriophage MS2 virus-like particles. *ACS Nano* **5**, 5729-45 (2011).
76. Goldinger, S.M. et al. Nano-particle vaccination combined with TLR-7 and -9 ligands triggers memory and effector CD8+T-cell responses in melanoma patients. *European Journal of Immunology* **42**, 3049-3061 (2012).
77. Braun, M. et al. Virus-like particles induce robust human T-helper cell responses. *European Journal of Immunology* **42**, 330-340 (2012).
78. Speiser, D.E. et al. Memory and effector CD8 T-cell responses after nanoparticle vaccination of melanoma patients. *J Immunother* **33**, 848-58 (2010).
79. Senti, G. et al. Use of A-type CpG oligodeoxynucleotides as an adjuvant in allergen-specific immunotherapy in humans: a phase I/IIa clinical trial. *Clin Exp Allergy* **39**, 562-70 (2009).
80. Storni, T. et al. Nonmethylated CG Motifs Packaged into Virus-Like Particles Induce Protective Cytotoxic T Cell Responses in the Absence of Systemic Side Effects. *The Journal of Immunology* **172**, 1777-1785 (2004).
81. Bachmann, M.W., CH), Storni, Tazio (VIGANELLO, CH), Maurer, Patrik (WINTERTHUR, CH), Tissot, Alain (ZURICH, CH), Schwarz, Katrin (SCHLIEREN, CH), Meijerink, Edwin (ZURICH, CH), Lipowsky, Gerd (ZURICH, CH), Pumpens, Paul (RIGA, LV), Cielens, Indulis (RIGA, LV), Renhofa, Regina (RIGA, LV). (CYTOS BIOTECNOLOGY AG (ZURICH - SCHLIEREN, CH), United States, 2012).
82. Bachmann, M.F. et al. 190 pp. (Cytos Biotechnology A.-G., Switz. . 2004).
83. Bachmann, M.F. et al. 197 pp. (Cytos Biotechnology A.-G., Switz. . 2004).
84. Bachmann, M.F.S., CH), Proba, Karl G. (Zurich, CH), Maurer, Patrik (Winterthur, CH), Meijerink, Edwin (Zurich, CH), Schwarz, Katrin (Schlieren, CH),. (Cytos Biotechnology AG (Zurich-Schlieren, CH), United States, 2008).

85. Lau, J.L. et al. Evolution and Protein Packaging of Small-Molecule RNA Aptamers. *ACS Nano* **5**, 7722-7729 (2011).
86. Prasuhn, D.E., Jr. et al. Plasma clearance of bacteriophage Qbeta particles as a function of surface charge. *J Am Chem Soc* **130**, 1328-34 (2008).
87. Boutureira, O. et al. Fluoroglycoproteins: ready chemical site-selective incorporation of fluorosugars into proteins. *Chem. Commun. (Cambridge, U. K.)* **46**, 8142-8144 (2010).
88. Salwiczek, M., Nyakatura, E.K., Gerling, U.I.M., Ye, S. & Koksch, B. Fluorinated amino acids: compatibility with native protein structures and effects on protein-protein interactions. *Chemical Society Reviews* **41**, 2135-2171 (2012).
89. Danielson, M.A. & Falke, J.J. Use of ^{19}F NMR to Probe Protein Structure and Conformational Changes. *Annual Review of Biophysics and Biomolecular Structure* **25**, 163-195 (1996).
90. Kitevski-LeBlanc, J.L. & Prosser, R.S. Current applications of ^{19}F NMR to studies of protein structure and dynamics. *Progress in Nuclear Magnetic Resonance Spectroscopy* **62**, 1-33 (2012).
91. Buer, B.C. & Marsh, E.N.G. Fluorine: A new element in protein design. *Protein Science* **21**, 453-462 (2012).
92. Yoder, N.C. & Kumar, K. Fluorinated amino acids in protein design and engineering. *Chemical Society Reviews* **31**, 335-341 (2002).
93. Cobb, S.L. & Murphy, C.D. ^{19}F NMR applications in chemical biology. *Journal of Fluorine Chemistry* **130**, 132-143 (2009).
94. Cametti, M., Crousse, B., Metrangolo, P., Milani, R. & Resnati, G. The fluorous effect in biomolecular applications. *Chemical Society Reviews* **41**, 31-42 (2012).
95. Boehm-Sturm, P., Mengler, L., Wecker, S., Hoehn, M. & Kallur, T. In Vivo Tracking of Human Neural Stem Cells with ^{19}F Magnetic Resonance Imaging. *PLoS ONE* **6**, e29040 (2011).
96. Rydzik, A.M. et al. Monitoring Conformational Changes in the NDM-1 Metallo-β-lactamase by ^{19}F NMR Spectroscopy. *Angewandte Chemie International Edition* **53**, 3129-3133 (2014).
97. Mehta, V.D., Kulkarni, P.V., Mason, R.P., Constantinescu, A. & Antich, P.P. Fluorinated Proteins as Potential ^{19}F Magnetic Resonance Imaging and Spectroscopy Agents. *Bioconjugate Chemistry* **5**, 257-261 (1994).
98. Gerig, J.T. Fluorine NMR of proteins. *Progress in Nuclear Magnetic Resonance Spectroscopy* **26**, 293-370 (1994).
99. Tsukiji, S., Miyagawa, M., Takaoka, Y., Tamura, T. & Hamachi, I. Ligand-directed tosyl chemistry for protein labeling in vivo. *Nat Chem Biol* **5**, 341-343 (2009).

100. Luchette, P.A., Prosser, R.S. & Sanders, C.R. Oxygen as a paramagnetic probe of membrane protein structure by cysteine mutagenesis and (19)F NMR spectroscopy. *J Am Chem Soc* **124**, 1778-81 (2002).
101. Jacobson, O., Kiesewetter, D.O. & Chen, X. Fluorine-18 Radiochemistry, Labeling Strategies and Synthetic Routes. *Bioconjugate Chemistry* **26**, 1-18 (2015).
102. Furter, R. Expansion of the genetic code: site-directed p-fluoro-phenylalanine incorporation in Escherichia coli. *Protein Sci* **7**, 419-26 (1998).
103. Jackson, J.C., Hammill, J.T. & Mehl, R.A. Site-Specific Incorporation of a ^{19}F-Amino Acid into Proteins as an NMR Probe for Characterizing Protein Structure and Reactivity. *Journal of the American Chemical Society* **129**, 1160-1166 (2007).
104. Wilkins, B.J. et al. Site-Specific Incorporation of Fluorotyrosines into Proteins in Escherichia coli by Photochemical Disguise. *Biochemistry* **49**, 1557-1559 (2010).
105. Hammill, J.T., Miyake-Stoner, S., Hazen, J.L., Jackson, J.C. & Mehl, R.A. Preparation of site-specifically labeled fluorinated proteins for ^{19}F-NMR structural characterization. *Nat. Protocols* **2**, 2601-2607 (2007).
106. Duewel, H., Daub, E., Robinson, V. & Honek, J.F. Incorporation of Trifluoromethionine into a Phage Lysozyme: Implications and a New Marker for Use in Protein ^{19}F NMR†. *Biochemistry* **36**, 3404-3416 (1997).
107. Feeney, J. et al. ^{19}F Nuclear Magnetic Resonance Chemical Shifts of Fluorine Containing Aliphatic Amino Acids in Proteins: Studies on Lactobacillus casei Dihydrofolate Reductase Containing (2S,4S)-5-Fluoroleucine‖. *Journal of the American Chemical Society* **118**, 8700-8706 (1996).
108. Wang, P., Fichera, A., Kumar, K. & Tirrell, D.A. Alternative Translations of a Single RNA Message: An Identity Switch of (2S,3R)-4,4,4-Trifluorovaline between Valine and Isoleucine Codons. *Angewandte Chemie International Edition* **43**, 3664-3666 (2004).
109. Hyean-Woo, L. et al. ^{19}F NMR Investigation of F1-ATPase of Escherichia coli Using Fluorotryptophan Labeling. *Journal of Biochemistry* **127**, 1053-1056 (2000).
110. Li, C. et al. ^{19}F NMR Studies of α-Synuclein Conformation and Fibrillation. *Biochemistry* **48**, 8578-8584 (2009).
111. Kitevski-LeBlanc, J.L., Al-Abdul-Wahid, M.S. & Prosser, R.S. A Mutagenesis-Free Approach to Assignment of ^{19}F NMR Resonances in Biosynthetically Labeled Proteins. *Journal of the American Chemical Society* **131**, 2054-2055 (2009).
112. Wu, Z., Tang, L.-J., Zhang, X.-B., Jiang, J.-H. & Tan, W. Aptamer-Modified Nanodrug Delivery Systems. *ACS Nano* **5**, 7696-7699 (2011).
113. Daya, S. & Berns, K.I. Gene Therapy Using Adeno-Associated Virus Vectors. *Clinical Microbiology Reviews* **21**, 583-593 (2008).

114. Rust, M.J., Lakadamyali, M., Zhang, F. & Zhuang, X. Assembly of endocytic machinery around individual influenza viruses during viral entry. *Nat Struct Mol Biol* **11**, 567-573 (2004).
115. Seisenberger, G. et al. Real-time single-molecule imaging of the infection pathway of an adeno-associated virus. *Science* **294**, 1929-32 (2001).
116. Pereira, C.F., Rossy, J., Owen, D.M., Mak, J. & Gaus, K. HIV taken by STORM: Super-resolution fluorescence microscopy of a viral infection. *Virology Journal* **9**, 84-84 (2012).
117. Giepmans, B.N., Adams, S.R., Ellisman, M.H. & Tsien, R.Y. The fluorescent toolbox for assessing protein location and function. *Science* **312**, 217-24 (2006).
118. Suomalainen, M. et al. Microtubule-dependent plus- and minus end-directed motilities are competing processes for nuclear targeting of adenovirus. *J Cell Biol* **144**, 657-72 (1999).
119. Ayala-Nunez, N.V., Wilschut, J. & Smit, J.M. Monitoring virus entry into living cells using DiD-labeled dengue virus particles. *Methods* **55**, 137-43 (2011).
120. Johnson, I. The Molecular Probes Handbook: A Guide to Fluorescent Probes and Labeling Technologies, 11th Edition (Life Technologies Corporation, 2010).
121. Fields, B.N., Knipe, D.M. & Howley, P.M. Fields' Virology (Wolters Kluwer Health/Lippincott Williams & Wilkins, 2007).
122. Pearce, O.M.T. et al. Glycoviruses: Chemical glycosylation retargets adenoviral gene transfer. *Angewandte Chemie-International Edition* **44**, 1057-1061 (2005).
123. Wickham, T.J. Targeting adenovirus. *Gene Ther* **7**, 110-4 (2000).
124. Kanerva, A. et al. Targeting adenovirus to the serotype 3 receptor increases gene transfer efficiency to ovarian cancer cells. *Clin Cancer Res* **8**, 275-80 (2002).
125. Kanerva, A. & Hemminki, A. Modified adenoviruses for cancer gene therapy. *Int J Cancer* **110**, 475-80 (2004).
126. Eulitt, P.J. et al. Enhancing mda-7/IL-24 therapy in renal carcinoma cells by inhibiting multiple protective signaling pathways using sorafenib and by Ad.5/3 gene delivery. *Cancer Biol Ther* **10**, 1290-305 (2010).
127. Rein, D.T., Breidenbach, M. & Curiel, D.T. Current developments in adenovirus-based cancer gene therapy. *Future oncology (London, England)* **2**, 137-143 (2006).
128. Leopold, P.L. et al. Fluorescent Virions: Dynamic Tracking of the Pathway of Adenoviral Gene Transfer Vectors in Living Cells. *Human Gene Therapy* **9**, 367-378 (1998).
129. Nakano, M.Y. & Greber, U.F. Quantitative Microscopy of Fluorescent Adenovirus Entry. *Journal of Structural Biology* **129**, 57-68 (2000).

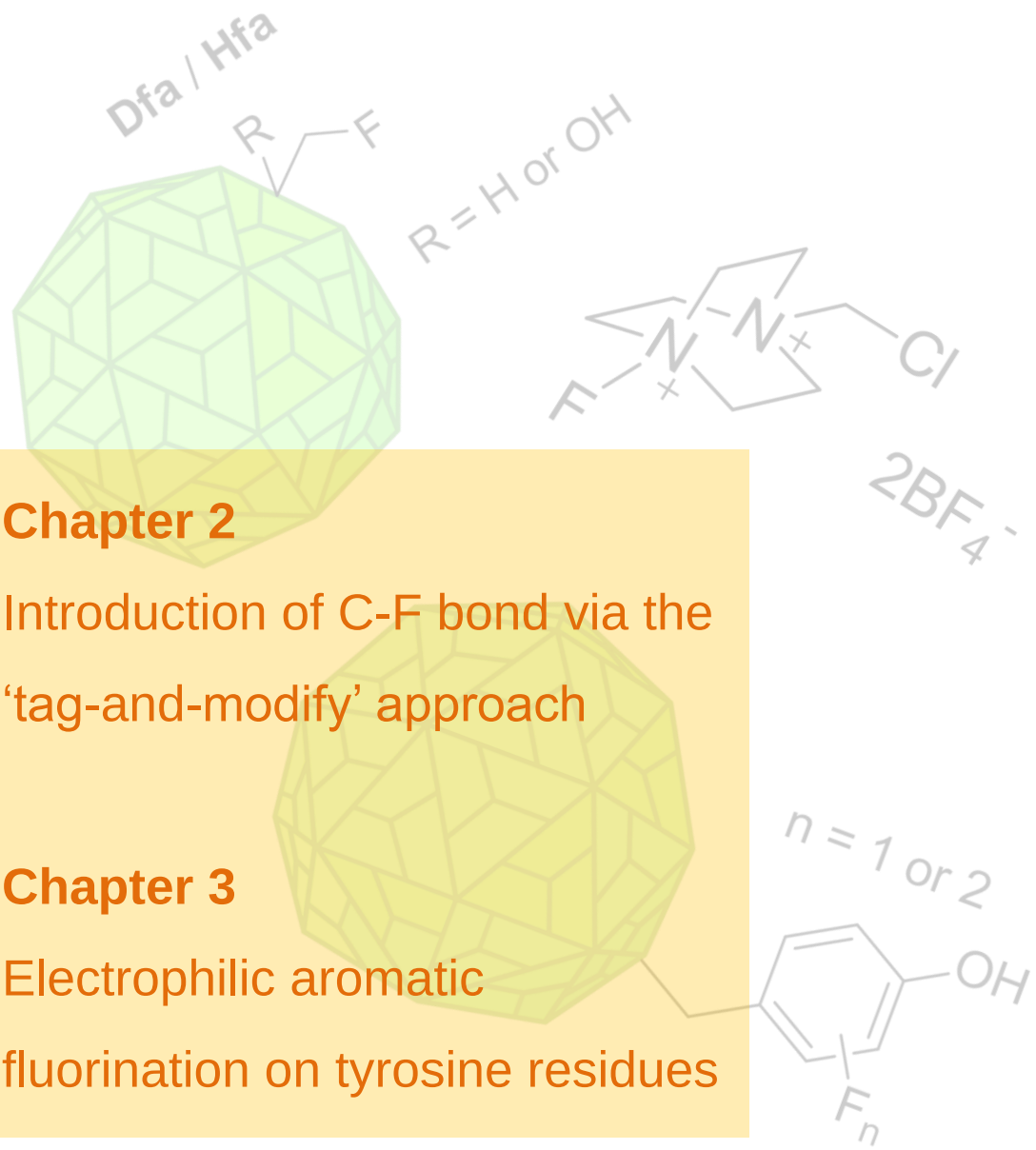
130. Flipse, J., Wilschut, J. & Smit, J.M. Molecular mechanisms involved in antibody-dependent enhancement of dengue virus infection in humans. *Traffic* **14**, 25-35 (2013).
131. Guirakhoo, F., Heinz, F.X., Mandl, C.W., Holzmann, H. & Kunz, C. Fusion activity of flaviviruses: comparison of mature and immature (prM-containing) tick-borne encephalitis virions. *J Gen Virol* **72 (Pt 6)**, 1323-9 (1991).
132. Putnak, R. et al. Dengue-1 virus envelope glycoprotein gene expressed in recombinant baculovirus elicits virus-neutralizing antibody in mice and protects them from virus challenge. *Am J Trop Med Hyg* **45**, 159-67 (1991).
133. Mukhopadhyay, S., Kuhn, R.J. & Rossmann, M.G. A structural perspective of the flavivirus life cycle. *Nat Rev Microbiol* **3**, 13-22 (2005).
134. Hulo, C. et al. ViralZone: a knowledge resource to understand virus diversity. *Nucleic Acids Res* **39**, D576-82 (2011).
135. Sabin, A.B. Research on dengue during World War II. *Am J Trop Med Hyg* **1**, 30-50 (1952).
136. Gubler, D.J. Dengue and dengue hemorrhagic fever. *Clin Microbiol Rev* **11**, 480-96 (1998).
137. Midgley, C.M. et al. An In-Depth Analysis of Original Antigenic Sin in Dengue Virus Infection. *Journal of Virology* **85**, 410-421 (2011).
138. Halstead, S.B., Mahalingam, S., Marovich, M.A., Ubol, S. & Mosser, D.M. Intrinsic antibody-dependent enhancement of microbial infection in macrophages: disease regulation by immune complexes. *The Lancet Infectious Diseases* **10**, 712-722 (2010).
139. Rodenhuis-Zybert, I.A., Wilschut, J. & Smit, J.M. Dengue virus life cycle: viral and host factors modulating infectivity. *Cell Mol Life Sci* **67**, 2773-86 (2010).
140. Gubler DJ, Kuno G, Markoff L. Flaviviruses. In: Knipe DM, Howley PM, editors. Fields Virology, 5th edn. Philadelphia, PA: Wolters Kluwer and Lippincott Williams; 2007, pp. 1154–1252.
141. Beltramello, M. et al. The human immune response to Dengue virus is dominated by highly cross-reactive antibodies endowed with neutralizing and enhancing activity. *Cell Host Microbe* **8**, 271-83 (2010).
142. Mathew, A. et al. B-cell responses during primary and secondary dengue virus infections in humans. *J Infect Dis* **204**, 1514-22 (2011).
143. Schieffelin, J.S. et al. Neutralizing and non-neutralizing monoclonal antibodies against dengue virus E protein derived from a naturally infected patient. *Virol J* **7**, 28 (2010).
144. Dejnirattisai, W. et al. Cross-reacting antibodies enhance dengue virus infection in

humans. *Science* **328**, 745-8 (2010).

145. Guzman, M.G. & Vazquez, S. The complexity of antibody-dependent enhancement of dengue virus infection. *Viruses* **2**, 2649-62 (2010).
146. Whitehead, S.S., Blaney, J.E., Durbin, A.P. & Murphy, B.R. Prospects for a dengue virus vaccine. *Nat Rev Microbiol* **5**, 518-28 (2007).
147. Rodrigo, W.W., Jin, X., Blackley, S.D., Rose, R.C. & Schlesinger, J.J. Differential enhancement of dengue virus immune complex infectivity mediated by signaling-competent and signaling-incompetent human Fcgamma RIA (CD64) or FcgammaRIIA (CD32). *J Virol* **80**, 10128-38 (2006).
148. Moi, M.L., Lim, C.K., Takasaki, T. & Kurane, I. Involvement of the Fc gamma receptor IIA cytoplasmic domain in antibody-dependent enhancement of dengue virus infection. *J Gen Virol* **91**, 103-11 (2010).
149. Boutureira, O, Schombs, M., et al. (Unpublished results)
150. Teare, H. et al. Radiosynthesis and Evaluation of [18F]Selectfluor bis(triflate). *Angewandte Chemie International Edition* **49**, 6821-6824 (2010).
151. de Visser, P.C. et al. A novel, base-labile fluorous amine protecting group: synthesis and use as a tag in the purification of synthetic peptides. *Tetrahedron Letters* **44**, 9013-9016 (2003).
152. Holzberger, B., Rubini, M., Möller, H.M. & Marx, A. A Highly Active DNA Polymerase with a Fluorous Core. *Angewandte Chemie International Edition* **49**, 1324-1327 (2010).
153. Gardner, C.L. et al. Deliberate attenuation of chikungunya virus by adaptation to heparan sulfate-dependent infectivity: a model for rational arboviral vaccine design. *PLoS Negl Trop Dis* **8**, e2719 (2014).
154. Kobayashi, Y. & Suzuki, Y. Compensatory evolution of net-charge in influenza A virus hemagglutinin. *PLoS One* **7**, e40422 (2012).
155. Shayakhmetov, D.M. & Lieber, A. Dependence of adenovirus infectivity on length of the fiber shaft domain. *J Virol* **74**, 10274-86 (2000).
156. Lee, Y.C., Stowell, C.P. & Krantz, M.J. 2-Imino-2-methoxyethyl 1-thioglycosides: new reagents for attaching sugars to proteins. *Biochemistry* **15**, 3956-3963 (1976).

PART I

Introduction of single F atoms on virus-like particles with Selectfluor™



CHAPTER

2 Introduction of C-F bond *via* the 'tag-and-modify' approach

2.1 Introduction

2.1.1 Selectfluor™ as a fluoro-labelling reagent for proteins

Fluorine can be introduced to organic molecules from nucleophilic 'F^-' [1] or electrophilic 'F^+' [2, 3] sources. The use of nucleophilic sources usually requires conditions incompatible with protein structure and function such as elevated temperature, organic solvents and pHs that can greatly affect the structure and function of proteins. Electrophilic sources like F-TEDA (Selectfluor™) (**Figure 2.1**), on the other hand, can be used under milder conditions and in aqueous solution.[4] Although other methods like metal-catalytic[5] or radical[6] fluorination are available, the conditions required are too harsh for proteins.

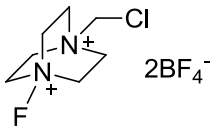


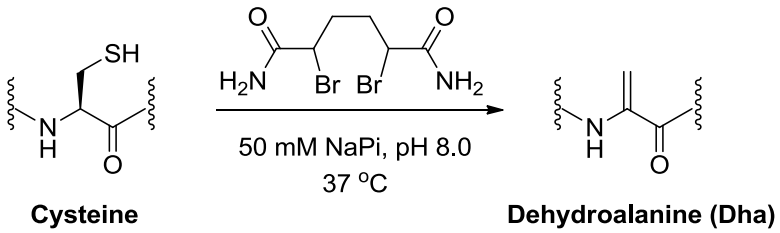
Figure 2.1: Selectfluor™ BF_4^- (IUPAC: 1-chloromethyl-4-fluorodiazoniabicyclo[2.2.2]octane bis(tetrafluoroborate))

Selectfluor™ is relatively mild, stable and non-toxic when compared with other electrophilic fluorine sources.[7] Although it is still unknown whether the exact mechanism of fluorination proceeds *via* nucleophilic S_N2 substitution or single-electron transfer (SET),[4] the features of Selectfluor™ make it a viable reagent for protein labelling. As mentioned in the

previous chapter, the direct labelling reagents available usually introduce extra chains or functional groups to the target site. Selectfluor™, on the other hand, introduces fluorine atoms directly to organic molecules. Furthermore, the ^{18}F-labelled Selectfluor™ has been successfully synthesised recently,[8] in turn greatly enhancing the application of this reagent to PET, an important imaging technique for biological studies.

2.1.2 The 'tag-and-modify' approach

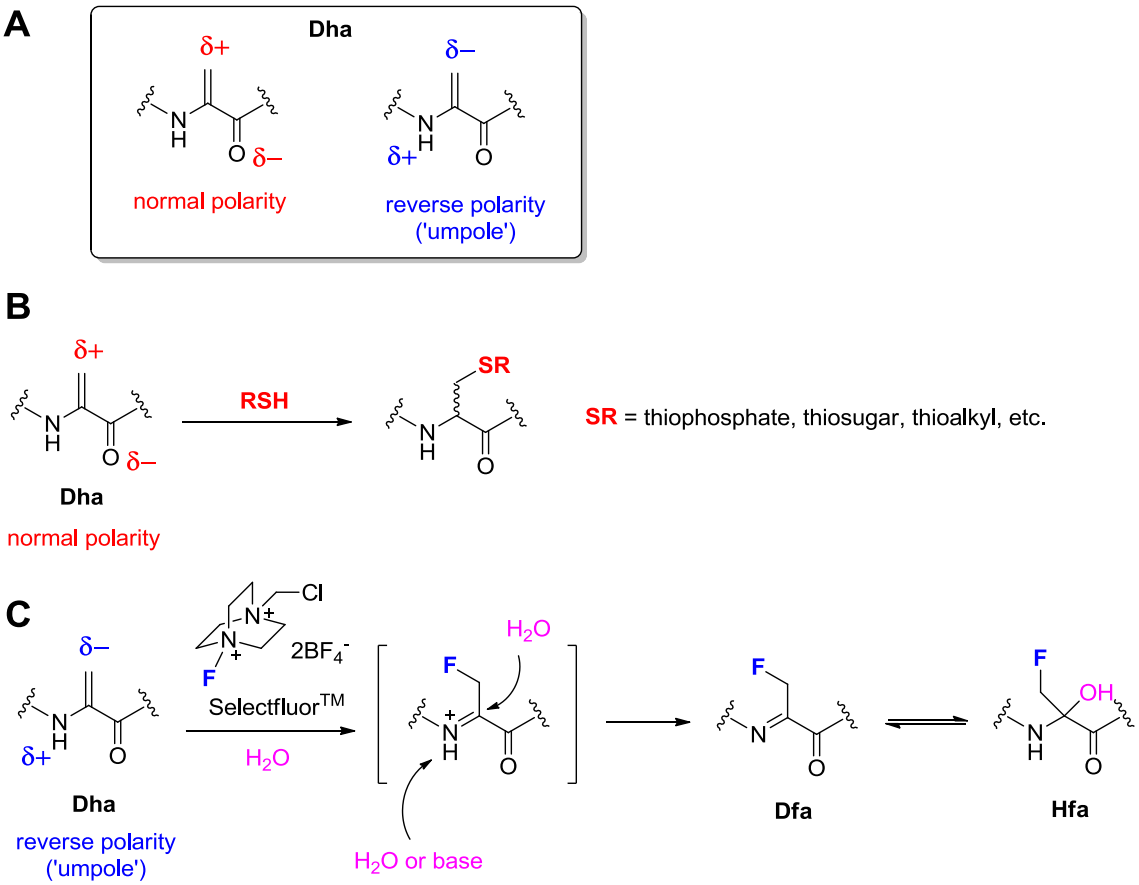
Dehydroalanine (Dha), the simplest amino acid containing an olefin, can be formed easily from cysteine residue on proteins by reagents such as α,α′-di-bromoadipyl(bis)amide (DBHDA) in alkaline buffer (**Scheme 2.1**).[9] This creates a 'tag' for subsequent installation of a wide variety of protein modifications, which can be achieved by treatment with reagents containing a thiol group. The olefin thus acts as an electrophile in this case (**Scheme 2.2B**).[10] By mutating a specific site on a protein to a cysteine residue, Dha can be formed at the desired position. The 'tag-and-modify' approach can thus become site-selective and enable mimicking of natural post-translational modifications.[11] In the past, our group successfully introduced different modifications to several proteins such as subtilisin *Bacillus lentus*,[10] histone,[12] p38α,[13] cAbLys,[14] etc. using this approach.



Scheme 2.1: Conversion of cysteine to dehydroalanine with the DBHDA reagent.

As discovered by previous group members, the electron-rich olefin on Dha can act as a soft carbon nucleophile for the fluorine electrophile Selectfluor™,[15] instead of acting as an electrophile for Michael-type additions.[9, 10, 16] The olefin thus requires a reversed polarity (i.e.

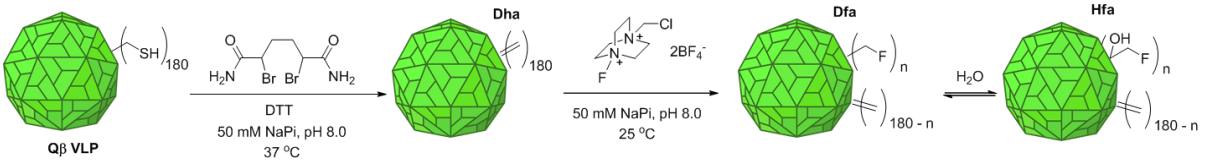
'umpole') to act as a nucleophile[17, 18] (**Scheme 2.2A**). C-F bond formation was confirmed to be on the β carbon of the small molecule model of Dha, leading to the formation of (*N*,α)-**d**ehydro-β-**f**luoro-**a**lanine (Dfa). The Dfa can further be hydrated to form α-**h**ydroxy-β-**f**luoro-**a**lanine (Hfa) in aqueous solution (**Scheme 2.2B**). The equilibrium between Dfa and Hfa was found to be structure and condition dependent, in other words, it varies from protein to protein. Furthermore, the fluorinated amino acids were found to be stable throughout isolation and purification.[15]



Scheme 2.2: (A) Dha usually acts as a conjugate electrophile (red), but the enamine moiety (blue) can also allow it to act as a nucleophile with suitable electrophiles; (B) Addition of thio-substrate on Dha; (C) C-F bond formation on β carbon of Dha can be achieved with Selectfluor™, giving a mixture of Dfa and Hfa in aqueous solution.

2.1.3 Labelling of Qβ VLP with Selectfluor™ *via* the 'tag-and-modify' approach

Previously, our group successfully introduced a fluoro-mannose on the surface of Qβ VLP.[19] However, the sugar group is relatively bulky. It is therefore imperative to develop a method for introduction of a smaller imaging group or even an atom in order to minimise the effect on the size and structure of the particle. The labelling strategy using Selectfluor™ provides a possible solution for the problem mentioned above.



Scheme 2.3: Introduction of C-F bond on Qβ VLP *via* the 'tag-and-modify' approach.

The 'tag-and-modify' approach makes use of the cysteine residue on the surface of the particle (**Scheme 2.3**). Since the two cysteine residues on Qβ monomer are engaged in disulphide bonding, it is necessary to introduce an extra cysteine residue to each monomer. Three different mutants (N10C, K16C and T18C) were proposed based on their surface accessibility, which can be easily found from the VIrus Particle ExploreR (VIPER) database (**Figure 2.2**).[20] On the other hand, a previous study has shown that the deletion of one cysteine residue in Qβ monomer (i.e. C74S) did not affect the structure and stability of the particle. The cysteine left (i.e. Cys-80) was also shown to be reactive to chemical reagents such as Ellman's reagent after dithiothreitol (DTT) treatment.[21] This mutant was thus also proposed for Dha formation.

Since ^{18}F-labelled Selectfluor™ has been developed recently, it is expected that the use of this approach can make Qβ VLP a potential candidate for use as a tracer in PET imaging. This can further increase the potential of Qβ VLP as a nano-carrier of drugs or a vaccine for presenting specific antigens as its circulation pathway can be traced easily once it is injected into the body.

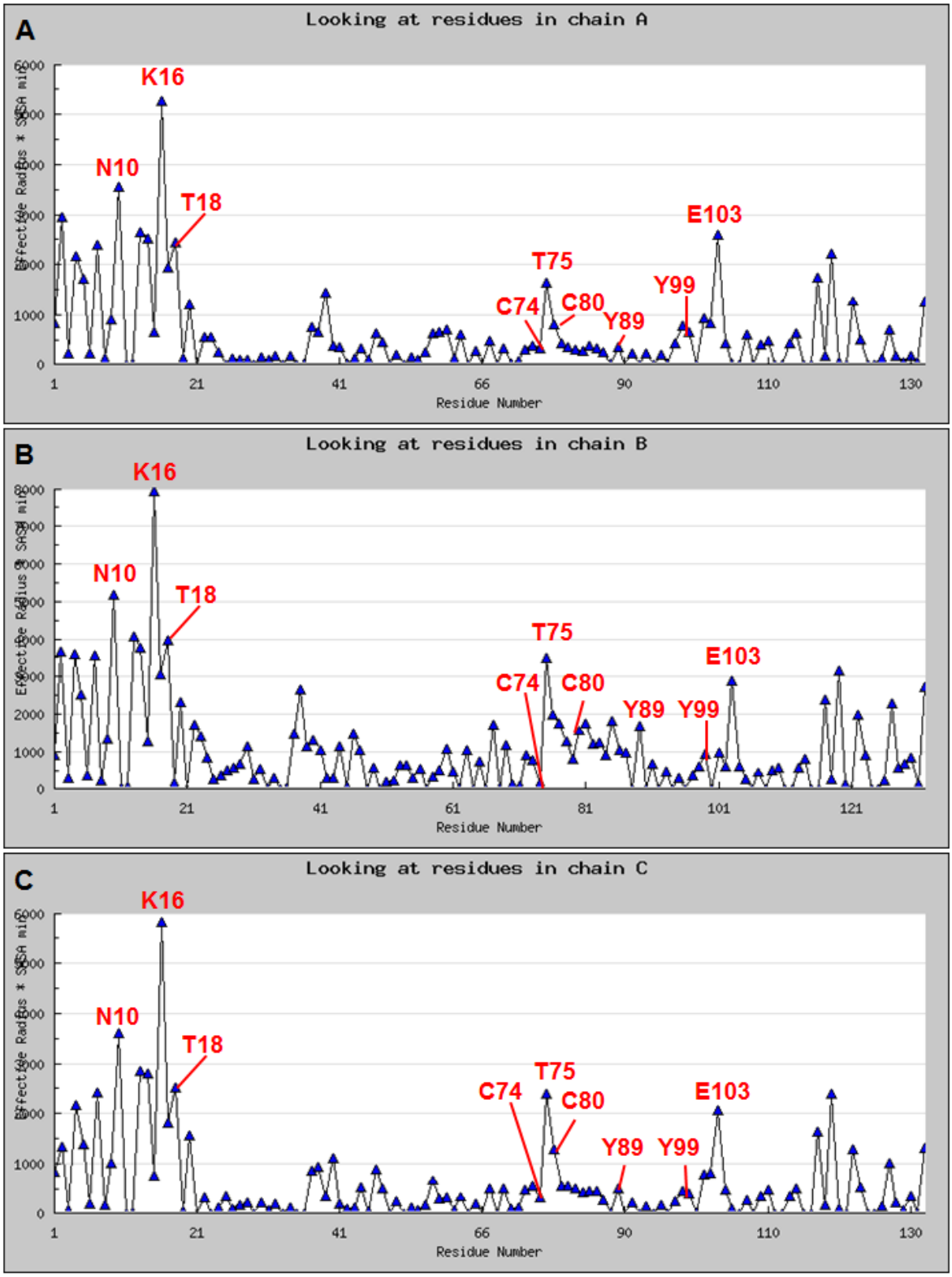


Figure 2.2: Accessible surface profile of Qβ VLP showing the residues in (A) conformer A; (B) conformer B and; (C) conformer C.[20] Accessibility is calculated based on the product of effective radius and solvent accessible surface area (SASA) of the residue. Effective radius is defined by the radius at which a residue located minus inner radius of the virus particle.

2.2 Results and Discussion

2.2.1 Introduction of C-F bond on Qβ mutants *via* the 'tag-and-modify' approach

2.2.1.1 Mutagenesis, expression and purification of Qβ VLPs

The four mutants (i.e. N10C, K16C, T18C and C74S) proposed were obtained *via* site-directed mutagenesis of the plasmid encoding the wild type (WT) Qβ. The DNA and protein sequences of the mutants are listed below (Mutated sites are labelled with red; the remaining cysteine residues are labelled with blue):

(A) Qβ N10C

Forward DNA sequence

ATGGCAAAAT TAGAGACTGT TACTTTAGGT **TGC**ATCGGGA AAGATGGAAA ACAAACTCTG
GTCCTCAATC CGCGTGGGGT AAATCCCACT AACGGCGTTG CCTCGCTTTC ACAAGCGGGT
GCAGTTCCTG CGCTGGAGAA GCGTGTTACC GTTTCGGTAT CTCAGCCTTC TCGCAATCGT
AAGAACTACA AGGTCCAGGT TAAGATCCAG AACCCGACCG CTTGCACTGC AAACGGTTCT
TGTGACCCAT CCGTTACTCG CCAGGCATAT GCTGACGTGA CCTTTTCGTT CACGCAGTAT
AGTACCGATG AGGAACGAGC TTTTGTTCGT ACAGAGCTTG CTGCTCTGCT CGCTAGTCCT
CTGCTGATCG ATGCTATTGA TCAGCTGAAC CCAGCGTATT AA

Amino acid sequence

AKLETVTLG**C** IGKDGKQTLV LNPRGVNPTN GVASLSQAGA VPALEKRVTV SVSQPSRNRK
NYKVQVKIQN PTA**C**TANGS**C** DPSVTRQAYA DVTFSFTQYS TDEERAFVRT ELAALLASPL
LIDAIDQLNP AY

Calculated average isotopic mass = 14112 (N-terminal Met cleaved)

(B) Qβ K16C

Forward DNA sequence

ATGGCAAAAT TAGAGACTGT TACTTTAGGT AACATCGGGA AAGATGGA**TG** **T**CAAACTCTG
GTCCTCAATC CGCGTGGGGT AAATCCCACT AACGGCGTTG CCTCGCTTTC ACAAGCGGGT
GCAGTTCCTG CGCTGGAGAA GCGTGTTACC GTTTCGGTAT CTCAGCCTTC TCGCAATCGT
AAGAACTACA AGGTCCAGGT TAAGATCCAG AACCCGACCG CTTGCACTGC AAACGGTTCT
TGTGACCCAT CCGTTACTCG CCAGGCATAT GCTGACGTGA CCTTTTCGTT CACGCAGTAT
AGTACCGATG AGGAACGAGC TTTTGTTCGT ACAGAGCTTG CTGCTCTGCT CGCTAGTCCT
CTGCTGATCG ATGCTATTGA TCAGCTGAAC CCAGCGTATT AA

Amino acid sequence

AKLETVTLGN IGKDG**C**QTLV LNPRGVNPTN GVASLSQAGA VPALEKRVTV SVSQPSRNRK
NYKVQVKIQN PTA**C**TANGS**C** DPSVTRQAYA DVTFSFTQYS TDEERAFVRT ELAALLASPL
LIDAIDQLNP AY

Calculated average isotopic mass = 14098 (N-terminal Met cleaved)

(C) Qβ T18C

Forward DNA sequence

ATGGCAAAAT TAGAGACTGT TACTTTAGGT AACATCGGGA AAGATGGAAA ACAA**TGT**CTG
GTCCTCAATC CGCGTGGGGT AAATCCCACT AACGGCGTTG CCTCGCTTTC ACAAGCGGGT
GCAGTTCCTG CGCTGGAGAA GCGTGTTACC GTTTCGGTAT CTCAGCCTTC TCGCAATCGT
AAGAACTACA AGGTCCAGGT TAAGATCCAG AACCCGACCG CTTGCACTGC AAACGGTTCT
TGTGACCCAT CCGTTACTCG CCAGGCATAT GCTGACGTGA CCTTTTCGTT CACGCAGTAT
AGTACCGATG AGGAACGAGC TTTTGTTCGT ACAGAGCTTG CTGCTCTGCT CGCTAGTCCT
CTGCTGATCG ATGCTATTGA TCAGCTGAAC CCAGCGTATT AA

Amino acid sequence

AKLETVTLGN IGKDGKQ**C**LV LNPRGVNPTN GVASLSQAGA VPALEKRVTV SVSQPSRNRK
NYKVQVKIQN PTA**C**TANGS**C** DPSVTRQAYA DVTFSFTQYS TDEERAFVRT ELAALLASPL
LIDAIDQLNP AY

Calculated average isotopic mass = 14125 (N-terminal Met cleaved)

(D) Qβ C74S

Forward DNA sequence

ATGGCAAAAT TAGAGACTGT TACTTTAGGT AACATCGGGA AAGATGGAAA ACAAACTCTG
GTCCTCAATC CGCGTGGGGT AAATCCCACT AACGGCGTTG CCTCGCTTTC ACAAGCGGGT
GCAGTTCCTG CGCTGGAGAA GCGTGTTACC GTTTCGGTAT CTCAGCCTTC TCGCAATCGT
AAGAACTACA AGGTCCAGGT TAAGATCCAG AACCCGACCG CTAGCACTGC AAACGGTTCT
TGTGACCCAT CCGTTACTCG CCAGGCATAT GCTGACGTGA CCTTTTCGTT CACGCAGTAT
AGTACCGATG AGGAACGAGC TTTTGTTCGT ACAGAGCTTG CTGCTCTGCT CGCTAGTCCT
CTGCTGATCG ATGCTATTGA TCAGCTGAAC CCAGCGTATT AA

Amino acid sequence

AKLETVTLGN IGKDGKQTLV LNPRGVNPTN GVASLSQAGA VPALEKRVTV SVSQPSRNRK
NYKVQVKIQN PTASTANGSC DPSVTRQAYA DVTFSFTQYS TDEERAFVRT ELAALLASPL
LIDAIDQLNP AY

Calculated average isotopic mass = 14107 (N-terminal Met cleaved)

The proteins were expressed in *E. coli* BL21(DE3) and purified by PEG8000 and sucrose gradient centrifugation based on a previously reported protocol.[22] Fractions containing the intact particles were combined according to the results from the native agarose and SDS-PAGE gels (**Figure 2.3**). Native agarose gel can be used as a preliminary tool for studying the net charge of the particles. It is also useful in cases where there is a mixture of particles and aggregates, when two different bands can be observed. SDS-PAGE gel cannot distinguish between particles and aggregates as both of them are reduced into the same monomeric species, but SDS-PAGE gel can show the purity of the fractions much clearer than native agarose gel. The fractions were then concentrated and analysed by liquid chromatography-mass spectrometry (LC-MS) (**Figure 2.4**) and dynamic light scattering (DLS) (**Figure 2.5**). The results are summarised in **Table 2.1**.

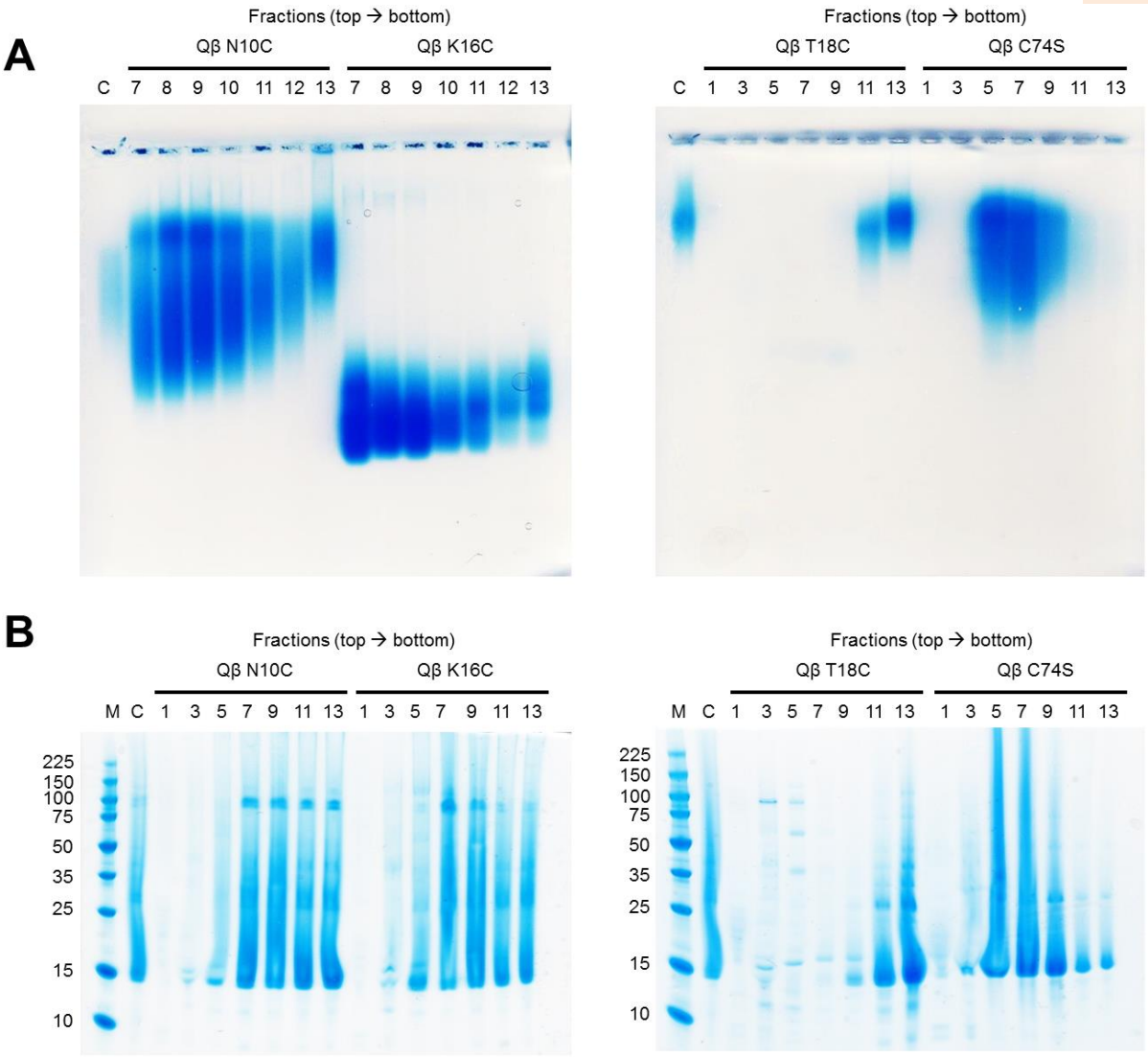


Figure 2.3: Coomassie Blue stained (A) native agarose and; (B) reducing SDS-PAGE gels showing the fractions obtained after sucrose gradient centrifugation of the Qβ N10C, K16C, T18C and C74S mutants. C: Qβ WT control; M: protein marker.

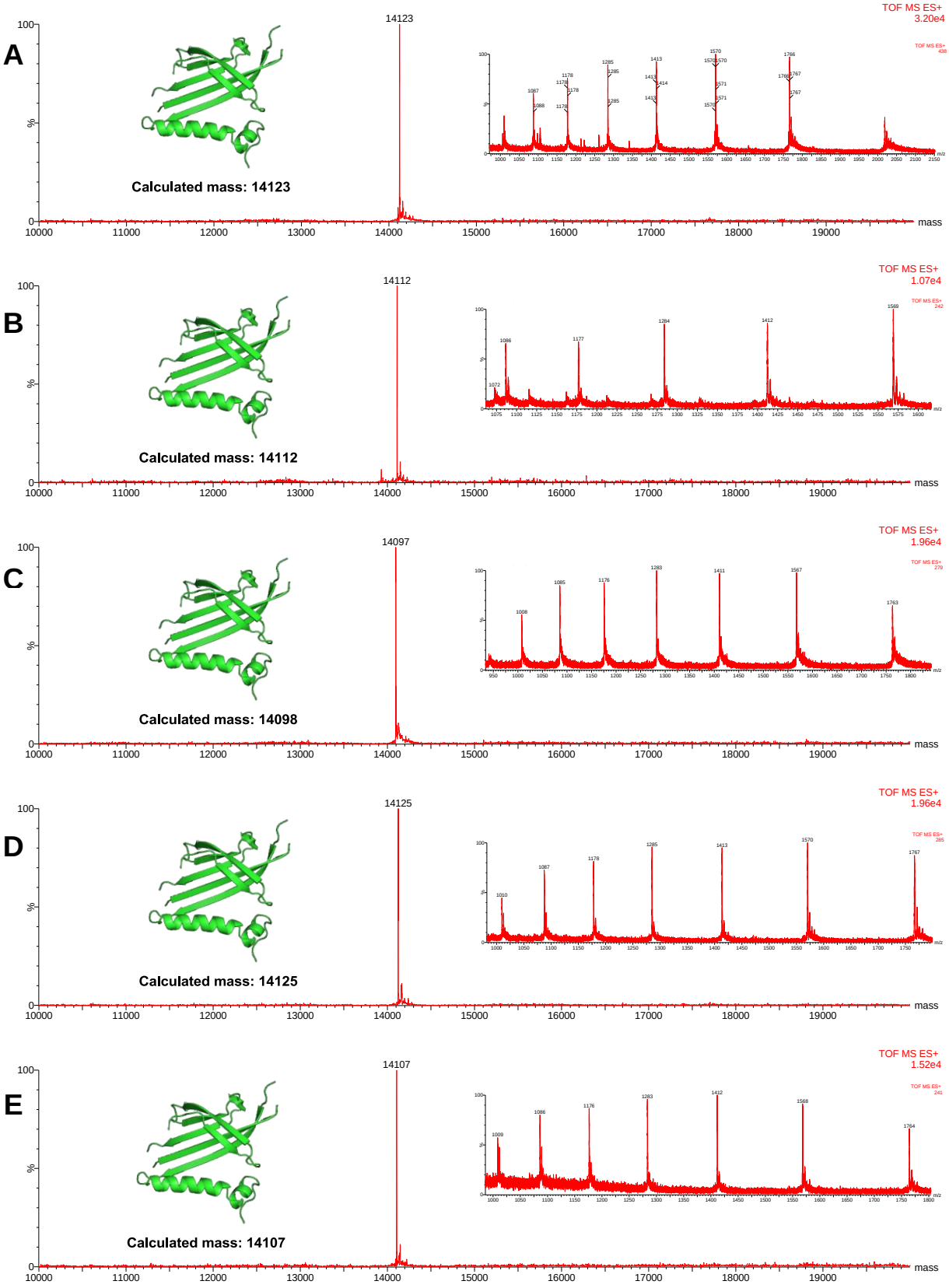


Figure 2.4: The deconvoluted mass spectra and the corresponding raw ion series of the purified (A) Qβ WT; (B) Qβ N10C; (C) Qβ K16C; (D) Qβ T18C and; (E) Qβ C74S VLPs. The VLPs were broken down to monomers for LC-MS analysis.

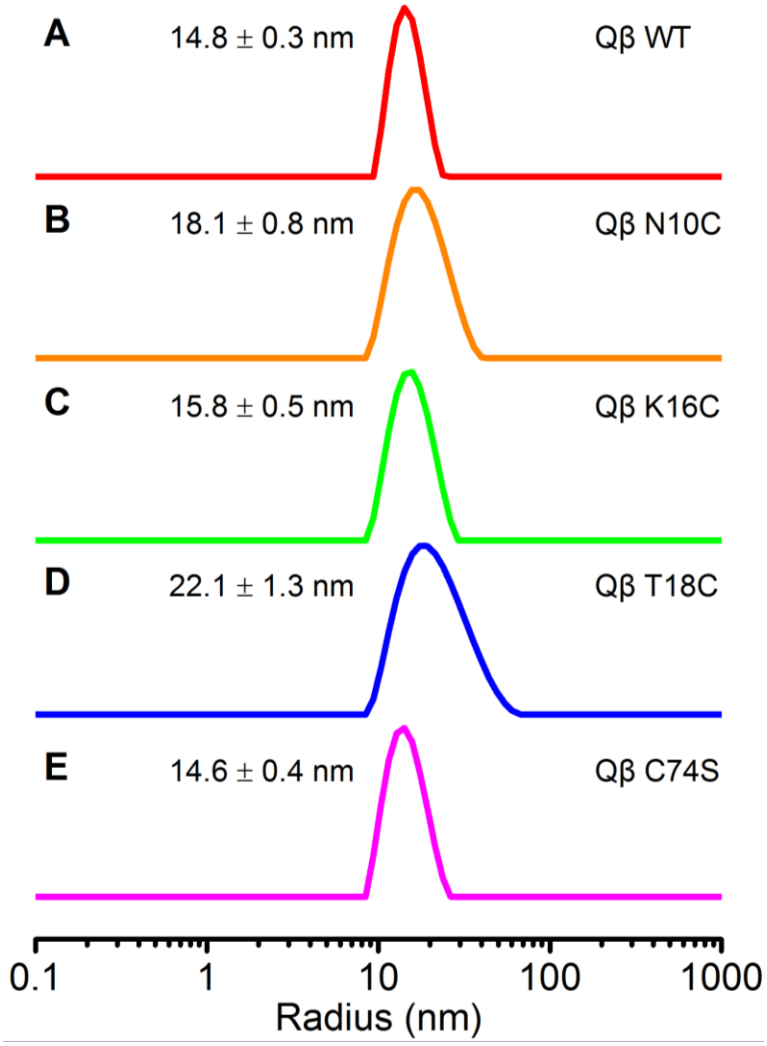


Figure 2.5: DLS results of the purified (A) Qβ WT; (B) Qβ N10C; (C) Qβ K16C; (D) Qβ T18C and; (E) Qβ C74S VLPs.

Table 2.1: The expression yield and the size of Qβ wild type and mutants

Qβ mutant	Yield (mg / L of media)	Radius of particle (nm)
WT	38.9	14.8 ± 0.3
N10C	29.3	18.1 ± 0.8 †
K16C	30.1	15.8 ± 0.5
T18C	11.2	22.1 ± 1.3 †
C74S	42.1	14.6 ± 0.4

† Unresolved peaks were identified by the software, indicating the presence of some aggregates.

As shown by the results, both Qβ WT and mutants were expressed successfully. The K16C species moved much quicker in the agarose gel as compared to the other species, which is consistent with the loss of the positively charged lysine residue. The mass of the expressed protein was proved to be correct by LC-MS. From the DLS results, both K16C and C74S mutants have a similar size as the wild type, which is consistent to the literature.[21] The N10C and T18C mutants have a larger average size, likely due to the presence of aggregates. This is not surprising as an extra cysteine residue is introduced which may affect the structure of the particle due to the formation of extra disulphide bonds. Furthermore, in the case of Qβ C74S, particles can still be formed even the disulphide bonds were removed, which is consistent with the literature.[21]

2.2.1.2 Conversion of cysteine to dehydroalanine

In order to convert cysteine to Dha, the expressed Qβ mutants were first treated with DTT to remove the disulphide linkages on the protein so as to expose the free thiol group. The DTT was then removed by desalting columns before treating with DBHDA under different conditions (**Table 2.2**) to avoid any side reactions. DTT is required even for the C74S species as the 'free single cysteine' can slowly form disulphide bonds with each other, which can significantly reduce the conversion. On the other hand, the wild type protein was used as a control to study the reactivity of the disulphide bond. The Dha conversion was calculated based on the intensity ratio of the starting material and product peaks from LC-MS.

As shown by the results, both the WT protein and T18C mutant gave very low conversion, while the other three mutants, especially Qβ N10C, showed greater reactivity. It was found that increasing the amount of DBHDA used did not improve conversion; indeed, it had the opposite effect. Furthermore, no peak was observed from all the species when the

temperature was increased from 37 °C to 45 °C, likely due to degradation of the protein. The effect of pH was also considered since basic conditions are required for deprotonation of the sulphonium intermediate during the reaction (**Scheme 2.4A**). However, a lot of adducts were observed in the spectra when a higher pH was used. Finally, more DTT was added before the reaction to make sure no aggregates were left when DBHDA was added. Nevertheless, only the C74S mutant showed Dha formation. A control experiment was also performed for all the species to confirm that DMF, the solvent used to solubilise the DBHDA, has no effect on the protein.

Table 2.2: First attempts of Dha formation on Qβ wild type and mutants

Entry	DTT (equiv.)	DBHDA (equiv.)	Temp (°C)	Time (h)	pH [§]	Dha conversion (%)				
						N10C	K16C	T18C	C74S	WT
1	25	500	37	1	7.5	58[†]	35[†]	0	36	7[†]
2	25	1000	37	1	7.5	0[†]	0[†]	0[†]	0	0[†]
3	25	1000	37	2	7.5	0[†]	0[†]	13[†]	0	0[†]
4	25	1000	37	4	7.5	0[†]	0[†]	NP	0	0[†]
5	25	1000	37	20	7.5	NP	18[†#]	0	0	11[†]
6	25	1000	45	4	7.5	NP	NP	NP	NP	NP
7	25	500	37	1	8.5	0*	0*	0*	0*	0*
8	25	500	37	20	8.5	0*	0*	0*	0*	0*
9	100	500	37	20	7.5	NP	0[†#]	NP	32[#]	NP
10	25	0	37	1	7.5	NP	0	0	0	0

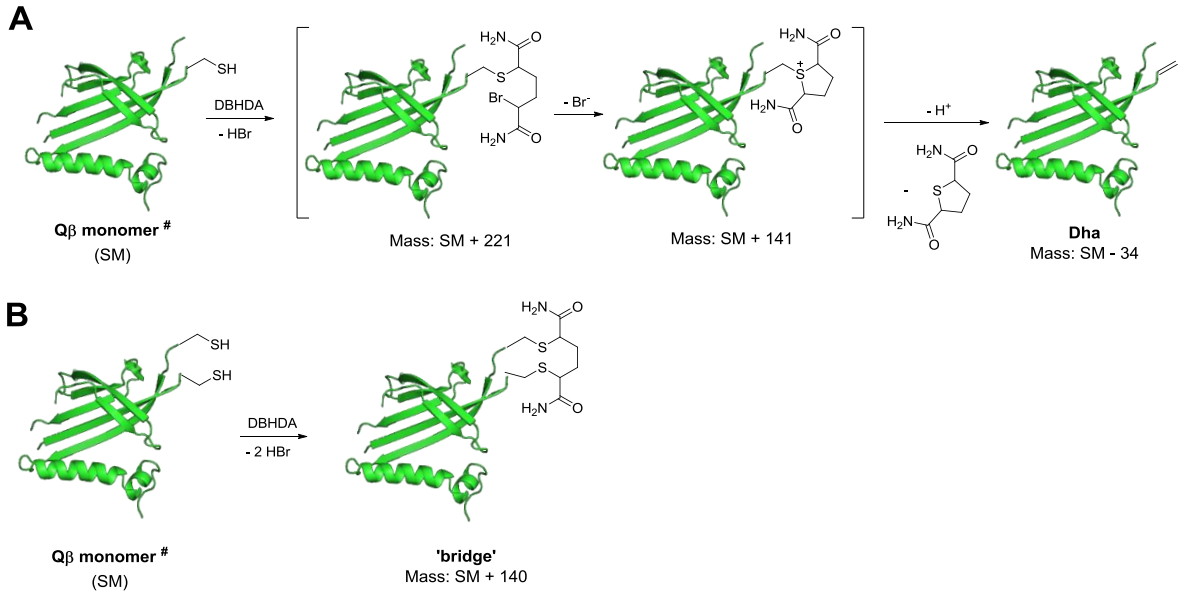
[§] TBS buffer (50 mM Tris, 150 mM NaCl)

[†] A side product with a mass of starting material + 140 was observed

[#] A side product with a mass of starting material + 120 was observed

* A lot of protein adducts with various masses were observed

NP: the Qβ protein peak was not observed in the spectrum (i.e. protein degraded)



Scheme 2.4: Two hypotheses were proposed for the formation of the +140 adduct: (A) the sulphonium intermediate of the proposed mechanism of Dha formation or; (B) the formation of 'bridge' between two cysteine residues in the same monomer. # The monomer is shown only for simplicity; the actual reaction is carried out on a particle.

Despite the considerable amount of side products, N10C gave the best conversion to Dha of all the mutants. On the other hand, C74S gave a relatively clean product compared with other species, though a +120 adduct was observed in some of the cases. It is still unknown what the +120 adduct is, although more information is known about the +140 adduct, which is commonly observed in the other species. As shown by the proposed mechanism of Dha formation, the +140 adduct may come from the sulphonium intermediate, though it is known to be unstable (**Scheme 2.4A**).[12] However, it is interesting that C74S did not give any +140 adduct as the other species did. This suggested that the +140 adduct may be a disulphide bridge formed between two proximal cysteine residues during the reaction (**Scheme 2.4B**). Since the C74S mutant only possesses one cysteine per monomer, it cannot form such bridges as the cysteine residues from different monomers are too far away from each other. This can also explain why the conversion yield of the WT species was so low even it possesses Cys-80, since this residue would be blocked by bridge formation during the

reaction. Based on the lack of any such putative disulphide bridge formation for the C74S mutant, it was decided to focus on it only.

The reaction conditions for the C74S mutant were further optimised. It was found that the reaction can reach near completion when DBHDA was added in two separate portions. After adding 500 equiv. of DBHDA for 1 h, another 500 equiv. of DBHDA was added and the reaction incubated overnight. No side product was observed by adopting this method (data not shown). A quicker and simpler way was subsequently discovered, involving simply changing the buffer from TBS to sodium phosphate (50 mM, pH 8.0). Under these conditions, the reaction can be completed within an hour at 37 ^{o}C, and no side product was formed (**Figure 2.6**). It is possible that the amine group in Tris hinders the reaction by competing with the thiol group in attacking DBHDA. Furthermore, it was found that the presence of DTT during the reaction did not lead to the formation of side product if the reaction is completed within an hour. The step for removing DTT before the reaction can therefore be omitted.

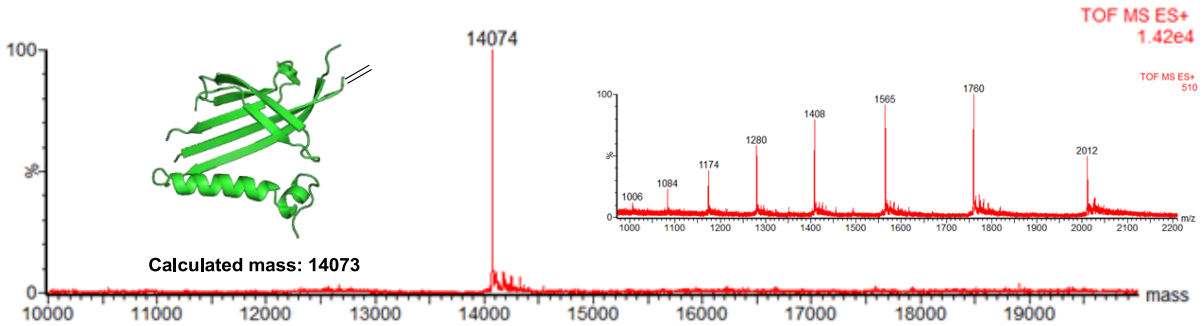


Figure 2.6: Deconvoluted mass spectrum and corresponding raw ion series of the purified Qβ C74S C80Dha VLPs. The VLPs were broken down to monomers for LC-MS analysis.

The protein recovery yield of this method was found to be extremely low (~5%) at first, and the LC-MS signal-to-noise ratio was very poor unless the reaction mixture was desalted using Vivaspin™ many times. The concentration of DMF was eventually revealed to be the

main factor. Since DBHDA is dissolved in DMF, the final concentration of DMF in the reaction mixture greatly depends on the concentration of protein. By reducing the protein concentration and hence the percentage of DMF to approximately 5%, both the protein recovery yield (~50%) and the noise signal were improved significantly.

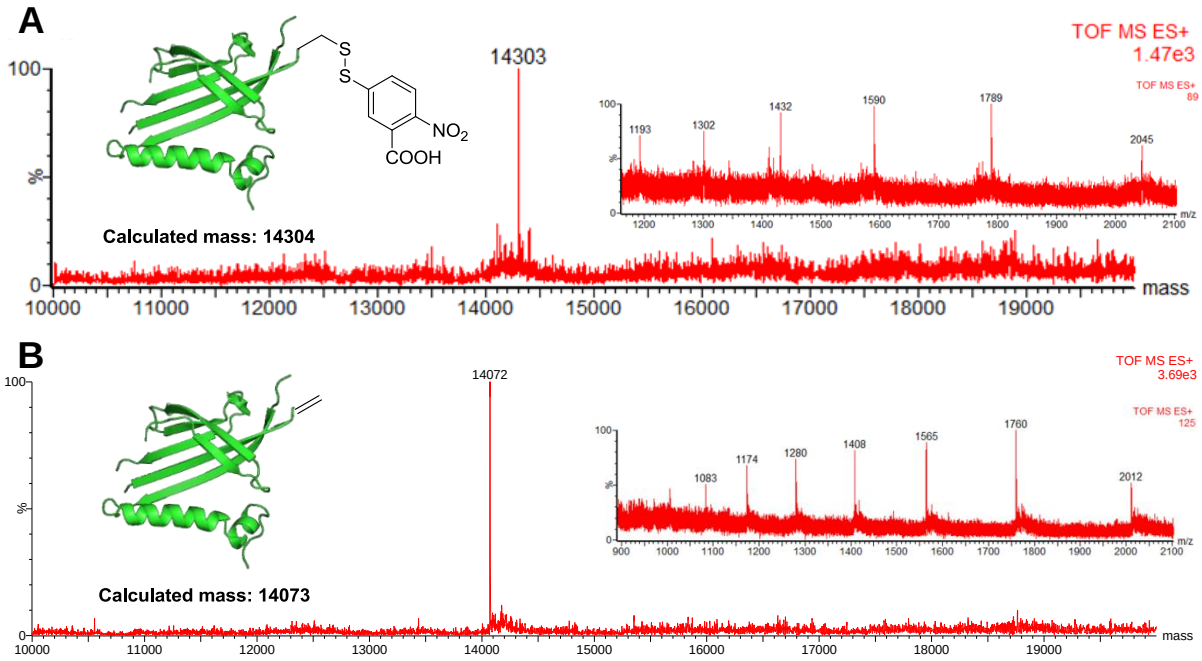


Figure 2.7: Deconvoluted mass spectra and corresponding raw ion series showing the reaction of Ellman's reagent (100 equiv.) with (A) Qβ C74S or; (B) Qβ C74S C80Dha VLPs. The VLPs were broken down to monomers for LC-MS analysis.

The formation of Dha was confirmed by the reaction with Ellman's reagent (**Figure 2.7**). No reaction would be observed if all the cysteine residues in the particle are converted to Dha as Ellman's reagent can only couple to the thiol group on free cysteine. On the other hand, the reactivity of Dha was tested by the reaction with 2-mercaptoethanol (**Figure 2.8**). It can be clearly observed that the Dha group can be modified completely by the thiol substrate. This is consistent with the 'tag-and-modify' approach reported previously in other proteins.[9, 12-14] To the best of our knowledge, this is the first time we showed that Qβ, as a complicated large particle, can be successfully modified using this approach.

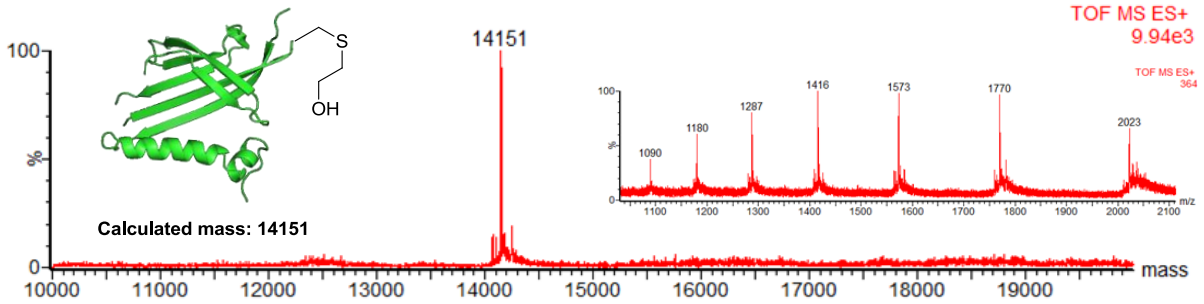


Figure 2.8: Deconvoluted mass spectrum and corresponding raw ion series showing the reaction of 2-mercaptoethanol (1000 equiv.) with Qβ C74S C80Dha VLPs. The VLPs were broken down to monomers for LC-MS analysis.

2.2.1.3 Fluorination of Qβ VLPs with Selectfluor™

2.2.1.3.1 Qβ WT VLPs

Qβ WT VLP was first tested with Selectfluor™ to see if any side reaction can occur (**Figure 2.9**). As shown by the mass spectra, no reaction was observed even if 40 equiv. of Selectfluor™ was added. This showed that the amino acids on the natural particle are inert towards Selectfluor™, which provided a good start for the 'tag-and-modify' approach since Dha would be the only reactive species. However, the noise of the LC-MS spectra increases with the amount of Selectfluor™ added, indicating that the presence of excess Selectfluor™ may interfere with the raw ion series. The exact mechanism is still unknown but it is proposed that Selectfluor™ tends to interact strongly with proteins which cannot be easily removed by columns. The noise can be slightly improved by removal of the excess Selectfluor™ after the reaction through standard desalting procedures.

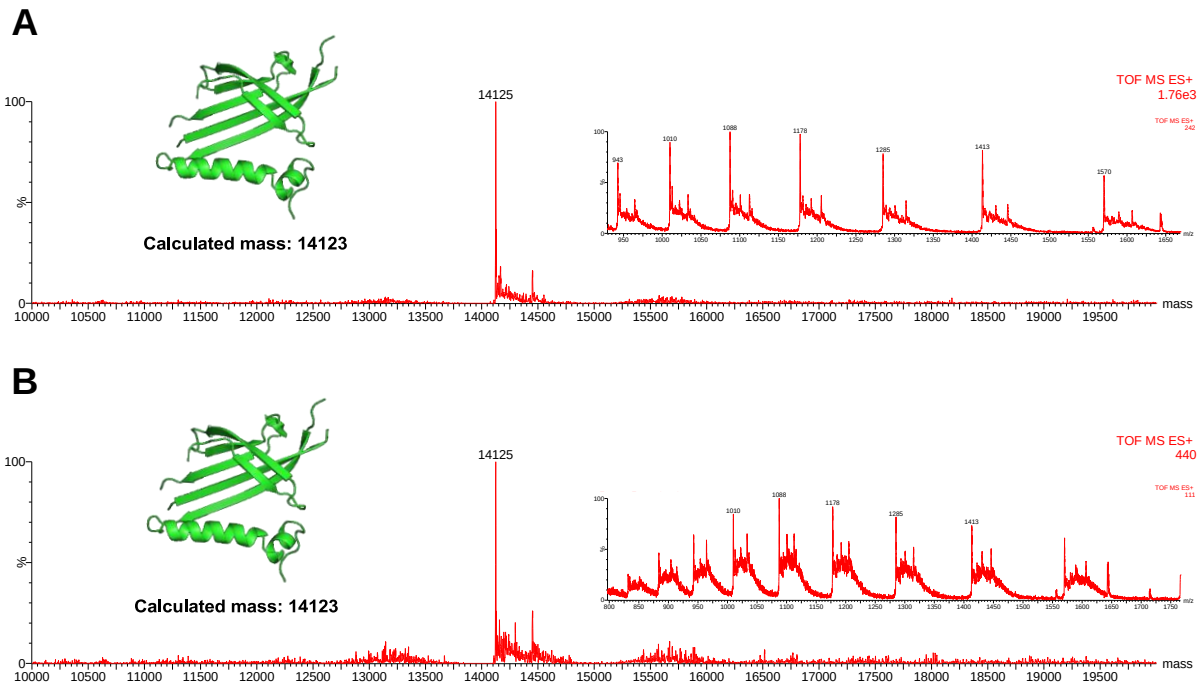


Figure 2.9: Deconvoluted mass spectra and corresponding raw ion series showing the reaction of Qβ WT VLPs with (A) 20 equiv. or; (B) 40 equiv. of Selectfluor™ BF_4^-. The VLPs were broken down to monomers for LC-MS analysis.

2.2.1.3.2 Qβ C74S C80Dha VLPs

Different equiv. of Selectfluor™ was added to Qβ C74S C80Dha to test for any fluorination. As shown by the spectra, fluorination started to occur when more than 5 equiv. of Selectfluor™ was added (**Figure 2.10**). Only a single fluorine addition was observed for reaction mixture with less than 10 equiv. of Selectfluor™. For mixture with more than 20 equiv. of Selectfluor™, the spectra became much noisier. A peak of Dha + 37 was observed in the 20 equiv. sample, which may indicate the formation of Hfa, the hydrated form of Dfa. The high level of noise in the spectrum for the 40 equiv. sample prevented proper analysis of the products, but all the starting material appeared to be consumed completely as the peak corresponding to Dha (mass: 14073) was not observed. **Table 2.3** summarises the percentage of conversion based on the intensity ratio of each peak.

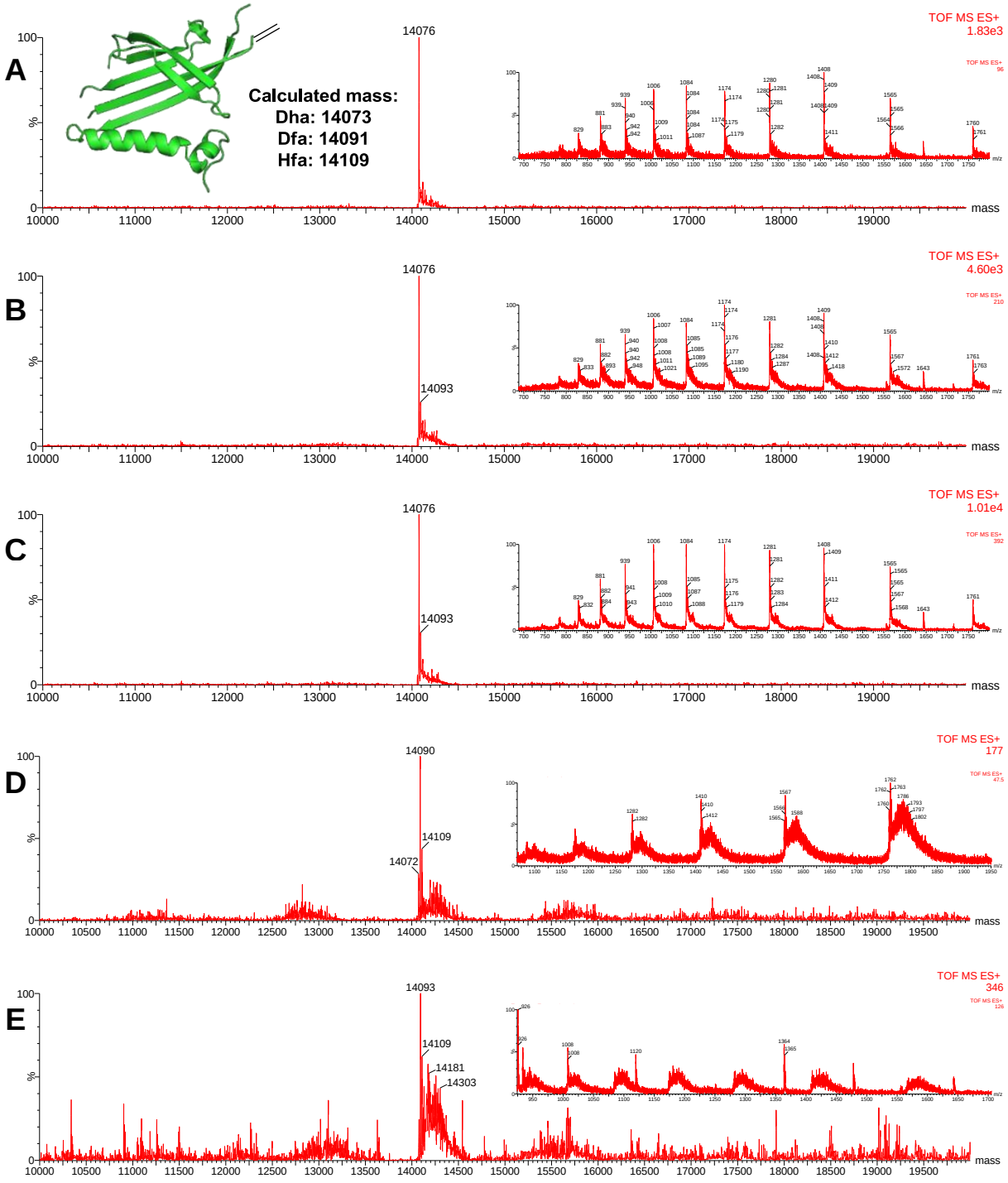


Figure 2.10: Deconvoluted mass spectra and corresponding raw ion series showing the reaction of Qβ C74S C80Dha VLPs with (A) 1 equiv.; (B) 5 equiv.; (C) 10 equiv.; (D) 20 equiv. or; (E) 40 equiv. of Selectfluor™ BF_4^-. The VLPs were broken down to monomers for LC-MS analysis.

Table 2.3 Optimisation experiments of Qβ C74S C80Dha VLPs with Selectfluor™

Entry	Selectfluor™		Co-solvent	Product Conversion (%)
	Anion	Equiv.		
1	BF_4^-	1	-	Dha (100)
2	BF_4^-	5	-	Dha (85) Dha + 18 (15)
3	BF_4^-	10	-	Dha (75) Dha + 18 (25)
4	BF_4^-	20	-	Dha (15) Dha + 18 (60) Dha + 36 (25)
5	BF_4^-	40	-	*
6	OTf^-	20	7% Acetone	Dha (55) Dha + 18 (30) Dha + 36 (15)
7	OTf^-	40	14% Acetone	*

General conditions: Qβ C74S C80Dha VLPs in NaPi (50 mM, pH 8) and Selectfluor™ in H_2O at 25 °C for 30 min unless otherwise indicated; * the signal-to-noise ratio was so high that the spectrum cannot be analysed properly

In the above experiments, Selectfluor™ with BF_4^- as the counter-ion was used for testing as this species is more soluble in water than with other counter-ions. As mentioned earlier in this chapter, ^{18}F-labelled Selectfluor™ has been successfully synthesised recently. From the reported synthetic route, the resulting ^{18}F species would have OTf^- as the counter-ion instead. A small amount of acetone is required to make it soluble in protein solution. To prove that this method can be applied for ^{18}F labelling, Qβ C74S C80Dha VLPs were tested with a ^{19}F version of Selectfluor™ OTf^- in the presence of acetone to see if similar results can be observed. Although the conversion is slightly lower, both the Dha + 18 and Dha + 36 peaks were clearly seen in the mass spectrum (**Figure 2.11**).

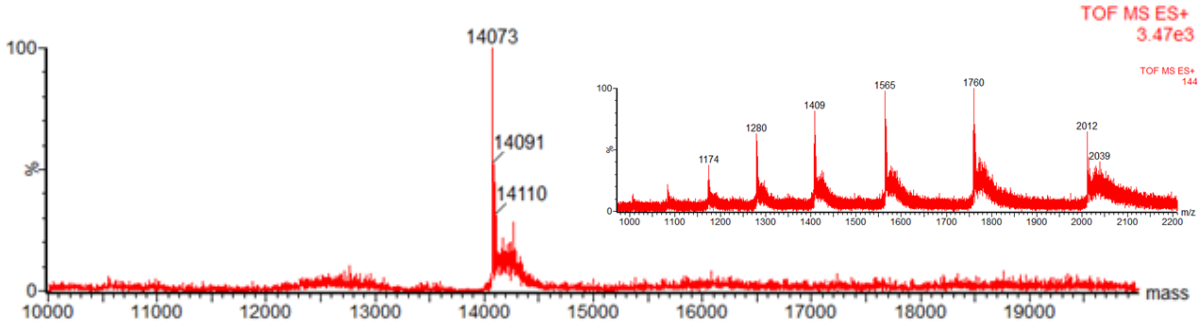
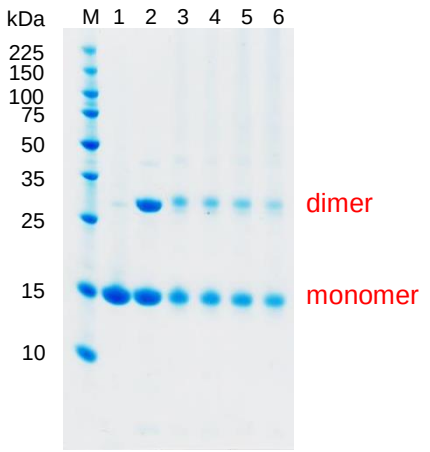


Figure 2.11: Deconvoluted mass spectrum and corresponding raw ion series showing the reaction of Qβ C74S C80Dha VLPs with 20 equiv. of Selectfluor™ OTf ⁻ with 7% acetone as co-solvent. The VLPs were broken down to monomers for LC-MS analysis.

To confirm if the protein is still stable after the reaction, the fluorinated VLPs were analysed by both SDS-PAGE (**Figure 2.12**) and DLS (**Figure 2.13**). The protein was fairly stable after fluorination as shown by the gel, but it can be observed that dimer interactions are much stronger after the formation of Dha. It is proposed that dimer formation is also enhanced during Dha formation and the dimers are not efficiently reduced in the SDS-PAGE gel.



