Novel Fluorinated Probe for Medical Imaging

A thesis submitted to the Board of the Faculty of Physical Sciences in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

at the

University of Oxford

Matthew D. Robinson

Hertford College, University of Oxford

Trinity Term 2014

Declaration

The work presented in this thesis was conducted in the Chemistry Research Laboratory at the University of Oxford between September 2010 and April 2014 under the supervisor of Professor Ben Davis and Professor Christopher Schofield. All the work is my own, expect where otherwise stated, and has not been submitted for any other degree at the University of Oxford or any other university.

Matthew D. Robinson

6th June 2014

A Novel Fluorinated Probe for Medical Imaging

A thesis submitted for the degree of Doctor of Philosophy, University of Oxford

Matthew D. Robinson, Hertford College, Trinity Term 2014

Medical imaging is an important area of scientific research with the principle aim of accurately diagnosing diseases to guide more effective treatment. Magnetic resonance imaging (MRI) provides advantageous high spatial resolution and deep tissue visualization. However, ^{1}H MRI can suffer from low signal-to-noise ratios due to strong background signals from abundant, endogenous molecules such as water and fat. ^{19}F MRI overcomes this drawback by directly detecting administered ^{19}F-labelled agents without any background signal.

This thesis reports the design and development of a new fluorinated agent for medical imaging. The bacteriophage Qβ, which comprises 180 copies of its coat protein, was used as the scaffold for the imaging agent. The fluorinated amino acid trifluoromethionine (TFM) was incorporated into the Qβ coat protein by the sense codon reassignment method using auxotrophic *E. coli*. This marks the first example of introducing fluorine into a virus by genetic modification. Reports of previous attempts to incorporate TFM into proteins have shown low to good incorporation levels (the maximum being 82%). Here, extensive optimisation of the protein production conditions resulted in unprecedentedly high incorporation levels (96%). Analysis of the modified Qβ-F by agarose gel electrophoresis, dynamic light scattering and transmission electron microscopy showed the coat protein self-assembles into virus-like particles (VLPs), essentially indistinguishable from the wild-type Qβ VLPs.

The intact Qβ-F VLP gave no ^{19}F NMR signal. This was rationalised in terms of the VLP's high molecular weight and resulting fast transverse relaxation. Stimulating disassembly of the VLP by addition of dithiothreitol (DTT), sodium dodecyl sulphate (SDS), or a combination of the two resulted in the development of ^{19}F NMR peaks. Multimers and monomers of the disassembled Qβ-F were distinguished by ^{19}F NMR and ^{19}F DOSY. This result was verified by measuring the ^{19}F NMR of a prepared sample of alkylated Qβ-F monomer and comparing its chemical shift value with that of peaks from the ^{19}F NMR of the disassembled Qβ-F VLP.

Finally, initial *in vivo* experiments were carried out in order to investigate the biodistribution of the Qβ-F particle. After injection of Qβ-F VLPs into rats, organs were dissected and homogenized. The Qβ-F particle was found to accumulate predominantly in the liver. ^{19}F NMR of the liver after sucrose density gradient ultracentrifugation only showed a signal on incubation with DTT, demonstrating that the Qβ-F particle remained intact *in vivo*.

Table of contents

List of Abbreviations

Å	Angstroms
aaRS	Aminoacyl tRNA synthetase
Ac	Acetyl
Aha	Azidohomoalanine
aq.	Aqueous
BCA	Bicinchoninic acid
Boc	*t*-Butoxycarbonyl
BSA	Bovine serum albumin
^{o}C	Degrees Celsius
CCMV	Cowpea chlorotic mottle virus
Conc.	Concentrated
CPMV	Cowpea mosaic virus
CT	Computed tomography
d	Doublet
dd	Doublet of doublets
δ	NMR chemical shift
Da	Daltons
DC	Dendritic cell
DC-SIGN	Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin
DCM	Dichloromethane
DEPT	Distortionless enhancement by polarisation transfer
DFM	Difluoromethionine
DLS	Dynamic light scattering
DNA	Deoxyribonucleic acid
dt	Doublet of triplets
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ES	Electrospray
ESI-MS	Electrospray ionisation mass spectrometry
Et	Ethyl
eq.	Equivalent(s)
FDG	18-fluorodeoxyglucose
g	Gram(s)
Hag	Homoallylglycine
HBV	Human hepatitis B virus
Hpg	Homopropargylglycine
Hz	Hertz
ICG	Indocyanine green
IPTG	Isopropyl-β-D-thiogalactoside
J	Coupling constant
k	Kilo
L	Litre

LB	Luria broth
Lit	Literature value
m	Multiplet or milli
M	Molar or mega
M^+	Positively charged molecular ion
µ	Micro
Me	Methyl
Met	Methionine
mol	Mole(s)
MPO	Myeloperoxidase
MRI	Magnetic resonance imaging
MW	Molecular weight
MWCO	Molecular weight cut off
m/z	Mass to charge ratio
n	Nano
NIR	Near-infrared
NMR	Nuclear magnetic resonance
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PDB	RCSB Protein Data Bank
PET	Positron emission tomography
PFPE	Perfluoropolyether
ppm	Parts per million
R_f	Retention factor
RF	Radiofrequency
RNA	Ribonucleic acid
RPM	Revolutions per minute
RT	Room temperature
s	singlet
SDS	Sodium dodecyl sulfate
t	triplet
T_1	Longitudinal relaxation time
T_2	Transverse relaxation time
TAE	Tris-acetate EDTA
TBS	Tris-buffered saline
^{t}Bu	Tertiary butyl
TCEP	Tris(2-carboxyethyl)phosphine
td	triplet of doublets
TEM	Transmission electron microscopy
TFA	Trifluoroacetic acid
TFM	Trifluoromethionine
tfm-Phe	trifluoromethyl-L-phenylalanine
TLC	Thin layer chromatography
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer ribonucleic acid
UV	Ultraviolet

V	Volts
v/v	Volume to volume ratio
VCAM-1	Vascular cell adhesion molecule 1
VLP	Virus-like particle
WT	Wild-type
w/v	Weight to volume ratio

Acknowledgements

I would like to thank Professor Ben Davis and Professor Chris Schofield for offering me the position of DPhil student in their groups. I would particularly like to thank Professor Chris Schofield for his support and encouragement during the course, and for many useful discussions which I feel I have learnt from immensely in terms of creative thinking.

Many people have helped me during my DPhil studies, members both past, and present, of the BGD group. Dr Al Aljibali taught me the essential skills of virus particle analysis and I have found his knowledge and enthusiasm for studying viruses very encouraging indeed. Thanks also goes to Al for taking the TEM micrographs at the Old Road Campus.

Dr Tim Claridge has been of invaluable help with the protein ^{19}F NMR work. Tim's world-class expertise in NMR is impressive and inspiring, and it has been a pleasure to work with him. Tim took many of the NMR measurements on the AVB500 and the AV600 spectrometers and has devoted a considerable amount of his time. For this I am very grateful. Dr Barbra Odell has also provided much help with ^{19}F NMR.

I worked with Dr Daniel Anthony towards the end of the project with the animal studies. Although it was brief, it was great to get to know Daniel and I enjoyed discussing the potential experiments we could carry out with Qβ-F. It is a shame that we didn't have enough time to carry out further experiments, such as determining the half-life of Qβ-F VLP; however, this will be an exciting starting point for the person who is to take up the Qβ baton!

I would like to thank the people who proof-read this thesis: Chris Schofield, Thomas Parsons, Bala Vijayakrishnan, Steve Robinson and Thomas Wright. Their careful attention to detail and generous offering of their time is much appreciated.

I have made many great friends during my DPhil years at the Chemistry Department. I would particularly like to thank the friends in my academic year: Bhaskar, Mahima, Phin, Lukas, Inga and Hua.

I would like to thank my fiancée, Marta Zarzecka, for her support and friendship over the past few years. Finally, I thank my parents, Mary and Steve Robinson, for their persistent encouragement and unfailing support in the work I do, during and prior to my DPhil course. The completion of this thesis has been made possible by their love.

1. Introduction

1.1 Introduction

Medical imaging is an exciting and practical area of research which has benefits to people with a variety of diseases. The field is characteristically interdisciplinary, with research and development projects concerning medical imaging expertise ranging from chemists to physicists, to radiologists and medical doctors. There are a number of imaging modalities presently available, which are being developed by many researchers. *In vivo* imaging requires administration of an appropriately distributed probe and a means to detect the probe *in vivo*.

The aim of the research described in this thesis was to develop a novel ^{19}F magnetic resonance imaging agent and to investigate its potential for use in medical imaging. In this introductory chapter, key imaging modalities currently used in research and in the clinic are introduced, then a review of available techniques for the fluorine-labelling of proteins is given. Finally, the aims of the work described in this thesis are outlined.

1.2 Medical Imaging

1.2.1 An introduction to medical imaging

Medical imaging is used to detect and locate diseased cells either with a view to removing them or to guide more accurate therapy, such as radiotherapy. However, the stage of a disease required for anatomical detection is usually late and recovery rates when diseases are detected at such late stages are low. Early diagnosis of diseases can be life-saving – early stage-detection is a key objective of medical imaging. Often, diseases can be imaged and diagnosed even before noticeable symptoms have developed, thus the treatment can be more effective. For example, the vascular cell adhesion molecule 1 (VCAM-1) has been shown to be up-regulated in cerebral vessels during early stages of brain tumors in both murine models

and human brain tissue.[1] Hence, MR imaging of nanoparticles conjugated to VCAM-1 antibodies has allowed pre-symptomatic *in vivo* imaging of brain metastasis by targeting VCAM-1.[1,2]

Such imaging as that highlighted above is termed tomographic i.e. provides anatomical images of sections (slices) through the body. In contrast, molecular imaging allows tissues and cells to be probed at a *molecular* level, and allows processes to be imaged through time as opposed to discrete "snap-shots".[3,4] In addition to enabling the detection of diseases at a curable stage, molecular imaging has the potential to provide crucial information which may contribute to a further understanding of diseases such as cancer.

Further examples of molecular imaging agents include those which are activated by specific enzymes prevalent in particular diseased states. Myeloperoxidase (MPO) plays a major role in the inflammatory response in animal models of stroke and has been identified as a key biomarker for inflammation. For example, gadolinium-based contrast agents containing 5-hydroxytryptamine side-groups have been shown to be oxidised by MPO resulting in polymerisation of the contrast agent and subsequent enhancement of neighbouring proton T_1 relaxivity.[5,6] Another prevalent condition is Alzheimer's disease, which is associated with amyloid-β plaques and tau protein aggregates.[7] Higuchi and co-workers recently developed a fluorescent compound, which selectively binds tau aggregates and so can be used as a molecular probe for tau amyloids.[8] Furthermore, the ligand was ^{11}C-labelled and so the tau aggregates were also visualized by positron emission tomography (PET) in a patient.

1.2.2 Imaging methods

Currently used molecular imaging techniques include PET imaging, computed tomography (CT), fluorescence-based imaging, and magnetic resonance imaging (MRI). These are now briefly reviewed.

PET uses biological molecules (tracers) which have been labelled with radioactive nuclei: ^{11}C, ^{13}N, ^{15}O or ^{18}F.[9] These tracers are injected intravenously into the body and their uptake is monitored. Pairs of gamma rays, generated when a positron from a radioactive nucleus collides with an electron, are detected by a scintillator and form an image.[10,11] The most commonly used tracer for PET imaging is 18-fluorodeoxyglucose (FDG),[12] which is used to diagnose tumors.[13] Cancer cells exhibit a high rate of glycolysis, even under normal oxygen concentrations (in normal cells oxidative phosphorylation predominates over glycolysis to generate energy when oxygen is present).[14] Hence glucose transporters, termed Glut-1-5, are upregluated and lead to a high FDG uptake.[15,16] Once in the cells, FDG is phosphorylated to 2-deoxyglucose-6-phosphate by hexokinases, which is also over-produced in cancer cells.[17] 2-Deoxyglucose-6-phosphate is unable to diffuse out of the cell and so the ^{18}F-labelled compound builds up inside the tumor cells. FDG can therefore provide a means to distinguish between cancerous and benign tissue.[18]

PET provides high sensitivity and allows imaging of biological processes. PET-FDG is primarily used to assess the response of tumors to therapy. However, PET scanners are expensive to purchase and to run, and are generally limited to hospitals associated with a research institute which can provide a cyclotron necessary to generate radioactive isotopes. Furthermore, the number of scans per patient is limited due to harmful radioactive exposure.[19]

CT is one of the most well-established imaging methods used in hospitals. In CT an X-ray emitter and detector are positioned on either side of the patient. Different tissue densities attenuate the X-ray beam by different amounts. The attenuation at each position along the detector is recorded. The emitter and detector move axially in a circular path around the patient to another position where the X-ray beams pass through a different combination of voxels. An algorithm is then used to determine the attenuation value for each voxel and so a

2D image of a slice through the body is generated.[20] Often CT contrast agents are used to enhance visualization of blood vessels and so ease interpretation. CT contrast agents contain elements with a high atomic number e.g. iodine[21] or bismuth,[22] which effectively absorb X-rays. Most CT contrast agents are based on a 2,4,6-triiodonated benzene ring.[23]

Fluorescence imaging uses dyes such as fluorescein and indocyanine green (ICG). Near-infrared (NIR) fluorescence imaging has proved useful for guiding surgery.[24,25] In the procedure, a NIR fluorescent dye is injected intravenously and the operative area of the patient is illuminated with visible and NIR light which are merged to give an anatomical image and a highlight of the tissue to be excised, respectively. For example, a folate-conjugated fluorescent agent, able to target the folate receptor-α, which is upregulated in ovarian cancer,[26] has been used to visualize ovarian cancer cells during operation and allowed image-guided excision of tumors.[27] However, a disadvantage of fluorescence imaging is its poor depth penetration (<8 mm) due to attenuation of excitation and emission signals.[28,29] Furthermore, high background signals can result from reflected excitation light, lowering the signal-to-noise ratio in the resulting image.[30]

There are many competing factors to be taken into account when employing an imaging modality as implied by the above discussion. Crucially, many imaging techniques suffer from low resolution; PET imaging, optical bioluminescence and fluorescence imaging give spatial resolutions of ~0.2-2 mm, 1-2 mm and 1-3 mm, respectively.[4] MRI, on the other hand, provides resolution of μm order and allows deep tissue visualization.[4] High resolution and the capability of visualizing deep below the surface are key advantages when imaging tumours, and so MRI provides a very powerful technique for molecular medical imaging.

1.2.3 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is a non-invasive imaging technique which is based on the principle of nuclear magnetic resonance.[31,32] When placed in a magnetic field, the spin of magnetic nuclei aligns either parallel (α) or anti-parallel (β) to the applied magnetic field (B_o) (Fig. 1.1).[33]

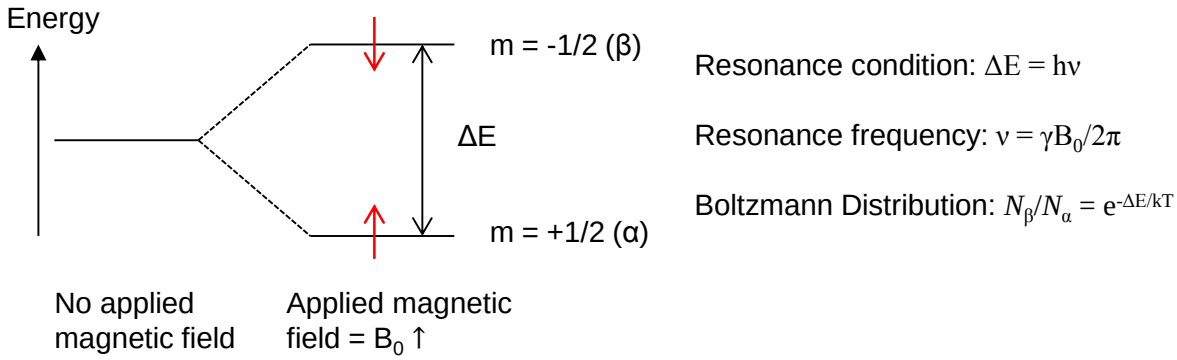


Figure 1.1 *The splitting of the energy of the nuclear magnetic moment in the presence of a magnetic field (B_0); and the resonance condition, resonance frequency, and Boltzmann distribution over the two resulting energy states. k = the Boltzmann constant.*

α spins are lower in energy than β spins and so, at equilibrium, there is an excess of α spins ($N_\beta/N_\alpha < 1$) (Fig. 1.2 A). This can be represented as a net magnetization vector along the *z* axis (Fig. 1.2 B). A radiofrequency (RF) pulse is applied which equalises the spin populations ($N_\beta/N_\alpha = 1$) (Fig. 1.2 C). The RF pulse also causes the spins to focus in the *y*-direction generating phase coherence (Fig. 1.2 C) resulting in the magnetization vector to now lie in the *xy* plane (Fig. 1.2 D).[34] The spins return to their equilibrium position ($N_\beta/N_\alpha < 1$), i.e. the magnetization vector returns to the *z* axis, by releasing energy in a process known as spin-lattice relaxation. This process is characterized by the time constant T_1. At the same time, the spins dephase in the *xy* plane (i.e. phase coherence is lost). This process is known as spin-spin relaxation and is characterized by the time constant T_2.

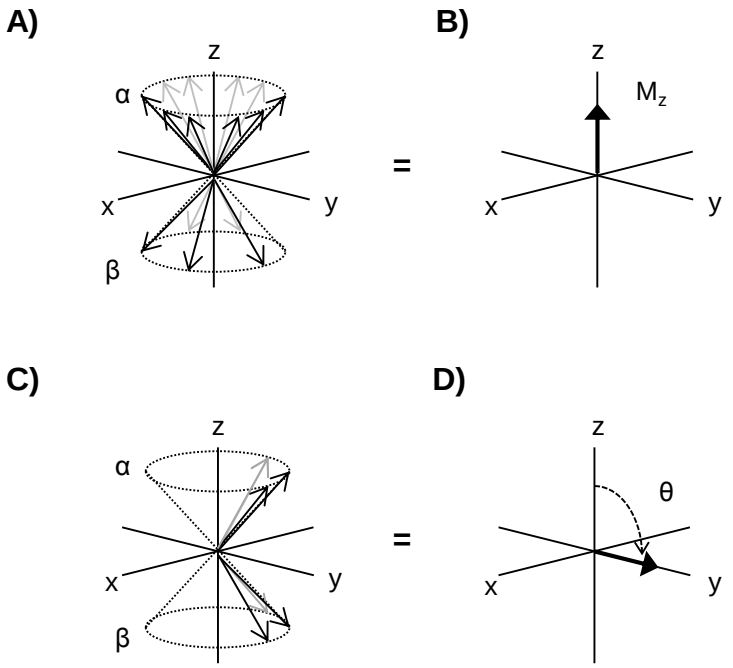


Figure 1.2 *At equilibrium, there is a slight excess of α spins (A) which gives rise to net magnetization in the z direction, M_z (B). After a RF pulse, the number of α and β spins are equalized and focused in the y direction (C). This causes the net magnetization vector to tip at an angle θ to the z axis and point along the y axis (D).*

The use of magnetic resonance in medical imaging began when Damadian showed an increase in T_1 and T_2 relaxation times for tumour tissue compared to normal tissue.[35] These initial experiments laid the foundations for the future development of contrast agents (*vide infra*).

Magnetic field gradients give the nuclear frequencies a spatial dependence, which is translated to the positioning in the final image.[31] The ^{1}H nuclide is highly NMR-sensitive (compared to other magnetic nuclei; see Table 1.1) and is abundant in living tissue (from molecules including water and fat molecules). As a result ^{1}H MRI benefits from high sensitivity. In data acquisition, the process of pulsing and allowing spins to relax is rapidly repeated to build up a signal. Therefore, spins with a short T_1 contribute greater to signal intensity.[36] Conversely, nuclei with a short T_2, i.e. fast transverse magnetization dephasing, lead to weaker signals. Nuclei in different tissue types have differing T_1 and T_2 times and so there is a contrast between the signal intensities they produce. To further enhance this contrast at the region of interest, contrast agents are targeted to these regions. The contrast

agent interacts with the neighbouring water protons, shortening their spin-lattice relaxation times (T_1) or spin-spin relaxation times (T_2), resulting in a higher or lower signal intensity, respectively, from proton spins in the vicinity of the contrast agent. However, even when using contrast agents, the abundance of background endogenous ^{1}H *in vivo* means contrast between the region of interest and benign tissue is still generally low. This can make image interpretation difficult and so limits its use for medical imaging.

1.2.4 ^{19}F Magnetic Resonance Imaging

^{19}F MRI is a promising alternative to ^{1}H MRI. Instead of using endogenous ^{1}H *in vivo* to create an image, ^{19}F MRI allows *direct* detection of an administered imaging probe. Since no detectable ^{19}F occurs in the body, high image contrast is achieved.[37,38] Furthermore, the ^{19}F nuclide has a similar gyromagnetic ratio, γ, to that of ^{1}H and hence: (i) its sensitivity is close to that of ^{1}H, and (ii) conventional ^{1}H MRI instruments can be used after facile modification of the RF coil (Table 1.1). Finally, ^{19}F has 100% natural abundance, making synthesis of ^{19}F-containing markers relatively inexpensive and simple.

Isotope	*Spin*	*Gyromagnetic ratio (10^7 rad T^{-1} s^{-1})*	*NMR frequency at 2.35 T (MHz)*	*Natural abundance (%)*	*Relative sensitivity*
^{1}H	1/2	26.752	100.000	99.985	1
^{2}H	1	4.107	15.351	0.015	1.45×10^{-6}
^{3}He	1/2	-20.380	76.181	1.40×10^{-4}	5.75×10^{-7}
^{7}Li	3/2	10.398	38.866	92.58	0.272
^{13}C	1/2	6.728	25.145	1.108	1.76×10^{-4}
^{14}N	1	1.934	7.228	99.63	1.00×10^{-3}
^{15}N	1/2	-2.712	10.137	0.37	3.86×10^{-6}
^{17}O	5/2	-3.628	13.562	0.037	1.08×10^{-5}
^{19}F	1/2	25.181	94.094	100	0.834
^{23}Na	3/2	7.080	26.466	100	9.27×10^{-2}
^{31}P	1/2	10.841	40.481	100	6.65×10^{-2}
^{39}K	3/2	1.250	4.672	93.1	4.75×10^{-4}
^{129}Xe	1/2	-7.452	27.856	26.44	5.71×10^{-3}

Table 1.1 *NMR properties of a selection of nuclei commonly exploited for NMR experiments. Reproduced by permission from John Wiley & Sons (39), copyright 2007.*

Ahrens *et. al.* demonstrated the promise of *in vivo* ^{19}F MRI when they used perfluorocarbons to label and track dendritic cells *in vivo* after administration into a mouse.[40] Overlay of the ^{1}H anatomical MR image in grayscale with the ^{19}F MR image in colour allows visualization of the biodistribution of the ^{19}F-labelled cells (Fig. 1.3). This overlay technique has also been used for imaging amyloid β plaques in the brain for presymptomatic diagnosis of Alzheimer disease and shows the potential of this technique.[41,42] The high contrast of these ^{19}F MR images demonstrate the advantage of imaging the ^{19}F nucleus over the ^{1}H nucleus, due to the absence of ^{19}F signal from animal bodies. However, perfluorocarbons have relatively long retention times, limiting their clinical applicability.[43]

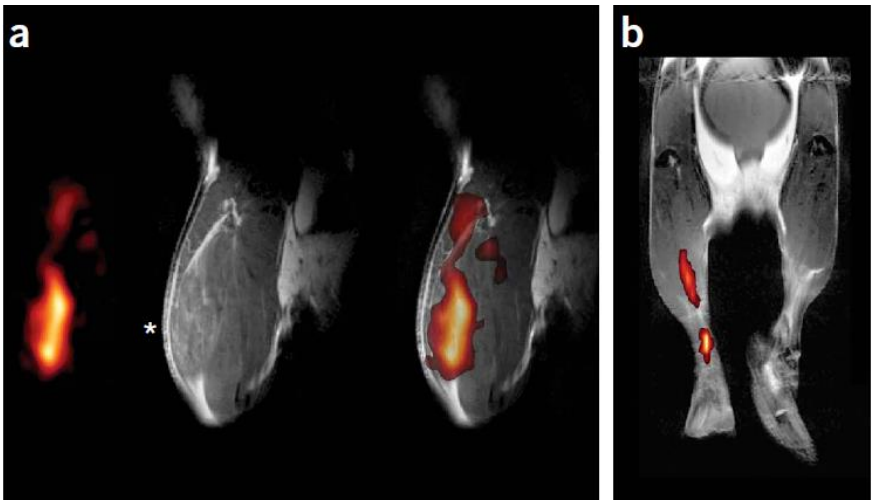


Figure 1.3 *In vivo MRI of PFPE-labeled DCs in mouse. The ^{19}F intensity is displayed on a 'hot-iron' intensity scale, and the ^{1}H images are shown in gray scale. (a) Mouse quadriceps after intramuscular injection of PFPE-labeled DCs (the asterisk indicates the injection site). Shown (from left to right) are ^{19}F and ^{1}H images and a 'composite' $^{19}F/^{1}H$ image. (b) Composite image of DC migration into the popliteal lymph node following a hind foot pad injection. Reprinted with permission from Macmillan Publishers Ltd: [Nature Biotechnology] (40), copyright (2005).*

The sensitivity of a ^{19}F agent can be enhanced by increasing the number of equivalent ^{19}F atoms in the imaging molecule. This has been elegantly shown by the construction of a molecule containing 27 symmetry-related fluorine atoms shown to give rise to a single ^{19}F NMR peak (^{19}FIT, Fig. 1.4).[44] Similarly, Whittaker and co-workers have constructed hyperbranched polymers which incorporate 12 ^{19}F atoms and allow a high enough molecular mobility in each branch to give the long T_2 times required for imaging.[45]

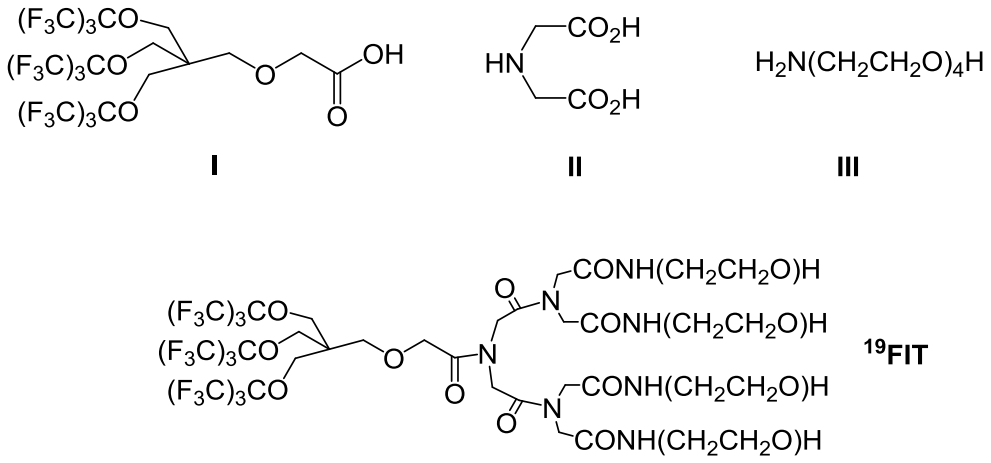


Figure 1.4 *Chemical structures of ^{19}FIT and its three building blocks (**I**, **II**, and **III**). Reproduced by permission from John Wiley & Sons: [Angewandte Chemie International Edition] [44], copyright (2009).*

^{19}F probes which are activated by a change in conditions or a biological event have been developed (termed smart probes). Imaging probes with a long T_2, and therefore minimal MRI signal, can be modified such that the T_2 is shortened sufficiently to give a signal. Two ways have been developed. First, the collapse of a self-assembled probe;[46] second the disconnection of a paramagnetic moiety (e.g. Gd^{3+}) which is otherwise in close enough proximity to ^{19}F to shorten its T_2.[47]

1.3 Viruses and virus-like particles

1.3.1 Virus-like particles

Viruses protect, transport and deliver viral nucleic acids (RNA) to host cells. Viruses can be produced without their RNA cargo, thus leaving a non-infectious, empty capsid shell (virus protein coat). These empty capsid shells are known as virus-like particles (VLPs).[48,49]

Viruses are arguably the most exquisite molecular architectures in the natural world due to their extensive structural refinement as a result of evolutionary development. A wide variety of VLP modification techniques have been investigated in the past decade and a half allowing the development of VLPs with useful medicinal applications, including imaging.[50–55] Some examples of commonly exploited VLPs are shown in Fig. 1.5.