Lane 1: Qβ C74S
Lane 2: Qβ C74S C80Dha
Lane 3: Selectfluor™ (20 equiv.) treated Qβ C74S C80Dha (-80 °C for 1 month)
Lane 4: Selectfluor™ (20 equiv.) treated Qβ C74S C80Dha (4 °C for 4 h)
Lane 5: Selectfluor™ (20 equiv.) treated Qβ C74S C80Dha (25 °C for 2 h)
Lane 6: Selectfluor™ (20 equiv.) treated Qβ C74S C80Dha (37 °C for 2 h)

Figure 2.12: Coomassie Blue stained reducing SDS-PAGE gel of Qβ C74S and its reaction products stored at different temperatures. M: protein marker.

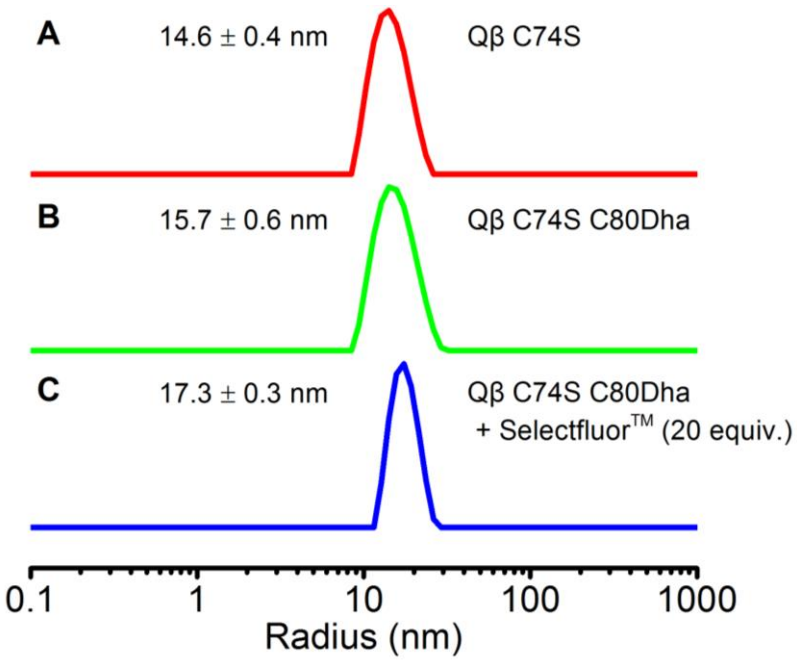


Figure 2.13: DLS results of the purified (A) Qβ C74S; (B) Qβ C74S C80Dha and; (C) Selectfluor™ (20 equiv.) treated Qβ C74S C80Dha VLPs.

2.2.1.4 Peptide mapping of the fluorinated particles

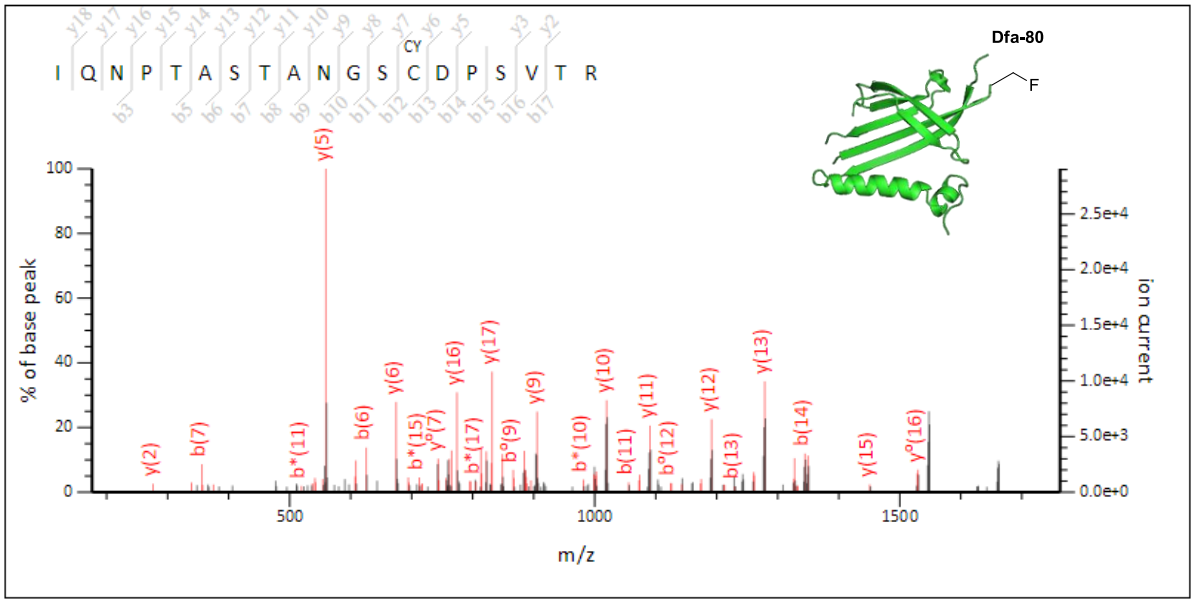
Although the mass spectra showed two peaks with masses of Dha + 18 and Dha + 36 respectively, it is possible that these peaks are contributed by other side reactions. The Selectfluor™ treated Qβ VLPs were thus further analysed by MS/MS to confirm the fluorination position. Ions score is calculated by -10*Log(P), where P is the probability that the observed match is a random event. As expected, both Dfa and Hfa modifications on Cys-80 were observed by MS/MS (**Table 2.4** and **Figure 2.14**). It is also known that Dfa can further react with 2-mercaptoethanol to form a coupled product (α-(hydroxyethylthio)-β-fluoro-alanine). As shown by the MS/MS spectrum, this product can be clearly observed when the protein was treated with 2-mercaptoethanol, further proving the presence of Dfa (**Table 2.5** and **Figure 2.15**). These results confirmed not only the direct formation of the C–F bond on Dha but also the stability of these residues both in the protein and the peptides derived from it after tryptic digestion.

Table 2.4: Key digested peptide fragments observed from Mascot analysis of fluorinated Qβ C74S C80Dha (protein sequence coverage: 100%)

Score	Mr(found)	Mr(Calc)	Error (ppm)	Expect	Modifications†	Peptide sequence
122.5	1883.8961	1883.8974	-0.67	5.6e-13	Cys80→Dha80	68-IQNPTASTANGS**C**DPSVTR-86
95.9	1901.8930	1901.8879	2.69	2.6e-10	Cys80→Dfa80	68-IQNPTASTANGS**C**DPSVTR-86
28.3	1919.8888	1919.8985	-5.06	0.0015	Cys80→Hfa80	68-IQNPTASTANGS**C**DPSVTR-86
71.6	2274.9701	2274.9717	-0.72	6.9e-08	Tyr89→Tyr_F89	87-QA**Y**ADVTFSFTQYSTDEER-105
90.3	2274.9725	2274.9717	0.35	9.3e-10	Tyr99→Tyr_F99	87-QAYADVTFSFTQ**Y**STDEER-105
86.8	2292.9522	2292.9623	-4.38	2.1e-09	Tyr99→Tyr_FF99	87-QAYADVTFSFTQ**Y**STDEER-105
66.0	2292.9603	2292.9623	-0.87	2.5e-07	Tyr89→Tyr_F89 Tyr99→Tyr_F99	87-QA**Y**ADVTFSFTQ**Y**STDEER-105

† Tyr_F: monofluorination on tyrosine residue; Tyr_FF: difluorination on tyrosine residue

A



B

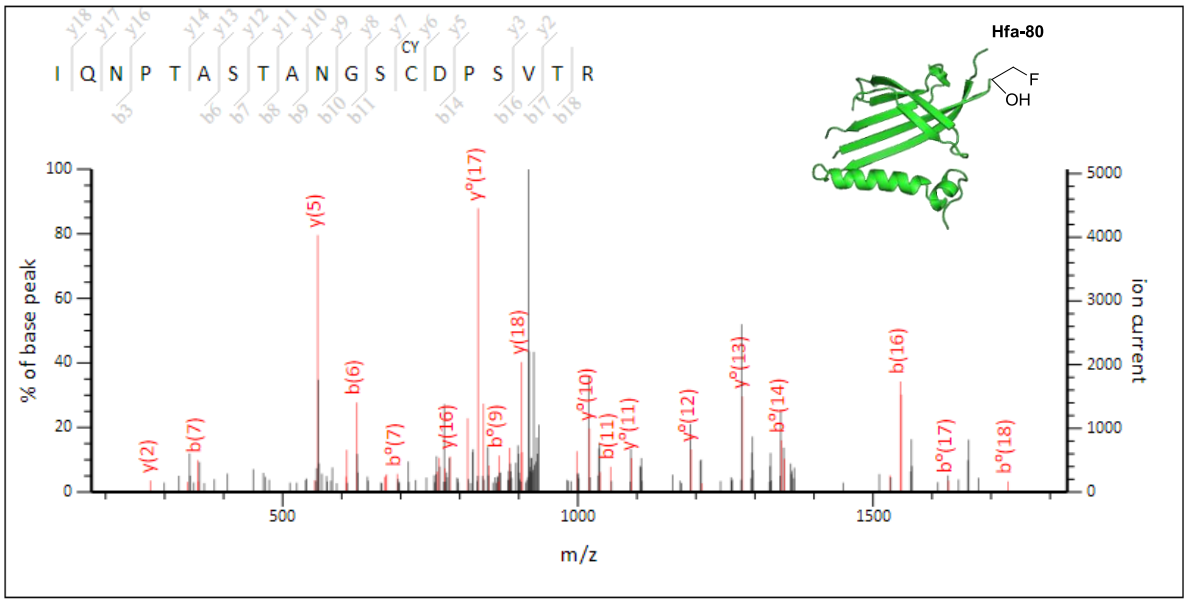


Figure 2.14: ESI-QTOF MS-MS spectra of peptide 68-IQNPTASTANGSCDPSVTR-86 obtained by tryptic digestion of Selectfluor™ (20 equiv.) treated Qβ C74S C80Dha that contains (A) Dfa-80 (Score = 95.9) or; (B) Hfa-80 (Score = 28.3).

Table 2.5: Key digested peptide fragments observed from Mascot analysis of fluorinated Qβ C74S C80Dha after the addition of 2-mercaptoethanol (protein sequence coverage: 100%)

Score	Mr(found)	Mr(Calc)	Error (ppm)	Expect	Modifications†	Peptide sequence
99.5	1961.9129	1961.9113	0.85	1.1e-10	Cys80→Dha_bME80	68-IQNPTASTANGS**C**DPSVTR-86
23.5	2434.1735	2434.1922	-7.70	0.0045	Cys80→Dfa_bME80	68-IQNPTASTANGS**C**DPSVTR-86
9.7	2374.1608	2374.1889	-11.8	0.11	Cys80→Hfa80	64-VQVKIQNPTASTANGS**C**DPSVTR-86
77.1	2274.9710	2274.9717	-0.29	2.0e-08	Tyr89→Tyr_F89	87-QA**Y**ADVTFSFTQYSTDEER-105
78.3	2274.9720	2274.9717	0.14	1.5e-08	Tyr99→Tyr_F99	87-QAYADVTFSFTQ**Y**STDEER-105
68.0	2292.9654	2292.9623	1.38	1.6e-07	Tyr99→Tyr_FF99	87-QAYADVTFSFTQ**Y**STDEER-105
84.3	2292.9630	2292.9623	0.30	3.7e-09	Tyr89→Tyr_F89 Tyr99→Tyr_F99	87-QA**Y**ADVTFSFTQ**Y**STDEER-105

† Dha_bME: 2-mercaptoethanol addition on Dha; Dfa_bME: 2-mercaptoethanol addition on Dfa; Tyr_F: monofluorination on tyrosine residue; Tyr_FF: difluorination on tyrosine residue

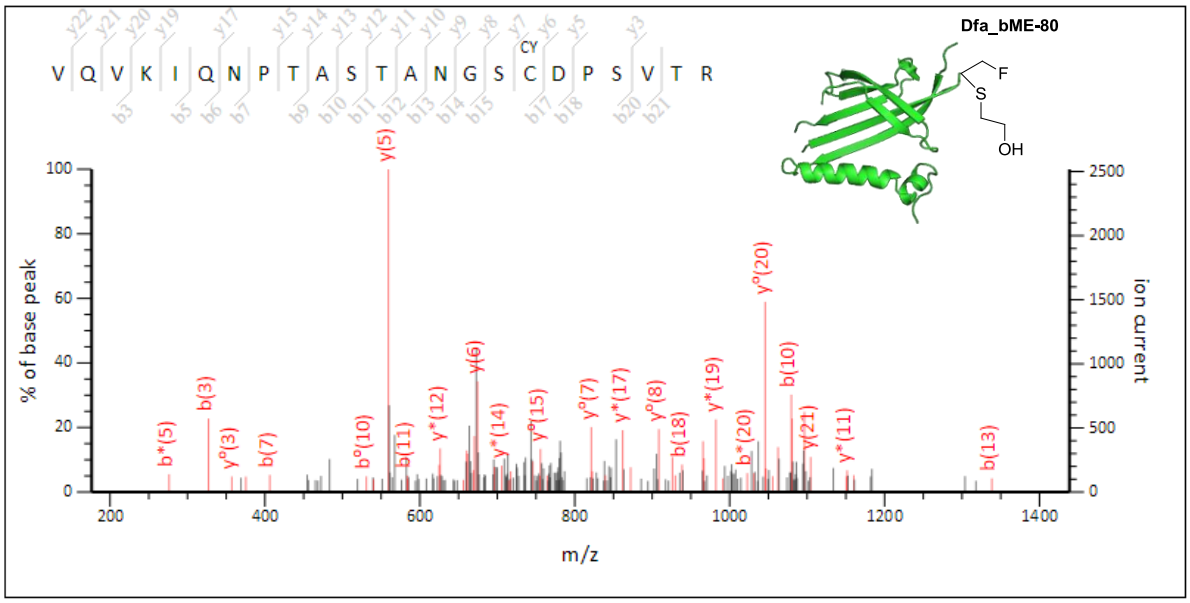


Figure 2.15: ESI-QTOF MS-MS spectrum of peptide 68-IQNPTASTANGSCDPSVTR-86 obtained by tryptic digestion of Selectfluor™ (20 equiv.) and 2-mercaptoethanol treated Qβ C74S C80Dha that contains the coupled product Dfa_bME-80 (Score = 23.5).

As shown by the peptide fragment table, fluorination on Tyr-89 and Tyr-99 was also observed as a side reaction, which will be further discussed in Chapter 3. The fluorination of tyrosine residue by Selectfluor™ has been observed before in other proteins.[15] It is known that this side reaction greatly depends on the accessibility of the tyrosine residues. For Qβ VLPs, all the tyrosine residues have already been shown to be inert on the WT species (**Figure 2.9**). It is therefore surprising that two out of four of the tyrosine residues (i.e. Tyr-89 and Tyr-99) were fluorinated. It is possible that the deletion of disulphide bonds slightly changed the protein structure, which in turns exposed these residues to the electrophilic fluorine source. In order to eliminate the side reaction, it was planned to remove these two tyrosine residues by mutagenesis.

2.2.2 Removal of reactive tyrosine residues

2.2.2.1 Mutagenesis, expression and purification of Qβ VLPs

To circumvent the side reactions on tyrosine residues, it was proposed to generate the Qβ mutant C74S Y89F Y99F, in which the two reactive tyrosine residues were mutated to phenylalanine. On the other hand, four more Qβ mutants were also expressed in which the Dha site was inserted in other position. This includes Qβ N10C C74S C80S Y89F Y99F (Qβ N10C*), Qβ K16C C74S C80S Y89F Y99F (Qβ K16C*), Qβ C74S T75C C80S Y89F Y99F (Qβ T75C*) and Qβ C74S C80S Y89F Y99F E103C (Qβ E103C*). The positions were chosen based on their accessibility (**Figure 2.2**). For all the mutants proposed, only one cysteine residue was left per monomer so as to avoid the formation of 'bridges' during Dha formation. The DNA and protein sequences of the mutants are listed below (The mutated sites are labelled with red; the remaining cysteine residues are labelled with blue):

(A) Qβ C74S Y89F Y99F

Forward DNA sequence

ATGGCAAAAT TAGAGACTGT TACTTTAGGT AACATCGGGA AAGATGGAAA ACAAACTCTG
GTCCTCAATC CGCGTGGGGT AAATCCCACT AACGGCGTTG CCTCGCTTTC ACAAGCGGGT
GCAGTTCCTG CGCTGGAGAA GCGTGTTACC GTTTCGGTAT CTCAGCCTTC TCGCAATCGT
AAGAACTACA AGGTCCAGGT TAAGATCCAG AACCCGACCG CT**AGC**ACTGC AAACGGTTCT
TGTGACCCAT CCGTTACTCG CCAGGCA**TTT** GCTGACGTGA CCTTTTCGTT CACGCAG**TTT**
AGTACCGATG AGGAACGAGC TTTTGTTCGT ACAGAGCTTG CTGCTCTGCT CGCTAGTCCT
CTGCTGATCG ATGCTATTGA TCAGCTGAAC CCAGCGTATT AA

Amino acid sequence

AKLETVTLGN IGKDGKQTLV LNPRGVNPTN GVASLSQAGA VPALEKRVTV SVSQPSRNRK
NYKVQVKIQN PTA**S**TANGS**C** DPSVTRQA**F**A DVTFSFTQ**F**S TDEERAFVRT ELAALLASPL
LIDAIDQLNP AY

Calculated average isotopic mass = 14075 (N-terminal Met cleaved)

(B) Qβ N10C C74S C80S Y89F Y99F

Forward DNA sequence

ATGGCAAAAT TAGAGACTGT TACTTTAGGT **TGC**ATCGGGA AAGATGGAAA ACAAACTCTG
GTCCTCAATC CGCGTGGGGT AAATCCCACT AACGGCGTTG CCTCGCTTTC ACAAGCGGGT
GCAGTTCCTG CGCTGGAGAA GCGTGTTACC GTTTCGGTAT CTCAGCCTTC TCGCAATCGT
AAGAACTACA AGGTCCAGGT TAAGATCCAG AACCCGACCG CT**AGC**ACTGC AAACGGTTCT
TCGGACCCAT CCGTTACTCG CCAGGCA**TTT** GCTGACGTGA CCTTTTCGTT CACGCAG**TTT**
AGTACCGATG AGGAACGAGC TTTTGTTCGT ACAGAGCTTG CTGCTCTGCT CGCTAGTCCT
CTGCTGATCG ATGCTATTGA TCAGCTGAAC CCAGCGTATT AA

Amino acid sequence

AKLETVTLG**C** IGKDGKQTLV LNPRGVNPTN GVASLSQAGA VPALEKRVTV SVSQPSRNRK
NYKVQVKIQN PTA**S**TANGS**S** DPSVTRQA**F**A DVTFSFTQ**F**S TDEERAFVRT ELAALLASPL
LIDAIDQLNP AY

Calculated average isotopic mass = 14048 (N-terminal Met cleaved)

(C) Qβ K16C C74S C80S Y89F Y99F

Forward DNA sequence

ATGGCAAAAT TAGAGACTGT TACTTTAGGT AACATCGGGA AAGATGGA**TG** **T**CAAACTCTG
GTCCTCAATC CGCGTGGGGT AAATCCCACT AACGGCGTTG CCTCGCTTTC ACAAGCGGGT
GCAGTTCCTG CGCTGGAGAA GCGTGTTACC GTTTCGGTAT CTCAGCCTTC TCGCAATCGT
AAGAACTACA AGGTCCAGGT TAAGATCCAG AACCCGACCG CT**AGC**ACTGC AAACGGTTCT
TCGGACCCAT CCGTTACTCG CCAGGCA**TTT** GCTGACGTGA CCTTTTCGTT CACGCAG**TTT**
AGTACCGATG AGGAACGAGC TTTTGTTCGT ACAGAGCTTG CTGCTCTGCT CGCTAGTCCT
CTGCTGATCG ATGCTATTGA TCAGCTGAAC CCAGCGTATT AA

Amino acid sequence

AKLETVTLGN IGKDG**C**QTLV LNPRGVNPTN GVASLSQAGA VPALEKRVTV SVSQPSRNRK
NYKVQVKIQN PTA**S**TANGS**S** DPSVTRQA**F**A DVTFSFTQ**F**S TDEERAFVRT ELAALLASPL
LIDAIDQLNP AY

Calculated average isotopic mass = 14034 (N-terminal Met cleaved)

(D) Qβ C74S T75C C80S Y89F Y99F

Forward DNA sequence

ATGGCAAAAT TAGAGACTGT TACTTTAGGT AACATCGGGA AAGATGGAAA ACAAACTCTG
GTCCTCAATC CGCGTGGGGT AAATCCCACT AACGGCGTTG CCTCGCTTTC ACAAGCGGGT
GCAGTTCCTG CGCTGGAGAA GCGTGTTACC GTTTCGGTAT CTCAGCCTTC TCGCAATCGT
AAGAACTACA AGGTCCAGGT TAAGATCCAG AACCCGACCG CT**AGCTGT**GC AAACGGTTCT
TCGGACCCAT CCGTTACTCG CCAGGCA**TTT** GCTGACGTGA CCTTTTCGTT CACGCAG**TTT**
AGTACCGATG AGGAACGAGC TTTTGTTCGT ACAGAGCTTG CTGCTCTGCT CGCTAGTCCT
CTGCTGATCG ATGCTATTGA TCAGCTGAAC CCAGCGTATT AA

Amino acid sequence

AKLETVTLGN IGKDGKQTLV LNPRGVNPTN GVASLSQAGA VPALEKRVTV SVSQPSRNRK
NYKVQVKIQN PTA**SC**ANGS**S** DPSVTRQA**F**A DVTFSFTQ**F**S TDEERAFVRT ELAALLASPL
LIDAIDQLNP AY

Calculated average isotopic mass = 14061 (N-terminal Met cleaved)

(E) Qβ C74S C80S Y89F Y99F E103C

Forward DNA sequence

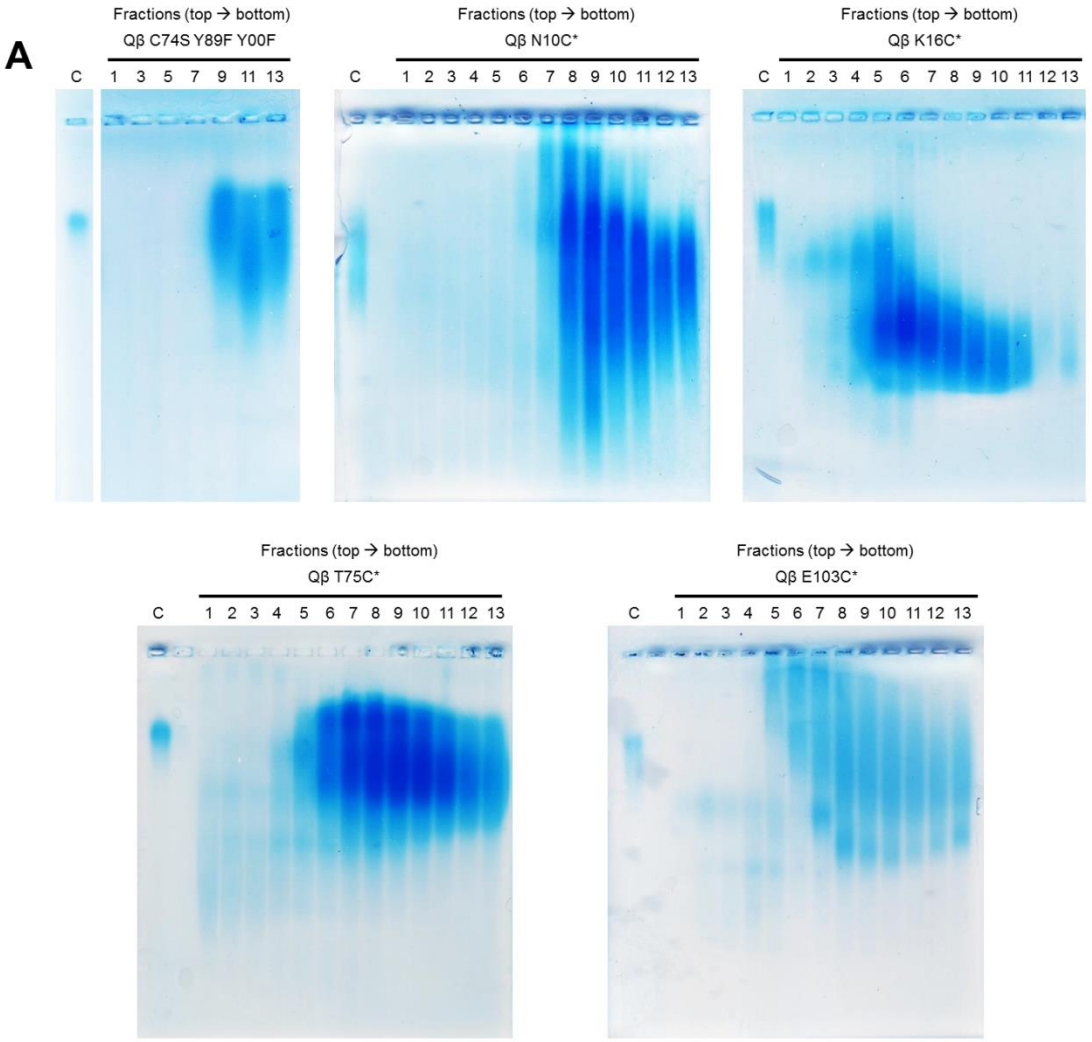
ATGGCAAAAT TAGAGACTGT TACTTTAGGT AACATCGGGA AAGATGGAAA ACAAACTCTG
GTCCTCAATC CGCGTGGGGT AAATCCCACT AACGGCGTTG CCTCGCTTTC ACAAGCGGGT
GCAGTTCCTG CGCTGGAGAA GCGTGTTACC GTTTCGGTAT CTCAGCCTTC TCGCAATCGT
AAGAACTACA AGGTCCAGGT TAAGATCCAG AACCCGACCG CT**AGC**ACTGC AAACGGTTCT
TCGGACCCAT CCGTTACTCG CCAGGCA**TTT** GCTGACGTGA CCTTTTCGTT CACGCAG**TTT**
AGTACCGAT**T** **GT**GAACGAGC TTTTGTTCGT ACAGAGCTTG CTGCTCTGCT CGCTAGTCCT
CTGCTGATCG ATGCTATTGA TCAGCTGAAC CCAGCGTATT AA

Amino acid sequence

AKLETVTLGN IGKDGKQTLV LNPRGVNPTN GVASLSQAGA VPALEKRVTV SVSQPSRNRK
NYKVQVKIQN PTA**S**TANGS**S** DPSVTRQA**F**A DVTFSFTQ**F**S TD**C**ERAFVRT ELAALLASPL
LIDAIDQLNP AY

Calculated average isotopic mass = 14033 (N-terminal Met cleaved)

The proteins were successfully expressed and purified using the same protocol as previously outlined. Fractions containing the intact particles were combined based on the results from the native agarose and SDS-PAGE gels (**Figure 2.16**). The fractions were then concentrated and analysed by LC-MS (**Figure 2.17**) and DLS (**Figure 2.18**). The results were summarised in **Table 2.6**. Since more mutations were introduced, the structure of the protein was expected to vary more. This is confirmed by the relatively large size of all the mutants expressed compared to Qβ WT VLPs.



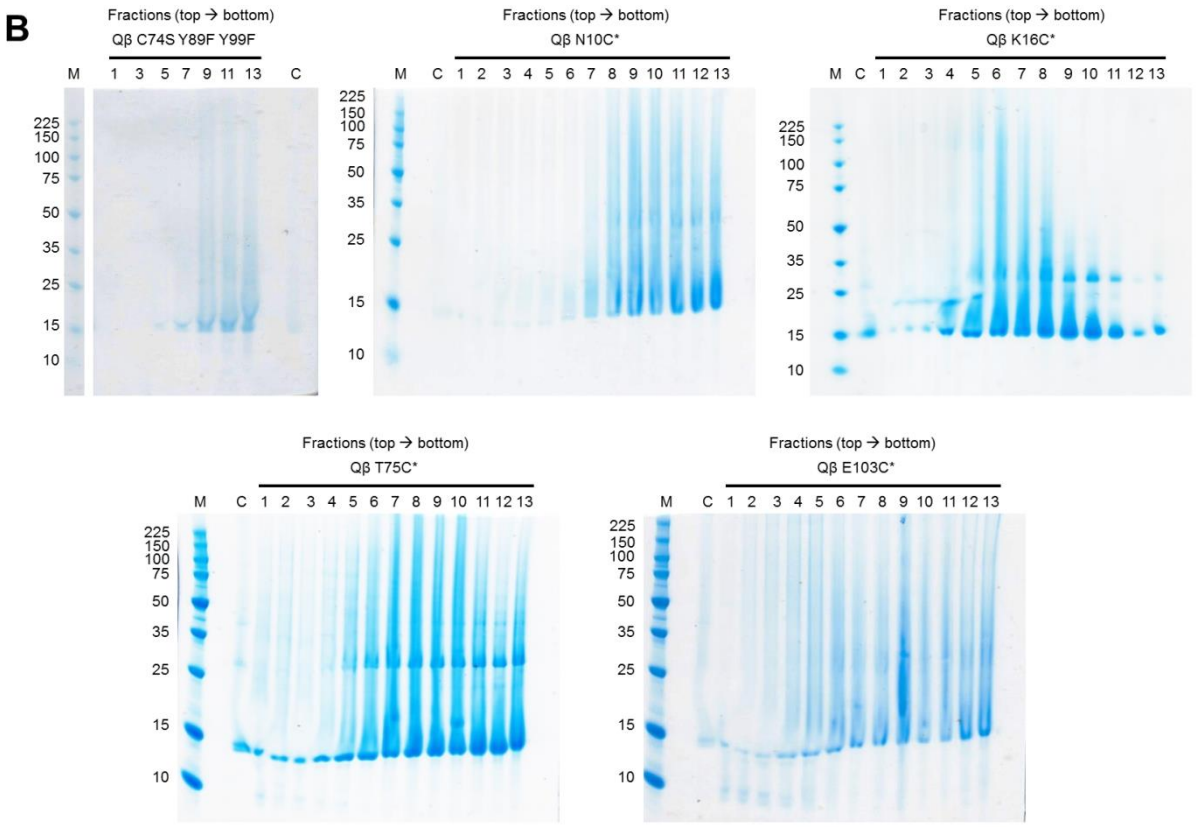


Figure 2.16: Coomassie Blue stained (A) native agarose and; (B) reducing SDS-PAGE gels showing the fractions obtained after sucrose gradient centrifugation of the Qβ C74S Y89F Y99F, N10C*, K16C*, T75C* and E103C* mutants. C: Qβ WT control; M: protein marker.

Table 2.6: The expression yield and the size of Qβ mutants

Qβ mutant	Yield (mg / L of media)	Radius of particle (nm)
C74S Y89F Y99F	34.1	16.6 ± 0.6
N10C*	20.9	22.8 ± 0.9 †
K16C*	32.0	21.7 ± 1.1
T75C*	35.3	16.5 ± 0.5
E103C*	35.6	16.9 ± 0.7

† Unresolved peaks were identified by the software, indicating the presence of some aggregates.

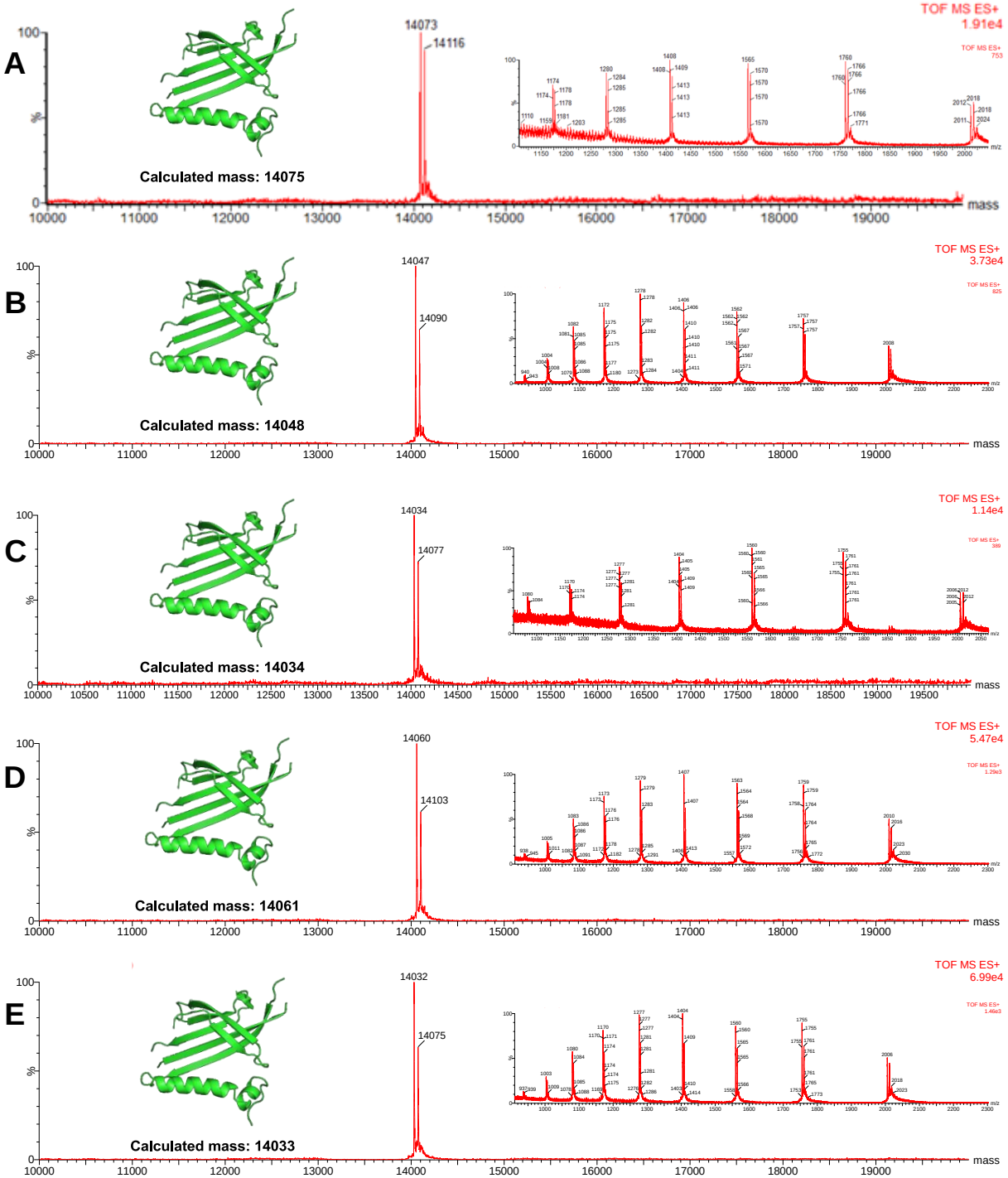


Figure 2.17: Deconvoluted mass spectra and corresponding raw ion series of the purified (A) Qβ C74S Y89F Y99F; (B) Qβ N10C*; (C) Qβ K16C*; (D) Qβ T75C* and; (E) Qβ E103C* VLPs. The VLPs were broken down to monomers for LC-MS analysis. The M + 42 ± 1 peak observed was likely to be the acetonitrile adduct from the LC-MS itself.[23, 24]

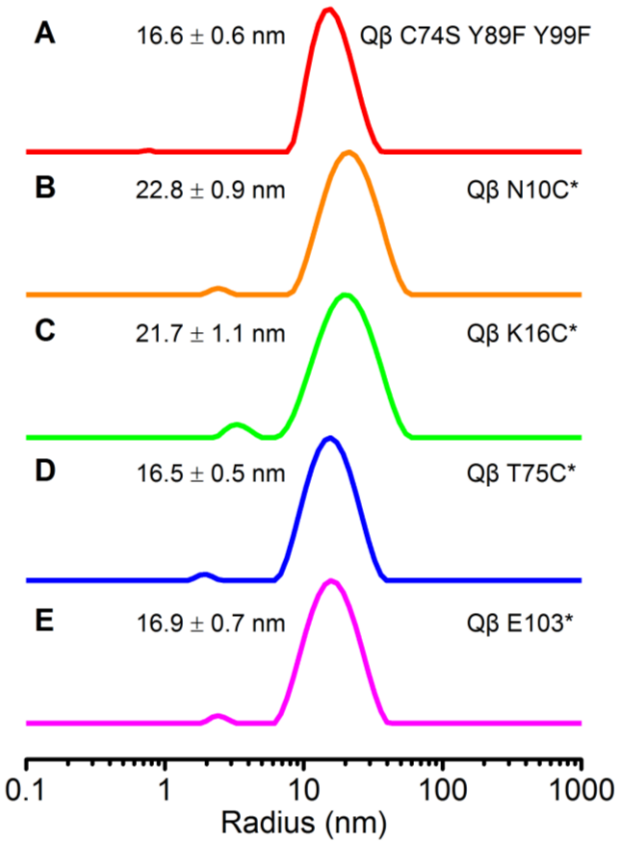


Figure 2.18: DLS results of the purified (A) Qβ C74S Y89F Y99F; (B) Qβ N10C*; (C) Qβ K16C*; (D) Qβ T75C* and (E) Qβ E103C* VLPs.

2.2.2.2 Conversion of cysteine to dehydroalanine

Conversion of cysteine to Dha on the newly expressed Qβ mutants was achieved by using the optimised protocol mentioned earlier. Although the peak corresponding to Dha can be observed from all the five species, only Qβ C74S Y89F Y99F (**Figure 2.19A**), Qβ N10C* (**Figure 2.19B**) and Qβ T75C* (**Figure 2.19D**) showed complete conversion. For Qβ K16C*, approximately 90% conversion was observed (**Figure 2.19C**), but it was found that the Dha species degraded very quickly at room temperature. It is thus not suitable for Selectfluor™ reaction which needs to be carried out at room temperature. For Qβ E103C*, a lot of unknown side products were observed beside the desired Dha peak (**Figure 2.19E**). The three species with complete conversion were tested with both Ellman's reagent (**Figure 2.20**) and 2-mercaptoethanol (**Figure 2.21**) to confirm the presence of Dha.

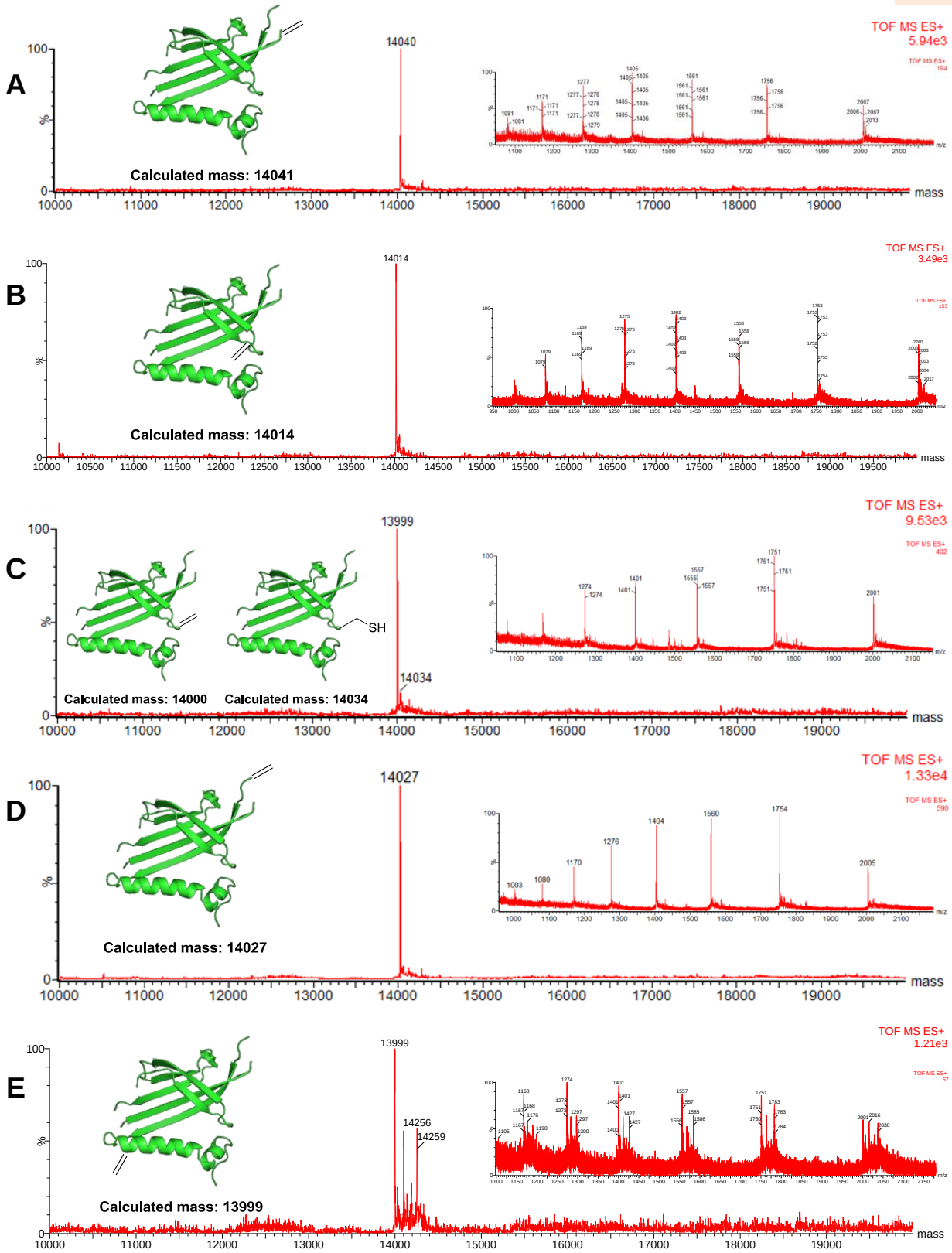
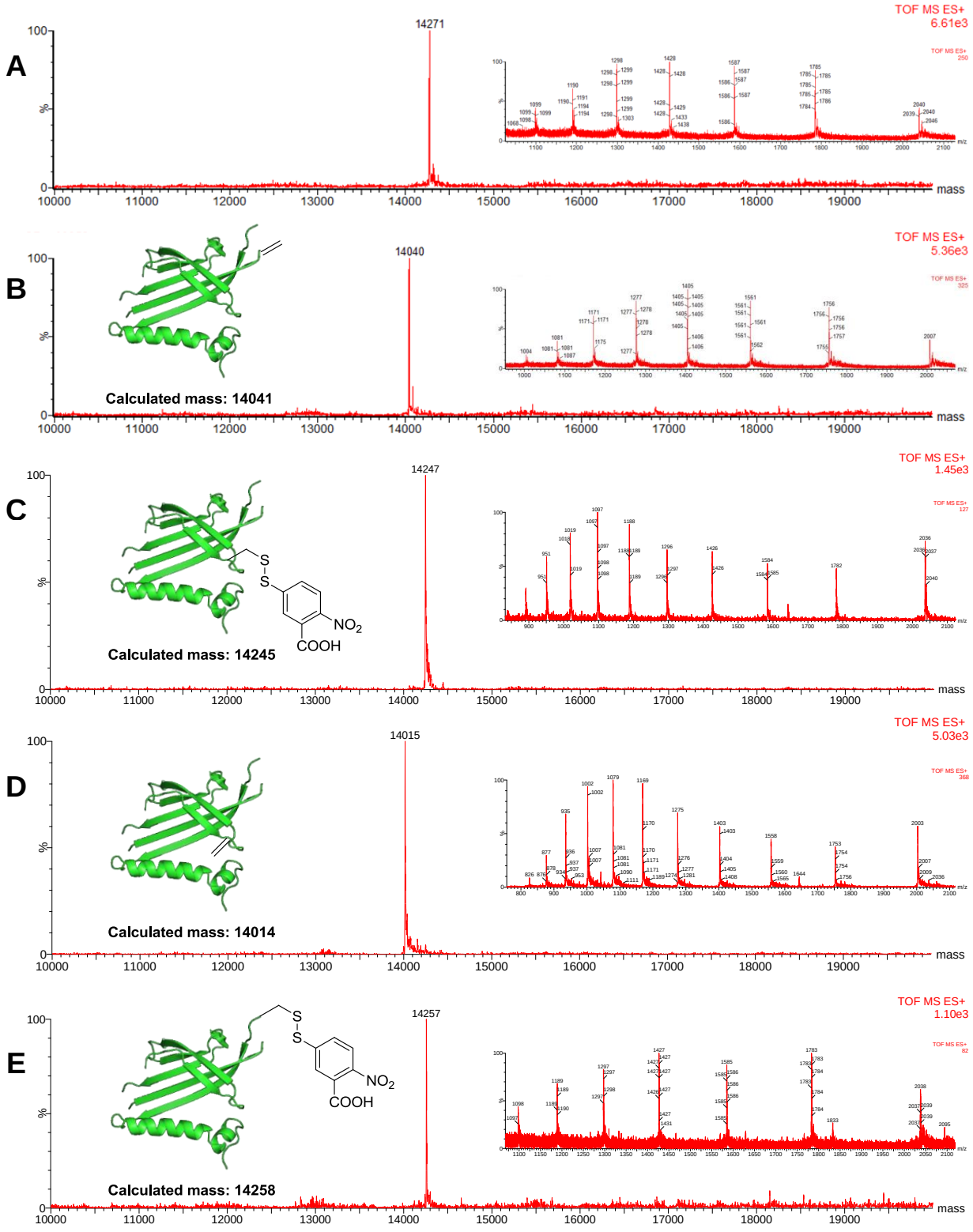


Figure 2.19: Deconvoluted mass spectra and corresponding raw ion series of the purified (A) Qβ C74S C80Dha Y89F Y99F; (B) Qβ N10Dha*; (C) Qβ K16Dha*; (D) Qβ T75Dha* and; (E) Qβ E103Dha* VLPs. The VLPs were broken down to monomers for LC-MS analysis.


A
TOF MS ES+
6.61e3
14271
mass
B
TOF MS ES+
5.36e3
14040
Calculated mass: 14041
mass
C
TOF MS ES+
1.45e3
14247
Calculated mass: 14245
mass
D
TOF MS ES+
5.03e3
14015
Calculated mass: 14014
mass
E
TOF MS ES+
1.10e3
14257
Calculated mass: 14258
mass

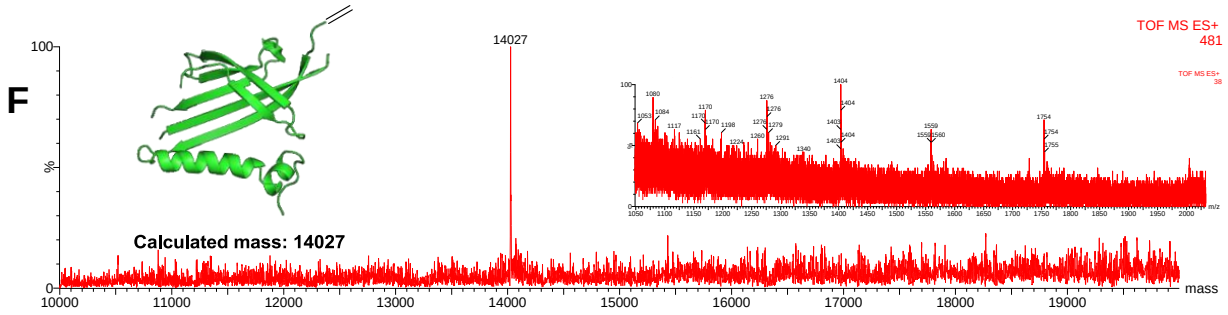


Figure 2.20: Deconvoluted mass spectra and corresponding raw ion series showing the reaction of Ellman's reagent (100 equiv.) with (A) Qβ C74S Y89F Y99F; (B) Qβ C74S C80Dha Y89F Y99F; (C) Qβ N10C*; (D) Qβ N10Dha*; (E) Qβ T75C* or; (F) Qβ T75Dha* VLPs. The VLPs were broken down to monomers for LC-MS analysis.

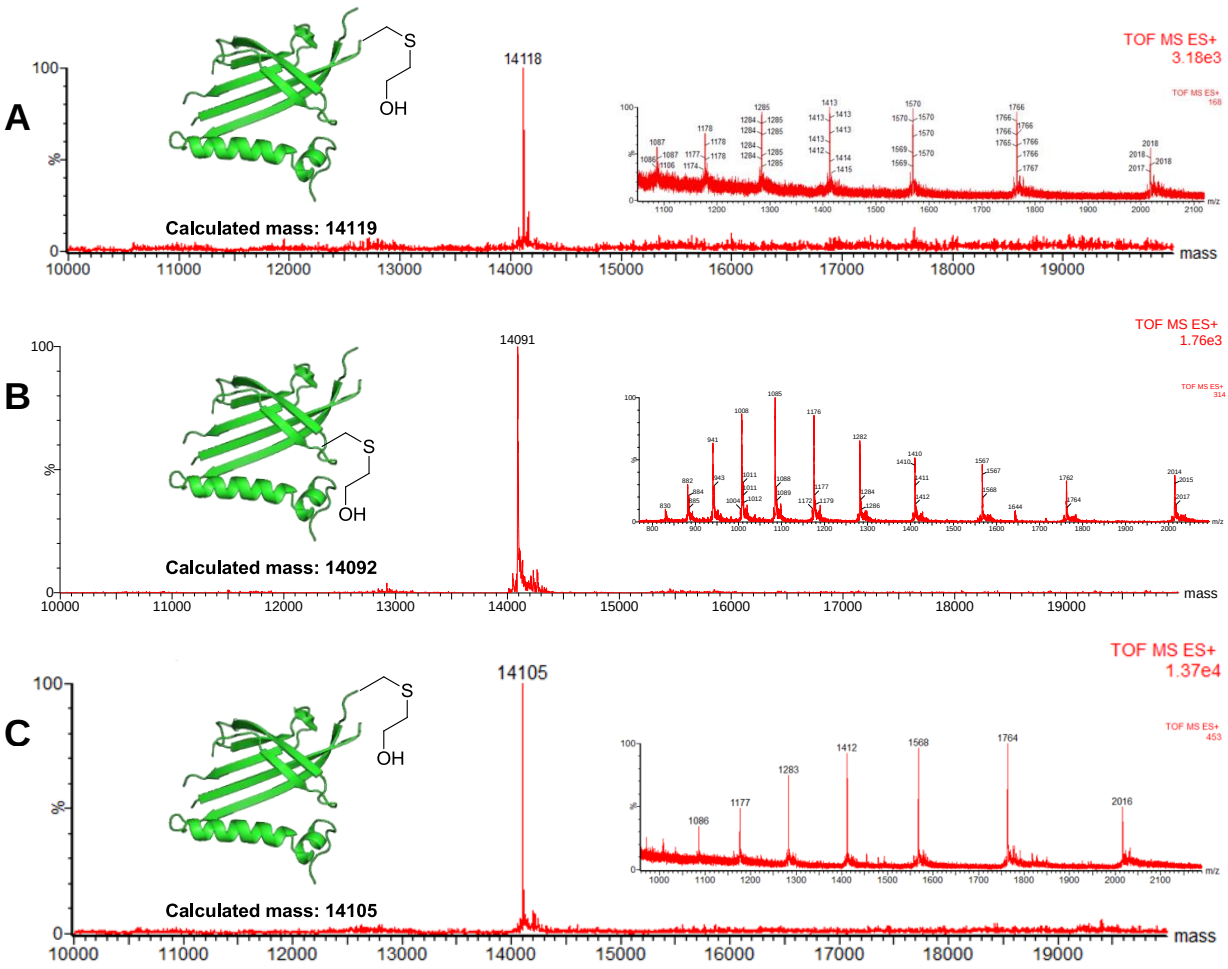


Figure 2.21: Deconvoluted mass spectra and corresponding raw ion series showing the reaction of 2-mercaptoethanol (1000 equiv.) with (A) Qβ C74S C80Dha Y89F Y99F; (B) Qβ N10Dha* or; (C) Qβ T75Dha* VLPs. The VLPs were broken down to monomers for LC-MS analysis.

2.2.2.3 Fluorination of Qβ VLPs with Selectfluor™

The three Dha species without reactive tyrosine residues were then tested with different equiv. of Selectfluor™ (**Table 2.7**). Surprisingly, fluorination of Qβ C74S C80Dha Y89F Y99F was not observed even when 1000 equiv. of Selectfluor™ was added. It is possible that the removal of tyrosine residues altered interactions such as hydrogen bonds between neighbouring residues, so that the Dha residue was hindered from the bulky Selectfluor™ reagent. On the other hand, the Qβ N10Dha* species was also found to be inert towards Selectfluor™, while a tiny peak with a mass of M + 18 was observed from the Qβ T75Dha* species (**Figure 2.22**). Interestingly, the intensity of the M + 18 peak did not increase with the amount of Selectfluor™ added to the Qβ T75Dha* species. The conversion yield is always in between 10 to 15%. It is difficult to tell whether this peak represents the formation of Dfa unless peptide mapping is done. There is a possibility that the two remaining tyrosine residues (i.e. Tyr-62 and Tyr-132) can become exposed after protein mutagenesis.

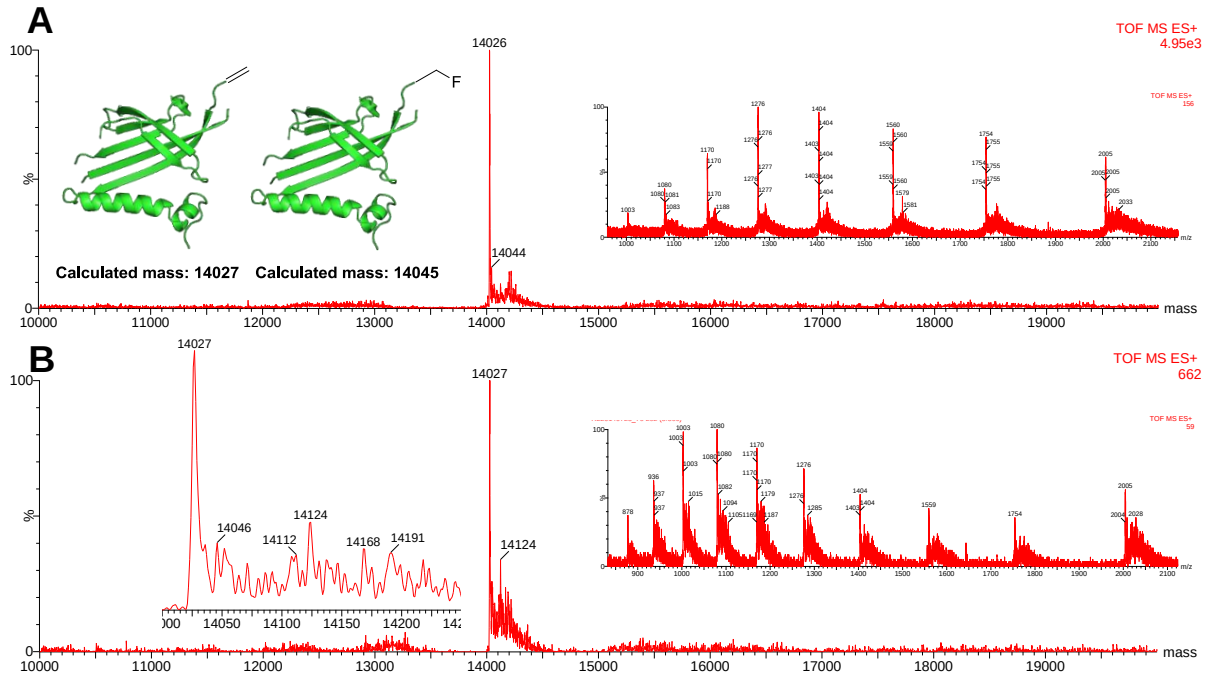


Figure 2.22: Deconvoluted mass spectra and corresponding raw ion series showing the reaction of Qβ C74S C80Dha VLPs with (A) 10 equiv. or; (B) 100 equiv. of Selectfluor™ BF_4^-. The VLPs were broken down to monomers for LC-MS analysis.

Table 2.7 Reaction of Qβ Dha species with Selectfluor™

Entry	Qβ species	Selectfluor™		Co-solvent	Product Conv. (%)
		Anion	Equiv.		
1	C74S C80Dha Y89F Y99F	BF_4^-	10	-	Dha (100)
2	C74S C80Dha Y89F Y99F	BF_4^-	20	-	Dha (100)
3	C74S C80Dha Y89F Y99F	BF_4^-	40	-	Dha (100)
4	C74S C80Dha Y89F Y99F	BF_4^-	500	-	*
5	C74S C80Dha Y89F Y99F	BF_4^-	1000	-	*
6	N10Dha*	BF_4^-	1	-	Dha (100)
7	N10Dha*	BF_4^-	5	-	Dha (100)
8	N10Dha*	BF_4^-	10	-	Dha (100)
9	N10Dha*	BF_4^-	40	-	*
10	T75Dha*	BF_4^-	1	-	Dha (90) Dha + 18 (10)
11	T75Dha*	BF_4^-	5	-	Dha (90) Dha + 18 (10)
12	T75Dha*	BF_4^-	10	-	Dha (85) Dha + 18 (15)
13	T75Dha*	BF_4^-	40	-	Dha (90) Dha + 18 (10)
14	T75Dha*	BF_4^-	100	-	*
15	T75Dha*	BF_4^-	200	-	*
16	T75Dha*	OTf^-	20	7% Acetone	Dha (90) Dha + 18 (10)

General conditions: Qβ Dha species in NaPi (50 mM, pH 8.0) and Selectfluor™ in H_2O at 25 °C for 30 min unless otherwise indicated; * the signal-to-noise ratio was so high that the spectrum cannot be analysed properly

The fluorinated VLPs were also analysed by both SDS-PAGE (**Figure 2.23**) and DLS (**Figure 2.24**) to study their stability. Again, it can be observed that dimer formations are also enhanced during the formation of Dha. Although the size of the particles was found to be slightly larger after the formation of Dha and the reaction with Selectfluor™, they were quite stable at room temperature and remained as intact particles.

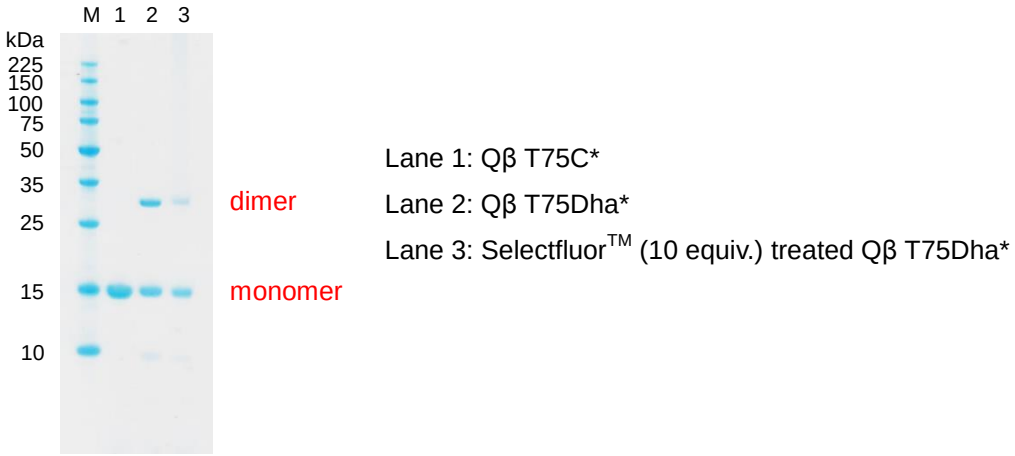


Figure 2.23: Coomassie Blue stained reducing SDS-PAGE gel of Qβ T75C* and its reaction products. M: protein marker.

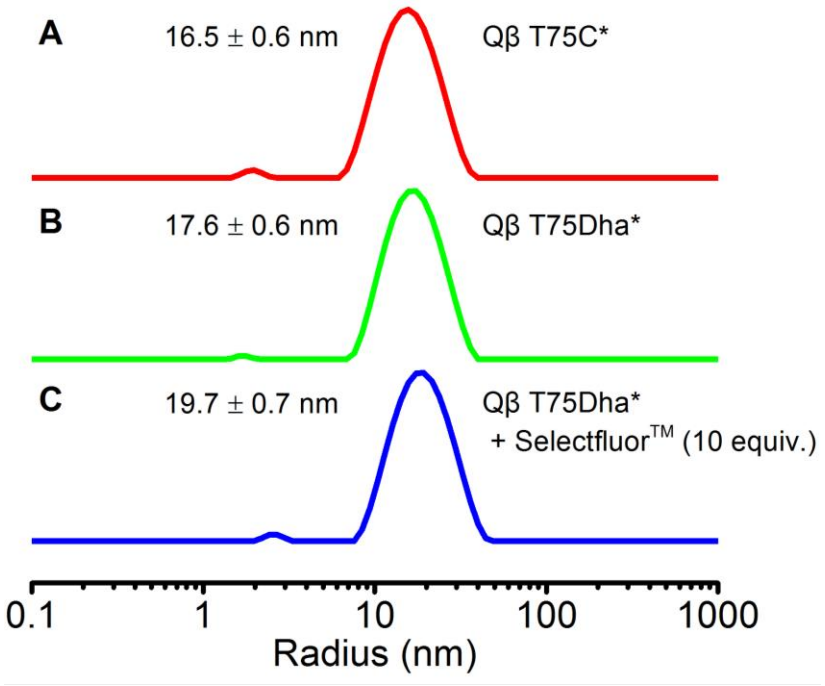


Figure 2.24: DLS results of the purified (A) Qβ T75C*; (B) Qβ T75Dha* and; (C) Selectfluor™ (10 equiv.) treated Qβ T75Dha* VLPs.

2.2.2.4 Peptide mapping of the fluorinated particles

To confirm the formation of Dfa, the fluorinated Qβ T75Dha* VLP was analysed by MS/MS. As expected, Dfa-75 was observed by MS/MS (**Table 2.8** and **Figure 2.25**). Side reactions on both the two remaining tyrosine residues (i.e. Tyr-62 and Tyr-132) were not observed, indicating that the peak observed in the mass spectra is solely contributed by Dfa formation. The presence of Dfa was further proved by the 2-mercaptoethanol trapping control as shown by the MS/MS spectrum (**Table 2.9** and **Figure 2.26**).

Table 2.8 Key digested peptide fragments observed from Mascot analysis of fluorinated Qβ T75Dha* (protein sequence coverage: 100%)

Score	Mr(found)	Mr(Calc)	Error (ppm)	Expect	Modifications	Peptide sequence
86.8	1869.8827	1869.8817	0.52	2.1e-09	Cys75→Dha75	68-IQNPTAS**C**ANGSSDPSVTR-86
79.8	1887.8788	1887.8723	3.44	1.1e-08	Cys75→Dfa75	68-IQNPTAS**C**ANGSSDPSVTR-86

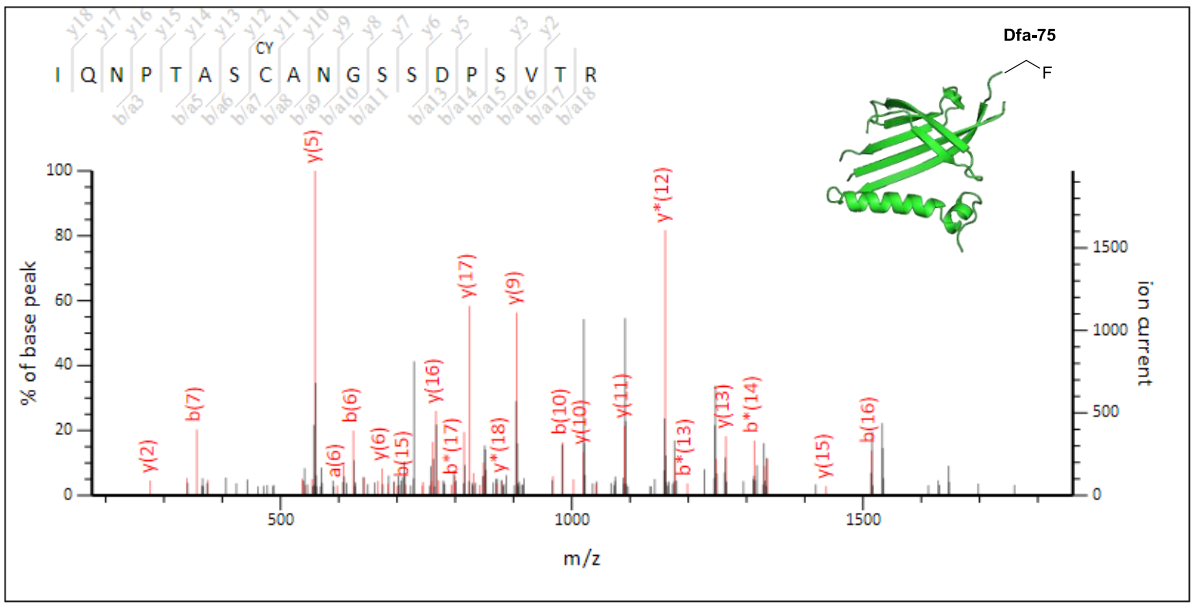


Figure 2.25: ESI-QTOF MS-MS spectrum of peptide 68-IQNPTASCANGSSDPSVTR-86 obtained by tryptic digestion of Selectfluor™ (10 equiv.) treated Qβ T75Dha* that contains Dfa-75 (Score = 79.8).

Table 2.9 Key digested peptide fragments observed from Mascot analysis of fluorinated Qβ T75Dha* after the addition of 2-mercaptoethanol (protein sequence coverage: 100%)

Score	Mr(found)	Mr(Calc)	Error (ppm)	Expect	Modifications†	Peptide sequence
99.2	1947.8955	1947.8956	-0.068	1.2e-10	Cys75→Dha_bME75	68-IQNPTAS**C**ANGSSDPSVTR-86
20.8	1965.8886	1965.8862	1.21	0.0083	Cys75→Dfa_bME75	68-IQNPTAS**C**ANGSSDPSVTR-86

† Dha_bME: 2-mercaptoethanol addition on Dha; Dfa_bME: 2-mercaptoethanol addition on Dfa

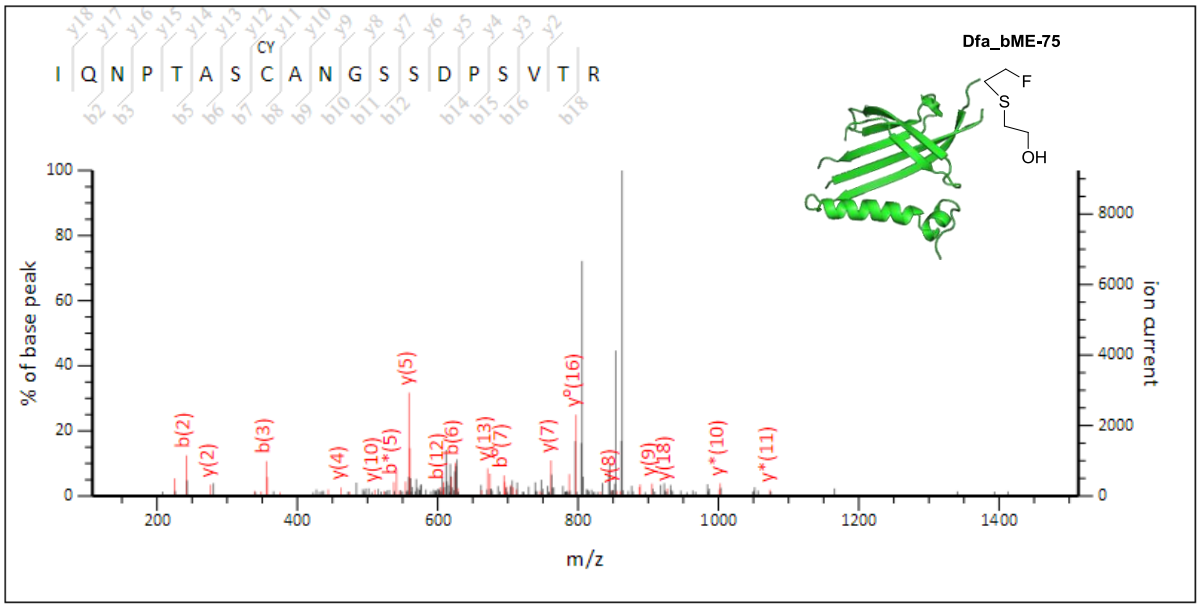


Figure 2.26: ESI-QTOF MS-MS spectrum of peptide 68-IQNPTASCANGSSDPSVTR-86 obtained by tryptic digestion of Selectfluor™ (10 equiv.) and 2-mercaptoethanol treated Qβ T75Dha* that contains the coupled product Dfa_bME-75 (Score = 20.8).

2.2.2.5 Accessibility of the Dha residue

It remains a mystery that only Qβ T75Dha* can react with Selectfluor™ but not the other two Qβ Dha species. One of the possible reasons would be the accessibility of the Dha residue on different species towards Selectfluor™. Since Selectfluor™ is a relatively bulky molecule, it can only be attacked by the residues exposed on the protein surface. To get a better idea about the reactivity of Dha in different species, the Qβ Dha VLPs were treated with the bulky thiol-functionalised *N*-acetylglucosamine (GlcNAc-SH) to see if any coupled product was formed. As shown by the mass spectra (**Figure 2.27**), the coupled products were

only observed in the cases of Qβ C74S C80Dha Y89F Y99F and Qβ T75Dha*, with the later case giving significantly higher conversion. Since Dha is likely less reactive to Selectfluor™ than GlcNAc-SH, these results gave us a hint about the lack of reactivity of Qβ C74S C80Dha Y89F Y99F and Qβ N10Dha* towards Selectfluor™ compared to Qβ T75Dha*.

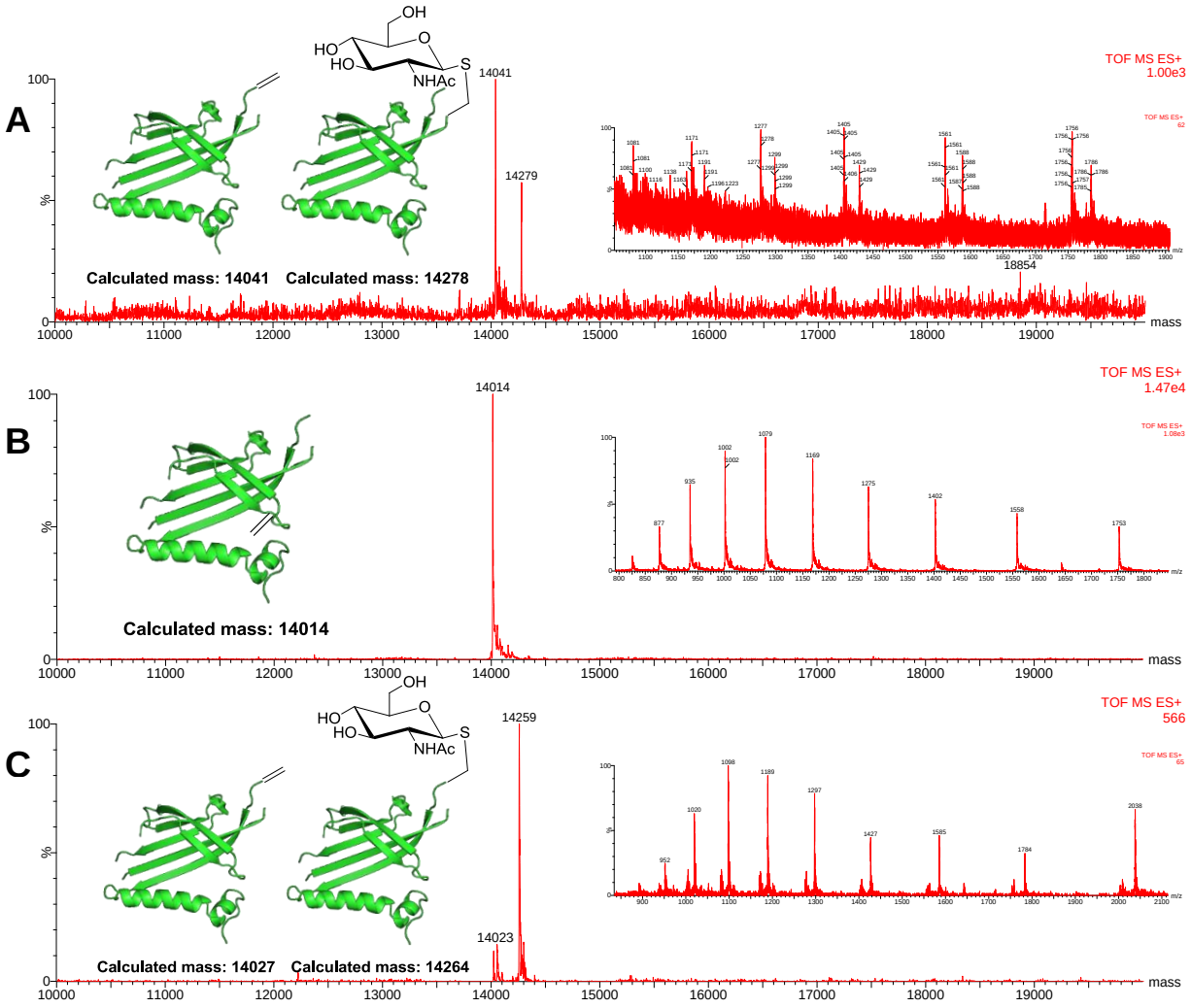


Figure 2.27: Deconvoluted mass spectra and corresponding raw ion series showing the reaction of GlcNAc-SH (100 equiv.) with (A) Qβ C74S C80Dha Y89F Y99F; (B) Qβ N10Dha* or; (C) Qβ T75Dha* VLPs. The VLPs were broken down to monomers for LC-MS analysis.

2.3 Conclusion

In this chapter, Dha species of Qβ VLPs, including Qβ C74S C80Dha, Qβ C74S C80Dha Y89F Y99F, Qβ N10Dha* and Qβ T75Dha*, were produced for the first time with complete conversion; the products exhibited good stability. Although fluorination of Dha on Qβ C74S C80Dha VLPs using the electrophilic fluorine source Selectfluor™ was achieved, side reactions on tyrosine residues were also observed. After removing the reactive tyrosine residues, pure Dfa formation can be achieved with Qβ T75Dha* VLPs, though with a relatively low conversion yield. However, to the best of our knowledge, this is the first demonstration of direct C-F bond formation on Qβ VLPs at a specific position using the 'tag-and-modify' approach, without affecting the stability of the particle. Further optimisation work will be carried out in the near future to increase the conversion.

In this section, Hfa (–NH–C(CH_2F)R–C(O)– where R = OH) and α-(hydroxyethylthio)-β-fluoro-alanine (Dfa_bME; R = –S(CH_2)$_2$OH) were shown as two examples of α,α-substituted amino acid that may be formed by the Selectfluor™ reaction. Indeed, as revealed previously by group members, other functional groups such as R = –OMe and R = –N_3 can be installed to the α-carbon by addition of a small amount of MeOH and NaN_3 respectively.[15] This not only confers conformational strains upon the peptide backbone,[25] but also allows direct modulation of extra protein properties in addition to C-F bond formation. Last but not least, the use of ^{18}F Selectfluor™ would provide a new way for direct C–^{18}F bond formation on Qβ VLP, which can be imaged by PET when it is applied as a nano-carrier of vaccines or drugs.

2.4 Experimental Section

2.4.1 General procedures

Chemicals were purchased from commercial suppliers and used without purification. NuPAGE® MES buffer and 12% Bis-Tris gel were purchased from Life Technologies™. Protein markers used were Perfect Protein™ Marker 10-225 kDa purchased from Novagen® . Vivaspin® concentrators were purchased from Life Technologies™. PD SpinTrap™ G-25, MiniTrap™ G-25, MidiTrap™ G-25 and PD10 desalting columns were purchased from GE Healthcare Life Sciences. Dialysis cassettes and devices (10,000 MW) were purchased from Thermo Scientific® . QIAprep Miniprep Kit was purchased from QIAGEN. The PfuUltra™ high-fidelity DNA polymerase AD was purchased from Agilent Technologies. Restriction enzymes used were purchased from New England Biolabs. Primers used were purchased from Sigma Alrich. The bacterial strains used were BL21(DE3) and XL10 gold purchased from Novagen and Agilent Technologies respectively. Unless mentioned specifically, all the Selectfluor™ used was Selectfluor™ BF_4^-.

PCR was done using an Applied Biosystems® 2720 Thermocycler. DNA sequencing was performed by Source BioScience. Ultracentrifugation was performed in Beckman Optima XL-80 Ultracentrifuge with SW 41 Ti rotor and Ultra-clear™ centrifuge tubes. Dynamic light scattering analysis was performed in Viscotek 801 DLS. Abs_{280} measurement for protein concentration was performed in NanoDrop ND-1000. Mass spectrometry analysis was performed in Waters LCT Premier™. Gel filtration was performed in GE Ä KTAFPLC ™.

2.4.2 Mutagenesis of Qβ

2.4.2.1 Primers used

Mutant	Template		Primer sequence (5' → 3')
N10C	WT	F	GTTACTTTAGGTTGCATCGGGAAAG
		R	CTTTCCCGATGCAACCTAAAGTAAC
K16C	WT	F	GTAACATCGGGAAAGATGGATGTCAAACTCTGGTCCTCAATCCG
		R	CGGATTGAGGACCAGAGTTTGACATCCATCTTTCCCGATGTTAC
T18C	WT	F	GAAAGATGGAAAACAATGTCTGGTCCTCAATCCGC
		R	GCGGATTGAGGACCAGACATTGTTTTCCATCTTTC
C74S	WT	F	CCGACCGCTAGCACTGCAAAC
		R	GTTTGCAGTGCTAGCGGTCGG
C74S C80S	C74S	F	AAACGGTTCTTCGGACCCATCCG
		R	CGGATGGGTCCGAAGAACCGTTT
C74S Y89F Y99F	C74S	F	GCCAGGCATTTGCTGACGTGACCTTTTCGTTCACGCAGTTTAGTACCG
C74S C80S Y89F Y99F (*)	C74S C80S	R	CGGTACTAAACTGCGTGAACGAAAAGGTCACGTCAGCAAATGCCTGGC
N10C*	*	F	GTTACTTTAGGTTGCATCGGGAAAG
		R	CTTTCCCGATGCAACCTAAAGTAAC
K16C*	*	F	GTAACATCGGGAAAGATGGATGTCAAACTCTGGTCCTCAATCCG
		R	CGGATTGAGGACCAGAGTTTGACATCCATCTTTCCCGATGTTAC
T75C*	*	F	CCGAAGAACCGTTTGCACAGCTAGCGGTCGGGTTC
		R	GAACCCGACCGCTAGCTGTGCAAACGGTTCTTCGG
E103C*	*	F	CACGCAGTTTAGTACCGATTGTGAACGAGCTTTTGTTCGTAC
		R	GTACGAACAAAAGCTCGTTCACAATCGGTACTAAACTGCGTG

2.4.2.2 Protocols

The DNA plasmid of Qβ was mixed with forward and reverse primers, dNTP, enzyme buffer and PfuUltra™ high-fidelity DNA polymerase AD according to the protocol from Agilent Quikchange kit. The PCR was performed using the following conditions: initial denaturation at 95 ℃ for 1 min; 18 cycles of denaturation at 95 ℃ for 1 min, annealing at 55 ℃ for 1 min, extension at 72 ℃ for 8 min; and then incubated at 72 ^{o}C for another 8 min

and cooled to 20 °C. The reaction mixture (50 µL) was incubated with DpnI (1 µL) at 37 °C for 2 h. The DpnI treated PCR product (2 µL) was added to XL-10 gold cell (50 µL) with 2-mercaptoethanol (2 µL) in a round bottomed tube and incubated on ice for 30 min. The tube was heat-shocked at 42 °C for 30 sec and immediately placed on ice for 2 min. LB medium (450 µL) was added to the tube and the cell culture was shaken at 37 °C for 1 h. The cell culture (200 µL) was spread on an ampicillin LB agar plate and the plate was incubated at 37 °C overnight. Colony was picked from the plate and grown in LB medium (with 100 µg/mL ampicillin) (8 mL) at 37 °C overnight. The cell culture was centrifuged at 7200 x g for 3 min and the DNA plasmid was extracted from the cells using the Miniprep Kit. The DNA plasmid was then analysed by DNA sequencing.

2.4.3 Preparation of glycerol stock from DNA plasmid

The DNA plasmid obtained from mutagenesis (2 µL) was added to BL21(DE3) cell (50 µL) in a round bottomed tube and incubated on ice for 30 min. The tube was heat-shocked at 42 °C for 30 sec and immediately placed on ice for 2 min. LB medium (450 µL) was added to the tube and the cell culture was shaken at 37 °C for 1 h. The cell culture (200 µL) was spread on an ampicillin LB agar plate and the plate was incubated at 37 °C overnight. A colony was picked from the plate and grown in LB medium (with 100 µg / mL ampicillin) (8 mL) at 37 °C overnight. The overnight culture (800 µL) was mixed with 50% glycerol (800 µL) and stored at -80 °C.

2.4.4 Expression of Qβ VLP

Glycerol stock of BL21(DE3) (1 µL) was added to LB medium (with 100 µg / mL ampicillin) (20 mL). The mixture was shaken at 37 ^{o}C for 16 h with a speed of 250 rpm. The overnight culture (8 mL) was added to two flasks each containing LB medium (with 100 µg / mL ampicillin) (750 mL). The cell culture was shaken at 37 ^{o}C with a speed of 250 rpm until OD_{600} reached ~0.6. 1 mM IPTG was then added and the protein was expressed at 37 ^{o}C for 6 h. The cells were then harvested by centrifugation at 9,500 x g for 10 min at 4 ^{o}C. The cell pellet was resuspended in TBS buffer (50 mM Tris, 150 mM NaCl, pH 7.5) (20 mL). After sonication (5 min, 20% power, 2 sec ON/OFF), the lysed solution was pelleted by centrifugation at 58,000 x g for 30 min at 4 ^{o}C.

2.4.5 Purification of Qβ VLP

2.4.5.1 Sucrose gradient purification

The supernatant obtained was mixed with an equal volume of 1:1 v/v butanol:chloroform by inversion and vortexing. The mixture was centrifuged at 3000 x g for 30 min at 4 ^{o}C. The aqueous layer on top was mixed and rocked with 10% w/v PEG8000 and 100 mM NaCl for 30 min at 4 ^{o}C. The mixture was then centrifuged at 3000 x g for 30 min at 4 ^{o}C, and the pellet was resuspended in TBS (10 mL). The above purification steps were repeated once depending on the trial. The resulting solution was further purified by sucrose gradient centrifugation (10%, 15%, 20%, 25%, 30%, 35%, 35% and 40% sucrose solution) at 170,000 x g for 13 h at 4^{o}C. 1 mL fractions from each tube were collected and analysed by SDS-PAGE and native agarose gel. Fractions containing Qβ particles were combined together and dialysed (10,000 MW) thrice in TBS or 1 mM DTT in TBS (for species with a single cysteine per monomer) at 4 ^{o}C. The resulting solution was analysed by DLS, LCT mass spectrometry, BCA assay and UV absorbance.

2.4.5.2 Size-exclusion chromatography (for VLPs showing aggregate peak in DLS)

Self-packed 10/30 column with GE Sephacryl S-500 HR resin was used for the chromatography. The column was equilibrated by 1.1 column volume of water and TBS respectively. Protein sample (500 µL) was loaded into a 1 mL loop and injected into the column. The protein was eluted by TBS with a flow rate of 0.2 - 0.5 mL/min depending on the back pressure (maximum set to 0.2 MPa). The protein was collected in small fractions (500 µL). Fractions containing the desired product was combined and concentrated by Vivaspin® (30,000 MW). A typical example of chromatogram is shown in **Figure S.1**.

2.4.6 Characterization of Qβ VLPs

2.4.6.1 LC-MS analysis

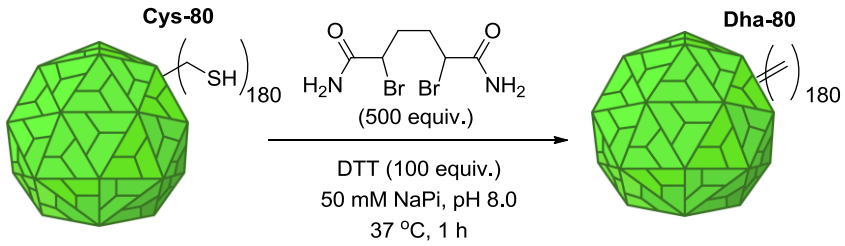
Unless otherwise specified, the sample was treated with 4.5 M urea or 4.5 M urea and 50 mM DTT (for species with disulphide bonds) in TBS or sodium phosphate buffer at 25 ^{o}C for 1 h before LC-MS analysis. The ion signals obtained during the period of t = 1.50 to t = 2.25 were integrated and analysed (**Figure S.2**). The mass spectra were reconstructed from the ion series and deconvoluted using the MaxEnt1 algorithm in the MassLynx software (v 4.1 from Waters).

2.4.6.2 DLS analysis

The protein solution was centrifuged at 13,000 x g for 20 min before taking the supernatant for analysis. The temperature of the machine was set to 20 ^{o}C. The viscosity values for TBS and sodium phosphate buffer were set to 1.0185 cP and 1.0190 cP respectively. The reflective index values for TBS and sodium phosphate buffer were set to 1.3345 and 1.3330 respectively. At least 30 measurements were carried out for each analysis (10 seconds for each run). Intensity distribution was shown.

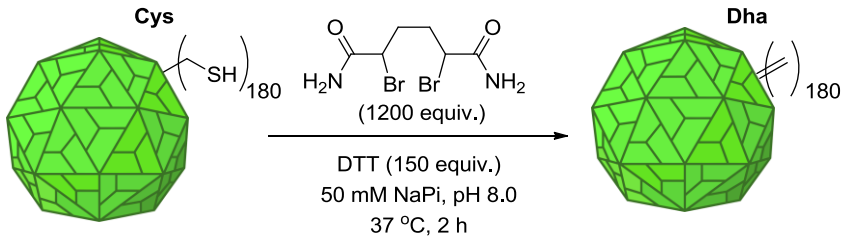
2.4.7 Dehydroalanine formation on Qβ VLPs

2.4.7.1 Qβ C74S



Qβ C74S VLPs (1 mg / mL) in TBS with 1 mM DTT were buffer exchanged using PD SpinTrap™ G-25 / MiniTrap™ G-25 / MidiTrap™ G-25 / PD10 desalting column into sodium phosphate buffer (50 mM, pH 8.0) depending on the scale. The protein solution was mixed with DTT (100 equiv.) and shaken at 25 °C for 30 min. DBHDA in DMF (100 mg/mL) (500 equiv.) was added to the protein solution and shaken at 37 °C for 1 h. The final concentration of DMF in the reaction mixture should not be more than 5%. The solution can be diluted with sodium phosphate buffer if necessary. The DBHDA was removed by washing with sodium phosphate buffer in Vivaspin™ (10,000 MW) several times.

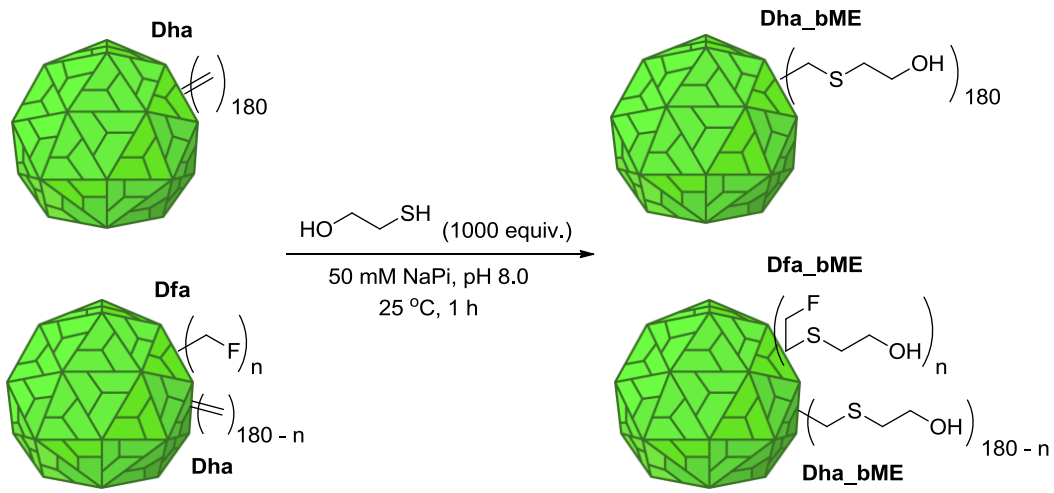
2.4.7.2 Qβ C74S Y89F Y99F / Qβ N10C* / Qβ T75C*



Qβ VLPs (1 mg / mL) in TBS with 1 mM DTT were buffer exchanged using PD SpinTrap™ G-25 / MiniTrap™ G-25 / MidiTrap™ G-25 / PD10 desalting column into sodium phosphate buffer (50 mM, pH 8.0) depending on the scale. The protein solution was mixed with DTT (100 equiv.) and shaken at 25 °C for 30 min. DBHDA in DMF (100 mg/mL) (600 equiv.) was added to the protein solution and shaken at 37 °C for 1 h. The final

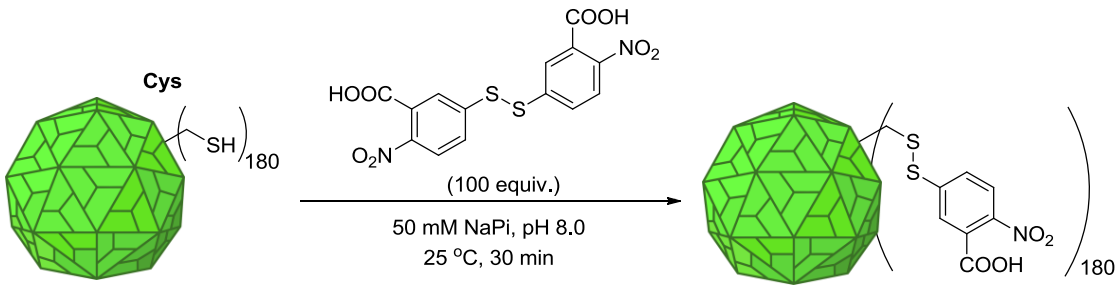
concentration of DMF in the reaction mixture should not be more than 5%. The solution can be diluted with sodium phosphate buffer if necessary. The protein solution was mixed with DTT (50 equiv.) and shaken at 25 °C for 10 min. DBHDA in DMF (100 mg/mL) (600 equiv.) was added to the protein solution and shaken at 37 °C for 1 h. The final concentration of DMF in the reaction mixture should not be more than 5%. The DBHDA was removed by washing with sodium phosphate buffer in Vivaspin™ (10,000 MW) several times.

2.4.8 Reaction with 2-mercaptoethanol



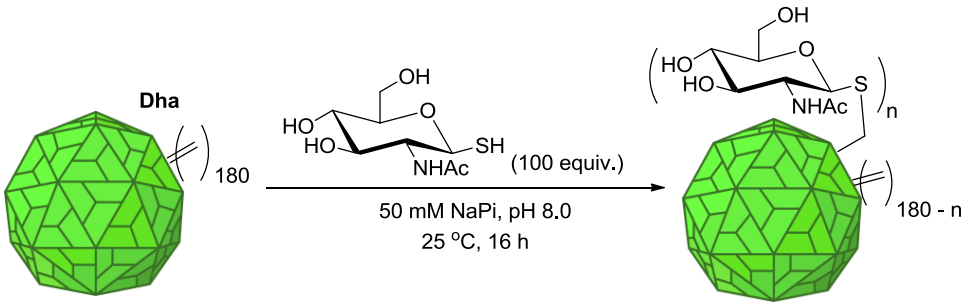
Qβ Dha VLPs or their fluorinated version (1 mg / mL) in sodium phosphate buffer (50 mM, pH 8.0) were mixed with 2-mercaptoethanol (1000 equiv.) and 2 M (for MS/MS analysis) or 4.5 M (for LC-MS analysis) urea at 25 °C for 1 h.

2.4.9 Reaction with Ellman's reagent



Qβ VLPs (1 mg / mL) in sodium phosphate buffer (50 mM, pH 8.0) were mixed with freshly prepared Ellman's reagent (100 mg / mL in DMF) (100 equiv.) and shaken at 25 °C for 30 min. The Ellman's reagent was then removed by washing with sodium phosphate buffer in Vivaspin™ (10,000 MW) several times.

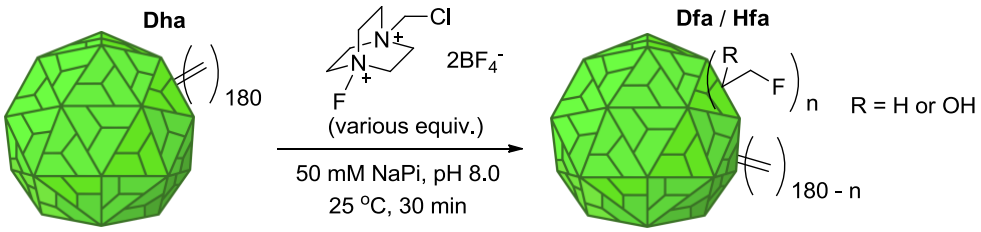
2.4.10 Reaction with GlcNAc-SH



Qβ Dha VLPs (1 mg / mL) in sodium phosphate buffer (50 mM, pH 8.0) were mixed with freshly prepared GlcNAc-SH (30 mg / mL in water) (100 equiv.) and shaken at 25 °C for 16 h.

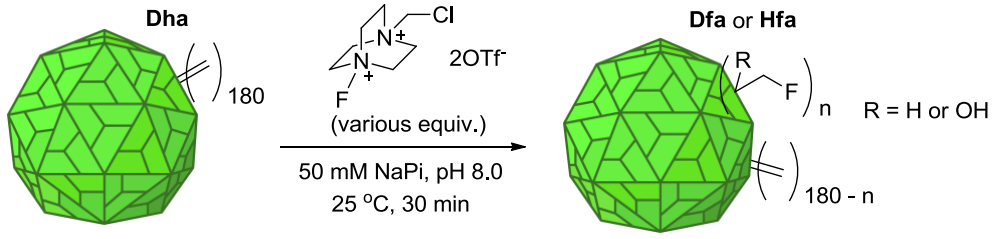
2.4.11 Fluorination of Qβ VLPs with Selectfluor™

2.4.11.1 Selectfluor™ BF_4^-



Qβ VLPs (1 mg / mL) in sodium phosphate buffer (50 mM, pH 8.0) were mixed with freshly prepared Selectfluor™ BF_4^- (5 mg/mL in H_2O) (various equiv.) and shaken at 25 °C for 30 min. The excess Selectfluor™ was removed by PD SpinTrap™ G-25 / MiniTrap™ G-25 / MidiTrap™ G-25 / PD10 desalting column with sodium phosphate buffer (50 mM, pH 8.0) depending on the scale.

2.4.11.2 Selectfluor™ OTf -



Qβ VLPs (1 mg / mL) in sodium phosphate buffer (50 mM, pH 8.0) were mixed with freshly prepared Selectfluor™ OTf - (5.5 mg/mL in acetone) (10 or 20 equiv.) and shaken at 42 °C for 30 min. The final concentration of acetone was 7% and 14% respectively for 10 and 20 equiv. of Selectfluor™ OTf -. The excess Selectfluor™ was removed by PD SpinTrap™ G-25 / MiniTrap™ G-25 / MidiTrap™ G-25 / PD10 desalting column with sodium phosphate buffer (50 mM, pH 8.0) depending on the scale.

2.4.12 MS/MS analysis (Target Discovery Institute, University of Oxford)

Qβ VLPs in sodium phosphate buffer (50 mM, pH 8.0) were mixed with 2 M urea, followed by the addition of 2-mercaptoethanol (1000 equiv.) if necessary. The reaction mixture was incubated at 25 °C for 2 h. Trypsin (0.2 mg/mL) (1:50 w/w) was added to the mixture and shaken at 37 °C overnight. 0.4% of formic acid was then added to the sample, followed by purification using ZipTip®. The purified sample was submitted to Target Discovery Institute (Dr. Benedikt M. Kessler) for MS/MS analysis. The raw data were analysed by Mascot using the following parameters: enzyme: trypsin; mass values: monoisotopic; protein mass: unrestricted; peptide charge: +2, +3, +4; peptide mass tolerance: ± 50 ppm; fragment mass tolerance: ± 0.6 Da; mas missed cleavages: 2; instrument type: ESI-QUAD-TOF.

2.5 References

1. Mascaretti, O.A. Modern methods for the monofluorination of aliphatic organic compounds. *Aldrichim. Acta* **26**, 47-58 (1997).
2. Rauniyar, V., Lackner, A.D., Hamilton, G.L. & Toste, F.D. Asymmetric Electrophilic Fluorination Using an Anionic Chiral Phase-Transfer Catalyst. *Science* **334**, 1681-1684 (2011).
3. Lee, E. et al. A Fluoride-Derived Electrophilic Late-Stage Fluorination Reagent for PET Imaging. *Science* **334**, 639-642 (2011).
4. Nyffeler, P.T., Durón, S.G., Burkart, M.D., Vincent, S.P. & Wong, C.-H. Selectfluor: Mechanistic Insight and Applications. *Angewandte Chemie International Edition* **44**, 192-212 (2005).
5. Furuya, T., Kamlet, A.S. & Ritter, T. Catalysis for fluorination and trifluoromethylation. *Nature* **473**, 470-477 (2011).
6. Liu, W. et al. Oxidative Aliphatic C-H Fluorination with Fluoride Ion Catalyzed by a Manganese Porphyrin. *Science* **337**, 1322-1325 (2012).
7. Banks, R.E., Besheesh, M.K., Mohialdin-Khaffaf, S.N. & Sharif, I. N-Halogeno compounds. Part 18. 1-Alkyl-4-fluoro-1,4-diazoniabicyclo[2.2.2]octane salts: user-friendly site-selective electrophilic fluorinating agents of the N-fluoroammonium class. *Journal of the Chemical Society, Perkin Transactions 1*, 2069-2076 (1996).
8. Teare, H. et al. Radiosynthesis and Evaluation of [18F]Selectfluor bis(triflate). *Angewandte Chemie International Edition* **49**, 6821-6824 (2010).
9. Chalker, J.M. et al. Methods for converting cysteine to dehydroalanine on peptides and proteins. *Chemical Science* **2**, 1666-1676 (2011).
10. Bernardes, G.J.L., Chalker, J.M., Errey, J.C. & Davis, B.G. Facile Conversion of Cysteine and Alkyl Cysteines to Dehydroalanine on Protein Surfaces: Versatile and Switchable Access to Functionalized Proteins. *Journal of the American Chemical Society* **130**, 5052-5053 (2008).
11. Chalker, J.M., Bernardes, G.J.L. & Davis, B.G. A "Tag-and-Modify" Approach to Site-Selective Protein Modification. *Accounts of Chemical Research* **44**, 730-741 (2011).
12. Chalker, J.M., Lercher, L., Rose, N.R., Schofield, C.J. & Davis, B.G. Conversion of Cysteine into Dehydroalanine Enables Access to Synthetic Histones Bearing Diverse Post-Translational Modifications. *Angewandte Chemie International Edition* **51**,

1835-1839 (2012).
13. Chooi, K.P. et al. Synthetic Phosphorylation of p38α Recapitulates Protein Kinase Activity. *Journal of the American Chemical Society* **136**, 1698-1701 (2014).
14. Gunnoo, S.B. et al. Creation of a gated antibody as a conditionally functional synthetic protein. *Nat Commun* **5** (2014).
15. Boutureira, O, Schombs, M., et al. (Unpublished results)
16. Wang, J., Schiller, S.M. & Schultz, P.G. A Biosynthetic Route to Dehydroalanine-Containing Proteins. *Angewandte Chemie* **119**, 6973-6975 (2007).
17. Baldwin, J.E., Adlington, R.M. & Moss, N. Synthetic Approaches to 2-Methoxycysteine containing Peptides. The conversion of [(5S)-5-amino-5-carboxy-2-oxapentanoyl]-2-methoxy-(2S)-cysteinyl-(2R)-valine into 10-Oxa-6α-methoxyisopenicillin N by the Enzyme Isopenicillin N Synthase. *Tetrahedron* **45**, 2841-2856 (1989).
18. Shin, C.-g., Abe, C. & Yonezawa, Y. Total Synthesis of Bistratamide G, a Metabolite of the Philippines Ascidian Lissoclinum bistratum, from Dehydrotripeptides. *Chemistry Letters* **33**, 664-665 (2004).
19. Boutureira, O. et al. Fluoroglycoproteins: ready chemical site-selective incorporation of fluorosugars into proteins. *Chem. Commun. (Cambridge, U. K.)* **46**, 8142-8144 (2010).
20. Shepherd, C.M. et al. VIPERdb: a relational database for structural virology. *Nucleic Acids Res* **34**, D386-9 (2006).
21. Fiedler, J.D. et al. Engineered Mutations Change the Structure and Stability of a Virus-Like Particle. *Biomacromolecules* **13**, 2339-2348 (2012).
22. Ribeiro-Viana, R. et al. Virus-like glycodendrinanoparticles displaying quasi-equivalent nested polyvalency upon glycoprotein platforms potently block viral infection. *Nat Commun* **3**, 1303 (2012).
23. Li, B. et al. Histone H3 Lysine 36 Dimethylation (H3K36me2) Is Sufficient to Recruit the Rpd3s Histone Deacetylase Complex and to Repress Spurious Transcription. *Journal of Biological Chemistry* **284**, 7970-7976 (2009).
24. Simon, M.D. et al. The Site-Specific Installation of Methyl-Lysine Analogs into Recombinant Histones. *Cell* **128**, 1003-1012 (2007).
25. Toniolo, C., Formaggio, F., Kaptein, B. & Broxterman, Q.B. You Are Sitting on a Gold Mine! *Synlett* **2006**, 1295-1310 (2006).

CHAPTER

3 Electrophilic aromatic fluorination on tyrosine residues

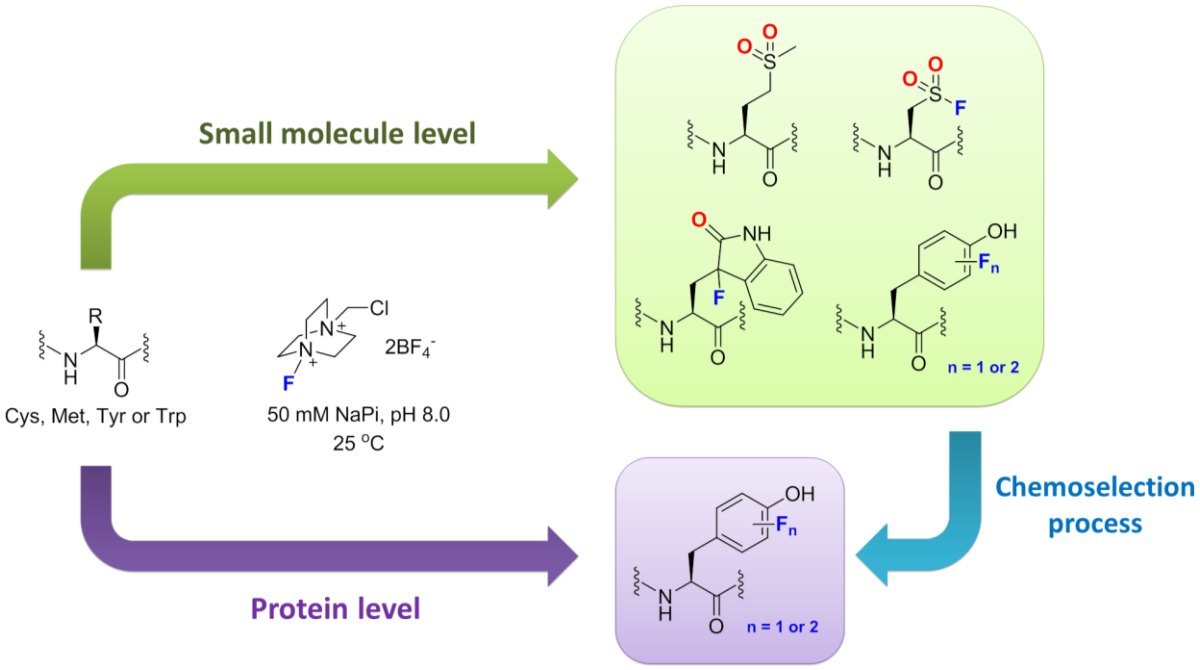
3.1 Introduction

3.1.1 The 'side reaction' of Selectfluor™

As mentioned in the previous chapter, fluorination on tyrosine residues was observed as a side reaction of the 'tag-and-modify' approach. As shown previously by our group members, side reactions of Selectfluor™ can indeed occur on tryptophan, tyrosine, cysteine and methionine residues (**Scheme 3.1**).[1] These include both fluorination and oxidation, depending on the type of the residue. In complicated macromolecules like proteins, however, the reactivity of these residues greatly depends on their accessibility. For example, Qβ WT VLP is inert to Selectfluor™ even it possesses both cysteine and tyrosine residues. This can likely be accounted by the presence of disulphide bonds and the inaccessibility of the tyrosine residues. Nevertheless, some of the tyrosine residues can become reactive when the protein structure is changed by mutagenesis (i.e. Qβ C74S C80Dha VLP). It would be interesting to exploit these 'side reactions' as new methods for achieving direct chemoselective fluorination on proteins by manipulating the amino acid sequence.

In this chapter, we investigate the direct chemoselective electrophilic aromatic fluorination on tyrosine residues. Based on the amino acid model, both 3-fluoro-tyrosine and the *ipso*-fluorination product were observed.[1] Thus, a maximum of two fluorine atoms can be added to a tyrosine residue. Since the reaction involves the addition of only fluorine atoms, it

can minimise the change on the structure of the protein which is a common problem caused by the prosthetic groups of other fluorinating reagents.

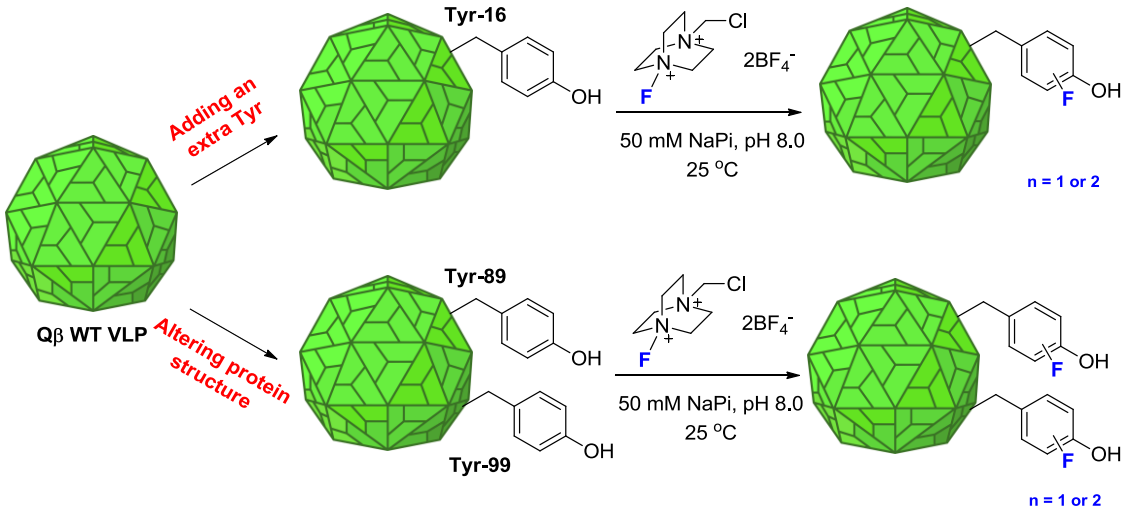


Scheme 3.1: Among the 20 natural amino acids, only cysteine, methionine, tyrosine and tryptophan can undergo reaction with Selectfluor™ at the small molecule level. Chemoselective fluorination on tyrosine residues can be achieved in the protein level by manipulating the amino acid sequence.

3.1.2 Direct labelling of Qβ VLP with Selectfluor™

To introduce C-F bond on Qβ VLPs *via* direct labelling with Selectfluor™, two approaches can be used (**Scheme 3.2**). The first method is to add an extra tyrosine residue to the protein at a highly exposed position. Since Lys-16 is the most exposed position of Qβ in all the three different chains (**Figure 2.2**), the mutant Qβ K16Y was proposed. Alternatively, as has already been shown in the previous chapter, the protein structure can be altered so as to make some of the tyrosine residues accessible for reaction. Qβ C74S C80Dha is a good candidate in which both Tyr-89 and Tyr-99 can be readily fluorinated, as previously demonstrated. It also possesses a Dha tag which can be coupled to other groups before applying Selectfluor™ to the protein, allowing the potential for bio-orthogonal reactions on

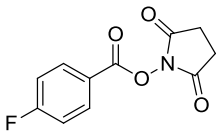
the same protein. On the other hand, the mutant Qβ C74S C80S was also proposed as a candidate for demonstrating pure tyrosine fluorination using this method. The Cys-80 was removed as to prevent any undesired side reactions.



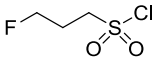
Scheme 3.2: Direct fluorination of tyrosine residue(s) on Qβ VLP can be achieved by either adding an extra tyrosine residue or altering the protein structure *via* mutagenesis.

3.1.3 Direct labelling of Qβ VLP with other fluorinating reagents for comparison

Similar to the 'tag-and-modify' approach, ^{18}F Selectfluor™ can also be applied for tyrosine labelling to render the particle applicable in PET imaging. To compare our novel method with other existing ^{18}F labelling methods, the ^{19}F versions of two commonly used ^{18}F fluorinating reagents were also applied to Qβ VLP (**Figure 3.1**). *N*-**s**uccinimidyl 4-**f**luoro**b**enzoate (SFB) is a commonly used ^{18}F labelling reagent in which the N$^{\varepsilon}$-Lys and the N-terminus of protein can be non-selectively fluorinated.[2] On the other hand, the ^{18}F prosthetic reagent 4-**f**luoro-**p**ropane-1-**s**ulphonyl **c**hloride (FPSC) has also been shown before as a promising choice for labelling amines on peptides.[3] The labelling efficiency and the properties of the labelled particles using different reagents were investigated in order to assess the potential of our proposed method.



N-succinimidyl 4-fluorobenzoate (SFB)



4-fluoro-propane-1-sulphonyl chloride (FPSC)

Figure 3.1: Two commonly used ^{18}F reagent for peptide/protein labelling

3.2 Results and Discussion

3.2.1 Mutagenesis, expression and purification of Qβ VLPs

Both Qβ K16Y and Qβ C74S C80S mutants can be easily obtained *via* site-directed mutagenesis of the plasmid coding the wild type Qβ. The DNA and protein sequences of the mutants are listed below (The mutated sites are labelled with red; the remaining tyrosine residues are labelled with blue):

(A) Qβ K16Y

Forward DNA sequence

ATGGCAAAAT TAGAGACTGT TACTTTAGGT AACATCGGGA AAGATGGA**TA** **T**CAAACTCTG
GTCCTCAATC CGCGTGGGGT AAATCCCACT AACGGCGTTG CCTCGCTTTC ACAAGCGGGT
GCAGTTCCTG CGCTGGAGAA GCGTGTTACC GTTTCGGTAT CTCAGCCTTC TCGCAATCGT
AAGAACTACA AGGTCCAGGT TAAGATCCAG AACCCGACCG CTTGCACTGC AAACGGTTCT
TGTGACCCAT CCGTTACTCG CCAGGCATAT GCTGACGTGA CCTTTTCGTT CACGCAGTAT
AGTACCGATG AGGAACGAGC TTTTGTTCGT ACAGAGCTTG CTGCTCTGCT CGCTAGTCCT
CTGCTGATCG ATGCTATTGA TCAGCTGAAC CCAGCGTATT AA

Amino acid sequence

AKLETVTLGN IGKDG**Y**QTLV LNPRGVNPTN GVASLSQAGA VPALEKRVTV SVSQPSRNRK
N**Y**KVQVKIQN PTACTANGSC DPSVTRQA**Y**A DVTFSFTQ**Y**S TDEERAFVRT ELAALLASPL
LIDAIDQLNP A**Y**

Calculated average isotopic mass = 14158 (N-terminal Met cleaved)

(B) Qβ C74S C80S

Forward DNA sequence

ATGGCAAAAT TAGAGACTGT TACTTTAGGT AACATCGGGA AAGATGGAAA ACAAACTCTG
GTCCTCAATC CGCGTGGGGT AAATCCCACT AACGGCGTTG CCTCGCTTTC ACAAGCGGGT
GCAGTTCCTG CGCTGGAGAA GCGTGTTACC GTTTCGGTAT CTCAGCCTTC TCGCAATCGT
AAGAACTACA AGGTCCAGGT TAAGATCCAG AACCCGACCG CTAGCACTGC AAACGGTTCT
TCGGACCCAT CCGTTACTCG CCAGGCATAT GCTGACGTGA CCTTTTCGTT CACGCAGTAT
AGTACCGATG AGGAACGAGC TTTTGTTCGT ACAGAGCTTG CTGCTCTGCT CGCTAGTCCT
CTGCTGATCG ATGCTATTGA TCAGCTGAAC CCAGCGTATT AA

Amino acid sequence

AKLETVTLGN IGKDGKQTLV LNPRGVNPTN GVASLSQAGA VPALEKRVTV SVSQPSRNRK
NYKVQVKIQN PTASTANGSS DPSVTRQAYA DVTFSFTQYS TDEERAFVRT ELAALLASPL
LIDAIDQLNP AY

Calculated average isotopic mass = 14091 (N-terminal Met cleaved)

The proteins were expressed in *E. coli* BL21(DE3) and purified using the same method as for other Qβ species. Fractions containing the intact particles were combined based on the results from the native agarose and SDS-PAGE gels (**Figure 3.2**) and were subsequently analysed by LC-MS (**Figure 3.3**) and DLS (**Figure 3.4**). The results are summarised in **Table 3.1**. Although both Qβ K16Y and Qβ C74S C80S were expressed successfully, the former protein was obtained only in very low yield in comparison to other mutants. The size of the Qβ K16Y could not be measured properly by DLS due to protein aggregation. It is possible that the addition of an extra tyrosine residue has a detrimental effect on the protein structure. Further experiments on this mutant were thus aborted. On the other hand, the size of Qβ C74S C80S was found to be very similar to that of Qβ C74S C80Dha. It is proposed that these two species have a similar structure and hence comparable chemical properties towards Selectfluor™.

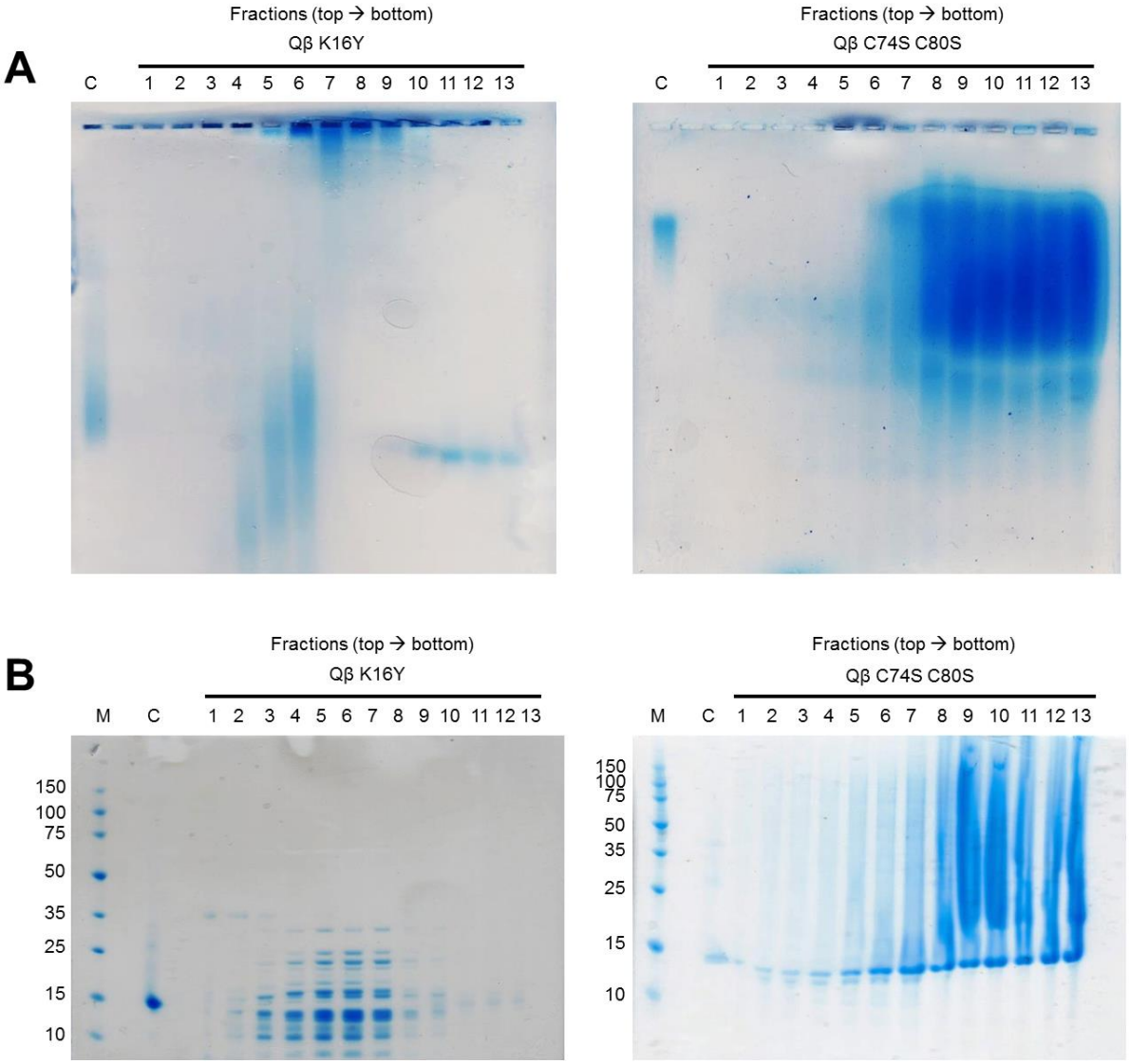


Figure 3.2: Coomassie Blue stained (A) native agarose and; (B) reducing SDS-PAGE gels showing the fractions obtained after sucrose gradient centrifugation of the Qβ K16Y and Qβ C74S C80S mutants. C: Qβ WT control; M: protein marker.

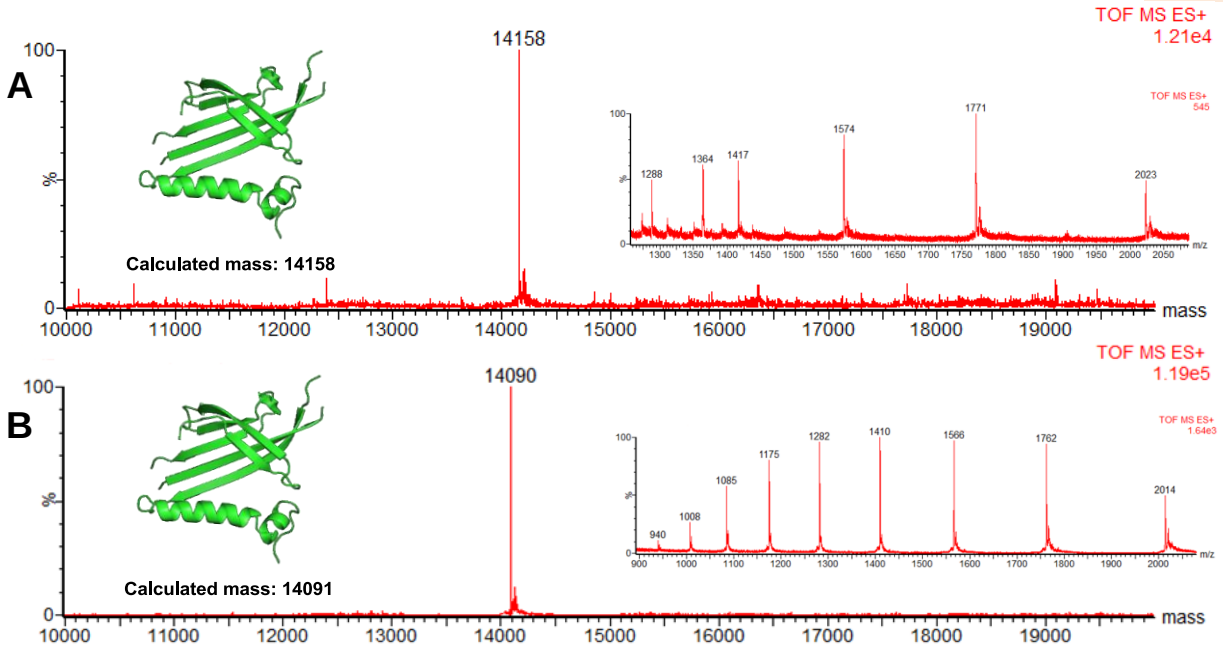


Figure 3.3: Deconvoluted mass spectra and corresponding raw ion series of the purified (A) Qβ K16Y and; (B) Qβ C74S C80S VLPs. The VLPs were broken down to monomers for LC-MS analysis.

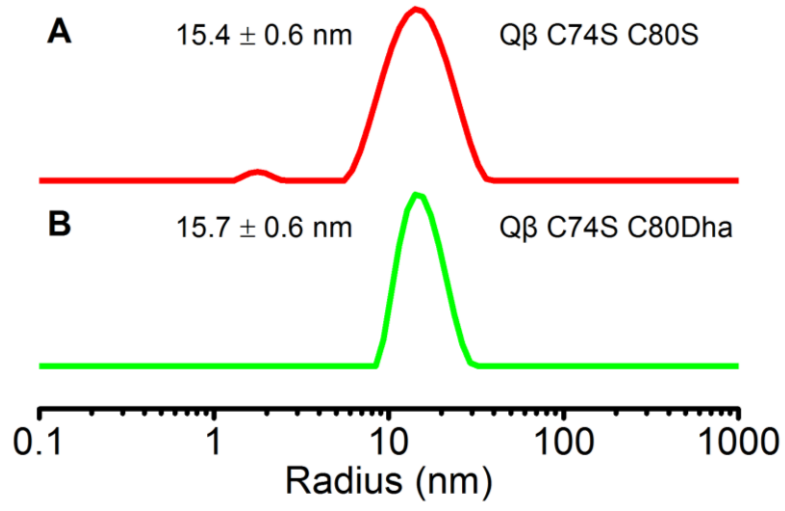


Figure 3.4: DLS results of the purified (A) Qβ C74S C80S and; (B) Qβ C74S C80Dha VLPs.

Table 3.1: The expression yield and the size of Qβ mutants

Qβ Mutant	Yield (mg / L of media)	Radius of particle (nm)
K16Y	1.3	Aggregates
C74S C80S	68.4	15.4 ± 0.6

3.2.2 Fluorination of Qβ VLPs with Selectfluor™

3.2.2.1 Qβ C74S C80S

The VLPs were tested with Selectfluor™ to see if any reaction can occur (**Table 3.2**). As shown by the mass spectra, fluorination (i.e. M + 18, M + 36 and M + 54) was observed (**Figure 3.5**). Similarly to the other Qβ species tested before, the noise of the spectra increases with the amount of Selectfluor™ added, which poses difficulties for spectrum analysis. The addition of 20 equiv. of Selectfluor™ gave approximately 65% conversion based on the consumption of starting materials. This is lower than the analogous conversion obtained from Qβ C74S C80Dha (i.e. 85%) in which Dha is also a reactive species beside the tyrosine residues.

Table 3.2: Optimisation experiments of Qβ C74S C80S with Selectfluor™

Entry	Selectfluor™ BF_4^- (equiv.)	Product conversion (%)
1	1	SM (85) SM + 18 (15)
2	5	SM (75) SM + 18 (25)
3	10	SM (55) SM + 18 (30) SM + 36 (15)
4	20	SM (35) SM + 18 (40) SM + 36 (15) SM + 54 (10)
5	40	*

General conditions: Qβ C74S C80S in NaPi (50 mM, pH 8.0) and Selectfluor™ in H_2O at 25 °C for 30 min unless otherwise indicated; SM: starting material; * the signal-to-noise ratio was so high that the spectrum cannot be analysed properly

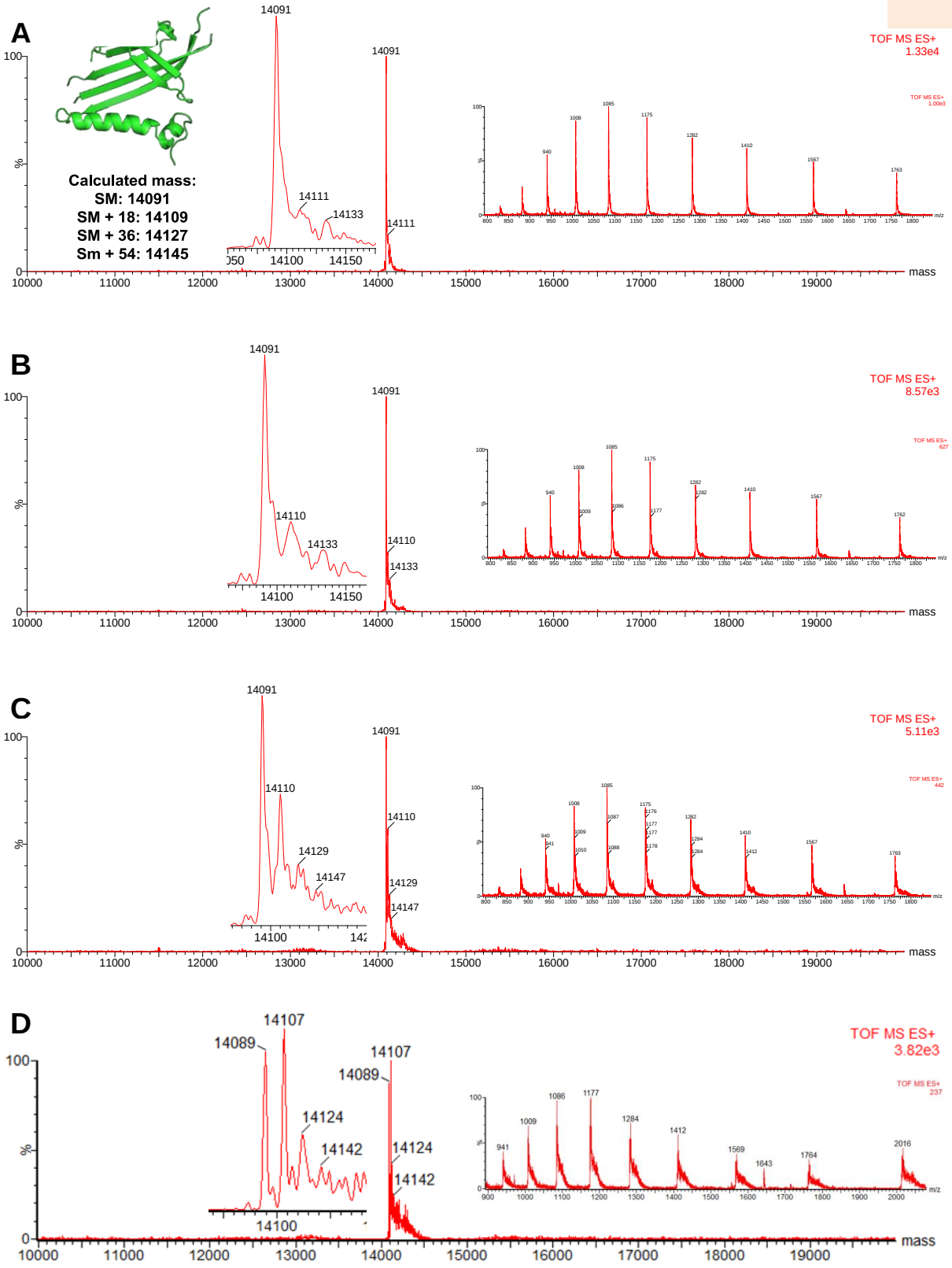
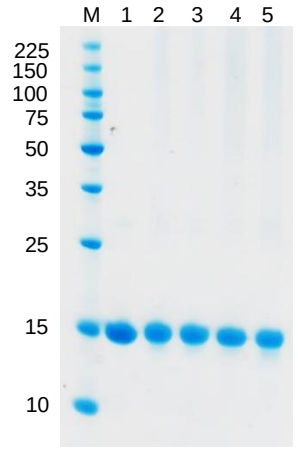


Figure 3.5: Deconvoluted mass spectra and corresponding raw ion series showing the reaction of Qβ C74S C80S VLPs with (A) 1 equiv.; (B) 5 equiv.; (C) 10 equiv. or; (D) 20 equiv. of Selectfluor™ BF_4^-. The VLPs were broken down to monomers for LC-MS analysis.



Lane 1: Qβ C74S C80S
Lane 2: Selectfluor™ (20 equiv.) treated Qβ C74S C80S (-80 °C for 1 month)
Lane 3: Selectfluor™ (20 equiv.) treated Qβ C74S C80S (4 °C for 4 h)
Lane 4: Selectfluor™ (20 equiv.) treated Qβ C74S C80S (25 °C for 2 h)
Lane 5: Selectfluor™ (20 equiv.) treated Qβ C74S C80S (37 °C for 2 h)

Figure 3.6: Coomassie Blue stained reducing SDS-PAGE gel of Qβ C74S C80S and its fluorinated product stored at different temperatures. M: protein marker.

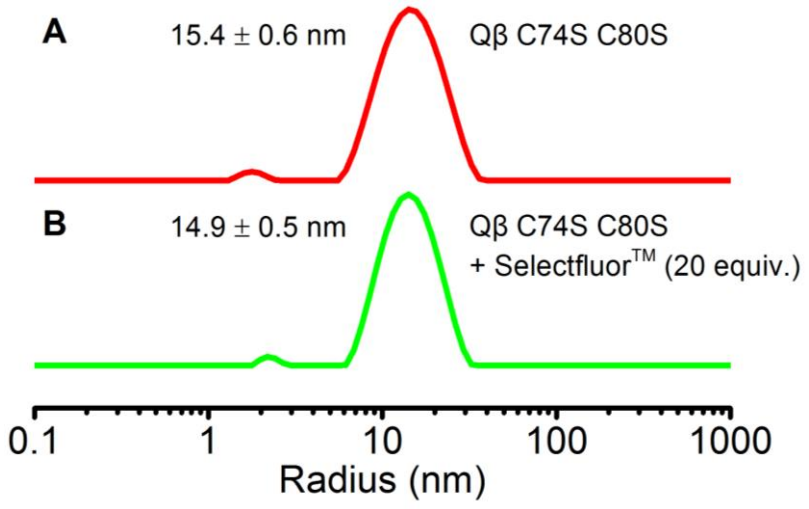


Figure 3.7: DLS results of the purified (A) Qβ C74S C80S and; (B) Selectfluor™ (20 equiv.) treated Qβ C74S C80S VLPs.

As shown by the SDS-PAGE gel (**Figure 3.6**), the protein is quite stable at different temperatures. Interestingly, the size of the particles (**Figure 3.7**) did not increase but instead decreased slightly after fluorination, in contrast to previous observations on both Qβ C74S C80Dha and Qβ T75Dha*. It is possible that positions 75 and 80, which are both located in a loop region, are of greater importance in determining particle shape as compared to the positions where the reactive tyrosine residues are located.

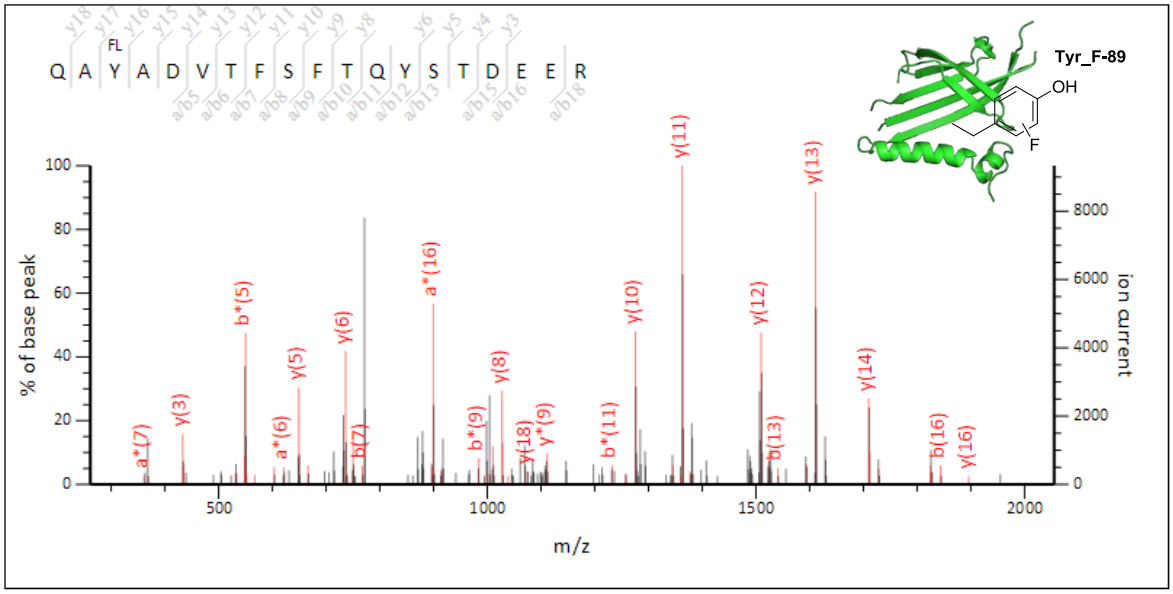
To confirm the fluorination position, the Selectfluor™ treated VLPs were further analysed by MS/MS (**Table 3.3** and **Figure 3.8-3.9**). As expected, both Tyr-89 and Tyr-99 were readily fluorinated, but only Tyr-99 gave doubly fluorinated product. Reaction was not observed on the other two remaining tyrosine residues (i.e. Tyr-62 and Tyr-132). This mutant is thus the first example showing pure C_{sp2}–F bond formation on specific tyrosine residues (i.e. Tyr-89 and Tyr-99) with no other side reactions. These fluorinated tyrosine residues were also shown to be inert under tryptic digestion. Combining with the DLS results mentioned earlier, the labelling method reported here is indeed a very powerful and useful technique.

Table 3.3: Key digested peptide fragments observed by Mascot analysis of fluorinated Qβ C74S C80S (protein sequence coverage: 100%)

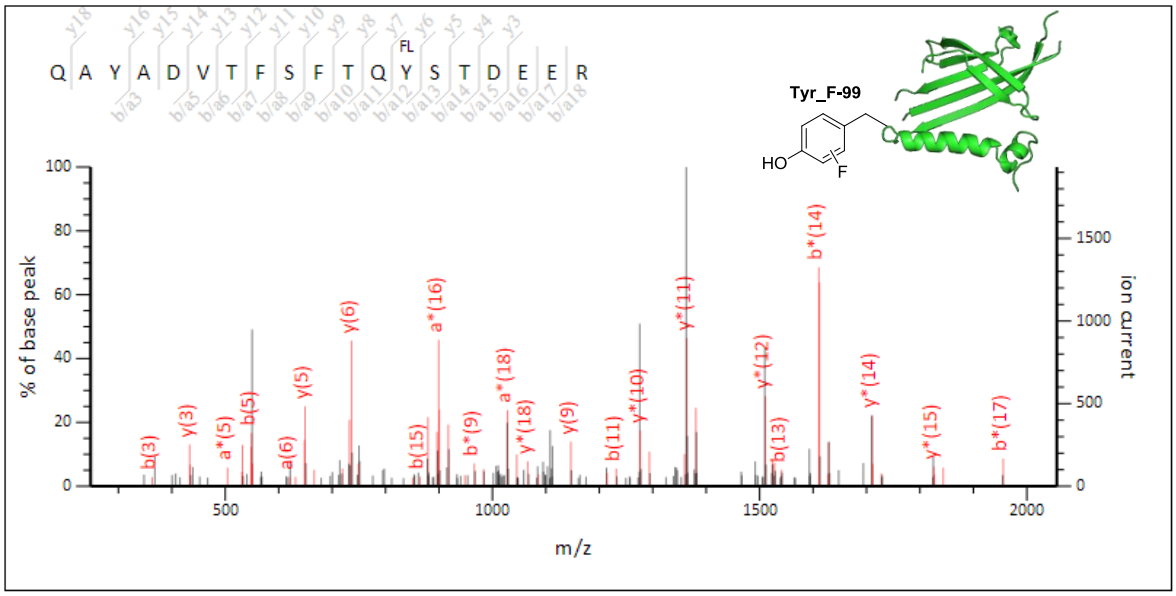
Score	Mr(found)	Mr(Calc)	Error (ppm)	Expect	Modifications†	Peptide sequence
69.0	2274.9715	2274.9717	-0.078	1.3e-07	Tyr89→Tyr_F89	87-QA**Y**ADVTFSFTQYSTDEER-105
57.4	2274.9732	2274.9717	0.67	1.8e-06	Tyr99→Tyr_F99	87-QAYADVTFSFTQ**Y**STDEER-105
70.2	2766.2378	2766.2373	0.17	9.5e-08	Tyr99→Tyr_FF99	87-QAYADVTFSFTQ**Y**STDEERAFVR-109
69.7	2292.9613	2292.9623	-0.44	1.1e-07	Tyr89→Tyr_F89 Tyr99→Tyr_F99	87-QA**Y**ADVTFSFTQ**Y**STDEER-105
85.6	2310.9564	2310.9529	1.53	2.7e-09	Tyr89→Tyr_F89 Tyr99→Tyr_FF99	87-QA**Y**ADVTFSFTQ**Y**STDEER-105

† Tyr_F: monofluorination on tyrosine residue; Tyr_FF: difluorination on tyrosine residue

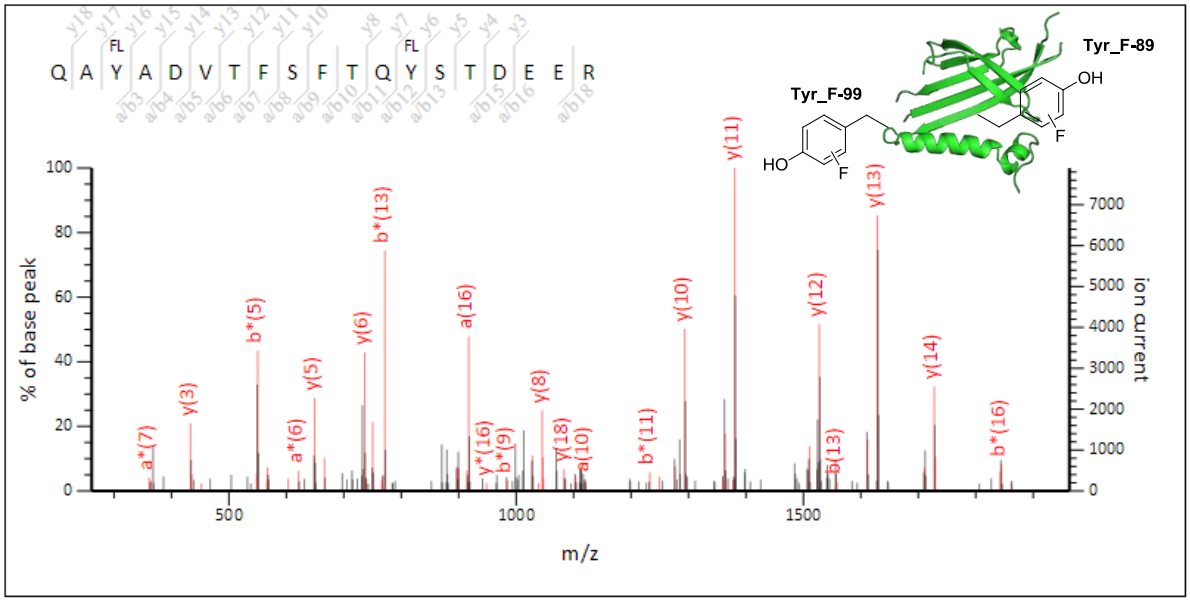
A


FL
Q A Y A D V T F S F T Q Y S T D E E R
Tyr_F-89
OH
F
% of base peak
ion current
m/z
100
80
60
40
20
0
8000
6000
4000
2000
500
1000
1500
2000
a*(7)
y(3)
b*(5)
a*(6)
y(5)
y(6)
b(7)
a*(16)
b*(9)
y(8)
y(18)
y*(9)
b*(11)
y(10)
y(11)
y(12)
b(13)
y(13)
y(14)
b(16)
y(16)

B


FL
Q A Y A D V T F S F T Q Y S T D E E R
Tyr_F-99
HO
F
% of base peak
ion current
m/z
100
80
60
40
20
0
1500
1000
500
0
500
1000
1500
2000
b(3)
y(3)
a*(5)
b(5)
a(6)
y(5)
y(6)
b(15)
a*(16)
b*(9)
a*(18)
y*(18)
y(9)
b(11)
y*(10)
y*(11)
y*(12)
b(13)
b*(14)
y*(14)
y*(15)
b*(17)

C



D

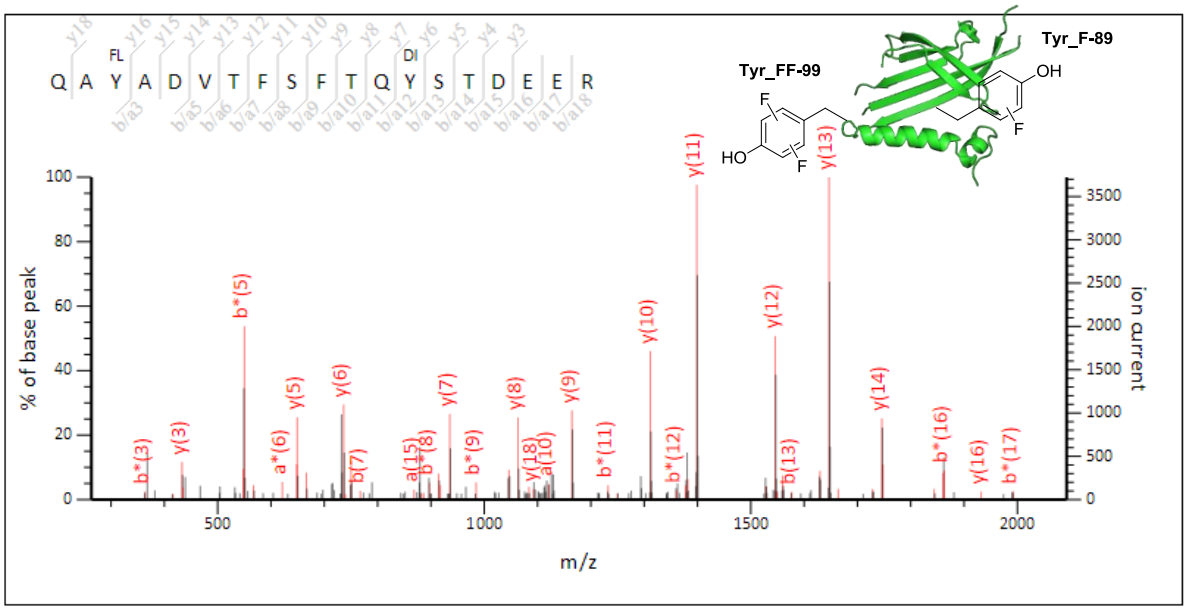


Figure 3.8: ESI-QTOF MS-MS spectra of peptide 87-QAYADVTFSFTQYSTDEER-105 obtained by tryptic digestion of Selectfluor™ (20 equiv.) treated Qβ C74S C80S that contains (A) single fluorination on Tyr-89 (Score = 69.0); (B) single fluorination on Tyr-99 (Score = 57.4); (C) single fluorination on both Tyr-89 and Tyr-99 (Score = 69.7) or; (D) single fluorination on Tyr-89 and double fluorination on Tyr-99 (Score = 85.6).

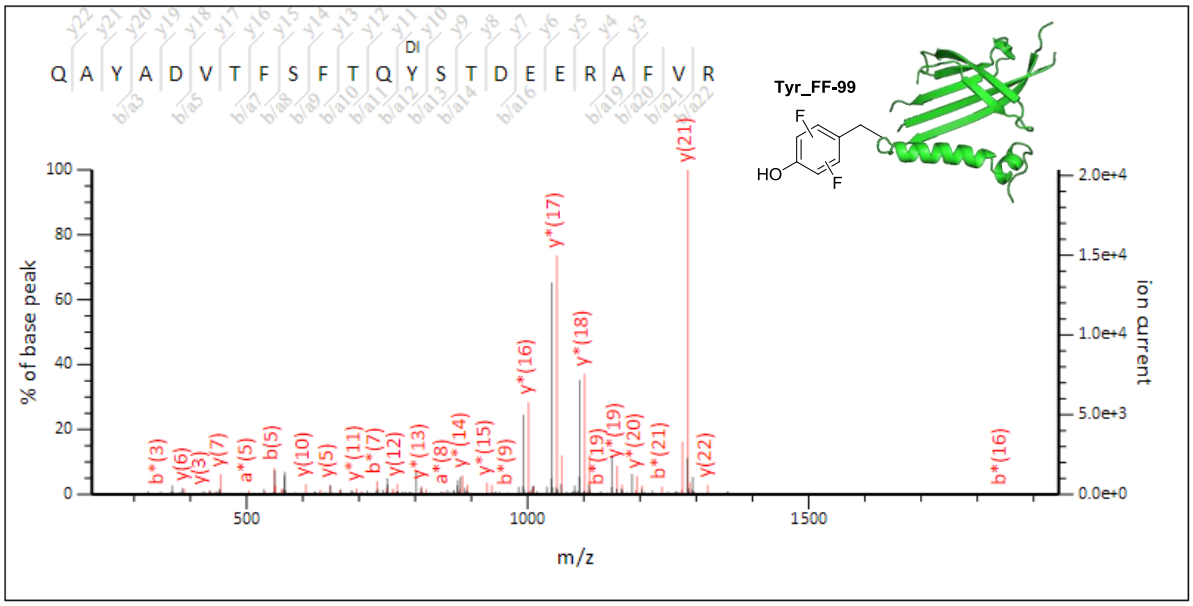


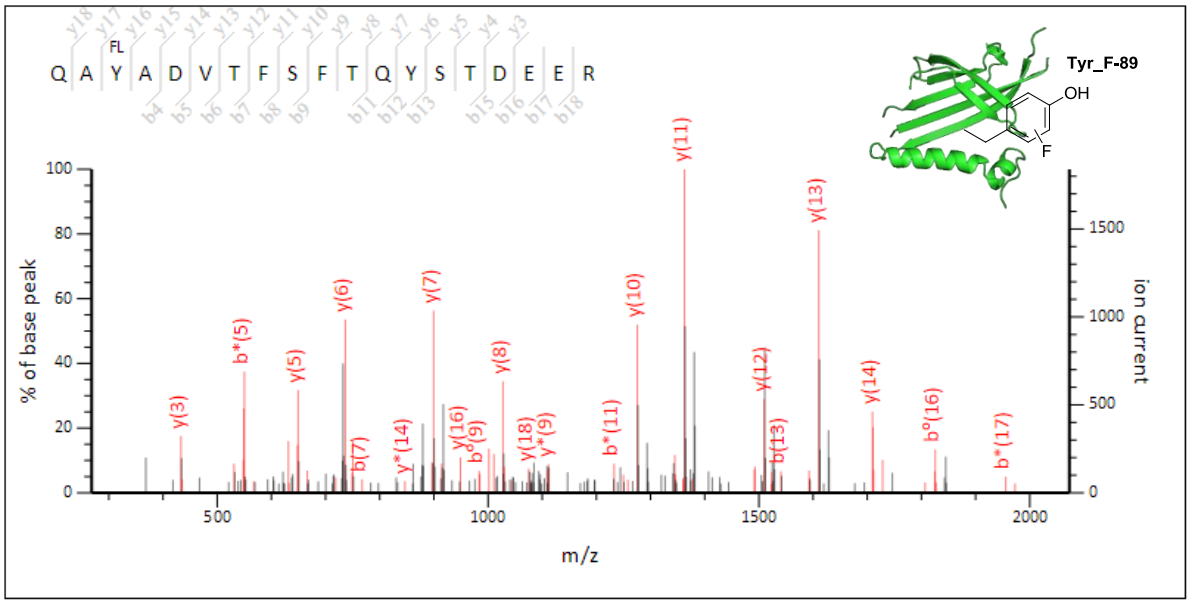
Figure 3.9: ESI-QTOF MS-MS spectrum of peptide 87-QAYADVTFSFTQYSTDEERAFVR-109 obtained by tryptic digestion of Selectfluor™ (20 equiv.) treated Qβ C74S C80S that contains double fluorination on Tyr-99 (Score = 70.2).

3.2.2.2 Qβ C74S C80Dha

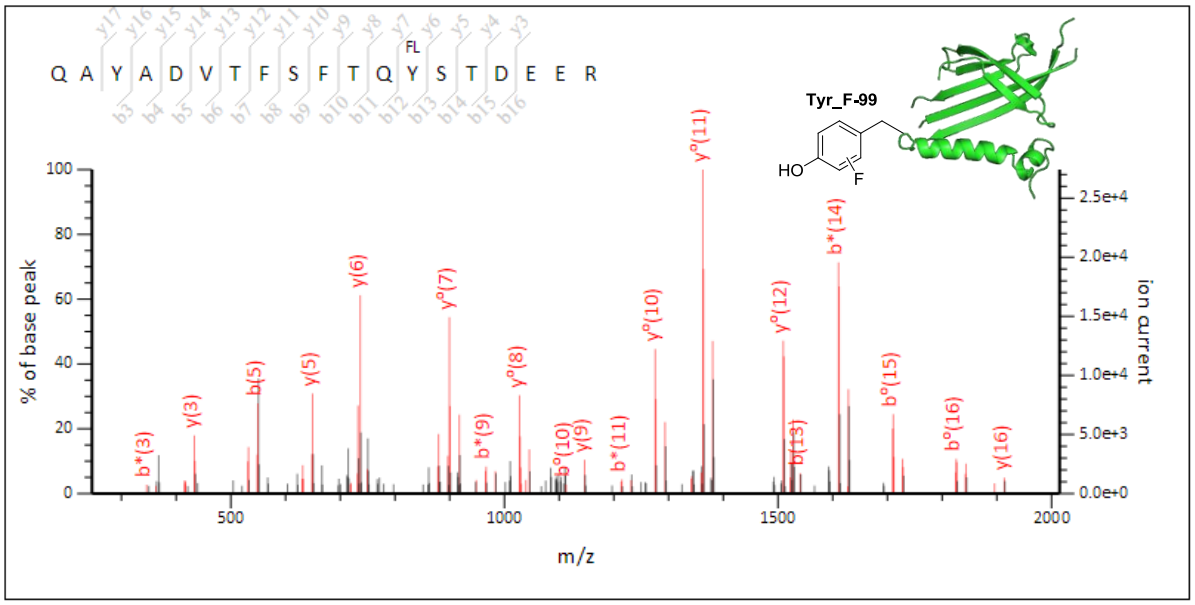
As mentioned in the previous chapter, side reactions were observed on both Tyr-89 and Tyr-99 when Qβ C74S C80Dha was treated with Selectfluor™ (**Table 2.4**). The MS/MS data were further analysed to identify the level of fluorination (**Figure 3.10**). Again, double fluorination was only observed on Tyr-99. However, it is difficult to compare the level of fluorination on Dha and that on tyrosine residues based on the protein mass spectra and MS/MS results as the peptides ionise differently.

From the point of view of the 'tag-and-modify' approach, fluorination on tyrosine residues is perhaps considered to be an undesired side reaction. Nonetheless, the existence of Dha on the protein is indeed a bonus for the direct labelling approach. The Dha provides potential for the coupling of other thiol-containing groups to the protein where the tyrosine residues can be fluorinated on the same protein. This kind of orthogonal dual-labelling strategy could be very useful for studying protein functions or mechanisms.[4-6]

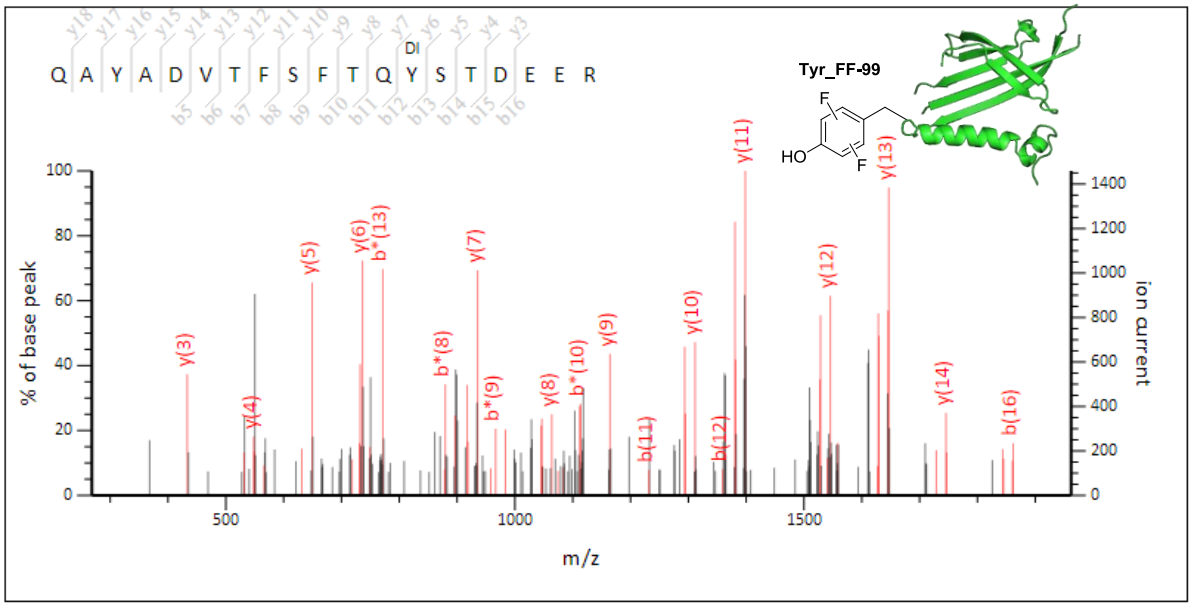
A


Q A Y A D V T F S F T Q Y S T D E E R
FL
Tyr_F-89
OH
F
% of base peak
ion current
m/z
100
80
60
40
20
0
500
1000
1500
2000
1500
1000
500
0

B


Q A Y A D V T F S F T Q Y S T D E E R
FL
Tyr_F-99
HO
F
% of base peak
ion current
m/z
100
80
60
40
20
0
500
1000
1500
2000
2.5e+4
2.0e+4
1.5e+4
1.0e+4
5.0e+3
0.0e+0

C



D

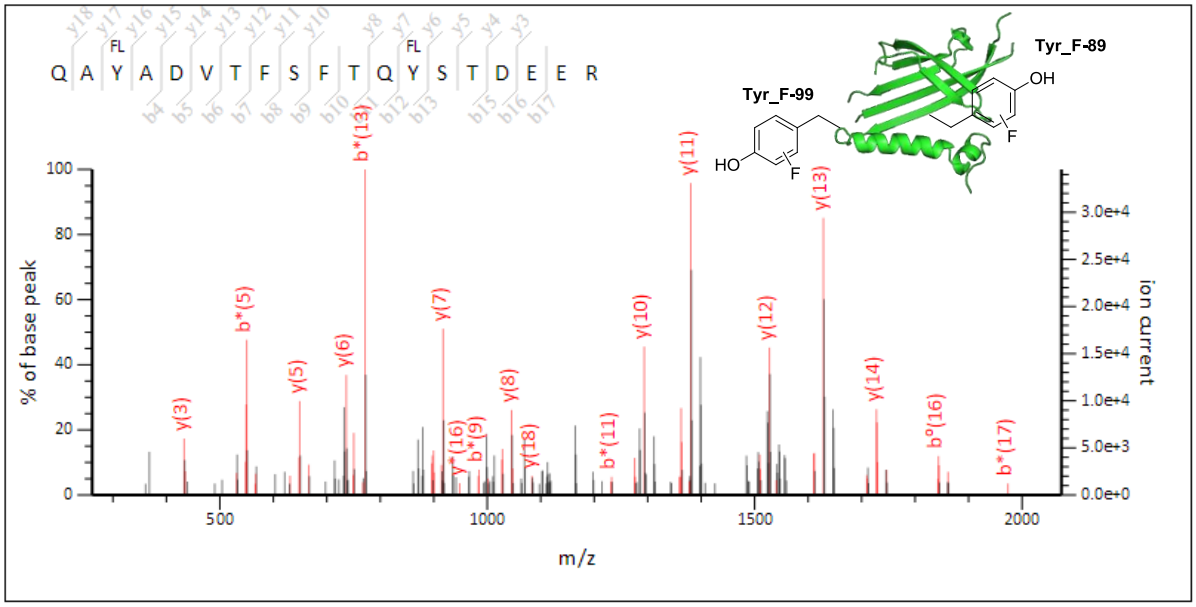
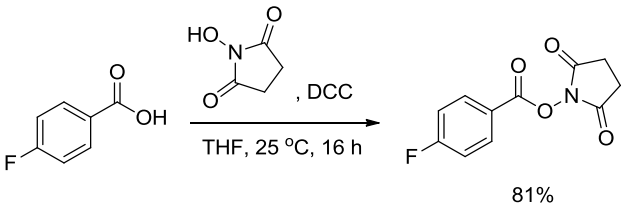


Figure 3.10: ESI-QTOF MS-MS spectra of peptide 87-QAYADVTFSFTQYSTDEER-105 obtained by tryptic digestion of Selectfluor™ (20 equiv.) treated Qβ C74S C80Dha that contains (A) single fluorination on Tyr-89 (Score = 71.6); (B) single fluorination on Tyr-99 (Score = 90.3); (C) double fluorination on Tyr-99 (Score = 86.8) or; (D) single fluorination on both Tyr-89 and Tyr-99 (Score = 66.0).

3.2.3 Direct labelling of Qβ VLPs with other common ^{18}F reagents

Qβ VLPs were also treated with other commonly used ^{18}F reagents (i.e. SFB and FPSC) for comparison. In here, the ^{19}F versions of the reagents were used instead for safety reasons. As mentioned earlier in this chapter, both reagents react with the amine groups on the protein. Qβ WT VLP was thus chosen as a model for the experiments as each monomer contains eight reactive amine groups including the N-terminal amine.



Scheme 3.3: Synthesis of SFB.

SFB can be synthesised easily from commercially available 4-fluorobenzoic acid and *N*-hydroxysuccinimide in good yield (**Scheme 3.3**). Since it is poorly soluble in water, the white SFB powder was added directly to the protein solution. As shown by the mass spectrum, up to five amine groups were labelled when approximately 100 equiv. of SFB was added (**Figure 3.11**). No starting material was observed, indicating that at least one amine group was labelled in each monomer. The mass spectrum does not have much noise when compared to those obtained from Selectfluor™ treated sample, indicating that SFB can be removed efficiently with columns. Nonetheless, when the sample was analysed by DLS, a significant amount of aggregates (22.4% area) were observed apart from the fluorinated particles (77.6% area) (**Figure 3.12**). It is possible that the prosthetic group of the reagent interrupted the ionic interactions of the labelled lysine residues with other neighbouring residues and hence the particle formation. The reagent may be useful for some simple proteins, but it may not be applicable to large protein particles such as Qβ VLPs in which

both the protein and particle structural integrity should be maintained. This provides a strong motive for using the approach reported in this chapter in which only a single fluorine atom is added.