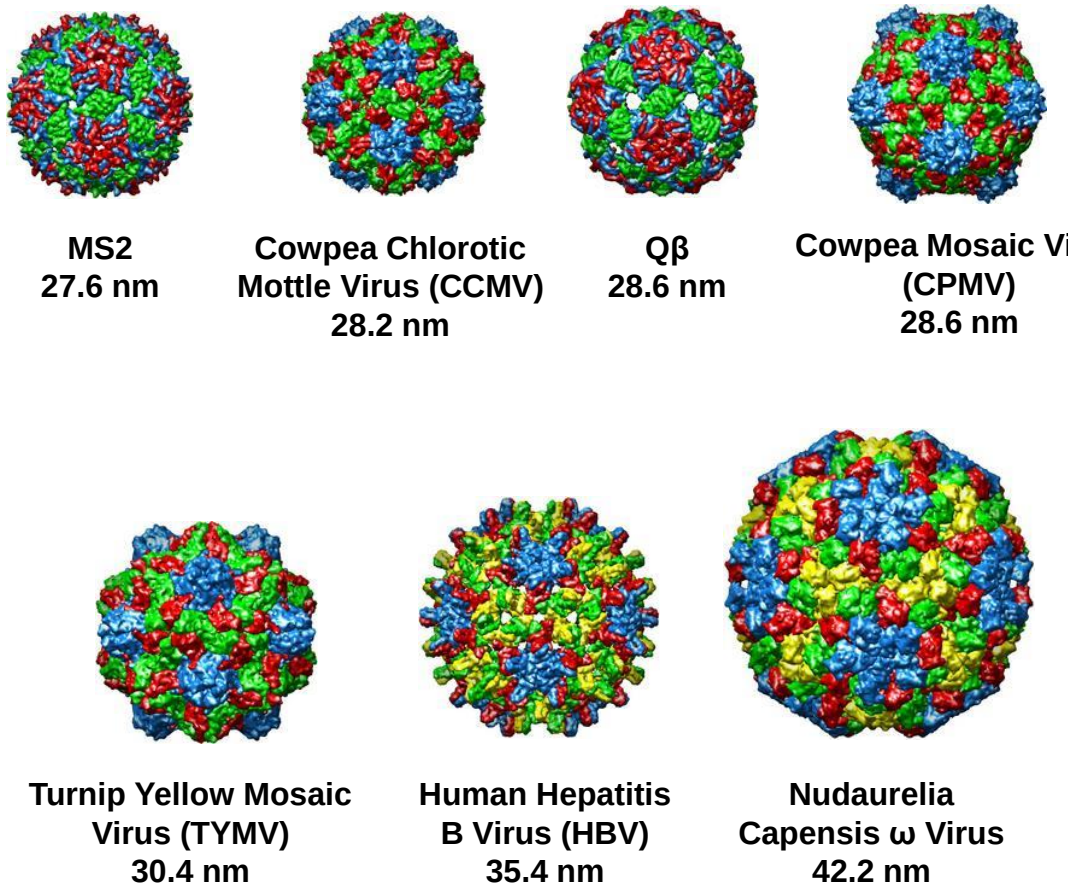


Figure 1.5 *A range of commonly studied virus-like particles with their names and average diameters. Images and data are from the VIPER database http://viperdb.scripps.edu.*[126]

Wang *et. al.* used the CPMV virion to demonstrate the use of VLPs as robust scaffolds for chemical modification, by chemically modifying cysteine residues with the 5-maleimidofluorescein dye moiety (Fig. 1.6 A).[56] The same group later showed that lysine residues could be modified in a similar way using the *N*-succinimidyl ester of fluorescein (Fig. 1.6 B).[57] Using these bioconjugation strategies the Alexa Fluor reagent has also been attached to VLP surfaces.[58,59]

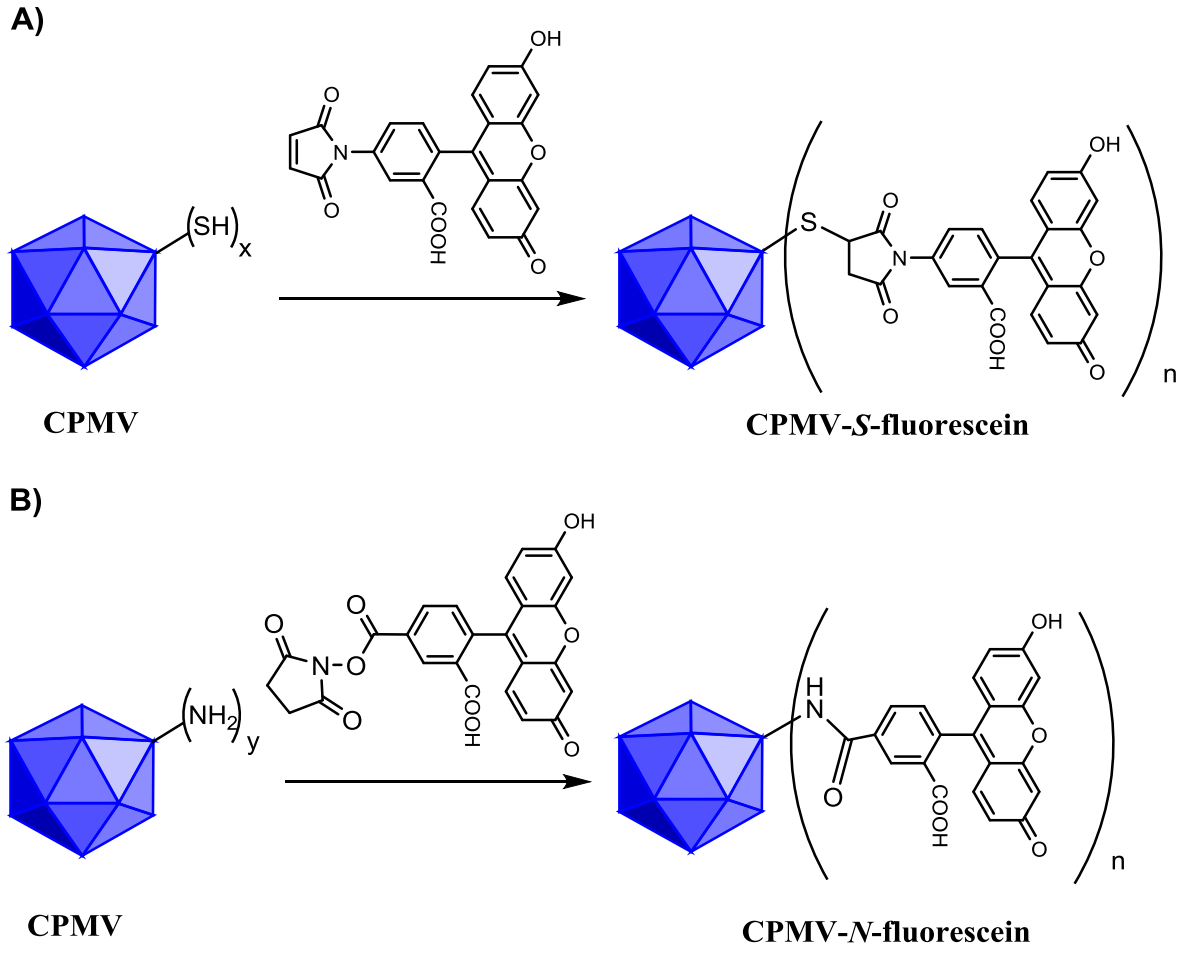


Figure 1.6 *Modification of the CPMV virus-like particle with 5-maleimidofluorescein dye via either: attack of a maleimide group by the sufhydryl group of cysteines (A), or attack of a N-hydroxysuccinimidyl ester by the amino group of lysines (B).*

Seminal work by Sharpless and co-workers has advanced the toolkit for on-surface virus conjugation with the copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition reaction, which allows facile coupling of an azide to an alkyne.[60–63] The addition of glycan ligands,[64,65] Gd^{3+} complexes for MRI contrast,[66] and AlexaFluor488[67] to the surface of VLPs has been enabled by this methodology.

Another versatile technique for virus modification is the incorporation of an unnatural amino acid baring a chemically reactive functional group, e.g. an alkene,[68] alkyne,[69] or azide[69] (Fig. 1.7). The unnatural amino acid can be installed by site-specific mutation and so this technique allows higher selectivity than modification of surface lysine and cysteine residues.

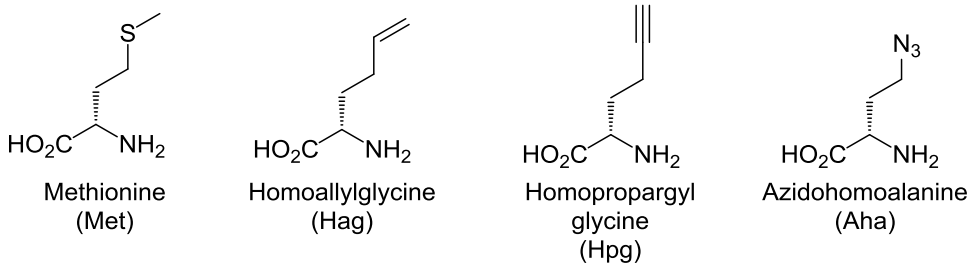


Figure 1.7 *Methionine and a selection of its analogues which are typically incorporated into proteins for bioconjugation reactions.*

1.3.2 Bacteriophages

Bacteriophages (phages) are viruses containing nucleic acids which infect, and multiply in bacteria.[70] Discovered by the English bacteriologist Frederick Twort in 1915,[71] phages were initially shown by d'Hérelle to be effective in the treatment of infections by specific bacteria.[72] However, a lack of understanding of the microbiology of phages, difficulties with production and purification, and the emergence of antibiotics, caused a diminution in research and development of phages as antibacterial agents.

The first single-stranded RNA bacteriophage, f2, was discovered in 1961.[73] Since then, much research has been carried out on the structure and biology of RNA bacteriophages and most recently a resurgence in phage therapy has developed.[74] With antibiotic resistance becoming an increasing problem, research into the use of phages as an alternative approach to combatting bacterial infections has shown increasing promise. In particular, phage cocktails (formulations containing 10-50 distinct phages) are being developed to broaden the infection types which the phage therapy can cure.[75]

1.3.3 Bacteriophage Qβ

Qβ is an *E. coli* single-stranded RNA bacteriophage, its capsid being a multimeric protein construct assembled from 180 protein monomers which self-assemble to form a virus with icosahedral symmetry.[76–78] The Qβ RNA genome codes for four proteins: the coat

protein,[79,80] a maturation protein (A2 protein, ~1 per virion),[81] a read-through protein (A1 protein, 3 to 5 per virion) resulting from occasional read-through of a UGA termination codon, and RNA replicase (Fig. 1.8).[82]

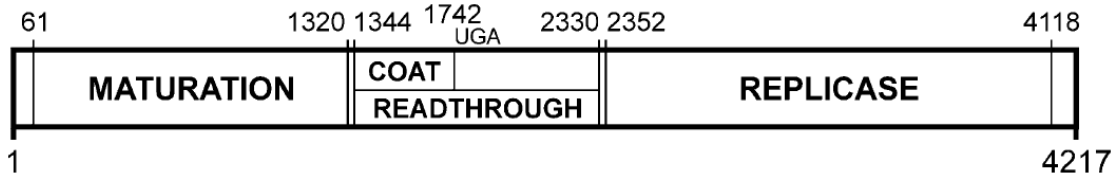


Figure 1.8 *The genetic map of Qβ. Reproduced with permission from The Springer Index of*

Takamatsu and Iso elucidated the capsid substructure of Qβ, by controlled dissociation and analysis of the resulting components by gel electrophoresis.[83] They isolated pentamers and hexamers of the coat protein as distinct chemical entities by addition of the denaturant sodium dodecyl sulphate (SDS). Crystallographic data show that hydrogen bonding is also involved in the association of adjacent coat proteins in the Qβ particle.[84] Takamatsu and Iso concluded that the Qβ capsid shell is composed of ~12 pentamers and ~20 hexamers.[83] After addition of the reducing agent β-mercaptoethanol and heating at 90 ^{o}C for 5 minutes to separate samples of pentamers and hexamers, electrophoresis showed a single dominant peak corresponding to the coat protein monomer molecular weight. It was therefore concluded that the coat proteins are held together to form these pentamers and hexamers principally by disulphide bonds.[83] Stockley and co-workers also reported pentameric species stabilised by intermolecular disulphide bonds, by using mass spectrometry.[85] Interestingly, they proposed that disulphide bonding was not as extensive at the 3-fold axis (i.e. in hexamers) due to potential flexibility in the A and C subunits, implied by the lack of electron density in the crystal structure data for residues 76-79 (*vide infra*).[84]

A crystal structure of Qβ phage has been solved to 3.5 Å resolution in 1996.[84] This structure reveals that the coat protein assumes three slightly different conformations within the capsid, termed A, B, and C (Fig. 1.9). Hence, Qβ has a triangulation number of $T = 3$ (the

number of different conformations of the subunit for each virus[76]). Each coat protein has 7 β strands and 2 α helices. The FG loop of each coat protein contains two cysteine residues, one at position 74 and the other at position 80.

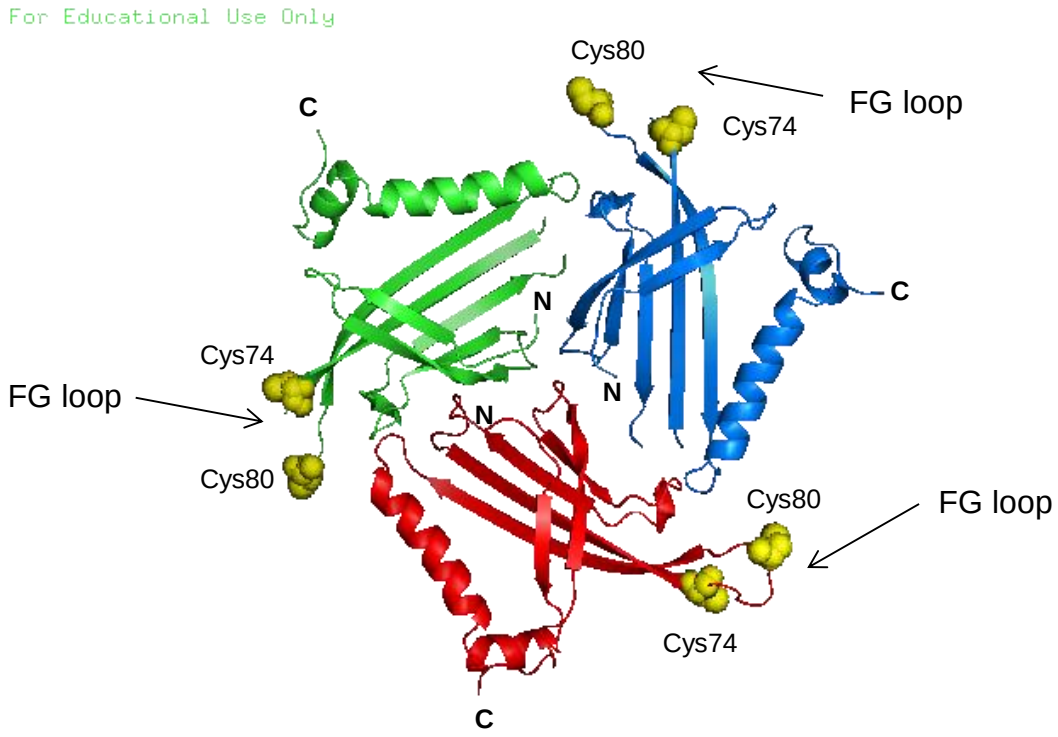
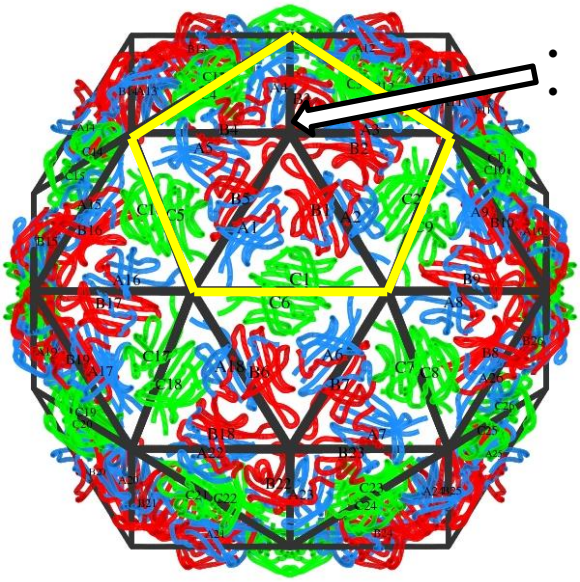


Figure 1.9 *The three conformations of the Qβ coat protein: A (blue), B (red), C (green). FG loops are labelled but are not resolved in the published X-ray diffraction data for protein conformations A and C. N- and C-termini are labelled, and cysteine residues are shown as yellow spheres. Drawn in Pymol from the Qβ crystal structure [84]. Protein Data Bank number 1QBE.*

The 180 Qβ coat proteins form a virus capsid with icosahedral symmetry (Fig. 1.10). At the vertices of the icosahedron, 5 B-conformer coat proteins are bonded together by disulphide bonds arranged around a 5-fold symmetry axis (Fig. 1.10 A). There are 12 such vertices, making 60 coat proteins in the B conformation per capsid. At the centre of a triangular face of the icosahedron, 3 A-conformer and 3 C-conformer coat proteins are bonded together by disulphide bonds in an alternating arrangement around a 3-fold symmetry axis (Fig. 1.10 B). There are 20 such faces, making 60 coat proteins in the A conformation and 60 coat proteins in the B conformation per capsid.

Finn and co-workers used systematic mutations to study the importance of different amino acids in the coat protein on the structure and stability of the Qβ particle.[86] Single and

A)



- Vertices of icosahedron
- 5-fold symmetry axis
- 5 B coat proteins disulphide bonded

B)

- Centre of triangular faces
- 3-fold symmetry axis
- 3 A coat proteins & 3 C coat proteins disulphide bonded

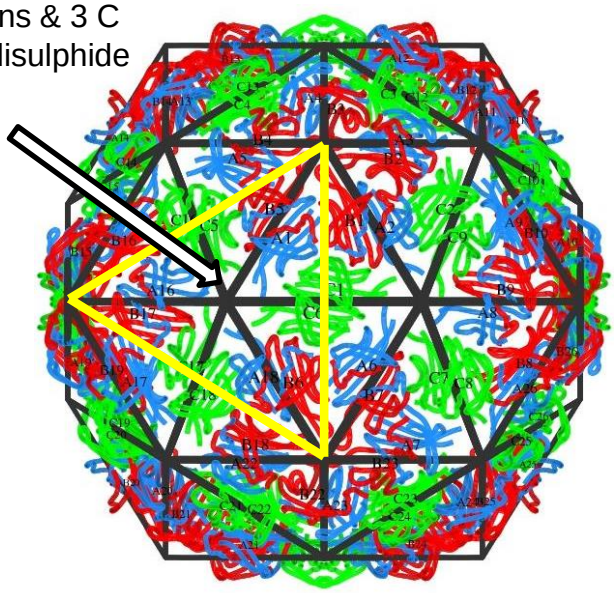


Figure 1.10 *The icosahedral structure of Qβ with 5-fold symmetry axes at a vertex encompassing 5 B coat proteins (A), and 6-fold symmetry axes at a triangular face encompassing 3 A and 3 C coat proteins (B). The image of the VLP is from the VIPER database http://viperdb.scripps.edu.*[126]

double mutants C74S and C74S/C80S were both observed to form particles but with melting temperatures ~20 ^{o}C less than that of wild-type particles.[86] This result demonstrates the importance of disulphide bonds to the capsid stability. Non-covalent interdimer bonding was shown not to have a significant effect and particles still formed correctly. However, non-

covalent *intradimer* bonding was crucial, particularly hydrophobic interactions. Non-covalent dimers are arranged such that the α-helices lie on top of the F and G β-sheets of both coat proteins. The interior region of this arrangement includes phenylanaline residues at positions 94 and 96 on the G-strand, helping to make the interior hydrophobic. Substituting the phenylanaline residues in these positions resulted in no particles being formed. Finally, substitution of leucine at position 35 for tryptophan resulted in VLPs having three structural classes, all smaller than wild-type Qβ (28 nm). 49% of the population had an average diameter of 16.5 nm, 39% with an average diameter of 20 nm and 12% were elongated (23.5 X 20 nm).

1.4 Incorporation of unnatural amino acids

1.4.1 Introduction

Chemical methods for peptide synthesis involve the successive amide bond formation between amino acids using solid phase peptide synthesis[87] and the use of peptide ligase enzymes.[88] Peptides can be ligated onto recombinantly expressed proteins.[89] However, these techniques are not practical for biosynthesizing large proteins or for producing large quantities of protein.

In vivo methods allow co-translational incorporation of unnatural amino acids, generating modified proteins in high quantities. There are two strategies used for the incorporation of unnatural amino acids: i) sense codon reassignment (residue-specific incorporation)[90–93] and ii) non-sense codon suppression (site-specific incorporation).[94–96]

1.4.2 Sense codon reassignment

Sense codon reassignment experiments originated with the landmark work of Cohen and co-workers, where methionine is replaced by selenomethionine, a methionine analogue, in

expressed protein.[97,98] In this technique, auxotrophic bacterial strains, which cannot biosynthesize a given natural amino acid, are used. Expression media deficient in this particular natural amino acid is doped with an unnatural analogue of the amino acid. During protein expression, wild-type aminoacyl-tRNA synthetases charge the unnatural amino acid onto the corresponding tRNAs. As a result, only unnatural amino acids which are sufficiently structurally similar to the natural analogue can be incorporated. Overall, the sense codon of the amino acid being replaced is reassigned to an unnatural analogue. Thus, a protein is produced in which the natural amino acid is substituted with its unnatural analogue.

Bacteria are initially grown in media containing all 20 natural amino acids, since most unnatural amino acids cannot support bacterial growth without their natural analogue. Once cells are grown to the logarithmic stage, cells are washed with media absent of the natural amino acid to be replaced before being resuspended in fresh minimal media, containing only 19 natural amino acids. A chemically defined medium is used for this purpose, such as M9 or derivatives thereof. The minimal medium is then supplemented with the unnatural amino acid to be incorporated before inducing protein expression. Tight control over expression of the recombinant protein is generally required since leaky expression in the growth stage, when all 20 amino acids are present, will lead to lower levels of unnatural amino acid incorporation. This control is achieved by using a strongly repressed promoter.

Van Hest *et. al.* showed successful incorporation of the unsaturated methionine analogues, homoallylglycine (Hag)[68] and homopropargylglycine (Hpg).[99] These residues represent useful chemical handles for a variety of post-translational modifications.[100] Similarly, incorporation of azidohomoalanine (Aha),[101] has provided a useful tool for protein bioconjugation. Together, *in vivo* incorporation of these amino acids (Hpg and Aha) have enabled the use of the Cu(I)-catalyzed [3 + 2] cycloaddition reaction on proteins (*vide supra*). Davis and co-workers have incorporated Hpg into the TIM-barrel protein *Ss*βG and the virus-

like particle Qβ to make fluoroglycoproteins.[102] Glycans bearing a fluorine atom and an azide group were coupled to the proteins using Cu(I)-catalyzed [3 + 2] cycloaddition *via* Hpg, incorporated into the proteins *via* reassignment of the Met ATG codon. However, no ^{19}F NMR studies on the resulting proteins were reported. As such, novel ^{19}F NMR on fluorinated Qβ are presented in this work (see Chapter 4). Virus-like "glycodendrinanoparticles" were developed by Ribeiro-Viana *et. al.* by conjugating mannose dendrimers onto Qβ virus-like particle *via* incorporation of Hpg.[103] These "glycodendrinanoparticles", presenting up to 1620 glycans per particle, were able to block binding to the carbohydrate-binding receptor DC-SIGN, hence inhibiting Ebola infection.

Finally, Honek *et. al.* have used the sense codon reassignment method to incorporate fluorinated methionine analogues into proteins as a ^{19}F NMR probe for protein structure dynamics. This work is discussed in more detail in Section 1.4.5.

1.4.3 Non-sense codon suppression

Non-sense codon suppression is a technique which allows incorporation of an unnatural amino acid at a specific site in the protein sequence. Key to this technique is the use of an aminoacyl-tRNA synthetase (aaRS)/tRNA pair which is orthogonal to the host (usually *E. coli*) translation system.

Site-directed mutagenisis is used to insert a stop codon at the desired site of unnatural amino acid incorporation in the protein sequence.[104] Non-sense suppressor tRNAs (tRNAs which insert natural amino acids in response to a stop codon) can be generated by mutation at the anticodon loop to recognize the stop codon. The suppressor tRNA must not be a substrate for *E. coli* aaRS; otherwise the suppressor tRNA would get charged with endogenous amino acids. The tRNA is charged with the unnatural amino acid either chemically or enzymatically

by an orthoganol aaRS. The orthoganol aaRS must only recognise the suppressor tRNA and not any of the *E. coli* tRNA.

To achieve orthoganolity, aaRS/tRNA pairs are often imported from archaea or eukaryotes. For example, the evolved *Methanococcus jannaschii* tyrosyl-tRNA/synthetase pair was the first orthogonal *E. coli* pair derived from archaea.[105] Spicer and Davis used this aaRS/tRNA pair to genetically incorporate *p*-iodophenylalanine into the maltose binding protein using UAG stop-codon suppression.[106] They used this aryl halide residue to perform Suzuki-Miyaura coupling on the protein surface.

1.4.4 Incorporation of multiple unnatural amino acids

Incorporation of multiple amino acids into proteins is a relatively new achievement in the field of genetic code expansion and provides a technique for more versatile protein modification. This method uses nonsense-codon suppression and requires two mutually orthogonal aaRS/tRNA pairs and two blank codons to incorporate the two different unnatural amino acids. Schultz and co-workers demonstrated this method by incorporating two different unnatural amino acids into myoglobin.[107] Schultz used one aaRS/tRNA pair containing a quadruplet codon, AGGA, derived from the archaebacterium *P. horikoshii*. In other work, Wan *et. al.* mutated the PylRS-pylT pair to suppress the UAA (ochre) codon.[108] They used this with a mutated *M. jannaschii* TyrRS which was able to charge *p*-azidophenylalanine and install it into the protein in response to the UAG (amber) codon at a specific site.

1.4.5 Incorporation of fluorinated amino acids

Fluorinated amino acids were synthesized and used in growth experiments as early as the 1940s.[109,110] The first fluorinated amino acid shown to be incorporated into bacterial proteins

was *p*-fluorophenylanaline in 1959.[97] Early studies in this field focused on the change in *E. coli* growth in the presence of fluorinated amino acids and the effect on activity of fluorinating enzymes.[111–113] However, Sykes *et. al.* set the stage for the use of fluorinated amino acids as ^{19}F NMR probes when they incorporated *m*-fluorotyrosine into alkaline phosphatase.[114] The chemical shift of ^{19}F has a much wider range than ^{1}H and is extremely sensitive to the local environment.[115] Together, these two features make ^{19}F NMR a very useful tool for analysing changes in protein structure. Thus, Sykes *et. al.* observed 11 distinct resonances in the ^{19}F NMR, in line with 11 tyrosine sites in each subunit of the dimeric alkaline phosphatase protein. They tentatively assigned tyrosine residues buried within the protein based on peak broadening resulting from restricted bond rotation.[114,116] During the same period Pratt *et. al.* successfully incorporated 4-, 5-, and 6-fluorotryptophan into lactose permease.[117]

^{19}F NMR has been used extensively for analysis of protein structure and conformational changes.[118] Significant for the work of this thesis is that of Honek and co-workers in the 1990s who demonstrated the use of fluorinated methionine analogues as ^{19}F NMR protein probes. Honek and co-workers achieved 70% incorporation of trifluoromethionine (TFM) into phage lysozyme LaL using sense codon reassignment.[119] They found that TFM did not support growth of the *E. coli* auxotroph (B834(DE3)) and so they initially grew the cells in 0.1 mM L-methionine before washing the cells to remove L-methionine and transferring to fresh minimal media supplemented only with L-TFM. Lysozyme contains three methionine residues (positions 1, 14, and 107), however four peaks were observed in the ^{19}F NMR spectrum (Fig. 1.11). Integration revealed that peaks A and B each corresponded to one TFM whereas peaks C and D together corresponded to the third TFM. Using mutagenesis to systematically replace one methionine position with a leucine, they were able to assign the peaks. They found that trifluoromethionine at position 107 gave rise to two resonances.[120]

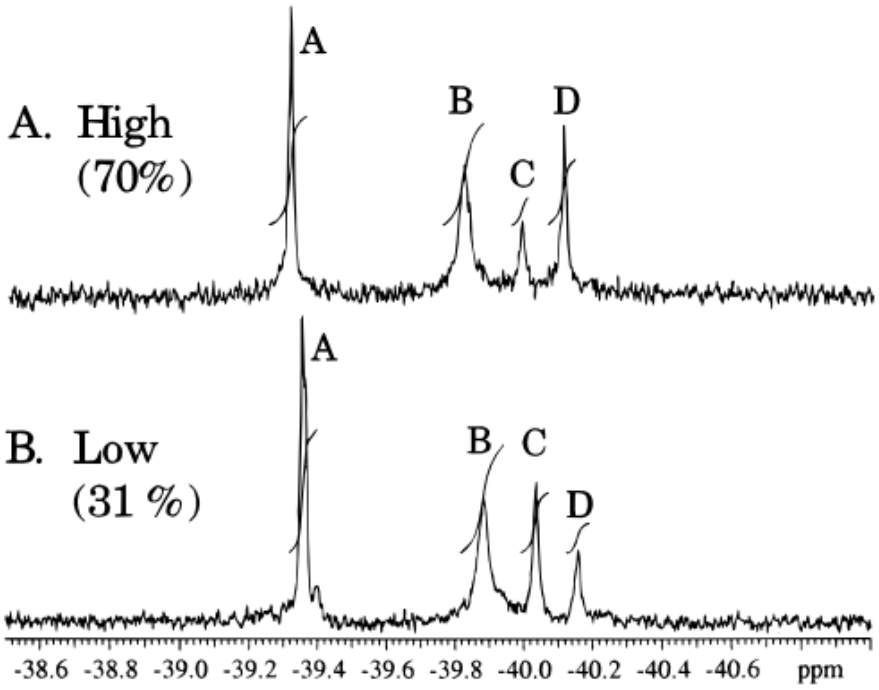


Figure 1.11 *^{19}F NMR spectra of high (A) and low (B) level incorporated trifluoromethionine-labelled lysozyme. Reprinted with permission from [119]. Copyright 1997 American Chemical Society.*

Interestingly, resonances for position 107 at both C *and* D was only observed when TFM, and not methionine, was at position 14. Incorporation of TFM into mutant M14L gave a peak at C but not D. From this, they reasoned that the increased size of TFM compared to methionine lead to a conformational change at position 14 in the protein that was relayed to position 107, giving rise to a changed environment at the latter position, hence the additional peak.