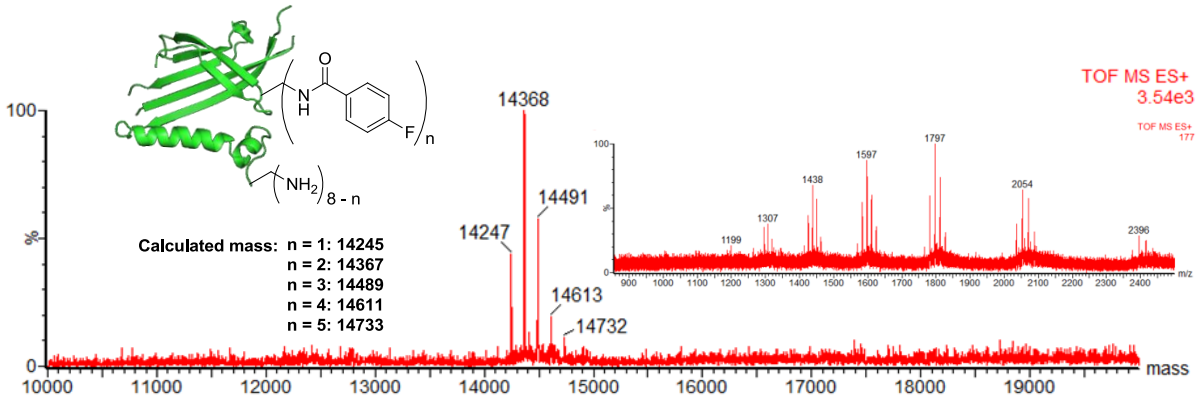


Figure 3.11: Deconvoluted mass spectrum and corresponding raw ion series showing the reaction of Qβ WT VLP with SFB (100 equiv.) in NaPi (50 mM, pH 8.0) at 25 °C.

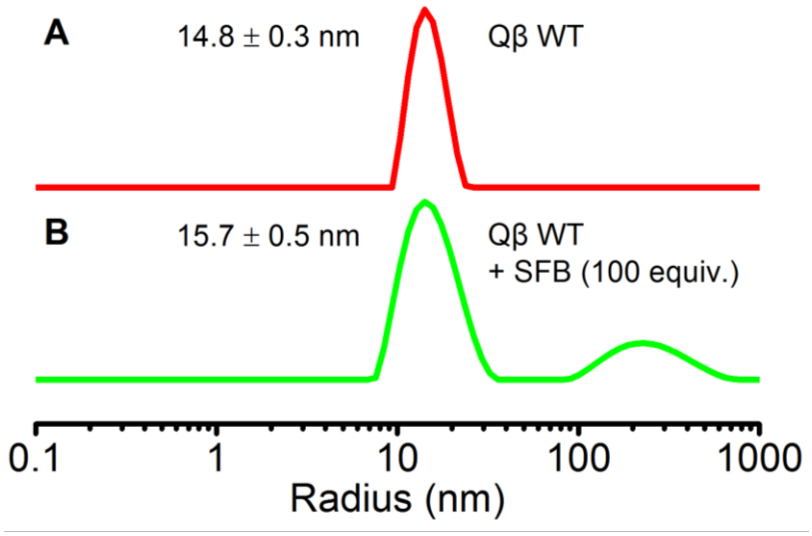


Figure 3.12: DLS results of the purified (A) Qβ WT and; (B) SFB (100 equiv.) treated Qβ WT VLPs.

On the other hand, Qβ WT VLPs were also treated with FPSC, another commonly used ^{18}F reagent for which the ^{19}F version is commercially available. Qβ WT VLP was found to be inert towards this reagent even when 3000 equiv. of it was added (**Figure 3.13**). Although a

tiny peak with a mass matching the singly coupled product was observed, its signal was very similar to the background noise. In addition, some unidentified side products were observed, further complicating the spectrum. This again served to emphasise that a good fluorinating reagent should be both efficient and at the same time causing a minimum effect on the protein (or particle) structure, both of which can be accomplished by the Selectfluor™ method.

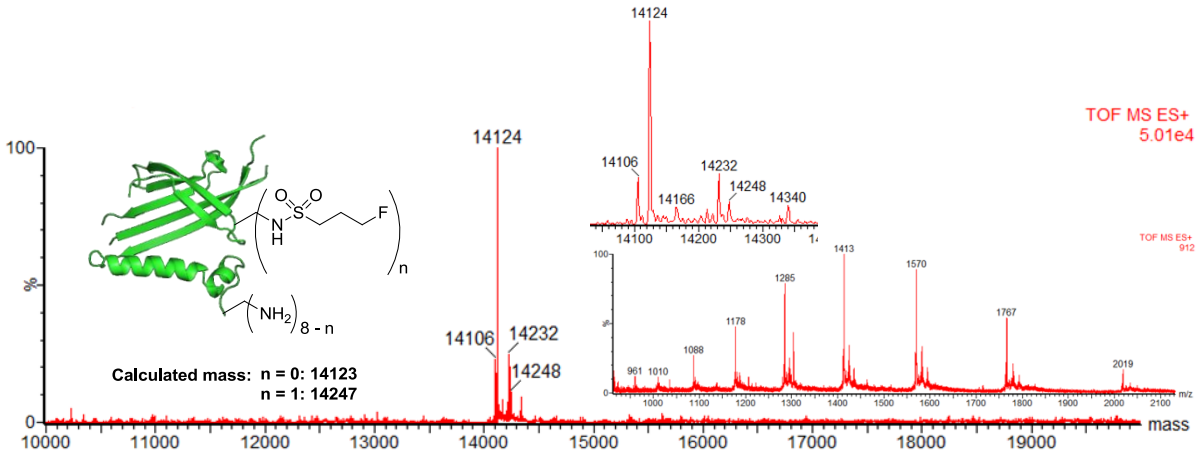


Figure 3.13: Deconvoluted mass spectrum and corresponding raw ion series showing the reaction of Qβ WT VLP with FPSC (3000 equiv.) in NaPi (50 mM, pH 8.0) at 25 °C.

3.3 Conclusion

Two Qβ species were reported in this chapter for which specific tyrosine residues can be selectively labelled with Selectfluor™. This provides the first examples for Qβ VLPs of C_{sp2}-F bond formation *via* direct chemoselective electrophilic aromatic fluorination. Qβ C74S C80S VLPs showed that Tyr-89 and Tyr-99 can be selectively labelled, with no other side reactions being observed. Qβ C74S C80Dha VLPs, on the other hand, opened up the potential for combining with other coupling strategies to establish an orthogonal way for protein labelling. By manipulating the amino acid sequence or the structure of the protein, it is believed that other tyrosine residues could also be selectively labelled.

The methodology disclosed here provides an easy platform to chemoselectively label proteins carrying canonical amino acids with only F atoms in a very straightforward manner that complements current strategies using more sterically demanding prosthetic groups. This advantage was clearly demonstrated by comparing Qβ VLPs treated with Selectfluor™ and other commonly used fluorinating reagents. Furthermore, it is very likely that the particles can be labelled by radioactive ^{18}F atom in the same way by using ^{18}F Selectfluor™. This can perhaps provide an alternative for the biomedical community including those lacking the expertise and resources required for other fluorination methods similarly to the widely used, commercially available radio-iodination kits for protein tracking.

There is one question that remains unanswered after demonstrating both possible labelling approaches of Selectfluor™. For Qβ C74S C80Dha VLP, although it has been shown that both Dha and tyrosine residues can be labelled, it is difficult to tell which residue is more reactive towards fluorination (**Figure 3.14**). When Dha-80 was changed to Ser-80, the tyrosine residues still react readily with Selectfluor™ and the level of fluorination was

significant (> 65%). On the other hand, when the tyrosine residues were removed, the protein surprisingly became inert towards Selectfluor™. Pure Dha fluorination could only be achieved when the position of Dha had changed to position 75, which is believed to be more accessible. However, the level of fluorination was still low (~15%). From the above results, it can perhaps be deduced that the tyrosine residues on Qβ C74S C80Dha VLP are more reactive than Dha. However, the structural change of the protein after each round of mutagenesis should also be considered. These changes can be clearly observed from the particle sizes measured by DLS. It is therefore tricky to compare different Qβ mutants as the accessibility of amino acids is not the same. More advanced technologies such as X-ray protein crystallography may be required to answer this question, which is beyond the scope of this thesis.

To conclude, both 'tag-and-modify' and direct labelling approaches have their own advantages and disadvantages. It is believed that other VLPs can also be labelled in the same way, which could lead to further application in VLP-based vaccine or drug delivery.

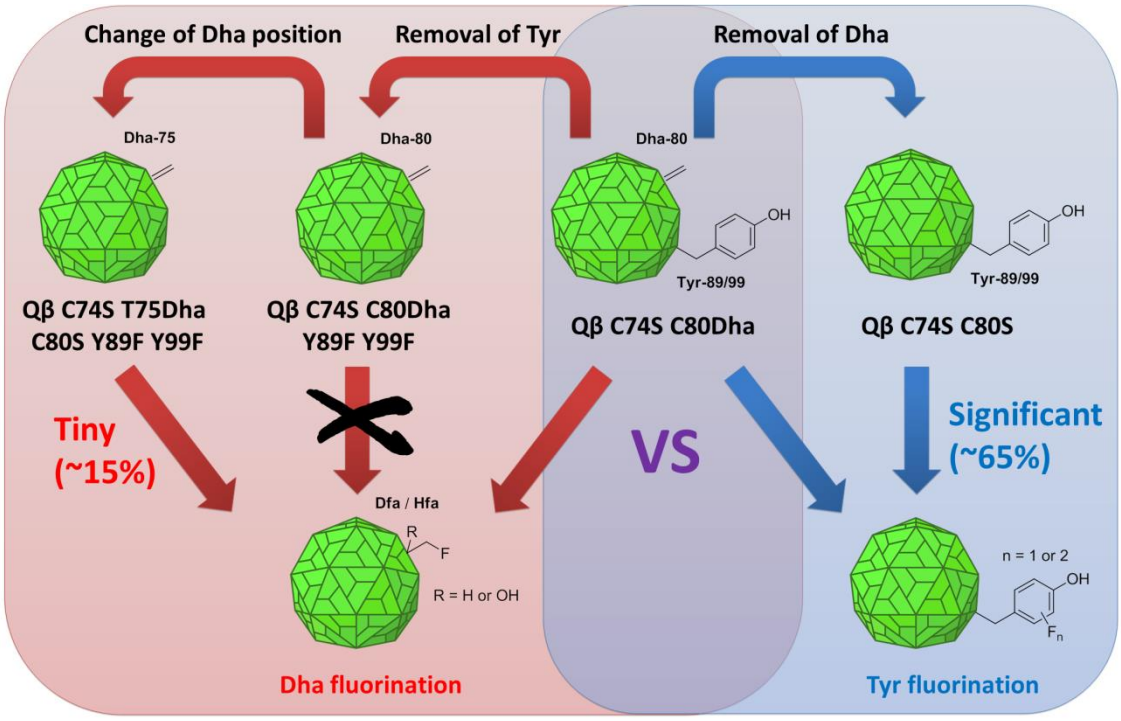


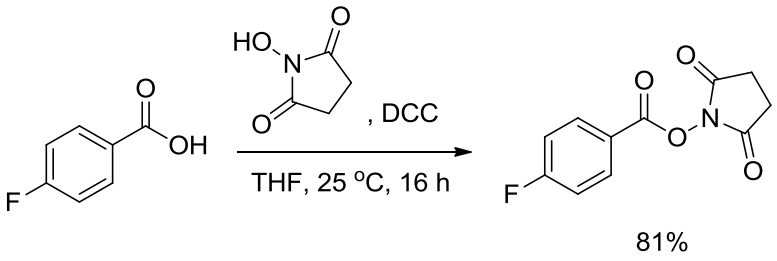
Figure 3.14: Fluorination of different Qβ species with Selectfluor™.

3.4 Experimental Section

3.4.1 General procedures

Refer to Chapter 2.

3.4.2 Synthesis of *N*-succinimidyl 4-fluorobenzoate (SFB)



81%

4-fluorobenzoic acid (500 mg, 3.7 mmol) and *N*-hydroxysuccinimide (500 mg, 4.3 mmol) were added to a 50 mL RB flask, dried under vacuum and filled with argon. Anhydrous THF (10 mL) was added to the flask, followed by the addition of DCC (760 mg, 3.7 mmol). The reaction mixture was stirred at 25 ^{o}C for 16 h. The precipitates formed were filtered, and the filtrate was evaporated under vacuum. The residue was purified by silica column chromatography using CH_2Cl_2 as the eluent. The fractions containing the product were combined and evaporated under vacuum to give SFB as a white powder (708 mg, 81%); R_f = 0.81 (0.5:9.5 MeOH/CH_2Cl_2, v/v); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H = 2.91 (s, 4H, NHS ring), 7.19 (m, 2H, protons ortho to F), 8.16 (m, 2H, protons meta to F); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C = 25.78 (NHS C-H); 116.33 (C-Ar); 116.55 (C-Ar); 121.50 (C-Ar); 133.47 (C-Ar); 133.57 (C-Ar); 161.05 (C=O); 166.99 (d, J 257.74, C-F), 169.33 (NHS C=O); LRMS m/z (ESI$^+$): measured: 260.0 $(M + Na)^+$, calculated: 260.0, measured: 276.0 $(M + K)^+$, calculated: 276.0; IR: 1777, 1725, 1597, 1503, 1420, 1374, 1248, 1206, 1152, 1078, 1067, 1005, 989, 851, 816 cm^{-1}.

3.4.3 Mutagenesis of Qβ

3.4.3.1 Primers used

Mutant	Template		Primer sequence (5' → 3')
K16Y	WT	F	GTAACATCGGGAAAGATGGATATCAAACTCTGGTCCTCAATCCG
		R	CGGATTGAGGACCAGAGTTTGAGATCCATCTTTCCCGATGTTAC
C74S	WT	F	CCGACCGCTAGCACTGCAAAC
		R	GTTTGCAGTGCTAGCGGTCGG
C74S C80S	C74S	F	AAACGGTTCTTCGGACCCATCCG
		R	CGGATGGGTCCGAAGAACCGTTT

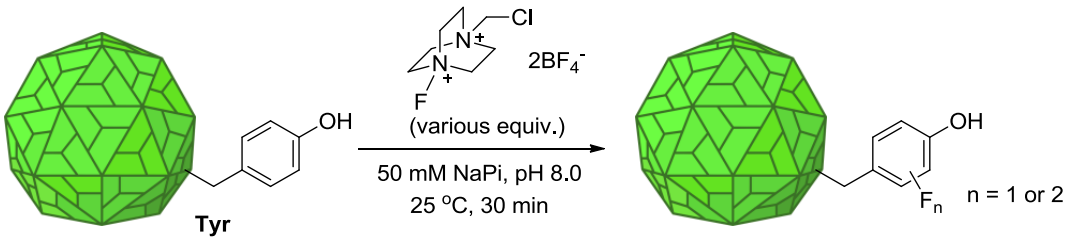
3.4.3.2 Protocols

Refer to Chapter 2.

3.4.4 Expression, purification and characterisation of Qβ VLPs

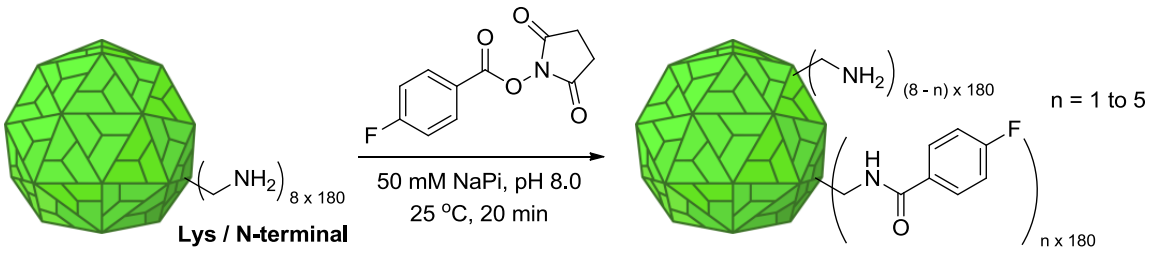
Refer to Chapter 2.

3.4.5 Fluorination of Qβ VLPs with Selectfluor™



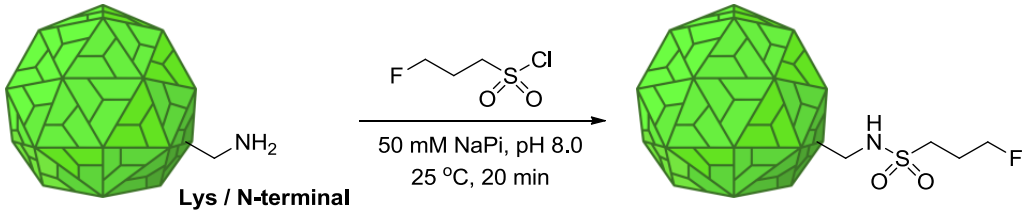
Qβ VLPs (1 mg / mL) in TBS were buffer exchanged using PD SpinTrap™ G-25 / MiniTrap™ G-25 / MidiTrap™ G-25 / PD10 desalting column into sodium phosphate buffer (50 mM, pH 8.0) depending on the scale. The protein solution was then mixed with freshly prepared Selectfluor™ BF_4^- (5 mg/mL in H_2O) (various equiv.) and shaken at 25 °C for 30 min. The excess Selectfluor™ was removed by PD SpinTrap™ G-25 / MiniTrap™ G-25 / MidiTrap™ G-25 / PD10 desalting column with sodium phosphate buffer (50 mM, pH 8.0) depending on the scale.

3.4.6 Fluorination of Qβ WT VLPs with SFB



Qβ WT VLPs in TBS were buffer exchanged using PD SpinTrap™ G-25 into sodium phosphate buffer (50 mM, pH 8.0). SFB powder (100 equiv.) was added directly to the protein solution. The reaction mixture was shaken at 25 °C for 20 min, and then purified by Micro Bio-Spin™ P6 column using sodium phosphate buffer (50 mM, pH 8.0) as the elution buffer.

3.4.7 Fluorination of Qβ WT VLPs with FPSC



Qβ WT VLPs in TBS were buffer exchanged using PD SpinTrap™ G-25 into sodium phosphate buffer (50 mM, pH 8.0). Liquid FPSC (3000 equiv.) was added directly to the protein solution. The reaction mixture was shaken at 25 °C for 20 min, and then purified by Micro Bio-Spin™ P6 column using sodium phosphate buffer (50 mM, pH 8.0) as the elution buffer.

3.4.8 MS/MS analysis (Target Discovery Institute, University of Oxford)

Qβ VLPs in sodium phosphate buffer (50 mM, pH 8.0) were mixed with 2 M urea and incubated at 25 ^{o}C for 2 h. Trypsin (0.2 mg/mL) (1:50 w/w) was added to the mixture and shaken at 37 ^{o}C for overnight. 0.4% of formic acid was then added to the sample, followed by purification using ZipTip®. The purified sample was submitted to Target Discovery Institute (Dr. Benedikt M. Kessler) for MS/MS analysis. The raw data were analysed by Mascot using the following parameters: enzyme: trypsin; mass values: monoisotopic; protein mass: unrestricted; peptide charge: +2, +3, +4; peptide mass tolerance: ± 50 ppm; fragment mass tolerance: ± 0.6 Da; mas missed cleavages: 2; instrument type: ESI-QUAD-TOF.

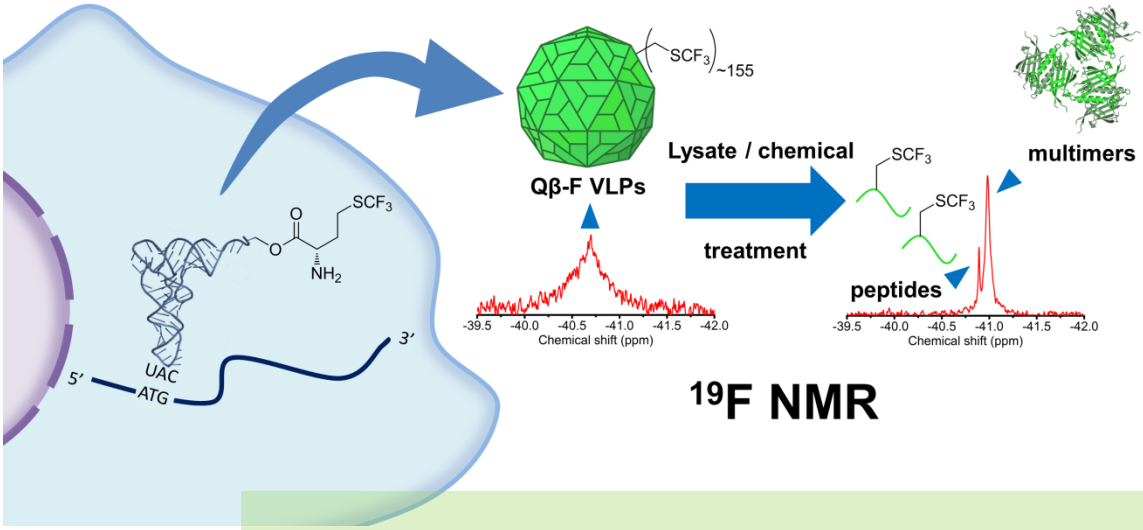
3.5 References

1. Boutureira, O, Schombs, M., et al. (Unpublished results)
2. Jacobson, O., Kiesewetter, D.O. & Chen, X. Fluorine-18 Radiochemistry, Labeling Strategies and Synthetic Routes. *Bioconjugate Chemistry* **26**, 1-18 (2015).
3. Löser, R. et al. Use of 3-[(18)F]fluoropropanesulfonyl chloride as a prosthetic agent for the radiolabelling of amines: Investigation of precursor molecules, labelling conditions and enzymatic stability of the corresponding sulfonamides. *Beilstein journal of organic chemistry* **9**, 1002-1011 (2013).
4. Kanerva, A. & Hemminki, A. Modified adenoviruses for cancer gene therapy. *Int J Cancer* **110**, 475-80 (2004).
5. Eulitt, P.J. et al. Enhancing mda-7/IL-24 therapy in renal carcinoma cells by inhibiting multiple protective signaling pathways using sorafenib and by Ad.5/3 gene delivery. *Cancer Biol Ther* **10**, 1290-305 (2010).
6. Dobbs, A.P. & Kimberley, M.R. Fluorous phase chemistry: a new industrial technology. *Journal of Fluorine Chemistry* **118**, 3-17 (2002).

PART

II

Monitoring the disassembly of virus-like particles by ^{19}F NMR



Chapter 4

Establishment of a 19F detection system for particle disassembly

Chapter 5

Towards the application in biological context

CHAPTER

4 Establishment of a ^{19}F detection system for particle disassembly

4.1 Introduction

4.1.1 Monitoring the disassembly of fluorinated Qβ VLP by ^{19}F NMR spectroscopy

In the previous chapters, Qβ VLPs were labelled with Selectfluor™ by either a 'tag-and-modify' or tyrosine labelling approach. As ^{18}F Selectfluor™ is also available, it is expected that the methods developed can greatly enhance the application of PET imaging to the particles. Together with other methods such as fluorescence microscopy[1, 2] and transmission electron microscopy (TEM),[3, 4] all these techniques provide robust data on the position of particles relative to target cells and organs. However, in order to harness these particles as drug carriers, it is vital to detect not only their location, but also the state of assembly, to establish whether the VLPs have actually dropped their cargo. It would thus be useful to develop a convenient way for monitoring the multimeric state of VLPs to aid their design towards such functional goals. In this chapter, a novel methodology involving ^{19}F labelling of Qβ VLPs is reported.

As explained in Chapter 1, ^{19}F-protein labelling is invaluable since ^{19}F generally has a very low background when biological samples are analysed. Most importantly, ^{19}F is NMR-active with a wide chemical shift range, making it sensitive to the local environment.[5] The high sensitivity of ^{19}F renders it easy to detect, which is useful for monitoring both structural and interaction changes. Hence, 'background-free' virus tracking *in vivo* can be

envisaged using ^{19}F NMR spectroscopy without obscurity from the complex mixture of biomacromolecules in the cell.

4.1.2 Biosynthetic fluorine labelling of Qβ VLP

In order to render Qβ VLP a sensitive imaging agent, the particle should have high fluorine content and all the ^{19}F atoms should possess the same chemical shift. In other words, site-specific labelling with magnetic equivalent fluorine atoms is required. Although site-specific labelling can be achieved using the 'tag-and-modify' approach mentioned in Chapter 2, the number of fluorine atoms incorporated per particle was not sufficient for detection by NMR spectroscopy or MRI. We envisaged that labelling VLPs with the ^{19}F-containing unnatural amino acid trifluoromethionine (Tfm), which contains three equivalent fluorine atoms, should provide sufficiently strong signals for interrogation by NMR spectroscopy (**Figure 4.1**).

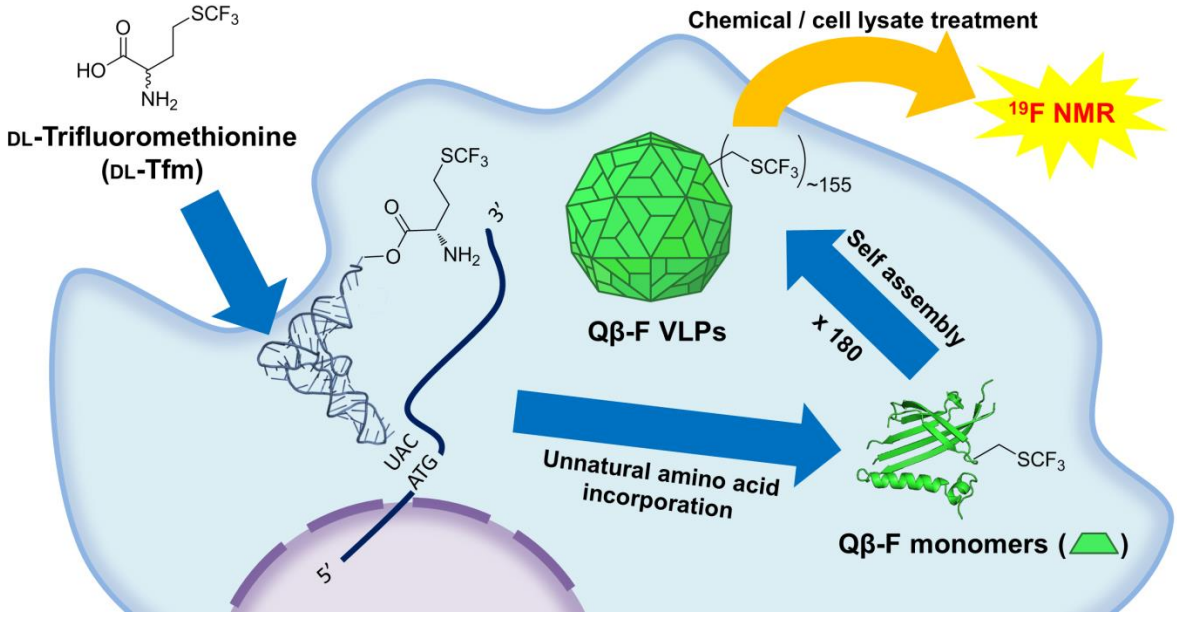


Figure 4.1: Strategy for genetic incorporation of Tfm to Qβ K16M VLP in *E. coli* auxotroph B834(DE3). The purified VLPs were analysed by ^{19}F NMR spectroscopy after chemical or biological treatment.

One approach for introducing Tfm to proteins involves the commandeering of a 'sense' codon for methionine and subsequent reassignment to incorporate it as a methionine analogue, with the help of bacteria strains which are auxotrophic to methionine. The use of *E. coli* auxotroph B834(DE3) has been reported in the past for incorporating Tfm into different proteins such as phage lysozyme LaL,[6] dihydrofolate reductase,[7] DNA polymerase,[8] etc. Since the first methionine residue of Qβ protein is cleaved during expression, it has no naturally occurring methionine residues. By converting the surface exposed Lys-16 to methionine (i.e. Qβ K16M), it was predicted that this can provide both a convenient and accessible probe site (**Figure 2.2**).[9]

4.1.3 Disassembly of Qβ VLP

To study the disassembly of Qβ VLP, it is important to first understand how the particle is formed. As mentioned in Chapter 1, the particle is self-assembled *via* strong protein-protein interactions,[10, 11] and is further stabilised by the presence of inter-subunit disulphide bonds.[12] When the particle is treated with denaturant such as SDS, it will break down into twelve pentamers and twenty hexamers, in which the monomers are held only by intermolecular disulphide linkages. By adding reductants such as DTT or 2-mercaptoethanol, the multimers can further dissociate into monomers (**Figure 4.2**).[13]

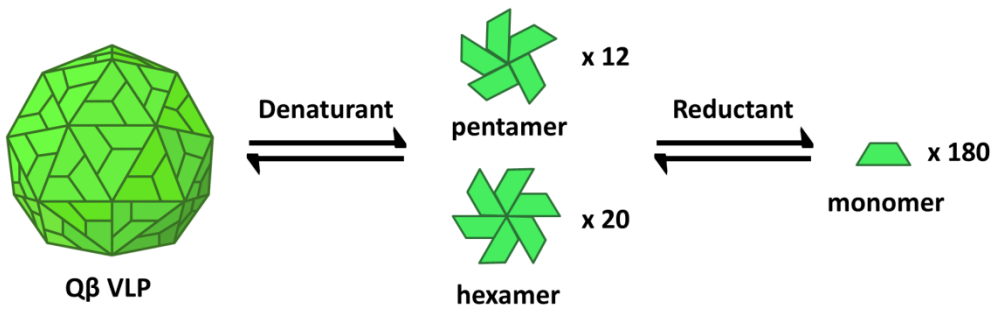


Figure 4.2: Disassembly of Qβ VLP under the treatment of denaturant and reductant.

In order to develop a powerful ^{19}F NMR detection tool for the disassembly of Qβ VLP, it is crucial that different multimeric states of the VLP have a significant difference in the ^{19}F NMR spectra, which includes easily observable changes such as chemical shift, peak width and peak intensity. It is particularly vital that the particle has a totally different spectrum from the multimeric species to indicate its disassembly hence the release of cargo.

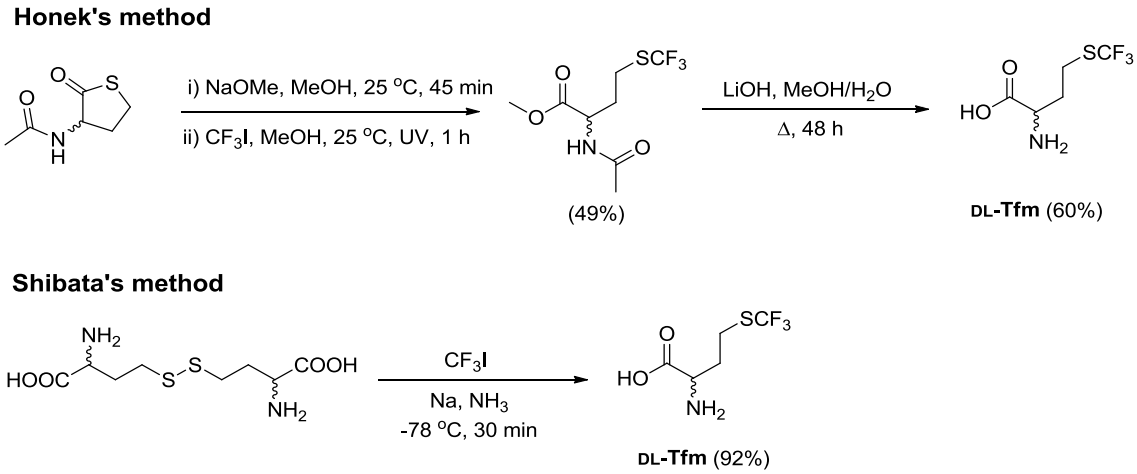
4.2 Results and Discussion

4.2.1 Establishment of the ^{19}F NMR detection system

4.2.1.1 Synthesis of DL-trifluoromethionine

Two different methods were used to synthesise DL-Tfm (**Scheme 4.1**). The Honek's method starts from *N*-acetyl-DL-homocysteine thiolactone.[6, 8] Thiolactone was first converted to thiolactate in the presence of sodium methoxide. The thiol group then reacted with trifluoromethyl iodide to form the fully protected racemic amino acid. The amino acid was then deprotected by heating under reflux in LiOH solution. The overall yield of Honek's route was 29% as a significant amount of material was lost during the first step due to the formation of disulphide bonds following addition of sodium methoxide. Tributylphosphine was added as a reducing agent but it failed to improve the yield of the reaction.

Shibata's method, on the other hand, provides a simple and efficient way for synthesising the amino acid.[14] The method starts from DL-homocystine, which reacts with trifluoromethyl iodide under the Birch reduction conditions in a one-pot process. The reaction gave a high yield and the amino acid can be purified easily using an ion-exchange (DOWEX H^+) column.



Scheme 4.1: Synthetic routes of DL-Tfm.

4.2.1.2 Mutagenesis, expression and purification of Qβ K16Tfm (Qβ-F) VLP

Qβ K16M mutant was obtained *via* site-directed mutagenesis of the plasmid coding the wild type Qβ. The DNA and protein sequences of the mutant are listed below:

Forward DNA sequence

ATGGCAAAAT TAGAGACTGT TACTTTAGGT AACATCGGGA AAGATGGAAT GCAAACTCTG
GTCCTCAATC CGCGTGGGGT AAATCCCACT AACGGCGTTG CCTCGCTTTC ACAAGCGGGT
GCAGTTCCTG CGCTGGAGAA GCGTGTTACC GTTTCGGTAT CTCAGCCTTC TCGCAATCGT
AAGAACTACA AGGTCCAGGT TAAGATCCAG AACCCGACCG CTTGCACTGC AAACGGTTCT
TGTGACCCAT CCGTTACTCG CCAGGCATAT GCTGACGTGA CCTTTTCGTT CACGCAGTAT
AGTACCGATG AGGAACGAGC TTTTGTTCGT ACAGAGCTTG CTGCTCTGCT CGCTAGTCCT
CTGCTGATCG ATGCTATTGA TCAGCTGAAC CCAGCGTATT AA

Amino acid sequence

AKLETVTLGN IGKDGMQTLV LNPRGVNPTN GVASLSQAGA VPALEKRVTV SVSQPSRNRK
NYKVQVKIQN PTACTANGSC DPSVTRQAYA DVTFSFTQYS TDEERAFVRT ELAALLASPL
LIDAIDQLNP AY

Calculated average isotopic mass for K16M = 14126 (N-terminal Met cleaved);

Calculated average isotopic mass for K16Tfm = 14180 (N-terminal Met cleaved)

The protein was expressed in the methionine auxotroph *E. coli* B834(DE3), which is different from other Qβ mutants mentioned in the previous chapter. L-Methionine was added

only during the growing stage of the bacteria, while Tfm was added before IPTG induction. It was essential to remove the methionine in the medium as much as possible by repeated washing since it can compete with Tfm for incorporation during protein translation. The particles were purified using the same way as other Qβ species. Fractions containing the intact particles were combined according to the results from the native agarose and SDS-PAGE gels (**Figure 4.3**), followed by LC-MS (**Figure 4.4**) and DLS (**Figure 4.5**) analysis. The expression yield of Qβ K16Tfm (Qβ-F) was relatively low (~ 5 mg / L of medium) when compared with the other Qβ species (30-40 mg / L of medium) reported in the previous chapters.

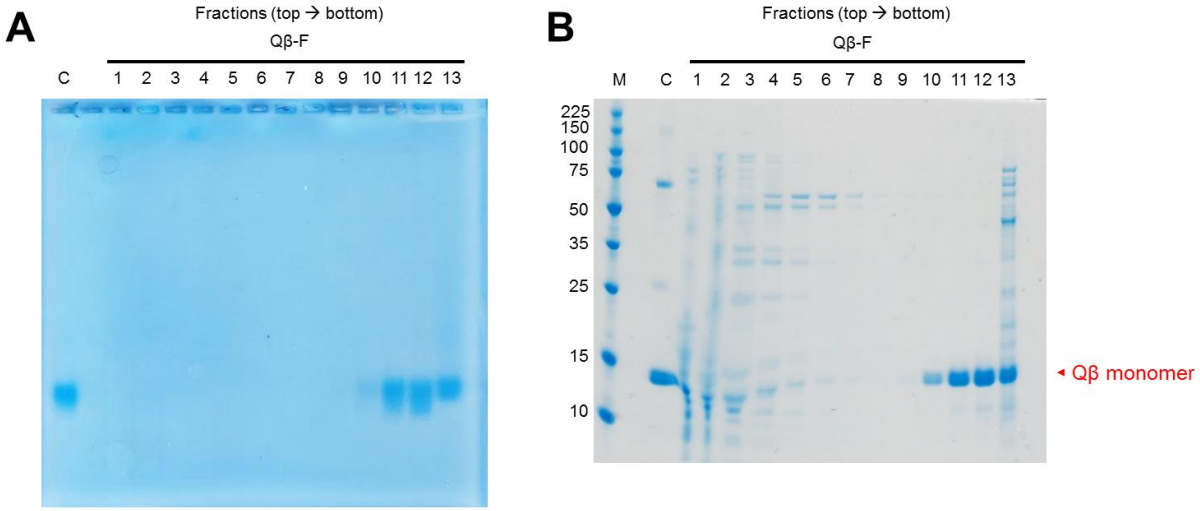


Figure 4.3: Coomassie Blue stained (A) native agarose and; (B) reducing SDS-PAGE gels showing the fractions obtained after sucrose gradient centrifugation of Qβ-F VLP. C: Qβ K16M control; M: protein marker.

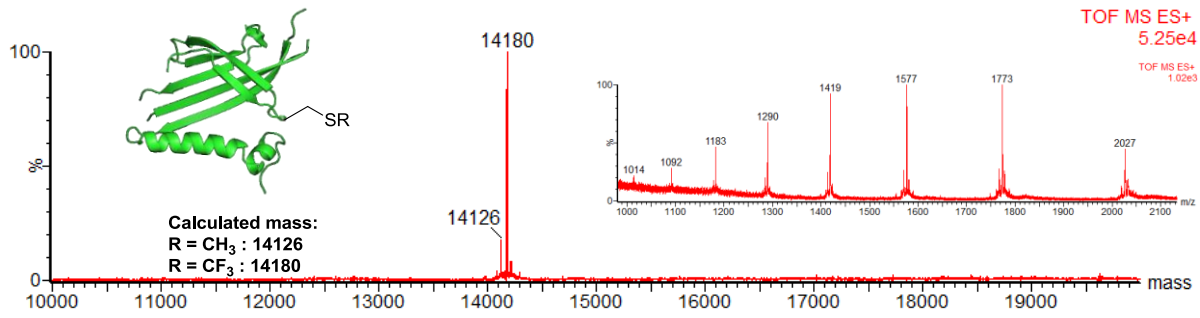


Figure 4.4: The deconvoluted mass spectrum and the corresponding raw ion series of the purified Qβ-F VLPs. The VLPs were broken down to monomers for LC-MS analysis.

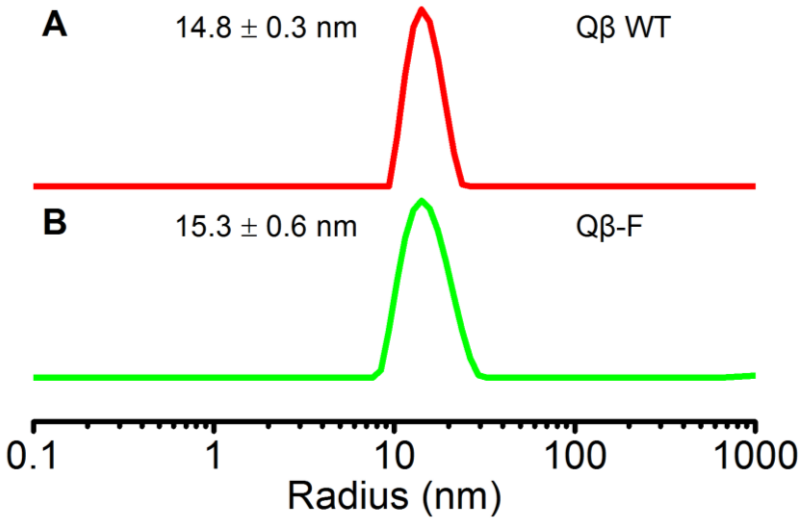


Figure 4.5: DLS results of the purified (A) Qβ WT and; (B) Qβ-F VLPs.

As shown by the mass spectrum, only around 85% Tfm incorporation was achieved. A lot of effort was spent optimising the expression conditions and improve the incorporation yield (**Table 4.1**), but to no avail. The parameters tested include expression time, expression temperature, amount of $_{DL}$-Tfm and $_{L}$-Met added, number of rounds of washing, etc. Indeed, the incorporation yield obtained is similar to the highest one reported in other proteins so far.[8] It has been shown before by crystallography that the conformation of *E. coli* methionyl-tRNA synthetase bound with Tfm is different from that bound with methionine.[15] The poor efficiency of methionyl-tRNA synthetase for incorporating Tfm renders the residual methionine in the medium very competitive even when has been removed by repeated washing. On the other hand, methionine analogues such as difluoromethionine (Dfm) and azidohomoalanine (Aha) are believed to give a similar enzyme conformation as that of methionine,[15] thus resulting in complete incorporation.[11, 16-18]

Table 4.1: Expression trials of Qβ-F VLPs

Entry	Scale † (mL)	L-Met (mM)	Washing *	30 °C incubation	DL-TFM (mM)	Expression time (h)	Expression temp (°C)	Rounds of purification	Protein yield (mg / L of medium)	Incorporation yield (%)
1	180	0.1	0.65V (4 times)	Yes	6	12	37	2	0.1	75
2	180	0.1	0.65V (4 times)	Yes	2	12	37	2	0.3	80
3	180	0.05	0.65V (4 times)	Yes	2	3	37	1	3.8	80
4	180	0.1	0.65V (4 times)	Yes	2	3	37	1	15.6	88
5	180	0.1	0.65V (4 times)	No	2 #	3	37	1	20.7	86
6	180	0.1	0.65V (4 times)	No	2	3	37	1	13.2	82
7	180	0.1	0.65V (4 times)	No	6	3	37	1	1.8	89
8	180	0.1	0.65V (4 times)	No	2	16	25	1	5.3	70
9	1500	0.1	0.52V (5 times)	No	2 #	3	37	2	0.3	87
10	1500	0.1	0.52V (5 times)	No	2 #	3	37	1	3.0	85
11	1500	0.1	0.22V (6 times)	No	2 #	3	37	1	2.8	86
12	1500	0.1	0.22V (6 times)	No	1.7 #	3	37	1	5.4	83

Growing conditions: SelenoMet™ base supplemented with SelenoMet™ nutrient mix, 100 μg/mL ampicillin, 37 °C, grew until OD_{600} ~ 0.6; † Volume of SelenoMet™ base medium used; * V: volume of original medium used for growing; # DL-TFM was added before the cell suspension.

To verify this hypothesis, a simple competitive experiment was performed by adding both Tfm and Aha to the medium before IPTG induction. Surprisingly, 100% incorporation of Aha resulted as shown by the mass spectrum (**Figure 4.6**). No Tfm or methionine incorporation was observed. The Aha also acted as a positive control to ensure the conditions used for unnatural amino acid incorporation was working well. Since the residual methionine in the pellet can never be removed completely, it was difficult to increase the incorporation yield of Tfm unless the enzyme was manually evolved to suit Tfm specifically.

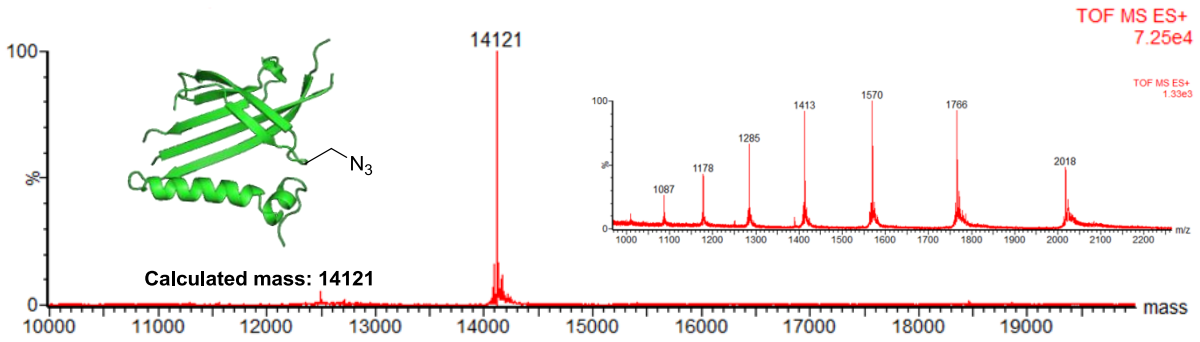


Figure 4.6: Deconvoluted mass spectrum and corresponding raw ion series of the purified Qβ K16M mutant in the presence of 2 mM DL-TFM and 1 mM L-Aha during IPTG induction. The VLPs were broken down to monomers for LC-MS analysis.

4.2.1.3 ^{19}F NMR analysis of Qβ-F VLPs

When analysed by ^{19}F NMR spectroscopy, the intact Qβ-F VLPs gave a large, broad peak with a shift of – 40.68 ppm (FWHH = ~240 Hz) (**Figure 4.7**). Controlled disassembly of the particles was achieved through titrated addition of denaturant and reductant and monitored by ^{19}F NMR. Denaturant SDS (0.2 M) gave a single sharp peak with a shift of – 40.90 ppm (FWHH = 11 Hz). This significant change of chemical shift and peak width suggested clear change of particle structure, which was expected (**Figure 4.2**). Further treating the same sample with reductant DTT (50 mM) led to formation of monomers, characterised by similar chemical shift and peak width to the multimers (δ_F = – 40.90 ppm;

FWHH = 7 Hz) (refer to Chapter 4.2.2 for the difference between 'multimers' and 'monomers').

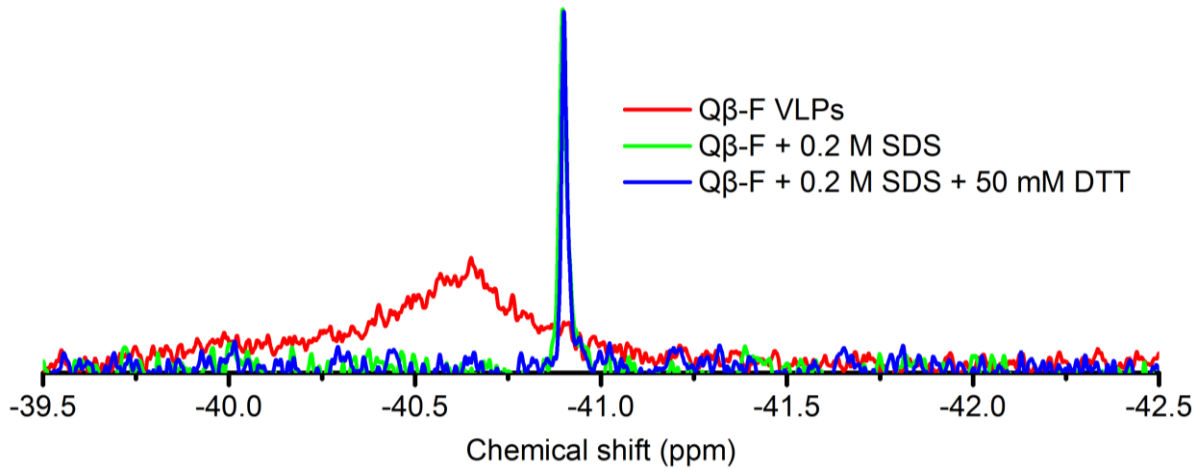


Figure 4.7: ^{19}F NMR spectra of Qβ-F VLPs under different conditions. 1,1,1-Trifluroacetone (TFAcetone) was added to all the samples as an internal reference. The reference peak was fixed to – 86.40 ppm relative to $CFCl_3$. Intensity rescaled for comparison (refer to **Figure S.3A** in Appendix).

The R_2 relaxation rates of Qβ-F VLP and SDS+DTT-treated Qβ-F VLP were also measured (**Figure 4.8**) to get more information about the rough sizes of the species. The VLP associated peak was found to have an R_2 almost two orders of magnitude larger than the monomer one. These values are consistent with the peak widths observed in the NMR spectra. For large macromolecule such as Qβ VLP, the transverse relaxation time (T_2) is expected to be relatively short, thus giving a broad peak and a large R_2 value.[19] When Qβ VLP is broken down into smaller multimers or monomers, the peak will become much sharper. The above results revealed that intact Qβ-F VLP and the species broken down from it had a significant difference in chemical shift, peak width and peak intensity. This provided a good start for developing the disassembly detection method.

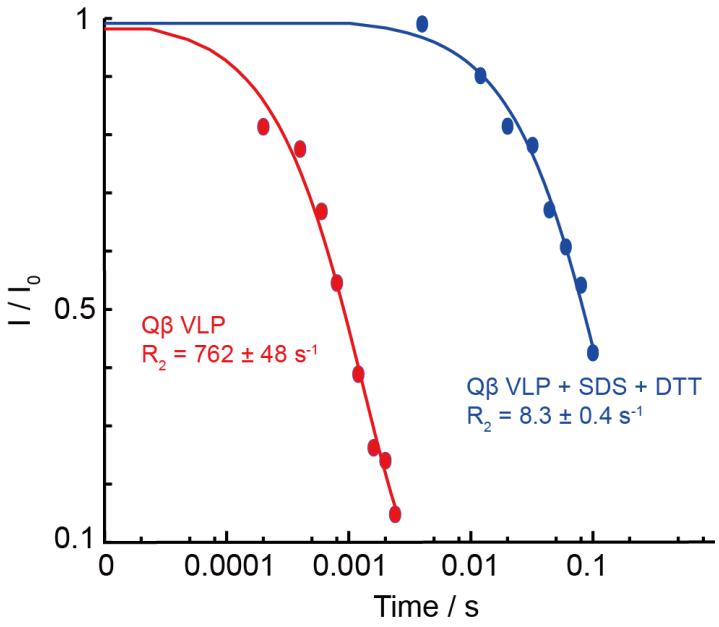


Figure 4.8: The transverse relaxation times (T_2) of the two resonances arose from Qβ-F VLP and SDS+DTT-treated Qβ-F VLP determined at 298K. The R_2 relaxation rates were calculated based on the data.

4.2.2 Investigating the difference between Qβ 'multimers' and 'monomers'

4.2.2.1 Mutagenesis, expression and purification of Qβ K16Tfm C74S C80S (Qβ-CS)

As shown by the NMR spectra, the addition of DTT to SDS-treated Qβ-F did not cause much change in the spectrum. It is possible that the difference between multimers and monomers are not resolved. Another possibility is that SDS and DTT may not be powerful enough to break all the particles into monomers.

To get a better idea, a control Qβ species (Qβ K16M C74S C80S) with no disulphide bonds was proposed. It is believed that this species can produce pure monomers when it is treated with SDS. The mutant was generated *via* site-directed mutagenesis in which all the cysteine residues were converted to isosteric but non-reactive serine. The DNA and protein sequences of the mutant are listed below:

Forward DNA sequence

ATGGCAAAAT TAGAGACTGT TACTTTAGGT AACATCGGGA AAGATGGAAT GCAAACTCTG
GTCCTCAATC CGCGTGGGGT AAATCCCACT AACGGCGTTG CCTCGCTTTC ACAAGCGGGT
GCAGTTCCTG CGCTGGAGAA GCGTGTTACC GTTTCGGTAT CTCAGCCTTC TCGCAATCGT
AAGAACTACA AGGTCCAGGT TAAGATCCAG AACCCGACCG CTAGCACTGC AAACGGTTCT
TCGGACCCAT CCGTTACTCG CCAGGCATAT GCTGACGTGA CCTTTTCGTT CACGCAGTAT
AGTACCGATG AGGAACGAGC TTTTGTTCGT ACAGAGCTTG CTGCTCTGCT CGCTAGTCCT
CTGCTGATCG ATGCTATTGA TCAGCTGAAC CCAGCGTATT AA

Amino acid sequence

AKLETVTLGN IGKDGMQTLV LNPRGVNPTN GVASLSQAGA VPALEKRVTV SVSQPSRNRK
NYKVQVKIQN PTASTANGSS DPSVTRQAYA DVTFSFTQYS TDEERAFVRT ELAALLASPL
LIDAIDQLNP AY

Calculated average isotopic mass for K16M C70S C80S = 14094 (N-terminal Met cleaved);

Calculated average isotopic mass for K16Tfm C74S C80S = 14148 (N-terminal Met cleaved)

Qβ K16Tfm C74S C80S (Qβ-CS) VLPs were expressed and purified in the same way as that of Qβ-F. Fractions containing the intact particles were combined according to the results from the native agarose and SDS-PAGE gels (**Figure 4.9**), followed by LC-MS (**Figure 4.10**) and DLS (**Figure 4.11**) analysis. The incorporation efficiency of Tfm in Qβ-CS was found to be lower (~75%) than that of Qβ-F.

4.2.2.2 Comparison of Qβ-F and Qβ-CS VLPs by gel electrophoresis

Both Qβ-F and Qβ-CS VLPs were analysed by Native-PAGE and non-reducing SDS-PAGE under different conditions. Native-PAGE can provide information about the disassembly of particles qualitatively, while non-reducing SDS-PAGE can give more details about the multimeric state of the protein.

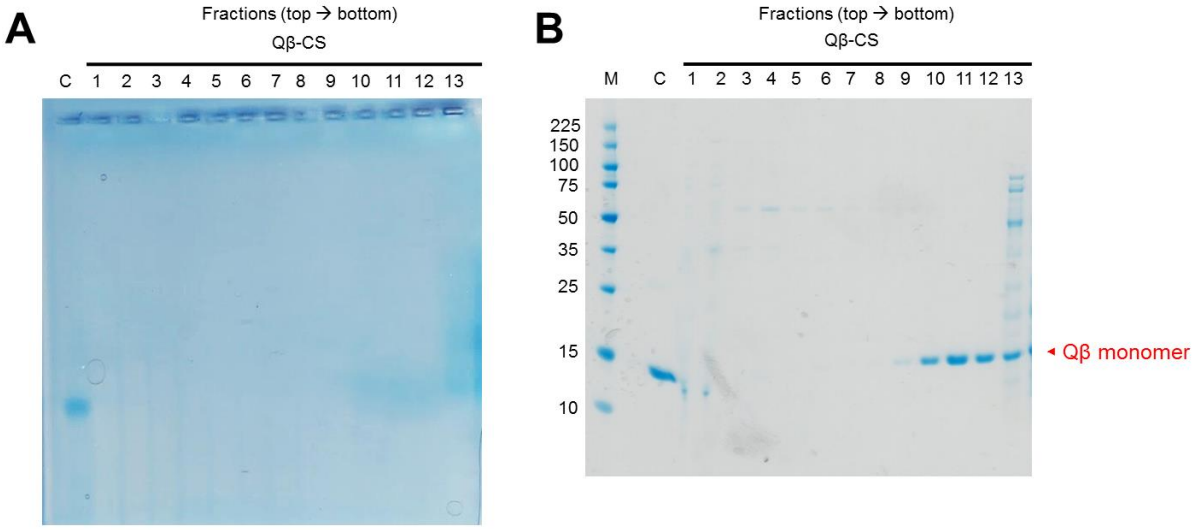


Figure 4.9: Coomassie Blue stained (A) native agarose and; (B) reducing SDS-PAGE gels showing the fractions obtained after sucrose gradient centrifugation of Qβ-CS VLP. C: Qβ K16M control; M: protein marker.

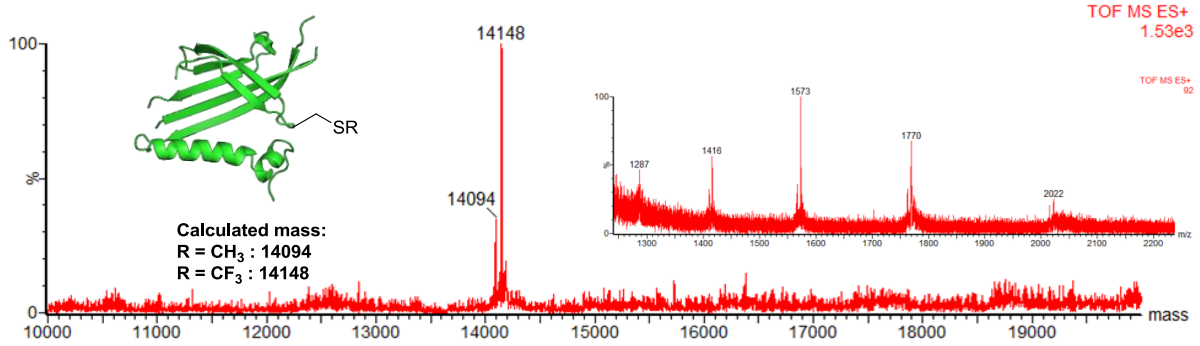


Figure 4.10: Deconvoluted mass spectrum and corresponding raw ion series of the purified Qβ-CS VLPs. The VLPs were broken down to monomers for LC-MS analysis.

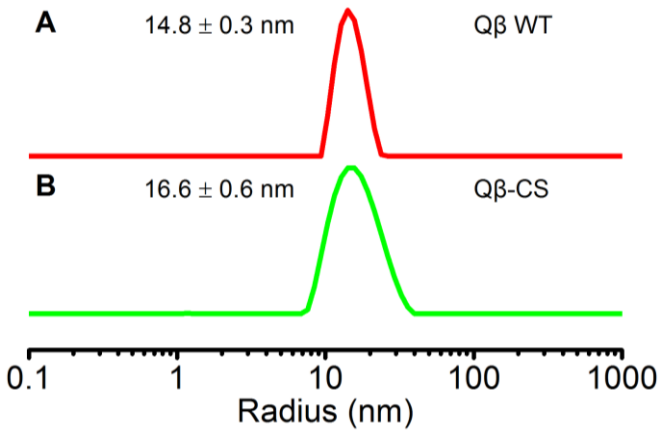


Figure 4.11: DLS results of the purified (A) Qβ WT and; (B) Qβ-CS VLPs.

As shown by the Native-PAGE (**Figure 4.12A**), the presence of VLPs was represented by two distinct bands at around 1100-1300 kDa. It is unknown why the particles gave two bands instead of one. It is possible that the particle is not stable and broken apart in the gel as it should have a mass of around 2550 kDa. Nonetheless, as shown by the Qβ K16M control, this appears to be a general observation for all the intact Qβ VLPs. When the particles were treated with SDS, smearing bands were observed. It was difficult to tell the mass of the multimers as the additional negative charge on SDS can interfere with the native gel result. However, it was still possible to compare between different species if they are treated with the same amount of SDS. As shown by the gel, it is obvious that the Qβ-F species gave a smearing band with a larger mass and hence a higher multimeric state than that of Qβ-CS.

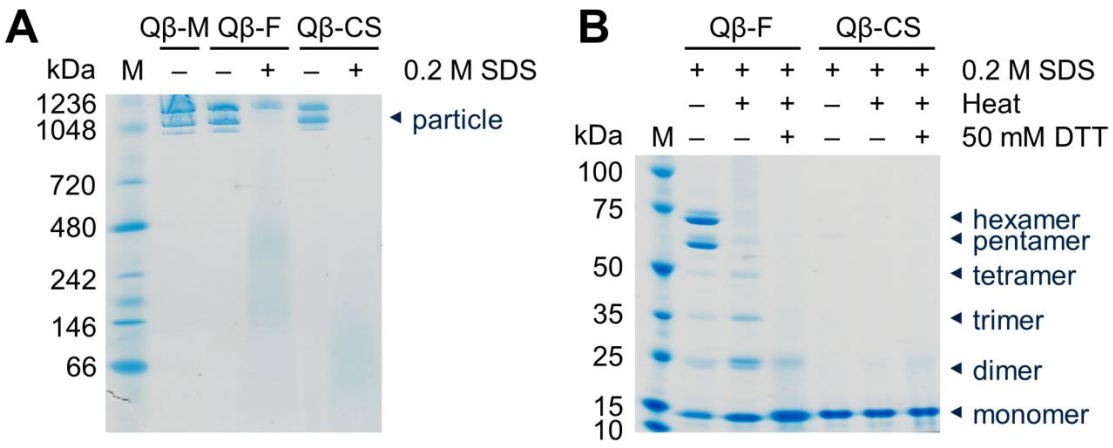


Figure 4.12: Coomassie Blue stained (A) Native-PAGE and; (B) non-reducing SDS-PAGE gels showing the multimeric states of Qβ-F and Qβ-CS under different conditions. M: protein marker; Qβ-M: Qβ-K16M intact VLPs; Heat: incubated at 90 ^{o}C for 60 min.

The multimeric species were then further analysed by non-reducing SDS-PAGE (**Figure 4.12B**). Since reductant was not included in the loading dye, the multimeric state of the protein can be clearly seen. From the results, bands representing hexamers and pentamers were observed when Qβ-F was treated with SDS, which is consistent with our hypothesis

(**Figure 4.2**). The intensity of these bands reduced following heating and disappeared after the addition of DTT. Although a tiny dimer band was still observed after SDS+DTT treatment, the intense monomer band indicated that the disassembly approach *via* chemical treatment worked well. This also provided the first supporting evidences to our hypothesis that ^{19}F NMR spectroscopy cannot distinguish between multimers and monomers. For Qβ-CS, only pure monomeric species were observed when treated with SDS. It can thus act as a suitable monomeric control for the following experiments.

4.2.2.3 Comparison of Qβ-F and Qβ-CS VLPs by ^{19}F NMR spectroscopy

As shown by the ^{19}F NMR spectra (**Figure 4.13**), the chemical shifts of Qβ-CS VLPs in the absence and presence of SDS were – 40.63 ppm (FWHH = ~200 Hz) and – 40.88 ppm (FWHH = 6 Hz), respectively. Both the chemical shifts and FWHH are similar to those obtained from Qβ-F VLPs (**Figure 4.7**). The green peak observed in here essentially represents Qβ monomer as revealed by SDS-PAGE (**Figure 4.12B**). Combined with the data from the two Qβ species, these new findings provided additional evidence that monomers cannot be distinguished from multimers in ^{19}F NMR spectroscopy.

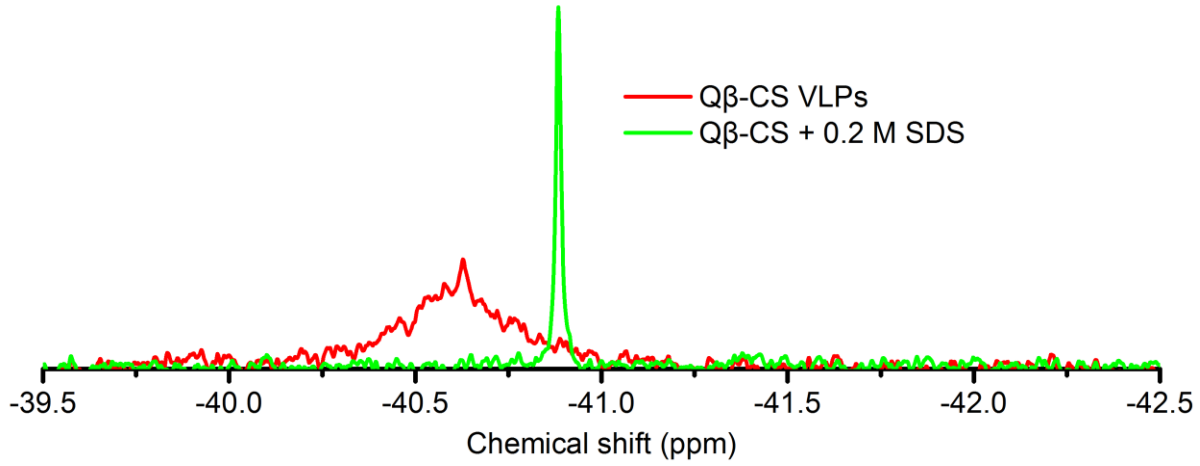


Figure 4.13: ^{19}F NMR spectra of Qβ-CS VLPs under different conditions. TFAcetone was added to all the samples as an internal reference. The reference peak was fixed to – 86.40 ppm relative to $CFCl_3$. Intensity rescaled for comparison (refer to **Figure S.3B** in Appendix).

4.2.2.4 Comparison of Qβ-F and Qβ-CS VLPs by ^{19}F diffusion NMR analysis

Qβ-F and Qβ-CS VLPs were further characterised by ^{19}F diffusion NMR experiments under different conditions (**Figure 4.14**). Diffusion NMR spectroscopy is a powerful technique which combines radio-frequency pulses used in routine NMR measurement with magnetic field gradients that encode spatial information. The diffusion coefficient of the species corresponding to a particular peak can be obtained by employing this technique.

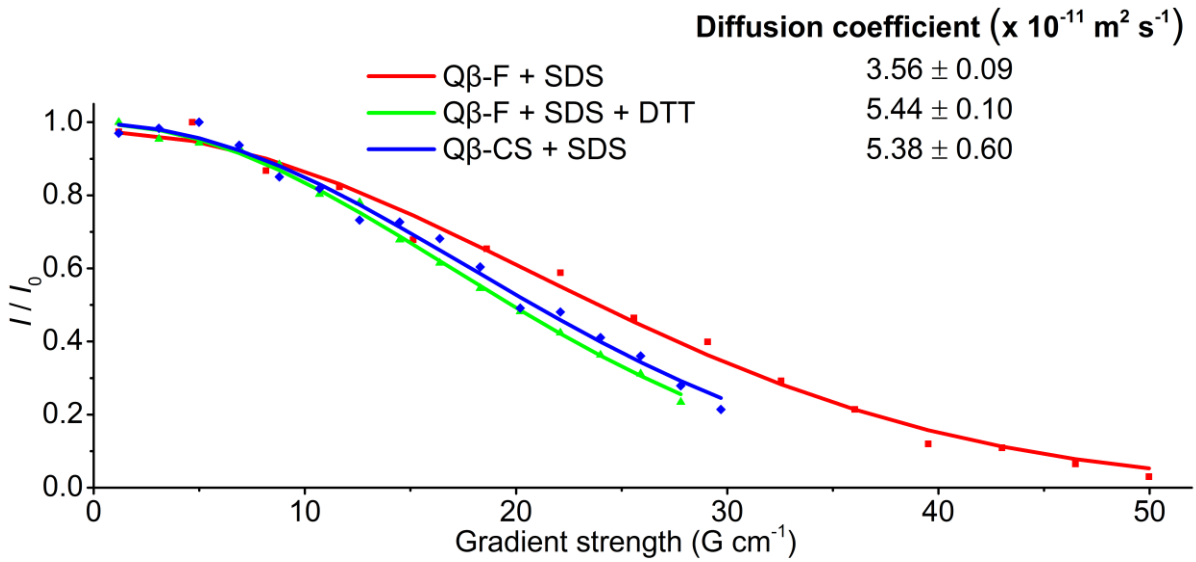


Figure 4.14: I / I_0 versus gradient strength of various Qβ species. Diffusion coefficients were calculated using TopSpin™ and normalised according to TFAcetone reference to eliminate environmental or solvent effects (refer to **Table S.1** in Appendix). Low intensity of broad resonance for intact VLPs precluded determination of coefficients. I: observed intensity; I_0: reference intensity (gradient strength = 0).

Despite the similar chemical shifts of 'multimers' (Qβ-F+SDS) and 'monomers' (Qβ-F+SDS+DTT) in 1D ^{19}F NMR spectra, their diffusion coefficients were strikingly different. Notably, adding DTT in addition to SDS to Qβ-F VLPs resulted in an increase of diffusion coefficient, which is then directly comparable to diffusion coefficient of SDS-treated Qβ-CS, as expected. These two species were thus predicted to have the same

multimeric state, which was believed to be monomers. Altogether, the results from gel electrophoresis, ^{19}F NMR and ^{19}F diffusion analysis provided strong evidence that resonances around ~ – 40.9 ppm can represent Qβ protein with different multimeric states.

4.2.3 Choice of wordings: are they really 'monomers' and 'particles'?

4.2.3.1 'Monomer' versus 'small multimer'

The ^{19}F diffusion experiments provided a powerful clue for answering the identity of the – 40.9 ppm peak, but it also raised up a new question at the same time. The diffusion coefficient of the 'monomer' species, which was obtained by either treating Qβ-F with SDS+DTT or Qβ-CS with SDS, is smaller than the expected value for a 14 kDa protein. The diffusion coefficient of a protein can be estimated from its molecular weight using the equation proposed by Young *et al.*:[20]

$$D = 8.34 \times 10^{-12} \left(\frac{T}{\eta M^{1/3}} \right) \quad \text{---------- (1)}$$

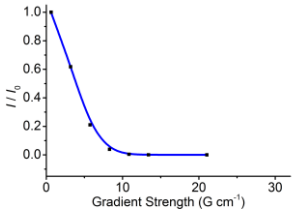
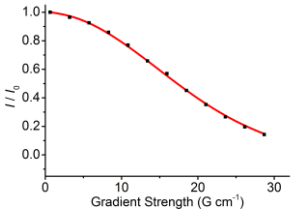
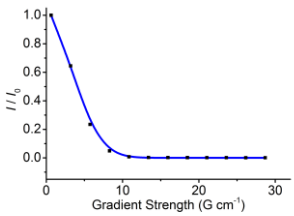
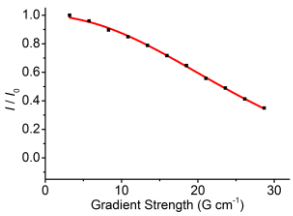
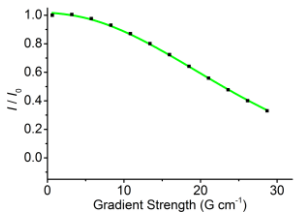
where diffusion coefficient D is in $m^2\ s^{-1}$; temperature T is in K; the solvent viscosity η is in cP; and molecular weight M is in $g\ mol^{-1}$.

The diffusion coefficient of Qβ monomer calculated by equation (**1**) is $1.15 \times 10^{-10}\ m^2\ s^{-1}$, which is around twice the measured value. By using the same equation, it can be easily shown that the 'monomer' and 'multimer' species are indeed multimers formed by around 10 and 34 monomers respectively. These results were somehow surprising at the beginning as they were different from the proposed disassembly theory of Qβ (**Figure 4.2**). Although clear monomers were observed in the non-reducing SDS-PAGE gel, this did not mean that the protein existed as pure monomers in solution. The pores in the gel can eliminate interactions between proteins when they pass through as long as they are not covalently linked. It is thus possible that the 'monomer' observed was actually something bigger than expected.

The second possibility which could account for the observed diffusion coefficient is the

presence of concentrated SDS (0.2 M) in the solution. SDS not only affects the viscosity of the solution, but can also interact with proteins so as to lower their diffusion coefficients.[21] Since it is impossible to obtain a diffusion coefficient of Qβ monomer without the presence of SDS, a protein with a similar mass was employed to address the effect of SDS. A simple ^{1}H diffusion experiment with hen egg-white lysozyme (HEL) was performed (**Table 4.2**). The measured diffusion coefficient of HEL was 9.23 x 10^{-11} m^2 s^{-1}, which is slightly smaller than the reported value in the literature (~1 x 10^{-10} m^2 s^{-1}).[22] This can be either due to the different calibrations between machines or the formation of protein dimers through non-covalent interactions in concentrated protein sample. When the protein was treated with SDS (0.2 M), the diffusion coefficient dropped to 5.10 x 10^{-11} m^2 s^{-1}, which is very similar to the value observed for Qβ 'monomer'. Interestingly, SDS itself also gave a diffusion coefficient similar to that of the protein, which indicated strong interaction between them. On the other hand, the diffusion coefficient of water molecule (i.e. HDO residue) only dropped slightly. This revealed that the effect of SDS on solution viscosity was not really significant.

Table 4.2: ^{1}H diffusion NMR analysis of HEL in the absence and presence of 0.2 M SDS

	Diffusion coefficient (x 10^{-11} m^2 s^{-1})		
	HDO	**HEL**	**SDS**
HEL in D_2O	192 ± 6	9.23 ± 0.06	
			---
HEL in D_2O + 0.2 M SDS	177 ± 3	5.10 ± 0.05	5.35 ± 0.03
			

From the results outlined above, it is clear that SDS has a significant effect on the diffusion coefficient of protein. However, evidence is indirect and it cannot fully confirm that the 'monomer' species observed is really a monomer. Conservatively, it is more appropriate to refer to the 'multimer' and 'monomer' as 'large multimer' and 'small multimer' respectively.

4.2.3.2 'Particle peak' versus 'particle-associated peak'

The same question also arose from the 'particle peak'. A broad peak was observed from both intact Qβ-F and Qβ-CS, but there is no direct evidence that the peak represents the particle itself. One possible solution is to measure the diffusion coefficient of the broad peak, but this was challenging owing to its large R_2 value (**Figure 4.8**). Thus, instead of plotting the whole diffusion graph, only two gradient strength points (i.e. 2% and 98%) were recorded to shorten the experimental time (**Figure 4.15**). A short diffusion gradient length (δ) was also used so as to increase the signal intensity. When compared with the SDS-treated Qβ-F VLP, the intact VLP showed a diffusion coefficient around 0.76 times smaller (i.e. ~ 2.7 x 10^{-11} m^2 s^{-1}). This value is larger than the one calculated by equation (**1**) (i.e. ~ 2.0 x 10^{-11} m^2 s^{-1}). Since Qβ is a perfect sphere with known hydrodynamic radius, the diffusion coefficient of the VLP can also be calculated using the Stokes–Einstein equation:

$$D = \frac{k_B T}{6 \pi \eta r} \quad \text{---------- (2)}$$

where diffusion coefficient D is in m^2 s^{-1}; Boltzmann constant k_B is in J K^{-1}; temperature T is in K; the solvent viscosity η is in Pa s; and hydrodynamic radius r is in m.

The diffusion coefficient calculated by equation (**2**) is even smaller (~1.6 x 10^{-11} m^2 s^{-1}) than the one calculated by equation (**1**), further indicating that the broad peak is unlikely coming from the particle itself. One possible reason for explaining this would be the presence of multimers in the particle sample, but this is unlikely as shown from the clean

DLS spectrum (**Figure 4.5**). A more complicated hypothesis would be that the 'particle peak' was indeed monomeric species that was in exchange with the intact particles. More experiments will be done in the future to address this problem. Even though the broad peak may not be coming directly from the particle itself, it is still a precious indicator which only appears in the presence of intact particles. This is sufficient to render the system applicable for monitoring particle disassembly. However, the term 'particle peak' will now be referred to as a 'particle-associated peak'.

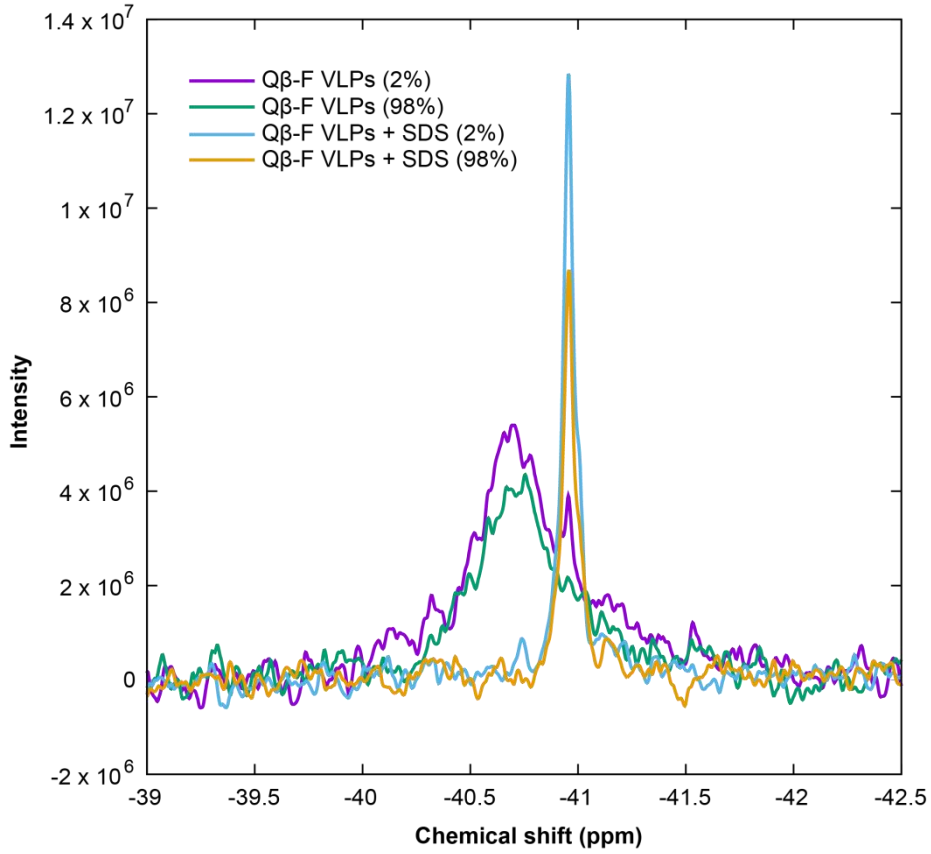


Figure 4.15: The ^{19}F NMR spectra of intact and SDS-treated Qβ-F VLPs with 2% and 98% gradient strength, with the maximum being 66 G cm^{-1}. Values for the measurement were Δ = 500 ms, δ = 0.7 ms.

4.3 Conclusion

In this chapter, Qβ VLPs bearing more than 450 fluorine atoms per particle were produced *via* biosynthetic labelling. These fluorinated VLPs have a similar structure to wild type Qβ as shown by DLS measurement. When analysed by ^{19}F NMR spectroscopy, a broad particle-associated peak was observed. This broad peak changed to an intense sharp peak when the VLPs were broken down into multimers by chemical treatment. It was found that ^{19}F NMR spectroscopy cannot distinguish between large and small multimers as proved by the Qβ-CS control. In spite of that, the method developed here has a large potential in monitoring the disassembly of Qβ VLP as long as a sharp change of signal is observed during particle break-down. The next step would be to test if this kind of fluorinated particles can be applied to real biological system with good sensitivity.

4.4 Experimental Section

4.4.1 General procedures

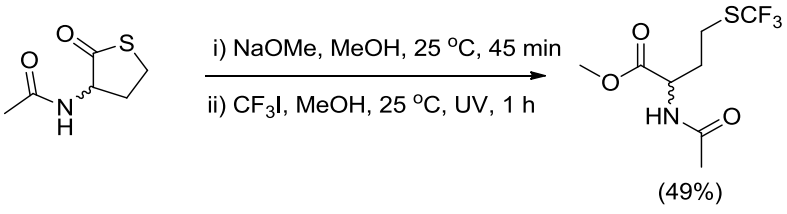
Chemicals were purchased from commercial suppliers and used without purification. NuPAGE® 12% Bis-Tris gel, NativePAGE™ 3-12% Bis-Tris gel and the corresponding buffer were purchased from Life Technologies™. Protein markers used were Perfect Protein™ Marker 10-225 kDa purchased from Novagen®. Vivaspin® concentrators were purchased from Life Technologies™. PD SpinTrap™ G-25, MiniTrap™ G-25, MidiTrap™ G-25 and PD10 desalting columns were purchased from GE Healthcare Life Sciences. Micro Bio-Spin™ P-6 columns were purchased from Bio-Rad. Dialysis cassettes and devices

(10,000 MW) were purchased from Thermo Scientific®. QIAprep Miniprep Kit was purchased from QIAGEN. PfuUltra™ high-fidelity DNA polymerase AD was purchased from Agilent Technologies. Restriction enzymes used were purchased from New England Biolabs. Primers used were purchased from Sigma Alrich. The bacterial strains BL21(DE3) and B834(DE3) were purchased from Novagen® .

PCR was performed in Applied Biosystems® 2720 Thermocycler. DNA sequencing was performed by Source BioScience. Ultracentrifugation was performed in Beckman Optima XL-80 Ultracentrifuge with SW 41 Ti rotor and Ultra-clear™ centrifuge tubes. Dynamic light scattering analysis was performed in Viscotek 801 DLS. Abs_{280} measurement for protein concentration was performed in NanoDrop ND-1000. Gel filtration was performed in GE Ä KTAFPLC ™. Small molecule and protein mass spectrometry analysis was performed in Bruker µTOF and Waters LCT Premier™ respectively. Chemical NMR study was performed in Bruker AVIII400 HD Nanobay instrument running TOPSPIN 3 equipped with a 5 mm z-gradient multinuclear BBFO probe. Protein NMR study was performed in Bruker AVIII 600 equipped with a dedicated 5mm BB-F/1H Prodigy N2 cryoprobe. Infrared spectrometry was performed in Bruker Tensor 27. Thin layer chromatography (TLC) was performed using aluminium sheets coated with 60F254 silica gel and visualised using ninhydrin solution.