Honek and co-workers also investigated the use of difluoromethionine (DFM) for labelling proteins with a view to follow conformational changes by ^{19}F NMR.[121] They found that expression with L-DFM gave much higher protein yields and incorporation efficiency compared to expression with L-TFM. They achieved essentially 100% incorporation of DFM at all three methionine positions compared to 30-70% incorporation for expression with TFM (determined by mass spectrometry). Due to the diastereotopic relationship between the two fluorine atoms in DFM, the -$SCHF_2$ group gives rise to a double doublet pattern (Fig. 1.12 A). They used this splitting as a measure of internal amino acid packing since restricted

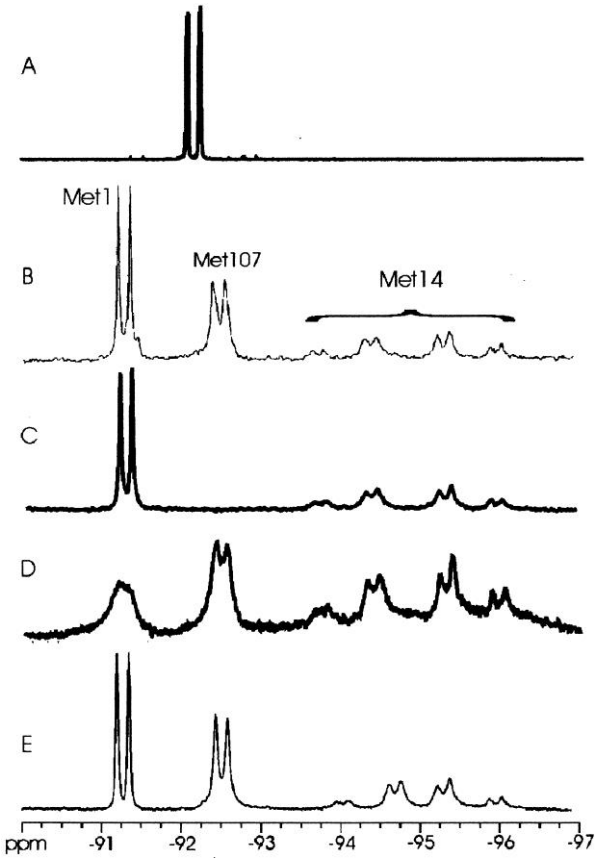


Figure 1.12 *^{19}F NMR spectra of: free DFM, demonstrating the ABX splitting pattern (A); completely labelled DFM-LaL lysozyme, indicating the identities of the three spin systems (B); DFM-LaL Met107Leu mutant, which lacks the Met107 signal (C); DFM-LaL in the presence of 1 mM GdEDTA$^-$ (D); DFM-LaL in the presence of 7 mM hexa-N-acetyl chitohexaose (E). Reprinted with permission from [121]. Copyright 1999 American Chemical Society.*

rotation about the S-C and C-C bonds of DFM leads to a greater chemical shift spread in the double doublet.[121]

In view of the differences in incorporation efficiency found between DFM and TFM, the crystal structures of these two amino acids bound to the *E. coli* methionyl-tRNA synthetase enzyme (MetRS) was solved and compared with that of natural methionine bound to the enzyme.[122] Binding of methionine by MetRS induces the rearrangement of several aromatic residues. W253, F300, and Y15 rock towards the bound methionine (Fig. 1.13). As a result, new stacking interactions between W229 and F304, and F300 and W253, are possible. The same conformational changes are induced on binding of DFM. However, in the case of TFM

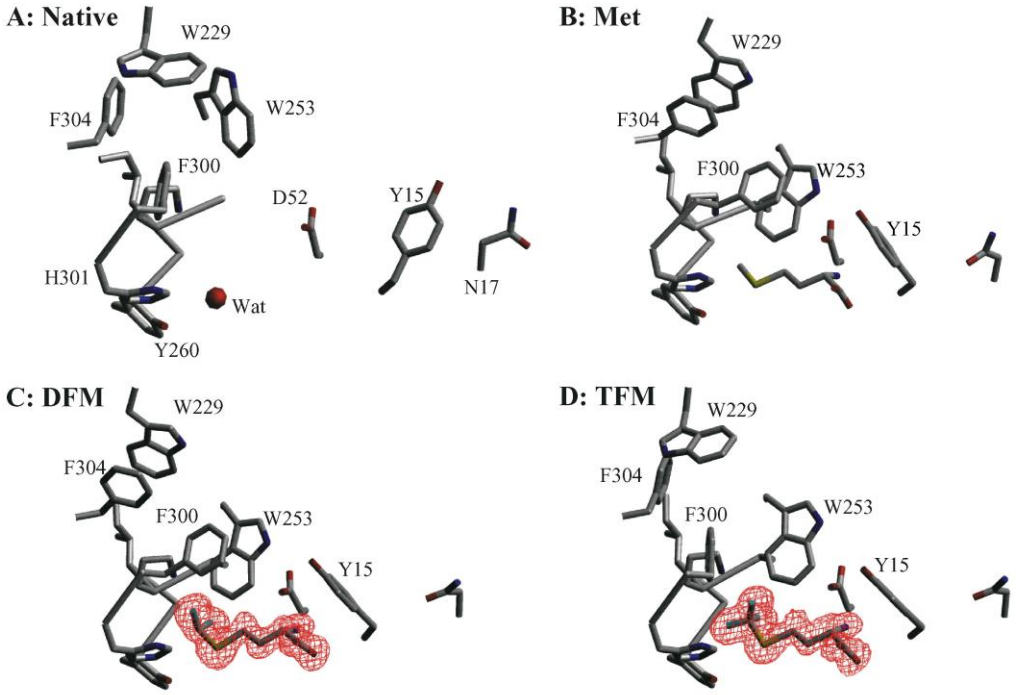


Figure 1.13 *Rearrangement of the methionine-binding site in free MetRS (A), and the complexes formed upon binding of methionine (B), DFM (C), and TFM (D). Reprinted with permission from [*[122]*]. Copyright 2003 Elsevier.*

binding, the movement of W253 is restricted due to the steric hindrance presented by the additional F atom. This is believed to be the cause of the much greater reduction of affinity of MetRS for TFM than DFM, compared to the native methionine.[122]

Furter was the first to incorporate a fluorinated amino acid *in vivo* by stop codon suppression.[123] A yeast phenylalanyl-tRNA synthetase and suppressor $tRNA^{Phe}$ pair was used, since these are orthogonal to the *E. coli* translation system. A *p*-fluoro-phenylalanine (*p*-F-Phe)-resistant *E. coli* strain, in which the PheRS excludes *p*-F-Phe from its binding site, was used. This was so that only the yeast tRNA/phenylalanyl-RS pair could deliver *p*-F-Phe. However, this did not result in site-*specific* incorporation, since 3-7% of all Phe residues were translated as *p*-F-Phe. Furthermore, the maximum efficiency was *p*-F-Phe at the stop codon site was 75%, the remainder being phenylalanine and lysine.

Mehl and co-workers were the first to site-specifically incorporate a fluorinated amino acid into a protein.[124] This required using an unnatural amino acid which is not recognised by natural *E. coli* tRNA synthetases. They selected, from a library of *Methanococcus jannaschii*

TyrRS mutants,[125] those mutants which supported *E. coli* growth in the presence of trifluoromethyl-L-phenylalanine (tfm-Phe). tfm-Phe was incorporated into the protein nitroreductase at position 124 in response to a TAG codon. No peaks corresponding to natural amino acid incorporation at position 124 were observed in ESI-MS analysis of the trypsin-digested fragment, thus demonstrating tfm-Phe had been selectively incorporated.

1.5 Thesis objectives

The initial aim of this research was to develop a new ^{19}F MR imaging agent with high sensitivity. Based on the literature review and accompanying discussions presented in this chapter, two key factors are required for a high-sensitivity imaging agent: 1) the molecule must have a high fluorine content and 2) all ^{19}F atoms should possess the same chemical shift. To achieve this incorporation of a fluorinated amino acid into the phage virus-like particle (VLP) Qβ was proposed for the following reasons. Firstly, the Qβ VLP consists of 180 copies of the coat protein thus a high payload of fluorine could potentially be introduced. Secondly, the Qβ coat proteins are related by symmetry in the iscosahedral VLP, thus the fluorinated amino acid in each coat protein should have the same chemical shift.

The first key objective set was to incorporate a fluorinated amino acid into the phage virus-like particle (VLP) Qβ. The sense codon reassignment method was proposed to incorporate the unnatural amino acid trifluoromethionine (TFM) into a mutated Qβ coat protein gene sequence containing only one methionine codon.

It was intended to analyse the resulting protein to establish whether the modified protein would maintain its ability to self-assemble into intact VLPs.

A further aim of this work was to investigate the potential of the fluorinated VLP as a ^{19}F MR imaging agent using ^{19}F NMR. It was also intended to carry out initial *in vivo* studies to explore the potential of the fluorinated VLP as a medical ^{19}F MRI probe.

1.6 References

1. Serres, S., Sarmiento, M., Hamilton, A., Mcateer, M. A. & Carbonell, W. S. Molecular MRI enables early and sensitive detection of brain metastases. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 6674–6679 (2012).

2. McAteer, M. A. *et al.* In vivo magnetic resonance imaging of acute brain inflammation using microparticles of iron oxide. *Nat. Med.* **13,** 1253–1258 (2007).

3. O'Farrell, A. C., Shnyder, S. D., Marston, G., Coletta, P. L. & Gill, J. H. Non-invasive molecular imaging for preclinical cancer therapeutic development. *Br. J. Pharmacol.* **169,** 719–735 (2013).

4. Kuchmiy, A. A., Efimov, G. A. & Nedospasov, S. A. Methods for in vivo molecular imaging. *Biochemistry-Moscow* **77,** 1339–1353 (2012).

5. Breckwoldt, M. O. *et al.* Tracking the inflammatory response in stroke in vivo by sensing the enzyme myeloperoxidase. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 18584–18589 (2008).

6. Rodríguez, E., Nilges, M., Weissleder, R. & Chen, J. W. Activatable magnetic resonance imaging agents for myeloperoxidase sensing: mechanism of activation, stability, and toxicity. *J. Am. Chem. Soc.* **132,** 168–177 (2010).

7. Ballatore, C., Lee, V. M. Y. & Trojanowski, J. Q. Tau-mediated neurodegeneration in Alzheimer's disease and related disorders. *Nat. Rev. Neurosci.* **8,** 663–672 (2007).

8. Maruyama, M. *et al.* Imaging of tau pathology in a tauopathy mouse model and in Alzheimer patients compared to normal controls. *Neuron* **79,** 1094–1108 (2013).

9. Phelps, M. E. PET: A biological imaging technique. *Neurochem. Res.* **16,** 929–940 (1991).

10. Rohren, E. M., Turkington, T. G. & Coleman, R. E. Radiology Clinical Applications of PET in Oncology 1. *Radiology* **231,** 305–332 (2004).

11. Gambhir, S. S. Molecular imaging of cancer with positron emission tomography. *Nat. Rev. Cancer* **2,** 683–693 (2002).

12. Phelps, M. E., Huang, S. C., Hoffman, E. J. & Selin, C. Tomographic Measurement of Local Cerebral Glucose Metabolic Rate in Humans with Validation of Method. *Ann. Neurol.* **6,** 371–388 (1978).

13. Som, P. *et al.* A Fluorinated Glucose Analog, 2-fluoro-2-deoxy-D-glucose (F-18): Nontoxic Tracer for Rapid Tumor Detection. *J. Nucl. Med.* **21,** 670–675 (1980).

14. Gatenby, R. A. & Gillies, R. J. Why do cancers have high aerobic glycolysis? *Nat. Rev. Cancer* **4,** 891–899 (2004).

15. Brown, R. S. & Wahl, R. L. Overexpression of Glut-1 Glucose Transporter in Human Breast Cancer An Immunohistochemical Study. *Cancer* **72,** 2979–2985 (1993).

16. Macheda, M. L., Rogers, S. & Best, J. D. Molecular and cellular regulation of glucose transporter (GLUT) proteins in cancer. *J. Cell. Physiol.* **202,** 654–662 (2005).

17. Mathupala, S. P., Ko, Y. H. & Pedersen, P. L. Hexokinase-2 bound to mitochondria: Cancer's stygian link to the "Warburg Effect" and a pivotal target for effective therapy. *Semin. Cancer Biol.* **19,** 17–24 (2009).

18. Kelloff, G. J. *et al.* Progress and Promise of FDG-PET Imaging for Cancer Patient Management and Oncologic Drug Development Management and Oncologic Drug Development. *Clin. Cancer Res.* **11,** 2785–2808 (2005).

19. ww.nhs.co.uk. Radiation. at <http://www.nhs.uk/Conditions/Radiation/Pages/Introduction.aspx>

20. McCullough, E. C. X-ray-transmission computed tomography. *Med. Phys.* **4,** 85–98 (1977).

21. Olsson, B., Aulie, A., Sveen, K. & Andrew, E. Human Pharmacokinetics of Iohexol A New Nonionic Contrast Medium. *Invest. Radiol.* **18,** 177–182 (1983).

22. Rabin, O., Manuel Perez, J., Grimm, J., Wojtkiewicz, G. & Weissleder, R. An X-ray computed tomography imaging agent based on long-circulating bismuth sulphide nanoparticles. *Nat. Mater.* **5,** 118–122 (2006).

23. Singh, J. & Daftary, A. Iodinated contrast media and their adverse reactions. *J. Nucl. Med. Technol.* **36,** 69–74 (2008).

24. Vahrmeijer, A. L., Hutteman, M., van der Vorst, J. R., van de Velde, C. J. H. & Frangioni, J. V. Image-guided cancer surgery using near-infrared fluorescence. *Nat. Rev. Clin. Oncol.* **10,** 507–518 (2013).

25. Van den Berg, N. S. *et al.* Concomitant radio- and fluorescence-guided sentinel lymph node biopsy in squamous cell carcinoma of the oral cavity using ICG-(99m)Tc-nanocolloid. *Eur. J. Nucl. Med. Mol. Imaging* **39,** 1128–1136 (2012).

26. Kalli, K. R. *et al.* Folate receptor alpha as a tumor target in epithelial ovarian cancer. *Gynecol. Oncol.* **108,** 619–626 (2008).

27. Van Dam, G. M. *et al.* Intraoperative tumor-specific fluorescence imaging in ovarian cancer by folate receptor-α targeting: first in-human results. *Nat. Med.* **17,** 1315–1319 (2011).

28. Chin, P. T., Beekman, C. A., Buckle, T., Josephson, L. & van Leeuwen, F. W. Multispectral visualization of surgical safety-margins using fluorescent marker seeds. *Am. J. Nucl. Med. Mol. Imaging* **2,** 151–162 (2012).

29. Ntziachristos, V. Going deeper than microscopy: the optical imaging frontier in biology. *Nat. Methods* **7,** 603–614 (2010).

30. Weissleder, R. & Ntziachristos, V. Shedding light onto live molecular targets. *Nat. Med.* **9,** 123–128 (2003).

31. Lauterbur, P. C. Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance. *Nature* **242,** 190–191 (1973).

32. Hornak, J. P. The Basics of MRI. (1996). at <http://www.cis.rit.edu/htbooks/mri/inside.htm>

33. Hore, P. *Nuclear Magnetic Resonance*. (Oxford University Press, 2006).

34. Claridge, T. D. W. *High-Resolution NMR Techniques in Organic Chemistry*. (Elsevier, 2009).

35. Damadian, R. Tumor detection and nuclear magnetic resonance. *Science* **171,** 1151–1153 (1971).

36. Lauffer, R. B. Paramagnetic metal-complexes as water proton relaxation agents for NMR imaging - Theory and design. *Chem. Rev.* **87,** 901–927 (1987).

37. Holland, G. N., Bottomley, P. A. & Hinshaw, W. S. 19F magnetic resonance imaging. *J. Magn. Reson.* **28,** 133–136 (1977).

38. McFarland, E., Koutcher, J. A., Rosen, B. R., Teicher, B. & Brady, T. J. In vivo 19F NMR imaging. *J. Comput. Assist. Tomogr.* **9,** 8–15 (1985).

39. De Graff, R. A. *In vivo NMR spectroscopy: Principles and Techniques*. (Wiley-Blackwell, 2007).

40. Ahrens, E. T., Flores, R., Xu, H. Y. & Morel, P. A. In vivo imaging platform for tracking immunotherapeutic cells. *Nat. Biotechnol.* **23,** 983–987 (2005).

41. Higuchi, M. *et al.* 19F and 1H MRI detection of amyloid beta plaques in vivo. *Nat. Neurosci.* **8,** 527–533 (2005).

42. Yanagisawa, D. *et al.* Preferred Features of a Fluorine-19 MRI Probe for Amyloid Detection in the Brain. *J. Alzheimers. Dis.* **39,** 617–631 (2014).

43. Meyer, K. L., Carvlin, M. J., Mukherji, B., Sloviter, H. A. & Joseph, P. M. Fluorinated Blood Substitute retention in the Rat Measured by Fluorine-19 Magnetic Resonance Imaging. *Invest. Radiol.* **27,** 620–627 (1992).

44. Jiang, Z.-X., Liu, X., Jeong, E.-K. & Yu, Y. B. Symmetry-Guided Design and Fluorous Synthesis of a Stable and Rapidly Excreted Imaging Tracer for 19F MRI. *Angew. Chem. Int. Ed. Engl.* **48,** 4755–4758 (2009).

45. Thurecht, K. J. *et al.* Functional Hyperbranched Polymers: Toward Targeted in Vivo 19F Magnetic Resonance Imaging Using Designed Macromolecules. *J. Am. Chem. Soc.* **132,** 5336–5337 (2010).

46. Takaoka, Y. *et al.* Self-assembling nanoprobes that display off/on 19F nuclear magnetic resonance signals for protein detection and imaging. *Nat. Chem.* **1,** 557–561 (2009).

47. Mizukami, S. *et al.* Paramagnetic Relaxation-Based 19F MRI Probe To Detect Protease Activity. *J. Am. Chem. Soc.* **130,** 794–795 (2008).

48. Zeltins, A. Construction and Characterization of Virus-Like Particles: A Review. *Mol. Biotechnol.* **53,** 92–107 (2013).

49. Pattenden, L. K., Middelberg, A. P. J., Niebert, M. & Lipin, D. I. Towards the preparative and large-scale precision manufacture of virus-like particles. *Trends Biotechnol.* **23,** 523–529 (2005).

50. Kushnir, N., Streatfield, S. J. & Yusibov, V. Virus-like particles as a highly efficient vaccine platform: diversity of targets and production systems and advances in clinical development. *Vaccine* **31,** 58–83 (2012).

51. Plummer, E. M. & Manchester, M. Viral nanoparticles and virus-like particles: platforms for contemporary vaccine design. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **3,** 174–196 (2010).

52. Roldão, A., Mellado, M. C. M., Castilho, L. R., Carrondo, M. J. T. & Alves, P. M. Virus-like particles in vaccine development. *Expert Rev. Vaccines* **9,** 1149–1176 (2010).

53. Vacher, G., Kaeser, M. D., Moser, C., Gurny, R. & Borchard, G. Recent advances in mucosal immunization using virus-like particles. *Mol. Pharm.* **10,** 1596–1609 (2013).

54. Kang, S.-M., Song, J.-M., Quan, F.-S. & Compans, R. W. Influenza vaccines based on virus-like particles. *Virus Res.* **143,** 140–146 (2009).

55. Sapsford, K. E. *et al.* Functionalizing nanoparticles with biological molecules: developing chemistries that facilitate nanotechnology. *Chem. Rev.* **113,** 1904–2074 (2013).

56. Wang, Q., Lin, T., Tang, L., Johnson, J. E. & Finn, M. G. Icosahedral Virus Particles as Addressable Nanoscale Building Blocks. *Angew. Chem. Int. Ed. Engl.* **41,** 459–462 (2002).

57. Wang, Q., Kaltgrad, E., Lin, T., Johnson, J. E. & Finn, M. G. Natural Supramolecular Building Blocks: Wild-Type Cowpea Mosaic Virus. *Chem. Biol.* **9,** 805–811 (2002).

58. Lewis, J. D. *et al.* Viral nanoparticles as tools for intravital vascular imaging. *Nat. Med.* **12,** 354–360 (2006).

59. Shriver, L. P., Koudelka, K. J. & Manchester, M. Viral nanoparticles associate with regions of inflammation and blood brain barrier disruption during CNS infection. *J. Neuroimmunol.* **211,** 66–72 (2009).

60. Rostovtsev, V. V, Green, L. G., Fokin, V. V & Sharpless, K. B. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angew. Chem. Int. Ed. Engl.* **41,** 2596–2599 (2002).

61. Wang, Q. *et al.* Bioconjugation by Copper(I)-Catalyzed Azide-Alkyne [3 + 2] Cycloaddition. *J. Am. Chem. Soc.* **125,** 3192–3193 (2003).

62. Sen Gupta, S. *et al.* Accelerated Bioorthogonal Conjugation: A Practical Method for the Ligation of Diverse Functional Molecules to a Polyvalent Virus Scaffold. *Bioconjug. Chem.* **16,** 1572–1579 (2005).

63. Hong, V., Presolski, S. I., Ma, C. & Finn, M. G. Analysis and Optimization of Copper-Catalyzed Azide-Alkyne Cycloaddition for Bioconjugation. *Angew. Chem. Int. Ed. Engl.* **48,** 9879–9883 (2009).

64. Kaltgrad, E. *et al.* On-Virus Construction of Polyvalent Glycan Ligands for Cell-Surface Receptors. *J. Am. Chem. Soc.* **130,** 4578–4579 (2008).

65. Yin, Z. *et al.* Boosting Immunity to Small Tumor-Associated Carbohydrates with Bacteriophage Qβ Capsids. *ACS Chem. Biol.* **8,** 1253–1262 (2013).

66. Prasuhn, D. E., Yeh, R. M., Obenaus, A., Manchester, M. & Finn, M. G. Viral MRI contrast agents: coordination of Gd by native virions and attachment of Gd complexes by azide-alkyne cycloaddition. *Chem. Comm.* 1269–1271 (2007).

67. Pokorski, J. K., Hovlid, M. L. & Finn, M. G. Cell Targeting with Hybrid Qβ Virus-Like Particles Displaying Epidermal Growth Factor. *ChemBioChem* **12,** 2441–2447 (2011).

68. Van Hest, J. C. M. & Tirrell, D. A. Efficient introduction of alkene functionality into proteins in vivo. *FEBS Lett.* **428,** 68–70 (1998).

69. Strable, E. *et al.* Unnatural Amino Acid Incorporation into Virus-Like Particles. *Bioconjug. Chem.* **19,** 866–875 (2008).

70. Calendar, R. & Abedon, S. T. *The Bacteriophages*. (Oxford University Press, 2006).

71. Twort, F. W. An investigation on the nature of ultra-microscopic viruses. *Lancet* **2,** 1241–1243 (1915).

72. D'Hérelle, F. *The Bacteriophage and its Behaviour*. 490–497, 540–549 (Williams and Wilkins, 1926).

73. Loeb, T. & Zinder, N. D. A Bacteriophage Containing RNA. *Proc. Natl. Acad. Sci. U. S. A.* **47,** 282–289 (1961).

74. Kutateladze, M. & Adamia, R. Bacteriophages as potential new therapeutics to replace or supplement antibiotics. *Trends Biotechnol.* **28,** 591–595 (2010).

75. Chan, B. K., Abedon, S. T. & Loc-Carrillo, C. Phage cocktails and the future of phage therapy. *Future Microbiol.* **8,** 769–83 (2013).

76. Caspar, D. & Klug, A. Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.* **27,** 1–24 (1962).

77. Overby, L. R., Barlow, G. H., Doi, R. H., Jacob, M. & Spiegelman, S. Comparison of Two Serologically Distinct Ribonucleic Acid Bacteriophages I. Properties of the Viral Particles. *J. Bacteriol.* **91,** 442–448 (1966).

78. Overby, L. R., Barlow, G. H., Doi, R. H., Jacob, M. & Spiegelman. Comparison of Two Serologically Distinct Bacteriophages II . Properties of the Nucleic Acids and Coat Proteins. *J. Bacteriol.* **92,** 739–745 (1966).

79. Maita, T. & Konigsberg, W. The Amino Acid Sequence of the Qβ Coat Protein. *J. Biol. Chem.* **246,** 5003–5024 (1971).

80. Konigsberg, W., Maita, T. & Katze, J. Amino-acid Sequence of the Qβ Coat Protein. *Nature* **227,** 271–273 (1970).

81. Karnik, S. & Billeter, M. The lysis function of RNA bacteriophage Qβ is mediated by the maturation (A2) protein. *Embo J.* **2,** 1521–1526 (1983).

82. Horiuchi, K., Webster, R. E. & Matsuhashi, S. Gene Products of Bacteriaphage Qβ. *Virology* **45,** 429–439 (1971).

83. Takamatsu, H. & Iso, K. Chemical evidence for the capsomeric structure of phage Qβ. *Nature* **298,** 819–824 (1982).

84. Golmohammadi, R., Fridborg, K., Bundule, M., Valegard, K. & Liljas, L. The crystal structure of bacteriophage Qβ at 3.5 Å resolution. *Structure* **4,** 543–554 (1996).

85. Ashcroft, A. E. *et al.* Engineering thermal stability in RNA phage capsids via disulphide bonds. *J. Nanosci. Nanotechnol.* **5,** 2034–2041 (2005).

86. Fiedler, J. D. *et al.* Engineered Mutations Change the Structure and Stability of a Virus-Like Particle. *Biomacromolecules* **13,** 2339–2348 (2012).

87. Kent, S. B. H. Chemical Synthesis of Peptides and Proteins. *Annu. Rev. Biochem.* **57,** 957–989 (1988).

88. Jackson, D. Y. *et al.* A Designed Peptide Ligase for Total Synthesis of Ribonuclease A with Unnatural Catalytic Residues. *Science* **266,** 243–247 (1994).

89. Chang, T. K., Jackson, D. Y., Burnier, J. P. & Wells, J. A. Subtiligase: A tool for semisynthesis of proteins. *Proc. Natl. Acad. Sci. U. S. A.* **91,** 12544–12548 (1994).

90. Budisa, N. *et al.* Toward the experimental codon reassignment in vivo: protein building with an expanded amino acid repertoire. *FASEB J.* **13,** 41–51 (1999).