4.4.2 Synthesis of DL-trifluoromethionine (DL-Tfm)

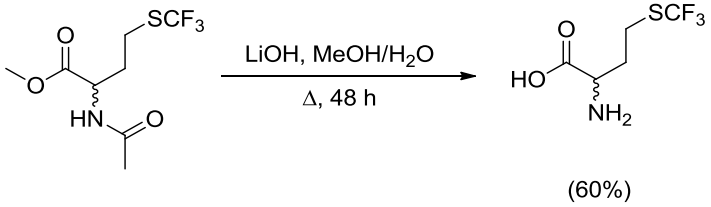
4.4.2.1 *N*-Acetyl-DL-(trifluoromethyl)homocysteine methyl ester



A 100 mL two-necked RB flask with a cold finger condenser attached was flame dried. *N*-Acetyl-DL-homocysteine thiolactone (1.0 g, 6.3 mmol) was added to the flask and dried by

vacuum. The powder was dissolved by anhydrous methanol (25 mL), followed by the addition of 25% sodium methoxide in methanol (3.6 mL, 16 mmol). The reaction mixture was stirred at 25 ^{o}C for 45 min under argon atmosphere. Trifluoromethyl iodide (2.1 mL, 25 mmol) was condensed slowly through the cold finger. The reaction was stirred at 25 ^{o}C for 1 h under UV light (ex: 365 nm, 110 W). The solvent was removed by vacuum. The yellow residue was dissolved in dichloromethane (60 mL) and washed with saturated $NaHCO_3$ (60 mL) and brine (60 mL). The organic layer was dried with anhydrous $MgSO_4$, filtered and evaporated under vacuum. The crude product was purified by silica column chromatography (1:1 EtOAc/PE) to give the target compound as a pale yellow oil (0.79 g, 49%); R_f = 0.6 (1:1 EtOAc/PE, v/v); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H = 2.05 (m, 4H, acetate and H_β), 2.28 (m, 1H, H_β), 2.90 (m, 2H, H_γ), 3.78 (s, 3H, COOMe), 4.72 (td, J 5.00, 7.78, 1H, H_α), 6.22 (d, J 6.84, 1H, NH); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C = 23.25 (NCO**C**$\mathbf{H_3}$), 26.10 (unresolved q, $^3J_{C\gamma\text{-}F}$ 2.18, C_γ), 33.30 (C_β), 51.30 (C_α), 52.95 (COO**C**$\mathbf{H_3}$), 131.01 (unresolved q, $^1J_{C\text{-}F}$ 306.23, CF_3), 170.37 (**C**$OOCH_3$), 172.14 (N**C**OCH_3); ^{19}F NMR (375 MHz, $CDCl_3$) δ_F = -41.29 (CF_3); LRMS m/z (ESI^+): measured: 282.0 ($M + Na^+$), calculated: 282.0; IR: 3269, 3071, 2957, 2365, 2329, 1744, 1655, 1544, 1440, 1375, 1307, 1279, 1217, 1160, 1103 cm^{-1}. The data agreed with the reported values in the literature.[6]

4.4.2.2 DL-Tfm (Honek's method)[6, 8]

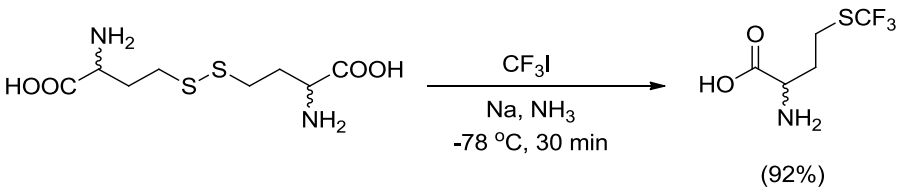


(60%)

N-Acetyl-DL-(trifluoromethyl)homocysteine methyl ester (0.72 g, 2.8 mmol) was dissolved in methanol (12 mL) in a 100 mL RB flask, followed by the addition of lithium hydroxide solution (0.67 g, 28 mmol in 12 mL H_2O). The reaction mixture was stirred under

reflux for 48 h. The solvent was removed by vacuum. The residue was dissolved in H_2O (20 mL) and neutralised by DOWEX® 50WX8 (H^+) 100-200 until pH < 7. The resulting suspension was loaded onto an ion exchange column (DOWEX® 50WX8 (H^+) 100-200). The column was prewashed with H_2O until the solution was neutral. The product was eluted by 5% NH_4OH (aq). Fractions containing the product were combined and evaporated under vacuum, followed by freeze-drying to give DL-Tfm as a white powder (0.34 g, 60%); R_f = 0.52 (3:1:1 EtOAc/AcOH/H_2O v/v/v); M.p. = 227-229 ^oC; ^1H NMR (400 MHz, D_2O) δ_H = 2.23 (m, 2H, H_β), 3.06 (m, 2H, H_γ), 3.78 (t, J 6.44, 1H, H_α); ^{13}C NMR (100 MHz, D_2O) δ_C = 25.42 (C_γ), 31.36 (C_β), 53.50 (C_α), 129.89 (q, J 328.97, CF_3), 174.25 (COOH); ^{19}F NMR (375 MHz, D_2O) δ_F = -41.18 (CF_3); LRMS m/z (ESI^+): measured: 204.1 ($M + H^+$), calculated: 204.0; IR: 2945, 2558, 2144, 1617, 1587, 1504, 1452, 1411, 1348, 1319, 1277, 1255, 1160, 1096 cm^{-1}. The data agreed with the reported values in the literature.[6]

4.4.2.3 DL-Tfm (Shibata's method)[14]



A 100 mL two-necked RB flask with a cold finger condenser attached was dried with a heat gun. DL-homocystine (1.0 g, 3.7 mmol) was added and dried under vacuum. The flask was cooled to -78 ^oC by dry ice-acetone bath. Ammonia (~20 mL) was condensed to the flask through the cold finger to dissolve the powder under stirring. Sodium (0.36 g, 16 mmol) was cut into pieces and added to the solution until the solution turned deep blue. Trifluoromethyl iodide (0.76 mL, 9.3 mmol) was condensed through the cold finger and the reaction mixture was stirred at -78 ^oC for 30 min. The flask was removed from the dry ice-acetone bath and the ammonia was left to evaporate. The resulting residue was dissolved

in H_2O (30 mL) and acidified by DOWEX® 50WX8 (H^+) 100-200 until pH ≈ 2. The resulting suspension was loaded onto an ion exchange column (DOWEX® 50WX8 (H^+) 100-200). The column was prewashed with H_2O until the solution was neutral. The product was eluted by 0.5% w/w NH_4OH (aq). Fractions containing the product were combined and evaporated under vacuum, followed by freeze-drying to give DL-Tfm as a white powder (1.4 g, 92%).

4.4.3 Mutagenesis of Qβ

4.4.3.1 Primers used

Mutant	Template		Primer sequence (5' → 3')
K16M or	WT or	F	GTAACATCGGGAAAGATGGAATGCAAACTCTGGTCCTCAATCCG
K16M C74S C80S	C74S C80S	R	CGGATTGAGGACCAGAGTTTGCATTCCATCTTTCCCGATGTTAC

4.4.3.2 Protocols

Refer to Chapter 2.

4.4.4 Preparation of B834(DE3) cell glycerol stock from DNA plasmid

Refer to Chapter 2. (The competent cell was changed to B834(DE3) instead)

4.4.5 Expression of Qβ VLPs

4.4.5.1 Qβ K16M

Refer to Chapter 2.

4.4.5.2 Qβ-F and Qβ-CS (optimised protocol)

Glycerol stock of B834(DE3) cells (1 µL) was added to SelenoMet™ base (25 mL) supplemented with nutrient mix (20X), 0.1 mM L-methionine and 100 µg / mL ampicillin. The mixture was shaken at 37 °C for 16 h with a speed of 250 rpm. The overnight culture (12 mL) was added to a flask containing SelenoMet™ base (750 mL) supplemented with nutrient mix (20X), 0.1 mM L-methionine and 100 µg / mL ampicillin. The cell culture was

shaken at 37 ^{o}C with a speed of 250 rpm until OD_{600} reached ~0.6. The cells were centrifuged down at 9500 x g for 10 min at 4 ^{o}C. The cell pellet (combined from 2 flasks) was resuspended in SelenoMet^{TM} base (300 mL) supplemented with 100 µg / mL ampicillin. The centrifugation and resuspension steps were repeated for 3 times and then 2 times with 400 mL base medium. The final cell pellet was resuspended in SelenoMet^{TM} base (20 mL) supplemented with 100 µg / mL ampicillin. The resuspended cell culture was added to a flask containing preheated SelenoMet^{TM} base (750 mL) supplemented with nutrient mix (20X), 100 µg / mL ampicillin and 1.67 mM DL-Tfm. The cell culture was shaken at 37 ^{o}C with a speed of 250 rpm for 1 h. 1 mM IPTG was added and the protein was expressed at 37 ^{o}C for 3 h. The cells were then centrifuged down at 9500 x g for 10 min at 4 ^{o}C. The cell pellet was resuspended in TBS buffer (50 mM Tris, 150 mM NaCl, pH 7.5) (20 mL). After sonication (5 min, 20% power, 2 sec ON/OFF), the lysed solution was pelleted down at 58,000 x g for 30 min at 4 ^{o}C.

4.4.6 Purification and characterisation of Qβ VLPs

Refer to Chapter 2.

4.4.7 NMR analysis of Qβ VLPs

4.4.7.1 1D ^{19}F NMR of Qβ VLPs with chemical treatment

Qβ VLP protein solution was diluted in TBS with 10% D_2O. 0.05 mM of TFAcetone was added to the solution as a reference. After analysing the original sample, the sample was treated with 0.2 M SDS and analysed again. For Qβ-F VLP, the sample was further treated with 50 mM DTT and heated at 90 ^{o}C for 60 min, followed by analysis. ^{19}F NMR spectra were typically acquired at 298 K with acquisition times of 0.5 s and recovery delays of 2 s. The number of transients collected varied according to sample concentration and resonance linewidth, ranging from 64 to 4096 per spectrum.

4.4.7.2 ^{19}F diffusion analysis of Qβ VLPs

The same Qβ VLP protein samples were used for diffusion analysis. Diffusion NMR measurements were acquired at 298 K using a bipolar stimulated-echo sequence incorporating a longitudinal eddy current delay (5 ms). Parameters were optimised for the species of interest (protein or TFAcetone internal reference). Values for TFAcetone were Δ = 100 ms, δ = 4 ms with gradients of 2 – 40% of maximum 66 G cm^{-1}. 16 transients were collected for 16 gradient increments. Typical values for protein measurements were Δ = 200 ms, δ = 5 ms with gradients from 2% up to 80% of maximum 66 G cm^{-1}. The number of transients collected varied according to sample, ranging from 320-1024 for 8 or 16 gradient increments. For each experiment the peak of interest was placed on-resonance. All experiments employed smoothed-square profile gradient pulses (Bruker SMSQ10.100) and data fitting to the Stejskal-Tanner equation was performed using Bruker TOPSPIN 3:

$$I = I_0 \ e^{-D\gamma^2 g^2 \delta^2 (\Delta - \frac{\delta}{3})} \quad \text{---------- (2)}$$

where observed intensity I and reference intensity I_0 have no unit; diffusion coefficient D is in m^2 s^{-1}; gyromagnetic ratio of the observed nucleus γ is in Hz G^{-1}; gradient strength g is in G m^{-1}; the length of the gradient δ is in s; and the diffusion time Δ is in s.

4.4.7.3 Measurement of R_2 relaxation rate

The transverse relaxation times of the two resonances were determined at 298K. For the low molecular weight resonance, the signal intensity was measured after a varying number of (τ-π-τ)n elements where τ = 1 ms and n is an even integer which is varied from 2 to 50. Each spectrum was recorded with 512 transients, an acquisition time of 1 second and 59,524 points. For the VLP resonance, owing to the very rapid relaxation, the signal intensity was determined after a (τ-π-τ) Hahn-echo sequence, whose duration varied from 100 to 1200 μs. In this case

each spectrum was recorded with 4800 transients, an acquisition time of 0.25 seconds and 4006 points.

4.5 References

1. Banerjee, D., Liu, A.P., Voss, N.R., Schmid, S.L. & Finn, M.G. Multivalent Display and Receptor-Mediated Endocytosis of Transferrin on Virus-Like Particles. *Chembiochem* **11**, 1273-1279 (2010).
2. Welsher, K. & Yang, H. Multi-resolution 3D visualization of the early stages of cellular uptake of peptide-coated nanoparticles. *Nat Nano* **9**, 198-203 (2014).
3. Lewis, J.D. et al. Viral nanoparticles as tools for intravital vascular imaging. *Nat Med* **12**, 354-360 (2006).
4. Ewers, H. et al. Single-particle tracking of murine polyoma virus-like particles on live cells and artificial membranes. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15110-15115 (2005).
5. Kitevski-LeBlanc, J.L. & Prosser, R.S. Current applications of ^{19}F NMR to studies of protein structure and dynamics. *Progress in Nuclear Magnetic Resonance Spectroscopy* **62**, 1-33 (2012).
6. Duewel, H., Daub, E., Robinson, V. & Honek, J.F. Incorporation of Trifluoromethionine into a Phage Lysozyme: Implications and a New Marker for Use in Protein ^{19}F NMR†. *Biochemistry* **36**, 3404-3416 (1997).
7. de Visser, P.C. et al. A novel, base-labile fluorous amine protecting group: synthesis and use as a tag in the purification of synthetic peptides. *Tetrahedron Letters* **44**, 9013-9016 (2003).
8. Holzberger, B., Rubini, M., Möller, H.M. & Marx, A. A Highly Active DNA Polymerase with a Fluorous Core. *Angewandte Chemie International Edition* **49**, 1324-1327 (2010).
9. Shepherd, C.M. et al. VIPERdb: a relational database for structural virology. *Nucleic Acids Res* **34**, D386-9 (2006).
10. Golmohammadi, R., Fridborg, K., Bundule, M., Valegard, K. & Liljas, L. The crystal structure of bacteriophage Q beta at 3.5 angstrom resolution. *Structure* **4**, 543-554 (1996).

11. Strable, E. et al. Unnatural amino acid incorporation into virus-like particles. *Bioconjugate Chemistry* **19**, 866-875 (2008).
12. Valegard, K., Fridborg, K. & Liljas, L. Crystallization and preliminary X-ray diffraction studies of the bacteriophage Qbeta. *Acta Crystallogr D Biol Crystallogr* **50**, 105-9 (1994).
13. Takamatsu, H. & Iso, K. Chemical evidence for the capsomeric structure of phage Q[beta]. *Nature* **298**, 819-824 (1982).
14. Yasui, H., Yamamoto, T., Tokunaga, E. & Shibata, N. Robust synthesis of trifluoromethionine and its derivatives by reductive trifluoromethylation of amino acid disulfides by CF3I/Na/Liq.NH3 system. *Journal of Fluorine Chemistry* **132**, 186-189 (2011).
15. Crepin, T. et al. Use of Analogues of Methionine and Methionyl Adenylate to Sample Conformational Changes During Catalysis in Escherichia coli Methionyl-tRNA Synthetase. *Journal of Molecular Biology* **332**, 59-72 (2003).
16. Cleve, P., Robinson, V., Duewel, H.S. & Honek, J.F. Difluoromethionine as a Novel ^{19}F NMR Structural Probe for Internal Amino Acid Packing in Proteins. *Journal of the American Chemical Society* **121**, 8475-8478 (1999).
17. Hovlid, M.L. et al. Encapsidated Atom-Transfer Radical Polymerization in Qβ Virus-like Nanoparticles. *ACS Nano* **8**, 8003-8014 (2014).
18. Yoder, N.C., Yüksel, D., Dafik, L. & Kumar, K. Bioorthogonal noncovalent chemistry: fluorous phases in chemical biology. *Current Opinion in Chemical Biology* **10**, 576-583 (2006).
19. Fiaux, J., Bertelsen, E.B., Horwich, A.L. & Wuthrich, K. NMR analysis of a 900K GroEL-GroES complex. *Nature* **418**, 207-211 (2002).
20. Dafik, L., d'Alarcao, M. & Kumar, K. Fluorination of Mammalian Cell Surfaces via the Sialic Acid Biosynthetic Pathway. *Bioorganic & medicinal chemistry letters* **18**, 5945-5947 (2008).
21. Dafik, L., d'Alarcao, M. & Kumar, K. Modulation of cellular adhesion by glycoengineering. *J Med Chem* **53**, 4277-84 (2010).
22. Yarema, K.J., Mahal, L.K., Bruehl, R.E., Rodriguez, E.C. & Bertozzi, C.R. Metabolic Delivery of Ketone Groups to Sialic Acid Residues: APPLICATION TO CELL SURFACE GLYCOFORM ENGINEERING. *Journal of Biological Chemistry* **273**, 31168-31179 (1998).

CHAPTER

5 Towards the application in biological context

5.1 Introduction

In the previous chapter, a ^{19}F NMR detection system for particle disassembly has been described. To apply this system in biological context, cell lysate was chosen as a starting point (**Figure 5.1**). The advantage of using cell lysate is that the efficiency of cell uptake of the particles is no longer a factor. However, it was still important to study the internalisation of Qβ VLPs in different cells as to choose the best candidate that can be used for cell study in the next stage.

Two types of cell lysate were chosen from bacterial and mammalian sources respectively. For bacterial cells, *E. coli* was selected as it is the natural host of Qβ bacteriophage. For mammalian cells, THP-1 monocyte, THP-1 derived macrophage and A549 epithelial cells were chosen for the screening process. By labelling Qβ VLPs with fluorescent dyes, the efficiency of internalisation of VLPs into different cells can be easily studied. The type of mammalian cell with the most promising VLPs uptake ability was chosen for lysate study.

It is expected that both the particle and multimer peaks would be broadened *in vivo* due to macromolecular crowding,[1] protein-protein[2] and charge[3] interactions. If a peak can be observed in cell lysate, it is very likely that the same can be seen in real cells given that the concentration of particles is the same. It is thus again the cellular uptake efficiency which determines the signal intensity.

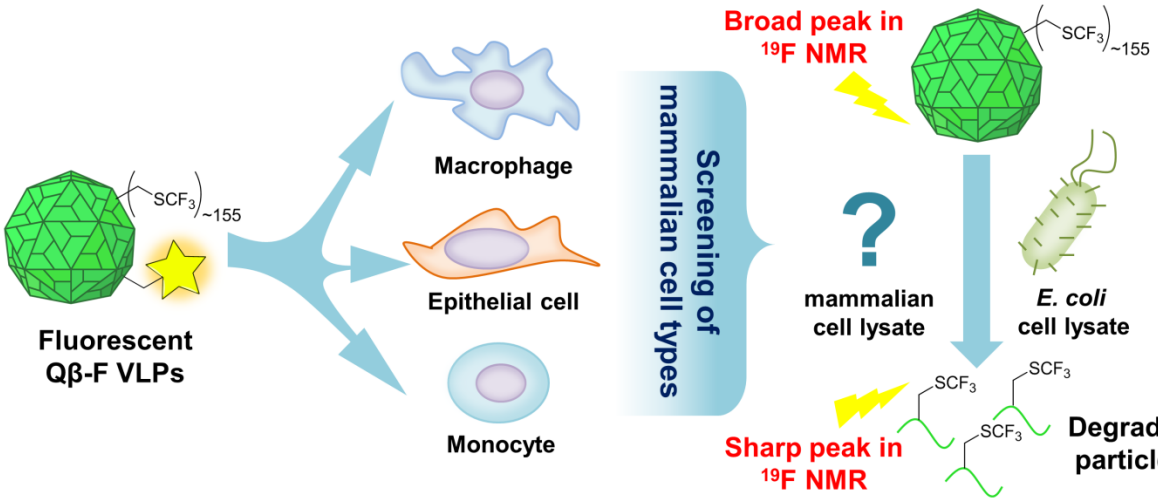


Figure 5.1: Multimodal tracking of VLP response to cells. VLPs were screened for both internalisation and stability to the corresponding internal cellular fluid (lysate) using combined fluorescent and ^{19}F-labeling.

5.2 Results and Discussion

5.2.1 Labelling of Qβ VLPs with fluorescent dyes

In order to study the internalisation of VLPs in different cells, both Qβ-F and Qβ-CS VLPs were first labelled with fluorescein. Both fluorescein isothiocyanate (FITC) and 5/6-carboxyfluorescein succinimidyl ester (fluo-NHS) were tested. Interestingly, only fluo-NHS was found to be coupled successfully by LC-MS. The same results were observed in other Qβ species (Qβ WT and Qβ K16M) in which only dyes carrying an NHS ester group can be coupled (data not shown). It is unknown that why Qβ VLPs have such a selectivity towards the coupling group, which is rarely observed in other simple proteins.

As revealed by the mass spectra, approximately 1-2 fluorescein molecules were coupled in each monomer for both the two species (namely Qβ-F-fluo and Qβ-CS-fluo) (**Figure 5.2**). This is sufficient for detection under flow cytometry or fluorescence microscopy. If too many

fluorescein molecules are added, the fluorescence signal will be quenched due to close proximity of fluorophores.[4]

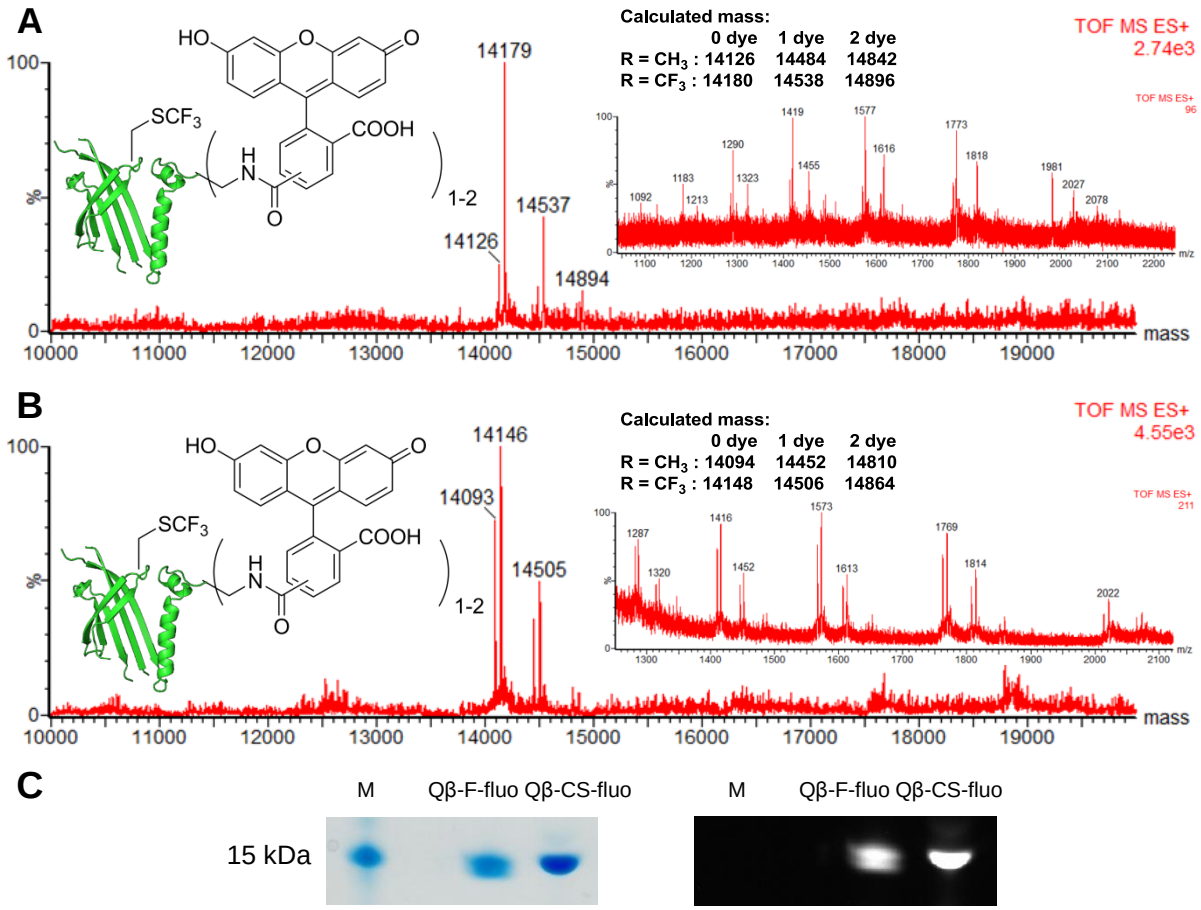


Figure 5.2: (A-B) The deconvoluted mass spectra and the corresponding raw ion series of (A) Qβ-F or; (B) Qβ-CS VLPs labelled with fluorescein-NHS (15 equiv.). The VLPs were broken down to monomers for LC-MS analysis; (C) Coomassie Blue stained (left) and 488 nm illuminated (right) reducing SDS-PAGE gel of the labelled Qβ species. M: protein marker.

5.2.2 Internalisation study of Qβ VLPs in mammalian cells

5.2.2.1 Flow cytometry

To study the uptake of VLPs by mammalian cells, labelled VLPs were incubated with various cell types for 1 h, 6 h or 24 h. As the cells started to die after 24 h when they were mixed with VLPs, it was not possible to obtain later time points (**Figure 5.3**). For this

experiment, only Qβ-F-fluo VLPs were tested in order to give approximate idea about the uptake efficiency. It is expected that Qβ-CS-fluo VLPs would give a similar result. Since there may be some free dye remaining in the protein solution, a control experiment was done in which free fluorescein dye (with no coupling group) was added to the cell culture. As a result, only cells which gave a stronger fluorescence signal than that of the free dye control were counted as fluorescein-positive cells (refer to **Figure S.4** in Appendix for the original histograms).

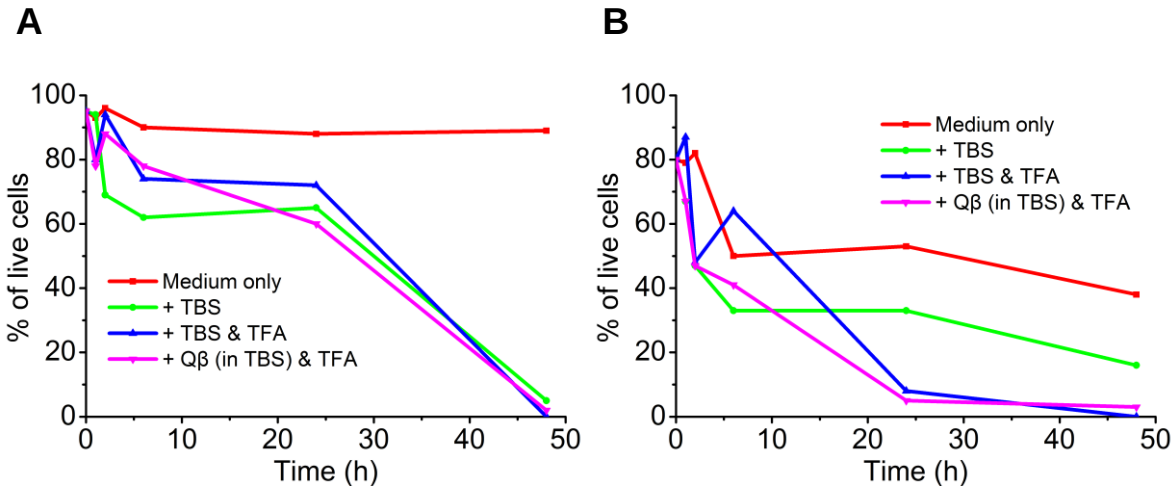


Figure 5.3: Survival graphs of (A) THP-1 monocytes and; (B) THP-1 derived macrophages under different conditions.

The results revealed that only THP-1 derived macrophage readily engulfed the particles, regardless of the time point (**Figure 5.4**). This is consistent with the phagocytic property of THP-1 derived macrophage reported in the literature.[5, 6] Although both THP-1 and A549 showed an increased uptake after 24 h, the status of the cells could not be determined at that time point, which may affect the accuracy of the measurement.

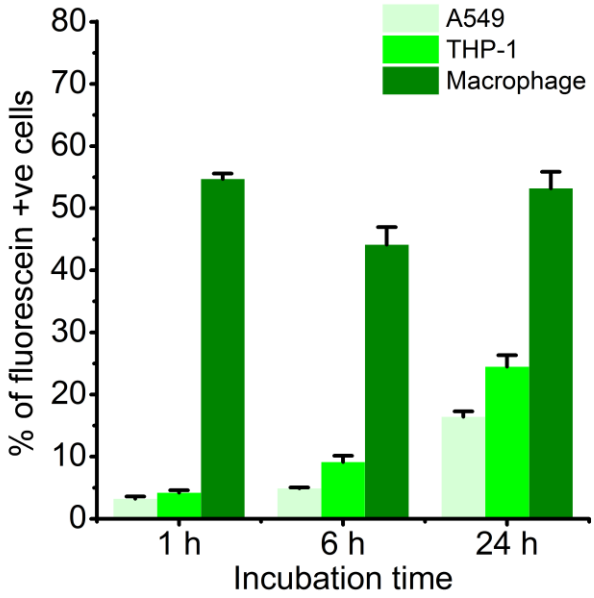


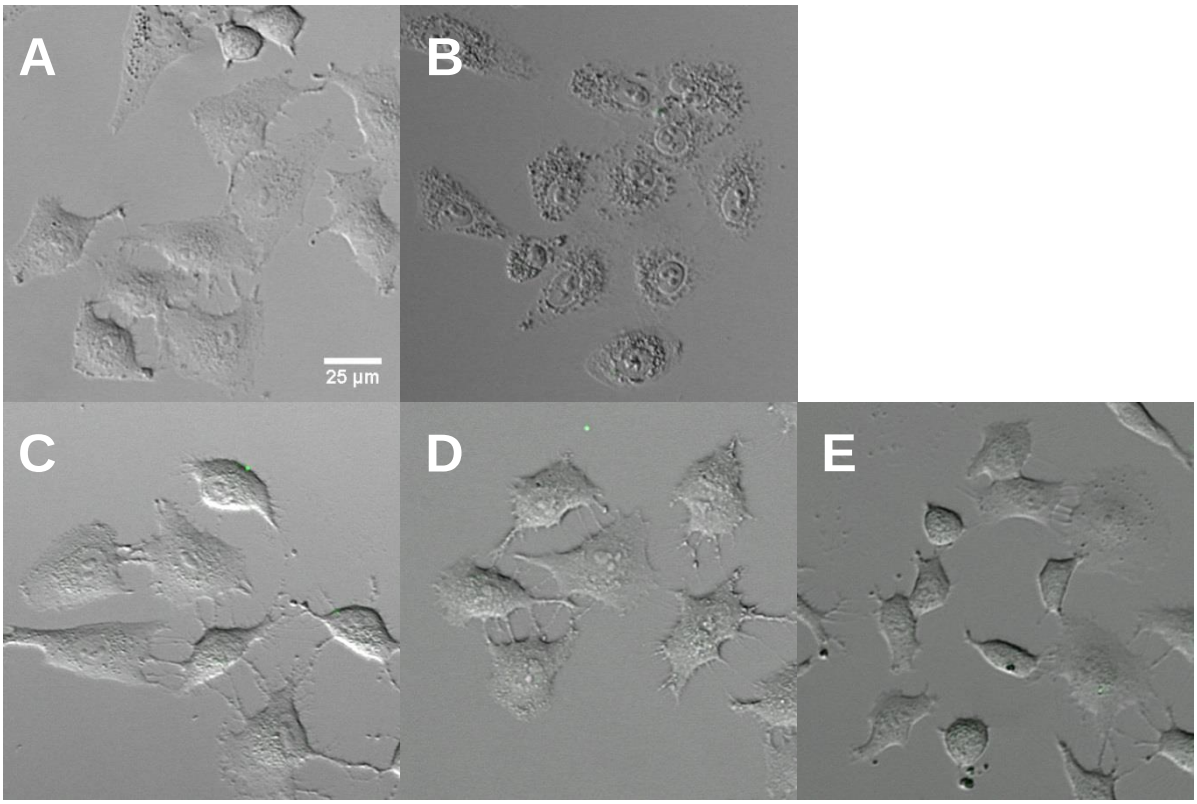
Figure 5.4: A graph showing the uptake of Qβ-F-fluo VLPs by mammalian cells after different time points studied by flow cytometry. The graph was plotted based on the raw histograms (refer to **Figure S.4** in Appendix).

5.2.2.2 Fluorescence microscopy

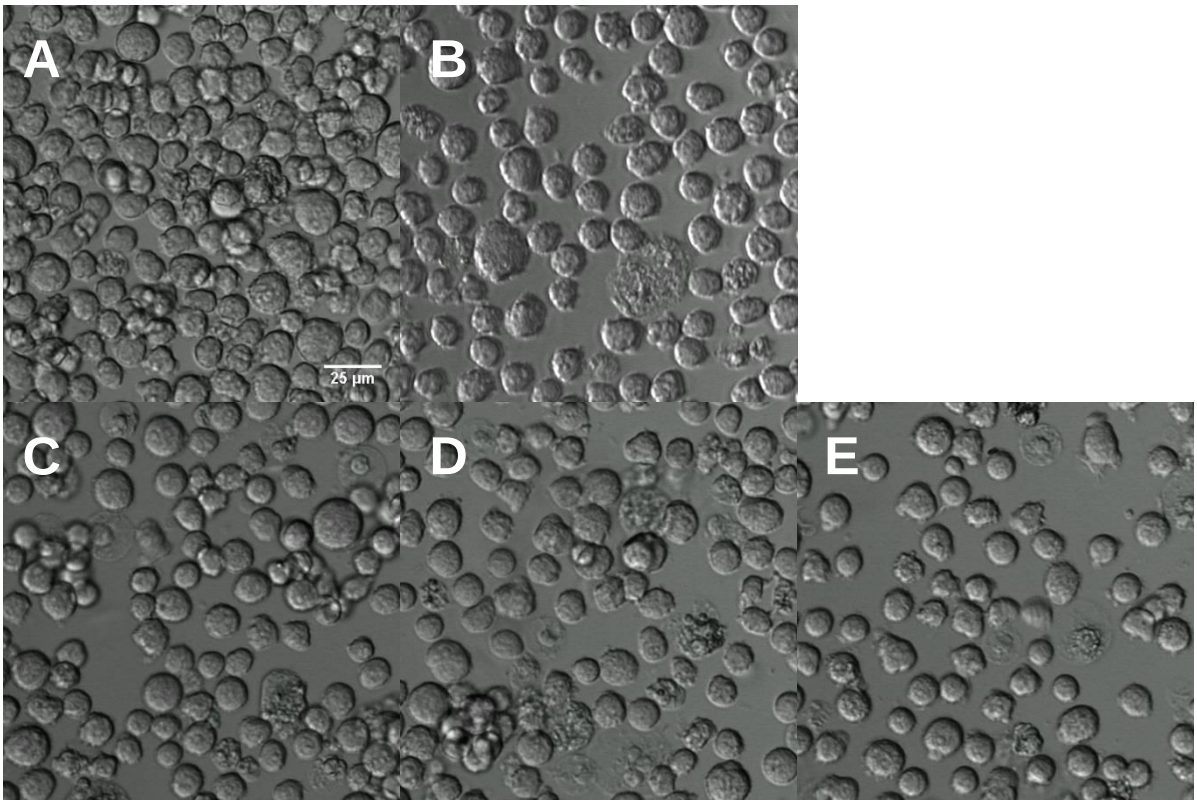
Although flow cytometry can provide excellent quantitative details about the percentage of fluorescein positive cells, it cannot distinguish between particles attached on the cell surface and those being internalised. To answer this question, the mammalian cells treated with Qβ-F-fluo VLPs were further studied by fluorescence microscopy.

Simple live-cell imaging was first performed to get a rough idea about the internalisation of particles (**Figure 5.5**). Similarly to the results from flow cytometry, only THP-1 derived macrophages showed the green fluorescence signal from the labelled VLPs. It was also apparent from the data that the particles were engulfed but not simply attaching on the cell surface. Interestingly, no fluorescence signal was observed in the 1 hour sample, which was contradictory to the flow cytometry results. It is possible that live-cell imaging was not sensitive enough to detect cells with tiny amount of fluorescent particles.

A549 epithelial cells



THP-1 monocytes



THP-1 derived macrophages

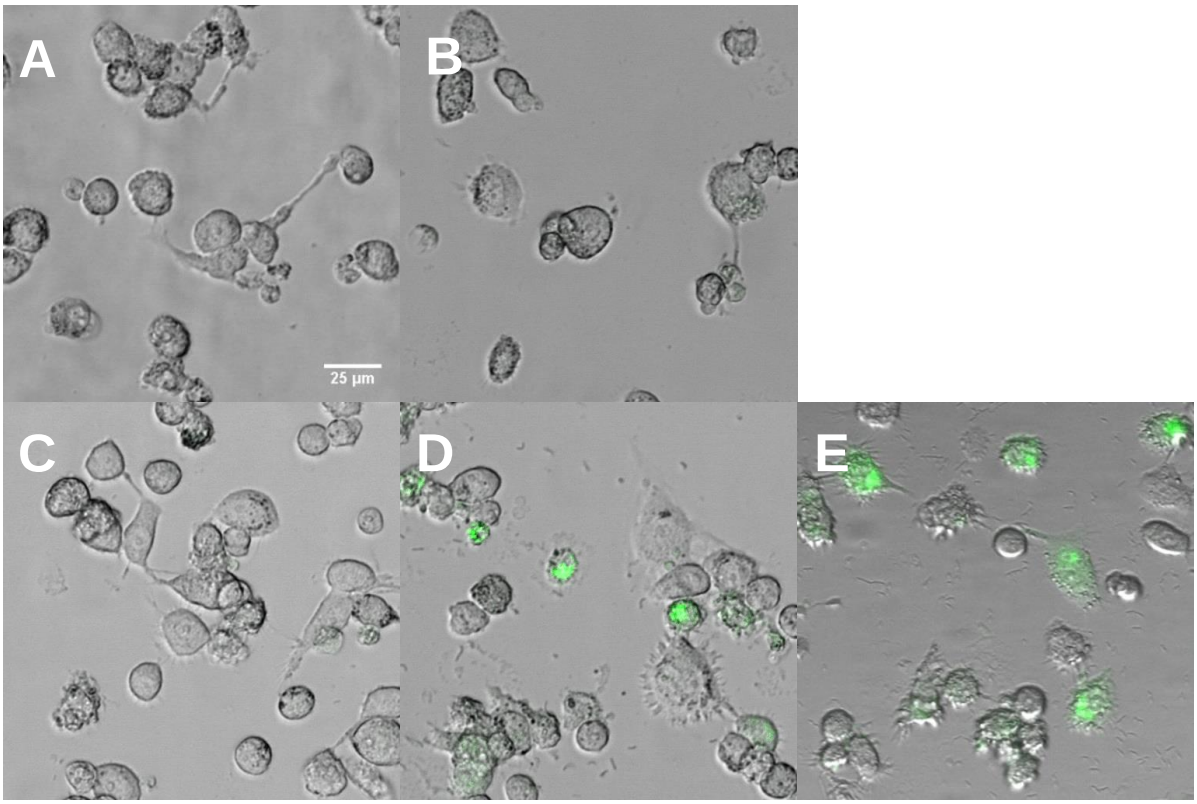


Figure 5.5: Microscopy images of mammalian cells treated with (A) nothing; (B) fluorescein dye; (C-E) Qβ-F-fluo VLPs for 1 h, 6 h and 24 h respectively. Green: Qβ-F-fluo VLPs

To improve the resolving power of the microscope, the cells were fixed before analysis. This allowed a more powerful objective lens requiring oil immersion to be used. Furthermore, antibodies targeting different cellular organelles can be applied so as to give a more detailed image about the sub-cellular localisation of the particles. For instance, by labelling the plasma membrane of the macrophage cells with fluorescent anti-CD14, the uptake of labelled VLPs can be clearly seen. As revealed by the images (**Figure 5.6**), it was obvious that both Qβ-F-fluo and Qβ-CS-fluo had already been internalised after 20 minutes of incubation, which is consistent with the flow cytometry results.

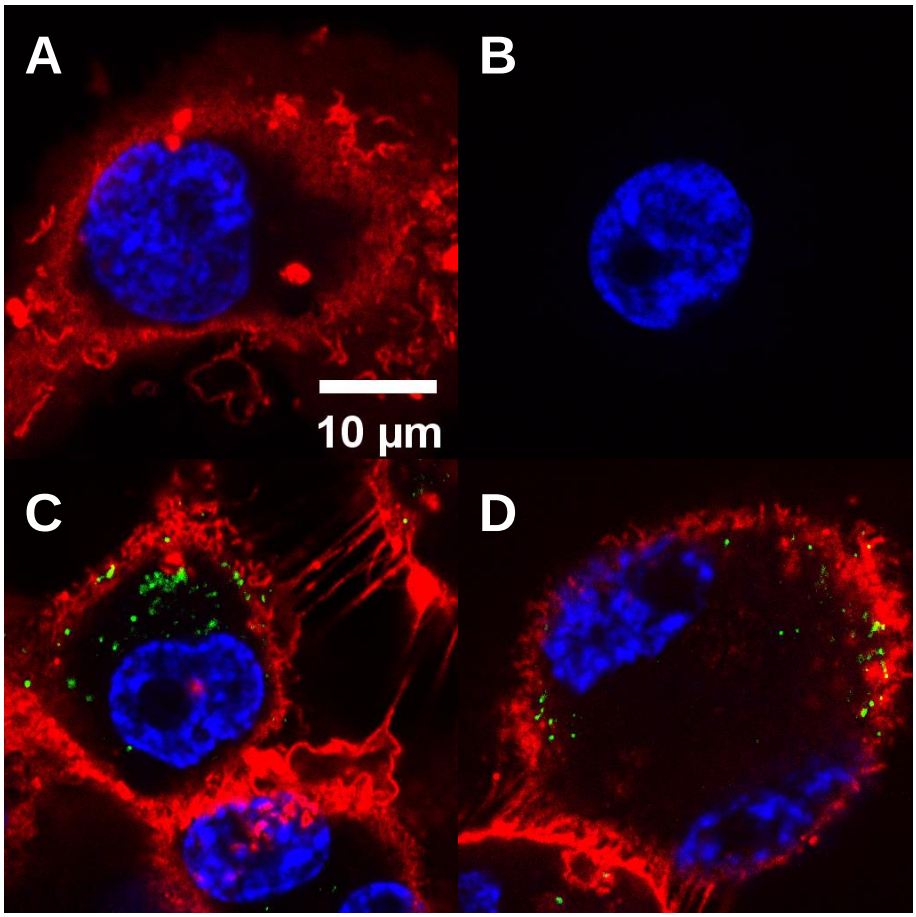


Figure 5.6: Microscopy images of fixed THP-1 derived macrophage treated with (A) TBS for 6 h; (B) TBS for 6 h (without antibody); (C) Qβ-F-fluo for 20 min or; (D) Qβ-CS-fluo for 20 min. Red: CD14 marker; Green: Qβ-F-fluo or Qβ-CS-fluo VLPs; Blue: DAPI. Refer to **Figure S.5** in Appendix for the full set of images.

To get a rough idea about the subcellular localisation of the particles, the fixed macrophage cells were labelled with antibodies against early endosomes and lysosomes. Some of the particles were found to be co-localised with early endosomes, while none of them were found in lysosomes (**Figure 5.7**). It is possible that the acidic environment of lysosome could attenuate the signal of fluorescein.[7] In addition, isotype control experiments were performed to prove that both antibodies were working fine (refer to **Figure S.6** in Appendix).

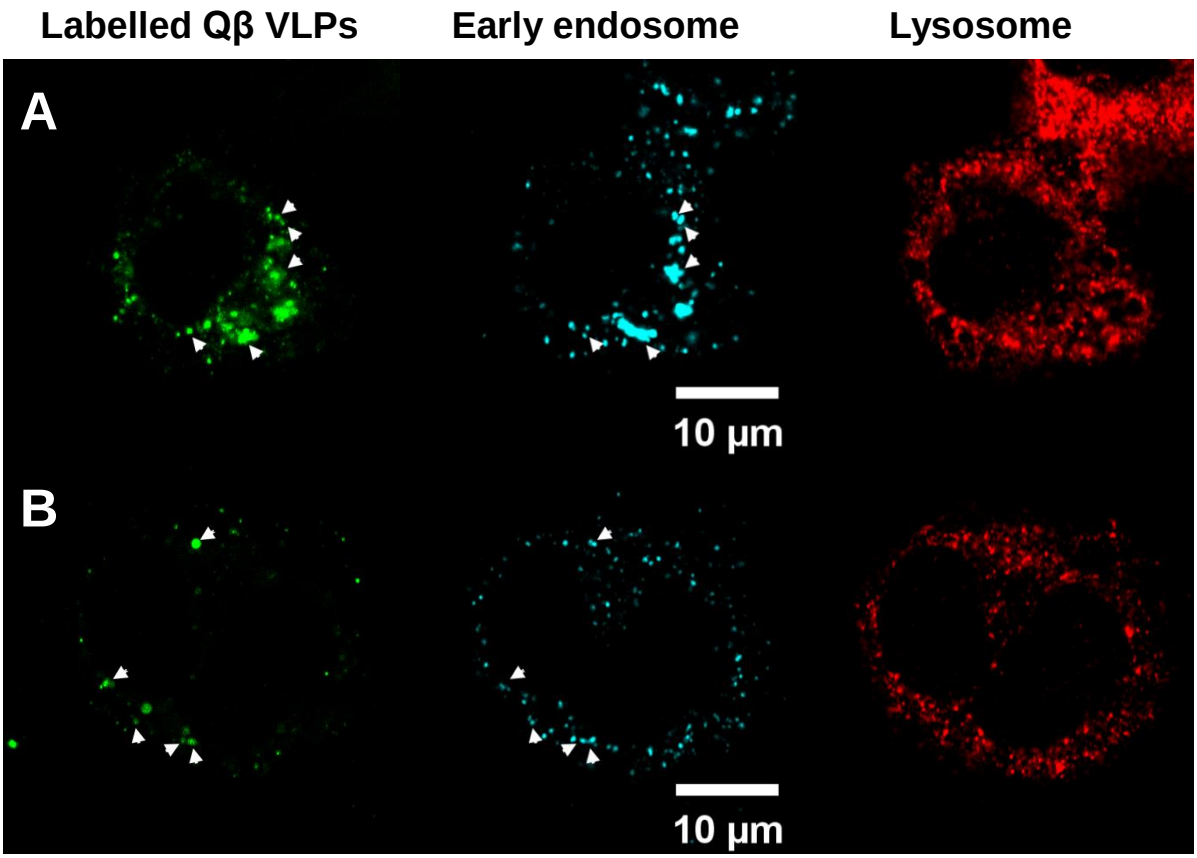
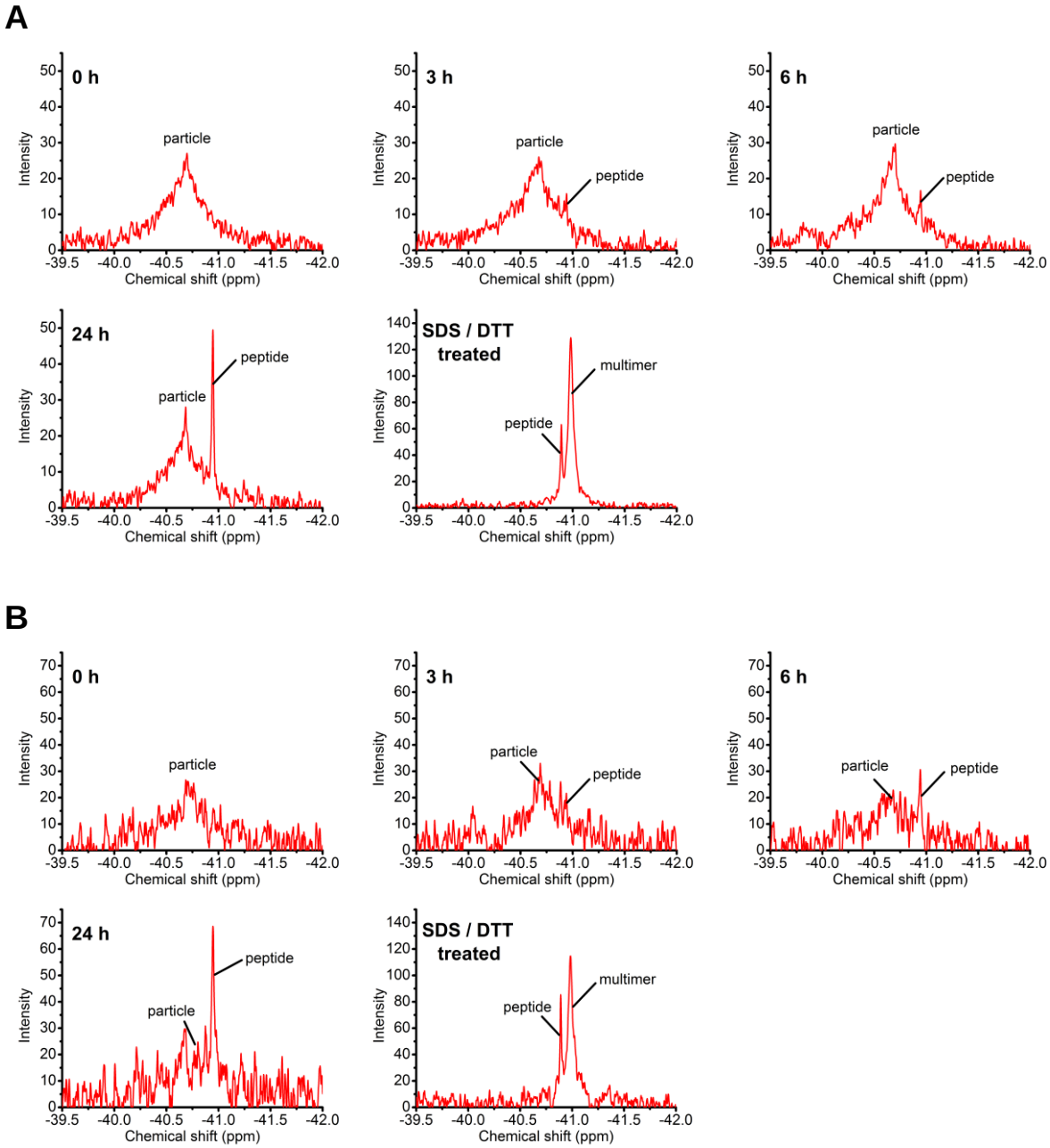


Figure 5.7: Microscopy images of fixed THP-1 derived macrophage treated with (A) Qβ-F-fluo for 20 min or; (B) Qβ-CS-fluo for 20 min. Green: Qβ-F-fluo or Qβ-CS-fluo VLPs; Teal: early endosome marker; Red: lysosome marker. The overlapping regions are indicated by white arrows. Refer to **Figure S.6** in Appendix for the full set of images.

5.2.3 ^{19}F NMR study of Qβ VLPs treated with cell lysate

To probe the more detailed behavior of VLPs in the milieu of the cell interior, both Qβ-F and Qβ-CS VLPs were incubated with cell lysates and analysed by ^{19}F NMR spectroscopy. In addition to the THP-1-derived macrophage that was selected from the cell screening in the previous part, *E. coli* lysate was also chosen for the study. Surprisingly, even the broad particle-associated peak proved to be sufficiently observable in both lysates (**Figure 5.8**), owing to the particular suitability of the pairing of ^{19}F NMR spectroscopy and our system. A small sharp peak with a chemical shift of – 40.95 ppm (FWHH = 10 Hz) was observed after 3 to 6 hours of incubation and can be clearly seen after 24 hours.


A
0 h
particle
3 h
particle
peptide
6 h
particle
peptide
24 h
peptide
particle
SDS / DTT
treated
multimer
peptide
B
0 h
particle
3 h
particle
peptide
6 h
particle
peptide
24 h
peptide
particle
SDS / DTT
treated
multimer
peptide
Intensity
Chemical shift (ppm)

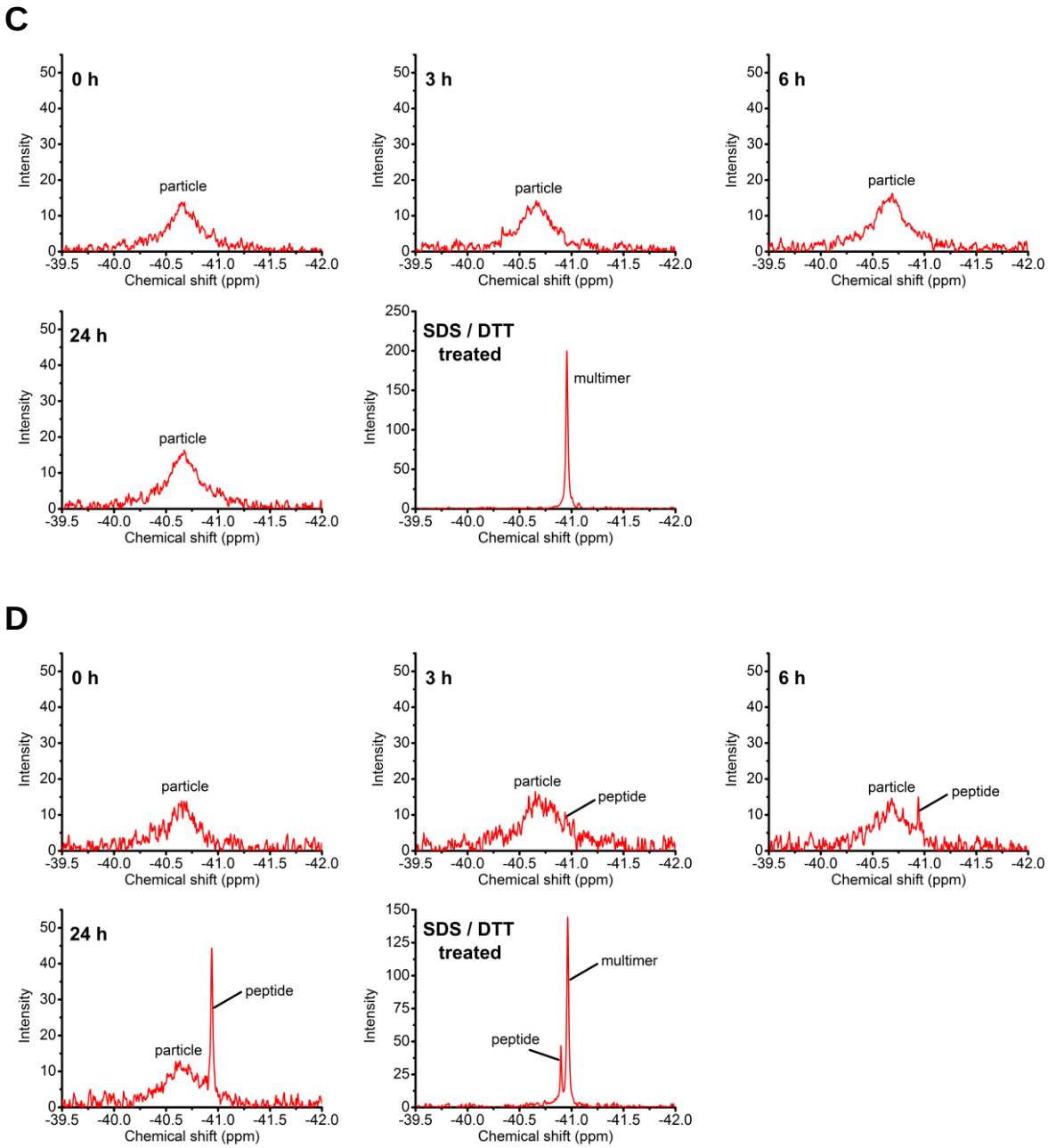
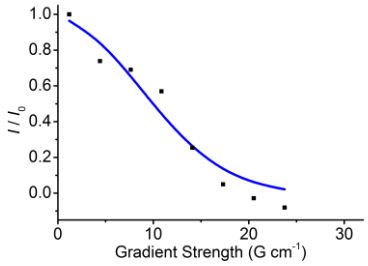
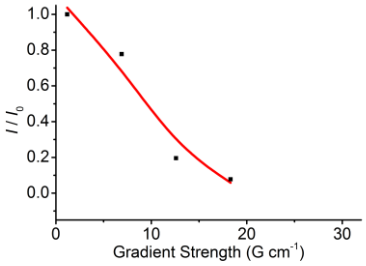
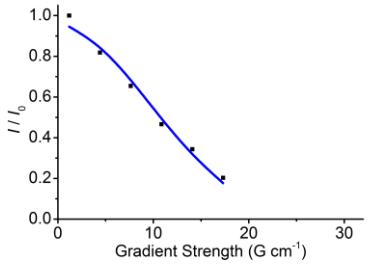
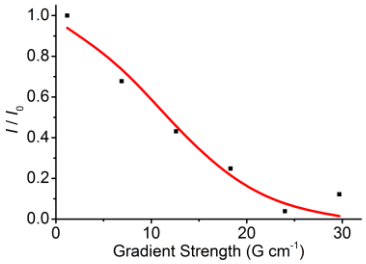
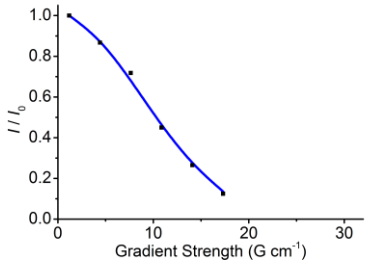
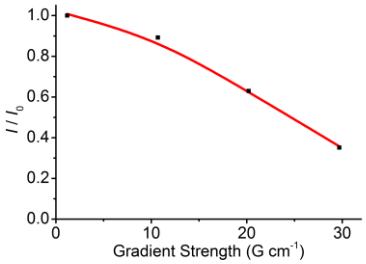


Figure 5.8: ^{19}F NMR spectra of (A) Qβ-F VLPs incubated with *E. coli* cell lysate; (B) Qβ-CS VLPs incubated with *E. coli* cell lysate; (C) Qβ-F VLPs incubated with THP-1 derived macrophage cell lysate and; (D) Qβ-CS VLPs incubated with THP-1 derived macrophage cell lysate at 37 ^{o}C for 0 h, 6 h, 24 h and followed by SDS+DTT treatment.

The sharp peak was originally thought to be coming from Qβ multimers as it had a similar chemical shift and peak width to the peak produced by SDS/DTT treatment. However, when the 24 h sample was treated with SDS and DTT, a new sharp peak arose (δ_F = – 40.98 ppm) accompanying the loss of the broad peak. The original sharp peak was found to be shifted downfield to – 40.90 ppm. The new sharp peak appeared after SDS+DTT treatment was most likely coming from Qβ multimers. This was proved by its similar diffusion coefficient (3.95 x 10^{-11} m^2 s^{-1}) to that of large multimers (**Table 5.1**). Theoretically, it should have a diffusion coefficient closer to that of small multimers (~ 5 x 10^{-11} m^2 s^{-1}) as DTT was present in the sample. The unexpected small value could be explained by the increased viscosity in the lysate solution. On the other hand, the peak appeared before SDS+DTT treatment gave a substantially large diffusion coefficient (2.98 x 10^{-10} m^2 s^{-1}). By using the equation (**1**) proposed by Young *et al*,[8] the species contributing to this should have a molecular mass of approximately 800 Da, which is likely to be the fluorine-containing peptides degraded from the Qβ particles by the cell lysate. The diffusion coefficient of the peptides was also found to be reduced by 1.8 times after SDS treatment (**Table 5.1**), which was consistent with the HEL results mentioned in the previous chapter (**Table 4.2**).

The lysate experiment revealed not only that the system developed can be applied to biological context, but also that it can distinguish between protein and peptide in complex media. Furthermore, it has been shown by a previous group member that the mechanism of VLP disassembly is not a cooperative catastrophic collapse but a gradual drift in population in response to environmental stress.[9] By monitoring the signal increase at different time points, half-life curves showing the degradation of particles in different types of lysates can be plotted (**Figure 5.9**).

Table 5.1: The I / I_0 versus gradient strength plots and diffusion coefficient values of various Qβ species and their corresponding references after lysate treatment.

	Diffusion coefficient (x 10^{-11} m^2 s^{-1})		
	TFAcetone	**Protein / peptide**	**Protein after normalisation** †
The sharp peak observed after lysate treatment	66.3 ± 3.1 	28.9 ± 3.6 	29.8 ± 1.9
The sharp peak at -40.90 ppm after further SDS & DTT treatment	59.2 ± 1.5 	15.0 ± 0.9 	16.6 ± 1.1
The sharp peak at -40.98 ppm after further SDS & DTT treatment	62.8 ± 2.1 	3.77 ± 0.02 	3.95 ± 0.11

† The diffusion coefficient of protein is normalised based on a single diffusion coefficient of TFAcetone (6.6 x 10^{-10} m^2 s^{-1}) to eliminate any variations in viscosity.

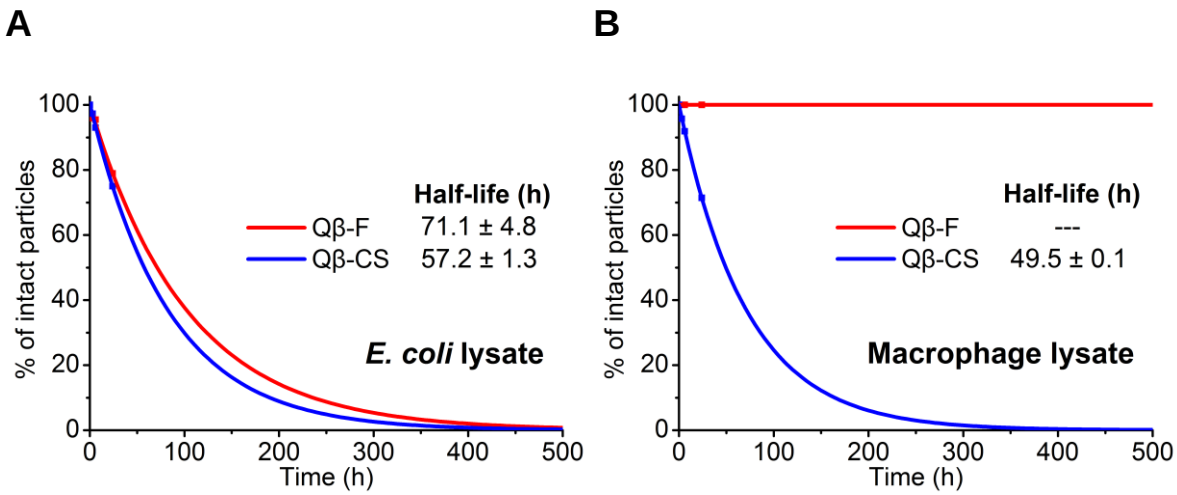


Figure 5.9: Decay curves of Qβ-F & Qβ-CS VLPs in (A) *E. coli* or; (B) THP-1 derived macrophage cell lysates.

In turn, this system allowed us to analyse the designed 'flimsy' VLP system Qβ-CS. Comparison of the degradation processes of Qβ-F and Qβ-CS revealed a significantly shorter half-life for Qβ-CS in both lysate systems, consistent with *in vitro* data and with the designed removal of disulfide bonds mediated by Cys-74 and Cys-80. Though the disulphide bonds are not necessary for particle assembly, they make a significant contribution on the stability of particles. Interestingly, whereas both Qβ-F and Qβ-CS were degraded by *E.coli* lysate, the more stable Qβ-F was not significantly degraded by the macrophage lysate.

5.2.4 Introduction of mannose targeting groups on Qβ VLPs

Once the system was proven to work well in cell lysate, the next step was to apply it to real cells. However, as mentioned at the beginning of this chapter, the biggest challenge for this would be to make sure that the cells can take up enough particles to give a detectable signal. Introduction of targeting groups on the VLP surface is one of the possible ways for increasing the uptake efficiency. In here, only mammalian cells were considered as the potential targets. Although *E. coli* is the natural host of Qβ bacteriophage, the virus itself

never enters to the host cell but only inject viral RNA to the cell *via* a small hole created by the virus.[5, 6]

Carbohydrates have been widely employed as targeting agents due to their critical role in cellular trafficking.[10] D-Mannose was known to target both mannose receptors (MR) and dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin (DC-SIGN) on the membrane of macrophages and dendritic cells respectively.[11-14] By incorporating mannose on the surface of VLPs, it was believed that the uptake efficiency by macrophage could be further increased (**Figure 5.10**).

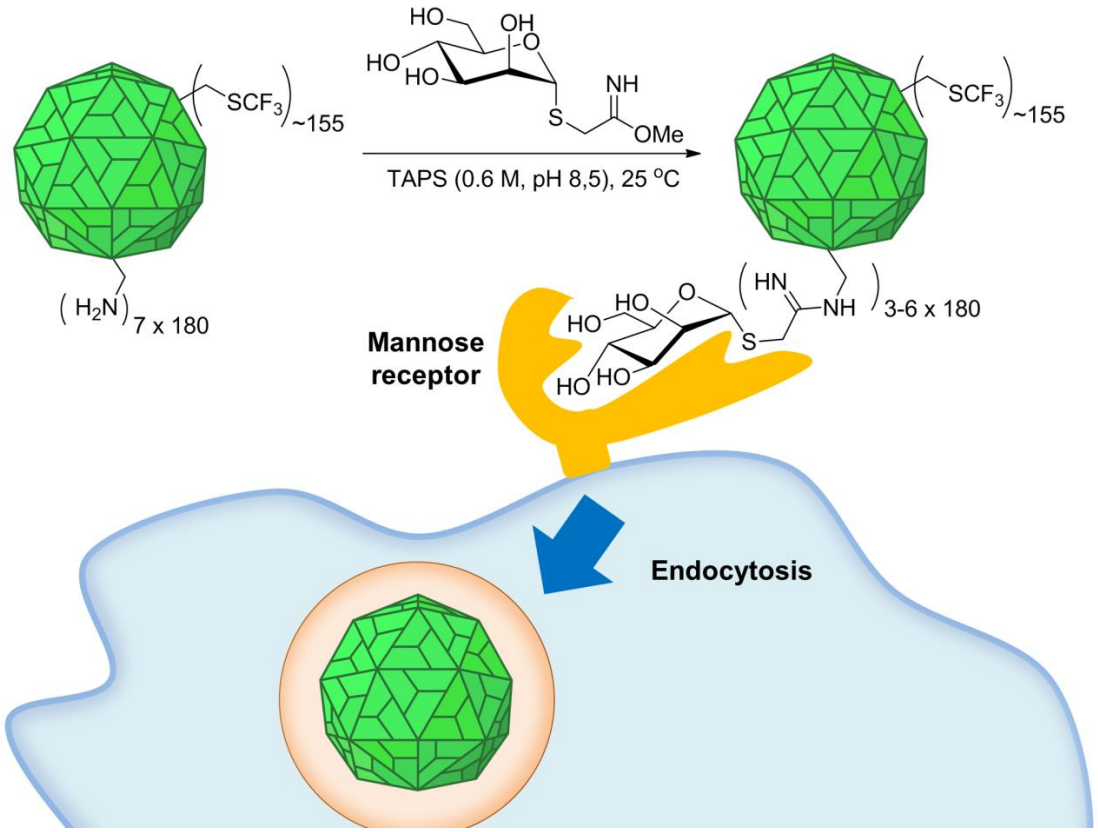
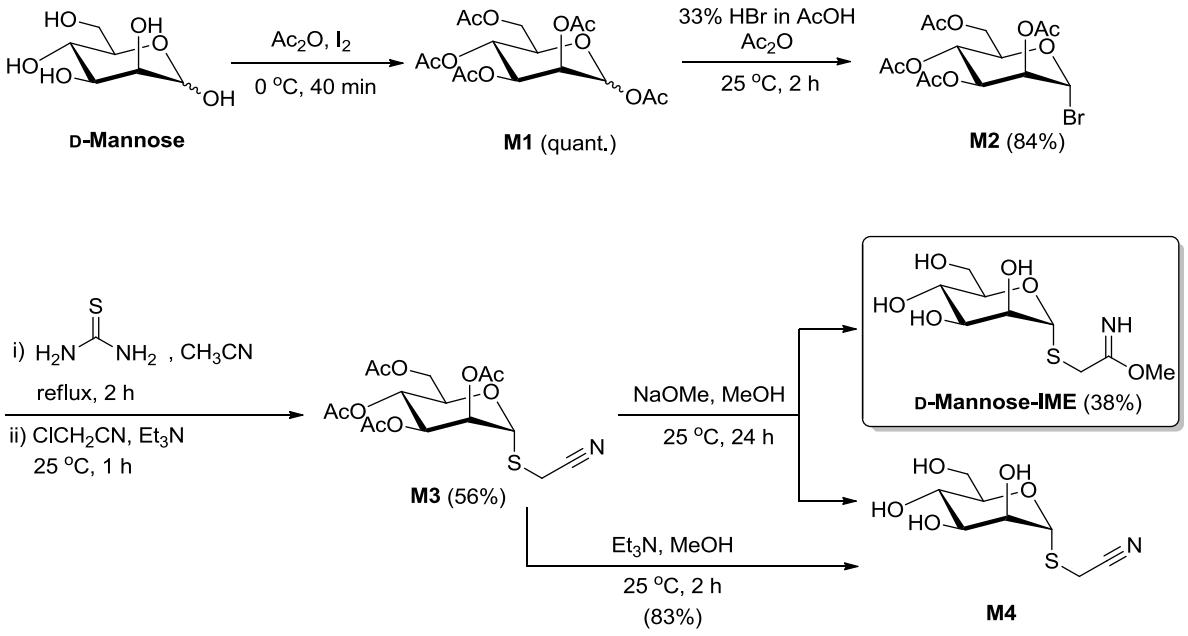


Figure 5.10: Schematic diagram showing the uptake of mannose-decorated Qβ-F VLPs by cells with mannose receptors on the surface. It was believed that decorating Qβ-F with D-mannose-IME could improve the uptake efficiency of macrophage.

To label the Qβ VLPs with D-mannose, a coupling group should first be introduced to the sugar. In previous study, D-mannose has been successfully linked to the lysine residues on the surface of adenovirus with the 2-imino-2-methoxyethyl (IME) group.[14] This coupling group is advantageous in minimising the change of pI on the labelled protein (refer to Chapter 6 for more details).[15]

D-Mannose-IME can be readily synthesised by following the previous published protocols (**Scheme 5.1**).[14, 16] When the compound **M3** was synthesised, it was further activated to form D-mannose-IME by adding sodium methoxide in methanol. This reaction formed an equilibrium mixture of D-mannose-IME and **M4**. In order to calculate the ratio of D-mannose-IME and **M4** by NMR spectroscopy, pure **M4** was also synthesised to act as a reference. From the NMR spectrum, the ratio of D-mannose-IME and **M4** was around 1:1, which is consistent with the literature.[15]



Scheme 5.1: Synthetic route of mannose-IME from mannose

Qβ-F VLPs were labelled with D-mannose-IME in TAPS buffer with an alkaline pH (**Figure 5.10**). The mass shift on SDS-PAGE gel (**Figure 5.11**) showed that the reaction was successful. LC-MS (**Figure 5.12A**) further confirmed that up to six lysine residues were coupled per monomer after 4 h of reaction. To study the internalisation of D-mannose-labelled Qβ-F (Qβ-F-man) VLPs in different cells, the fluorescein-labelled Qβ-F-fluo was also coupled with D-mannose-IME (**Figure 5.12B**).

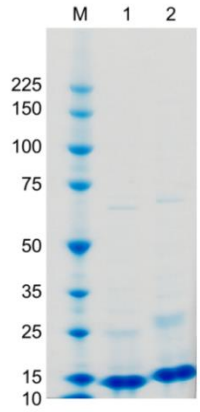
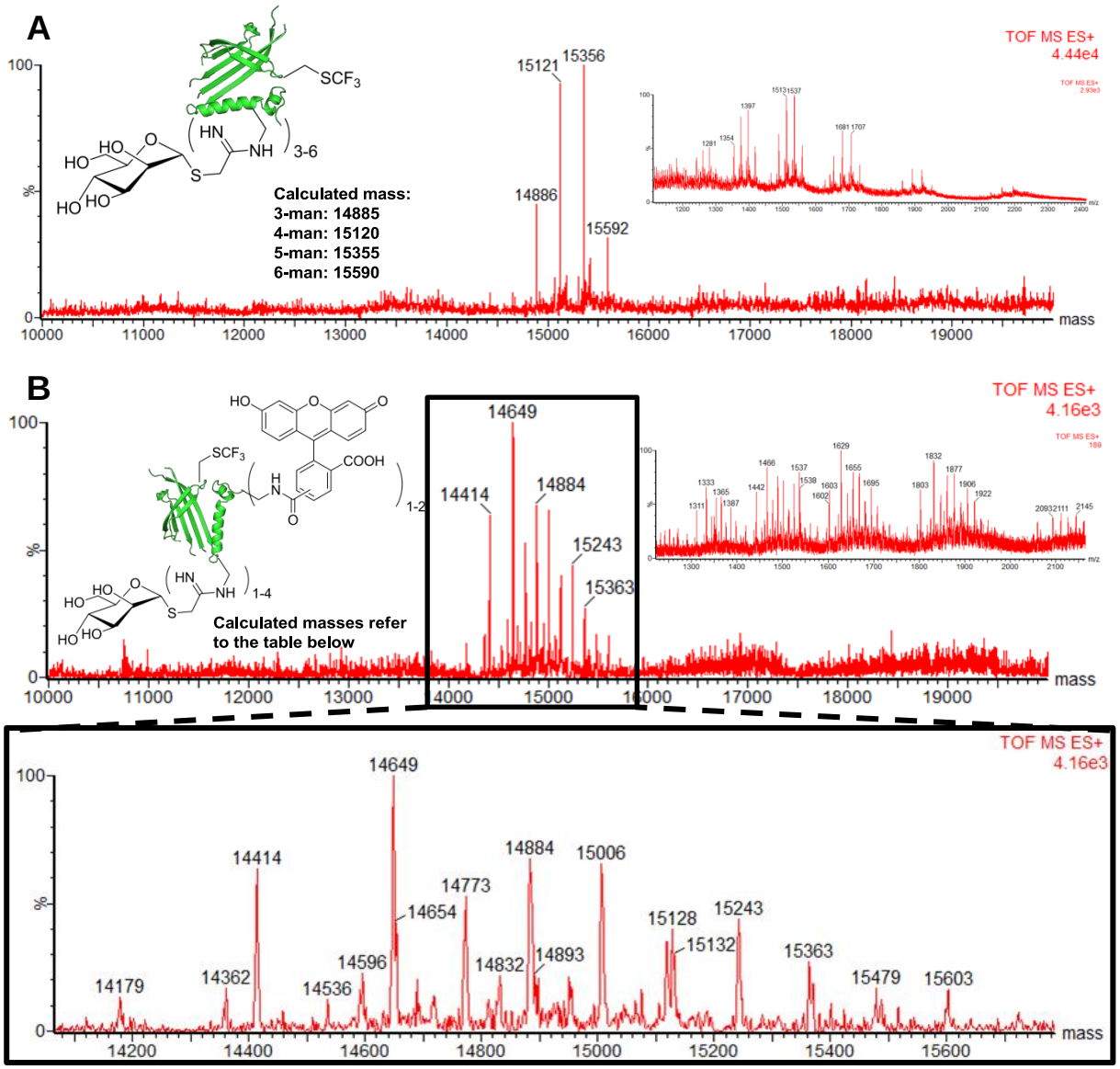


Figure 5.11: Coomassie Blue stained reducing SDS-PAGE gel showing Qβ-F before (lane 1) and after (lane 2) mannose modification. M: protein marker.

The uptake efficiency of the dual-labelled Qβ-F (namely Qβ-F-fluo-man) VLPs by different mammalian cells was studied using flow cytometry (**Figure 5.13**). Unfortunately, it was found that the macrophage internalisation ability did not improve when compared with Qβ-F-fluo VLPs. Surprisingly, the percentage of fluorescein-positive THP-1 monocytes increased significantly after 6 and 24 h of incubation. However the reasons for this are not understood as monocytes should not possess any DC-SIGN or MR. In addition, the microscopy experiments did not exhibit any fluorescent signal (**Figure 5.14**). It was likely that only tiny amounts of particles were engulfed by an unknown mechanism. Alternatively, the surface of THP-1 may have an unknown receptor that could interact strongly with the D-mannose-labelled particles. In conclusion, the mannose labelling approach did not work as well as expected and additional targeting groups will need to be studied in the future.



No. of fluo	No. of man	Calc. mass
0	0	14180
0	1	14415
1	0	14538
0	2	14650
1	1	14773
0	3	14885
2	0	14896

No. of fluo	No. of man	Calc. mass
1	2	15008
0	4	15120
2	1	15131
1	3	15243
2	2	15366
1	4	15478
2	3	15601

Figure 5.12: Deconvoluted mass spectra and corresponding raw ion series of (A) Qβ-F or; (B) Qβ-F-fluo labelled with D-mannose-IME (800 equiv.). The VLPs were broken down to monomers for LC-MS analysis.

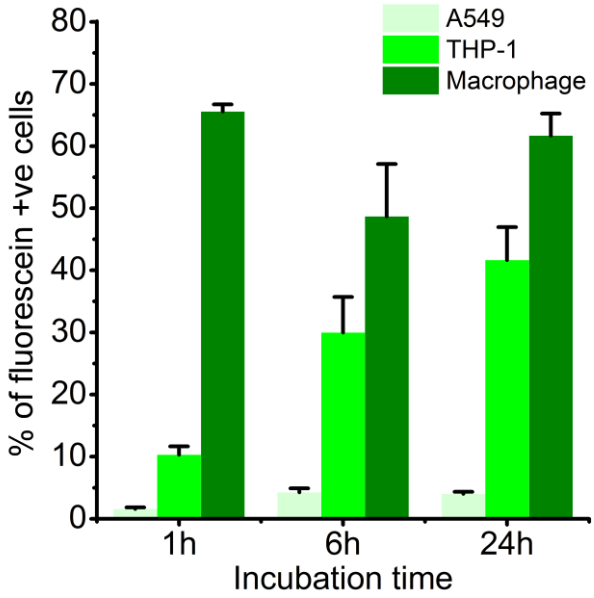
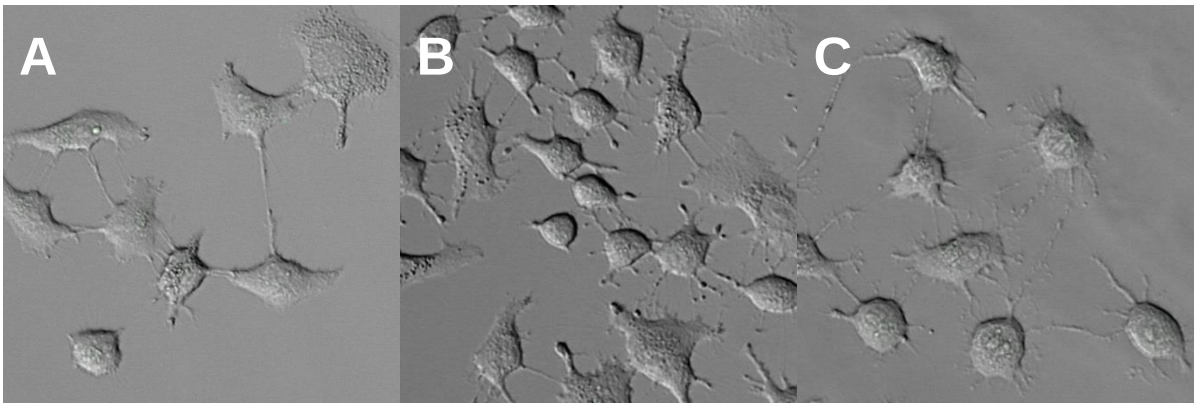
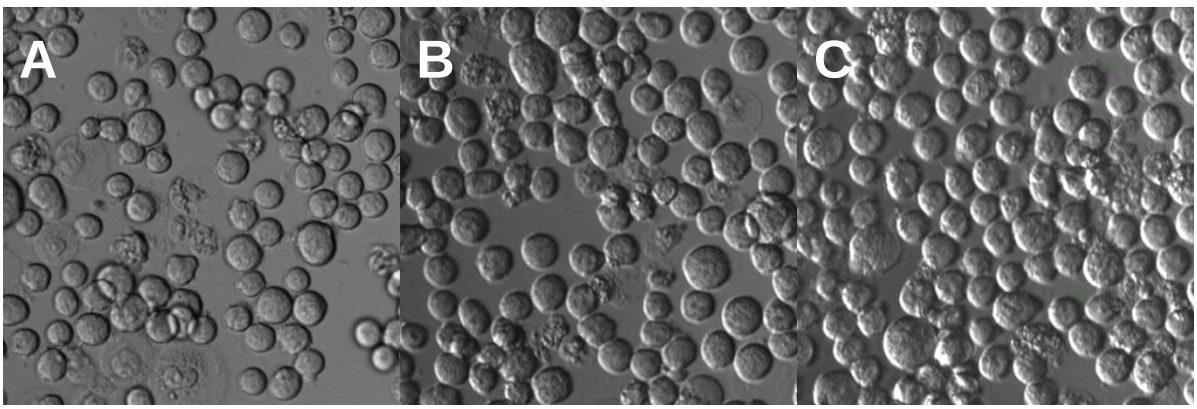


Figure 5.13: The uptake of Qβ-F-fluo-man VLPs by mammalian cells after different time points was studied by flow cytometry. The graph was plotted based on the raw histograms (refer to **Figure S.4** in Appendix).

A549 epithelial cells



THP-1 monocytes



THP-1 derived macrophages

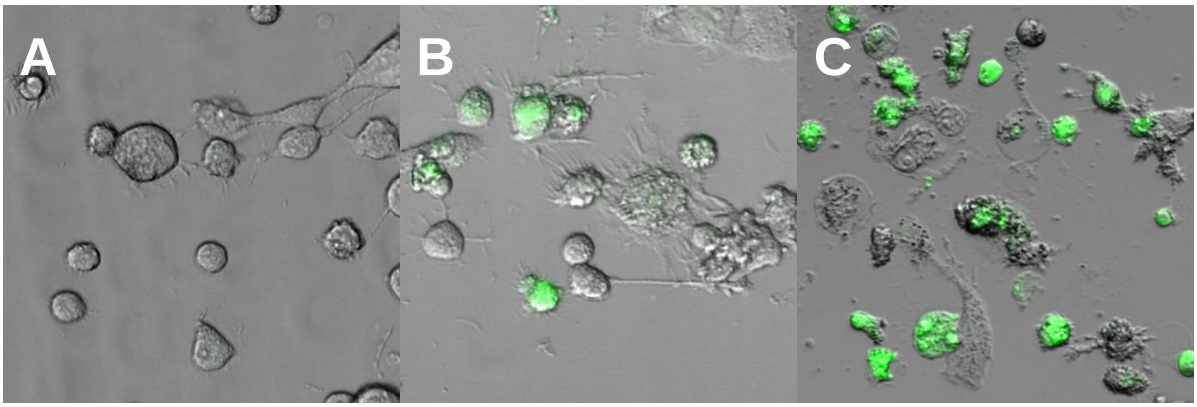


Figure 5.14: Microscopy images of mammalian cells treated with Qβ-F-fluo-man for (A) 1 h; (B) 6 h or; (C) 24 h. Green: Qβ-F-fluo VLPs.

5.2.5 In-cell ^{19}F NMR study of Qβ VLPs

Although the mannose targeting approach failed to work, it was still possible that the unlabelled Qβ VLPs could be internalised efficiently by the THP-1 derived macrophages. To analyse the VLP disassembly in real cells by ^{19}F NMR spectroscopy, the VLPs were incubated with the cells for 1 – 2 days. The mammalian cells were then collected and washed with buffer several times to remove all the particles outside the cells. As shown by the ^{19}F NMR spectra, neither the broad and sharp peaks were not observed for all the three Qβ species tested (**Table 5.2**).

Table 5.2: ^{19}F NMR study of cells incubated with different Qβ species

Sample	Incubation time	Intact particles observed?	Multimer / peptide observed?	Fluoride observed?
THP-1 only	1 d / 2 d	---	---	No
THP-1 + Qβ-F-man	1 d / 2 d	No	No	No
Macrophage only	1 d / 2 d	---	---	No
Macrophage + Qβ-F	1 d / 2 d	No	No	No
Macrophage + Qβ-CS	1 d	No	No	1 d

† the peak was observed after heating with 0.2 M SDS and 50 mM DTT

Interestingly, a fluoride peak with a chemical shift of around – 119.6 ppm was observed in the Qβ-CS samples (**Figure 5.15)**. It is possible that the particles underwent a different degradation mechanism inside the cells to produce a fluoride peak. To confirm the source of the peak, several control experiments were performed (**Table 5.3**). It was found that the fluoride peak can indeed be observed when the cell medium and TBS buffer were mixed and incubated for 5 days. Also, the presence of particles seemed to speed up the release of the fluoride. This phenomenon was not only observed in fluorinated VLPs, but also in non-fluorinated Qβ K16M VLPs. Although the exact mechanism for causing the fluoride peak was still unknown, the peak was certainly not coming from the degraded particles.

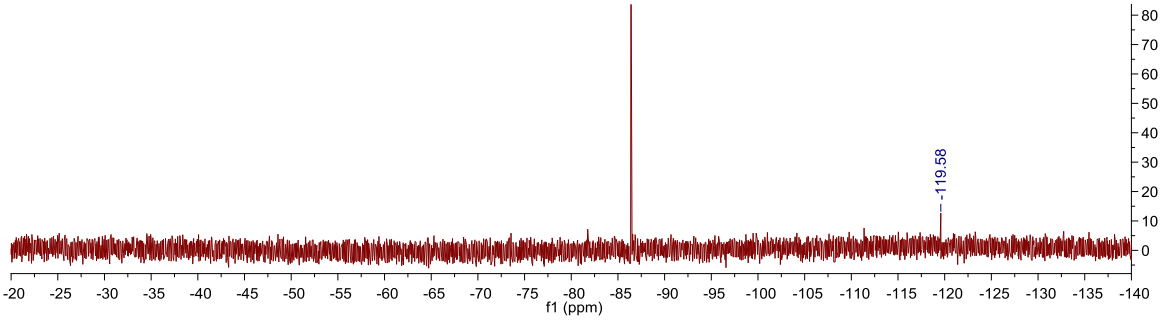


Figure 5.15:^{19}F NMR spectrum of Qβ-CS mixed with THP-1 derived macrophage in medium after 1 day followed by washing. TFAcetone was added to the sample as an internal reference. The reference peak was fixed to – 86.40 ppm relative to $CFCl_3$.

Table 5.3: Control experiments for identifying the origin of the fluoride peak

Sample [#]	Incubation time	Intact particles observed?	Multimer / peptide observed?	Fluoride observed?
THP-1 only	1 h – 3 d	---	---	No
Macrophage only	1 h – 3 d	---	---	No
Medium only	1 d – 5 d	---	---	No
TBS only	1 d – 5 d	---	---	No
Medium + TBS	1 d – 5 d	---	---	After 5 d
Medium + Qβ-F	1 h – 4 d	Yes	SDS / DTT †	After 2 d
Medium + Qβ K16M	3 h – 2 d	---	---	After 1 d

[#] all the samples were analysed directly without washing

† the peak was observed after heating with 0.2 M SDS and 50 mM DTT

5.3 Conclusion

In summary, the ^{19}F NMR detection system for Qβ VLPs was successfully applied to the complicated bacterial and mammalian cell lysates. Notably, fluorine-containing peptides degraded from the particles by internal machinery of the cell lysate can be discriminated from both the particle-associated broad peak and the multimer sharp peak produced by chemical treatment. The ease of the NMR technique also allows time-point monitoring of the particle disassembly in different types cell lysates. This not just provides a convenient way for comparing the 'digesting' ability of different lysate system, but can also act as a simple tool for assessing the stability of different VLP species. This is particularly crucial for drug nano-carriers as the efficiency of drug release can greatly affect the drug bioavailability and hence the dose required. Notably, it has also been suggested that antibacterial phage strategies may benefit from novel mutant phages,[17] including designed variants such as the Qβ-CS species reported here. It is thus believed that the established novel system has a great potential in the future.

To apply this system to real cells or even the whole body, more effort is obviously required to improve the detection signal. This can either be achieved by increasing the uptake efficiency of the target cells or by improving the detection system itself, such as adding more fluorine atoms on the particle surface. Until today, only the internal proteins over-expressed by the cell itself can be effectively studied by in-cell NMR spectroscopy.[2, 3, 18-21] In these cases, a sufficient amount of proteins is present in the cell to give a strong detectable signal. It will be therefore a great break-through if in-cell NMR spectroscopy can be achieved by applying external proteins. Once an NMR detection system is fully developed, it can possibly be applied to MRI for *in vivo* study.

5.4 Experimental Section

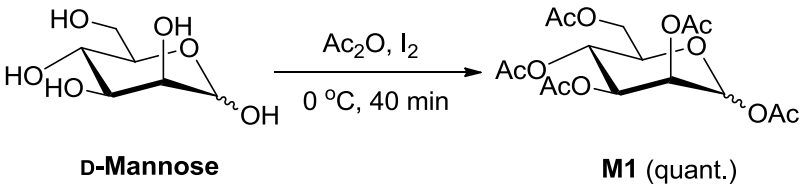
5.4.1 General procedures

Chemicals were purchased from commercial suppliers and used without purification. NuPAGE® MES buffer and 12% Bis-Tris gel were purchased from Life Technologies™. Protein markers used were Perfect Protein™ Marker 10-225 kDa purchased from Novagen® . PD SpinTrap™ G-25, MiniTrap™ G-25, MidiTrap™ G-25 and PD10 desalting columns were purchased from GE Healthcare Life Sciences. Micro Bio-Spin™ P-6 columns were purchased from Bio-Rad. Antibodies were purchased from Abcam® and Biolegend® .

Small molecule and protein mass spectrometry analysis was performed in Bruker μTOF and Waters LCT Premier™ respectively. Chemical NMR study was performed in Bruker AVIII400 HD Nanobay instrument running TOPSPIN 3 equipped with a 5 mm z-gradient multinuclear BBFO probe. Protein NMR study was performed in Bruker AVIII 600 equipped with a dedicated 5mm BB-F/1H Prodigy N2 cryoprobe. Infrared spectrometry was performed in Bruker Tensor 27. Thin layer chromatography (TLC) was performed using aluminium sheets coated with 60F254 silica gel and visualised using ninhydrin solution. Microscopy analysis was performed in Leica SP8 STED and ZEISS LSM880 microscopy. Flow cytometry was performed in BD FACSCalibur™.

5.4.2 Synthesis of D-mannose-IME

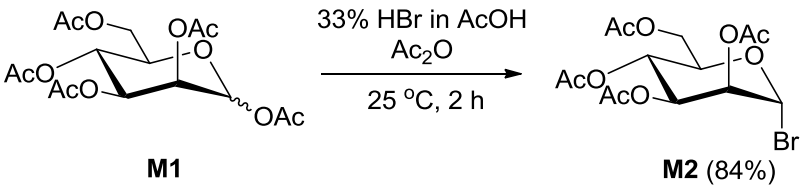
5.4.2.1 Penta-*O*-acetyl-α-D-mannopyranose and penta-*O*-acetyl-β-D-mannopyranose (M1)



D-Mannose **M1** (quant.)

Iodine (0.112 g, 0.43 mmol) was added to a stirred solution of D-Mannose (2.0 g, 11.1 mmol) in acetic anhydride (10 mL, 110 mmol) at 0 °C. The reaction mixture was stirred at 0 °C for 40 min and then warmed to 25 °C. The mixture was then diluted with ethyl acetate (50 mL) and washed with ice cold sodium thiosulphate (2 x 25 mL). The upper organic layer was washed with sat. $NaHCO_3$ (4 x 25 mL) and dried with magnesium sulphate. The solution was evaporated to dryness under vacuum to give compound **M1** as a yellow oil, which was used for the next step without further purification; R_f = 0.63 (1:1 PE/EtOAc, v/v); ^{1}H NMR (250 MHz, $CDCl_3$, α-anomer) δ_H = 2.00 (s, 3H, CH_3), 2.05 (s, 3H, CH_3), 2.09 (s, 3H, CH_3), 2.16 (s, 3H, CH_3), 2.17 (s, 3H, CH_3), 3.97-4.18 (stack, 2H, H_5, H_6), 4.28 (dd, J 4.90, 12.38, 1H , H_6), 5.23-5.37 (stack, 3H, H_2, H_3, H_4), 6.08 (d, J 1.81, 1H, H_1); ^{1}H NMR (250 MHz, $CDCl_3$, β-anomer) δ_H = 2.10 (s, 3H, CH_3), 3.80 (ddd, J 2.35, 5.29, 9.78, 1H, H_5), 5.12 (dd, J 3.27, 9.99, 1H, H_3), 5.48 (dd, J 1.06, 3.24, 1H, H_2), 5.85 (d, J 1.12, 1H, H_1). The data agreed with the reported values in the literature.[16]

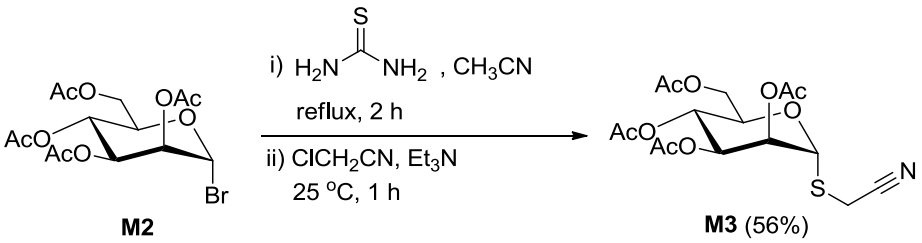
5.4.2.2 Tetra-*O*-acetyl-α-D-mannopyranosyl bromide (M2)



M1 **M2** (84%)

33% HBr in acetic acid (10.8 mL, 66 mmol HBr) was added to a solution of compound **M1** (4.38 g, 11.2 mmol) in acetic anhydride (1.4 mL) and stirred at 25 ^{o}C for 2 h. The reaction mixture was then diluted with ethyl acetate (100 mL) and washed with ice cold water (50 mL), sat. $NaHCO_3$ (50 mL) and brine (50 mL). The crude product was dried over magnesium sulphate and evaporated under vacuum to give compound **M2** as a yellow syrup (3.86 g, 84%); R_f = 0.55 (7:3 PE/EtOAc, v/v); $[\alpha]_D^{22}$ = +128 ($CHCl_3$, C = 1.0); ^{1}H NMR (200 MHz, $CDCl_3$) δ_H = 1.99 (s, 3H, CH_3), 2.06 (s, 3H, CH_3), 2.09 (s, 3H, CH_3), 2.16 (s, 3H, CH_3), 4.10-4.17 (m, 1H, H_6), 4.17-4.26 (m, 1H, H_5), 4.32 (dd, J 4.59, 11.96, 1H, H_6), 5.35 (t, J 9.98, 1H, H_4), 5.43 (dd, J 1.62, 3.41, 1H, H_2), 5.70 (dd, J 3.42, 10.17, 1H, H_3), 6.28 (d, J 1.28, 1H, H_1). The data agreed with the reported values in the literature.[16]

5.4.2.3 Cyanomethyl tetra-*O*-acetyl-α-D-thiomannopyranoside (M3)

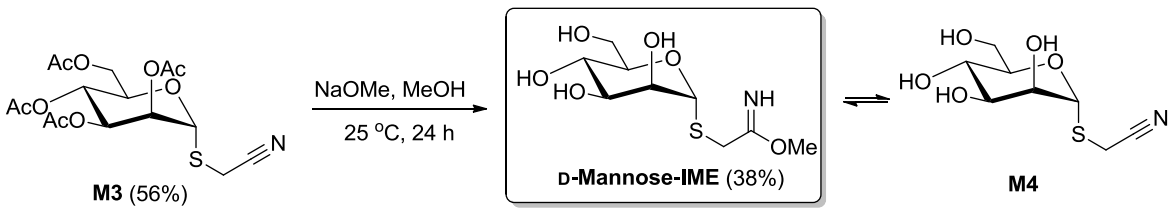


Compound **M2** (3.80 g, 9.2 mmol) and thiourea (0.74 g, 9.7 mmol) were dissolved in anhydrous acetonitrile (18.5 mL) at 25 ^{o}C. The solution was heated to reflux for 2 h and then cooled down to 25 ^{o}C. Triethylamine (3.2 mL, 23.1 mmol) was added to the reaction mixture and stirred at 25 ^{o}C for another 20 min. The solvent of the reaction mixture was removed under vacuum. The residue was dissolved in ethyl acetate (100 mL) and washed with water (50 mL), 1 N HCl (50 mL) and brine (50 mL). The crude product was dried over magnesium sulphate and evaporated under vacuum to give a yellow powder. The powder was recrystallized from hot ethanol to afford **M3** as a white powder (2.09 g, 56%); R_f = 0.49 (2:8 PE/EtOAc, v/v); M.p = 130-131 ^{o}C; $[\alpha]_D^{22}$ = +80 ($CHCl_3$, C = 1.0); ^{1}H NMR (400 MHz,

$CDCl_3$) δ_H = 1.97 (s, 3H, CH_3), 2.04 (s, 3H, CH_3), 2.08 (s, 3H, CH_3), 2.16 (s, 3H, CH_3), 3.31 (d, J 17.18, 1H, SCH_2C), 3.46 (d, J 17.18, 1H, SCH_2C), 4.13 (m, 1H, H_6), 4.31 (m, 2H, H_5, H_6), 5.16 (dd, J 3.47, 10.00, 1H, H_3), 5.33 (m, 2H, H_2, H_4), 5.44 (d, J 1.29, 1H, H_1); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C = 15.81 (S**CH$_2$**CN), 20.63 (OC(O)**CH$_3$**), 20.71 (OC(O)**CH$_3$**), 20.78 (OC(O)**CH$_3$**), 20.84 (OC(O)**CH$_3$**), 62.06 (C_6), 65.87 (C_4), 69.33 (C_3), 69.64 (C_2), 70.00 (C_5), 82.48 (C_1), 115.78 (SCH_2**CN**), 169.62 (O**C**(O)CH_3), 169.77 (O**C**(O)CH_3), 169.86 (O**C**(O)CH_3), 170.58 (O**C**(O)CH_3); LRMS m/z (ESI^+): measured: 404.1 ($M + H^+$), calculated: 404.1; IR: 2970, 2359, 1740, 1406, 1366, 1325, 1249, 1218, 1205, 1130, 1117, 1101, 1068, 1048, 1015 cm^{-1}. The data agreed with the reported values in the literature.[14]

5.4.2.4 2-Imino-2-methoxyethyl-α-D-thiomannopyranoside (D-mannose-IME)

5.4.2.4.1 For chemical characterisation



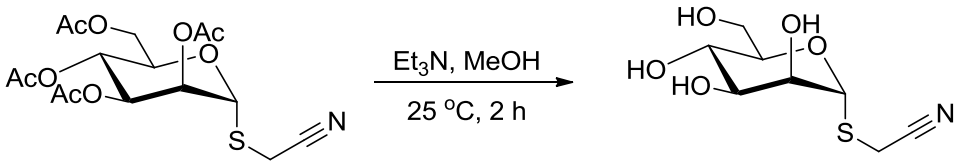
Compound **M3** (0.80 g, 2.0 mmol) was dissolved in anhydrous methanol (20 mL) with a methanolic solution of 25% sodium methoxide (0.18 mL, 0.80 mmol NaOMe). The reaction mixture was stirred at 25 °C for 24 h. The reaction mixture was evaporated by blowing nitrogen to a final volume of 5 mL. A few drops of acetic acid were added to the solution until the pH reached 7 on wet pH paper. The product was crushed out by adding diethyl ether (80 mL), and then centrifuged down at 7000 x g for 5 min. The pellet was resuspended with diethyl ether (20 mL) and centrifuged at 7000 x g for 5 min. The pellet was dried under vacuum to give a mixture of D-mannose-IME and **M4** as a white powder (0.35 g, 71%; 1.18:1 D-mannose-IME/**M4**, mol/mol); R_f of D-mannose-IME = 0.24; R_f of **M4** = 0.51 (8:2 EtOAc/MeOH, v/v); ^{1}H NMR (400 MHz, MeOD) δ_H = 3.61-3.64 (stack, 2H, H_3, H_4), 3.70

(m, 1H, H_6), 3.83-3.76 (stack, 2H, H_5, H_6), 3.90 (m, 1H, H_2), 5.16 (s, 1H, H_1); ^{13}C NMR (100 MHz, MeOD) δ_C = 62.63 (C_6), 68.70 (C_4), 73.07 (C_2), 73.15 (C_3), 73.51 (C_5), 86.07 (C_1); LRMS m/z (ESI$^+$) of D-mannose-IME: measured: 268.1 (M + H$^+$), calculated: 268.1; LRMS m/z (ESI$^+$) of **M4**: measured: 236.1 (M + H$^+$), calculated: 236.1.

5.4.2.4.2 For protein reaction

Compound **M3** was added to a vial or RB flask, dried under vacuum and filled with argon. The powder was dissolved in anhydrous methanol (20 mg / mL) followed by the addition of 25% sodium methoxide in methanol (0.4 equiv.). The reaction mixture was stirred at 25 °C for 24 h. The reaction was dried under nitrogen or by vacuum. The residue was dissolved in TAPS (0.6 M, pH 8.5) to prepare a stock solution.

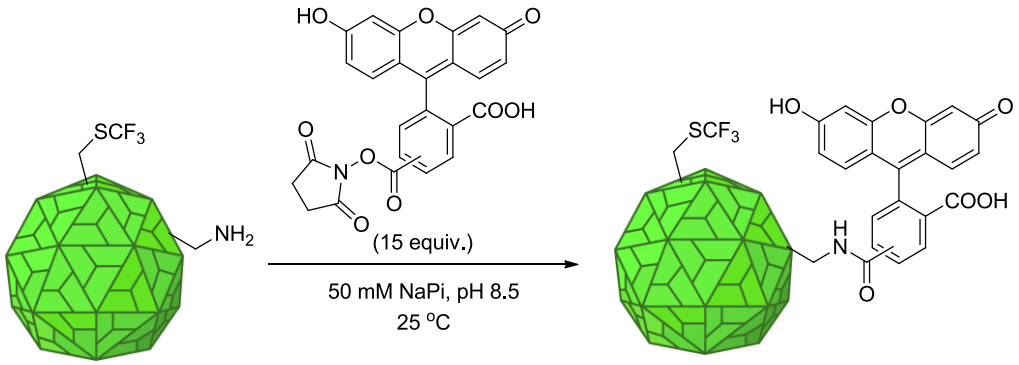
5.4.2.5 Cyanomethyl α-D-thiomannopyranoside (M4)



M3 **M4** (83%)

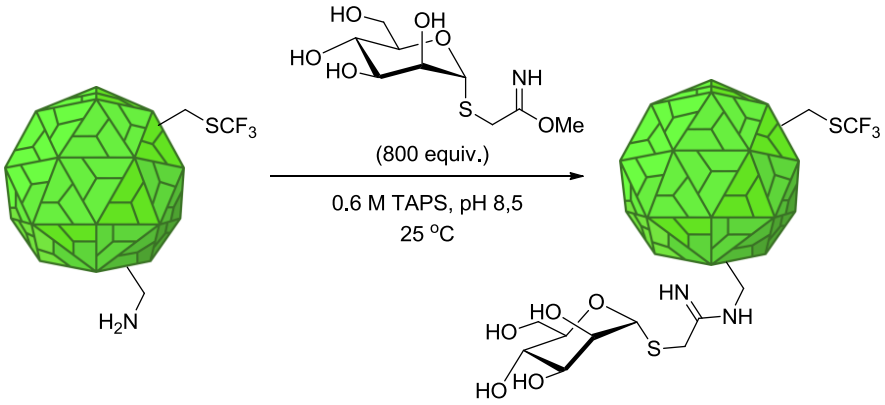
To a suspension of compound **M3** (0.2 g, 0.50 mmol) in methanol (3.9 mL), triethylamine (0.39 mL, 2.8 mmol) was added and stirred at 25 °C for 2 h. The solvent was removed under vacuum. The crude product was then co-evaporated with toluene (3 mL x 3) to give compound **M4** as a pale yellow powder (0.097 g, 83%); R_f = 0.44 (1:1 PE/EtOAc, v/v); ^{1}H NMR (400 MHz, MeOD) δ_H = 3.59 (dd, J 3.43, 9.32, 1H, H_3), 3.68 (t, J 8.44, 1H, H_4), 3.72 (dd, J 4.82, 10.63, 1H, H_6), 3.80 (m, 1H, H_5), 3.86 (dd, J 2.04, 11.63, 1H, H_6), 3.92 (dd, J 1.39, 3.41, 1H, H_2), 5.42 (d, J 0.88, 1H, H_1); ^{13}C NMR (100 MHz, MeOD) δ_C = 62.48 (C_6), 68.52 (C_4), 72.72 (C_2), 73.22 (C_3), 75.78 (C_5), 86.04 (C_1); LRMS m/z (ESI$^+$) of **M4**: measured: 236.1 (M + H$^+$), calculated: 236.1.

5.4.3 Labelling of Qβ VLPs with fluorescein-NHS



Qβ VLPs (1 mg / mL) were buffer exchanged using Micro Bio-Spin® P6, G25 minitrap / miditrap / PD10 columns into sodium phosphate buffer (50 mM, pH 8.5). Fluorescein-NHS (100 mg / mL in DMF) (15 equiv.) was added to the protein solution. The reaction mixture was shaken at 25 °C for 1 h in dark, followed by purification using Micro Bio-Spin® P6 or G25 minitrap / miditrap / PD10 columns depending on the scale. The buffer used for elution varied based on the following steps.

5.4.4 Labelling of Qβ VLPs with mannose-IME



Qβ VLPs (1 mg / mL) in TAPS (0.6 M, pH 8.5) were mixed with the mannose-IME solution (800 equiv.). The reaction mixture was shaken at 25 °C for 4 h. The reaction was done in dark if the particles were labelled with fluorophores. The reaction mixture was purified by Micro Bio-Spin® P6, G25 minitrap / miditrap / PD10 columns depending on the scale using TBS buffer.

5.4.5 Cell growing

5.4.5.1 A549

The cells were grown in DMEM (10% FBS, 1% Pen/Strep) at 37 ^{o}C with 5% CO_2. For splitting, medium in the cell culture flask was removed. The cells were washed with sterile PBS (~10 mL). Trypsin/EDTA (5 mL) was added and the cells were incubated at 37 ^{o}C for 10 min. Pre-warmed medium (10 mL) was added to the flask and the cell culture was centrifuged down (300 x g, 10 min) and resuspended with medium (5 mL). The mixture was added to a new flask containing medium (20 mL). The cell culture was incubated at 37 ^{o}C for a few days before next splitting.

5.4.5.2 THP-1 monocytes

The cells were grown in RPMI 1640 (10% FBS, 1% Pen/Strep) at 37 ^{o}C with 5% CO_2. They were split by taking a portion of the cell culture to a new flask with fresh medium. The cells can be harvested by centrifugation (300 x g, 3 min) and resuspended in certain volume of medium or buffer.

5.4.5.3 THP-1 derived macrophages

THP-1 cells seeded at 2 x 10^5 cells / cm^2 were differentiated into macrophages by adding PMA (320 nM) and incubated at 37 ^{o}C with 5% CO_2 for 6 h. M2 polarisation was induced by adding IL-4 (20 ng / mL) and IL-13 (20 ng / mL). The cells were incubated at 37 ^{o}C with 5% CO_2 for at least 60 h before use (depending on the differentiation status).

5.4.6 Preparation of cell lysate

5.4.6.1 *E. coli*

Glycerol stock of BL21(DE3) with an ampicillin-resistant plasmid (1 µL) was added to LB medium (with 100 µg/mL ampicillin) (15 mL). The mixture was shaken at 37 ^{o}C for 16 h with a speed of 250 rpm. The overnight culture (1 mL) was added to eight flasks each

containing LB medium (with 100 µg/mL ampicillin) (750 mL). The cell culture was shaken at 37 ^{o}C with a speed of 250 rpm for 16 h. The cells were then centrifuged down at 9,000 x g for 10 min at 4 ^{o}C (F500 rotor). The cell pellet was resuspended in a minimum amount of TBS buffer (50 mM Tris, 150 mM NaCl, pH 7.5) (20 mL) and transferred to a beaker. The cell slurry was stirred for 15 min until the solution became homogeneous. After lysing the cell by a microfluidiser, the lysed solution was centrifuged down at 58,000 x g for 30 min at 4 ^{o}C (JLA 25.5 rotor). The supernatant was collected (85% w/v) and stored at – 80 ^{o}C.

5.4.6.2 THP-1 derived macrophage

Twelve T175 flasks of THP-1 derived macrophage (~8 x 10^5 cell / mL) (50 mL) were grown according to the protocols listed in the previous section. The medium was removed and the cells were washed with PBS (3 x 10 mL). The cells in each flask were treated with trypsin (10 mL) and incubated at 37 ^{o}C for 1 h. The remaining cells were detached by scraping. Fresh medium (10 mL) was added to each flask. The cell culture was transferred to a falcon tube and centrifuged down at 300 x g at 25 ^{o}C for 3 min. The pellet was resuspended with PBS and centrifuged again for thrice. The supernatant obtained from each wash was also combined, centrifuged and washed with PBS. The final combined cell pellet was frozen at – 80 ^{o}C for a few days. The pellet was then thawed and resuspended, followed by sonication (30%, 2 sec ON/OFF, 5 min). The cell solution was transferred to a 1.5 mL tube and centrifuged down at 17,000 x g at 4 ^{o}C for 30 min. The supernatant was collected (65% w/v) and used for NMR study.

5.4.7 LC-MS analysis

Unless otherwise specified, the sample was treated with 4.5 M urea or 4.5 M urea and 50 mM DTT (for species with disulphide bonds) in TBS or sodium phosphate buffer at 25 ^{o}C for 1 h before LCT mass spectrometry analysis. The ion signals obtained during the period of

t = 1.50 to t = 2.25 were integrated and analysed (**Figure S.2**). The mass spectra were reconstructed from the ion series and deconvoluted using the MaxEnt1 algorithm in the MassLynx software (v 4.1 from Waters).

5.4.8 Microscopy study of Qβ VLPs internalisation into different cells (live cell)

5.4.8.1 Adherent cells (A549 epithelial cell and THP-1 derived macrophage)

The cells were seeded on an 8-well microscopy slide where each well contained 160,000 cells (2 x 10^5 cells / cm^2) at 37 ^oC with 5% CO_2 for 16 h. Depending on the experiment, TBS, free dyes or labelled Qβ species were added to the wells and incubated for different time points at 37 ^oC. The medium was removed, and the wells were washed with PBS (3 x 200 µL). PBS (200 µL) was added to the wells followed by microscopy analysis.

5.4.8.2 Suspension cells (THP-1 monocyte)

The cells were added on an 8-well microscopy slide. For THP-1, each well contained 160,000 cells (2 x 10^5 cells / cm^2). Depending on the experiment, TBS, free dyes or labelled Qβ species were added to the wells and incubated for different time points. A few µL from each well was transferred to some microscopy slides for analysis.

5.4.9 Microscopy study of Qβ VLPs internalisation into fixed macrophage cell

Cover slips were placed on each well of a 24-well plate. THP-1 derived macrophage cells were seeded on the plate where each well contained 120,000 cells at 37 ^oC with 5% CO_2 for 16 h. The medium was removed and replaced with a fresh one (containing 20 ng / mL IL-4 and IL-13 and 320 nM PMA) (200 µL). Depending on the experiment, TBS, free dyes or labelled Qβ species were added to the wells and incubated for 20 min at 37 ^oC. The medium was removed, and the wells were washed with PBS (3 x 1 mL). 4% paraformaldehyde (1 mL) was added and the plate was incubated at 25 ^oC for 15 min in dark. After washing with PBS

(3 x 1 mL), permeabilising solution (0.1% Triton / PBS) (200 µL) was added and the plate was incubated at 25 °C for 15 min in dark. After washing with PBS (3 x 1 mL), blocking solution (2% BSA / PBS) (200 µL) was added and incubated at 25 °C for 1 h in dark with shaking. After removing the solution, primary antibodies (1 µg / mL rabbit anti-LAMP-1 and 1 µg / mL mouse anti-EEA1 in 2% BSA / PBS) (200 µL) were added and incubated at 25 °C for 1 h in dark with shaking. For isotype control, rabbit or mouse IgG isotype (1 µg / mL) (200 µL) was added instead. After washing with PBS (3 x 1 mL), secondary antibodies (1:2000 donkey anti-rabbit Alexa Fluor® 568 and 1:2000 goat anti-mouse Alexa Fluor® 647 in 2% BSA / PBS) (200 µL) were added and incubated at 25 °C for 1 h in dark with shaking. For membrane labelling, only primary antibody (1 µg / mL mouse anti-CD14 coupled to AlexaFluor647) was added, without the need of secondary antibody. After washing with PBS (3 x 1 mL), the cells were stained with DAPI solution at 25 °C for 20 min in dark with shaking. After washing with PBS (3 x 1 mL), the cover slip was placed on a microscopy plate with a drop of DAKO fluorescence mounting medium and left at 25 °C for 2 days in dark before storing at 4 °C. The microscopy plates were then analysed by confocal microscopy.

5.4.10 FACS study of Qβ VLPs internalisation into cells

5.4.10.1 Cell preparation

5.4.10.1.1 A549

The trypsin treated A549 cells (2 x 10^5 cells / cm^2) were transferred to a 96 well plate (100 µL / well). The plate was incubated at 37 °C with 5% CO_2 overnight.

5.4.10.1.2 THP-1

THP-1 cells (2 x 10^5 cells / cm^2) were transferred to a 96 well plate (100 µL / well). The cells were incubated at 37°C with 5% CO_2 overnight.

5.4.10.1.3 THP-1 derived macrophage

Macrophage grown in T75 flasks were dissociated by treating with trypsin (5 mL) at 37 ^{o}C with 5% CO_2 for 1 h. The cell mixture was centrifuged (300 x g, 3 min) and resuspended with RPMI 1640 (10% FBS, 1% Pen/Strep, 320 nM PMA, 20 ng / mL IL-4 and 20 ng / mL IL-13). The detached cells (2 x 10^5 cells / cm^2) were transferred to a 96 well plate (100 µL / well). The plate was incubated at 37 ^{o}C with 5% CO_2 overnight.

5.4.10.2 Qβ addition and FACS analysis

Different Qβ species (5 µL) (~100,000 MOI) were added to the wells. Control experiments were performed using TBS or fluorescein dye. 3 replicas were done for each sample. The plate was incubated at 37 ^{o}C with 5% CO_2 for different time points. For THP-1, the cells were transferred to a V-shaped 96 well plate and centrifuged (350 x g, 3 min, 4 ^{o}C). The pellet was washed with FACS buffer (100 µL). The washing process was repeated for twice before analysis. For macrophage and A549, the medium was removed and the cells were washed with PBS (3 x 100 µL). The cells were dissociated by treating with trypsin (5 mL) at 37 ^{o}C with 5% CO_2 for 10 min (A549) or 1 h (macrophage). The resulting cell mixture was diluted with FACS buffer (200 µL) before analysis.

5.4.11 Survival test of cells under conditions employed for NMR study

THP-1 monocytes or THP-1 derived macrophages (~5 x 10^5 cell / mL) were mixed with medium, TBS or Qβ species in the same ratio as those used for NMR study. The resulting mixtures (100 µL) were added to a 96 well plate and incubated at 37 ^{o}C with 5% CO_2. The number of live cells was counted at 1 h, 2 h, 6 h, 1 day and 2 days respectively.

5.4.12 NMR analysis of Qβ VLPs

5.4.12.1 1D ^{19}F NMR analysis of Qβ VLPs treated with cell lysate

Qβ species were concentrated with Vivaspin® (10 k MW) and resuspended with D_2O (50 μL). The concentrated protein solution was mixed with cell lysate (85% w/v for *E. coli*; 65% w/v for macrophage) (450 μL) and 0.05 mM of TFAcetone. The mixture was incubated at 37 ^{o}C for different time points and analysed by NMR spectroscopy. The samples were then treated with 0.2 M SDS, 50 mM DTT and heated at 90 ^{o}C for 60 min, followed by NMR analysis. ^{19}F NMR spectra were typically acquired at 298 K with acquisition times of 0.5 s and recovery delays of 2 s. The number of transients was 1536 for all the samples.

5.4.12.2 ^{19}F diffusion analysis of Qβ VLPs

Refer to Chapter 4.

5.4.12.3 In-cell NMR analysis

Qβ VLPs in TBS (~1.5 mg / mL) (250 μL) was added to THP-1 monocytes or THP-1 derived macrophage (~5 x 10^5 cell / mL) in a T175 flask with medium. Seven flasks were prepared in total for each cell. After 1 or 2 days, the cells were taken out from the flasks by simple pipetting (THP-1 monocytes) or by trypsin treatment and scraping (THP-1 derived macrophages). The cell culture was centrifuged (300 x g, 3 min) and then washed with TBS (10% D_2O) for 3 times. The pellet was resuspended with the remaining solution left and transferred to a sterile Shigemi NMR tube. The tube was left for settling overnight at 4 ^{o}C before analysis.

5.5 References

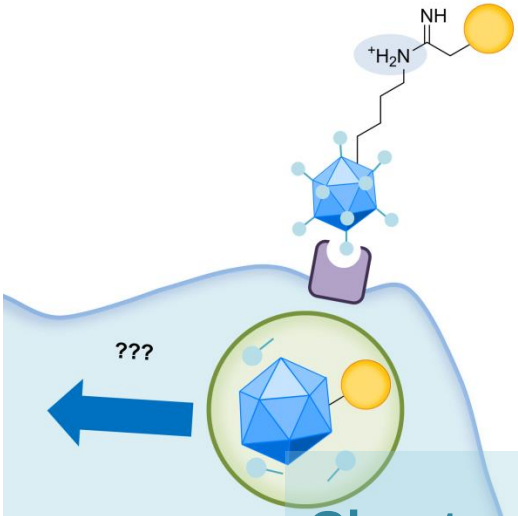
1. Ellis, R.J. Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Current Opinion in Structural Biology* **11**, 114-119 (2001).
2. Wang, Q., Zhuravleva, A. & Gierasch, L.M. Exploring Weak, Transient Protein–Protein Interactions in Crowded In Vivo Environments by In-Cell Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **50**, 9225-9236 (2011).
3. Crowley, P.B., Chow, E. & Papkovskaia, T. Protein Interactions in the Escherichia coli Cytosol: An Impediment to In-Cell NMR Spectroscopy. *ChemBioChem* **12**, 1043-1048 (2011).
4. in Principles of Fluorescence Spectroscopy (ed. Lakowicz, J.R.) 277-330 (Springer US, Boston, MA, 2006).
5. Calendar, R. The Bacteriophages (Oxford University Press, New York, 2006).
6. Tsukada, K. et al. Quantitative analysis of the bacteriophage Qβ infection cycle. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1790**, 65-70 (2009).
7. Chooi, K.P. et al. Synthetic Phosphorylation of p38α Recapitulates Protein Kinase Activity. *Journal of the American Chemical Society* **136**, 1698-1701 (2014).
8. Young, M.E., Carroad, P.A. & Bell, R.L. Estimation of diffusion coefficients of proteins. *Biotechnology and Bioengineering* **22**, 947-955 (1980).
9. Leung, L. C. R. et al. Monitoring the disassmebly of virus-like particles by ^{19}F NMR (unpublished)
10. Dwek, R.A., Butters, T.D., Platt, F.M. & Zitzmann, N. Targeting glycosylation as a therapeutic approach. *Nat Rev Drug Discov* **1**, 65-75 (2002).
11. Lennartz, M.R., Cole, F.S. & Stahl, P.D. Biosynthesis and processing of the mannose receptor in human macrophages. *J Biol Chem* **264**, 2385-90 (1989).
12. van Liempt, E. et al. Specificity of DC-SIGN for mannose- and fucose-containing glycans. *FEBS Lett* **580**, 6123-31 (2006).
13. Robinson, M.A. et al. LEAPT: Lectin-directed enzyme-activated prodrug therapy. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14527-14532 (2004).
14. Pearce, O.M.T. et al. Glycoviruses: Chemical Glycosylation Retargets Adenoviral Gene Transfer. *Angewandte Chemie International Edition* **44**, 1057-1061 (2005).
15. Lee, Y.C., Stowell, C.P. & Krantz, M.J. 2-Imino-2-methoxyethyl 1-thioglycosides: new reagents for attaching sugars to proteins. *Biochemistry* **15**, 3956-3963 (1976).
16. Beignet, J., Tiernan, J., Woo, C.H., Kariuki, B.M. & Cox, L.R. Stereoselective

synthesis of allyl-C-mannosyl compounds: use of a temporary silicon connection in intramolecular allylation strategies with allylsilanes. *J Org Chem* **69**, 6341-56 (2004).

17. Gunnoo, S.B. et al. Creation of a gated antibody as a conditionally functional synthetic protein. *Nat Commun* **5** (2014).
18. Ye, Y. et al. ^{19}F NMR Spectroscopy as a Probe of Cytoplasmic Viscosity and Weak Protein Interactions in Living Cells. *Chemistry – A European Journal* **19**, 12705-12710 (2013).
19. Shi, P. et al. In situ^{19}F NMR studies of an E. coli membrane protein. *Protein Science* **21**, 596-600 (2012).
20. Liu, J.J., Horst, R., Katritch, V., Stevens, R.C. & Wüthrich, K. Biased Signaling Pathways in β2-Adrenergic Receptor Characterized by ^{19}F-NMR. *Science* **335**, 1106-1110 (2012).
21. Li, C. et al. Protein ^{19}F NMR in Escherichia coli. *Journal of the American Chemical Society* **132**, 321-327 (2010).

PART III

Development of IME fluorescent dyes for retaining the infectivity of labelled viruses



Chapter 6

Development of IME-functionalised fluorescent dyes

Chapter 7

Labelling of viruses with IME-functionalised fluorescent dyes

CHAPTER

6 Development of IME-functionalised fluorescent dyes

6.1 Introduction

6.1.1 The effect of fluorescence labelling on virus infectivity

One of the most common ways to label viruses is by fluorescence labelling.[1] Most labelling procedures are quick and simple, involving the conjugation of small fluorophores into capsid proteins[2] or lipid membranes[3] on the viral surface. Although all reactive side chains on the capsid proteins can be labelled, it is crucial to control the amount of dye coupled. The infectivity of the virus can be largely reduced when its surface is covered with dye molecules, as this can greatly change the surface net charge and hinder the interaction sites responsible for cell entry.

The surface net charge of different viruses has been revealed to be a crucial component for guiding viral infection.[4-6] However, few studies have been undertaken to minimise the change of net charge, especially the charge on the coupled atom, following virus labelling. Lysine is one of the most common targets for virus labelling,[7] and it can easily be conjugated to commercially available dyes linked with different reactive groups such as an *N*-hydroxysuccinimide (NHS) ester, isothiocyanate (ITC) or thiocyanate. The pKaH of the lysine side chain is around 10.5, but coupling to the aforementioned reactive groups forms coupled products with pKaH of around – 0.5 (**Figure 6.1**).[8, 9] As a consequence, the nitrogen

atom on the side chain is almost completely deprotonated at physiological pH, resulting in a decrease of positive charge on the virus surface. This effect can be substantial if a lot of dyes are coupled to the virus.

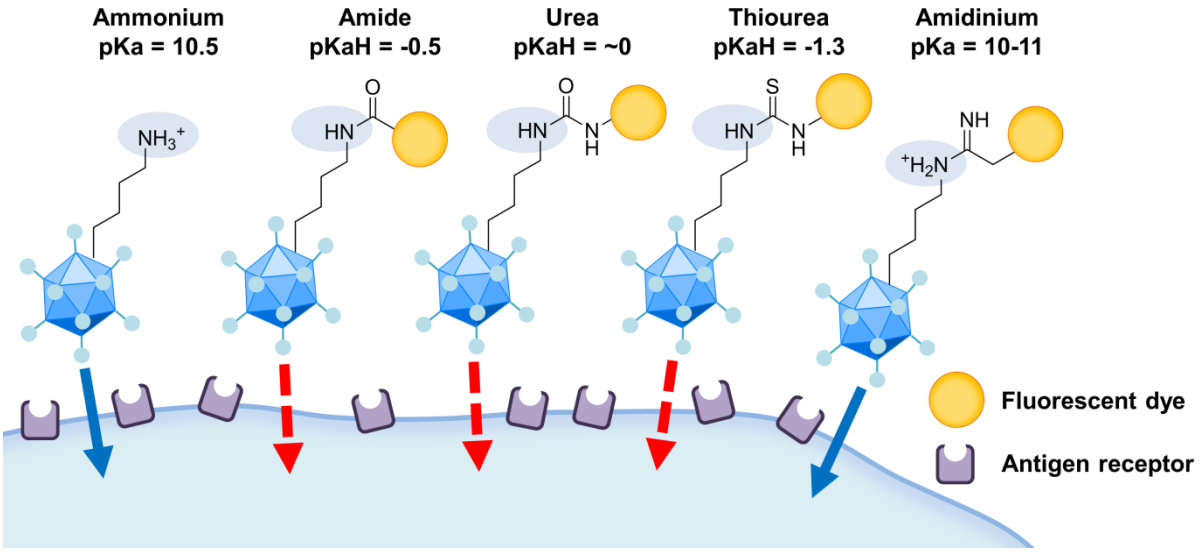


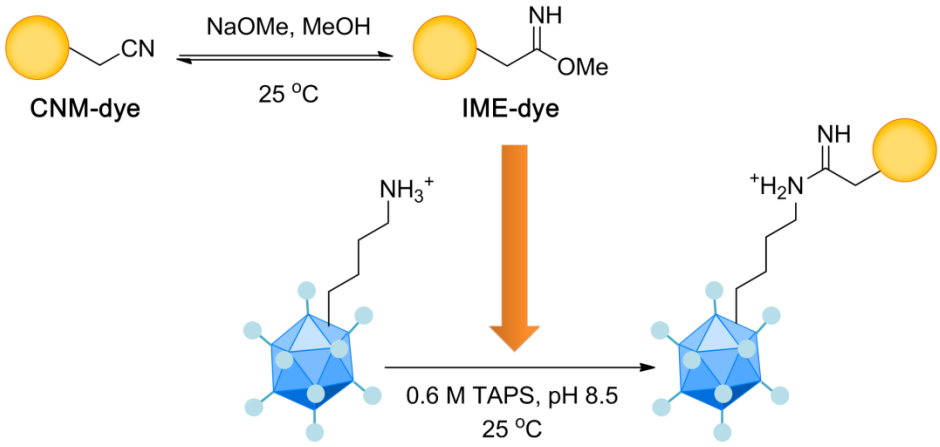
Figure 6.1: The surface net charge of a virus can be hugely altered by the coupling group used for dye labelling depending on the pKa value. In physiological pH, the amidine linkage formed from the 2-imino-2-methoxyethyl (IME) group retains the positive charge on the nitrogen atom.

6.1.2 Development of an IME-functionalised dye

The use of a 2-imino-2-methoxyethyl (IME) coupling group was proposed for minimising the effect on virus surface net charge following dye labelling. The IME group selectively reacts with primary amines (i.e. lysine residues and N-terminal amines) and has been widely used for coupling sugars to proteins.[10-12] The amidine linkage formed is known to be stable to neutral or acidic pH. Notably, it has a pKaH of around 10 – 11 , which is very similar to the unmodified lysine side chain and will therefore be protonated at physiological pH (**Figure 6.1**).[10] As demonstrated in the previous chapter, IME-sugars can be coupled effectively to big macromolecules like Qβ VLPs. It was therefore theorised that the functionalisation of viruses using IME-dyes would yield modified viruses that better mimic

the natural state.

IME-thioglycosides can be generated easily from cyanomethyl-thioglycosides (CNM-thioglycosides) *via* basic catalysis. By treating a CNM-thioglycoside with sodium methoxide in methanol, an equilibrium mixture of CNM- and IME-thioglycoside can be obtained in a 1:1 ratio.[10-12] Although this mixture can be used directly for protein modification, it should first be neutralised with acid to prevent protein damage. Alternatively, a strong buffer (pH 8.5) can be used for the protein reaction. It was proposed that this methodology could be adapted to incorporate fluorescent dyes into viruses. Proteins or viruses are expected to be labelled upon treatment with a mixture of CNM- and IME-dye in TAPS buffer (pKa = 8.44) (**Scheme 6.1**).



Scheme 6.1: The proposed IME-labelling strategy of viruses.

6.1.3 Rhodamine B as a fluorophore model

Rhodamine B was employed as a dye model to study the effect of the IME group as it is soluble in water, readily available, and has a high extinction coefficient and quantum yield.[13, 14] To compare the effect of different conjugating groups, rhodamine B dyes possessing an IME, NHS or ITC coupling group were synthesised or purchased (**Figure 6.2**).

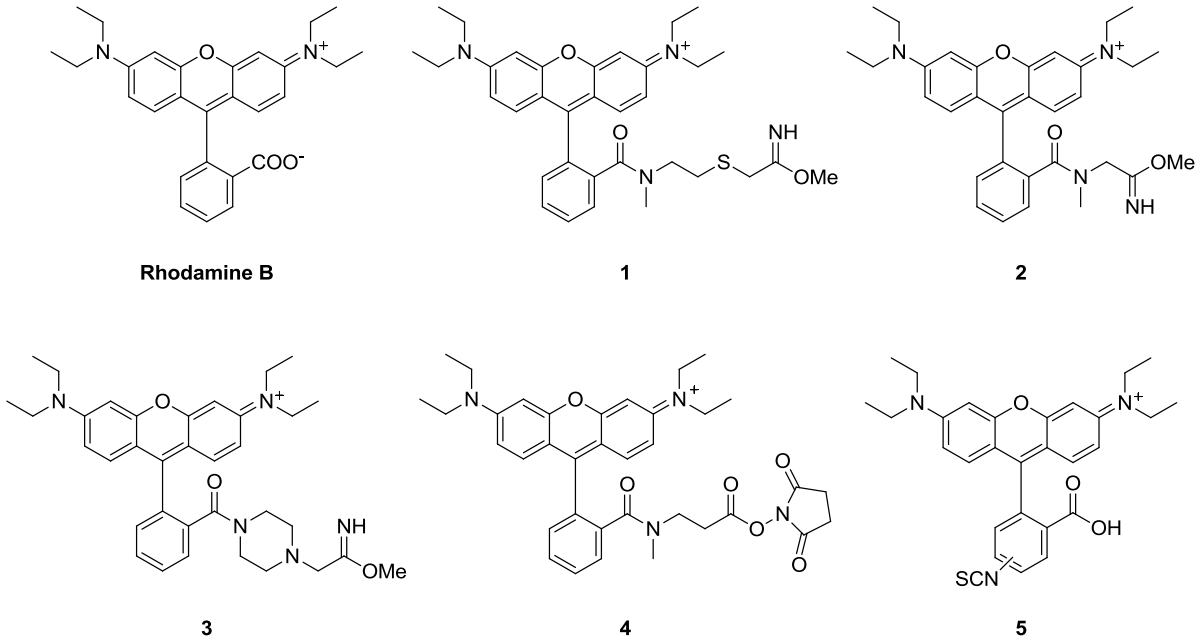
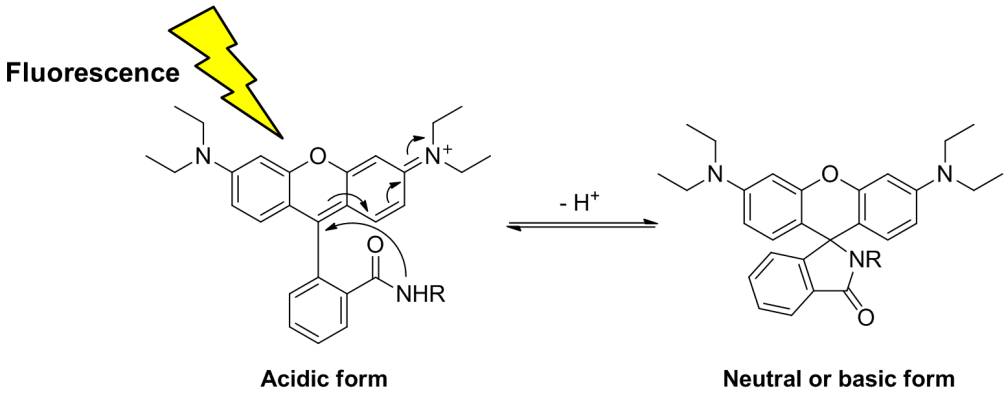


Figure 6.2: The proposed rhodamine B candidates. Only rhodamine B and dye **5** were commercially available.

For dye **1**, a thioether was introduced to mimic the original IME-thioglycoside. On the other hand, dye **2** was proposed to test if the sulphur atom is essential and the effect of the chain length. Dye **3** was suggested based on an effective strategy for synthesising rhodamine B derivatives previously reported.[15] Dyes **1** – **4** were functionalised with a secondary amide as primary amides are known to cyclise under neutral and basic conditions (i.e. in protein solution) to form non-fluorescent dyes (**Scheme 6.2**).[16]

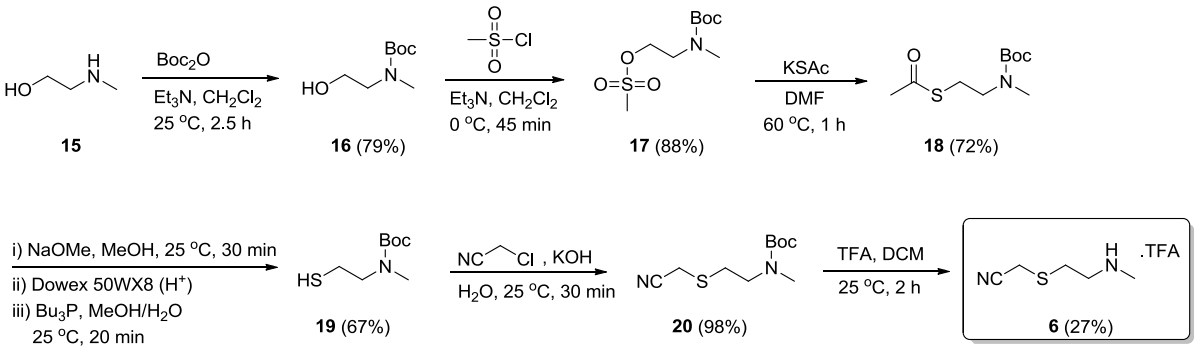


Scheme 6.2: Cyclisation of primary amide under neutral and basic condition impairs the fluorescent properties of rhodamine B dyes.

6.2 Results and Discussion

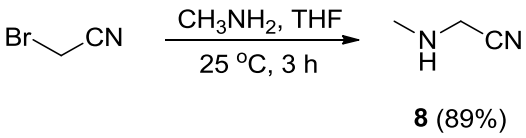
6.2.1 Synthesis of rhodamine B derivatives

Several linker precursors were prepared for adding the corresponding functional groups to rhodamine B. Linker **6** used for dye **1** was synthesised in 6 steps (**Scheme 6.3**). Protection of *N*-methylehanolamine (**15**) as a Boc-carbamate, followed by mesylation of the free hydroxyl group and displacement with potassium thioacetate yielded thioacetate **18** in good yields. Subsequent deprotection of the thioacetate yields thiol **19**,[17] which was then converted into linker **6** by alkylation with chloroacetonitrile,[18] followed by Boc-deprotection using TFA.[19] The CNM-dye **7** was formed by coupling the linker to rhodamine B.



Scheme 6.3: Synthesis of linker **6** for dye **1**.

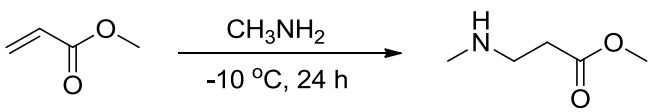
Linker **8** for dye **2** was synthesised by simple nucleophilic substitution between 2-bromoacetonitrile and methylamine as reported in the literature (**Scheme 6.4**).[20] Linker **8** was coupled to activated rhodamine B to give the CNM-dye **9** with a satisfactory yield compared with that of **7** and **11**.



8 (89%)

Scheme 6.4: Synthesis of linker **8** for dye **2**.

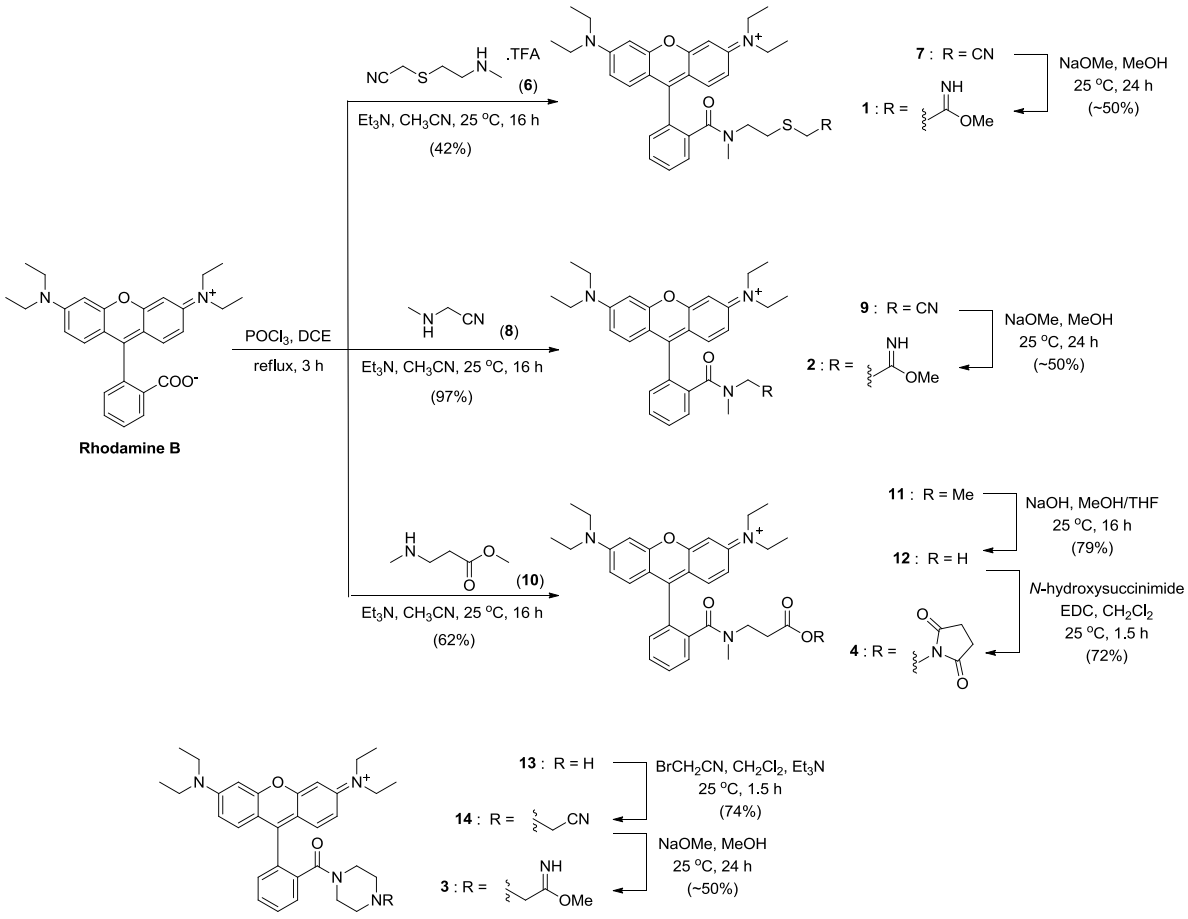
For dye **4**, an NHS group can be attached directly to the carboxylic group of rhodamine B without a linker. However, as the amides formed between lysine residues and NHS esters are primary amides, cyclization will occur and the fluorescent signal will be largely reduced (**Scheme 6.2**). Thus, a linker (**10**) containing a secondary amine and a methyl ester was first coupled to the dye before adding the NHS group. This linker can be easily synthesised from methylamine and methyl acrylate *via* a 1,4-conjugate addition (**Scheme 6.5**).[21, 22] Once the linker was coupled to rhodamine B, the resulting ester **11** was then hydrolysed to give carboxylic acid **12**. Interestingly, rotameric peaks were observed in both the ^{1}H and ^{13}C NMR spectra of **12**, which were not found in other species. Rotamers are often observed in secondary amides due to rotational restriction and they can rarely be distinguished by NMR at room temperature.[23] It is possible that the interaction between carboxylic acid and amide in **12** largely restricted the rotation of the amide bond, resulting in this unique property. Finally, the NHS dye **4** was formed by activation of carboxylic acid **12** using 1-ethyl-3-(3-dimethylamino-propyl)-carbodiimide (EDC), followed by coupling to *N*-hydroxy-succinimide.



10 (16%)

Scheme 6.5: Synthesis of linker **10** for dye **4**.

The syntheses of rhodamine B derivatives are shown in **Scheme 6.6**. For dyes **1**, **2** and **4**, the carboxylic group of rhodamine B was activated by phosphoryl chloride, followed by coupling an appropriate linkers.[24] Dye **3** was synthesised using a different strategy than that used in the synthesis of the other dyes. Piperazine **13** is prepared by treatment of piperazine and rhodamine B with trimethyl aluminium, as described in a previously reported procedure.[15] Piperazine **13** was then alkylated with bromoacetonitrile to yield the CNM dye **14** in 74% yield.[25]



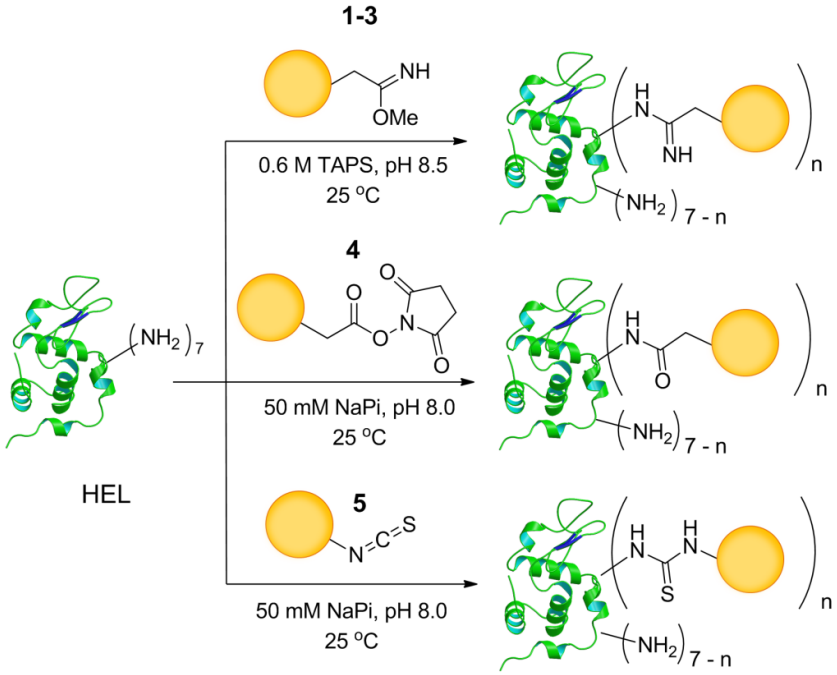
Scheme 6.6: Synthesis of rhodamine B derivatives with different conjugating groups.

Treatment of the CNM-dyes **7**, **9** and **14** with sodium methoxide in methanol yielded IME-dyes **1**, **2** and **3** respectively through a reversible reaction. The ratio of starting material to product was around 1:1 for all the three dyes based on analysis by TLC and mass

spectrometry, which was consistent with the previous results obtained with sugars from our group.[11, 26]

6.2.2 Labelling of HEL with the rhodamine B derivatives

The coupling ability of the dye derivatives was tested on the model protein hen egg white lysozyme (HEL) (**Scheme 6.7**). This protein possesses seven reactive amines which could be readily labelled. For the IME-dyes, the activated dye mixture was added to the protein solution in TAPS buffer directly without acidification. On the other hand, the reaction with NHS dye **4** or ITC dye **5** was carried out in sodium phosphate buffer, a common condition used for these conjugating groups. As revealed by both SDS-PAGE (**Figure 6.3**) and mass spectrometry (**Figure 6.4**), all the dyes were reactive except the IME-dye **3**. It is possible that the bulky piperazine group of dye **3** hindered the reaction from happening. Also, it was found that the NHS dye **4** and ITC dye **5** were more reactive than the IME-dyes synthesised. Dyes **4** and **5** were also found to be compatible with TAPS in later experiments (data not shown).



Scheme 6.7: HEL was labelled with the rhodamine B derivatives under different conditions

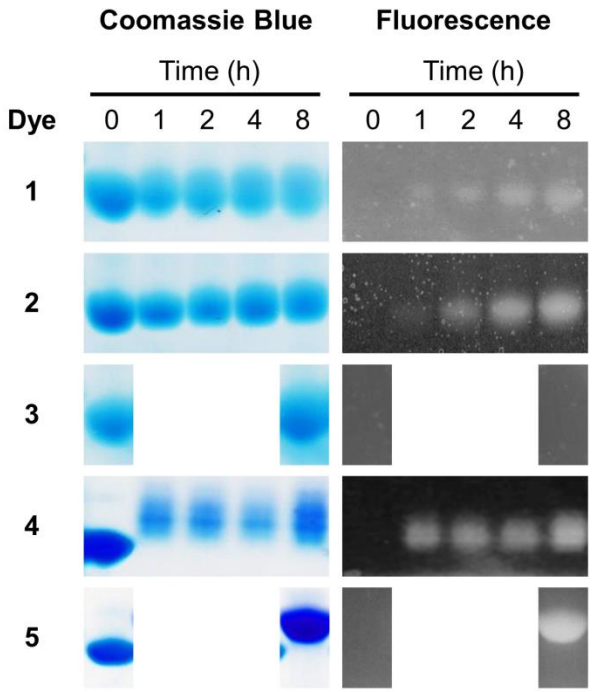
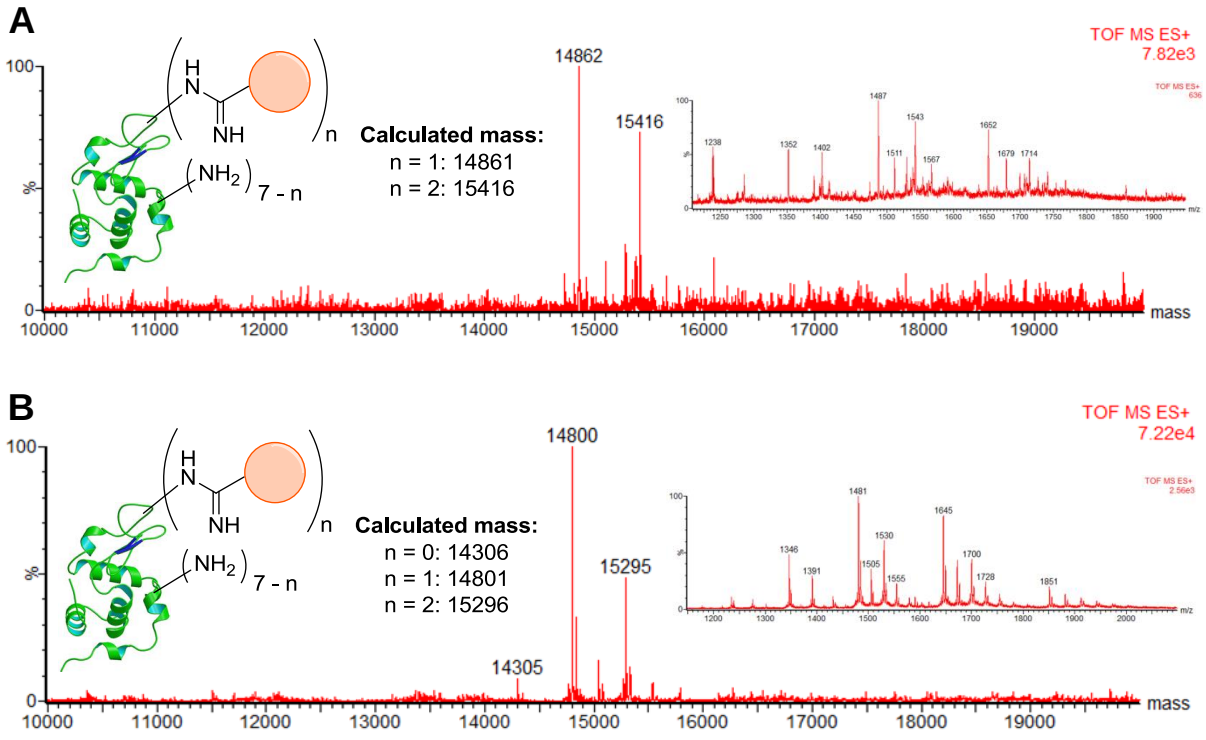


Figure 6.3: Reducing SDS-PAGE gel showing the reaction of HEL with rhodamine B derivatives (120 equiv.) at different time points. The gel was analysed by both Coomassie Blue staining and fluorescence scanning. The protein had a molecular mass of around 14 kDa.



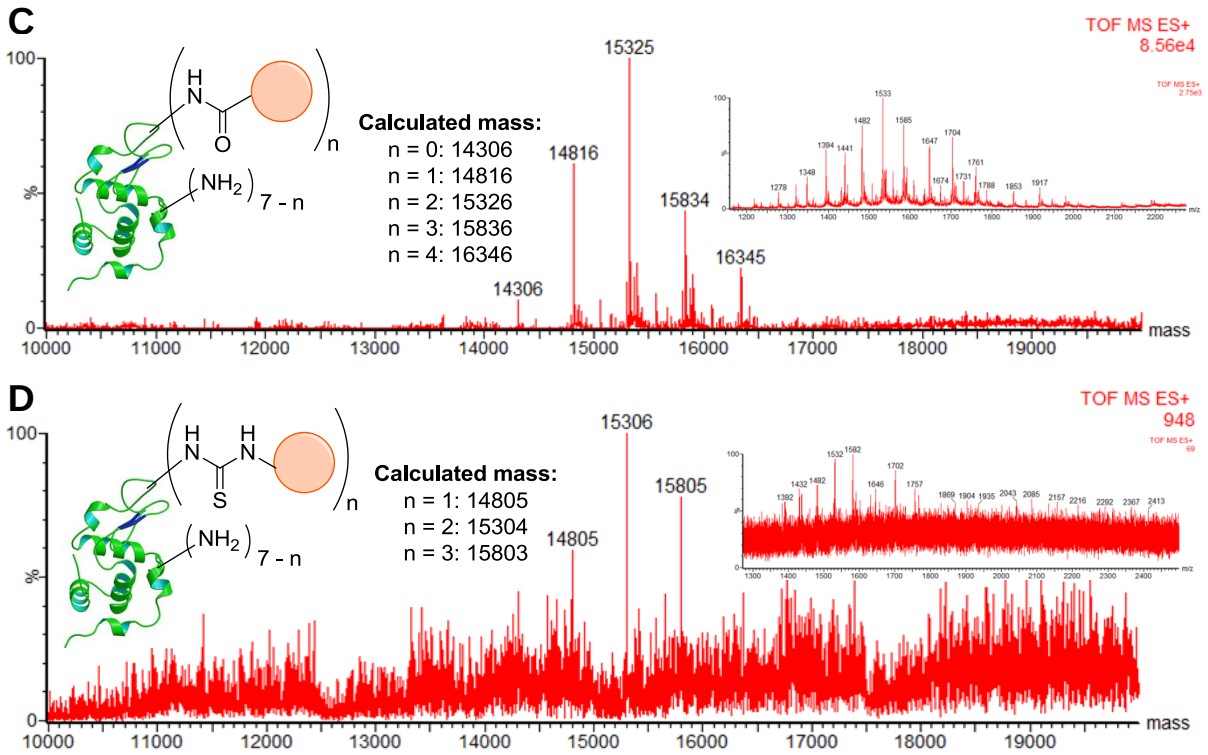


Figure 6.4: Deconvoluted mass spectra and corresponding raw ion series of the reaction between HEL and (A) **1** (60 equiv.); (B) **2** (60 equiv.); (C) **4** (120 equiv.) or; (D) **5** (120 equiv.) at 25 ^{o}C for 16 h.

6.3 Conclusion

To the best of our knowledge, this is the first time that an IME functional group has been coupled to a fluorescence dye. Pleasingly, two out of the three proposed IME-dyes worked perfectly on the simple protein HEL. As revealed by the positive results of dye **1** and **2**, the sulphur atom next to the IME group is not necessary for the conjugation. However, it was found that the presence of the thioether can increase the water solubility of the dye, which is crucial for bioconjugation. Due to the noxious nature of the thiols used in the synthesis of dye **1**, further modification of the dye to replace the thioether with an ether group could be useful.

The next stage would be to apply these dyes to real viruses to see if they can work in the same way, and whether the IME-dyes can minimise the effect of labelling on virus infectivity compared to dyes with other conjugating groups.

6.4 Experimental Section

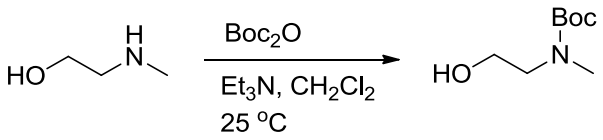
6.4.1 General procedures

Chemicals were purchased from commercial suppliers and used without purification. NuPAGE® MES buffer and 12% Bis-Tris gel were purchased from Life Technologies™. Protein markers used were Perfect Protein™ Marker 10-225 kDa purchased from Novagen® . PD SpinTrap™ G-25, MiniTrap™ G-25, MidiTrap™ G-25 and PD10 desalting columns were purchased from GE Healthcare Life Sciences.

Small molecule and protein mass spectrometry analysis was performed in Bruker μTOF and Waters LCT Premier™ respectively. Chemical NMR study was performed in Bruker AVIII400 HD Nanobay instrument running TOPSPIN 3 equipped with a 5 mm z-gradient multinuclear BBFO probe. Infrared spectrometry was performed in Bruker Tensor 27. Thin layer chromatography (TLC) was performed using aluminium sheets coated with 60F254 silica gel and visualised using ninhydrin solution or UV lamp. Flash chromatography was performed using Geduran SI 60 (40-63 μm) silica (Merck).

6.4.2 Synthesis of linkers

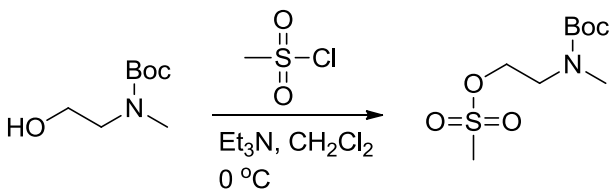
6.4.2.1 2-[*N*-(tert-Butyloxycarbonyl)-*N*-methylamino]ethanol (16)



15 **16** (99%)

The protocol was modified from that of Chalker, *et al.*[17] N-methylethanolamine (**15**) (5.6 mL, 69.7 mmol) was added to a 250 mL RB flask and dissolved in CH_2Cl_2 (100 mL). Et_3N (12.4 mL, 89.0 mmol) was added and the stirred solution was cooled to 0 °C. Boc_2O (13.3 mL, 57.9 mmol) was added and the reaction was stirred at 0 °C for 5 min and then at 25 °C for 2.5 h. The reaction was cooled to 0 °C and quenched with 1 M HCl (200 mL). The mixture was then diluted with CH_2Cl_2 (100 mL) and the organic layer was collected. The organic layer was washed with 1 M HCl (2 x 100 mL) while the aqueous layer was extracted with CH_2Cl_2 (2 x 50 mL). The combined organic layers were dried (Na_2SO_4), filtered and concentrated under reduced pressure to give alcohol **16** as a colourless liquid (12.2 g, 99%); R_f = 0.75 (9:1 DCM/MeOH, v/v); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H = 1.45 (s, 9H, Boc), 2.90 (s, 3H, NMe), 3.38 (t, J 4.98, 2H, **CH$_2$**CH$_2$OH), 3.73 (t, J 5.26, 2H, CH$_2$**CH$_2$**OH); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C = 28.53 (C**Me$_3$**), 35.65 (NMe), 51.59 (**CH$_2$**CH$_2$OH), 61.79 (CH$_2$**CH$_2$**OH), 80.11 (**C**Me$_3$); LRMS m/z (ESI$^+$): measured: 198.1 (M + Na$^+$), calculated: 198.1; IR: 3426, 2976, 1669, 1482, 1393, 1366, 1150 cm^{-1}.

6.4.2.2 2-[*N*-(tert-Butyloxycarbonyl)-*N*-methylamino]ethyl methanesulfonate (17)

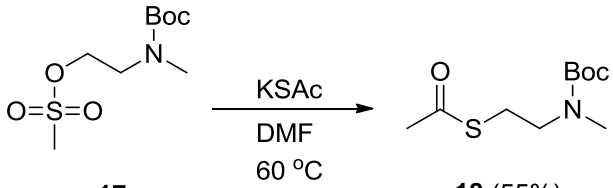


16 **17** (87%)

The protocol was modified from that of Chalker, *et al.*[17] Alcohol **16** (12.2 g, 69.6 mmol)

was added to a 250 mL RB flask and dissolved in CH_2Cl_2 (120 mL). Et_3N (12.6 mL, 90.5 mmol) was added and the stirred solution was cooled to 0 ^oC. Methanesulfonyl chloride (5.7 mL, 73.1 mmol) was added dropwise and CH_2Cl_2 (10 mL) was used to rinse the flask. The flask was stoppered and stirred at 0 ^oC for 1 h. The reaction was quenched at 0 ^oC with 1 M HCl (100 mL). The mixture was then diluted with CH_2Cl_2 (50 mL) and washed with 1 M HCl (2 x 100 mL). The organic layer was dried (Na_2SO_4), filtered and concentrated under reduced pressure to give mesylate **17** as a yellow oil (15.4 g, 87%). The product was unstable and thus used immediately in the next step; R_f = 0.81 (EtOAc); LRMS m/z (ESI^+): measured: 254.1 ($M + Na^+$), calculated: 254.1.

6.4.2.3 2-[*N*-(tert-Butyloxycarbonyl)-*N*-methylamino]acetylethanthiol (18)

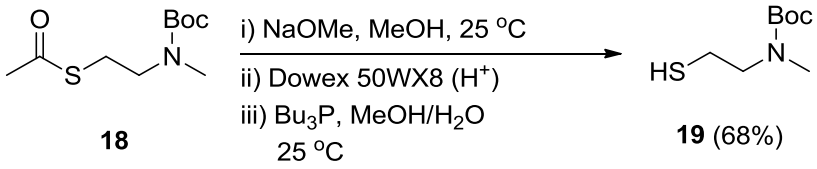


17 **18** (55%)

The protocol was modified from that of Chalker, *et al.*[17] Mesylate **17** (15.4 g, 60.8 mmol) was added to a 250 mL RB flask and dissolved in DMF (190 mL). Potassium thioacetate (10.4 g, 91.2 mmol) was added and the reaction mixture was stirred vigorously. The flask was stoppered and heated at 60 ^oC for 2 h, and then cooled to 25 ^oC. The reaction mixture was diluted with EtOAc (300 mL) and washed with H_2O (3 x 400 mL) and brine (300 mL). The organic layer was collected, while the aqueous layer was further extracted with EtOAc (3 x 200 mL). All the organic layers were combined, dried (Na_2SO_4), filtered and concentrated under reduced pressure. The resulting residue was further purified by silica column chromatography (20% EtOAc in PE) to thioacetate **18** as a yellow oil (7.86 g, 55%); R_f = 0.85 (1:1 EtOAc/PE, v/v); ^1H NMR (400 MHz, $CDCl_3$) δ_H = 1.45 (s, 9H, Boc), 2.33 (s, 3H, SAc), 2.89 (s, 3H, NMe), 3.00 (t, J 7.13, 2H, $CH_2\mathbf{CH_2}SAc$), 3.34 (t, J 7.17, 2H,

CH_2 CH_2SAc); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C = 27.35 (**CH_2**SAc), 28.53 (C**Me_3**), 30.78 (SC(O)**CH_3**), 34.73 (NMe), 48.46 (**CH_2**NMe), 79.83 (**C**Me_3); LRMS m/z (ESI^+): measured: 256.1 ($M + Na^+$), 272.1 ($M + K^+$), calculated: 256.1, 272.1; IR: 2976, 1689, 1480, 1391, 1365, 1163, 1129 cm^{-1}.