91. Link, A. J. & Tirrell, D. A. Reassignment of sense codons in vivo. *Methods* **36,** 291–298 (2005).

92. Schoffelen, S., Lambermon, M. H. L., van Eldijk, M. B. & van Hest, J. C. M. Site-Specific Modification of Candida antarctica Lipase B via Residue-Specific Incorporation of a Non-Canonical Amino Acid. *Bioconjug. Chem.* **19,** 1127–1131 (2008).

93. Johnson, J. A., Lu, Y. Y., Van Deventer, J. A. & Tirrell, D. A. Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications. *Curr. Opin. Chem. Biol.* **14,** 774–780 (2010).

94. Liu, C. C. & Schultz, P. G. Adding New Chemistries to the Genetic Code. *Annu. Rev. Biochem.* **79,** 413–444 (2010).

95. Beatty, K. E. & Tirrell, D. A. Noncanonical Amino Acids in Protein Science and Engineering. *Nucleic Acids Mol. Biol.* **22,** 127–153 (2009).

96. O'Donoghue, P., Ling, J., Wang, Y.-S. & Söll, D. Upgrading protein synthesis for synthetic biology. *Nat. Chem. Biol.* **9,** 594–598 (2013).

97. Munier, R. & Cohen, G. N. Incorporation d'analogues structuraux d'aminoacides dans les proteines bacteriennes au cours de leur synthese in vivo. *Biochim. Biophys. Acta* **31,** 378–391 (1959).

98. Cowie, D. B. & Cohen, G. N. Biosynthesis by Escherichia coli of active altered proteins containing selenium instead of sulfur. *Biochim. Biophys. Acta* **26,** 252–261 (1957).

99. Van Hest, J. C. M., Kiick, K. L. & Tirrell, D. A. Efficient incorporation of unsaturated methionine analogues into proteins in vivo. *J. Am. Chem. Soc.* **122,** 1282–1288 (2000).

100. Floyd, N., Vijayakrishnan, B., Koeppe, A. R. & Davis, B. G. Thiyl Glycosylation of Olefinic Proteins: S-Linked Glycoconjugate Synthesis. *Angew. Chem. Int. Ed.* **48,** 7798–7802 (2009).

101. Kiick, K. L., Saxon, E., Tirrell, D. A. & Bertozzi, C. R. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 19–24 (2002).

102. Boutureira, O. *et al.* Fluoroglycoproteins: ready chemical site-selective incorporation of fluorosugars into proteins. *Chem. Comm.* **46,** 8142–8144 (2010).

103. Ribeiro-Viana, R. *et al.* Virus-like glycodendrinanoparticles displaying quasi-equivalent nested polyvalency upon glycoprotein platforms potently block viral infection. *Nat. Commun.* **3,** 1–8 (2012).

104. Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C. & Schultz, P. G. A General Method for Site-Specific Incorporation of Unnatural Amino Acids into Proteins. *Science* **244,** 182–188 (1989).

105. Wang, L., Magliery, T. J., Liu, D. R. & Schultz, P. G. A New Functional Suppressor tRNA/Aminoacyl-tRNA Synthetase Pair for the in Vivo Incorporation of Unnatural Amino Acids into Proteins. *J. Am. Chem. Soc.* 5010–5011 (2000).

106. Spicer, C. D. & Davis, B. G. Palladium-mediated site-selective Suzuki-Miyaura protein modification at genetically encoded aryl halides. *Chem. Comm.* **47,** 1698–1700 (2011).

107. Anderson, J. C. *et al.* An expanded genetic code with a functional quadruplet codon. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 7566–7571 (2004).

108. Wan, W. *et al.* A Facile System for Genetic Incorporation of Two Different Noncanonical Amino Acids into One Protein in Escherichia coli. *Angew. Chem. Int. Ed. Engl.* **49,** 3211–3214 (2010).

109. Mitchell, H. K. & Niemann, C. The Competitive Inhibition of the Metabolism of Alpha-Amino Acids by their Halogenated Analogs. *J. Am. Chem. Soc.* **69,** 1232 (1947).

110. Boyer, P., Evans, J. & Phillips, P. H. The toxicity and effect of certain fluorinated derivatives of phenylalanine and tyrosine on the basal metabolic rate of the rat. *J. Pharmacol. Exp. Ther.* **73,** 176–185 (1941).

111. Rennert, O. & Anker, H. On the Incorporation of 5,5,5-Trifluoroleucine into Proteins of E. coli. *Biochemistry* **2,** 471–476 (1963).

112. Browne, D. T., Kenyon, G. L. & Hegeman, G. D. Incorporation of monofluorotryptophans into protein during the growth of Escherichia Coli. *Biochem. Biophys. Res. Commun.* **39,** 13–19 (1970).

113. Pratt, E. A. & Ho, C. Incorporation of Fluorotryptophans into Proteins of Escherichia coli. *Biochemistry* **14,** 3035–3040 (1975).

114. Sykes, B. D., Weingarten, H. I. & Schlesinger, M. J. Fluorotyrosine Alkaline Phosphatase from Escherichia coli: Preparation, Properties, and Fluorine-19 Nuclear Magnetic Resonance Spectrum. *Proc. Natl. Acad. Sci. U. S. A.* **71,** 469–473 (1974).

115. Gerig, J. T. Fluorine NMR of Proteins. *Prog. Nucl. Magn. Reson. Spectrosc.* **26,** 293–370 (1994).

116. Hull, W. E. & Sykes, B. D. Fluorotyrosine Alkaline Phosphatase. 19F Nuclear Magnetic Resonance Relaxation Times and Molecular Motion of the Individual Fluorotyrosines. *Biochemistry* **13,** 3431–3437 (1974).

117. Pratt, E. & Ho, C. Incorporation of Fluorotryptophans into Proteins of Escherichia coli. *Biochemistry* **14,** 3035–3040 (1975).

118. Danielson, M. A. & Falke, J. J. Use of F-19 NMR to probe protein structure and conformational changes. *Annu. Rev. Biophys. Biomol. Struct.* **25,** 163–195 (1996).

119. 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 19F NMR. *Biochemistry* **36,** 3404–3416 (1997).

120. Duewel, H. S., Daub, E., Robinson, V. & Honek, J. F. Elucidation of Solvent Exposure, Side-Chain Reactivity, and Steric Demands of the Trifluoromethionine Residue in a Recombinant Protein. *Biochemistry* **40,** 13167–13176 (2001).

121. Vaughan, M. D., Cleve, P., Robinson, V., Duewel, H. S. & Honek, J. F. Difluoromethionine as a Novel 19F NMR Structural Probe for Internal Amino Acid Packing in Proteins. *J. Am. Chem. Soc.* **121,** 8475–8478 (1999).

122. Crepin, T. *et al.* Use of Analogues of Methionine and Methionyl Adenylate to Sample Conformational Changes During Catalysis in Escherichia coli Methionyl-tRNA Synthetase. *J. Mol. Biol.* **332,** 59–72 (2003).

123. Furter, R. Expansion of the genetic code: Site-directed p-fluoro-phenylalanine incorporation in Escherichia coli. *Protein Sci.* **7,** 419–426 (1998).

124. Jackson, J. C., Hammill, J. T. & Mehl, R. A. Site-specific incorporation of a F-19-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity. *J. Am. Chem. Soc.* **129,** 1160–1166 (2007).

125. Wang, L., Brock, A., Herberich, B. & Schultz, P. G. Expanding the genetic code of Escherichia coli. *Science* **292,** 498–500 (2001).

126. Carrillo-Tripp, M. *et al.* VIPERdb2: an enhanced and web API enabled relational database for structural virology. *Nucleic Acids Res.* **37,** D436–442 (2009).

127. Friedman, S. D. & Weber, H. *The Springer Index of Viruses*. 795–798 (Springer New York, 2011).

2. Development of fluorinated Qβ

2.1 Rational design of the fluorinated virus-like particle

As described in Chapter 1, the initial aim of this project was to genetically incorporate a fluorinated amino acid into bacteriophage Qβ so the resultant virus-like particle (VLP) could be used as a ^{19}F NMR/MR imaging agent. The two methods for unnatural amino acid incorporation *in vivo*, sense codon reassignment and non-sense codon suppression, are described in Chapter 1. An advantage of the latter method is that it allows site-specific incorporation, since the gene is mutated to insert a stop codon at the precise position where modification is desired. In contrast, sense codon reassignment leads to global replacement of the natural amino acid being replaced. In the context of fluorinated amino acid incorporation, the fluorinated amino acids at different sites would potentially have different fluorine chemical shifts[1–4] and hence lead to an overall reduction in intensity of the ^{19}F signal of the protein. Multiple peaks can also give rise to ghost images making image interpretation difficult.[5]

The advantage of the sense codon reassignment method, however, is that in practical terms it is much simpler as it does not require development of a suppressor tRNA/aaRS pair and instead exploits the natural protein translational machinery of the host. The wild-type Qβ gene sequence[6] (Fig. 2.1 A) does not contain Met residues, therefore it was envisaged that mutation of the sequence to include one Met residue at a specific site could enable 'pseudo-site specific' methionine analogue incorporation *via* the sense codon reassignment methodology. Thus, replacement of one natural residue at a specific site by a fluorinated analogue in each monomer of Qβ should lead to 180 fluroinated residues in the assembled VLP in the same chemical environment and therefore with the same chemical shift (Fig. 2.1 B).

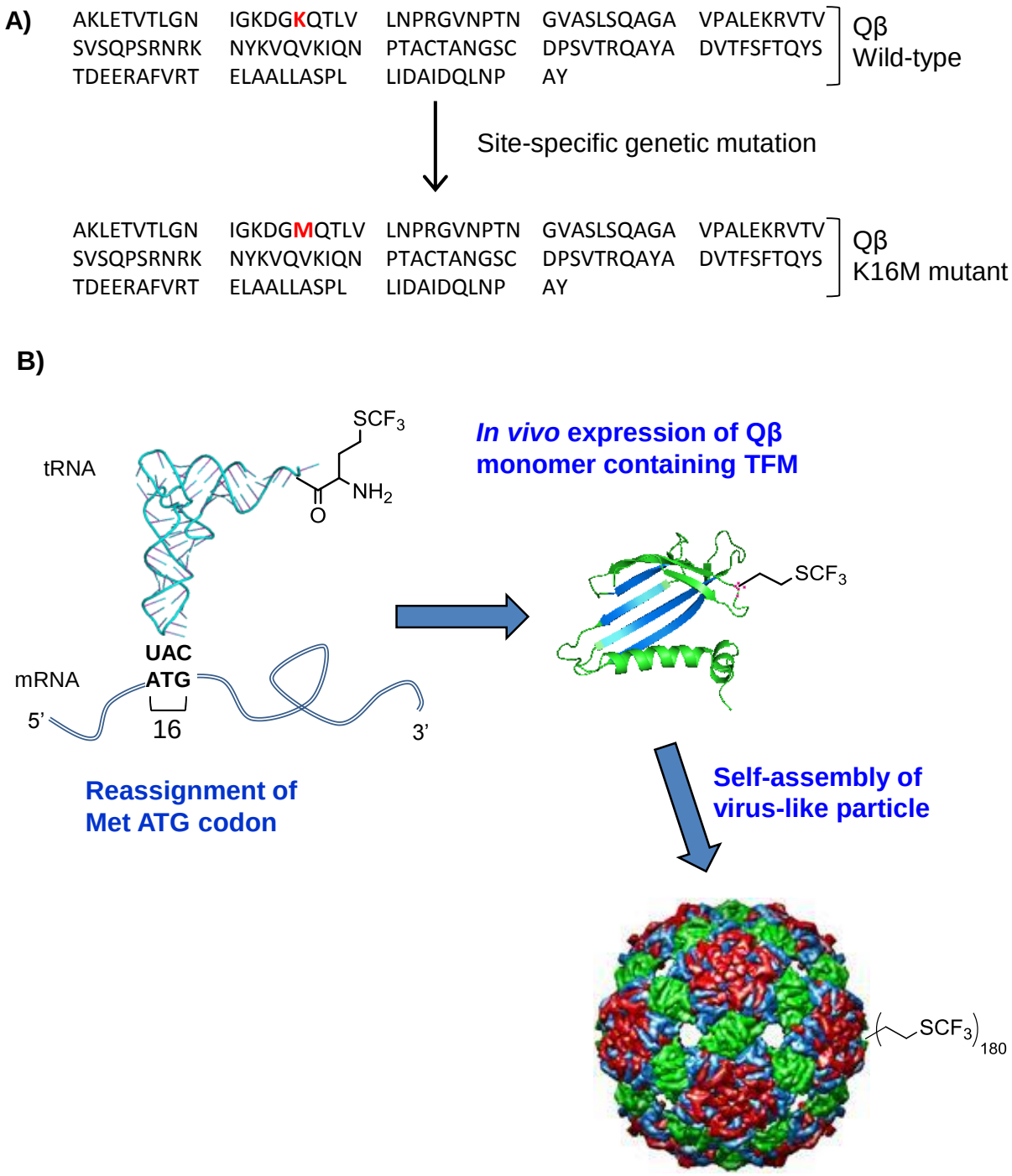


Figure 2.1 *A: Sequence of the wild-type and K16M mutant Qβ coat proteins; B: Strategy for the genetic incorporation of trifluoromethionine (TFM) into bacteriophage Qβ.*

A further reason for choosing methionine as the residue to be substituted is the presence of the relatively flexible carbon chain which should result in a smaller effective rotational correlation time.[7,8] Therefore, modifying the terminus of the methionine side chain with ^{19}F nuclei should give a sharper NMR peak than fluorinated analogues of (most) other amino

acid residues, making it more easily detectable in larger species e.g. the VLP used in this project.

Difluoromethionine (DFM) has been shown to incorporate into lambda bacteriophage lysozyme with a much higher efficiency than trifluoromethionine (TFM) under the same conditions.[3] Despite this advantage, it was decided to use TFM in the current work since the two fluorine atoms in DFM are diastereotopically related and therefore are chemically inequivalent. The ^{19}F signal of DFM gives rise to an ABX splitting pattern (two sets of quartets) (Figure 2.2). In view of the imaging application of the VLP, this peak splitting would reduce the signal-to-noise ratio, and hence the sensitivity, per ^{19}F nucleus of the VLP. Additionally, since the DFM has a third fewer ^{19}Fs compared to TFM, this effect on its own would reduce the overall sensitivity of the Qβ-F VLP. Since the primary goal of the project was to develop VLPs for imaging applications, engineering the particle such that it has maximal sensitivity is a key objective. Hence, work using DFM was not pursued.

E. coli can only efficiently utilise the L-enantiomer of methionine and therefore, likely, the L-enantiomers of methionine analogues that are compatible with its protein biosynthetic machinery. However, synthesis of pure L-TFM is expensive and very time consuming. Therefore it was decided that the racemic form of TFM (i.e. DL-TFM) would be synthesized. The quantity of the racemic mixture added would simply be double that of the quantity of the L-enantiomer estimated to be required for a given protein production. Thus a compromise is made between the quantity of amino acid which can be produced and the difficulty of the amino acid synthesis.

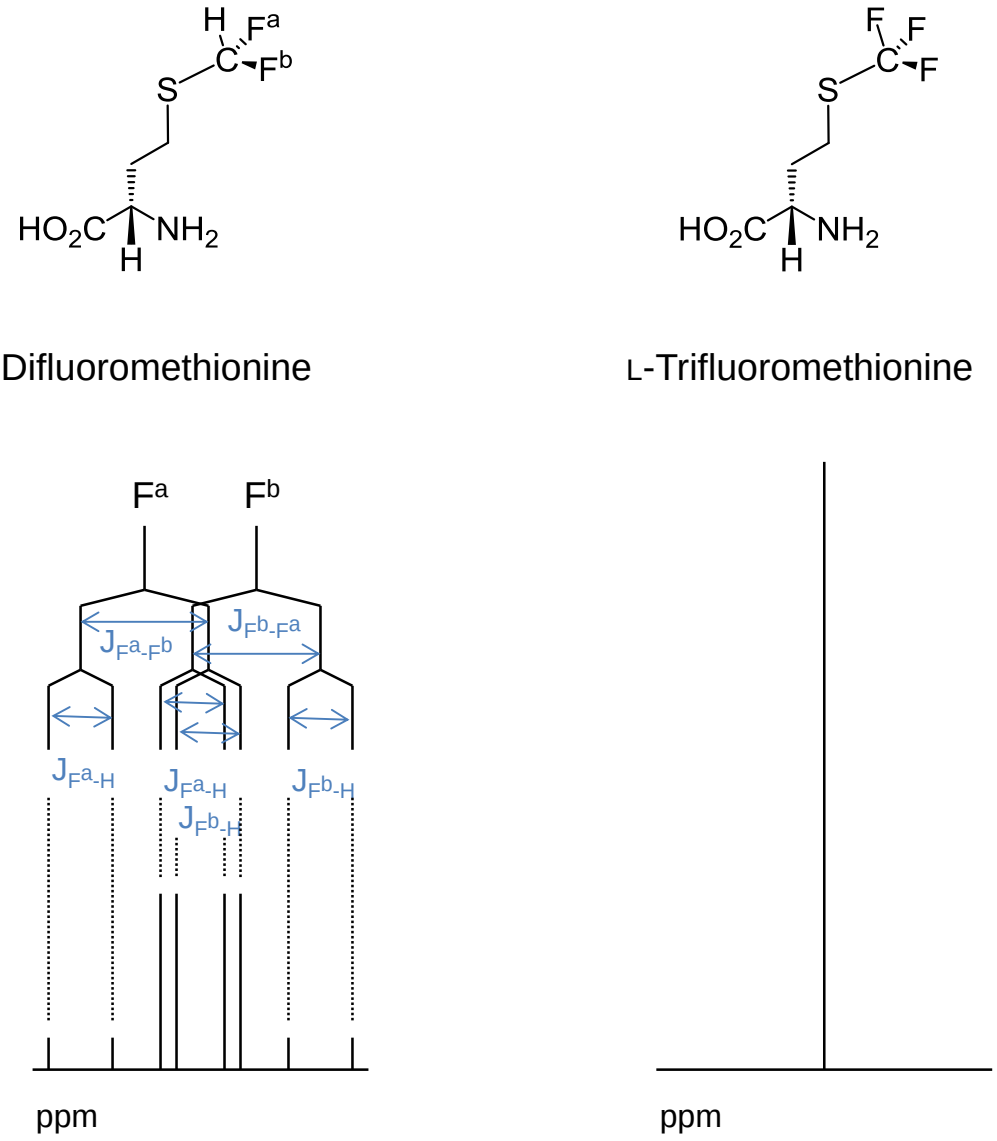


Figure 2.2 *Structures and schematic ^{19}F NMR spectra of DFM and TFM, and the ABX splitting pattern for DFM.*

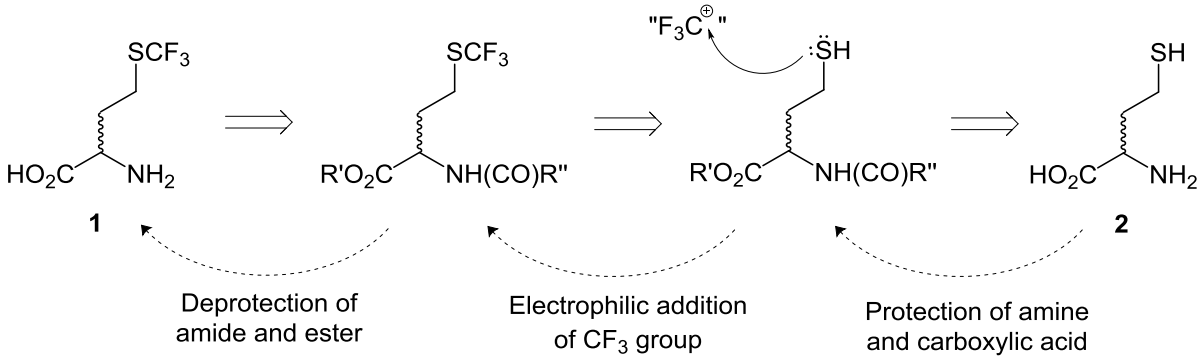
2.2 Synthesis of DL-trifluoromethionine

A synthesis of TFM was required for incorporation into the Qβ K16M variant (see Fig. 2.1). Several approaches to synthesize DL-TFM were investigated - these are discussed below.

2.2.1 Synthesis of DL-TFM using Togni's reagent

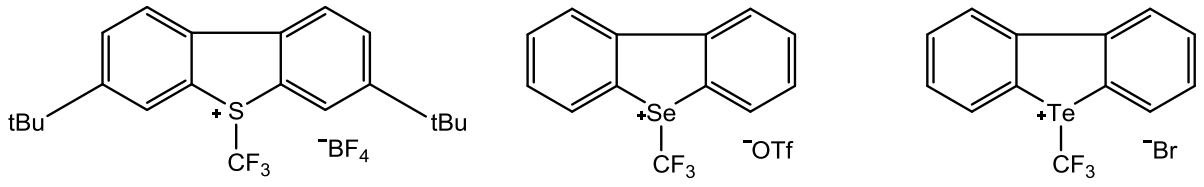
A synthetic route towards DL-TFM **1** *via* direct introduction of a trifluoromethyl group to homocysteine **2** was initially proposed (Scheme 2.1). Most sources of the CF_3 synthon are

nucleophilic; the most common being Ruppert's reagent (Me_3SiCF_3)[9] which transfers a $^-CF_3$ *via* an anionic pentacoordinated silicon species generated by activation of the reagent with a fluoride anion. However, addition to the nucleophilic sulfur of homocysteine **2** would require an electrophilic source of CF_3. Protection of the amine and carboxylic acid was deemed necessary to prevent their reaction with an electrophilic CF_3 reagent.



Scheme 2.1 *Retrosynthesis of DL-TFM via trifluoromethylation of homocysteine.*

For some time, the only class of electrophilic sources of the CF_3^+ synthon had been the *S*-, *Se*, and *Te*-(trifluoromethyl)dibenzotihio-, -seleno, and -tellurophenium salts, which are relatively difficult to synthesize (Fig. 2.3).[10–12]

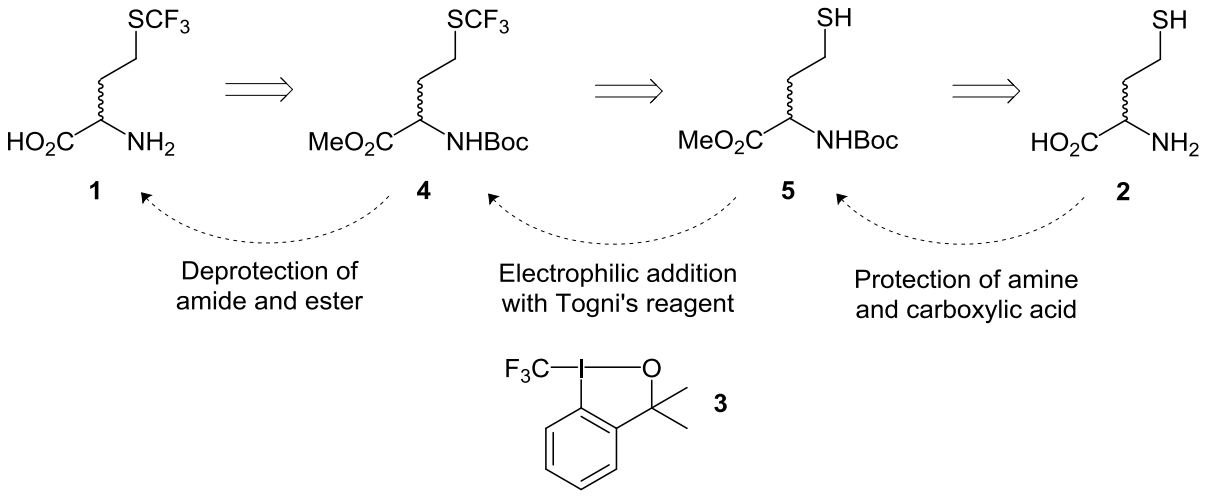


I **II** **III**

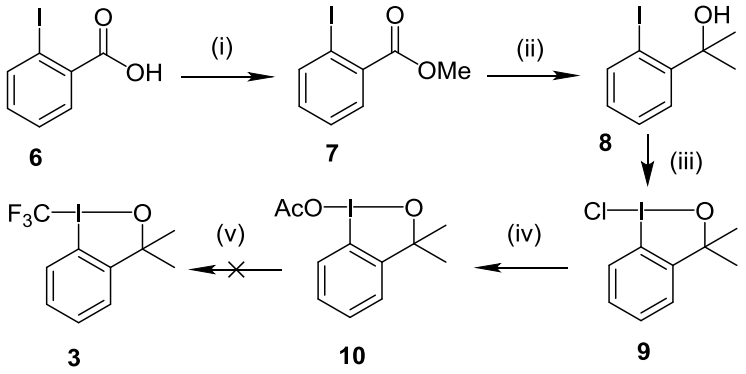
Figure 2.3 *Electrophilic trifluoromethylating agents: 3,7-di-tert-butyl-S-(trifluoromethyl)dibenzothiophenium tetrafluoroborate (**I**), Se-(trifluoromethyl)dibenzoselenophenium triflate (**II**), and Te-(trifluoromethyl)dibenzotellurophenium bromide (**III**).*

Recently, a new electrophilic trifluoromethylating reagent **3** (Scheme 2.2), which is derived from easily accessible and inexpensive precursors, has been reported by Togni and co-workers.[13,14] Scheme 2.2 shows the planned use of Togni's reagent **3** in the synthesis of the target, fluorinated amino acid **1**. Thus, the synthesis of Togni's reagent **3** was attempted

(Scheme 2.3). Unfortunately, despite many different conditions attempted, the synthesis of **3** was not achieved and in each case only intermediate **10** was isolated (see pages 70-71 of the Experimental section for details).



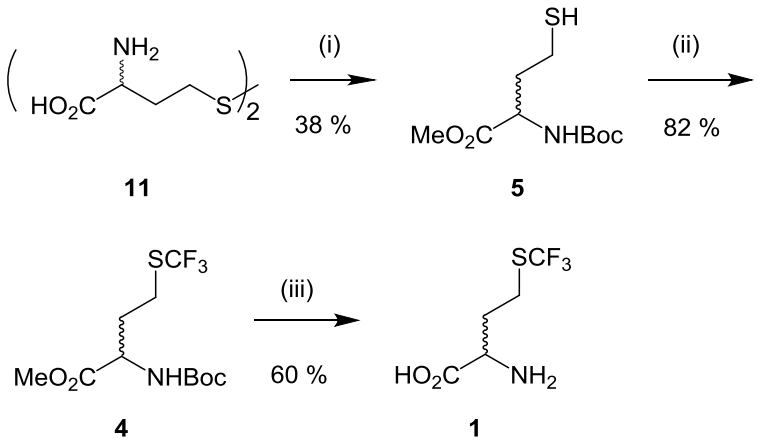
Scheme 2.2 *Retrosynthesis of DL-trifluormethionine **1** via trifluoromethylation of homocysteine **2** with Togni's reagent **3**.*



Scheme 2.3 *Reagents and conditions: **(i)** MeOH, conc. H_2SO_4, reflux overnight **(ii)** (a) Et_2O, MeMgBr, -20 ^oC, 16 hours (b) Reflux 1.5 hours (c) 0 ^oC, NH_4Cl (aq) **(iii)** $C(CH_3)_3(OCl)$, CCl_4, 30 minutes. Attempted **(iv)** & **(v)** (a) MeCN, KOAc (b) -17 ^oC, Me_3SiCF_3, TBAT in MeCN, 16 hours (c) -12 ^oC, Me_3SiCF_3 (d) 3 hours RT.*

In order to consider the reaction with Togni's reagent **3** for future amino acid production, **3** was purchased from Sigma-Aldrich. Since DL-homocysteine **2** is significantly more expensive than DL-homocystine **11**, the synthesis of the target amino acid **1** was started from the disulphide **11** (Scheme 2.4). Synthesis of **1** was achieved *via* a slight modification of the

published route.[15] It was decided to protect the amine and the carboxylic acid groups of DL-homocystine **11** as the Boc carbamate and methyl ester, respectively, to ease purification of the crude reduced product by flash chromatography. Hence, **11** was methyl and Boc protected using standard conditions, followed by reduction of the disulphide with tributylphosphine to give compound **5** (38%). Subsequent reaction with Togni's reagent **3** in methanol gave **4** in 82% yield. Hydrolysis of the carbamate in dichloromethane and trifluoroacetic acid followed by hydrolysis of the methyl ester in LiOH in THF gave the desired amino acid **1** in 60% yield.