6.4.2.4 2-[*N*-(tert-Butyloxycarbonyl)-*N*-methylamino]ethanthiol (19)

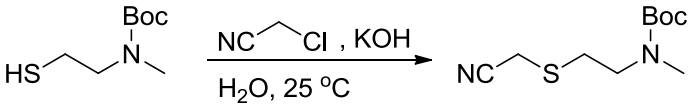


18 **19** (68%)

The protocol was modified from that of Chalker, *et al.*[17] Thioacetate **18** (7.86 g, 33.7 mmol) was added to a 250 mL RB flask and dissolved in MeOH (100 mL) under nitrogen. 25% NaOMe in MeOH (11.6 mL, 50.5 mmol) was added and the reaction mixture was stirred at 25 °C for 30 min. The reaction mixture neutralised using Dowex-50WX8 (H^+) as indicated by wet pH paper. After removing the resin by filtration, the resin was washed with MeOH (25 mL) and dried under reduced pressure. The resulting residue was redissolved in MeOH (90 mL) and H_2O (15 mL). Tributylphosphine (5.1 mL, 20.7 mmol) was added dropwise and the reaction mixture was stirred at 25 °C for 20 min. The mixture was diluted with Et_2O (200 mL) and washed with H_2O (100 mL). The organic layer was collected, while the aqueous layer was further extracted with Et_2O (2 x 100 mL). The combined organic extract were washed with brine (200 mL), dried (Na_2SO_4), filtered and dried under reduced pressure. The resulting residue was further purified by silica column chromatography (5% EtOAc in PE) to give thiol **19** as a yellow oil (4.38 g, 68%); R_f = 0.41 (1:4 EtOAc/PE, v/v); ^1H NMR (400 MHz, $CDCl_3$) δ_H = 1.31 (t, J 8.19, 1H, SH), 1.45 (s, 9H, Boc), 2.65 (m, 2H, **CH_2**SH), 2.88 (s, 3H, NMe), 3.37 (t, J 7.25, 2H, **CH_2**CH_2SH); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C = 22.68 (**CH_2**SH), 28.55 (C**Me_3**), 34.96 (NMe), 52.32 (**CH_2**CH_2SH), 79.83 (**C**Me_3);

LRMS m/z (ESI^+): measured: 214.1 ($M + Na^+$), calculated: 214.1; IR: 2975, 2563, 1688, 1480, 1390, 1365, 1162, 1135 cm^{-1}.

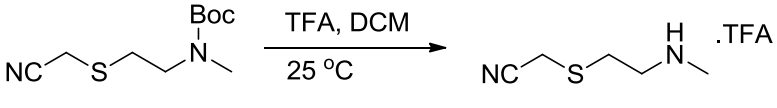
6.4.2.5 2-[2-[*N*-(tert-Butyloxycarbonyl)-*N*-methylamino]ethanthio]acetonitrile (20)



19 **20** (98%)

Potassium hydroxide (3.08 g, 54.9 mmol) was dissolved in H_2O (60 mL). Thiol **19** (4.38 g, 22.9 mmol) was added to the solution, followed by dropwise addition of chloroacetonitrile (2.8 mL, 43.5 mmol). After stirring at 25 °C for 30 min, the reaction mixture was extracted with chloroform (3 x 100 mL), dried (Na_2SO_4) and concentrated under reduced pressure to give nitrile **20** as a colourless liquid (5.16 g, 98%); R_f = 0.37 (1:4 EtOAc/PE, v/v); ^1H NMR (400 MHz, $CDCl_3$) δ_H = 1.47 (s, 9H, Boc), 2.90 (m, 5H, NMe and CH_2**CH_2**S), 3.42 (m, 4H, CH_2CN and **CH_2**CH_2S); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C = 17.14 (**CH_2**CN), 28.51 (C**Me_3**), 30.28 (CH_2**CH_2**S), 34.57 (NMe), 46.48 (**CH_2**CH_2S), 79.30 (**C**Me_3); HRMS m/z (ESI^+): measured: 253.0987 ($M + Na^+$), calculated: 253.0981; IR: 2974, 1683, 1481, 1392, 1366, 1160, 1129 cm^{-1}.

6.4.2.6 2-(*N*-methylaminoethanthio)acetonitrile (6)

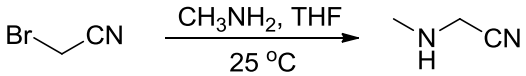


20 **6** (27%)

Nitrile **20** (2.79 g, 12.1 mmol) was dissolved in CH_2Cl_2 (10 mL), followed by dropwise addition of trifluoroacetic acid (10 mL, 103 mmol). The reaction mixture was stirred at 25 °C for 2 h. The mixture was then diluted with H_2O and washed with Et_2O (2 x 25 mL). The aqueous layer was neutralised with $NaHCO_3$ and extracted with CH_2Cl_2 (3 x 30 mL). The pH

of the aqueous layer was further increased to 10 by adding NaOH and extracted with CH_2Cl_2 (3 x 30 mL). The combined organic extracts were concentrated under reduced pressure. The resulting residue was further purified by silica column chromatography (10% MeOH in EtOAc with 1% NH_3) to give amine **6** as a colourless liquid (0.43 g, 27%) ; R_f = 0.48 (1:4:0.05 MeOH/EtOAc/NH_3, v/v/v); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H = 2.76 (s, 3H, NMe), 3.14 (t, J 7.01, 2H, CH_2**CH_2**S), 3.29 (t, J 7.06, 2H, **CH_2**CH_2S), 3.42 (s, 2H, **CH_2**CN); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C = 16.82 (**CH_2**CN), 28.00 (CH_2**CH_2**S), 33.97 (NMe), 48.31 (**CH_2**CH_2S), 126.52 (CN); HRMS m/z (ESI$^+$): measured: 131.0638 (M + H$^+$), calculated: 131.0637; IR: 3110, 1677, 1427, 1203, 1130 cm^{-1}.

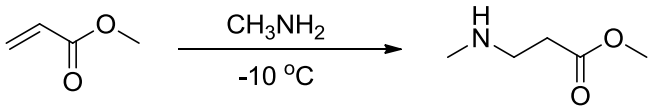
6.4.2.7 2-(methylamino)acetonitrile (8)



8 (89%)

Methylamine (2.0 M in THF) (57.1 mL, 115 mmol) was added to a dry 250 mL RB flask filled with argon and cooled to 0 °C. 2-bromoacetonitrile (2 mL, 28.7 mmol) was added to the vigorously stirred solution at < 10 °C. The reaction mixture was allowed to warm up to room temperature and stirred for 3 h. The reaction mixture was filtered. The flask was washed with Et_2O (40 mL x 2) and filtered. The filtrate was combined and the solvent was removed under low vacuum to give amine **8** as a yellow oil (1.80 g, 89%); R_f = 0.50 (1:9 MeOH/CH_2Cl_2 v/v); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H = 1.30 (s, 1H, NH), 2.52 (s, 3H, CH_3), 3.55 (s, 2H, CH_2); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C = 35.47 (CH_3), 39.20 (CH_2), 117.68 (CN); IR: 3342, 2946, 2804, 2240, 1453, 1420, 1333, 1125, 978, 870, 773 cm^{-1}.

6.4.2.8 Methyl 3-(methylamino)propanoate (10)

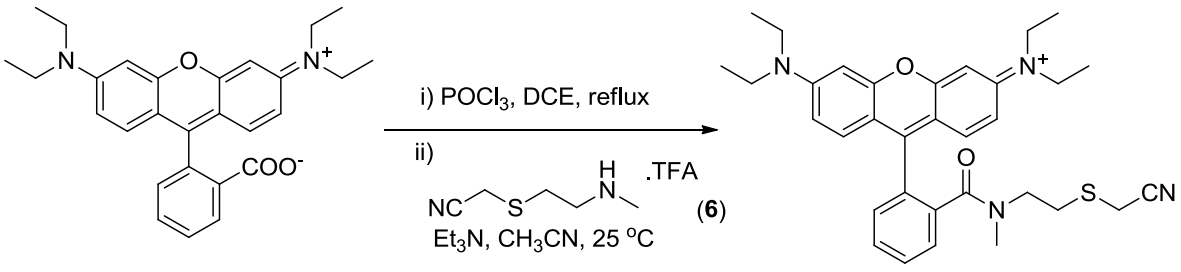


10 (16%)

Methylamine (33% wt in EtOH) (21.2 mL, 0.17 mol) was added to a 100 mL RB flask and cooled to -20 °C. Methyl acrylate (9.1 mL, 0.10 mol) was added slowly to the solution with a dropping funnel and stirred at -10°C for 24 h. Solvents were removed *in vacuo* and the residue was distilled under reduced pressure (46 °C, 15 mbar) to give amine **10** as a colourless liquid (1.83 g, 16%); R_f = 0.29 (1:4 MeOH/CH_2Cl_2, v/v); ^{1}H NMR (400 MHz, MeOD) δ_H = 2.37 (s, 3H, NMe), 2.53 (t, J 6.76, 2H, **CH$_2$**COOCH$_3$), 2.81 (t, J 6.73, 2H, **CH$_2$**NHMe), 3.36 (s, 3H, COO**CH$_3$**); ^{13}C NMR (100 MHz, MeOD) δ_C = 34.45 (**CH$_2$**COOCH$_3$), 35.88 (NMe), 47.74 (**CH$_2$**NHMe), 49.85 (COO**CH$_3$**), 174.45 (**C**OOCH$_3$); LRMS m/z (ESI$^+$): measured: 118.2 (M + H$^+$), calculated: 118.1; IR: 2954, 1733, 1437, 1168, 1038 cm^{-1}.

6.4.3 Synthesis of rhodamine derivatives

6.4.3.1 Rhodamine B derivative (7)



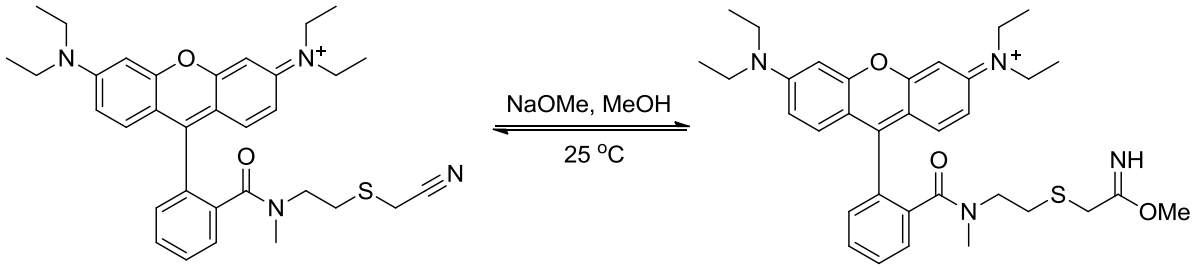
Rhodamine B

7 (42%)

Rhodamine B (100 mg, 0.23 mmol) was dissolved in 1,2-dichloroethane (5 mL). Phosphoryl chloride (0.31 mL, 3.28 mmol) was added to the solution and stirred at 110 °C for 3 h. The solvent and the excess phosphoryl chloride were removed under reduced pressure. The residue was dissolved in dry CH_3CN (5 mL), followed by the addition of Et_3N

(0.08 mL, 0.56 mmol) and amine **6** (119 mg, 0.92 mmol). After stirring at 25 ^{o}C for 16 h, the solvent was removed under reduced pressure. The resulting residue was purified by silica column chromatography (CH_2Cl_2, then 5% MeOH in CH_2Cl_2, then 20% MeOH in CH_2Cl_2). The fractions containing the product were evaporated under vacuum. The resulting residue was dissolved in $CHCl_3$ (25 mL) and washed with H_2O (25 mL). The organic layer was dried (Na_2SO_4), filtered and concentrated under reduced pressure to give the CNM-dye **7** as a purple solid (54.2 mg, 42%); R_f = 0.78 (1:9 MeOH/ CH_2Cl_2, v/v); ^{1}H NMR (400 MHz, MeOD) δ_H = 1.29 (t, J 7.09, 12H, $NCH_2\mathbf{CH_3}$), 2.49 (t, J 6.53, 2H, $NCH_2\mathbf{CH_2}$), 2.85 (s, 3H, NMe), 3.28 (s, 2H, $S\mathbf{CH_2}CN$), 3.49 (t, J 6.50, 2H, $N\mathbf{CH_2}CH_2$), 3.67 (q, J 7.18, 8H, $N\mathbf{CH_2}CH_3$), 6.95 (d, J 2.42, 2H), 7.05 (dd, J 2.44, 9.55, 2H), 7.27 (d, J 9.53, 2H), 7.51 (m, 1H), 7.64 (m, 1H), 7.75 (m, 2H); ^{13}C NMR (100 MHz, MeOD) δ_C = 12.83 ($NCH_2\mathbf{CH_3}$), 16.52 ($S\mathbf{CH_2}CN$), 29.94 ($NCH_2\mathbf{CH_2}$), 38.04 (NMe), 45.78 ($N\mathbf{CH_2}CH_2$), 46.89 ($N\mathbf{CH_2}CH_3$), 97.37, 111.45, 115.00, 115.24, 117.98, 128.56, 130.91, 131.37, 131.54, 133.35, 137.45, 156.73, 157.26, 159.36, 171.01; HRMS m/z (ESI^{+}): measured: 555.2771 (M^{+}), calculated: 555.2788; IR: 2920, 2851, 2160, 1586, 1464, 1411, 1336, 1273, 1178, 1072 cm^{-1}.

6.4.3.2 Activation of IME-rhodamine B (1)

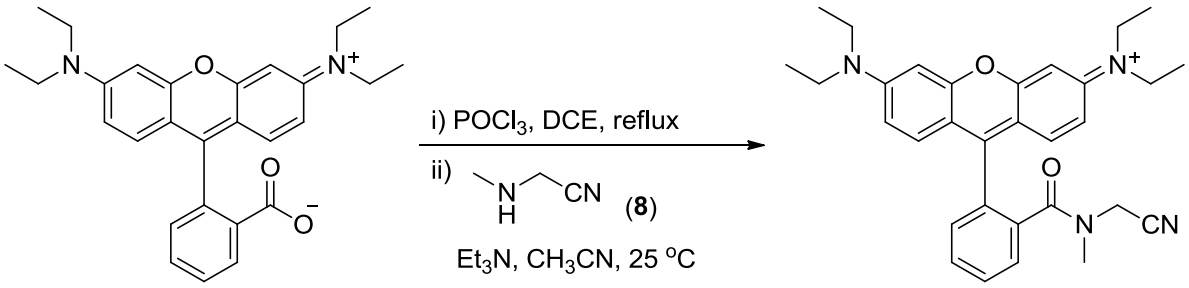


7 **1**

CNM-dye **7** (1 equiv.) was suspended in MeOH (250 µL per 5 mg of dye) in a vial/flask under argon. 25% sodium methoxide in methanol (1 equiv.) was added to the solution and stirred at 25 ^{o}C for 24 h to give an 1:1 equilibrium mixture of CNM-dye **7** and IME-dye **1**.

The solvent was removed by nitrogen and reduced pressure. The mixture was dissolved in DMSO (100 – 200 mg / mL) and used directly for labelling without further purification; R_f of **1** = 0.52 (1:9 MeOH/ CH_2Cl_2, v/v); HRMS m/z (ESI^+): measured: 587.3034 (M^+), calculated: 587.3050; IR: 2928, 2160, 1683, 1616, 1589, 1512, 1400, 1339, 1264, 1203, 1180, 1118, 1073 cm^{-1}.

6.4.3.3 Rhodamine B derivative (9)

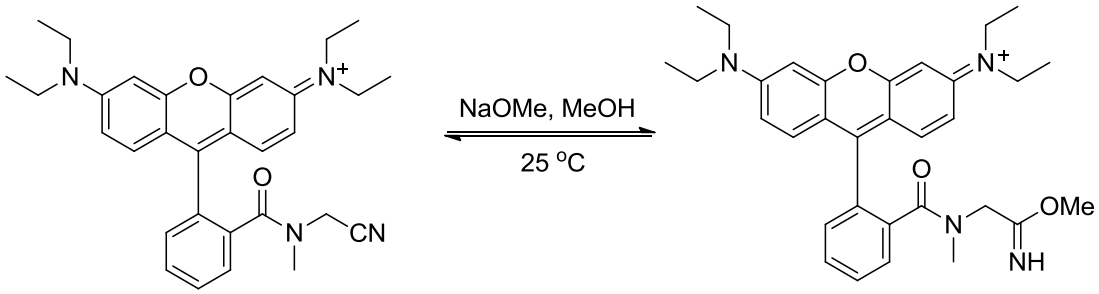


Rhodamine B **9** (97%)

Rhodamine B (500 mg, 1.13 mmol) was dissolved in 1,2-dichloroethane (30 mL). Phosphoryl chloride (1.53 mL, 16.4 mmol) was added to the solution and stirred at 110 °C for 3 h. The solvent and the excess phosphoryl chloride were removed under reduced pressure. Amine **8** (317 mg, 4.52 mmol) was added to the flask and dried under reduced pressure. The mixture was dissolved in dry CH_3CN (30 mL), followed by the addition of Et_3N (0.24 mL, 1.69 mmol). After stirring at 25 °C for 16 h, the solvent was removed under reduced pressure. The resulting residue was dissolved in $CHCl_3$ (40 mL) and washed with HCl (40 mL x 3). The organic layer was dried (Na_2SO_4), filtered and concentrated under reduced pressure to give the CNM-dye **9** as a purple solid (585 mg, 97%); R_f = 0.64 (1:9 MeOH/ CH_2Cl_2, v/v); ^1H NMR (400 MHz, MeOD) δ_H = 1.31 (t, J 7.11, 12H, $NCH_2\mathbf{CH_3}$), 3.08 (s, 3H, NMe), 3.68 (q, J 7.14, 8H, $N\mathbf{CH_2}CH_3$), 4.23 (s, 2H, NCH_2CN), 6.95 (d, J 2.24, 2H), 7.09 (dd, J 2.41, 9.57, 2H), 7.27 (d, J 9.52, 2H), 7.58 (m, 1H), 7.73 (m, 1H), 7.80 (m, 2H); ^{13}C NMR (100 MHz, MeOD) δ_C = 12.83 ($NCH_2\mathbf{CH_3}$), 35.82 ($N\mathbf{CH_2}CN$), 38.39 (NMe),

46.87 (N**C**$\mathbf{H_2}$$CH_3$), 97.43, 114.72, 115.44, 115.75, 128.92, 131.28, 131.63, 131.89, 132.65, 132.72, 135.88, 156.65, 157.26, 159.24, 170.71; HRMS m/z (ESI^+): measured: 495.27397 (M^+), calculated: 495.27545; IR: 2976, 2361, 1640, 1584, 1528, 1467, 1410, 1333, 1272, 1245, 1177, 1130, 1071, 1009, 976, 921, 821, 682 cm^{-1}.

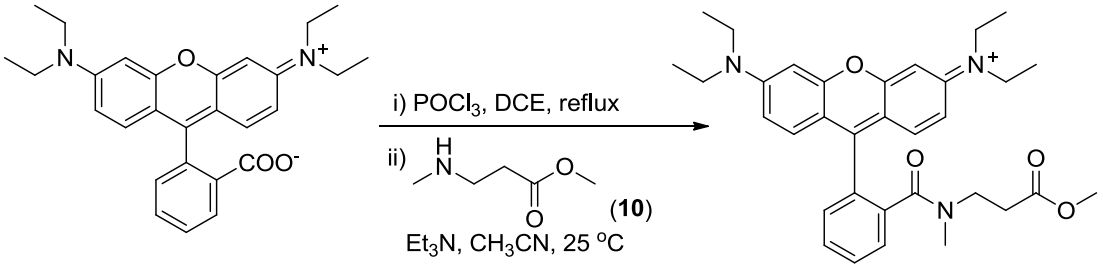
6.4.3.4 Activation of IME-rhodamine B (2)



9 **2**

CNM-dye **9** (1 equiv.) was suspended in MeOH (250 µL per 5 mg of dye) in a vial/flask under argon. 25% sodium methoxide in methanol (1 equiv.) was added to the solution and stirred at 25 °C for 16 h to give an 1:1 equilibrium mixture of CNM-dye **9** and IME-dye **2**. The solvent was removed by nitrogen followed by reduced pressure. The mixture was dissolved in DMSO (100 – 200 mg / mL) and used directly for labelling without further purification; R_f of **2** = 0.44 (1:9 MeOH/ CH_2Cl_2, v/v); LRMS m/z (ESI^+): measured: 527.3 (M^+), calculated: 527.3.

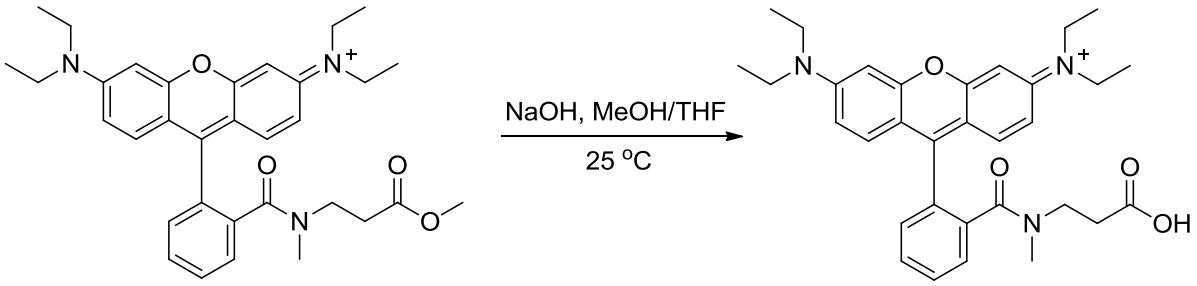
6.4.3.5 Rhodamine B ester (11)



Rhodamine B **11** (62%)

Rhodamine B (200 mg, 0.45 mmol) was dissolved in 1,2-dichloroethane (5 mL). Phosphoryl chloride (0.61 mL, 6.52 mmol) was added to the solution and stirred at 110 ^{o}C for 3 h. The solvent and the excess phosphoryl chloride were removed under reduced pressure. The residue was dissolved in dry CH_3CN (5 mL), followed by the addition of Et_3N (0.08 mL, 0.56 mmol) and amine **10** (215 mg, 1.84 mmol). After stirring at 25 ^{o}C for 16 h, the solvent was removed under reduced pressure. The residue was dissolved in CH_2Cl_2 (25 mL), washed with water (25 mL) and sodium bicarbonate (25 mL) and concentrated under reduced pressure. The resulting residue was further purified by silica column chromatography (CH_2Cl_2, then 5% MeOH in CH_2Cl_2) to give ester **11** as a purple solid (152 mg, 62%); R_f = 0.32 (1:9 MeOH/CH_2Cl_2, v/v); ^{1}H NMR (400 MHz, MeOD) δ_H = 1.30 (t, J 7.11, 12H, $NCH_2\mathbf{CH_3}$), 2.05 (t, J 6.68, 2H, $NCH_2\mathbf{CH_2}$), 2.86 (s, 3H, NMe), 3.41 (t, J 6.70, 2H, $N\mathbf{CH_2}CH_2$), 3.52 (s, 3H, $COOCH_3$), 3.68 (m, 8H, $N\mathbf{CH_2}CH_3$), 6.96 (d, J 2.43, 2H), 7.06 (dd, J 2.49, 9.55, 2H), 7.28 (d, J 9.54, 2H), 7.50 (m, 1H), 7.64 (m, 1H), 7.75 (m, 2H); ^{13}C NMR (100 MHz, MeOD) δ_C = 12.81 ($NCH_2\mathbf{CH_3}$), 31.92 ($NCH_2\mathbf{CH_2}$), 38.90 (NMe), 44.70 ($N\mathbf{CH_2}CH_2$), 46.88 ($N\mathbf{CH_2}CH_3$), 52.10 ($COO\mathbf{CH_3}$), 97.36, 114.73, 115.21, 128.77, 130.96, 131.44, 131.67, 133.27, 137.52, 157.23, 159.33, 170.76, 173.20; LRMS m/z (ESI^{+}): measured: 542.3 (M^{+}), calculated: 542.3.

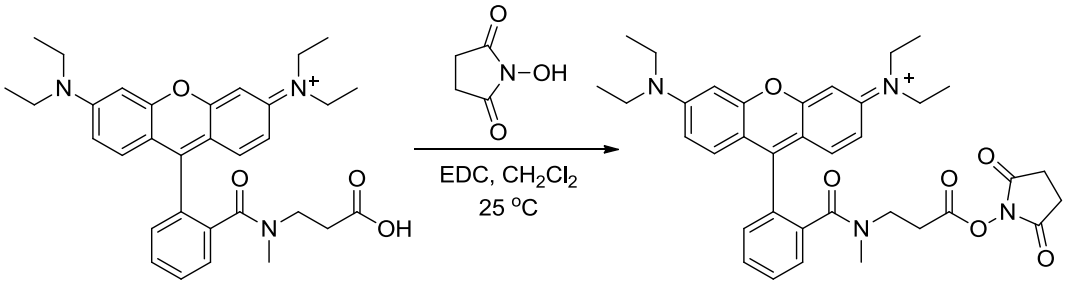
6.4.3.6 Rhodamine B carboxylic acid (12)



11 **12** (79%)

Ester **11** (137 mg, 0.25 mmol) was suspended in MeOH (3 mL) and THF (3 mL). 1 M NaOH (0.5 mL, 0.50 mmol) was added to the solution and stirred at 25 °C for 16 h. The solvent was then removed under reduced pressure. The resulting residue was purified by silica column chromatography (5% MeOH in CH_2Cl_2, then 10% MeOH in CH_2Cl_2, then 20% MeOH in CH_2Cl_2) to give acid **12** as a purple solid (105 mg, 79%); R_f = 0.39 (1:4 MeOH/CH_2Cl_2, v/v); ^{1}H NMR (400 MHz, MeOD) δ_H = 1.29 (t, J 7.08, 12H, $NCH_2\mathbf{CH_3}$), 1.94 (t, J 7.36, 1H, $NCH_2\mathbf{CH_2}$, rotamer), 2.29 (t, 1H, $NCH_2\mathbf{CH_2}$, rotamer), 2.67 (s, 1.5H, NMe, rotamer), 2.89 (s, 1.5H, NMe, rotamer), 3.40 (m, 2H, $N\mathbf{CH_2}CH_2$, rotamer x 2), 3.66 (m, 8H, $N\mathbf{CH_2}CH_3$), 6.93 (dd, J 2.44, 6.03, 2H), 7.05 (ddd, J 2.48, 4.77, 9.55, 2H), 7.26 (dd, J 5.76, 9.53, 2H), 7.47 (ddd, J 1.62, 5.97, 16.40, 1H), 7.69 (m, 3H); ^{13}C NMR (400 MHz, MeOD) δ_C = 12.81 ($NCH_2\mathbf{CH_3}$), 32.47 (NMe, rotamer), 34.81 ($NCH_2\mathbf{CH_2}$, rotamer), 37.31 ($NCH_2\mathbf{CH_2}$, rotamer); 38.63 (NMe, rotamer), 45.87 ($N\mathbf{CH_2}CH_2$, rotamer), 46.86 (NCH_2CH_3), 49.99 ($N\mathbf{CH_2}CH_2$, rotamer), 97.19, 97.42, 114.75, 115.19, 115.39, 127.56, 130.81, 131.29, 131.18, 133.39, 157.23, 159.28; HRMS m/z (ESI$^+$): measured: 528.2838 (M^+), calculated: 528.2857; IR: 3382, 1584, 1467, 1410, 1335, 1273, 1247, 1178, 1130, 1071 cm^{-1}.

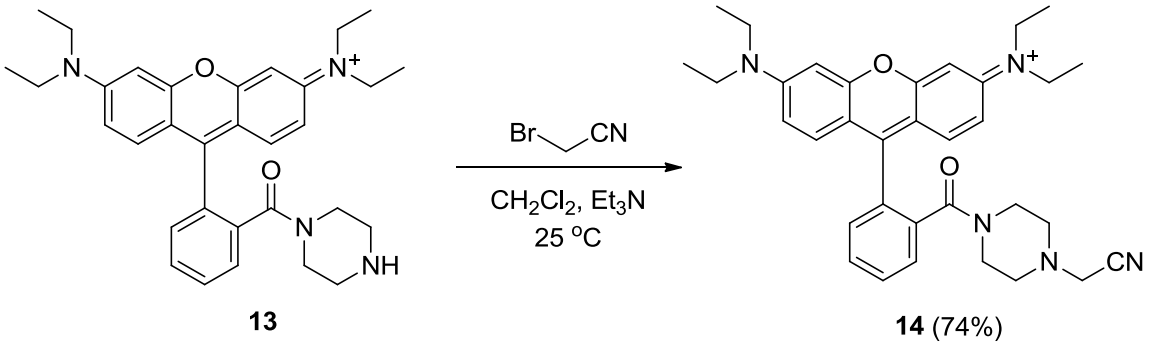
6.4.3.7 Rhodamine B derivative (4)



12 **4** (72%)

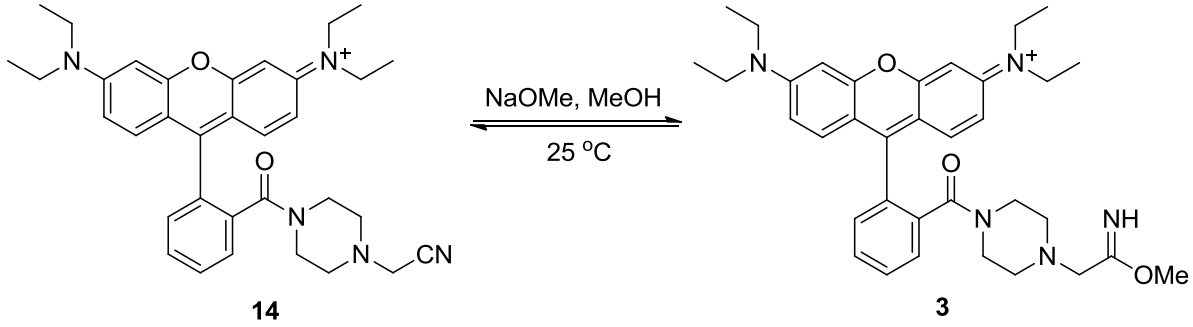
Acid **12** (22 mg, 0.042 mmol), N-hydroxysuccinimide (6.5 mg, 0.057 mmol) and 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) (22.7 mg, 0.118 mmol) were dissolved in dry CH_2Cl_2 (0.5 mL). The reaction mixture was stirred at 25 ^oC for 1.5 h. The resulting mixture was diluted with CH_2Cl_2 (10 mL) and washed with H_2O (20 mL). The aqueous layer was further extracted with CH_2Cl_2 (10 mL) twice. The combined organic layer was washed with 1 N HCl (20 mL), saturated $NaHCO_3$ (20 mL) and brine (20 mL). The organic layer was dried (Na_2SO_4), filtered and concentrated under reduced pressure to give NHS-dye **4** as a purple solid (18.9 mg, 72%). The product was dissolved in DMSO (100 – 200 mg / mL) and used directly for labelling without further purification; R_f = 0.74 (1:9 MeOH/CH_2Cl_2, v/v); ^1H NMR (400 MHz, MeOD) δ_H = 1.32 (t, J 7.11, 12H, $NCH_2\mathbf{CH_3}$), 2.52 (t, J 6.59, 2H, $NCH_2\mathbf{CH_2}$), 2.85 (s, 4H, NHS), 2.98 (s, 3H, NMe), 3.60 (m, 10H, $N\mathbf{CH_2}CH_3$ & $N\mathbf{CH_2}CH_2$), 6.78 (d, J 2.30, 2H), 6.92 (dd, J 2.25, 9.55, 2H), 7.23 (m, 2H), 7.34 (m, 1H), 7.65 (m, 3H); ^{13}C NMR (100 MHz, MeOD) δ_C = 12.69 ($NCH_2\mathbf{CH_3}$), 25.72 (NHS), 28.86 ($NCH_2\mathbf{CH_2}$), 38.86 (NMe), 43.77 ($N\mathbf{CH_2}CH_2$), 46.21 ($N\mathbf{CH_2}CH_3$), 96.54, 103.00, 104.59, 106.37, 114.18, 119.12, 124.47, 128.08, 130.30, 132.02, 155.78, 157.90, 169.25; HRMS m/z (ESI^+): measured: 625.3032 (M^+), calculated: 625.3021; IR: 2922, 2361, 1734, 1585, 1467, 1410, 1335, 1273, 1247, 1178, 1131, 1070 cm^{-1}.

6.4.3.8 Rhodamine B derivative (14)



Piperazine **13** (kindly provided by Dr. Yamamoto) (50 mg, 97.7 μmol)[25] was dissolved in CH_2Cl_2 (2.5 mL) in a 25 mL RB flask and cooled to 0 °C. Et_3N (24 μL, 171 μmol) was added to the solution and stirred at 0 °C for 5 min. Bromoacetonitrile (15 mg, 122 μmol) was added and the reaction mixture was stirred at 0 °C for 5 min and then at 25 °C for 1.5 h. Another portion of bromoacetonitrile (12 mg, 97.7 μmol) was added and stirred at 25 °C for 1 h. The reaction mixture was washed with saturated $NaHCO_3$ (10 mL x 2). The organic layer was dried (Na_2SO_4), filtered and concentrated under reduced pressure. The residue was purified by silica column chromatography (CH_2Cl_2, then 10% MeOH in CH_2Cl_2, then 20% MeOH in CH_2Cl_2). The fractions containing the product were combined and evaporated under vacuum to give CNM-dye **14** as a purple solid (40 mg, 74%); R_f = 0.71 (2:8 MeOH/ CH_2Cl_2, v/v); ^{1}H NMR (400 MHz, MeOD) δ_H = 1.31 (t, J 7.13, 12H, $NCH_2\mathbf{CH_3}$), 2.23 (m, 4H, $C(O)NCH_2\mathbf{CH_2}N$), 3.45 (m, 4H, $C(O)N\mathbf{CH_2}CH_2N$), 3.61 (s, 2H, $\mathbf{CH_2}CN$), 3.69 (m, 8H, $N\mathbf{CH_2}CH_3$), 6.96 (d, J 2.48, 2H), 7.08 (dd, J 2.49, 9.56, 2H), 7.28 (m, 2H), 7.53 (m, 1H), 7.65 (m, 1H), 7.76 (m, 2H); ^{13}C NMR (100 MHz, MeOD) δ_C = 12.82 ($NCH_2\mathbf{CH_3}$), 42.29 ($C(O)N\mathbf{CH_2}CH_2N$), 45.91 ($\mathbf{CH_2}CN$), 46.91 ($N\mathbf{CH_2}CH_3$), 48.19 ($C(O)NCH_2\mathbf{CH_2}N$), 52.44 ($C(O)NCH_2\mathbf{CH_2}N$), 97.48, 114.85, 115.26, 115.47, 128.89, 131.15, 131.32, 131.60, 132.06, 133.23, 136.70, 156.77, 157.28, 159.25, 169.42; HRMS m/z (ESI^+): measured: 550.31784 (M^+), calculated: 550.31765.

6.4.3.9 Activation of IME-rhodamine B (3)



14 **3**

CNM-dye **14** (1 equiv.) was suspended in MeOH (250 µL per 5 mg of dye) in a vial/flask under argon. 25% sodium methoxide in methanol (1 equiv.) was added to the solution and stirred at 25 °C for 24 h to give an 1:1 equilibrium mixture of CNM-dye **14** and IME-dye **3**. The solvent was removed by nitrogen followed by reduced pressure. The mixture was dissolved in DMSO (100 – 200 mg / mL) and used directly for labelling without further purification; R_f of **3** = 0.49 (1:9 MeOH/ CH_2Cl_2, v/v); LRMS m/z (ESI^+): measured: 582.3 (M^+), calculated: 582.3.

6.4.4 Coupling of rhodamine B derivatives with HEL

Rhodamine B derivative (**1**-**5**) (100 mg / mL in DMSO) (60 or 120 equiv.) was added to HEL solution (1 mg / mL in 0.6 M TAPS, pH 8.5). For dye **4** and **5**, the reactions were also performed in phosphate buffer (50 mM NaPi, pH 8.0) for testing purpose. The mixture was shaken at 25 °C for various time point and purified by Vivaspin® (5,000 MW) before analysing with LCT mass spectrometry. The labelled proteins were also analysed by SDS-PAGE and visualised *via* Coomassie Blue staining and fluorescence scanning.

6.4.5 LC-MS analysis

Proteins were analysed by LC-MS without further treatment. The mass spectra were reconstructed from the ion series and deconvoluted using the MaxEnt1 algorithm in the MassLynx software (v 4.1 from Waters).

6.5 References

1. Seisenberger, G. et al. Real-time single-molecule imaging of the infection pathway of an adeno-associated virus. *Science* **294**, 1929-32 (2001).
2. Suomalainen, M. et al. Microtubule-dependent plus- and minus end-directed motilities are competing processes for nuclear targeting of adenovirus. *J Cell Biol* **144**, 657-72 (1999).
3. Ayala-Nunez, N.V., Wilschut, J. & Smit, J.M. Monitoring virus entry into living cells using DiD-labeled dengue virus particles. *Methods* **55**, 137-43 (2011).
4. Gardner, C.L. et al. Deliberate attenuation of chikungunya virus by adaptation to heparan sulfate-dependent infectivity: a model for rational arboviral vaccine design. *PLoS Negl Trop Dis* **8**, e2719 (2014).
5. Kobayashi, Y. & Suzuki, Y. Compensatory evolution of net-charge in influenza A virus hemagglutinin. *PLoS One* **7**, e40422 (2012).
6. Shayakhmetov, D.M. & Lieber, A. Dependence of adenovirus infectivity on length of the fiber shaft domain. *J Virol* **74**, 10274-86 (2000).
7. Sunasee, R. & Narain, R. in Chemistry of Bioconjugates 1-75 (John Wiley & Sons, Inc., 2014).
8. Schiessl, W.C. et al. Experimental and Theoretical Approaches to the Protonation of Thiourea: A Convenient Nucleophile in Coordination Chemistry Revisited. *Zeitschrift für anorganische und allgemeine Chemie* **631**, 2812-2819 (2005).
9. Riddick, J.A., Bunger, W.B. & Weissberger, A. Techniques of Chemistry. Organic Solvents. Physical Properties and Methods of Purification (Wiley Interscience, 1970).
10. Lee, Y.C., Stowell, C.P. & Krantz, M.J. 2-Imino-2-methoxyethyl 1-thioglycosides: new reagents for attaching sugars to proteins. *Biochemistry* **15**, 3956-3963 (1976).
11. Pearce, O.M.T. et al. Glycoviruses: Chemical glycosylation retargets adenoviral gene transfer. *Angewandte Chemie-International Edition* **44**, 1057-1061 (2005).

12. Robinson, M.A. et al. LEAPT: Lectin-directed enzyme-activated prodrug therapy. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14527-14532 (2004).
13. Beija, M., Afonso, C.A.M. & Martinho, J.M.G. Synthesis and applications of Rhodamine derivatives as fluorescent probes. *Chemical Society Reviews* **38**, 2410-2433 (2009).
14. Snare, M.J., Treloar, F.E., Ghiggino, K.P. & Thistlethwaite, P.J. The photophysics of rhodamine B. *Journal of Photochemistry* **18**, 335-346 (1982).
15. Nguyen, T. & Francis, M.B. Practical Synthetic Route to Functionalized Rhodamine Dyes. *Organic Letters* **5**, 3245-3248 (2003).
16. Adamczyk, M. & Grote, J. EFFICIENT FLUORESCEIN SPIROLACTAM AND BIS-SPIROLACTAM SYNTHESIS. *Synthetic Communications* **31**, 2681-2690 (2001).
17. Chalker, J.M., Lercher, L., Rose, N.R., Schofield, C.J. & Davis, B.G. Conversion of Cysteine into Dehydroalanine Enables Access to Synthetic Histones Bearing Diverse Post-Translational Modifications. *Angewandte Chemie International Edition* **51**, 1835-1839 (2012).
18. Sendai, M. & Kishimoto, S. 47 pp. (Takeda Chemical Industries, Ltd., Japan . 1989).
19. Cheguillaume, A. et al. Solution synthesis and characterization of aza-beta(3)-peptides (N(alpha)-substituted hydrazino acetic acid oligomers). *J Org Chem* **66**, 4923-9 (2001).
20. Pospíšil, J. & Potáček, M. Microwave-assisted solvent-free intramolecular 1,3-dipolar cycloaddition reactions leading to hexahydrochromeno[4,3-b]pyrroles: scope and limitations. *Tetrahedron* **63**, 337-346 (2007).
21. Wang, Z., Li, D., Huang, A. & He, G. 11pp. (Nanjing Hengsheng Pharmaceutical Factory, Peop. Rep. China; Nanjing Lifenergy R & D Co., Ltd. . 2011).
22. EHRHARDT, C., MCQUIRE, Leslie Wighton , RIGOLLIER, Pascal, ROGEL, Olivier, SHULTZ, Michael, TOMMASI, Ruben Alberto. (2009).
23. Ototake, N., Nakamura, M., Dobashi, Y., Fukaya, H. & Kitagawa, O. Highly selective stereodivergent synthesis of separable amide rotamers, by using Pd chemistry, and their thermodynamic behavior. *Chemistry* **15**, 5090-5 (2009).
24. Mitronova, G.Y. et al. New Fluorinated Rhodamines for Optical Microscopy and Nanoscopy. *Chemistry – A European Journal* **16**, 4477-4488 (2010).
25. Parsons, T.B. et al. Optimal Synthetic Glycosylation of a Therapeutic Antibody. *Angewandte Chemie International Edition* **55**, 2361-2367 (2016).
26. van Kasteren, S.I. et al. Glyconanoparticles allow pre-symptomatic in vivo imaging of brain disease. *Proceedings of the National Academy of Sciences* **106**, 18-23 (2009).

CHAPTER

7 Labelling of viruses with IME-functionalised fluorescent dyes

7.1 Introduction

7.1.1 Adenovirus as a model for testing the synthesised dyes

Human adenovirus (AV) has an icosahedral structure with 12 fiber proteins extending from each vertex.[1] It readily infects cell types possessing coxsackie adenovirus receptors (CARs) through the interactions with the virus fiber protein.[1, 2] As AV has been successfully labelled with IME-sugars in the past,[3] it can act as an excellent virus model for testing the synthetic IME-dyes described in Chapter 6 (**Figure 7.1**). Each AV particle contains around 22,000 reactive lysine residues, in which most of them are located in the hexon capsid proteins. The virus is known to be stable in buffer containing a small amount of DMSO,[4] which is particularly useful in this case as the stock solution of the CNM- and IME-dye mixture is in DMSO.

An attenuated strain of human AV type 5 was used here in which the virus cannot replicate itself in cells. As a result, it can be safely handled in a normal laboratory. Furthermore, the strain contains a gene encoding the green fluorescence protein (GFP), which can be transfected to the target cells after successful infection. This virus infectivity can thus be estimated based on the fluorescence signal given out by the infected cells.

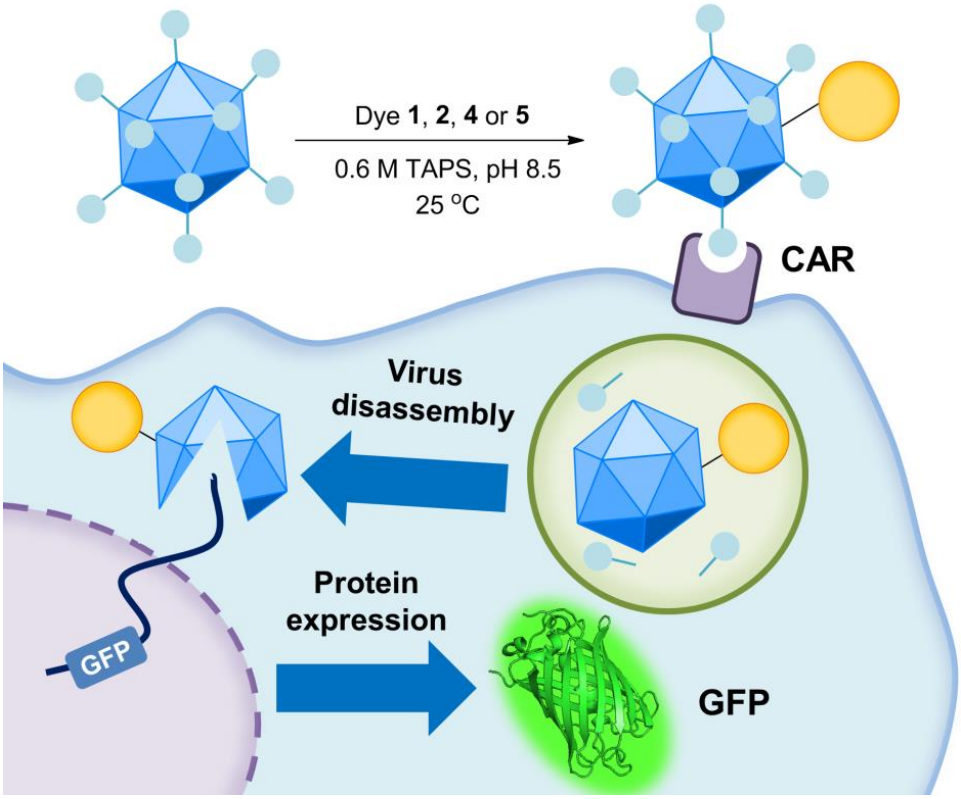


Figure 7.1: AV possessing a GFP gene was labelled with rhodamine B derivatives. The labelled virus was then incubated with CAR presenting A549 epithelial cells. GFP is expressed once the cells are successfully infected.

7.1.2 Dengue virus: a more complicated virus model

As mentioned in Chapter 1, dengue virus (DENV) is an enveloped (+)ssRNA virus belonging to the *Flaviviridae* family. Since the virus contains a lipid membrane, the labelling is expected to be very challenging. As DENV can cause life-threatening dengue fever (DF), all experiments must be performed in a laboratory specially equipped for bio-safety level 3. To study the antibody-dependent enhancement (ADE) phenomenon of the virus, one of the most interesting ways would be to investigate the difference of the virus infection pathway in the absence or presence of antibodies by microscopy (**Figure 1.10**). It is crucial to make sure that the virus infectivity does not change significantly after fluorescence labelling. The method described in this chapter is a possible solution to this problem.

In the past, the surface proteins and lipid of DENV have been successfully labelled with

Alexa Fluor® NHS esters[5] and 1,1'-dioctadecyl-3,3,3',3'-tetramethylindo-dicarbocyanine, 4-chlorobenzenesulfonate salt (DiD)[6-8] respectively. The former is believed to be a better choice as the fluorescence of DiD is greatly quenched when the dyes are closely packed on the virus membrane before the membrane fusion process, which renders the tracking of viruses before fusion difficult.[6-8] However, the coupled Alexa Fluor® dyes reduced the virus infectivity 30 folds, making them ineffective probes for virus infection studies.[5] Thus, there is an urgent need for the development of new dyes that have a minimal effect on virus infectivity.

7.2 Results and Discussion

7.2.1 Labelling of adenovirus with rhodamine B derivatives

The reaction condition used for HEL was applied to AV directly (**Figure 7.1**). As AV is a real virus with a complicated structure, it cannot easily be analysed by mass spectrometry like other protein models reported in previous chapters. The simplest way to study the reactivity of dyes is to observe the mass shift of the hexon capsid protein (~105 kDa), which is the most abundant protein in AV, on a SDS-PAGE gel before and after labelling. The proteins on the gel were visualised by silver staining as the virus stock solution was very dilute (~1 nM).

It was found that at least 1000 equiv. of dyes was required to give a significant mass shift on the gel. Interestingly, only AVs labelled with IME dye **1** or NHS dye **4** showed significant mass shifts (**Figure 7.2**). Also, the reactivity of **4** towards AV was similar to that of **1**, which was not the case in HEL. This revealed that the accessible reaction sites in a complicated virus structure can be very different from that of simple proteins. The relatively short linker

on dye **2** may account for its inert reactivity. Since AV has a more complicated structure than HEL, it is reasonable to suggest that the reactive group of the dye must be readily accessible in order to yield labelled virus. In addition, the ITC dye **5** was also unreactive towards AV. It is unknown whether this is due to the inaccessibility of the conjugating group or the inert reactivity of the group towards the virus, as was the case with the Qβ VLPs mentioned in Chapter 5.

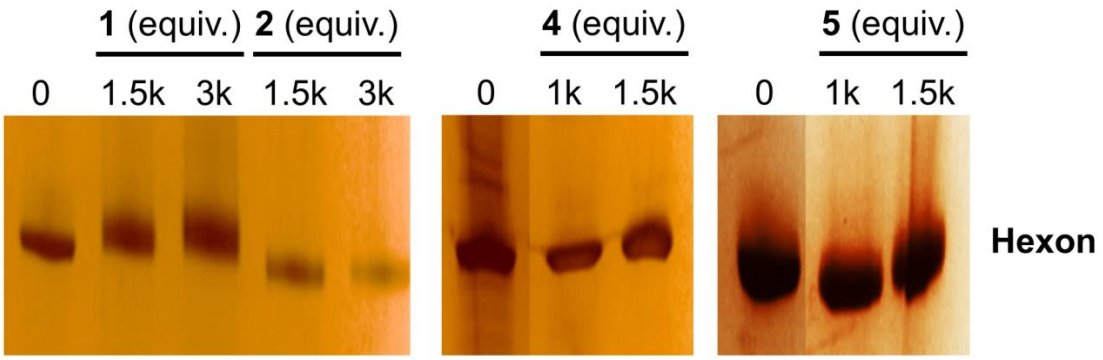


Figure 7.2: Silver-stained reducing SDS-PAGE gels showing the labelling of AV with different equiv. of rhodamine B derivatives. Successful labelling was indicated by the mass shift of the hexon protein (~105 kDa).

Since a large amount of dyes was added for the reaction, it is crucial to optimise the purification process so that clean labelled virus can be obtained with high recovery yield. The concentration of the labelled virus was determined by PicoGreen® assay. This assay works by measuring the amount of DNA in the solution after destroying the viral coat with SDS. By measuring the fluorescence intensity resulting from the DNA binding PicoGreen® dye, the amount of virus in an unknown sample can be calculated by comparing the value with a sample with known concentration.

In general, the recovery yield after virus purification decreased when more dyes were added (**Figure 7.3**). Although dialysis can greatly reduce the loss of virus, it cannot remove the dyes as efficiently as running a column. For Bio-Spin™ P-6 columns, both the viruses

and free dyes were lost after purification. G25 columns were found to be the best among the three methods in which a relatively clean sample with reasonable yield was obtained. The diluted virus solution obtained from the column was concentrated by passivated Vivaspin® . Passivated Vivaspin® was prepared by soaking the device in PBS with 1% BSA and previous report has shown that this can increase the protein recovery yield.[9] In addition, it was found that a better recovery yield can be obtained when virus is concentrated at 37 ^{o}C.

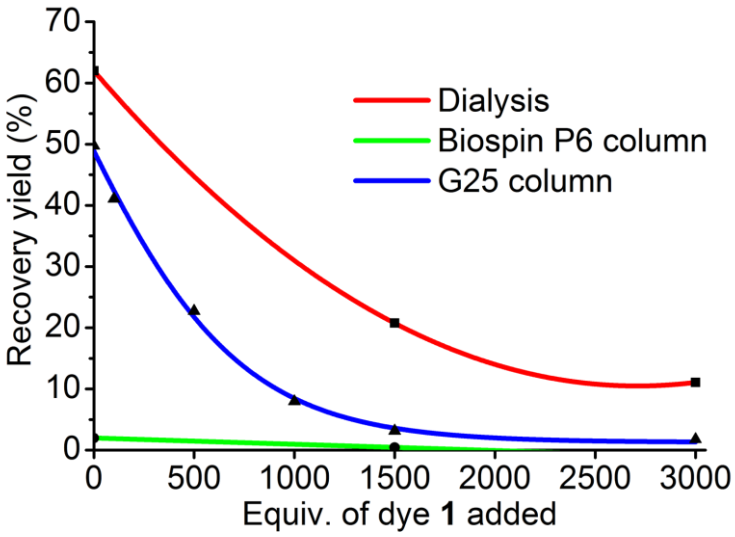
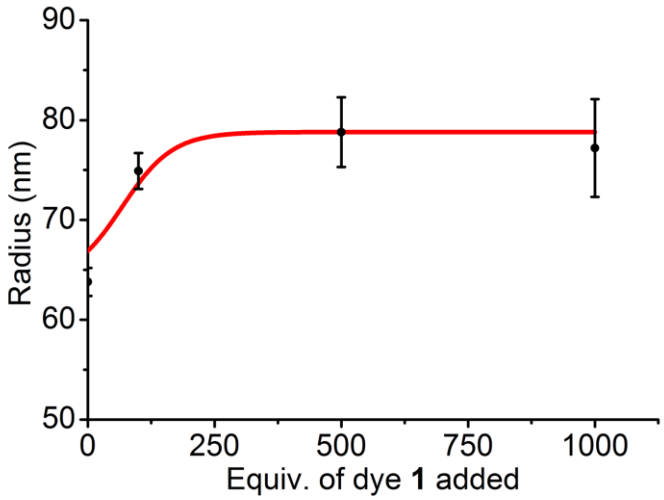


Figure 7.3: Graph showing the recovery yield determined by PicoGreen® assay after the purification of AV labelled with different equivalent of dye **1**. The virus purified by G25 column was further concentrated by passivated Vivaspin® before measurement.

The size of unlabelled AV measured by DLS (radius = 63.8 ± 1.4 nm) was comparable to the value previously reported by photon correlation spectroscopy (diameter = 120 ± 6 nm) (**Figure 7.4**).[3] For the labelled viruses, the radius increased significantly even when only 100 equiv. of dye was added. The size of virus reached a maximum when 500 equiv. of dye was added. Virus samples labelled with more than 1000 equiv. of dye were too dilute to be properly analysed (**Figure 7.3**).



1 (equiv.)	Radius (nm)
0	63.8 ± 1.4
100	74.9 ± 1.8
500	78.8 ± 3.5
1000	77.2 ± 4.9

Figure 7.4: Radius of AV labelled with different equiv. of dye **1** as measured by dynamic light scattering.

To test whether the virus labelled by the IME dye **1** was infective, the labelled AV was incubated with A549 epithelial cells that are known to express CAR on the cell surface. As shown by the microscopy images (**Figure 7.5**), the labelled viruses were transported to the nucleus (stained with DAPI) quickly after the incubation. Most of the viruses became associated with the nucleus after 20 min of infection. This matched well with the infection pathway shown in the literature.[10] The infection efficiency of AV relies on the effective transportation of AV genome to the cell nucleus. As expected, fluorescence signal from GFP was clearly observed after 24 h of virus treatment, indicating successful infection.

7.2.2 Fluorescence properties of the synthetic dyes

In order to compare the effect of different dyes on virus infectivity, it is crucial to make sure that the viruses are labelled with the same level of dyes. This cannot be simply assessed based on the equivalent of dyes added as the reactivity of different dyes may not be the same. One possible way to estimate the level of labelling is to measure the fluorescence intensity of different virus samples. Assuming the instrumental parameters are constant, the fluorescence

intensity of a fluorophore is proportional to the product of extinction coefficient (ε), quantum yield (Φ) and sample concentration.

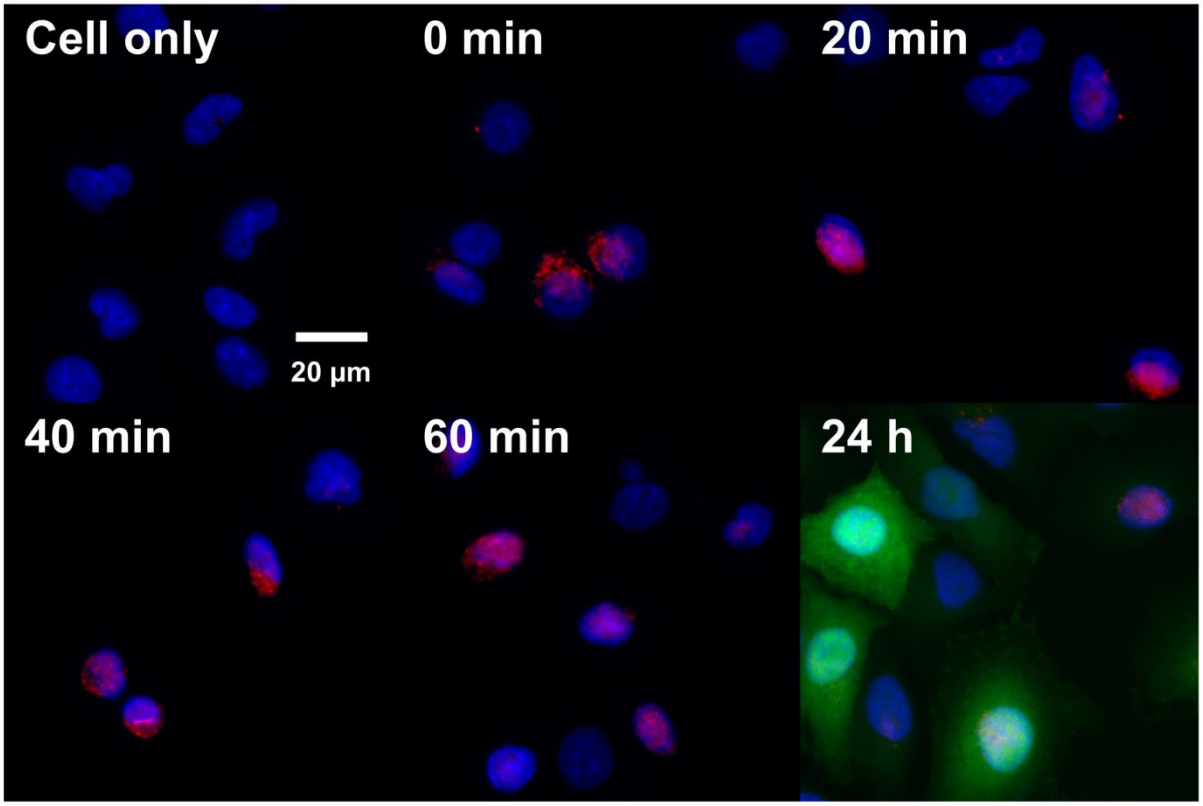
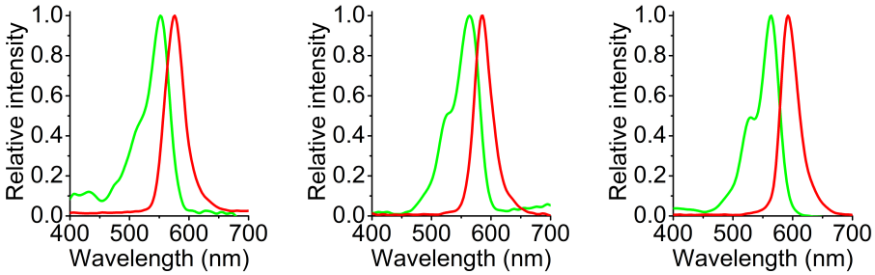


Figure 7.5: Microscopy images showing the infection of A549 cells with AV-dye **1** at different time points. Red: rhodamine B; Green: GFP; Blue: DAPI.

Different fluorescence properties of the screened dyes were measured and compared with unmodified rhodamine B (**Table 7.1**). A redshift of around 10 nm was observed in both excitation and emission spectra for dye **1** and **4**, yet the Stokes shift of **4** is slightly larger than that of **1**. The spectra also looked very similar after the dyes were coupled to HEL (refer to **Figure S.7** in Appendix), indicating that the synthesised dyes can retain their fluorescence properties after protein coupling.

Table 7.1: Fluorescence properties of rhodamine B derivatives in PBS (pH 7.5)

	Rho B	Dye **1**	Dye **4**
Absorption & emission † spectra			
Excitation max (nm)	552	564	564
Emission max (nm)	576	586	592
Stokes shift (nm)	24	22	28
ε (cm^{-1} M^{-1})			
at max	42733 ± 367	23700 ± 321	24233 ± 88
at 514 nm	25933 ± 273	12400 ± 153	14100 ± 100
Φ †	0.392 ± 0.005	0.233 ± 0.001	0.172 ± 0.002
Relative fluorescence	4.19	1.19	1.00

† excited at 514 nm.

The extinction coefficient of a fluorophore can be obtained by plotting a graph of absorbance versus concentration. The fluorescence quantum yield (Φ), on the other hand, was measured based on the comparative method of Williams *et al*. in which fluorophores with well characterised Φ were used as references.[11] Here, fluorescein and rhodamine B were used. Although the extinction coefficients for both **1** and **4** are similar, **1** has a slightly higher quantum yield than **4**. The relative fluorescence of **1** is thus higher than **4** under the same sample concentration. The measured fluorescence properties of unmodified rhodamine B are similar to those reported in the literature,[12] suggesting that the experimental conditions used were working fine.

7.2.3 Comparison of the infectivity of viruses labelled with IME- or NHS-dye

Next, the amount of dyes used for virus labelling was optimised. It is crucial to maximise the amount of dyes coupled so that the fluorescence signal is strong enough to be detected. Nonetheless, the infectivity of the virus would be reduced when the virus surface is crowded with dyes. To optimise the amount of dye used, the GFP fluorescence intensity of A549 cells infected by AV labelled with different equiv. of **1** was measured. By using flow cytometry, the percentage of GFP-positive cells can be counted (**Figure 7.6**). As expected, the infectivity of AV dropped significantly when more than 1000 equiv. of **1** was added. In addition, the recovery yield issue further restricted the amount of dye used to be around 1000 equiv. (**Figure 7.3**).

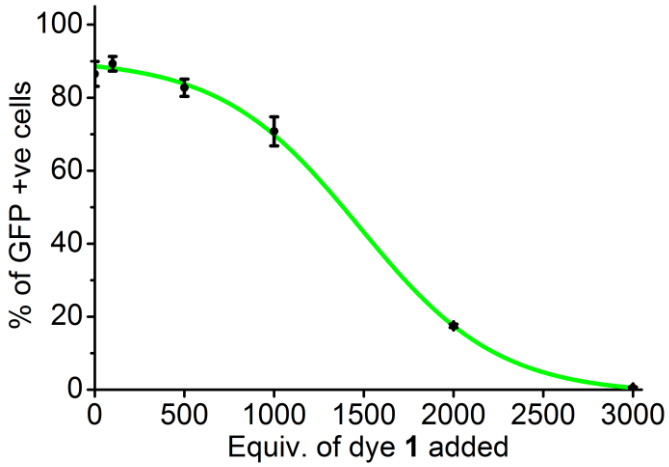


Figure 7.6: Graph showing the percentage of GFP-positive A549 epithelial cells after 24 h infection of AV (old batch) labelled with different equiv. of dye **1**. The graph was plotted based on the histograms obtained from flow cytometry (refer to **Figure S.8** in Appendix).

A new batch of AV was then labelled with dye **1** at the range of 1000 equiv. or with dye **4** at a wider range for comparison. By rendering the labelled AV samples at the same concentration, their fluorescence intensity was measured. As calculated from ε and Φ, the fluorescence intensity of AV-dye **1** should be around 1.19 time of AV-dye **4** under the same level of labelling (**Table 7.1**). As a result, two comparable pairs (1000 equiv. of **1** versus 1000 equiv. of **4**; 1250 equiv. of **1** versus 1500 equiv. of **4**) were identified (**Figure 7.7**).

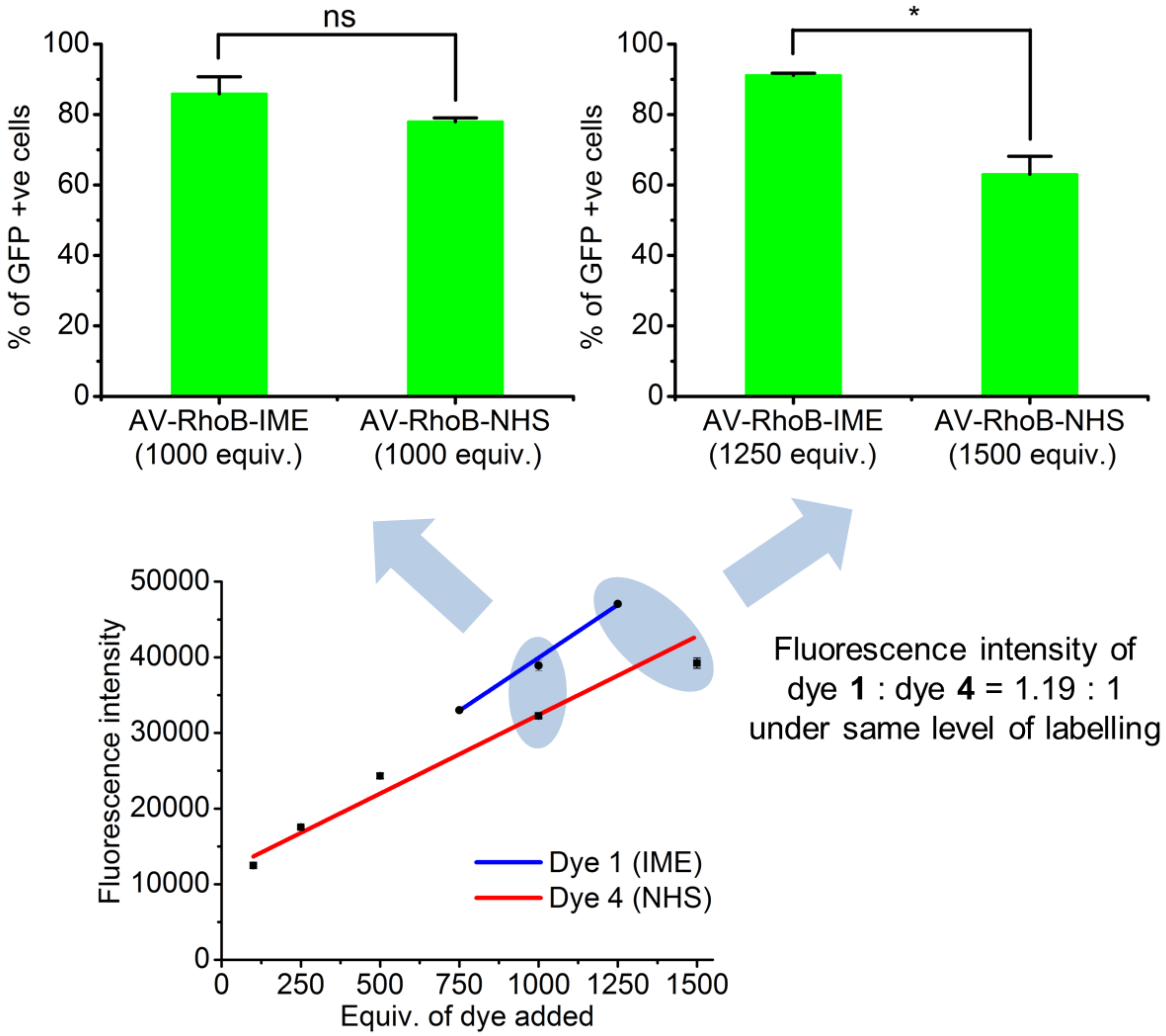


Figure 7.7: Fluorescence intensity of AV (new batch) labelled with different equiv. of dye **1** or **4** at the same concentration. The two comparable pairs are highlighted in blue. The infectivity of AV labelled with the same level of dye **1** or **4** was then measured based on the percentage of GFP-positive cells. The infectivity graph was plotted based on the histograms obtained from flow cytometry (refer to **Figure S.9** in Appendix).

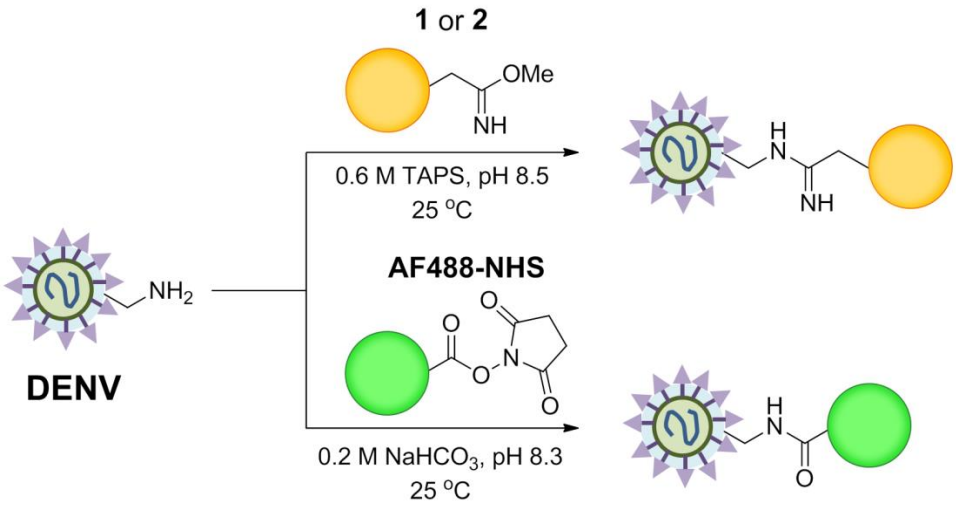
The GFP fluorescence and hence the infectivity of the corresponding labelled AV samples was then measured and compared. An increase of infectivity was observed for the new virus batch when compared with the old one (**Figure 7.6**) even the same amount of dye **1** was used. It is probably because the measurement of the new virus batch was performed immediately after the labelling reaction while the old batch was stored at 4 ^{o}C for 2 weeks before the measurement. Interestingly, no significant difference in infectivity was observed when AV was labelled with 1000 equiv. of **1** or **4**. However, **1** was able to retain the virus infectivity

better than **4** when an increasing amount of dye was added (**Figure 7.7**). This is indeed reasonable as a low level of labelling may not have a great effect on virus infectivity regardless of the dye derivative used. When the virus surface is covered by more and more dyes, the change of net charge would become more significant, and thus more dependent on the type of dye used.

The above results revealed that the IME group indeed helped retaining the infectivity of labelled virus much better than the NHS group, which was consistent with the hypothesis made at the beginning of this chapter. An interesting point of rhodamine B dyes is that it carries a positive charge at physiological pH. When the virus was labelled with rhodamine B-NHS, although the positive charge of the nitrogen atom on the lysine residue became neutral, the overall surface charge was unchanged. On the other hand, a net positive charge was added to the virus surface when the virus was labelled with rhodamine B-IME. It is possible that the charge on the dye is far away from the virus surface proteins involved in cell infection, thus it has a little effect on virus infectivity. The lysine residues, on the other hand, can be located in positions vital for virus infectivity. More dye derivatives will be tested in the future to further support this hypothesis.

7.2.4 Labelling of DENV with rhodamine B derivatives

So far only some preliminary studies have been performed of DENV labelling. DENV-2 was labelled with the IME-dye **1** and **2** using the same conditions as that of HEL (**Scheme 7.1**). A control experiment was also performed using the commercially available Alexa Fluor 488® NHS (AF488-NHS) dye with the previously published protocols.[5]



Scheme 7.1: DENV labelled with IME-dye **1** / **2** or AF488-NHS dye under different conditions.

To analyse the labelled viruses with confocal microscopy, they were first fixed on glass coverslips with poly-L-lysine. The primary antibody 3H5 for recognising the DENV E protein was added, followed by secondary antibodies conjugated to different fluorescent probes. In this case, secondary antibodies with AF-488 and AF-568 probes were used for DENV labelled with rhodamine B derivatives and AF488-NHS respectively. If the virus is successfully labelled, co-localisation of the two channels (i.e. conjugated dye and fluorescent antibody) will be observed in the same spot.

Nonetheless, as shown by the microscopy results (**Figure 7.8A-C**), only dots representing the antibodies (green for dye **1**/**2**; red for AF488-NHS) were observed for all the three dyes tested. Although a few dots representing the dyes were seen, they did not really co-localise with the antibodies. To confirm if the secondary antibodies were working properly, a control experiment with both the two secondary antibodies (i.e. goat anti-mouse AF488 and goat anti-mouse AF568) added was performed (**Figure 7.8D**). In this case, co-localisation of spots was observed, indicating that it was probably a problem with dye labelling.

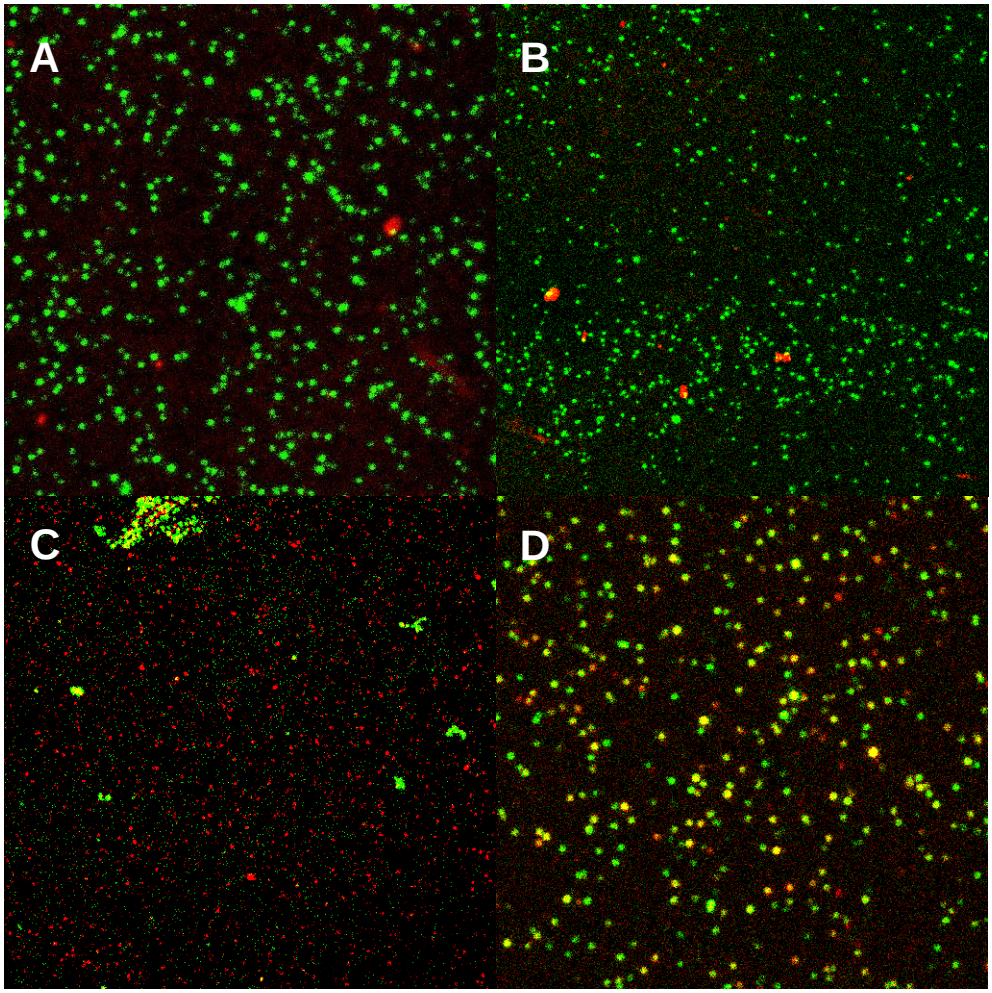


Figure 7.8: Microscopy images of (A) DENV labelled with AF488 + 3H5 + goat anti-mouse AF568 Ab; (B) DENV labelled with dye **1** + 3H5 + goat anti-mouse AF488 Ab; (C) DENV labelled with dye **2** + 3H5 + goat anti-mouse AF488 Ab and; (D) unlabelled DENV + 3H5 + goat anti-mouse AF488 Ab + goat anti-mouse AF568 Ab. Green: 488 channel; Red: 568 channel.

Assuming the dots representing the dyes were really labelled viruses, one of the possible reasons for explaining why they were not bound with antibodies would be the blocking of surface epitopes by the dyes. Two other different primary antibodies (anti-prM 2H2 and anti-E 4G2) were tested but the same results were observed (data not shown). These again proved that the labelling reactions were not successful.

To get a better idea about the virus stock (which was obtained from my collaborator), it was analysed by SDS-PAGE with silver staining (**Figure 7.9**). Surprisingly, a lot of bands

other than the virus proteins were observed, indicating that the sample had a lot of impurities, most probably coming from the culture medium and the mosquito host cell used for virus production. The intense band at around 55 kDa was originally thought to be the virus E protein (theoretical mass = 53 kDa). However, when the gel was analysed by western blot with the anti-E 3H5, no band was observed (lane S).

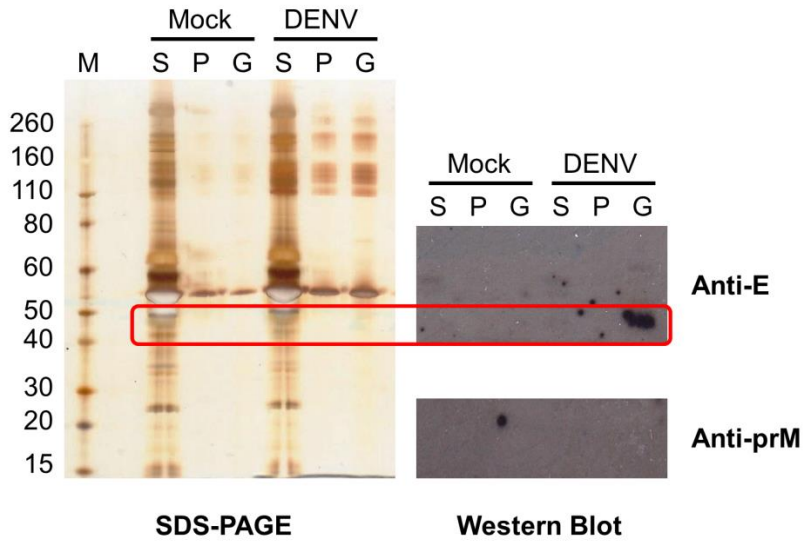


Figure 7.9: Silver-stained non-reducing SDS-PAGE of the DENV stock before and after purification. The gel was also analysed by ECL-visualised western blot with anti-E 3H5 and anti-prM 2H2 antibodies. Mock: control host cell extract without virus infection; M: protein marker; S: DENV stock; P: DENV after ultra-pelleting; G: DENV after ultra-pelleting and sucrose gradient purification.

To ensure that it was not a problem with virus concentration, the virus stock was purified and concentrated by ultra-pelleting and sucrose gradient to see if any bands can be observed in the western blot. Interestingly, the E protein band was observed after sucrose gradient, proving that the 3H5 antibody was working properly (lane G). However, the band representing prM protein was still not seen. Furthermore, the intense band at around 55 kDa on the SDS-PAGE was actually not the E protein after comparing the gel with the western blot.

All these results revealed that the concentration of DENV was still extremely low even after ultra-pelleting. Although most of the impurities were removed after purification, the amount of DENV was still incredibly tiny compared with that of other impurities. This explains why the viruses were not labelled, as most of the dyes were consumed by the protein impurities present in the sample. Future work will focus on further purifying and concentrating the virus samples.

7.3 Conclusion

In summary, the synthesised IME rhodamine B derivative **1** was shown to react readily with AV surface protein while retaining the virus infectivity to a greater extent compared to the NHS derivative. Also, this phenomenon was only observed at high level of dye labelling, which matches well with our hypothesis about the effect of surface net charge on virus infectivity. Nevertheless, more dye derivatives should be tested in the future in order to make a solid conclusion about this. In addition, it is believed that the dyes developed can be applied to other virus models for infection studies. Dengue virus, a more complicated virus model with a lipid membrane, was tested but no labelling was observed, likely due to the impurities present in the sample.

While rhodamine B was employed as a simple fluorophore model in this study, it is possible to functionalise other advanced fluorescence dyes like the Alexa Fluor® series with the IME linker proposed. This can further improve the fluorescence properties and the water solubility of the dye, which are essential for rendering our proposed method applicable to more complicated systems.

7.4 Experimental Section

7.4.1 General procedures

Chemicals were purchased from commercial suppliers and used without purification. NuPAGE® MOPS buffer and 4-12% Bis-Tris gel were purchased from Life Technologies™. Protein markers used were Perfect Protein™ Marker 10-225 kDa purchased from Novagen® . PD SpinTrap™ G-25, MiniTrap™ G-25, MidiTrap™ G-25 and PD10 desalting columns were purchased from GE Healthcare Life Sciences. Micro Bio-Spin™ P-6 column and Bio-Rad Clarity™ ECL reagent were purchased from Bio-Rad. SilverXpress® silver staining kit and Quant-iT™ PicoGreen® dsDNA assay kit were purchased from Life Technologies™. DENV work was performed in the biosafety category 3 laboratory at the WIMM who also grew and provided the virus (DENV-2 strain 16681). The primary antibodies used were D3-2H2-9-21 (ATCC® HB-114™), 3H5-1 (ATCC® HB-46™) and D1-4G2-4-15 (ATCC® HB-112™). The secondary antibodies used were goat anti-mouse Alex Fluor® 488 and Alexa Fluor® 568. The cell-line used for DENV infection was Vero-SF-ACF.