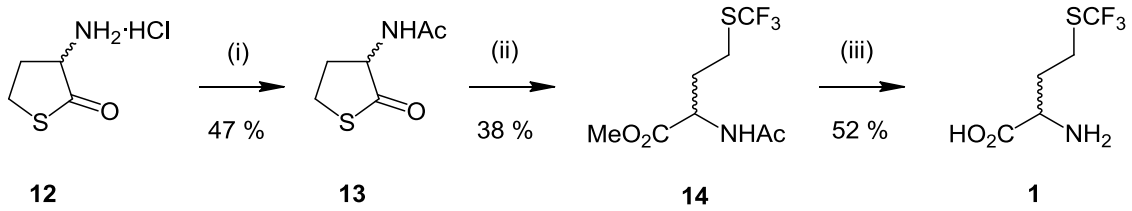


Scheme 2.4 *Reagents and conditions:* ***(i)*** *(a) CH_3COCl, MeOH (b) Boc_2O, Et_3N, CH_2Cl_2 (c) PBu_3, MeOH-H_2O;* ***(ii)*** *Togni's reagent* ***3****, MeOH;* ***(iii)*** *(a) CH_2Cl_2, TFA (b) LiOH-H_2O, THF, 2 hrs RT (c) DOWEX®.*

2.2.2 Synthesis of DL-TFM using CF_3I/hν (Honek's method)

Despite the high yield of the reaction with Togni's reagent **3**, the high cost of this reagent (£173.50/gram at the time of purchase) prompted me to investigate an alternative synthesis of the target amino acid for larger scale production. The synthesis by Honek and co-workers[1] was pursued (Scheme 2.5). The hydrochloride salt of homocysteine thiolactone **12** was neutralised with Et_3N in dichloromethane to liberate the free amine. Acetylation of this amine was achieved by addition of acetyl chloride to afford the protected thiolactone **13**. Methoxylation of the acetyl-protected thiolactone **13** generated the sodium thiolate *in situ*. Addition of excess trifluoromethyl iodide and subsequent ultra violet irradiation yielded the

protected amino acid **14 (**38%). The reaction is proposed to occur *via* a radical mechanism.[1] Deprotection of the methyl ester and the amide gave the free amino acid **1**. Conditions for the deprotection reported by Holzberger *et. al.*[16] (8 eq. NaOH heated under reflux for 3.5 hours) were found to give very low yields, and so the reaction was refluxed for 48 hours after which a higher recovery was obtained. After purification by DOWEX® column, **1** was isolated in 52% yield on gram scale (~ 1.2 g).



Scheme 2.5 *Reagents and conditions:* ***(i)*** *(a) 4 eq. Et_3N, CH_2Cl_2, 0 ^oC (b) 1.5 eq. AcCl;* ***(ii)*** *(a) NaOMe/MeOH (b) CF_3I, hv;* ***(iii)*** *(a) LiOH-H_2O, MeOH, 1hr (b) Reflux 48 hrs.*

Attempts were made at simplifying the route outlined in Scheme 2.5. It was thought that ultra violet radiation could be omitted by proceeding *via* nucleophilic attack of the thiolate to the electrophilic carbon of trifluoroiodomethane. DL-Homocysteine thiolactone hydrochloride **12** was treated with 5 M NaOH to open the lactone ring and generate the thiolate anion. The reaction mixture was adjusted to pH 8 before adding CF_3I. The product was found to be DL-homocysteine by ^1H and ^{13}C NMR, since there was a lack of the characteristic $^{13}CF_3$ quartet in the ^{13}C NMR. This result supports the proposal that the reaction with CF_3I and ultra violet light proceeds *via* a radical mechanism.

In another potential alternative route, the reaction was started from the unprotected thiolactone hydrochloride **12**, in an attempt to avoid the acetyl deprotection step. The same concentration of starting material **12** was used as in previous reactions starting from **13** (200 mM). However, after addition of NaOMe, a white precipitate formed. This may be due to polymerisation caused by reaction of the thiolactone with a free amine of another homocysteine thiolactone molecule (Fig. 2.4). TLC analysis (15% acetone in water) of the

product showed no target compound and only a spot at R_f = 0.0. Considering the high polarity of the solvent system used, such a low R_f is probably indicative of polymer formation. To reduce the likelihood of polymerisation, the reaction was repeated using a concentration of starting material ten times less than in the previous reaction (i.e. 20 mM). DL-TFM was isolated, however, as expected from such low dilution, the yield was low (23%) and so this route was not pursued.

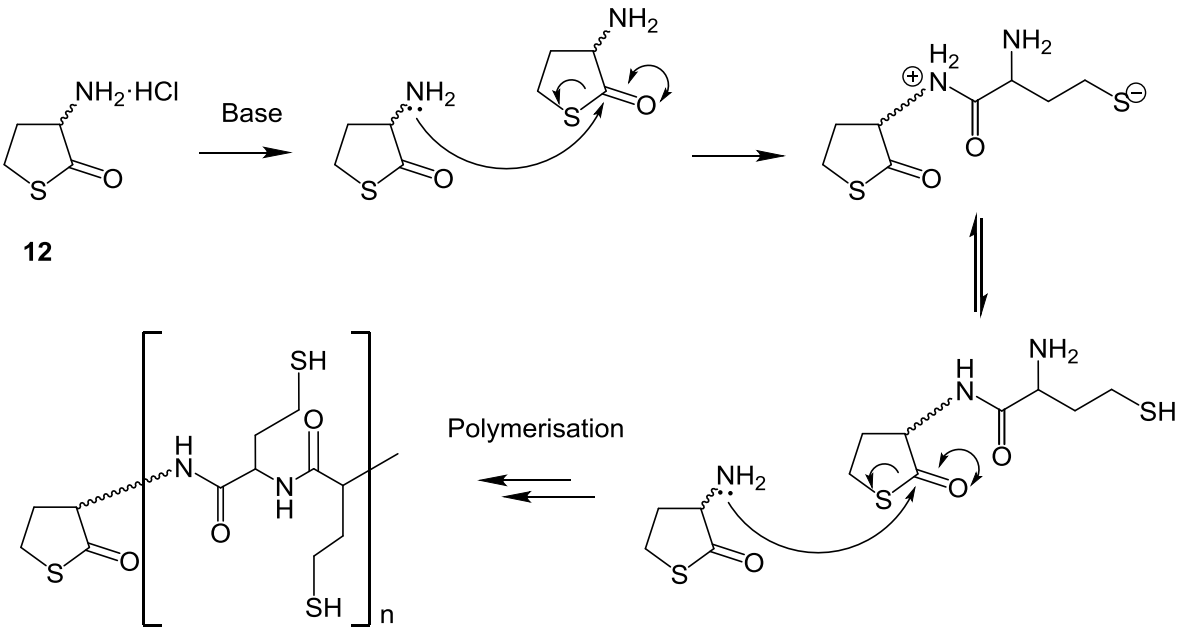
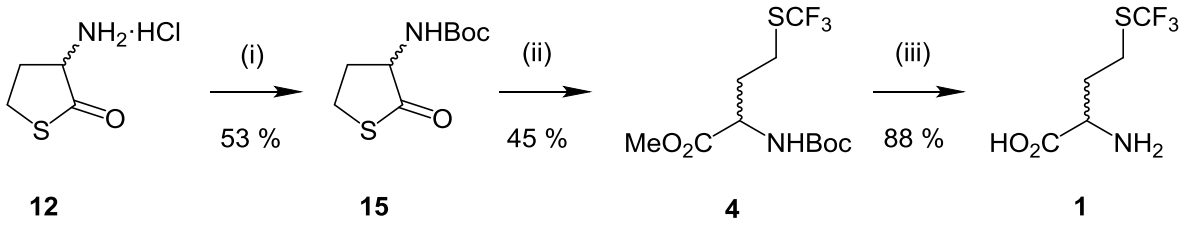


Figure 2.4 *Proposed mechanism for the polymerisation of homocysteine thiolactone.*

Deprotection of the amide group of **14** required harsh conditions which it was considered might be the cause of the modest yield (Scheme 2.5). To avoid the need for such conditions the amine of **12** was protected with a Boc group, which can be more efficiently deprotected than the acetamide group (Scheme 2.6). DL-Homocysteine thiolactone hydrochloride **12**, available from Alrdich (£0.66/gram), was Boc-protected by reacting with Boc anhydride (Boc_2O) in the presence of triethylamine in dichloromethane (Scheme 2.6 i). The crude product was recrystallized from hot ethyl acetate. The reaction of Scheme 2.6 i typically gave **15** in ~50-55% yield.

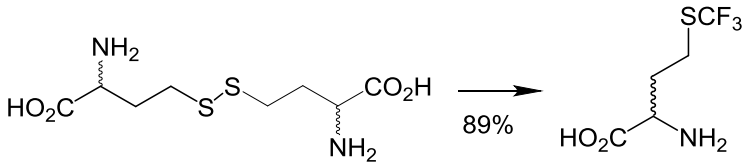


Scheme 2.6 *Reagents and conditions:* ***(i)*** *(a) Et_3N, CH_2Cl_2, 0 °C (b) Boc_2O;* ***(ii)*** *(a) NaOMe/MeOH (b) CF_3I, hν* ***(iii)*** *(a) CF_3CO_2H, CH_2Cl_2 (b) LiOH-H_2O, THF (c) DOWEX®.*

The Boc protected homocysteine thiolactone **15** was converted to the Boc and methoxyl protected trifluoromethionine **4** by treatment of the thiolactone **15** with NaOMe followed by addition of excess trifluoromethyl iodide and subsequent ultra violet irradiation (45% yield) (Scheme 2.6 ii). Deprotection of protected trifluoromethionine **4** was carried out under standard conditions: trifluoroacetic acid in dichloromethane to deprotect the carbamate, and lithium hydroxide in tetrahydrofuran to deprotect the methyl ester, to yield DL-TFM **1** (Scheme 2.6 iii, 88% yield).

2.2.3 Synthesis of DL-TFM using Birch reduction conditions

During the course of my work, a new synthesis of trifluoromethionine was reported by Yasui *et. al.*[17] This approach represented a great improvement as it is carried out in one pot from commercial homocystine and proceeds in high yield. Birch reduction conditions, sodium metal in liquid ammonia, were used to reduce the disulfide bond of DL-homocystine **16** (Scheme 2.7). Excess CF_3I gas was condensed into the reaction vessel and subsequently trapped by the thiol(ate) to form the trifluoromethylated product, the free trifluoromethionine amino acid **1** (Scheme 2.7). The target compound **1** was purified on a DOWEX® column eluted with 2% NH_4OH aqueous solution (89%). I was able to carry out this procedure on 10 gram scale which enabled ready access to sufficient material for subsequent protein expressions. This was the method of choice for the remainder of the research described herein.



16 **1**

Scheme 2.7 *Reagents and conditions: Na, CF_3I, liquid NH_3, -78 °C.*[17]

2.3 Incorporation of trifluoromethionine into bacteriophage Qβ

2.3.1 Initial expression studies

p75M plasmids carrying the Qβ K16M gene were kindly donated by Professor M. G. Finn. The plasmids were transformed into *E. coli* B834(DE3) cells, methionine auxotrophs. Initially, the procedure followed for cell growth was that published by previous researchers of the Davis group.[18] SelenoMet™ (Molecular Dimensions) minimal medium was used instead of the M9 minimal salts often used by other research groups for similar experiments.[19] Most unnatural amino acids do not support growth of auxotrophic cells and so cells are initially incubated with all 20 natural amino acids to generate a culture of *E. coli* in the logarithmic phase of growth, to grow enough cells to produce an appreciable quantity of protein and to enable efficient protein expression; typically cultures are grown until OD_{600} = 0.6 - 1.0. Using minimal medium in the initial out-growth, as opposed to LB (Luria Broth), also allows control over the concentration of L-methionine in the medium.

Cells were initially grown in SelenoMet™ medium supplemented with 0.27 mM L-Met until OD_{600} = 0.6 - 0.8.[18] At this point the cells were washed thrice with SelenoMet™ base containing no methionine, to remove residual methionine. The cells were then transferred to fresh, pre-warmed medium supplemented with 1.3 mM DL-TFM. The cells were incubated at 37 °C for 30 minutes and then at 30 °C for 30 minutes, to deplete the intracellular pool of methionine prior to initiation of protein expression. Protein expression was then induced by addition of 1 mM IPTG. After incubation at 37 °C for 6 hours, the cells were harvested by

centrifugation. The cells were lysed by sonication and to the lysate was added a 1:1 v/v mixture of n-butanol and chloroform. This mixture denatures host proteins and causes them, with membrane lipids, to coagulate. These components can then be easily separated from the VLPs (which remain in the aqueous phase) by low-speed centrifugation. After low-speed centrifugation, the aqueous layer was decanted and the VLPs were precipitated by addition of PEG8000 and NaCl. The precipitated VLPs were collected by centrifugation and dissolved in tris-buffered saline (TBS). Figure 2.5 shows a SDS-PAGE gel of samples at the different stages of purification. Qβ protein can be clearly seen in the cell lysate amongst other proteins (Lane 2). A significant amount of protein impurity is removed after solvent extraction and precipitation with of PEG8000 and NaCl (Lane 5). The streaking observed is likely due to different ordered oligomers and partially assembled VLPs (Lane 5).

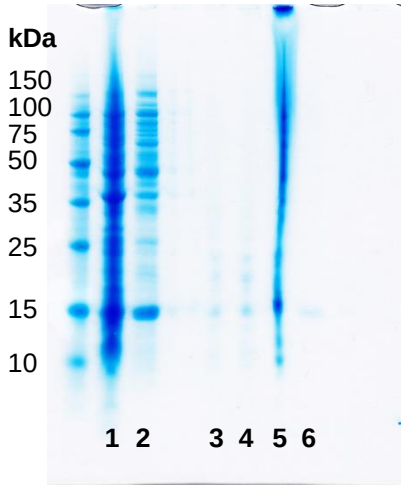


Figure 2.5 *SDS-PAGE. Lane 1: Cell debris; Lane 2: Cell lysate; Lane 3: Supernatant after PEG precipitate A; Lane 4: Supernatant after PEG precipitate B; Lane 5: Qβ-F protein in TBS before sucrose gradient; Lane 6: Qβ-F protein in TBS purified by sucrose density gradient (vide infra). Protein marker: 10-225 kDa.*

The final purification of the intact virus-like particles was achieved by sucrose density gradient ultracentrifugation. Sucrose density gradient ultracentrifugation is a technique used to purify large protein complexes (e.g. antibodies[20]), viruses,[21] and other cellular macromolecules. Solutions of different concentrations of sucrose, and therefore densities, are layered into an ultracentrifuge tube with the densest solution at the bottom (Fig. 2.6). The

tubes are held in swing buckets which move outwards to a horizontal position when the ultracentrifuge is in operation. Particles in a spinning rotor experience an acceleration equal to $\omega^2 r$, where ω is the angular velocity (radians per second) and r is the distance of the particle from the axis of rotation. Opposed to the force provided by this acceleration is a buoyant force, F_b, equal to $-\omega^2 r m_o$, where m_o is the mass of fluid displaced by the particle equal to $v\rho$, where v is the volume that each gram of the solute occupies in solution and ρ is the density of the solvent. The density of the solution increases towards the outer end of the tube (increasing concentration of sucrose). Provided the particle density is greater than the solvent density, the particles move outwards. When the particles reach a region where their density is equal to the solvent density, the particles form a layer in the tube. Hence, the VLPs can be readily separated from smaller species, including Qβ monomers. Furthermore, aggregates of VLPs are separated so a homogeneous sample of VLPs of the same size is obtained.

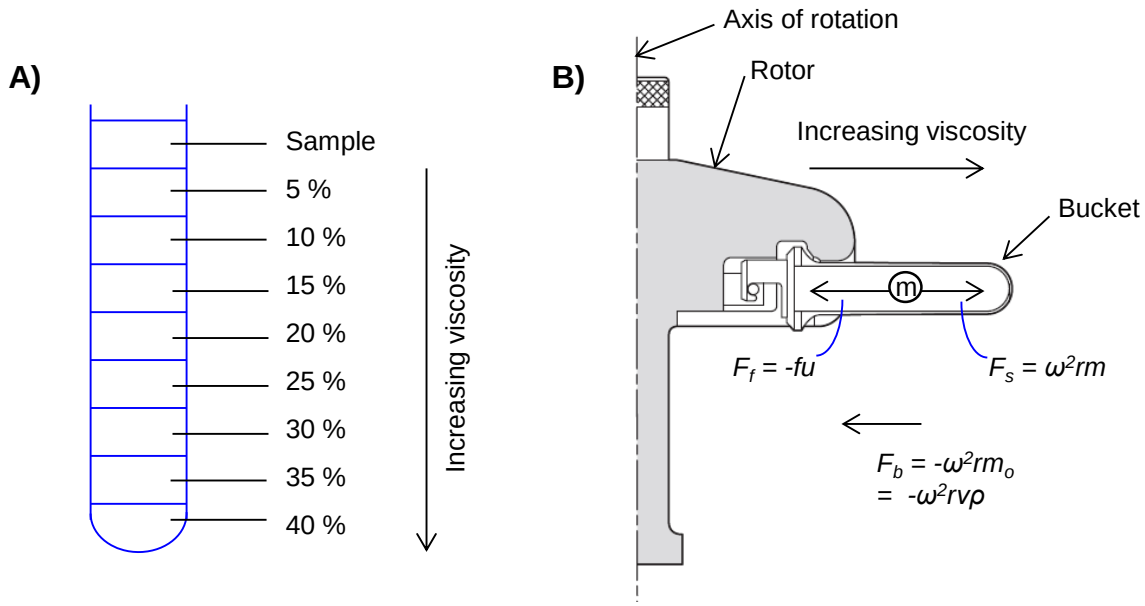


Figure 2.6 *Sucrose density gradient ultracentrifugation. A: An ultracentrifuge tube with different sucrose solutions labelled as sucrose/buffer w/w ratios. B: Profile of the Beckman ultracentrifuge rotor and bucket in its operating configuration (adapted from the Beckman Coulter SW40 Ti rotor manual).*

Gradients were fractionated and analysed by SDS-PAGE (Fig 2.7). Protein at ~14 kDa was found in the first two fractions, which corresponds to 5% w/w sucrose in TBS. It is likely that this protein corresponds to Qβ monomer which has not formed particles. Fractions 5 to 9, corresponding to 20% + w/w sucrose, clearly show pure Qβ, taken to be the intact VLPs due to the large sedimentation coefficient with respect to the monomers found in fractions 1 and 2. Fraction 13, at 40% w/w sucrose, contains protein impurities in addition to Qβ. The state of the Qβ in this fraction is likely aggregates due to its large sedimentation coefficient with respect to the VLPs. Further analysis of the VLP structure is presented in Chapter 3.

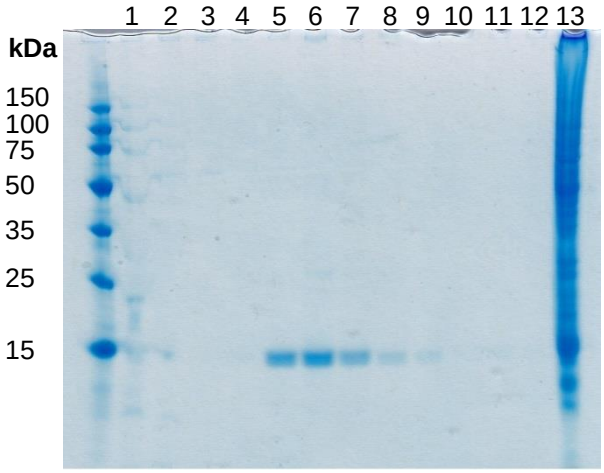


Figure 2.7 *Sucrose density gradient fractions for purification of Qβ-F from the top fraction (1) to the bottom fraction (13). Protein marker: 15-150 kDa.*

Fractions 5 to 9 were pooled and the VLPs were precipitated by addition of 2 M ammonium sulphate. The precipitate was pelleted by centrifugation and dissolved in TBS. The final purified protein was analysed by ESI-MS. Due to the large mass of the intact VLP, the VLP was reduced to the monomer by addition of TCEP (10 v/v % of 0.5 M TCEP to protein sample) before measuring its mass spectrum. The raw and deconvoluted mass spectrum of the purified protein is shown in Fig. 2.8 B. The raw mass spectrum shows the ion series for the protein which is deconvoluted to give the masses of the proteins detected by applying the MaxEnt algorithm to the mass spectrum (Masslynx software). The major peak in the ESI mass spectrum is mass 14179 Da, which corresponds to the calculated value of Qβ

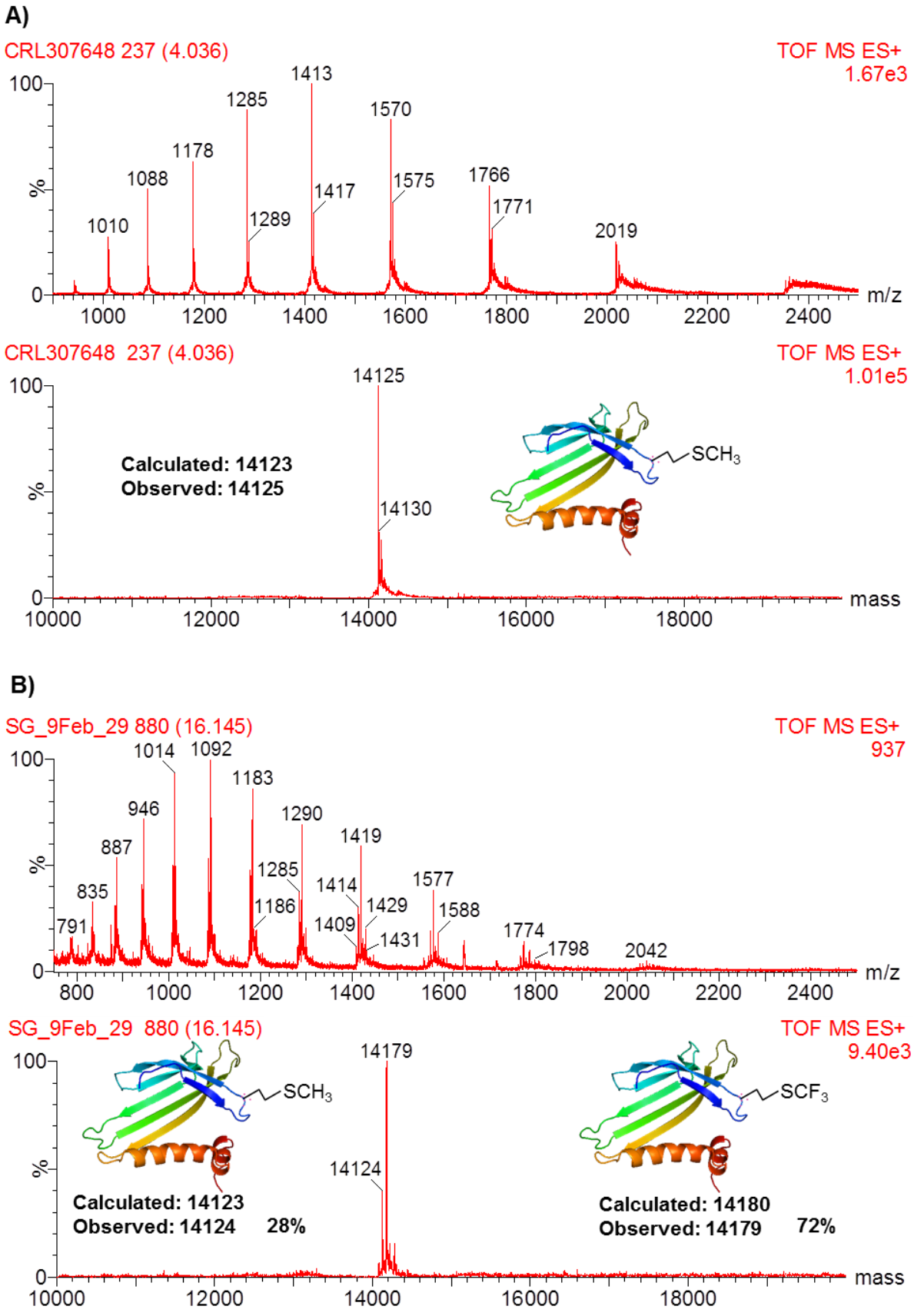


Figure 2.8 *Raw and deconvoluted ESI-MS of Qβ K16M expressed in the presence of methionine (A) and 1.3 mM DL-TFM (B).*

monomer with TFM substituted for methionine. This demonstrates successful incorporation of TFM into the Qβ protein monomer. It is noteworthy that this is the first example of genetic incorporation of fluorine into a virus. However, a minor peak at 14124 Da is also observed, which corresponds to the calculated mass of Qβ with methionine incorporated, as apposed to TFM. Verification that the peak at 14124 Da corresponds to Qβ-Met was achieved by expressing the mutant Qβ K16M in LB (i.e. in the presence of methionine). The VLP was purified in the same way as described above for Qβ-F and the ESI-MS showed a peak at 14125 Da (Fig. 2.8 A). The mass spectrum of Qβ K16M expressed in the presence of TFM clearly shows low levels of incorporation of the supplemented unnatural amino acid. The percentage incorporation (i.e. percentage of Qβ-F proteins in the sample) was calculated using the mass spectrum peak areas of the Qβ-Met and Qβ-F peaks and found to be 72% Qβ-F. Details of the verification of this method are presented in the experimental section (Section 2.6.5). In the light of this low level of incorporation, attempts were made to optimise the expression in order to maximise the incorporation of TFM in the Qβ protein.

2.3.2 Optimisation of incorporation of trifluoromethionine into bacteriophage Qβ

Incorporation of the unnatural amino acid results from misesterification of the $tRNA^{Met}$ by the enzyme methionyl-tRNA synthetase, during the translation process. Trifluoromethionine (TFM) competes with residual endogenous methionine for binding to methionyl-tRNA synthetase, and so the relative efficiency of these two competing processes has a direct effect on the percentage of incorporation of the unnatural amino acid. Therefore the effect of changing the concentrations of methionine and TFM on the level of incorporation of TFM in the Qβ protein were investigated.

The conditions for optimal growth of the *E. coli* cells and maximal incorporation of the unnatural amino acid requires a compromise between a sufficiently high concentration of methionine in the initial growth to produce a sufficient quantity of bacteria for appreciable

protein production, without supplementing an excess which will still be residual after washing steps (and hence lead to reduced unnatural amino acid incorporation levels). Furthermore, there is a compromise between having a high enough concentration of TFM to enable complete incorporation into the protein without the concentration being so high that the bacteria cells die before expression of the protein occurs (or at least before a reasonable quantity of protein has been produced). The toxicity of unnatural amino acid analogues has been well documented.[4,22–24]

The first report of incorporation of TFM into a protein was reported by Duewel *et. al.*[1] Their work has been discussed in Chapter 1. Their procedure was used to guide expression trials for optimisation in the current work. Duewel *et. al.* found that B834(DE3) cells did not grow in the presence of L-TFM.[1] To verify this result in the current work, B834(DE3) cells carrying the p75M/Qβ K16M plasmid were incubated in the presence of 0.01, 0.10, 0.20, 0.49, and 1.00 mM of DL-TFM. Potential cell growth was monitored by absorption of radiation at a wavelength of 600 nm. No increase in absorbance was observed at 3, 6, 30, or 71 hours post-inoculation. Since concentrations of L-Met as low as 0.05 mM are shown to support bacterial cell growth,[1] it was concluded that DL-TFM could not support growth in the absence of L-Met. Hence, cultures were thereafter grown initially in the presence of L–Met, following previous procedures.[18]

Next, the concentration of L–Met to be used was optimised. As mentioned above (page 45), previously published procedures from the Davis group use 0.27 mM L-Met for initial *E. coli* cell growth when incorporating methionine analogues,[18] hence this was used in the first expression. However, Duewel *et. al.* use 0.1 mM L-Met for growth of *E. coli* cells when incorporating trifluoromethionine.[1] Indeed, it has been reported that a concentration of L-Met as low as 0.05 mM can support cell growth (OD_{600} after 12 hrs: 1.070 for 0.05 mM L-Met, in the absence of TFM).[1] To investigate the optimum concentration of L-Met to use for cell

growth in this study, the OD_{600} of cultures grown in different concentrations of L-Met was monitored and recorded (Fig. 2.9). The OD_{600} in the exponential phase of the bacterial growth was much lower when using 0.05 mM L-Met (~0.2) compared to using 0.1 mM L-Met (~0.5) and 0.27 mM (~0.7). 0.1 mM gives sufficient cell growth, whereas it was thought that 0.27 mM could result in an excess of residual Met remaining after cell washing, which would subsequently lead to reduced incorporation of TFM. 0.05 mM clearly leads to very low cell densities. Considering these results, 0.1 mM L-Met was chosen and used for all subsequent amino acid incorporation expressions.