Absorption and emission spectra measurement were performed in Chirascan™ CD spectrometer. Western blot was performed in iBlot® from Life Technologies™. Microscopy analysis was performed in Leica SP8 STED. Flow cytometry was performed in BD FACSCalibur™.

7.4.2 Passivation of Vivaspin® for improving recovery yield

The Vivaspin® column was filled with MiliQ water and centrifuged (15,000 x g, 25 ^{o}C) to remove the water. The remaining water in the column was removed by a pipette. The column was filled with 1% BSA in PBS and incubated at 25 ^{o}C for at least 2 h. The BSA

solution was removed by a pipette. The column was rinsed with MiliQ water for 4 times. For the last time, the water was centrifuged down and removed. The column was then ready for protein concentration.

7.4.3 Labelling of AV with rhodamine B derivatives (purified by G25 columns)

AV stock at – 80 ^{o}C (8.13 x 10^{11} vp / mL in PBS with 50 mM HEPES, pH 7.5) was buffer exchanged *via* dialysis (10,000 MW) in sterile TAPS (0.6 M, pH 8.5) at 4 ^{o}C for overnight. Dye **1, 2, 4** or **5** was dissolved in sterile TAPS (0.6 M, pH 8.5, 20% DMSO). The dye solution (different equiv.) was mixed with the virus solution so that the final DMSO concentration was adjusted to around 7.3%. A control with only DMSO but no dye was also performed. The reaction mixture was shaken at 25 ^{o}C for 4 h in dark. A small aliquot of each sample was taken out for SDS-PAGE analysis (4-12% Bis-Tris gel, MOPS buffer). The gel was stained by SilverXpress® according to the protocol provided by the manufacturer. The dye was removed by G25 minitrap / miditrap / PD10 columns using sterile PBS (pH 7.5, with 50 mM HEPES) as the eluting buffer. The column was chosen depending on the reaction scale, and the reaction mixture should be diluted by at least 7 times after running the column to ensure complete removal of free dyes. The resulting virus solution was concentrated with passivated Vivaspin® (10,000 MW, 800 x g) at 25 ^{o}C. The concentrated virus samples were stored at 4 ^{o}C.

7.4.4 PicoGreen® assay for determining the concentration of AV

Standard solutions (100 μL) were prepared from the standard stock (100 μg / mL) with TE buffer (0.05% SDS). Sample solutions (100 μL) were prepared by diluting the samples with TE buffer (0.05% SDS) for 5 times. For the AV stock control, the sample was diluted for 20 times. The standard and sample solutions were heated at 56 ^{o}C for around 45 min. The

solutions (50 µL x 2) were then transferred to a 96 well plate. Half of the samples were diluted with TE buffer (50 µL) while the other half was mixed with PicoGreen® solution (50 µL) which prepared by mixing Quant-iT™ PicoGreen® dsDNA reagent (10 µL) and TE buffer (2 mL). The fluorescence intensity of each well was measured (excitation: 485 nm; emission: 520 nm). The concentration of AV in each sample was calculated by comparing the fluorescence intensity with the AV stock sample with known concentration.

7.4.5 Growing of A549 epithelial cells

Refer to Chapter 5.

7.4.6 Microscopy study of A549 infected by dye 1 labelled AV

A549 cells in DMEM (10% FBS, 1% Pen/Strep) (200 µL) were added to sterile cover slips on a 24-well plate so that each well contained 40,000 cells. The plate was incubated at 37 ^{o}C for overnight. The medium was removed, and the wells were washed with fresh DMEM (10% FBS, 1% Pen/Strep) (3 x 1 mL). Fresh DMEM (10% FBS, 1% Pen/Strep) (300 µL) was added to each well, followed by the addition of PBS, unlabelled or labelled virus (5 µL, 5,000 MOI). The plate was incubated at 37 ^{o}C for 10 min. The medium was removed, and the wells were washed with fresh DMEM (10% FBS, 1% Pen/Strep) (3 x 1 mL). Fresh DMEM (10% FBS, 1% Pen/Strep) (300 µL) was added to each well. The plate was incubated at 37 ^{o}C. At the time point of 0 min, 20 min, 40 min, 60 min and 24 h, the medium of the corresponding well was removed and the well was washed with PBS (3 x 1 mL) (0 h and 24 h for the PBS control). The coverslip was transferred to another 24-well plate and fixed with 2.5% PFA in PBS (500 µL) for 30 min in dark. The solution was removed and the wells were washed with PBS (3 x 1 mL). The coverslip was placed on a microscopy plate with a drop of Prolong® Diamond Antifade mountant with DAPI and left at 25 ^{o}C for overnight in dark. The slide was stored at 4 ^{o}C before microscopy analysis.

7.4.7 Infection of A549 cells with AV for FACS study

The trypsin treated A549 cells were counted and diluted with DMEM (no FBS). The cells (100 μL x 3) were then transferred to a 96 well plate so that each well contained 50,000 cells. PBS, unlabelled or labelled AV (1,000 MOI) (10 μL) was added to the corresponding well and incubated at 37 ^{o}C for 4 h. The medium from the wells was removed, followed by washing with DMEM (10% FBS) for several times. DMEM (10% FBS) (100 μL) was added to each well and incubated at 37 ^{o}C for 24 h. The medium from the wells was removed, followed by washing with PBS (150 μL) for twice. Trypsin (50 μL) was added to each well and incubated at 37 ^{o}C for 15 min. DMEM (10% FBS) (250 μL) was added to each well. The cell culture in each well was transferred to specialised tubes for FACS analysis.

7.4.8 Measurement of excitation spectra, emission spectra and quantum yields of dyes

The standard rhodamine B dye was dissolved in 94% EtOH. The samples were dissolved in PBS. The excitation and emission spectra were measured in the range of 400 – 700 nm. The emission spectra were excited at 514 nm for all the samples and standard. The absorbance at 514 nm of the sample was measured. The solution was then diluted with the corresponding solvent until the absorbance reached ~0.1. The emission spectra were recorded for absorbance of 0.1, 0.08, 0.06, 0.04, 0.02 and 0 respectively. The Φ of a sample can be calculated by the following equation:

$$\Phi_X = \Phi_{ST} \left(\frac{Grad_X}{Grad_{ST}}\right)\left(\frac{\eta_X^2}{\eta_{ST}^2}\right) \quad \text{---------- (3)}$$

in which the subscripts ST and X refer to standard and sample respectively, Φ is the fluorescence quantum yield, Grad is the gradient from the plot of integrated fluorescence intensity versus absorbance, and η is the refractive index of the solvent used. By including the solvent refractive indices, the equation can be applied to samples in different solvents. The integrated fluorescence intensity can be calculated by measuring the area under the

emission spectrum:

$$\text{Integrated fluorescence intensity} = \sum_{i=m}^{n} \frac{x_i\,(x_{i+\Delta i})}{2}\,\Delta i \quad \text{---------- (4)}$$

in which i is the emission wavelength, Δi is the wavelength difference between two successive measurements, m is the minimum emission wavelength recorded, n is the maximum emission wavelength recorded and x is the fluorescence intensity at a specific wavelength.

7.4.9 Measurement of extinction coefficient of dyes

Solutions of dye **1**, **4** and rhodamine B in PBS with concentrations of 100 μM, 80 μM, 60 μM, 40 μM and 20 μM were prepared and added to a 96 well plate. The absorbance of each sample at 514 nm and its absorption maximum was measured with three replicas. A graph of absorbance versus concentration was plotted. The extinction coefficient can be calculated from the gradient of the graph.

7.4.10 Labelling of DENV with fluorescent dyes

DENV-2 stock obtained from Dr. Kerstin Lühn in Weatherall Institute of Molecular Medicine (in RPMI with 15% FBS; ~8 x 10^6 pfu / mL) was buffer exchanged to $NaHCO_3$ (0.2 M, pH 8.3) or TAPS (0.6 M, pH 8.5) using PD SpinTrap™ G-25. Alexa Fluor® 488 NHS was dissolved in $NaHCO_3$ (0.2 M, pH 8.3) to give 7 mg / mL right before the labelling. DENV solution (200 μL in $NaHCO_3$ or TAPS) was added to a tube containing AF488-NHS solution (20 μL) or dye **1** / **2** powder (~1.6 mg) respectively. The reaction mixture was shaken at 25 °C for 2 h in dark. Hydroxylamine (1.5 M, pH 8.5, 20 μL) was added and shaken at 25 °C for 1 h in dark. The virus was purified by Micro Bio-Spin™ P-6 columns with HNE buffer (5 mM HEPES, 150 mM NaCl, 0.1 mM EDTA, pH 7.4) and stored at – 80 °C or used directly.

7.4.11 Microscopy study of labelled DENV

Cover slips were soaked with 0.01% poly-L-lysine (200 µL) for 5 min and dried in the air. They were then transferred to a 24-well plate, followed by the addition of virus solution (90 µL). The plate was incubated at 25 ^{o}C for 30 min. The cover slips on the 24-well plate were washed with PBS (1 mL). 2% paraformaldehyde (1 mL) was added to each well and incubated at 25 ^{o}C for 2 h. The wells were washed with PBS (3 x 1 mL), and covered with 2% BSA/PBS (200 µL) at 25 ^{o}C for 30 min. After removing the supernatant, 3H5 in 2% BSA/PBS (0.01 mg / mL) (200 µL) was added and shaken at 25 ^{o}C for 1 h. After washing with PBS (3 x 1 mL), goat anti-mouse Alexa Fluor® 488 or/and 568 in 2% BSA/PBS (1:1000) (200 µL) was added and shaken at 25 ^{o}C for 1 h. After washing with PBS (3 x 1 mL), the cover slips were placed microscopy plates with a drop of DAKO fluorescence mounting medium and left at 25 ^{o}C for 1 day in dark before storing at 4 ^{o}C. The microscopy plates were then analysed by confocal microscopy.

7.4.12 Concentration and purification of DENV

DENV-2 stock obtained from Dr. Kerstin Lühn in Weatherall Institute of Molecular Medicine (in RPMI with 15% FBS; ~8 x 10^{6} pfu / mL) or mock solution (12 mL) was loaded into ultraclear centrifugation tubes and pelleted down by ultracentrifugation (SW41Ti, 30,000 x g, 4 ^{o}C, 2 h). Ultraclear centrifugation tubes with 30% sucrose in HNE buffer (2.5 mL) at the bottom and pure HNE buffer (9 mL) on top was carefully prepared. The resulting pellet from the previous step was resuspended in HNE buffer (500 µL) and loaded on top of the sucrose cushion, followed by ultracentrifugation (SW41Ti, 77,000 x g, 4 ^{o}C, 15 h). 1 mL fractions were collected from the tubes.

7.4.13 SDS-PAGE and western blot analysis of purified DENV

DENV stock / resuspended pellet solution / purified DENV fractions and their corresponding mock controls were treated with 1% triton-X100 at 4 ^{o}C for 2 h with constant vortexing. The samples were loaded into two replicated non-reducing SDS-PAGE gels (4 – 12% Bis-Tris, MOPS buffer). One of the gels was visualised with SilverXpress® silver staining kit, while the other one was analysed by western blot. The proteins were transferred to a nitrocellulose membrane with iBlot® . The membrane was covered with 5% BSA/TBST at 25 ^{o}C for 1 h. After washing with TBST for 4 times (25 ^{o}C, 5 – 10 min each), the membrane was cut into two parts at around 35 kDa. The upper and lower parts of the membrane were covered with 4G2 (mouse anti-E) and 2H2 (mouse anti-prM) in 2% BSA/TBST (0.0005 mg / mL) respectively and shaken at 4 ^{o}C for 16 h. After washing with TBST for 4 times (25 ^{o}C, 5 – 10 min each), goat anti-mouse HRP in 2% BSA/TBST (1:10000) was added and incubated at 25 ^{o}C for 1 h. After washing with TBST for 4 times (25 ^{o}C, 5 – 10 min each), the membrane was treated with Bio-Rad Clarity^{TM} ECL reagent according to the protocol from the kit. The membrane was exposed to a film for 1 h before visualisation in a dark room.

7.5 References

1. San Martin, C. & Burnett, R.M. Structural studies on adenoviruses. *Curr Top Microbiol Immunol* **272**, 57-94 (2003).
2. Meier, O. & Greber, U.F. Adenovirus endocytosis. *The Journal of Gene Medicine* **6**, S152-S163 (2004).
3. Pearce, O.M.T. et al. Glycoviruses: Chemical glycosylation retargets adenoviral gene transfer. *Angewandte Chemie-International Edition* **44**, 1057-1061 (2005).
4. Triplett, J.W., Herring, B.P. & Pavalko, F.M. Adenoviral transgene expression enhanced by cotreatment with etoposide in cultured cells. *Biotechniques* **39**, 826, 828, 830, passim (2005).
5. Zhang, S.L., Tan, H.C., Hanson, B.J. & Ooi, E.E. A simple method for Alexa Fluor dye labelling of dengue virus. *J Virol Methods* **167**, 172-7 (2010).
6. van der Schaar, H.M. et al. Characterization of the Early Events in Dengue Virus Cell Entry by Biochemical Assays and Single-Virus Tracking. *Journal of Virology* **81**, 12019-12028 (2007).
7. van der Schaar, H.M. et al. Dissecting the cell entry pathway of dengue virus by single-particle tracking in living cells. *PLoS Pathog* **4**, e1000244 (2008).
8. Ayala-Nunez, N.V., Wilschut, J. & Smit, J.M. Monitoring virus entry into living cells using DiD-labeled dengue virus particles. *Methods* **55**, 137-43 (2011).
9. McRae, R., Naumann, C., Hoehne, K. and Zeidler, R. Treatment of Vivaspin concentrators for improved recovery of low-concentrated protein samples. *Sartorius Stedim Biotech Application note*.
10. Leopold, P.L. et al. Fluorescent Virions: Dynamic Tracking of the Pathway of Adenoviral Gene Transfer Vectors in Living Cells. *Human Gene Therapy* **9**, 367-378 (1998).
11. Williams, A.T.R., Winfield, S.A. & Miller, J.N. Relative fluorescence quantum yields using a computer-controlled luminescence spectrometer. *Analyst* **108**, 1067-1071 (1983).
12. Magde, D., Rojas, G.E. & Seybold, P.G. Solvent Dependence of the Fluorescence Lifetimes of Xanthene Dyes. *Photochemistry and Photobiology* **70**, 737-744 (1999).

CHAPTER

8 Conclusion and outlook

8.1 Conclusion

The three goals mentioned at the beginning of this thesis have been successfully achieved. Firstly, electrophilic fluorination of Qβ VLPs with Selectfluor^{TM} demonstrated that fluorine atoms can be easily installed on the particle surface without the need for large prosthetic groups. The modification has little effect on the particle structure as revealed by dynamic light scattering compared to particles labelled with SFB, a commonly used fluorinating reagent. Although site-selective fluorination could be achieved with the 'tag-and-modify' approach, the resulting conversion was only around 15%. On the other hand, chemoselective fluorination on tyrosine residues gave a conversion of > 65% with careful manipulation of the amino acid sequence.

Furthermore, a detection system using ^{19}F NMR for monitoring the disassembly of fluorinated Qβ VLPs was established. The particle-associated broad peak can be easily distinguished from the multimer sharp peak created by chemical treatment. Also, a sharp peak with a slightly different chemical shift corresponding to degraded peptides was observed when the particle was treated with bacterial or mammalian cell lysate, further revealing the sensitivity of the developed system to different disassembly states. Preliminary *in vitro* experiments were also performed in THP-1 derived macrophages but with no detectable signal, probably due to the low uptake efficiency of particles by the cells.

Last but not least, IME-functionalised rhodamine B dyes were developed and one of the derivatives was revealed to be a better candidate in retaining adenovirus infectivity compared to dyes possessing other coupling groups. Although more experiments are required to give a solid conclusion, it is clear that the coupling group of the dye has a larger effect on virus infectivity than expected, especially for cases with high levels of labelling.

8.2 Future Directions

8.2.1 Application of the fluorine-labelling system for *in vivo* study

Although the fluorination systems have been successfully demonstrated at protein or lysate level, the final goal is to apply the methods for *in vivo* study. For the Selectfluor™ labelling methods, the next stage is to label the particles with ^{18}F for PET imaging. Since the conditions used for ^{18}F labelling have been tested already, it is believed that this can be performed without much difficulty. By injecting the ^{18}F labelled particles into animal models, their circulation pathway could be easily revealed with the sensitive PET imaging technique.

On the other hand, more work is required for the ^{19}F NMR detection system before it can be applied to real animal models. The most challenging part is to solve the sensitivity problem of the system, in which a considerable number of fluorinated particles are required to give an acceptable signal. As shown by the preliminary results from real cells, the amount of particles internalised is too small to give a detectable signal. This may be even worse in more complicated *in vivo* systems when ^{19}F MRI is employed. There are two possible directions for improving the signal intensity, either by increasing the uptake efficiency of the target cells or by improving the detection system itself, for example by adding more fluorine atoms on the particle surface.

8.2.2 Development of $^{18}F/^{19}F$ double-labelled VLPs

Once the ^{18}F-labelling and ^{19}F-detection systems have been developed, it will be interesting to test if the two systems can be combined in a single particle, so that its circulation and disassembly status in the body can be traced at the same time. Since ^{18}F atom does not have a nuclear spin property and ^{19}F is non-radioactive, they will not interfere with each other during detection. For instance, the tyrosine residues on the Qβ-CS (i.e. K16Tfm C74S C80S) species reported can be further labelled with ^{18}F-Selectfluor, rendering it applicable for both NMR spectroscopy and PET imaging (**Figure 8.1**).

Indeed, there is a recent trend towards the development of multi-modality imaging techniques, as evidenced by many recent studies in this area, since each molecular imaging technique has its own advantages and disadvantages. Also, only one injection of the combined probe is needed to obtain information by both imaging techniques.[1] Furthermore, the pharmacokinetics can be kept constant when a combined probe is used, which can obviously facilitate drug development.[2] However, careful optimisation should be performed when combining two modalities especially when they have different sensitivities.[2] For instance, as the sensitivity of PET is at least 1000 times more than that of MRI, it is expected that the signal given out by one ^{18}F atom would be comparable to that from thousands ^{19}F atoms in the same time scale. Recently, PET/MRI has been successfully applied for studying immune[3] and cancer cells.[4]

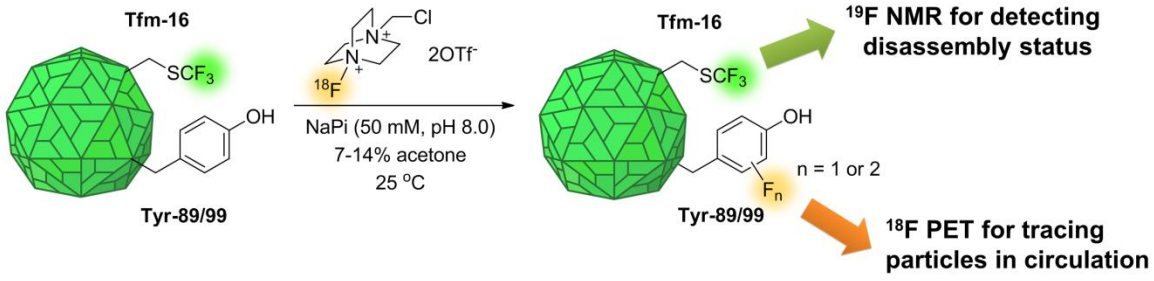


Figure 8.1: $^{18}F/^{19}F$ double labelled Qβ VLPs that can be detected by both ^{18}F PET and ^{19}F NMR.

8.2.3 Fluorine as a cell-targeting group

Highly fluorinated organic compounds such as polytetrafluoroethylene (or Teflon®) are characterised by their dual hydrophobic and lipophobic properties. This renders fluorinated species exhibiting reduced intermolecular interactions with non-fluorinated species but increased interactions with fluorinated species.[5, 6] For example, fluorous liquid liquid extraction is a technique which utilises a fluorous solvent to extract fluorous molecules from an organic or aqueous solvent.[5] This technique has been applied to a range of chemical compounds, including peptides.[7, 8]

One of the interesting yet challenging directions is to investigate if this kind of fluorine-fluorine interaction can be applied for cell-targeting. Previous studies have revealed the possibility for introducing fluorinated sialic acids on mammalian cell surfaces,[9-11] giving around 7 x 10^7 CF_3 groups per cell.[10] It was also found that the fluorinated cells exhibited reduced adhesion to non-fluorinated cells or proteins.[12] On the other hand, it is possible that these fluorinated cells can interact strongly with the fluorinated VLPs, hence rendering fluorine a cell-targeting group. It is however expected that a far greater number of fluorine atoms on the particle surface would be needed to make the interactions significant, for which more work is undoubtedly required. Also, careful optimisation is vital since aggregation between particles may occur if the fluorine interactions are too strong. The ultimate goal of the fluorine-labelling system is to develop a drug nano-carrier with tri-functionality, including targeting, circulation and drug release tracing (**Figure 8.2**).

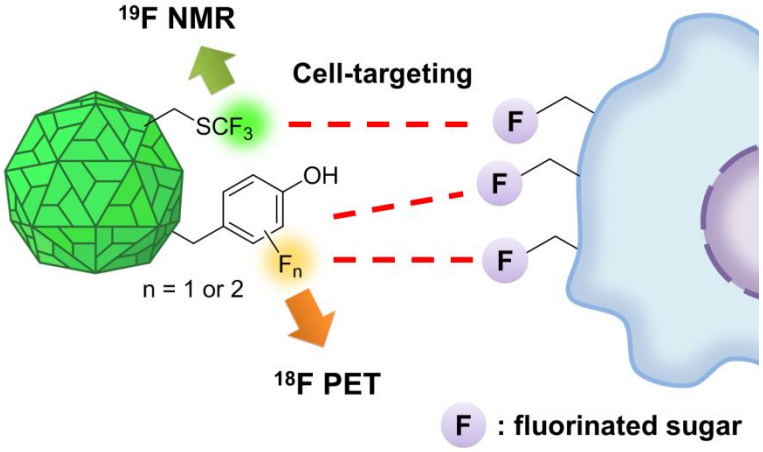


Figure 8.2 Fluorinated Qβ VLP as a drug nano-carrier with tri-functionality: targeting, circulation and drug release tracing.

8.2.4 Application of the IME coupling groups to different fluorescent dyes

As revealed by the results, the development of IME-functionalised dyes is still at an early beginning stage. To prove that IME coupling group is a good candidate for virus labelling, other coupling groups should be evaluated and compared. Also, it is possible that the IME coupling group only shows advantages in a certain type of virus such as adenovirus. More virus types are thus required for testing to confirm this as a general strategy for virus labelling. Other parameters such as the length of the linker and the introduction of ether groups should also be studied to further improve the reactivity and water solubility of the dye.

The ultimate goal of this project is to install the IME functional group to different fluorescent dyes, including the commercially available species such as the Alexa Fluor® series. This can perhaps provide an alternative choice for scientists, especially for those working in the virology field.

8.3 References

1. James, M.L. & Gambhir, S.S. A Molecular Imaging Primer: Modalities, Imaging Agents, and Applications. *Physiological Reviews* **92**, 897-965 (2012).
2. Louie, A. Multimodality imaging probes: design and challenges. *Chemical reviews* **110**, 3146-3195 (2010).
3. Lee, W.W. et al. PET/MRI of inflammation in myocardial infarction. *Journal of the American College of Cardiology* **59**, 153-163 (2012).
4. Kjær, A. et al. PET/MRI in cancer patients: first experiences and vision from Copenhagen. *Magnetic Resonance Materials in Physics, Biology and Medicine* **26**, 37-47 (2013).
5. Dobbs, A.P. & Kimberley, M.R. Fluorous phase chemistry: a new industrial technology. *Journal of Fluorine Chemistry* **118**, 3-17 (2002).
6. Yoder, N.C., Yüksel, D., Dafik, L. & Kumar, K. Bioorthogonal noncovalent chemistry: fluorous phases in chemical biology. *Current Opinion in Chemical Biology* **10**, 576-583 (2006).
7. de Visser, P.C. et al. A novel, base-labile fluorous amine protecting group: synthesis and use as a tag in the purification of synthetic peptides. *Tetrahedron Letters* **44**, 9013-9016 (2003).
8. Filippov, D.V. et al. Use of benzyloxycarbonyl (Z)-based fluorophilic tagging reagents in the purification of synthetic peptides. *Tetrahedron Letters* **43**, 7809-7812 (2002).
9. Yarema, K.J., Mahal, L.K., Bruehl, R.E., Rodriguez, E.C. & Bertozzi, C.R. Metabolic Delivery of Ketone Groups to Sialic Acid Residues: Application to Cell Surface Glycoform Engineering. *Journal of Biological Chemistry* **273**, 31168-31179 (1998).
10. Dafik, L., d'Alarcao, M. & Kumar, K. Fluorination of Mammalian Cell Surfaces via the Sialic Acid Biosynthetic Pathway. *Bioorganic & medicinal chemistry letters* **18**, 5945-5947 (2008).
11. Oetke, C. et al. Versatile Biosynthetic Engineering of Sialic Acid in Living Cells Using Synthetic Sialic Acid Analogues. *Journal of Biological Chemistry* **277**, 6688-6695 (2002).
12. Dafik, L., d'Alarcao, M. & Kumar, K. Modulation of cellular adhesion by glycoengineering. *J Med Chem* **53**, 4277-84 (2010).

Appendix

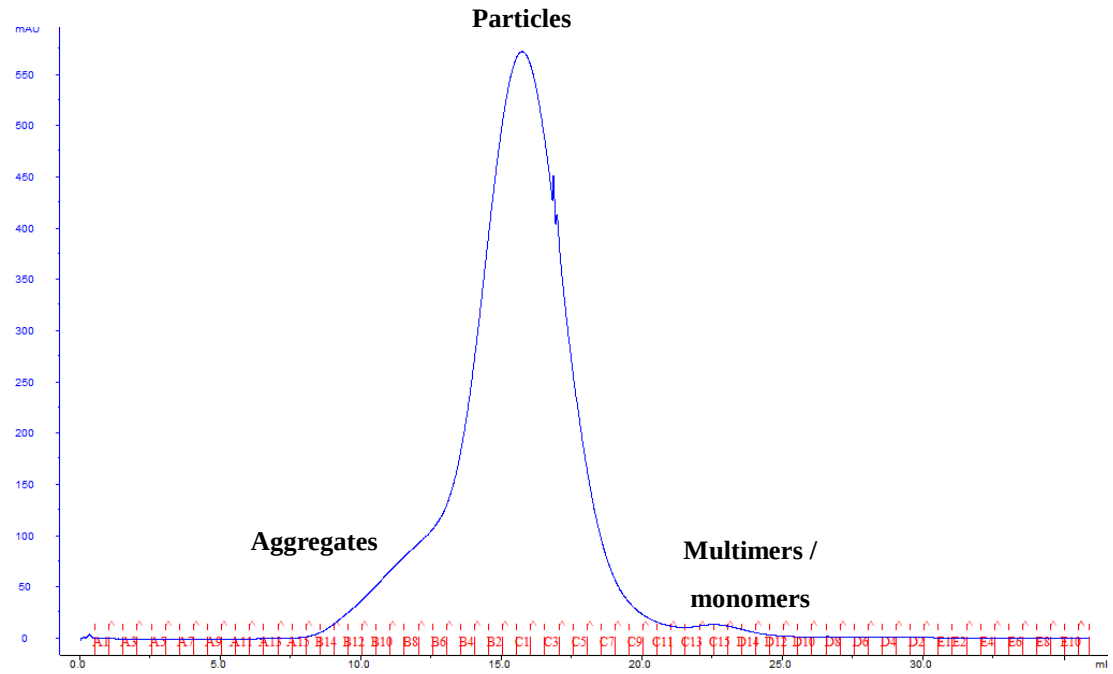


Figure S.1: A typical example of chromatogram obtained from size-exclusion chromatography of Qβ VLPs with significant amount of aggregates present after sucrose-gradient centrifugation.

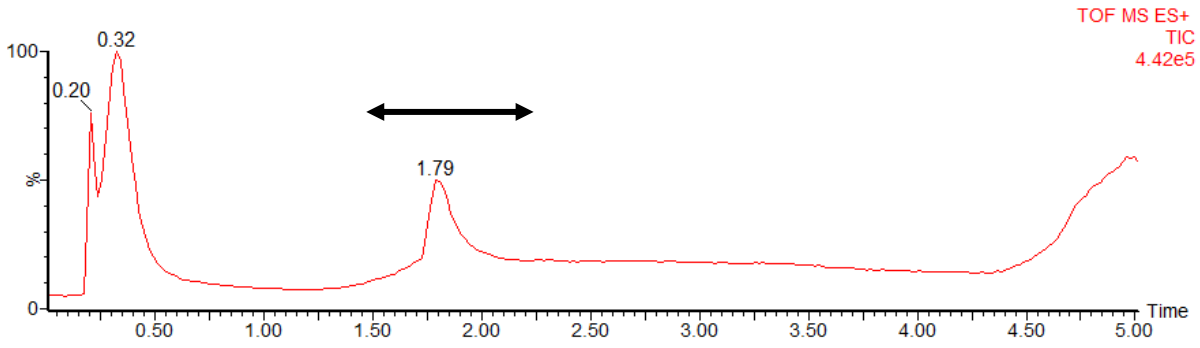
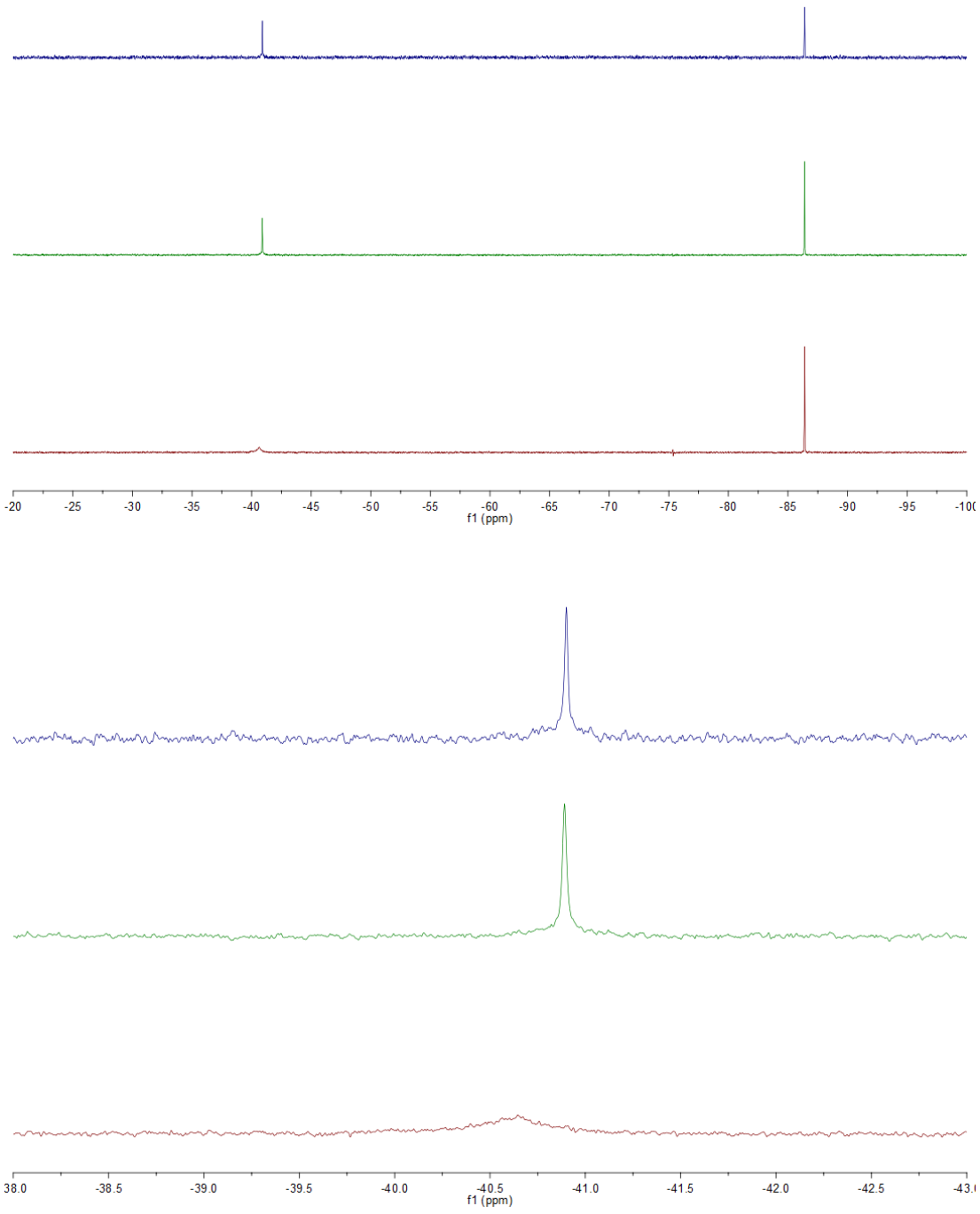


Figure S.2: A typical example of chromatogram obtained from LC-MS for protein analysis. All the proteins are assumed to be eluted between t = 1.50 min and t = 2.25 min due to gradient change (from 5% to 60% acetonitrile in 0.1% TFA) of elution buffer. The ion signals obtained during this period were integrated and analysed.

A


-20
-25
-30
-35
-40
-45
-50
-55
-60
-65
-70
-75
-80
-85
-90
-95
f1 (ppm)
38.0
-38.5
-39.0
-39.5
-40.0
-40.5
-41.0
-41.5
-42.0
-42.5
f1 (ppm)

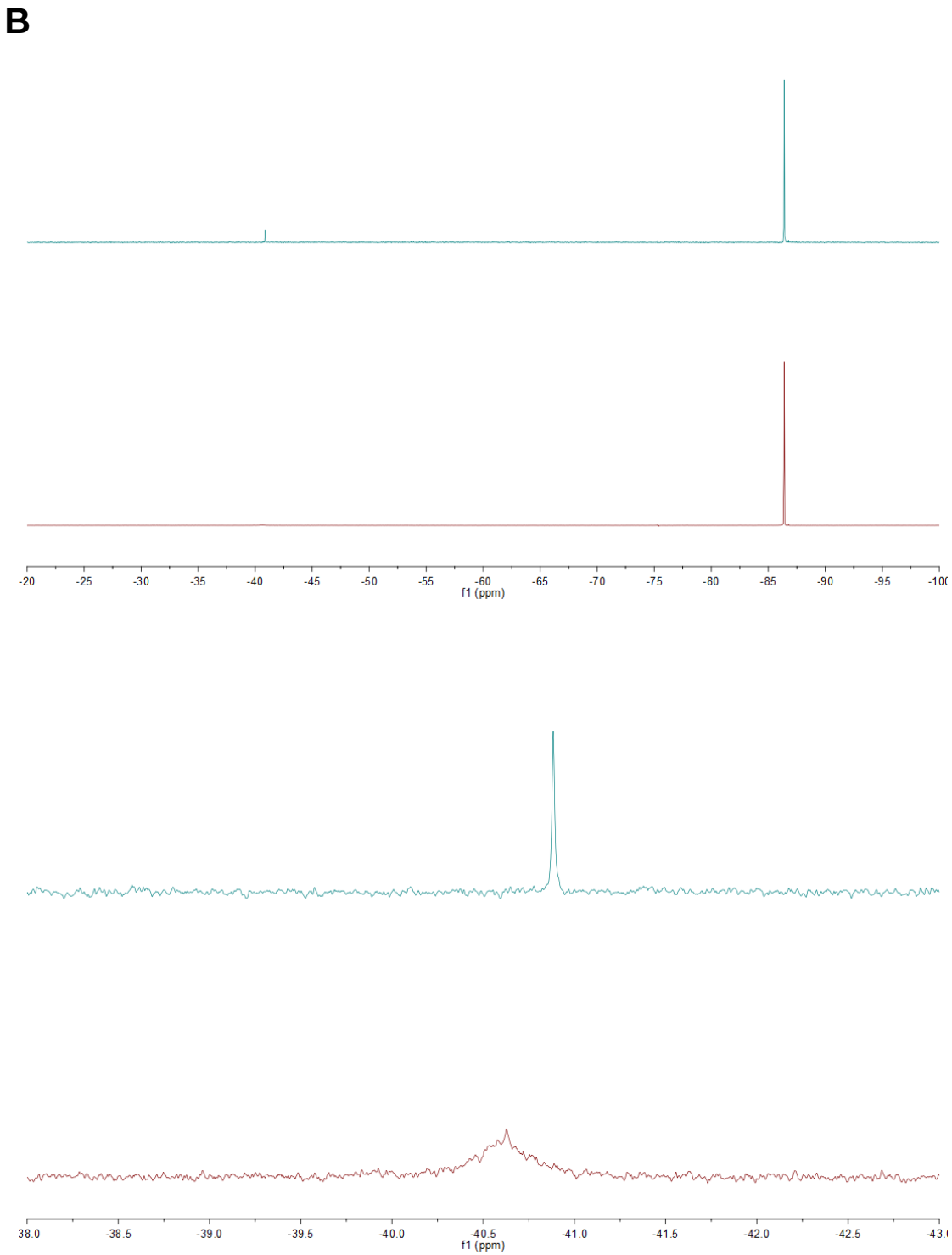


Figure S.3: ^{19}F NMR spectra of (A) Qβ-F and; (B) Qβ-CS VLPs under different conditions. Red: intact VLPs; Green/Teal: + 0.2 M SDS; Dark blue: + 0.2 M SDS and 50 mM DTT with heating at 90 ^{o}C. The reference peak (TFAcetone) was fixed to -86.40 ppm relative to $CFCl_3$. The spectra were adjusted in scale based on the number of scans to render the S/N ratio comparable.

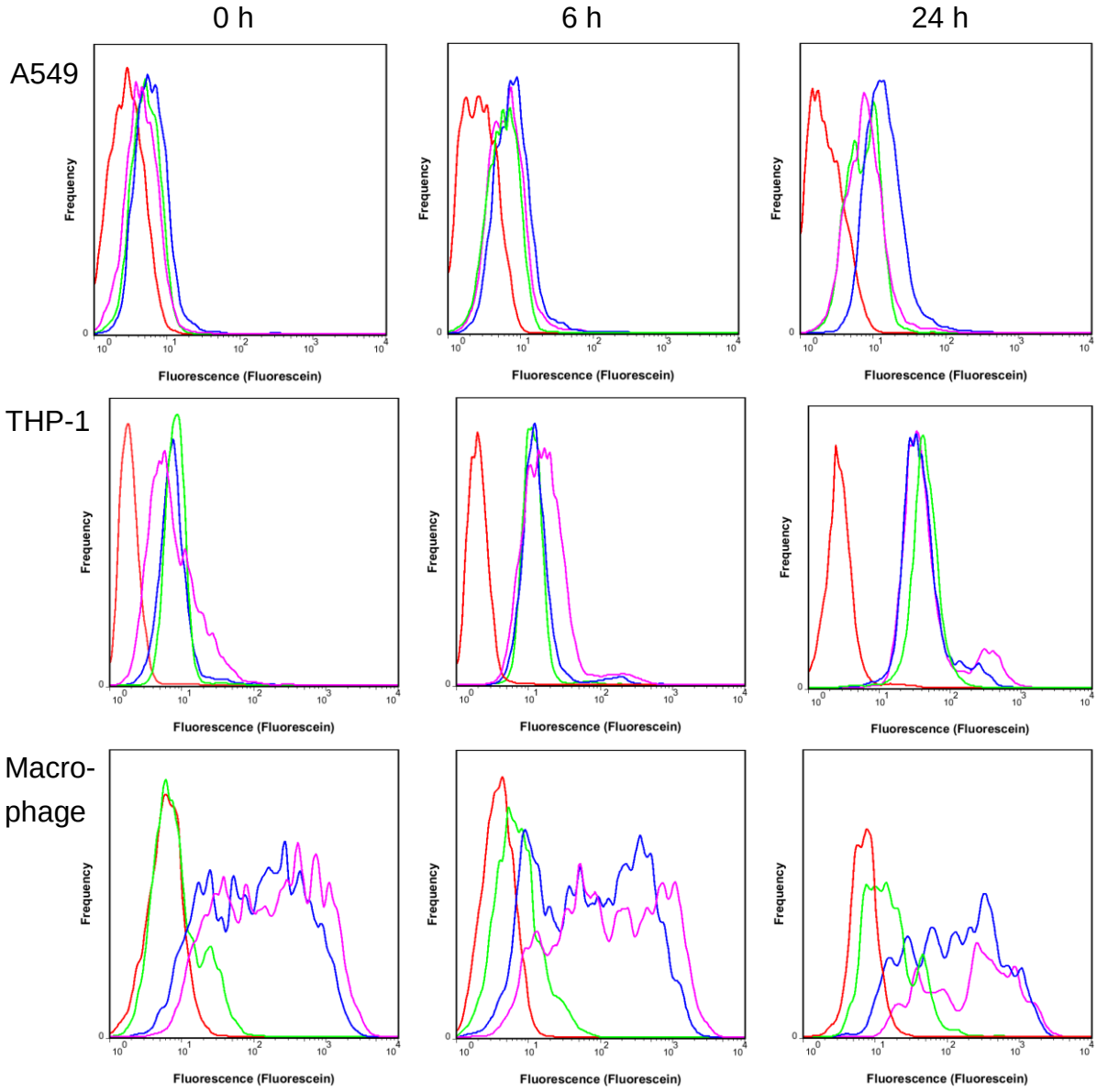
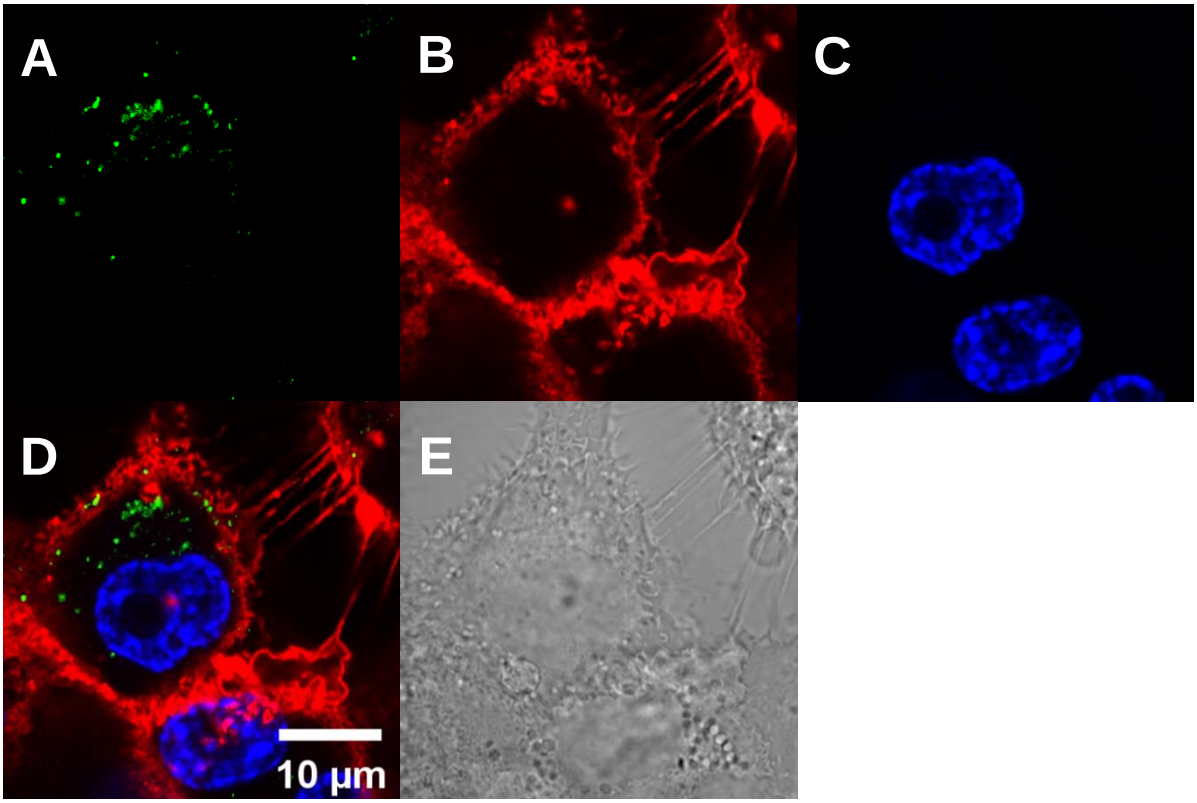
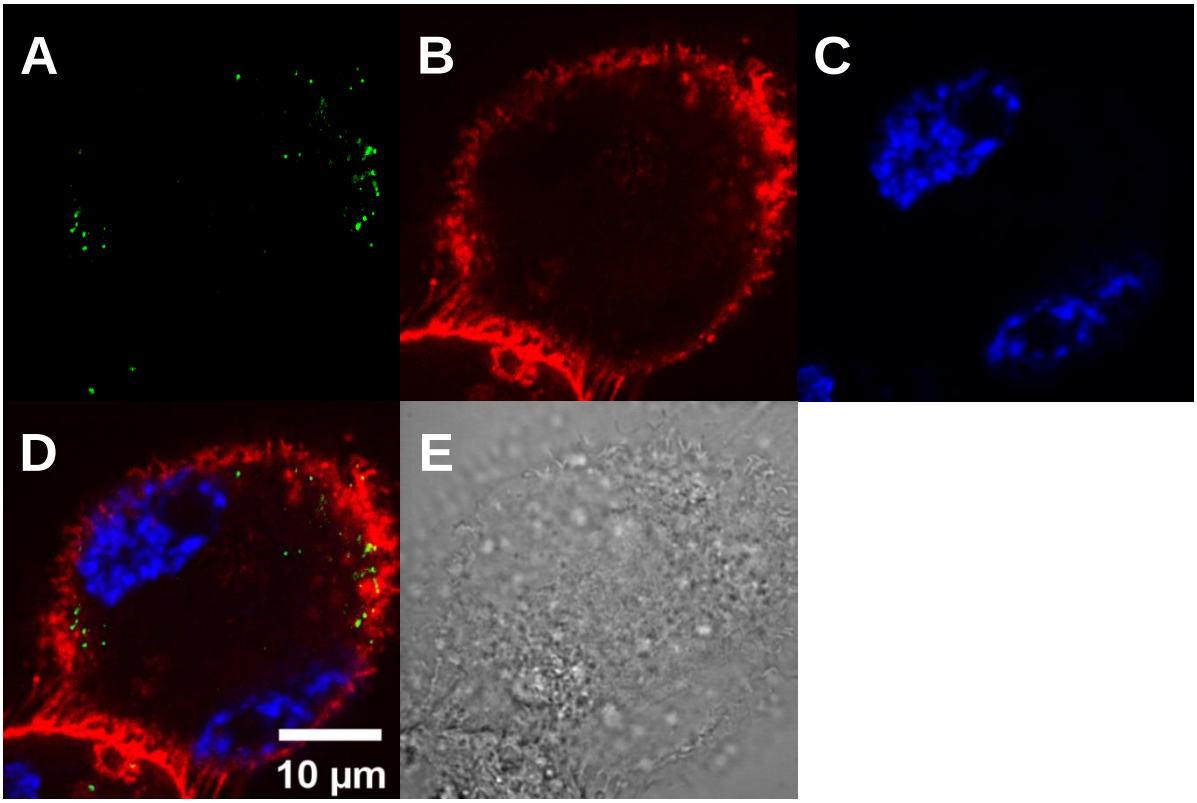


Figure S.4: Histograms from the flow cytometry experiment showing the internalisation of Qβ-F-fluo and Qβ-F-fluo-man VLPs into A549, THP-1 or THP-1 macrophage after 0 h, 6 h or 24 h incubation. Red: cell only; Green: cell + fluorescein dye; Blue: cell + Qβ-F-fluo; Purple: cell + Qβ-F-fluo-man.

Qβ-F-fluo VLPs incubated for 20 min



Qβ-CS-fluo VLPs incubated for 20 min



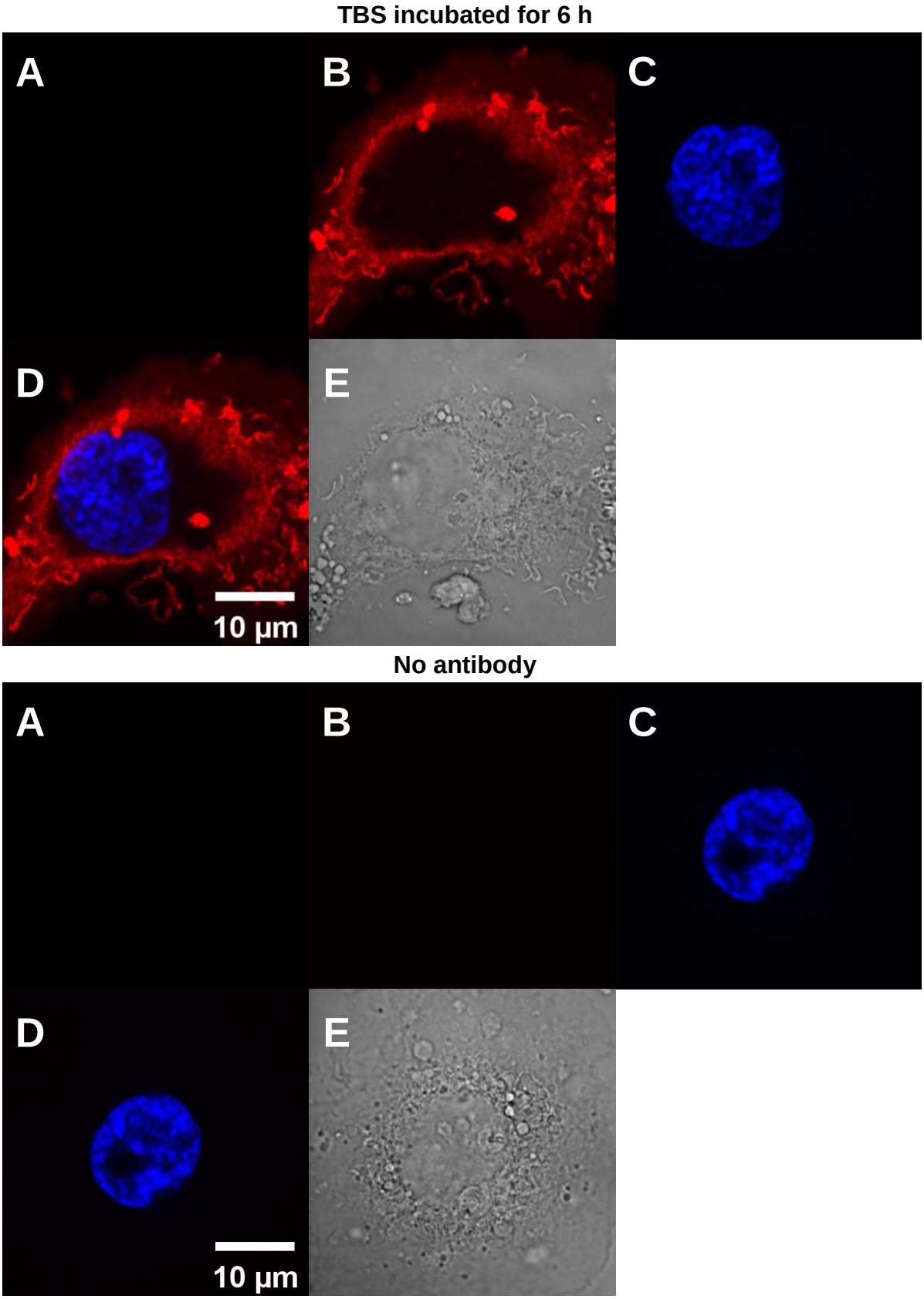
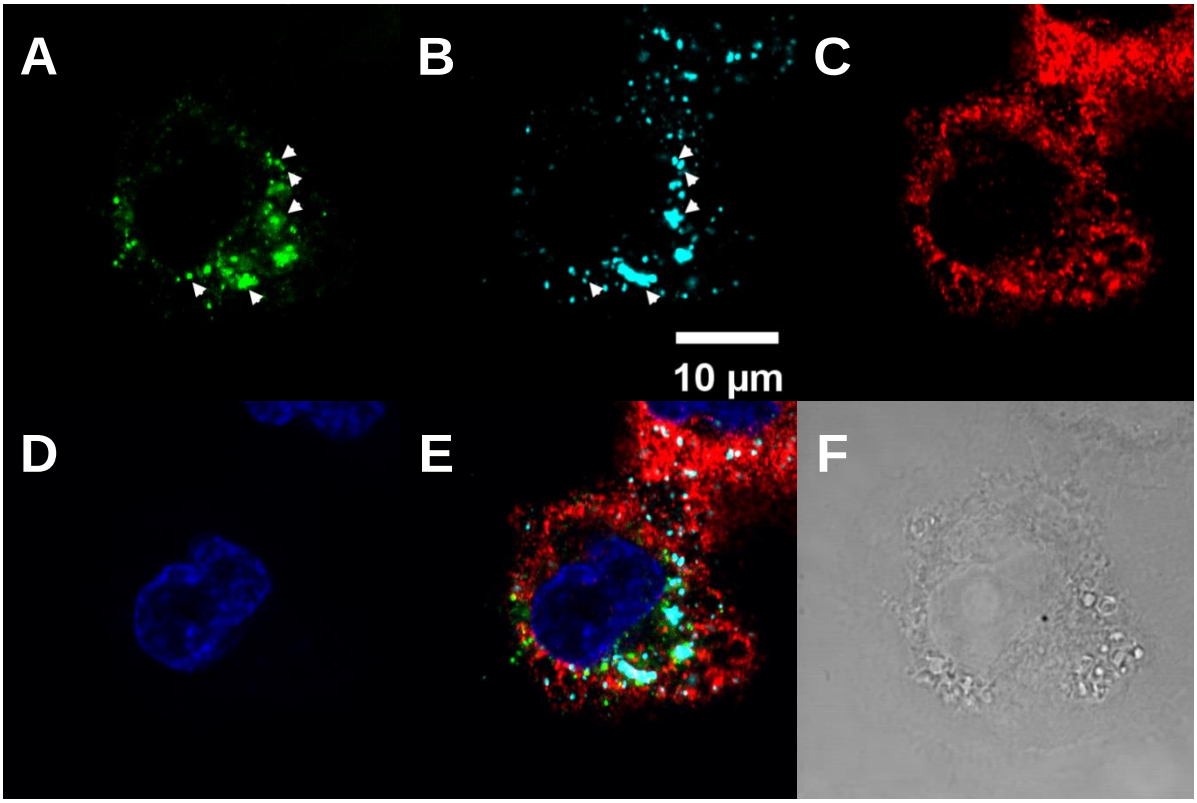
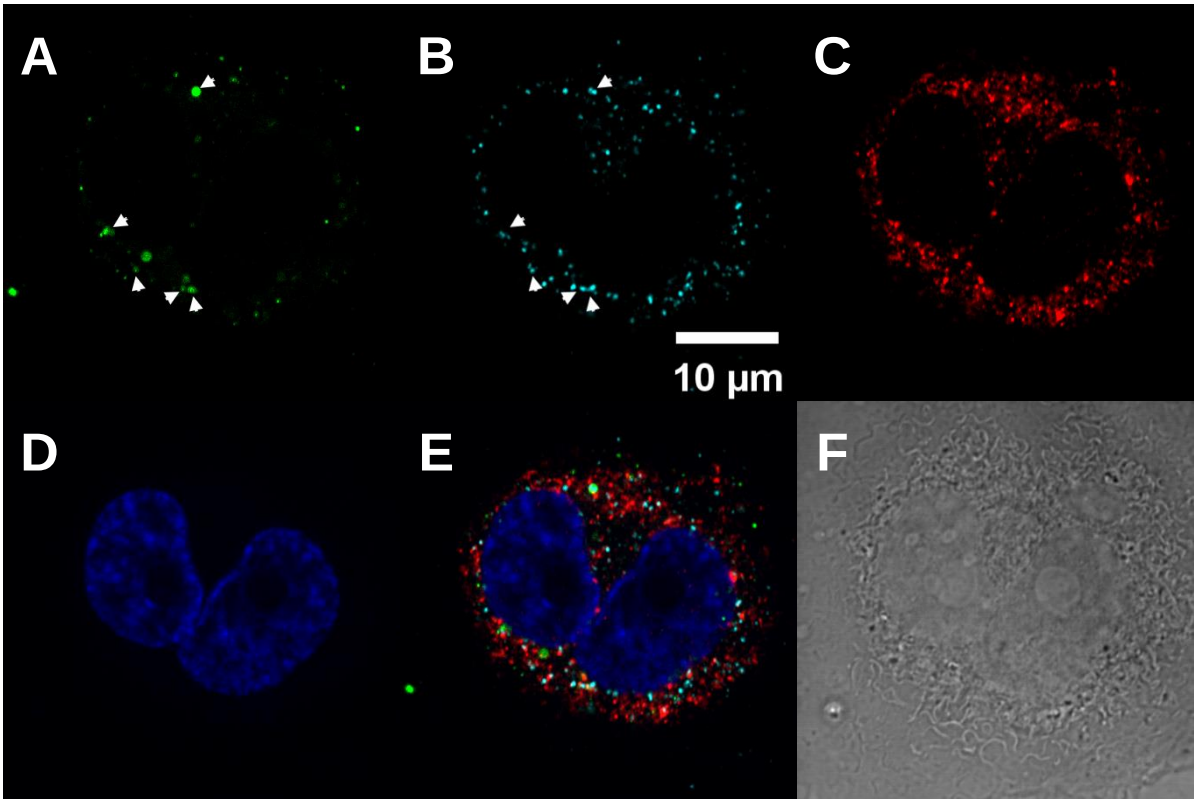


Figure S.5: Microscopy images of THP-1 derived macrophage under different conditions with membrane marker (CD14). (A) 488 channel (Qβ-F-fluo or Qβ-CS-fluo VLPs); (B) 647 channel (membrane with CD14); (C) DAPI channel (nucleus); (D) Merged image of (A)-(C); (E) Bright-field channel

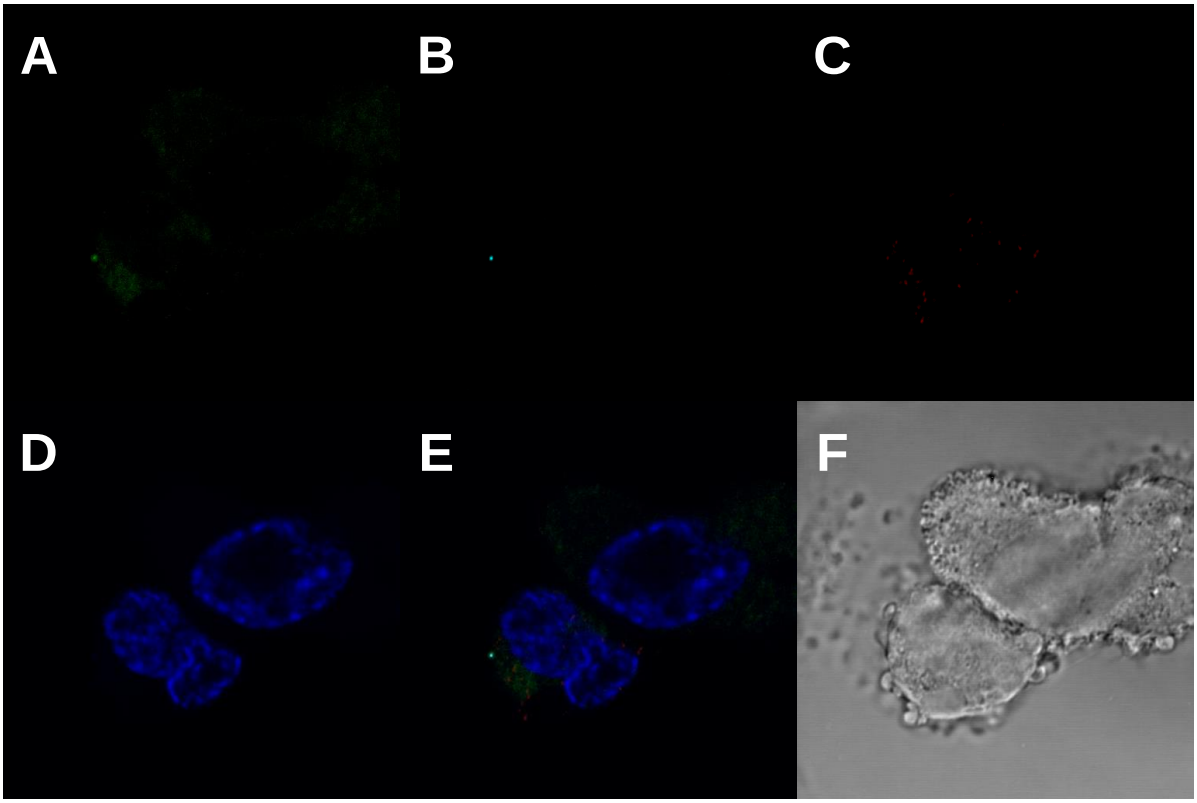
Qβ-F-fluo VLPs incubated for 20 min



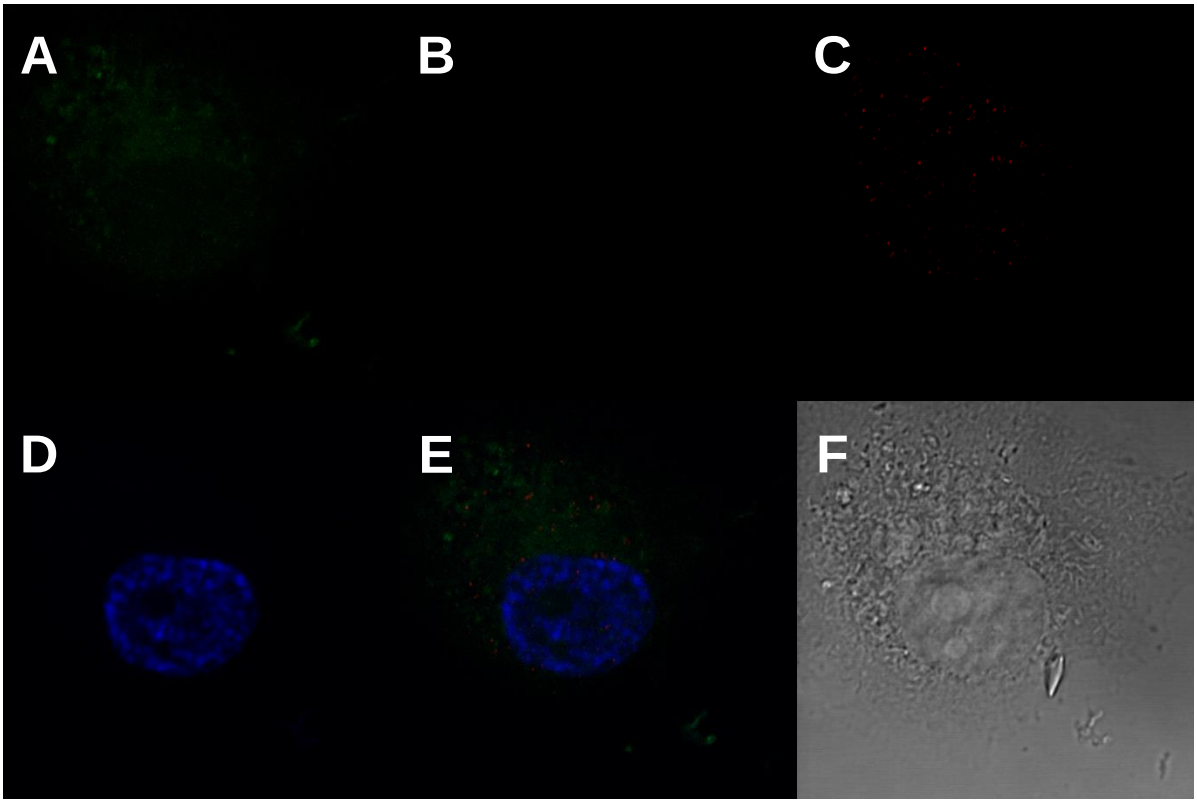
Qβ-CS-fluo VLPs incubated for 20 min



Mouse isotype control (for endosome)



Rabbit isotype control (for lysosome)



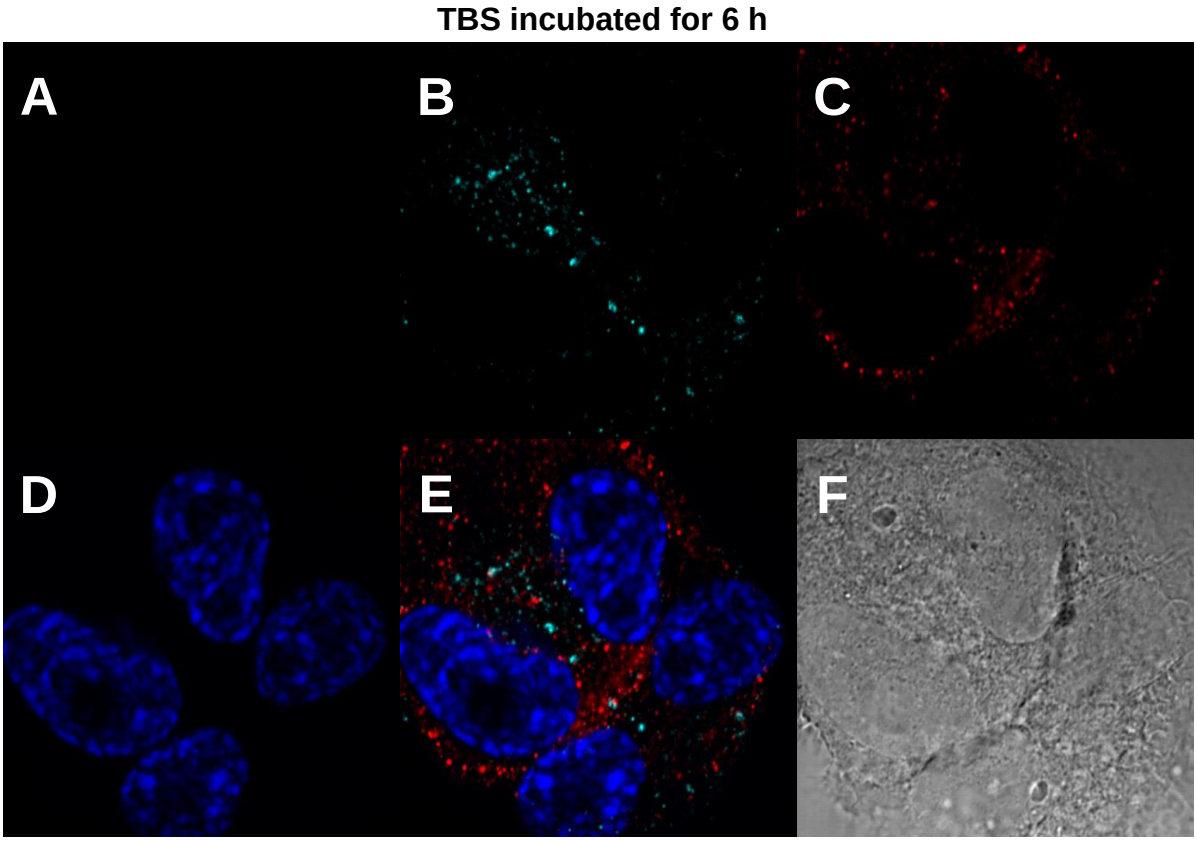
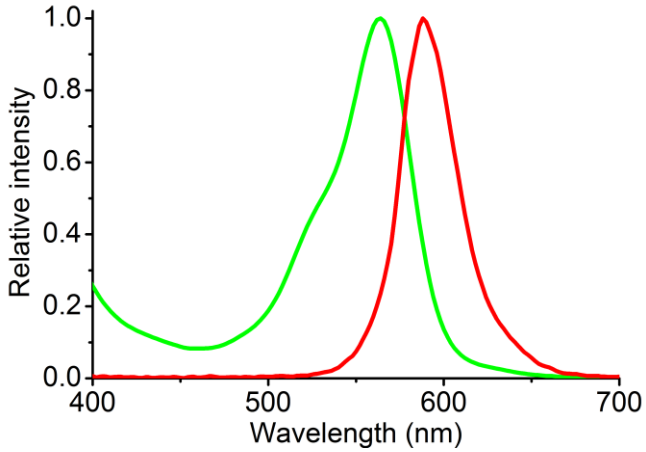


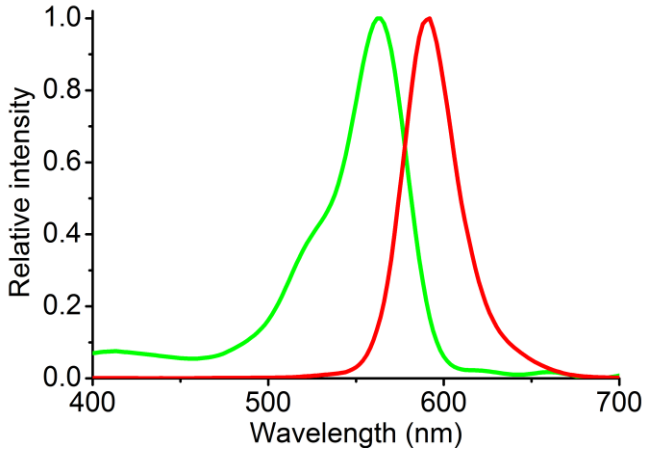
Figure S.6: Microscopy images of THP-1 derived macrophage under different conditions with endosome and lysosome markers. (A) 488 channel (Qβ-F-fluo or Qβ-CS-fluo VLPs); (B) 647 channel (early endosome); (C) 568 channel (lysosome); (D) DAPI channel (nucleus); (E) Merged image of (A)-(D); (F) Bright-field channel.

A



	HEL- dye **1**
Excitation max (nm)	564
Emission max (nm)	586
Stokes shift (nm)	22

B



	HEL- dye **4**
Excitation max (nm)	564
Emission max (nm)	592
Stokes shift (nm)	28

Figure S.7: Excitation and emission spectra of (A) HEL-dye **1** and (B) HEL-dye **4**.

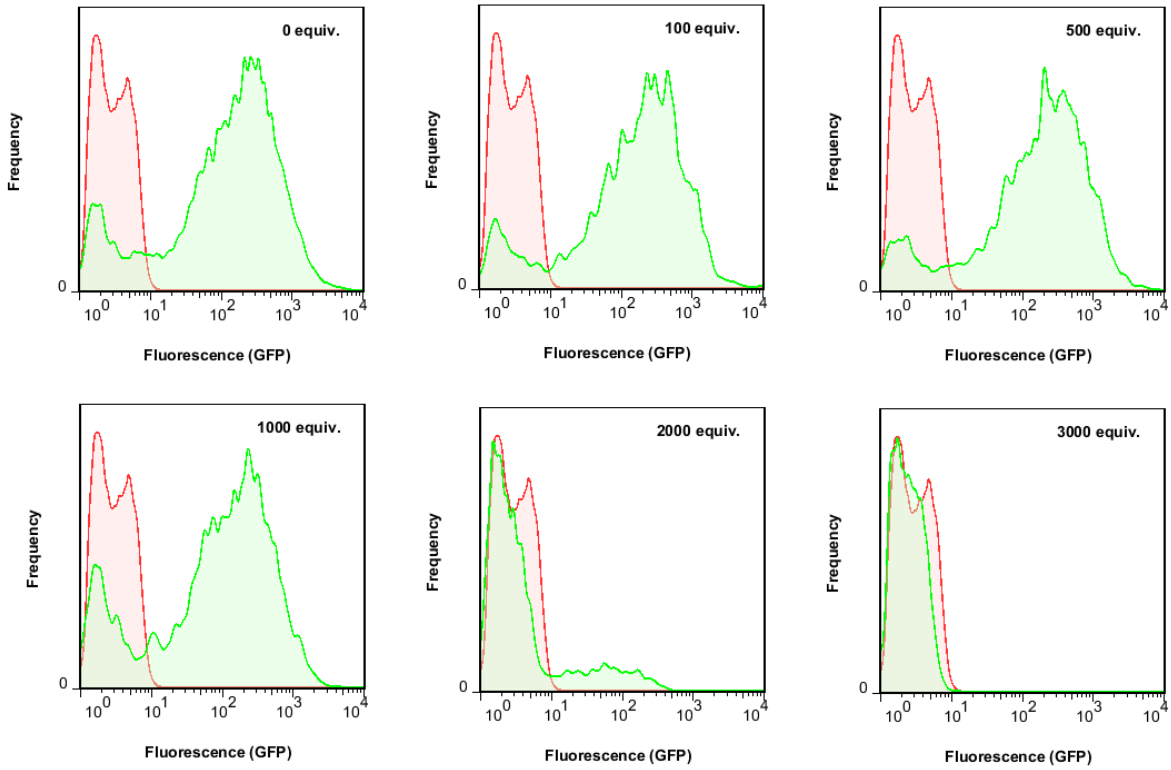


Figure S.8: Histograms from the flow cytometry experiment showing the GFP fluorescence signal of A549 epithelial cells infected by AV (old batch) labelled with different equiv. of dye **1**. Red: A549 cells without virus treatment; Green: A549 cells treated with AV. The experiment was performed after 24 h of infection.

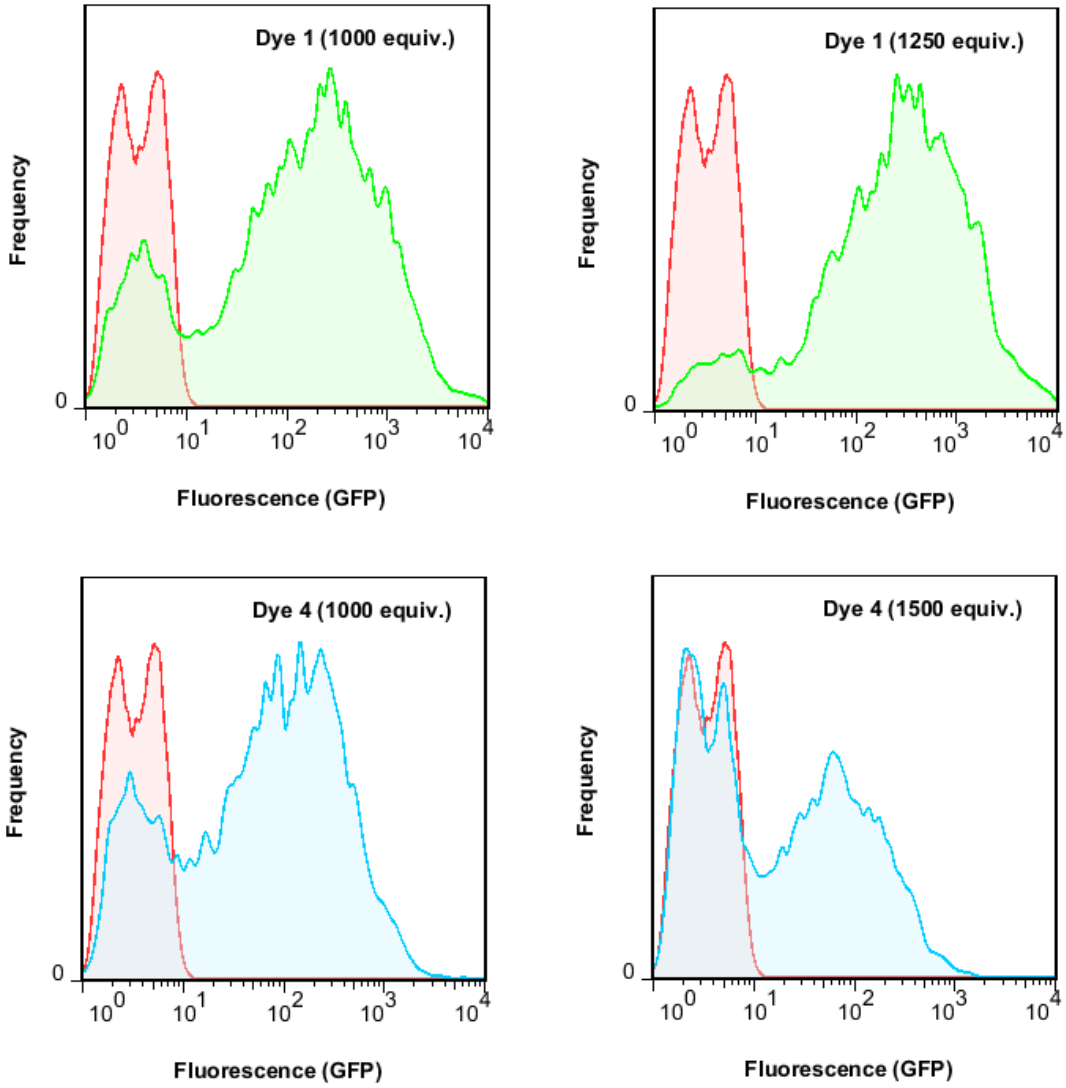
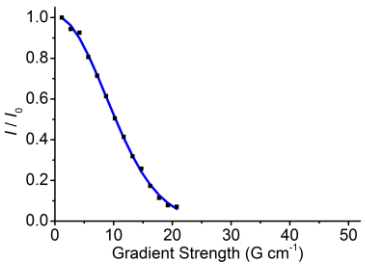
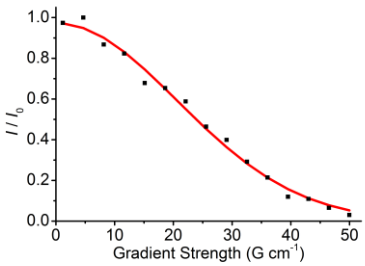
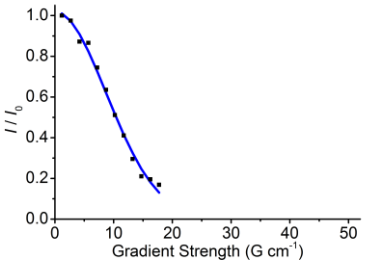
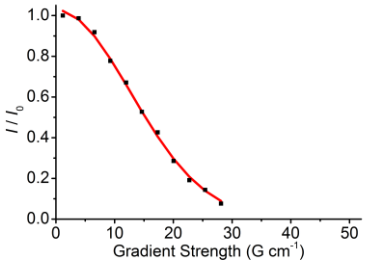
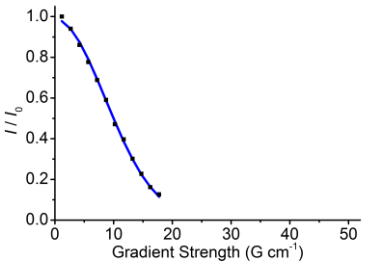
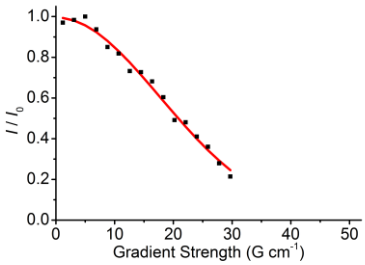


Figure S.9: Histograms from the FACS experiment showing the GFP fluorescence signal of A549 epithelial cells infected by AV (freshly prepared batch) labelled with different equivalent of dye **1** or **4**. Red: A549 cells without virus treatment; Green: A549 cells treated with AV-dye **1**; Blue: A549 cells treated with AV-dye **4**. The experiment was performed after 24 h of infection.

Table S.1: The I / I_0 versus gradient strength plots and diffusion coefficient values of various Qβ species and their corresponding references.

	Diffusion coefficient (x 10^{-11} m^2 s^{-1})		
	TFAcetone	**Protein**	**Protein after normalisation** †
Qβ-F VLPs + 0.2 M SDS	66.5 ± 0.6 	3.59 ± 0.09 	3.56 ± 0.09
Qβ-F VLPs + 0.2 M SDS + 50 mM DTT (90 °C heating)	66.3 ± 1.9 	5.48 ± 0.07 	5.44 ± 0.10
Qβ-CS VLPs + 0.2 M SDS	69.5 ± 1.2 	5.71 ± 0.74 	5.38 ± 0.60

† The diffusion coefficient of protein is normalised based on a single diffusion coefficient of TFAcetone (6.6 x 10^{-10} m^2 s^{-1}) to eliminate any variations in viscosity.