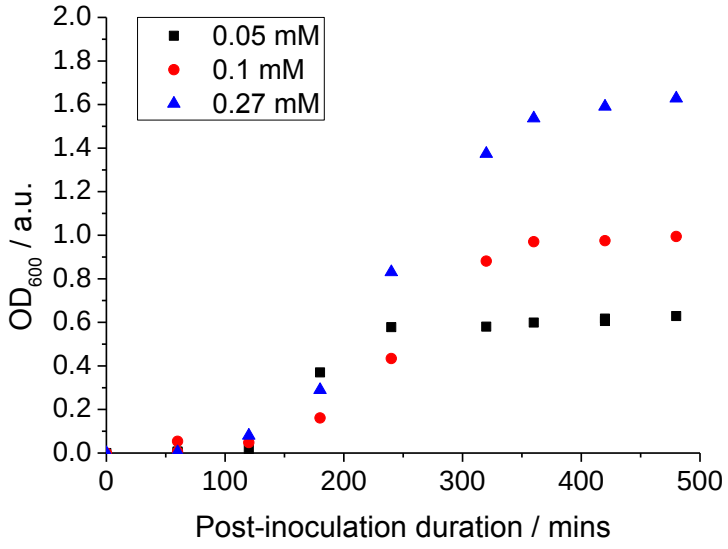


Figure 2.9 *Variation of OD_{600} of E. coli cultures grown in SelenoMet™ medium supplemented with different concentrations of L-Met.*

Having decided on the concentration of L-Met to use, the concentration of DL-TFM supplemented into the expression media was then investigated. Expression trials were conducted using smaller culture volumes than used previously (~160 mL compared to ~630 mL), so a sufficient quantity of protein to analyse by ESI-MS could be produced, but less unnatural amino acid would be required per expression. Once the conditions had been optimised, it was intended that larger scale expressions could then be used to produce a sufficient quantity of protein for further experiments.

Concentrations of 2 mM and 4 mM DL-TFM were trialled. The same concentration of L-Met (0.1 mM) in the initial growth stage was used in both cases so that the effect of doubling the concentration of DL-TFM could be determined. After purification of the Qβ-F protein using the same protocol as described above, the sample was analysed by ESI-MS (Fig. 2.10). The incorporation level of the Qβ K16M protein expressed in the presence of 2 mM and 4 mM DL-TFM was calculated to be 80% and 82%, respectively. Hence, as expected, increasing the concentration of DL-TFM has a positive effect on the incorporation. The protein yield for both expressions was ~0.3 mg/L, as determined by the Bradford assay.[25,26]

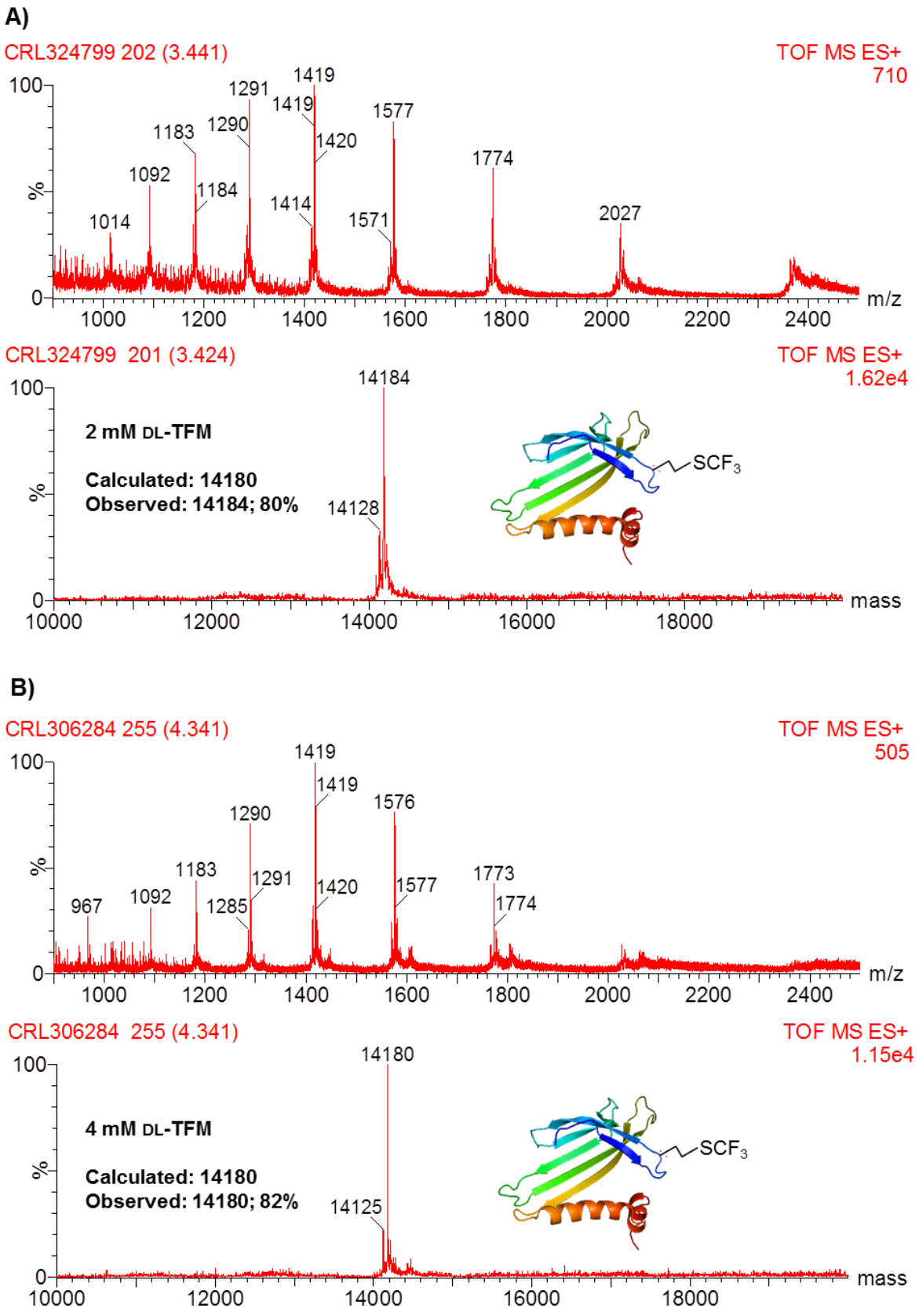


Figure 2.10 *Raw and deconvoluted ESI-MS of Qβ K16M expressed in the presence of 2 mM (A) and 4 mM (B) DL-TFM.*

Considering the positive affect on the TFM incorporation level with increasing concentration of DL-TFM, increasing the concentration of DL-TFM further was next investigated. The potential toxicity of DL-TFM to *E. coli* cells has been previously reported,[1] hence it was expected to observe a maximum concentration of DL-TFM that could be used in the medium before the cells die prior to producing protein. This maximum concentration was next determined empirically.

Cultures were supplemented with 6.5 mM, 8.5 mM, 10.5 mM, and 15.1 mM DL-TFM. Furthermore, the starvation time at 30 ^{o}C was increased from 30 minutes to 1 hour in an attempt to further deplete the endogenous pool of methionine, and hence increase the level of TFM incorporation.

The Qβ protein samples were purified as described previously (pages 46-48) and SDS-PAGE analysis of the sucrose density gradient fractions are shown in Fig. 2.11. The quantity of Qβ protein produced in the presence of 6.5 mM and 8.5 mM DL-TFM appears to be comparable. The quantity of Qβ VLPs (fractions 5-8) produced in the presence of 10.5 mM DL-TFM is negligible, if any; however small amounts of Qβ can be seen in the first two fractions, demonstrating that these are not fully formed particles (these are most likely monomers). The gel for the culture supplemented with 15.1 mM shows no Qβ was expressed, as neither bands in fractions corresponding to monomers nor to VLPs were observed. Interestingly, a protein of mass ~35 kDa was observed in fractions 11-14. The same protein band is seen in the bottom (aggregate) fraction of each of the other samples. This could possibly be from a particle formed from the A1 protein,[27] a 39 kDa protein resulting from ribosomal read-through of the coat protein UGA termination codon.[28] This would result in a particle with a higher density and hence explain the greater sedimentation coefficient compared to that of the coat protein VLP.

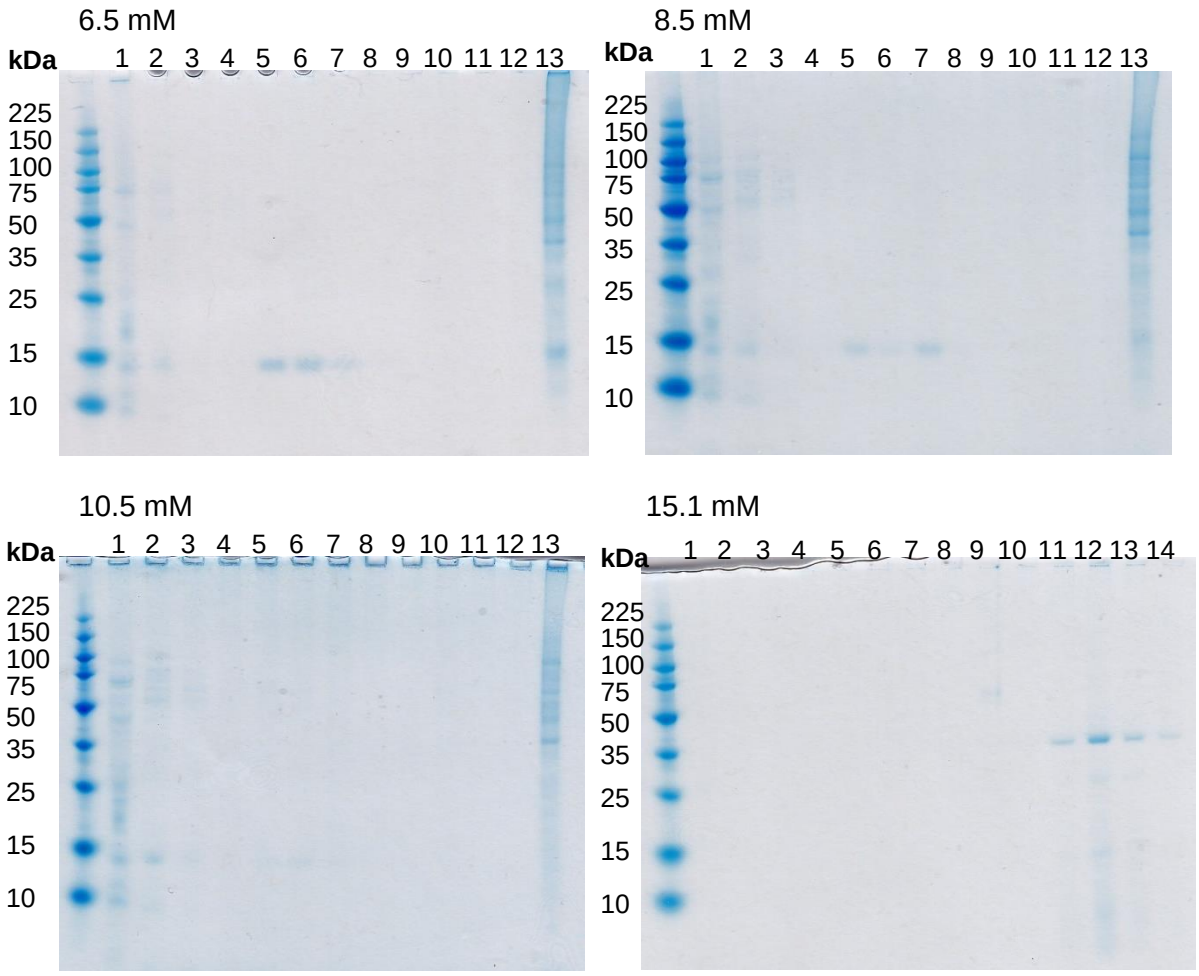


Figure 2.11 *SDS-PAGE of sucrose density gradient fractions of Qβ K16M expressed in SelenoMet™ medium supplemented with a concentration of* DL*-TFM as labelled. A 10-225 kDa ladder was used for each gel.*

Protein yields were calculated as 0.2 mg/L and 0.7 mg/L for the 6.5 mM-supplemented and 8.5 mM-supplemented cultures, respectively. Protein concentrations were calculated using the Bradford assay.[25,26] No protein was detected by Bradford assay for the 10.5 mM supplemented culture, presumably due to the low quantity of Qβ produced.

The ESI-MS of the Qβ protein expressed in 6.5 mM and 8.5 mM DL-TFM are shown in Figure 2.12 and the incorporation level was calculated to be 84% and 64%, respectively. It is not clear why the Qβ K16M protein expressed in 6.5 mM DL-TFM gives a higher incorporation of the unnatural amino acid compared to when expressed in 8.5 mM DL-TFM. However, based on these results 6.5 mM DL-TFM was selected as the concentration for optimum incorporation and was used in all subsequent expressions.

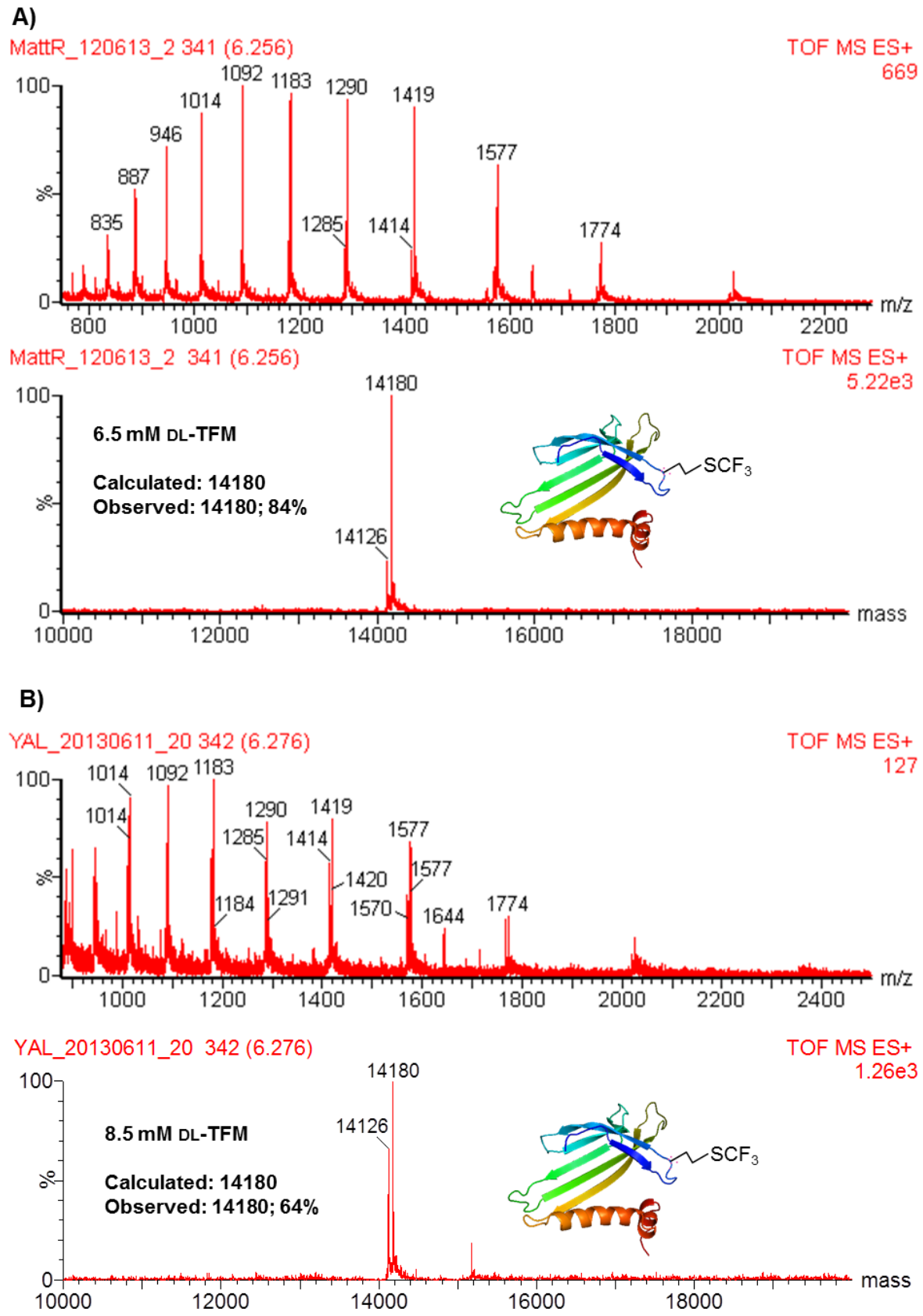


Figure 2.12 *Raw and deconvoluted ESI-MS of Qβ K16M expressed in the presence of 6.5 mM (A) and 8.5 mM (B)* DL*-TFM.*

In an attempt to further increase the level of incorporation of TFM into Qβ, varying expression periods were then explored. I proposed that once the intracellular supply of methionine (endogenous and that supplemented and remaining after cell washes) had been consumed, (almost) all Qβ protein biosynthesized after this point would incorporate TFM, since no other methionine source would be available. I therefore thought that extending the expression duration would increase the time for which the *E. coli* would be expressing pure TFM-incorporated protein and hence increase the proportion of TFM-incorporated protein.

In addition to 6 hours induction used previously, induction times of 10, 12, 14 and 16 hours were next investigated, each in the presence of 6.1 mM DL-TFM. SDS-PAGE gels of the sucrose density gradient fractions are shown in Fig. 2.13 and indicate that protein was expressed for induction times 10 and 12 hours, but no protein was expressed for induction times 14 and 16 hours. This could be due to protein being degraded by the *E. coli* when extended induction times are used. The protein yields for 10 and 12 hours were 0.32 mg/L and 0.51 mg/L, respectively, as determined by the Bradford assay.[25,26]

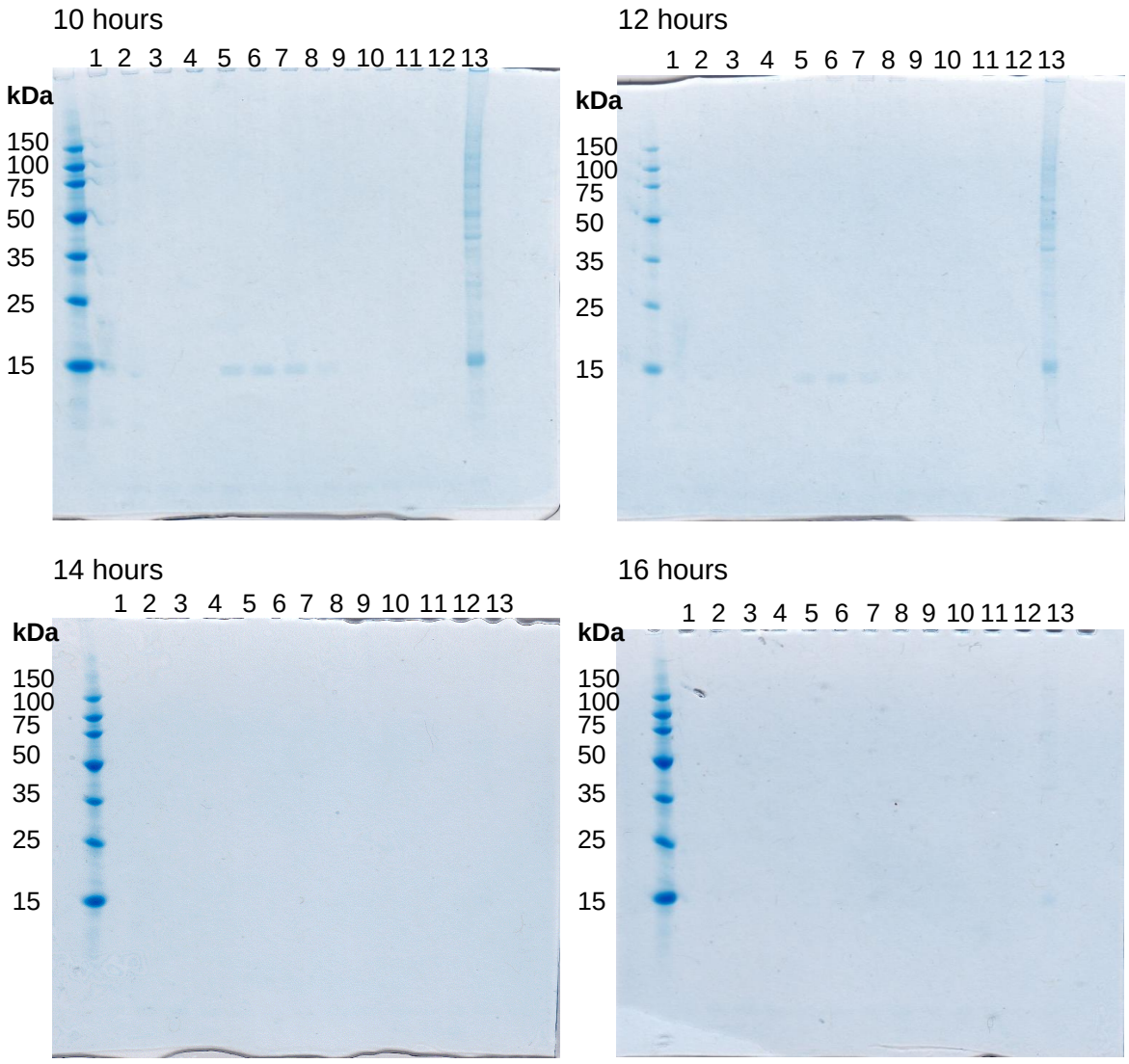


Figure 2.13 *SDS-PAGE of sucrose density gradient fractions of Qβ K16M expressed in the presence of 6.1 mM* DL*-TFM for durations as labelled. A 15-150 kDa protein marker was used for each gel.*

The ESI-MS of the Qβ proteins expressed for 10 and 12 hours are shown in Fig. 2.14. The results show an increase in the level of incorporation of TFM into Qβ on extending the duration of expression to 12 hours. More significantly, the ESI-MS for the Qβ protein expressed for 12 hours shows almost quantitative incorporation of TFM into the Qβ protein (the calculated incorporation level is 96%). This is an unprecedented result since the highest previous level of incorporation of TFM into any protein reported to date is 82%.[16]

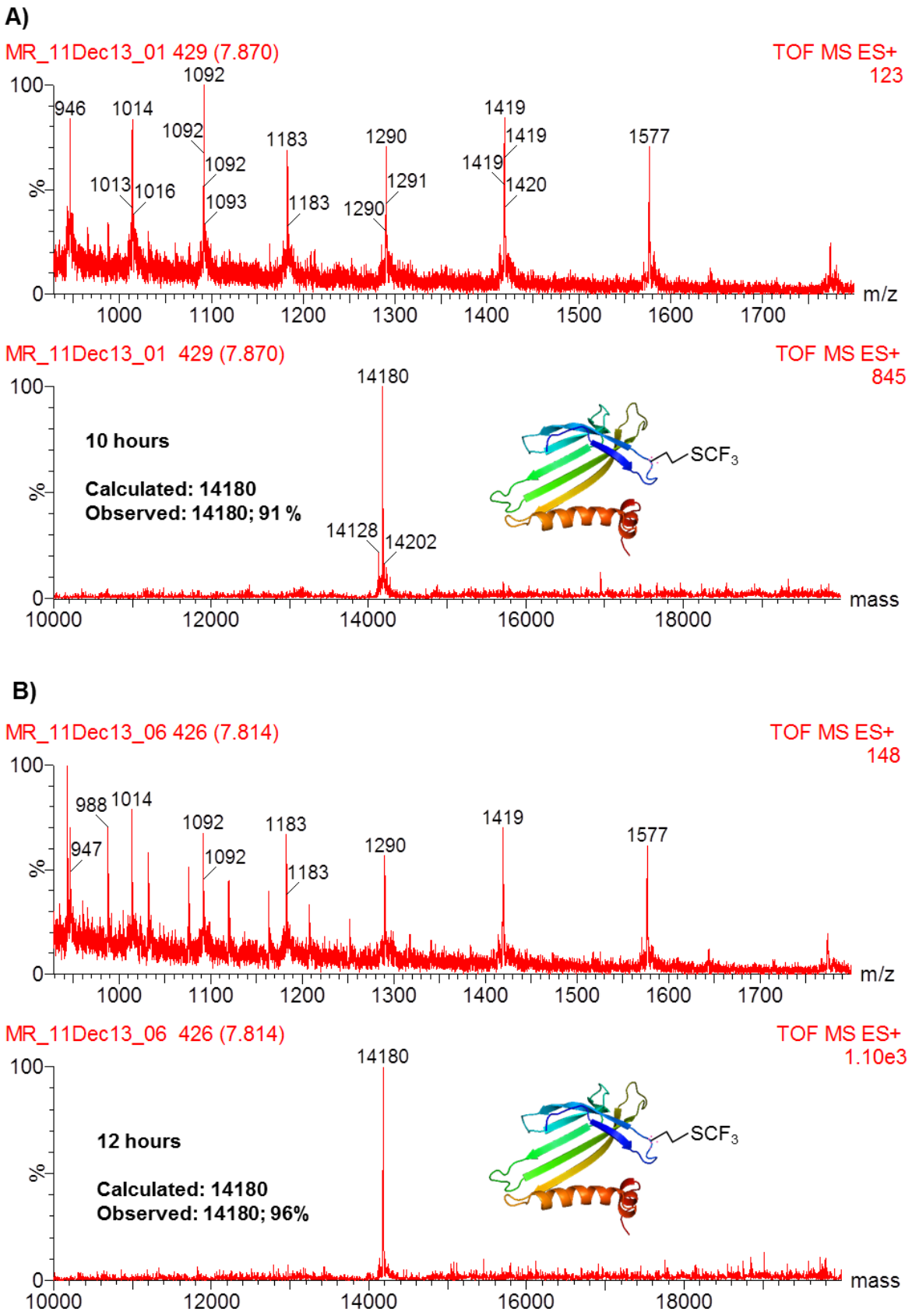


Figure 2.14 *Raw and deconvoluted ESI-MS of Qβ K16M expressed in the presence of 6.1 mM* DL-*trifluoromethionine with expression durations of 10 hours (A) and 12 hours (B).*

2.4 Conclusions

The unnatural amino acid trifluoromethionine has been successfully incorporated into the bacteriophage Qβ for the first time as demonstrated by ESI-MS analysis. This procedure involves initial production of large quantities of the pure DL-trifluoromethionine for expression trials. After investigating several different routes, a one-step method reported by Yasui *et. al.* was used.[17] This enabled production of 7-9 grams of material per reaction.

Previous attempts to incorporate trifluoromethionine with high levels of incorporation into a variety of proteins have proved challenging and until now the maximum level achieved was 82%[16] (see Table 2.1). This difficulty has been noted and rationalised in terms of the increased steric hindrance of the three fluorine atoms in trifluoromethionine as compared to the three hydrogen atoms in methionine, which leads to less efficient binding of trifluoromethionine to methionyl tRNA transferase than methionine (see discussion in Chapter 1 Section 1.4.5).[29]

Group	Year	Reference	Level of incorporation %
Honek	1997	[1]	70
Honek	1998	[30]	74
Honek	2001	[2]	70
Marx	2010	[16]	82

Table 2.1 *Previous reports of incorporation of trifluoromethionine into proteins using auxotrophic E. coli and the resulting level of incorporation.*

The incorporation of trifluoromethionine into bacteriophage Qβ was optimised by varying the initial concentration of L-methionine, the concentration of DL-trifluoromethionine, and the duration of protein expression. Pleasingly, the work led to optimum conditions which yielded protein with almost quantitative incorporation of trifluoromethionine into Qβ (96%). This

marks the highest incorporation level achieved for this amino acid into any protein, using the residue-specific method, to date.

Unnatural amino acid incorporation is an extensively used technique in chemical biology today and it is hoped that the methodology described in this chapter will be used for optimising the incorporation of other unnatural amino acids into proteins.

2.5 References

1. 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 19F NMR. *Biochemistry* **36,** 3404–3416 (1997).

2. Duewel, H. S., Daub, E., Robinson, V. & Honek, J. F. Elucidation of Solvent Exposure, Side-Chain Reactivity, and Steric Demands of the Trifluoromethionine Residue in a Recombinant Protein. *Biochemistry* **40,** 13167–13176 (2001).

3. Vaughan, M. D., Cleve, P., Robinson, V., Duewel, H. S. & Honek, J. F. Difluoromethionine as a Novel 19F NMR Structural Probe for Internal Amino Acid Packing in Proteins. *J. Am. Chem. Soc.* **121,** 8475–8478 (1999).

4. Budisa, N. *et al.* Efforts towards the Design of "Teflon" Proteins: In vivo Translation with Trifluorinated Leucine and Methionine Analogues. *Chem. Biodivers.* **1,** 1465–1475 (2004).

5. Srinivas, M. *et al.* Customizable, multi-functional fluorocarbon nanoparticles for quantitative in vivo imaging using 19F MRI and optical imaging. *Biomaterials* **31,** 7070–7077 (2010).

6. Konigsberg, W., Maita, T. & Katze, J. Amino-acid Sequence of the Qβ Coat Protein. *Nature* **227,** 271–273 (1970).

7. Serber, Z. *et al.* Methyl groups as probes for proteins and complexes in in-cell NMR experiments. *J. Am. Chem. Soc.* **126,** 7119–25 (2004).

8. Tugarinov, V. & Kay, L. E. Methyl Groups as Probes of Structure and Dynamics in NMR Studies of High-Molecular-Weight Proteins. *ChemBioChem* **6,** 1567–1577 (2005).

9. Ruppert, I., Schlich, K. & Volbach, W. The 1st CF3-substituted organyl(chloro)silanes. *Tetrahedron Lett.* **25,** 2195–2198 (1984).

10. Umemoto, T. & Ishihara, S. Power-Variable Trifluoromethylating Agents, (Trifluoromethyl)dibenzothio- and -Selenophenium Salt System. *Tetrahedron Lett.* **31,** 3579–3582 (1990).

11. Umemoto, T. & Ishihara, S. Power-Variable Electrophilic Trifluoromethylating Agents. S-, Se-, and Te-(trifluoromethyl)dibenzothio-, -seleno-, and -tellurophenium Salt System. *J. Am. Chem. Soc.* **115,** 2156–2164 (1993).

12. Umemoto, T. & Adachi, K. New Method for Trifluoromethylation of Enolate Anions and Applications to Regio-, Diastereo- and Enantioselective Trifluoromethylation. *J. Org. Chem.* **59,** 5692–5699 (1994).

13. Eisenberger, P., Gischig, S. & Togni, A. Novel 10-I-3 Hypervalent Iodine-Based Compounds for Electrophilic Trifluoromethylation. *Chem. Eur. J.* **12,** 2579–2586 (2006).

14. Kieltsch, I., Eisenberger, P. & Togni, A. Mild Electrophilic Trifluoromethylation of Carbon- and Sulfur-Centered Nucleophiles by a Hypervalent Iodine(III)-CF3 Reagent. *Angew. Chem. Int. Ed.* **46,** 754–757 (2007).

15. Eisenberger, P., Kieltsch, I., Koller, R., Stanek, K. & Togni, A. Preparation of a Trifluoromethyl Transfer Agent: 1-Trifluoromethyl-1,3-Dihydro-3,3-Dimethyl-1,2-Benziodoxole. *Org. Synth.* **88,** 168–180 (2011).

16. Holzberger, B., Rubini, M., Möller, H. M. & Marx, A. A Highly Active DNAPolymerase with a Fluorous Core. *Angew. Chem. Int. Ed.* **49,** 1324–1327 (2010).

17. 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. *J. Fluor. Chem.* **132,** 186–189 (2011).

18. Floyd, N., Vijayakrishnan, B., Koeppe, A. R. & Davis, B. G. Thiyl Glycosylation of Olefinic Proteins: S-Linked Glycoconjugate Synthesis. *Angew. Chem. Int. Ed.* **48,** 7798–7802 (2009).

19. Strable, E. *et al.* Unnatural Amino Acid Incorporation into Virus-Like Particles. *Bioconjug. Chem.* **19,** 866–875 (2008).

20. Challacombe, S. J., Russell, M. W. & Hawkes, J. Passage of intact IgG from plasma to the oral cavity via crevicular fluid. *Clin. Exp. Immunol.* **34,** 417–422 (1978).

21. Wang, Q., Kaltgrad, E., Lin, T., Johnson, J. E. & Finn, M. G. Natural Supramolecular Building Blocks: Wild-Type Cowpea Mosaic Virus. *Chem. Biol.* **9,** 805–811 (2002).

22. Houston, M. E., Harvath, L. & Honek, J. F. Syntheses of and Chemotactic Responses Elicited by fMet-Leu-Phe Analogs Containing Difluoro- and Trifluoromethionine. *Bioorg. Med. Chem. Lett.* **7,** 3007–3012 (1997).

23. Richmond, M. H. The Effect of Amino Acid Analogues on Growth and Protein Synthesis in Microorganisms. *Bacteriol. Rev.* **26,** 398–420 (1962).

24. Barker, D. G. & Bruton, C. J. The Fate of Norleucine as a Replacement for Methionine in Protein Synthesis. *J. Mol. Biol.* **133,** 217–231 (1979).

25. Bradford, M. M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **72,** 248–254 (1976).

26. Thermoscientific. Chemistry of Protein Assays. at <http://www.piercenet.com/method/chemistry-protein-assays#dyeassays>

27. Takamatsu, H. & Iso, K. Chemical evidence for the capsomeric structure of phage Qβ. *Nature* **298,** 819–824 (1982).

28. Horiuchi, K., Webster, R. E. & Matsuhashi, S. Gene Products of Bacteriaphage Qβ. *Virology* **45,** 429–439 (1971).

29. Crepin, T. *et al.* Use of Analogues of Methionine and Methionyl Adenylate to Sample Conformational Changes During Catalysis in Escherichia coli Methionyl-tRNA Synthetase. *J. Mol. Biol.* **332,** 59–72 (2003).

30. Duewel, H. S. & Honek, J. F. CNBr/Formic Acid Reactions of Methionine- and Trifluoromethionine-Containing Lambda Lysozyme: Probing Chemical and Positional Reactivity and Formylation Side Reactions by Mass Spectrometry. *J. Protein Chem.* **17,** 337–350 (1998).

31. Gottlieb, H. E., Kotlyar, V. & Nudelman, A. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *J. Org. Chem.* **62,** 7512–7515 (1997).

32. Carme Pampín, M., Estévez, J. C., Estévez, R. J., Maestro, M. & Castedo, L. Heck-mediated synthesis and photochemically induced cyclization of [2-(2-styrylphenyl)ethyl]carbamic acid ethyl esters and 2-styryl-benzoic acid methyl esters: total synthesis of naphtho[2,1f]isoquinolines (2-azachrysenes). *Tetrahedron* **59,** 7231–7243 (2003).

33. Brown, H. C., Okamoto, Y. & Ham, G. Rates of Solvolysis of the Halophenyldimethylcarbinyl Chlorides. The Effect of Halogen Substituents upon the Rates of Electrophilic Reactions. *J. Am. Chem. Soc.* **79,** 1906–1909 (1957).

34. Du, H., Zhao, B. & Shi, Y. Catalytic Asymmetric Allylic and Homoallylic Diamination of Terminal Olefins via Formal C-H Activation. *J. Am. Chem. Soc.* **130,** 8590–8591 (2008).

35. Amey, R. L. & Martin, J. C. Synthesis and Reaction of Substituted Arylalkoxyiodinanes: Formation of Stable Bromoarylalkoxy and Aryldialkoxy Heterocyclic Derivatives of Tricoordinate Organoiodine(III). *J. Org. Chem.* **44,** 1779–1784 (1979).

36. Benesch, R. & Benesch, R. E. Formation of Peptide Bonds by Aminolysis of Homocysteine Thiolactones. *J. Am. Chem. Soc.* **78,** 1597–1599 (1956).

37. Okumura, K. *et al.* Vitamin-B6 derivatives. 13. Synthesis of Tetrahydrothiazine Derivatives of Vitamin-B6 and Their Biological Properties. *J. Med. Chem.* **14,** 226–229 (1971).

38. Sambrook, J. & Russell, D. *Molecular Cloning: A Laboratory Manual*. (Cold Spring Harber Laboratory Press, 2001).

2.6 Experimental Procedures

2.6.1 General considerations for chemical synthesis and analysis

NMR: Proton nuclear magnetic resonance spectra (^{1}H NMR) were recorded on a Bruker DPX 200 (200 MHz), a Bruker DQX 400 (400 MHz), or by Dr. T. Claridge or Dr. B. Odell on a Bruker AVC 500 (500 MHz) spectrometer. Carbon nuclear magnetic resonance spectra (^{13}C NMR) were recorded on a Bruker DQX 400 (101 MHz), or by Dr. T. Claridge or Dr. B. Odell on a Bruker AVC 500 (126 MHz) and are proton decoupled. The chemical shifts of the spectra are calibrated by the spectrometer using the solvent peak. The spectra were processed using Advanced Chemistry Development/Labs (ACD/Labs) software. Chemical shifts are quoted on the δ scale in parts per million (ppm) and were referenced to residual solvent peaks using previously published solvent chemical shifts (^{1}H NMR: $CDCl_3$ = 7.26, D_2O = 4.79 and ^{13}C NMR: $CDCl_3$ = 77.16).[31] Integrations were also performed in the ACD/Labs software. The following splitting abbreviations have been used: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, dt = doublet of triplets, td = triplet of doublets.

Mass spectrometry: Low resolution mass spectrometry was carried out using a Micromass LCT Premier XE spectrometer. Samples were dissolved in MeOH prior to injection. Electrospray ionisation in positive (ESI+) and negative (ESI-) scan modes were used. m/z values were reported against their percentage abundance. Mass spectra of proteins were recorded using a Waters LCT Classic. This uses electrospray ionisation, a time-of-flight analyser and incorporates a HP1100 chromatography system.

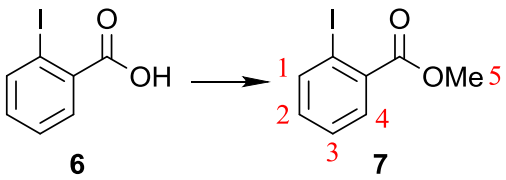
Melting points: Melting points were recorded using a Leica Galen III hot stage microscope equipped with a Testo 720 thermocouple and are uncorrected.

Chromatography: Thin layer chromatography (TLC) was carried out using Merk aluminium-backed sheets coated with Kieselgel 60F_{254} silica gel. Visualization was achieved using a UV lamp (λ = 254 nm) and/or ammonium molybdate (5% in 2M H_2SO_4), or potassium permanganate (5% in 1M NaOH), or ninhydrin (1% in ethanol) followed by heat gun treatment. Flash column chromatography was carried out using Fluka Kiegselgel 60 220-240 mesh silica gel. Dry loading onto TELOS (*Kinesis* Scientific Experts) was used for solid crude material. TELOS was added to a solvent solution containing the dissolved crude material to a 1:1 w/w ratio with respect to the crude material. The solvent was then removed on a rotary evaporator to give the TELOS powder with the crude material adhered. DOWEX® 50WX8-200 ion exchange resins were activated by washing with acetone, water, 1M HCl, water, MeOH, eluting with water until the filtrate was neutral (as determined by pH paper) and finally drying *in vacuo.*

Solvents and reagents: Unless otherwise stated, solvents and reagents were purchased from standard commercial suppliers and were used without further purification. Dry solvents were purchased from Acros Organics or Fluka and stored under argon over molecular sieves. "Petrol" refers to the fraction of light petroleum ether boiling in the range 40-60 ^{o}C. "Brine" refers to a saturated aqueous solution of sodium chloride.

2.6.2 Chemical synthesis

Methyl 2-iodobenzoate[32]

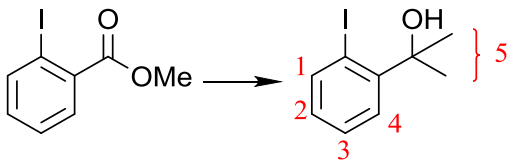


6 **7**

Standard esterification conditions were used to achieve the transformation from **6** to **7**. 2-Iodobenzoic acid **6** (27.0 g, 0.109 mol, 1.0 eq.) was added to a flame dried 200 mL round-

bottomed flask with a magnetic stirrer. 70 mL of anhydrous MeOH was added to the flask and the mixture was stirred until the 2-iodobenzoic acid **6** had dissolved. The reaction mixture was heated to 60 ^{o}C under reflux and under nitrogen. 5.5 mL of concentrated sulfuric acid was added to the reaction mixture dropwise. TLC (30% EtOAc in petrol) revealed the presence of starting material (R_f = 0.45) in the reaction mixture and so the reaction was stirred and heated overnight. After this a further 5.5 mL of concentrated sulfuric acid was added and left for 3 hours. TLC (30% EtOAc in petrol) revealed the product R_f = 0.75 against the starting material R_f = 0.45. The solvent was removed using a rotary evaporator and 20 mL of CH_2Cl_2 was added to dissolve the liquid product. 20 mL of H_2O was added and the organic layer was separated. The aqueous layer was extracted with CH_2Cl_2 (2 x 20 mL). A saturated aqueous solution of Na_2CO_3 was added to remove the excess H_2SO_4. The resulting organic fraction was dried over $MgSO_4$ and filtered. The remained CH_2Cl_2 was removed by rotary evaporation to give methyl 2-iodobenzoate **7** as a clear oil (26.8 g, 0.102 mol, 94%). MS m/z (ESI$^+$): $[M+Na]^+$ 284.94. ^{1}H NMR ($CDCl_3$, 400 MHz): δ = 3.93 (3H, s, H5 OC$\underline{H}_3$), 7.15 (1H, td, *J* = 7.5, 2.0 Hz, H2/H3), 7.40 (1H, td, *J* = 7.5, 1.0 Hz, H2/H3), 7.80 (1H, dd, *J* = 8.0, 1.5 Hz, H1/H4), 7.99 (1H, dd, *J* = 8.0, 1.0 Hz, H1/H4).

2-(2-iodophenyl)propan-2-ol[33]

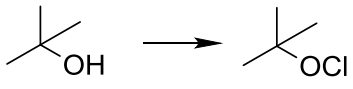


7 **8**

The procedure of Brown *et. al.* was used.[33] A 500 mL two-necked round-bottomed flask with an addition funnel was evacuated, flame dried and then flushed with nitrogen. Compound **7** (26.8 g, 0.102 mol, 1.0 eq.) was dissolved in 61.4 mL of Et_2O and added to the round-bottomed flask *via* a syringe. The flask was cooled to -20 ^{o}C using a NaCl/ice bath. 1M

MeMgBr in Bu_2O (224 mL, 0.224 mol, 2.2 eq.) was added to the addition funnel *via* a hypodermic needle. The MeMgBr solution was added dropwise over a period of 50 minutes. The reaction was left stirring for 16 hours. The reaction mixture was then refluxed for 1.5 hours, then cooled to 0 ^oC (ice-water bath) and then saturated NH_4Cl in H_2O (22.1 mL, 0.123 mol, 1.2 eq.) was added to quench the reaction. The organic phase was separated from the aqueous phase, and the aqueous phase was extracted with Et_2O. The combined organic phases were dried over $MgSO_4$, filtered, the solvent evaporated using a rotary evaporator and the residue was then dried under vacuum. Fractional distillation under vacuum was used to remove the remaining Et_2O and Bu_2O solvents. The crude product, a light yellow-coloured oil, was subjected to analysis by TLC and MS. This revealed 2-(2-iodophenyl)propan-2-ol **8** (R_f = 0.20, MS m/z (ESI^+): 285.00 [$M+Na^+$]), 1-(2-iodophenyl)ethanone (R_f = 0.35 MS m/z (ESI^+): 268.97 [$M+Na^+$]), and **7** (R_f = 0.50, MS m/z (ESI^+): 284.96 [$M+Na^+$]). The crude product was purified by column chromatography (5% EtOAc in Petrol) and 2-(2-iodophenyl)propan-2-ol **8** was isolated as a brown oil (8.76 g, 0.033 mol, 33%). MS m/z (ESI^+): [$M+Na^+$] 285.00. ^1H NMR ($CDCl_3$, 400 MHz): δ = 1.76 (6 H, s, H5 $C(C\underline{H}_3)_2$), 7.12 (1 H, td, *J* = 7.0, 2.0 Hz, H2/H3), 7.43 (1 H, m, *J* = 7.0, 2.0 Hz, H2/H3), 7.70 (1H, dd, *J* = 8.0, 1.0 Hz, H1/H4), 8.00 (1H, dd, *J* = 8.0, 1.0 Hz, H1/H4).

t-butyl hypochlorite[34]

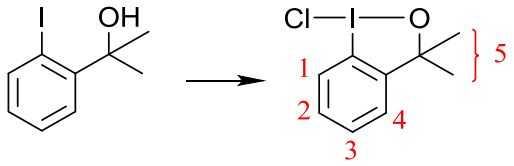


17 **18**

The procedure of Du *et. al.* was used.[34] An 8% commercial bleach solution (Fisher Scientific) was diluted with distilled water to make a 5% bleach solution. This solution (118 mL, 0.079 mol, 0.77 eq.) was placed in a 250 mL round-bottomed flask equipped with a magnetic stirrer. The flask was cooled to 0 ^oC (ice-water bath) and the solution was stirred rapidly.

After turning off the lights in the vicinity of the apparatus, a mixture of *t*-butyl alcohol **17** (9.73 mL, 0.102 mol, 1.0 eq.) and acetic acid (5.83 mL, 0.102 mol, 1.0 eq.) was added in a single portion to the rapidly stirring bleach solution, and stirring was continued for 5 minutes. The entire reaction mixture was poured into a 250 mL separation funnel. The yellow organic layer was washed with aqueous sodium carbonate (30 mL) and then with water (30 mL), dried over calcium chloride, and filtered to give *t*-butyl hypochloride **18** as a yellow liquid (5.24 g, 0.048 mol, 47%).

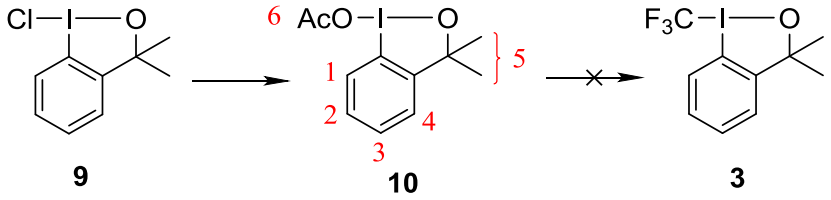
1-Chloro-1,3-dihydro-3,3-dimethyl-1,2-benziodoxole[35]



8 **9**

The procedure of Amey *et. al.* was used.[35] Compound **8** (8.76 g, 0.033 mol, 1.0 eq.) was added to a 50 mL round-bottomed flask, equipped with a magnetic stirrer, and was dissolved in 26.3 mL of CCl_4. The mixture was stirred until all of compound **8** had dissolved. Compound **18** (4.35 g, 0.040 mol, 1.2 eq.) was added to the mixture and the reaction was stirred. After 30 minutes a pale yellow powder was collected by filtration and recrystallised from CCl_4 followed by drying *in vacuo* to give 1-chloro-1,3-dihydro-3,3-dimethyl-1,2-benziodoxole **9** (6.88 g, 0.023 mol, 70 %). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 1.55 (6H, s, H5 C(C$\underline{H}_3$)$_2$), 7.16 (1H, dd, *J* = 7.0, 2.0 Hz, H4), 7.54 (2H, ddd, ^3J = 7.5, 7.0 Hz, ^4J = 1.5 Hz, H2 and H3), 8.02 (1H, dd, *J* = 8.0, 1.0 Hz, H1). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 29.29 (2C, s, C($\underline{C}H_3$)$_2$), 85.23 (1C, s, O$\underline{C}$(CH$_3$)$_2$), 114.72 (1C, s, aromatic), 126.19 (1C, s, aromatic), 128.52 (1C, s, aromatic), 130.54 (1C, s, aromatic), 131.03 (1C, s, aromatic), 149.59 (1C, s, aromatic).

Attempted synthesis of 1-trifluoromethyl-1,3-dihydro-3,3-dimethyl-1,2-benziodoxole[14]



The procedure outlined in the paper by Togni *et. al.*[14] was followed to try and achieve the above transformation. A 25 mL round-bottomed flask containing a magnetic stirrer was flame dried under vacuum. The flask was charged with compound **9** (0.100 g, 0.338 mmol, 1.0 eq.), 5 mL of anhydrous MeCN was added and the reaction mixture was stirred until all of the starting material had dissolved. KOAc (0.056 g, 0.568 mmol, 1.7 eq.) was added to the reaction mixture and stirring was continued for 1 hour at room temperature. Filtration and washing of the remaining white solid with MeCN gave a clear, almost colourless solution to which further MeCN (5 mL) was added. The solution was cooled to -17 ^{o}C using a NaCl-ice bath. Me_3SiCF_3 (0.08 mL, 0.520 mmol, 1.5 eq.) was added to the mixture followed by dropwise addition of tetrabutylammonium triphenyldifluorosilicate (TBAT; 0.547 mg, 0.001 mmol, 0.003 eq.) in 1 mL of MeCN. The reaction was stirred for 16 hours at -17 ^{o}C, then warmed to -12 ^{o}C, at which temperature further Me_3SiCF_3 was added (0.01 mL, 0.074 mmol, 0.2 eq.). The reaction mixture was warmed to room temperature over 3 hours and then stirred at room temperature for a further 3 hours. After the volatile components of the mixture had been removed using a rotary evaporator, the residue was dried *in vacuo*. Pentane (20 mL) was added to the remaining yellow-white solid. The solid was filtered off and collected, while the resulting solution was filtered through a pad of Al_2O_3 by vacuum filtration. The clear, colourless solution was transferred to a round-bottomed flask and the solvent was removed using a rotary evaporator. Analysis by ^{1}H NMR and ^{13}C NMR showed that the resulting white solid was 1-acetoxy-1,3-dihydro-3,3-dimethyl-1,2-benziodoxole **10**. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 1.52 (6H, s, H5 C(C$\underline{H}_3$)$_2$), 2.11 (3H, s, H6 COC$\underline{H}_3$), 7.18 (1H, dd, *J* = 7.0,

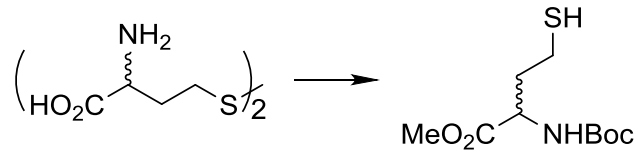
1.5 Hz, H4), 7.47 (2H, m, H2 and H3), 7.80 (1H, ddd, *J* = 8.0, 7.5, 1.0 Hz, H1). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 21.60 (1C, s, OCO<u>C</u>H_3), 29.33 (1C, s, C(<u>C</u>H_3)$_3$), 84.67 (1C, s, <u>C</u>$(CH_3)_3$), 115.83 (1C, s, aromatic), 126.31 (1C, s, aromatic), 130.02 (1C, s, aromatic), 130.06 (1C, s, aromatic), 130.52 (1C, s, aromatic), 149.52 (1C, s, aromatic), 177.45 (1C, s, O<u>C</u>OCH_3).

The solid that was collected before filtering through Al_2O_3 was anaylsed and also found to be compound **10**.

Several additional attempts were made to synthesize the target compound, using different conditions and methods. Firstly an excess amount of the reagents was used with respect to the starting material: 3.0 eq. of compound **9**, 3.0 eq. of KOAc, 3.0 eq. Me_3SiCF_3 (and then later in the reaction 0.66 eq.), and 0.009 eq. TBAT. Again, the resulting compound was **10**.

Secondly, a different paper[13] was followed in which the starting material was **10**, negating the first step in the scheme above. Compound **10** (0.544 g, 1.70 mmol, 1.0 eq.) was added to a flame dried 25 mL round-bottomed flask. Anhydrous MeCN (10 mL) was added and the mixture was stirred until all of compound **10** had dissolved. Me_3SiCF_3 (0.754 mL, 5.10 mmol, 3.0 eq. (twice the number of equivalents used in the paper)) was added using a syringe. CsF (3.00 mg, 0.017 mmol 0.01 eq.) in MeCN (1.3 mL) was added dropwise to the reaction mixture. The reaction was stirred for 22 hours. After removing the solvent using a rotary evaporator, pentane was added to the resulting oily solid. The precipitate was filtered off giving a clear, colourless solution. The solvent was removed using a rotary evaporator to give a white solid. However, ^{1}H NMR and ^{13}C NMR revealed this compound was the unreacted compound **10**.

***N*-Boc-DL-homocysteine methyl ester**



11 **5**

Anhydrous MeOH (90 mL) was added to a flame dried 250 mL round-bottomed flask equipped with a magnetic stirrer. The solvent was stirred and cooled to 0 ^{o}C and acetyl chloride (13.5 mL, 222 mmol, 30 eq.) was added dropwise over 5 minutes. The solution was stirred for an additional 10 minutes at 0 ^{o}C to give a concentrated solution of HCl. DL-Homocystine **11** (2.00 g, 7.45 mmol, 1.0 eq.) was then added in one portion and the flask was flushed with argon. The ice bath was removed and the reaction was stirred at room temperature for 24 hours. The solvent was then removed using a rotary evaporator to give the crude homocystine methyl ester hydrochloride as a thick yellow oil. This material was used immediately in the next step without purification. The crude ester was suspended in CH_2Cl_2 (90 mL) and cooled to 0 ^{o}C. Et_3N (5.19 mL, 37.3 mmol, 5.0 eq.) was added carefully followed by di-*tert*-butyl dicarbonate (Boc_2O, 3.58 g, 16.4 mmol, 2.2 eq.). The reaction was stirred at room temperature for 1.5 hours. The solvent was removed using a rotary evaporator and the resulting residue was redissolved in MeOH (45 mL) and H_2O (20 mL). Tributylphosphine (2.23 mL, 8.94 mmol, 1.2 eq.) was added dropwise to the stirred solution and then heated at 50 ^{o}C for 30 minutes. TLC (50% EtOAc in petrol, $KMnO_4$ stain) revealed reduction of the disulfide, with the disulfide eluting at R_f = 0.10 and *N*-Boc-DL-homocysteine methyl ester **5** at R_f = 0.50.

The reaction was diluted with Et_2O (113 mL) and H_2O (56 mL). The organic phase was separated and the aqueous phase was extracted with Et_2O (2 x 60 mL). The combined organic phases were washed with brine (120 mL), dried over $MgSO_4$, and filtered. The solvent was removed using a rotary evaporator and the residue was purified by column

chromatography eluting first with two column volumes (CV) of 5% EtOAc in petrol, then increasing the polarity of the solvent mixture in increments of 5% up to 20% EtOAc in petrol, with 2 CVs used for each solvent mixture. TLC analysis (30% EtOAc in petrol) of the fractions revealed *N*-Boc-DL-homocysteine methyl ester **5** (R_f = 0.30, MS m/z (ESI^+): 372.16 $[M+Na]^+$, 721.34 $[2M+Na]^+$) and *S*-Boc-*N*-Boc-DL-homocystine methyl ester (R_f = 0.40, MS m/z (ESI^+): 272.11 $[M+Na]^+$, 521.24 $[2M+Na]^+$). The fractions containing the pure product were combined and the solvent was removed using a rotary evaporator to give *N*-Boc-DL-homocysteine methyl ester **5** as a clear oil (0.707g, 2.83 mmol, 38%). MS m/z (ESI^+): 372.16 $[M+Na]^+$, 721.34 $[2M+Na]^+$. ^1H NMR (400 MHz, $CDCl_3$): δ = 1.43 (9H, s, Boc), 1.56 (1H, t, ^3J = 8.0 Hz, SH), 1.87-1.98 (1H, m, H_β), 2.04-2.16 (1H, m, H_β'), 2.49-2.66 (2H, m, H_γ), 3.76 (3H, s, $COCH_3$), 4.42-4.52 (1H, m, H_α), 5.06 (1H, d, ^3J = 7.5 Hz, NH). ^{13}C NMR (126 MHz, $CDCl_3$): δ = 20.86 (1C, s, C_γ), 28.44 (3C, s, $C(CH_3)_3$), 33.19 (1C, s, C_β), 52.38, 52.63 (2 x 1C, 2 x s, C_α/OCH_3), 80.81 (1C, s, $C(CH_3)_3$), 155.54 (1C, s, CONH), 172.98 (1C, s, $COOCH_3$).

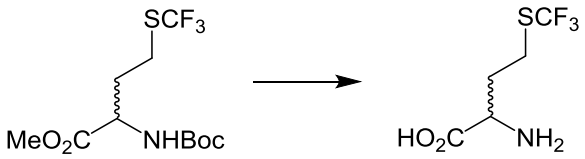
N-Boc-DL-(trifluoromethyl)homocysteine methyl ester

5 **4**

A 50 mL round bottomed-flask equipped with a magnetic stirrer was flame dried under vacuum and flushed with nitrogen gas. This vacuum/nitrogen cycle was repeated twice. The 50 mL round-bottomed flask was charged with Togni's reagent **3** (365 mg, 1.11 mmol, 1.10 eq.) before being applied to a high vacuum line for 1.5 hours. A second, 25 mL round-bottomed flask, equipped with a magnetic stirrer, was flame dried under vacuum and flushed with nitrogen. This cycle was repeated twice. A stock of **5** (0.690 g, 2.77 mmol) was put under high vacuum for 1.5 hours before being diluted with anhydrous MeOH (6.20 mL) to

give a solution of concentration 0.447 mmol/mL MeOH (as in the Organic Synthesis paper[31]). 2.25 mL (251 mg, 1.01 mmol, 1.00 eq.) of this solution was added to the 25 mL round-bottomed flask. 2.25 mL of anhydrous MeOH was added to the 50 mL round-bottomed flask. Both flasks were flushed with nitrogen again and then cooled to -78 ^{o}C (dry ice-acetone bath). The solution of **5** was added dropwise to the solution of **3** *via* a cannula over 2 minutes. The solution turned yellow immediately. The 25 mL round-bottomed flask was washed with anhydrous MeOH (2 x 1 mL) and the washing was transferred to the 50 mL round bottomed flask *via* the cannula. The solution was stirred at -78 ^{o}C for 30 minutes upon which it turned colourless. The dry ice-acetone bath was replaced with a water bath and the reaction was stirred for a further 30 minutes. TLC (20% EtOAc in petrol) revealed all of the initial amino acid had reacted and the presence of a new compound at R_f = 0.40. The solvent was removed using a rotary evaporator giving a pale yellow oil. The crude product was purified by column chromatography (10% EtOAc in petrol). TLC analysis of the fractions revealed compound **4** (R_f = 0.15) and the reduced form of Togni's reagent, **8** (R_f = 0.20, u.v. active on TLC plate). Fractions containing the titled compound **4** were combined and the solvent was removed using a rotary evaporator to give *N*-Boc-DL-(trifluoromethyl)homocysteine methyl ester as a clear oil **4** (0.261 g, 0.821 mmol, 82%). MS m/z (ESI$^+$): 340.10 $[M+Na]^+$, 657.22 $[2M+Na]^+$. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 1.41 (9H, s, Boc), 1.92-2.04 (1H, m, $\underline{H}_\beta$), 2.12-2.36 (1H, m, $\underline{H}_{\beta'}$), 2.91 (2H, t, ^3J = 7.5 Hz, $\underline{H}_\gamma$), 3.73 (3H, s, $CO_2C\underline{H}_3$), 4.34-4.46 (1H, dd, ^3J = 7.5, 4.5 Hz, $\underline{H}_\alpha$), 5.21 (1H, d, ^3J = 8.0 Hz, N$\underline{H}$COOC$(CH_3)_3$). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 25.92 (1C, unresolved q, $^3J_{C\gamma\text{-}F}$ = 2.5 Hz, $\underline{C}_\gamma$), 28.14 (3C, s, C$(\underline{C}H_3)_3$), 33.19 (1C, s, $\underline{C}_\beta$), 52.31, 52.52 (2 x 1C, 2 x s, $\underline{C}_\alpha$/O$\underline{C}H_3$), 80.25 (1C, s, $\underline{C}(CH_3)_3$), 130.9 (1C, q, $^1J_{C\text{-}F}$ = 306.0 Hz, $\underline{C}F_3$), 155.34 (1C, s, $\underline{C}$ONH), 172.17 (1C, s, $\underline{C}OOCH_3$). ^{19}F NMR (377 MHz, $CDCl_3$) δ = -41.27.

DL-Trifluoromethionine

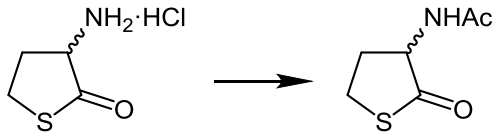


4 **1**

Compound **4** (5.26 g, 16.6 mmol, 1.0 eq.) was added to a 500 mL round-bottomed equipped with a magnetic stirrer. CH_2Cl_2 (60 mL) was added to the round-bottomed flask and the mixture was stirred to dissolve compound **4**. The flask was placed under high vacuum and then flushed with argon. The flask was placed under an argon atmosphere. Trifluoroacetic acid (60 mL, 784 mmol, 47.3 eq.) was added and the reaction was stirred at room temperature for 4 hours. The product formed is easily distinguished from the starting material by TLC (10 % EtOAc in petrol) at $R_f = 0.00$ with the starting material at $R_f \approx 0.2$. The solvent and excess trifluoracetic acid was removed using a rotary evaporator and the oily residue was dried briefly under high vacuum. The resulting residue was dissolved in THF (17 mL) and the stirred solution was cooled to 0 ^oC with an ice bath. A solution of 5 M LiOH was made up by adding H_2O to 7.34 g of $LiOH \cdot H_2O$ and making up to 35 mL. 5 M LiOH (33.2 mL) was added to the reaction mixture. The ice bath was removed and the reaction was stirred at room temperature for 2 hours. The reaction was then diluted with H_2O (60 mL). The reaction was neutralised by addition of DOWEX® 50WX8-200 (H^+) (added until pH < 7, as indicated by pH paper). The reaction suspension was poured into an empty column and washed with 500 mL H_2O (gravity flow) and the flow-through was discarded. The column was eluted with 5% NH_4OH (aq.) and the fractions were collected. Fractions containing the product (as indicated by TLC with 10% water in acetone using ninhydrin stain) were pooled and the solvent was removed using a rotary evaporator (with the water bath temperature set to 60 ^oC). The crude product was dry loaded onto 3.5 g of TELOS and purified by silica gel chromatography (10% water in acetone) to yield DL-trifluoromethionine **1** as a white crystalline solid (2.02 g, 9.94

mmol, 60% yield). Melting point: 227-230 ^{o}C (Lit: 227-229 ^{o}C[1]). ^{1}H NMR (400 MHz, D_2O): δ = 2.27 (2H, m, $\underline{H}_{\beta}$), 3.10 (2H, m, $\underline{H}_{\gamma}$). 3.85 (1H, t, *J* = 6.5 Hz, $\underline{H}_{\alpha}$).

***N*-Acetyl-*D,L*-homocysteine thiolactone**

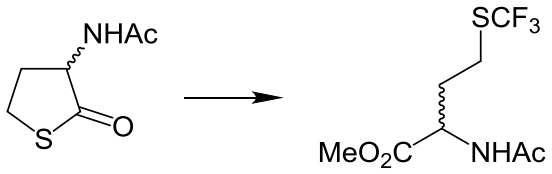


12 **13**

Compound **12** (7.00 g, 45.6 mmol, 1.0 eq.) was added to a 500 mL round-bottomed flask and sealed with a rubber septum. The reactant was put under high vacuum for 2 hours. 200 mL of CH_2Cl_2 was added and the mixture was stirred. The reaction was cooled to 0 ^{o}C (ice bath) and the reaction was left to cool for 5 minutes before carefully adding Et_3N (25.4 mL, 182 mmol, 4 eq.). CH_3COCl (4.86 mL, 68.3 mmol, 1.5 eq.) was added dropwise over 10 minutes upon which the solution turned yellow-brown immediately. The reaction was stirred for 10 minutes before removing the ice bath and was left stirring overnight. Analysis of the reaction mixture by TLC (EtOAc) showed all of the starting material (R_f = 0.15) had reacted and two new compounds were present eluting at R_f = 0.25 (the predominant product) and R_f = 0.45. The reaction was washed with H_2O-ice (2 x 50 mL), then $NaHCO_3$ (2 x 50 mL) and finally 1M HCl (2 x 50 mL). The organic phase was separated from the aqueous phase after each washing. The aqueous phases were combined and extracted with CH_2Cl_2 (3 x 50 mL). The combined organic phases were dried over $MgSO_4$ and filtered. The solvent was removed using a rotary evaporator and the crude product was purified by column chromatography eluting first with 8:2 EtOAc/petrol followed by 9:1 EtOAc/petrol, followed by pure EtOAc. *N*-Acetyl-D,L-homocysteine thiolactone **13** was isolated as a white solid (3.44 g, 21.6 mmol, 47%). Melting point: 109.5-111.5 ^{o}C (Lit: 111 ^{o}C[36]). MS m/z (ESI^{+}): 160.1 $[M+H]^{+}$, 182.1 $[M+Na]^{+}$, 341.1 $[2M+Na]^{+}$.^{1}H NMR (400 MHz, $CDCl_3$): 1.86-2.02 (1H, m, ^{1}J = 7.0, 5.5 Hz, $SCH_2C\underline{H}$), 1.91 (3H, s, $COC\underline{H}_3$), 2.54-2.65 (1H, m, $SCH_2C\underline{H}$'), 3.10-3.18 (1H, m, $SC\underline{H}_2$),

3.20-3.30 (1H, td, *J* = 6.5, 5.0 Hz, SC<u>H</u>'$_2$), 4.44-4.53 (1H, m, NHC<u>H</u>CO), 6.89-7.01 (1H, d, *J* = 7.0 Hz, N<u>H</u>). ^{13}C NMR (101 MHz, $CDCl_3$): 22.71 (1C, s, CO<u>C</u>H$_3$), 27.25 (1C, s, S<u>C</u>H$_2$), 31.00 (1C, s, SCH$_2$<u>C</u>H$_2$), 58.90 (1C, s, NH<u>C</u>HCO), 170.83 (1C, s, NH<u>C</u>O), 205.61 (1C, s, CH<u>C</u>OS).

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



13 **14**

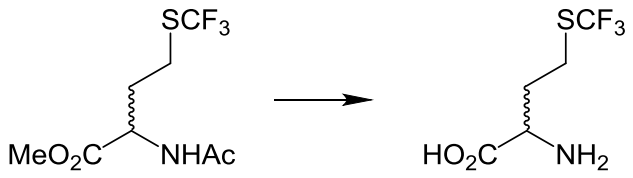
A cold finger condenser was attached to a 250 mL, two-necked, round-bottomed flask equipped with a magnetic stirrer. A rubber stopper was used to seal the neck opening, and the pipe fixed to the outlet of the cold finger was also sealed. The closed system was flame dried under vacuum followed by flushing with nitrogen gas. This process was repeated three times. Compound **13** (5.08 g, 31.9 mmol, 1.0 eq.) was added to the round-bottomed flask and the closed system was evacuated and flushed with nitrogen three times, then left under high vacuum for 1 hour. Anhydrous MeOH (137 mL) was added to **13** *via* a syringe, followed by NaOMe solution (25 wt. % in methanol, 18.3 mL, 79.8 mmol, 2.5 eq.). The reaction was stirred vigorously for 45 minutes after which TLC (EtOAc) revealed complete consumption of the starting material, which eluted at R_f = 0.25, and the formation of two new compounds, running at R_f = 0.10 and R_f = 0.45.

Dry ice and acetone were added to the cold finger condenser and CF_3I (~25 g, 128 mmol, 4.0 eq.) was introduced into the flask over 5 minutes. The solution was irradiated with long wavelength u.v. radiation for 1 hour. TCL (EtOAc) indicated formation of the product **14** at R_f = 0.55 with persistence of the compound at R_f = 0.10. Mass spectrometry indicated

that these products were **14** (m/z (ESI^+): 282.03 $[M+Na]^+$, 541.09 $[2M+Na]^+$) and the corresponding disulfide (m/z (ESI^+): 403.10 $[M+Na]^+$).

The solvent was removed using a rotary evaporator and the oily residue was dissolved in CH_2Cl_2 (300 mL) and washed successively with $NaHCO_3$ (100 mL) and brine (125 mL). The organic phase was separated and dried over $MgSO_4$ and filtered, and the solvent of the filtrate was removed using a rotary evaporator. The crude oil was purified by column chromatography using 1:1 EtOAc/Petroleum. N-Acetyl-DL-(trifluoromethyl)homocysteine methyl ester **14** was isolated as a pale orange oil (3.10 g, 12.0 mmol, 38%). MS m/z (ESI^+): 282.05 $[M+Na]^+$, 541.11 $[2M+Na]^+$. ^1H NMR (500 MHz, $CDCl_3$): δ = 1.99 (3H, s, $NHCOC\underline{H}_3$), 1.95-2.05 (1H, m, $\underline{H}_\beta$), 2.19-2.27 (1H, m, $\underline{H}_{\beta'}$), 2.87 (2H, t, ^3J = 7.5 Hz, $\underline{H}_\gamma$), 3.72 (3H, s, $CO_2C\underline{H}_3$), 4.66 (1H, td, J = 8.0, 5.0 Hz, $\underline{H}_\alpha$), 6.67 (1H, d, ^3J = 7.5 Hz, $N\underline{H}COCH_3$); ^{13}C NMR (126 MHz, $CDCl_3$): δ = 22.96 ($NHCO\underline{C}H_3$), 26.06 (unresolved q, $^3J_{C\gamma\text{-}F}$ = 2.0 Hz, $\underline{C}_\gamma$), 32.91 (1C, s, $\underline{C}_\beta$), 51.17 (1C, s, $CO_2\underline{C}H_3$), 52.72 (1C, s, $\underline{C}_\alpha$), 130.93 (1C, q, $^1J_{C\text{-}F}$ = 306.0 Hz, $\underline{C}F_3$), 170.58, 172.12 (2C, 2 x s, $NH\underline{C}OCH_3/\underline{C}O_2CH_3$); ^{19}F (377 MHz, $CDCl_3$): δ = - 41.37.

DL-Trifluoromethionine

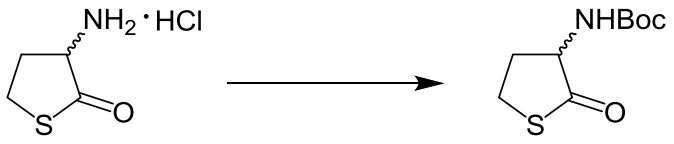


14 **1**

Compound **14** (2.90 g, 11.2 mmol, 1.0 eq.) was added to a 250 mL round-bottomed flask equipped with a magnetic stirrer, dissolved in MeOH (45 mL) and stirred at room temperature. To the stirring solution was added an aqueous solution of $LiOH{\cdot}H_2O$ (4.70 g, 112 mmol, 10 eq. in 45 mL H_2O). The reaction mixture was stirred vigorously (open to the atmosphere) for 1 hour. TCL analysis (EtOAc) revealed complete consumption of the starting

material **14** (R_f = 0.55) and the formation of a new compound (R_f = 0.00). A condenser was added and the reaction mixture was heated under reflux for 48 hours. TLC (3:1:1 EtOAc:AcOH:H_2O, ninhydrin stain) revealed formation of the product (R_f = 0.40). The solvent was removed using a rotary evaporator before diluting the crude with H_2O (25 mL). DOWEX® 50WX8-200 (H^+) was added to neutralize the solution (pH < 7, as indicated by pH indicator paper). The resulting suspension was loaded onto an ion exchange column (DOWEX® 50WX8-200 (H^+)) which was prepared by washing with two column volumes of H_2O. The product was eluted with 5% NH_4OH (aq.) and the fractions containing the titled compound **1** were combined. The solvent was removed using a rotary evaporator to give an off-white solid which was purified by column chromatography (90 g silica gel) [8:1.5:0.5 ^{i}PrOH:MeOH:NH_4OH; R_f = 0.30] to give DL-trifluoromethionine **1** as a white solid (1.19 g, 5.86 mmol, 52%). Melting point: 227-230 ^oC (Lit: 227-229 ^oC[1]). MS m/z (ESI^+): 204.04 $[M+H]^+$. ^1H NMR (500 MHz, D_2O) δ = 2.26 (2H, m, $\underline{H}_\beta$), 3.09 (2H, m, $\underline{H}_\gamma$), 3.82 (1H, t, ^3J = 6.0 Hz, $\underline{H}_\alpha$); ^{13}C NMR (126 MHz, D_2O): δ = 25.42 (1C, s, $\underline{C}_\gamma$), 31.22 (1C, s, $\underline{C}_\beta$), 53.44 (1C, s, $\underline{C}_\alpha$), 130.86 (1C, q, $^1J_{C\text{-}F}$ = 305.5 Hz, $\underline{C}F_3$), 173.89 (1C, s, $\underline{C}O_2H$); ^{19}F (377 MHz, D_2O): δ = -41.12.

N-Boc-D,L-homocysteine thiolactone

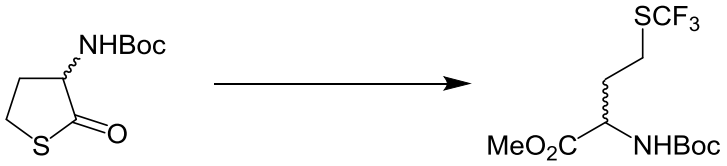


12 **15**

To a 2 L round-bottomed flask, equipped with a magnetic stirrer, compound **12** was added (10.0 g, 65.1 mmol, 1.0 eq.). CH_2Cl_2 (300 mL) was added to the flask and the suspension was stirred. Et_3N (33.6 mL, 241 mmol, 3.7 eq.) was added to the suspension and the reaction was stirred. Boc_2O (26.3 g, 121 mmol, 1.85 eq.) was added in one portion and the reaction was

stirred at room temperature for 4 hours. The reaction was diluted with 200 mL CH_2Cl_2 and washed with H_2O (200 mL) twice, after each wash the aqueous phase was separated from the organic phase. The organic phase was then dried over $MgSO_4$ and filtered. The solvent was removed using a rotary evaporator to give the crude product in the form of yellow-brown crystals. The crude product was recrystallized from hot EtOAc to yield *N*-Boc-D,L-homocysteine thiolactone **15** as a white crystalline solid (7.47 g, 34.4 mmol, 52.8% yield). Melting point: 135-136 ^oC (Lit: 124-134 ^oC[37]). ^1H NMR (400 MHz, $CDCl_3$): δ = 1.38 (9H, s, Boc), 1.87-2.04 (1H, m, $SCH_2C\underline{H}$), 2.68-2.82 (1H, m, $SCH_2C\underline{H}$'), 3.13-3.21 (1H, m, $SC\underline{H}_2$), 3.22-3.31 (1H, m, $SC\underline{H}'_2$), 4.17-4.31 (1H, m, $NHC\underline{H}CO$), 5.15 (1H, d, ^3J = 7.0 Hz, $N\underline{H}$). ^{13}C NMR (101 MHz, $CDCl_3$): 27.14(1C, s, $S\underline{C}H_2$), 28.28 (3C, s, $C(\underline{C}H_3)_3$), 31.84 (1C, s, $SCH_2\underline{C}H_2$), 60.46 (1C, s, $NH\underline{C}HCO$), 80.24 (1C, s, $\underline{C}(CH_3)_3$), 155.55 (1C, s, $NH\underline{C}O$), 205.37 (1C, s, $CH\underline{C}OS$).

***N*-Boc-D,L-(trifluoromethyl)homocysteine methyl ester**



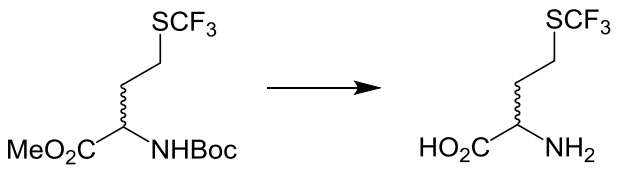
15 **4**

A two-neck 25 mL round-bottomed flask was equipped with a magnetic stirrer. A cold finger was attached to the top neck. Joints were lubricated with vacuum grease before fitting, and secured with clips. To the top of the cold finger was attached a piece of tubing, the end of which was plugged. The side-neck of the round-bottomed flask was plugged with a rubber septum. The closed system was evacuated and dried with a heat gun. The system was then flushed with nitrogen gas. Compound **15** (500 mg, 2.30 mmol, 1.0 eq.) was quickly added to the round-bottomed flask *via* the side-neck. The side-neck was quickly plugged again with a

rubber septum and the system was evacuated and then flushed with nitrogen gas. The system placed under high vacuum again for 1 hour. Anhydrous MeOH (14 mL), which had previously been cooled in an ice bath, was added to the flask *via* the side-neck and stirred. NaOMe 25 wt % in MeOH (1.32 mL, 5.75 mmol, 2.5 eq.) and the reaction mixture was stirred for 40 minutes. TLC (EtOAc) showed complete consumption of starting material (R_f = 0.80, u.v. active) and formation of a new species (R_f = 0.40, u.v. inactive, stained with ammonium molybdate hydrate). Dry ice and acetone were added to the cold finger. CF_3I (1.35 g, 6.90 mmol, 3.0 eq.) was slowly introduced into the reaction system *via* the cold finger. The reaction was stirred and irradiated with long wavelength u.v. radiation for 1 hour. A further portion of CF_3I (45 mg, 2.30 mmol, 1.0 eq.) was introduced into the reaction system *via* the cold finger, the reaction was stirred and irradiated with long wavelength u.v. radiation for 1 hour. TLC (EtOAc) indicated formation of a new product (R_f = 0.00, u.v. active). The reaction mixture was then transferred to a single-necked round-bottomed flask and the solvent was removed using a rotary evaporator. The brown-orange oily residue was taken up in CH_2Cl_2 (30 mL). The organic phase was washed with 5% $NaHCO_3$ (10 mL) and then with brine (12 mL). The organic phase was separated from the aqueous phase, dried over $MgSO_4$ and filtered. The solvent was removed using a rotary evaporator and the crude product was purified by silica gel chromatography (10% EtOAc in petrol, 10 g silica gel, 5 mL fraction volume) to yield *N*-Boc-D,L-(trifluoromethyl)homocysteine methyl ester **4** as a dull pink coloured oil (325 mg, 1.02 mmol, 44.5% yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 1.41 (9H, s, Boc), 1.92-2.04 (1H, m, $\underline{H}_\beta$), 2.15-2.28 (1H, m, $\underline{H}_\beta$'), 2.90 (2H, t, *J* = 7.5 Hz, $\underline{H}_\gamma$), 3.73 (3H, s, $CO_2C\underline{H}_3$), 4.32-4.44 (1H, dd, unresolved, $\underline{H}_\alpha$), 5.24 (1H, d, unresolved, $N\underline{H}COOC(CH_3)_3$). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 26.06 (1C, unresolved q, $^3J_{C\gamma\text{-}F}$ = 2.5 Hz, $\underline{C}_\gamma$), 28.30 (3C, s, $C(\underline{C}H_3)_3$), 33.33 (1C, s, $\underline{C}_\beta$), 52.50, 52.63 (2 x 1C, 2 x s, $\underline{C}_\alpha$/$O\underline{C}H_3$),

80.37 (1C, s, C(CH$_3$)$_3$), 130.9 (1C, q, $^1J_{C\text{-}F}$ = 306.0 Hz, CF$_3$), 155.49 (1C, s, CONH), 172.30 (1C, s, COOCH$_3$).

DL-Trifluoromethionine

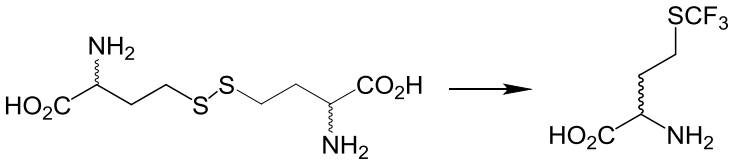


4 **1**

Compound **4** (4.21 g, 13.3 mmol, 1 eq.) was added to a 250 mL round-bottomed flask equipped with a magnetic stirrer. Compound **4** was dissolved in 45 mL of CH_2Cl_2. The flask was evacuated briefly and then flushed with argon. The vessel was then placed under an argon atmosphere. Trifluoroacetic acid (45 mL, 588 mmol, 44.3 eq.) was added to the reaction vessel and the reaction was stirred at room temperature for 3 hours. The solvent was then removed using a rotary evaporator and the residue was briefly dried under vacuum. The residue was dissolved in THF (13.5 mL) and the reaction was cooled to 0 ^oC using an ice bath. A solution of 5 M LiOH was added to the reaction (26.5 mL, 133 mmol, 10.0 eq.) and the reaction was stirred at room temperature for 1 hour. The reaction was diluted with H_2O (45 mL) and neutralized by the addition of DOWEX® 50WX8-200 (H^+) (pH < 7, as indicated by pH indicator paper). The resulting suspension was loaded onto an ion exchange column (DOWEX® 50WX8-200 (H^+)) which was prepared by washing with H_2O (500 mL). The product was eluted with 5% NH_4OH (aq.) and the fractions containing the titled compound **1** were combined. The solvent was removed using a rotary evaporator to give an off-white solid which was purified by column chromatography (9:1 Acetone:H_2O). Fractions contained the titled compound **1** were pooled and the solvent was removed using a rotary evaporator to give DL-trifluoromethionine **1** as a white solid (2.38 g, 11.7 mmol, 88%). Melting point: 227-229 ^oC (Lit: 227-229 ^oC[1]). ^1H NMR (400 MHz, D_2O) δ = 2.20 (2H, m, H_β), 3.04 (2H, m, H_γ),

3.74 (1H, t, ^{3}J = 6.5 Hz, $\underline{H}_{\alpha}$). ^{13}C NMR (101 MHz, D_2O): δ = 25.46 (1C, unresolved q, $^{3}J_{C\gamma\text{-}F}$ = 1.5 Hz, $\underline{C}_{\gamma}$), 31.61 (1C, s, $\underline{C}_{\beta}$), 53.57 (1C, s, $\underline{C}_{\alpha}$), 130.87 (1C, q, $^{1}J_{C\text{-}F}$ = 305.0 Hz, $\underline{C}F_3$), 174.87 (1C, s, $\underline{C}O_2H$).

***DL*-Trifluoromethionine**[17]



16 **1**

The procedure of Yasui *et. al.* was used.[17] A cold finger condenser was attached to a 500 mL three-necked round-bottomed flask equipped with a magnetic stirrer. A thermometer with a rubber seal was fitted to one side-neck, and a tap with tubing to an ammonia cylinder was fitted to the other side-neck. The system was dried with a heat gun under vacuum. Compound **16** (10.0 g, 37.6 mmol, 1.0 eq.) was added to the flask before evacuating the system and flushing with nitrogen gas. The flask was cooled to -78 ^{o}C (dry ice-acetone mixture). Ammonia was condensed into the flask to a volume of ≈200 mL. Sodium metal (3.6 g, 157 mmol, 4.2 eq.) was cut into pieces with a scalpel and added to the ammonia solution. The temperature of the solution was kept below -35 ^{o}C during the addition. A dark blue solution resulted, and CF_3I (18.3 g, 93.2 mmol, 2.5 eq.) was added to the reaction mixture. The reaction was stirred at -78 ^{o}C for 20 minutes, after which the flask was removed from the dry ice-acetone mixture and the ammonia was left to evaporate. The resulting residue was dissolved in 300 mL of MiliQ water. DOWEX® (H^+) was added to the solution until pH ≈ 2 (as indicated by pH paper). The slurry was poured into an empty column and washed with several column volumes of MiliQ water. The column was eluted with 0.5 % NH_4OH (aq.). The fractions containing the pure product were pooled and water was removed using a rotary evaporator (water bath at 60 ^{o}C) to give a white crystalline solid (6.71 g, 33.0 mmol, 88.6% yield). ^{1}H NMR (400 MHz, D_2O) δ = 2.20 (2H, m, $\underline{H}_{\beta}$), 3.04 (2H, m, $\underline{H}_{\gamma}$), 3.74 (1H, t, *J* = 6.5

Hz, $\underline{H}_{\alpha}$). ^{13}C NMR (101 MHz, D_2O): δ = 25.56 (1C, unresolved q, $^{3}J_{C\gamma\text{-}F}$ = 2.2 Hz, $\underline{C}_{\gamma}$), 31.71 (1C, s, $\underline{C}_{\beta}$), 53.75 (1C, s, $\underline{C}_{\alpha}$), 130.88 (1C, q, $^{1}J_{C\text{-}F}$ = 305.0 Hz, $\underline{C}F_3$), 174.97 (1C, s, $\underline{C}O_2H$). ^{19}F NMR (376 MHz, D_2O): δ = -41.06 ppm.

2.6.3General considerations for biological experiments

The biological techniques and procedures used herein are based upon those described in Molecular Cloning: A Laboratory Manual (3rd Edition).[38] All growth media, vessels and pipette tips were autoclaved at 121 °C for 1 hour to sterilise using a VWR Touchclave-R autoclave. Preparation of bacterial colonies was conducted under a laminar flow sterile cabinet (Kendro HERAsafe). MiliQ water refers to water purified using a Millipore Mili-Q purification system with a 0.22 µm outlet filter. Deionized water refers to water deionized using an Millipore ELIX water purification system. Plasmids used in this work code for ampicillin resistance. Stock solutions of 100 mg/mL of ampicillin solutions were prepared, filtered through a 0.4 µm filter and stored at 4 °C.

Growth Media

Luria Broth (LB) – 25 g of LB powder was dissolved in 1L of deionized H_2O and autoclaved.

SelenoMet™ medium – 21.6 g of SelenoMet™ base (Molecular Dimensions) was dissolved in 1 L deionized H_2O and autoclaved. Separately, 5.1 g SelenoMet™ Nutrient Mix (Molecular Dimensions) was dissolved in 50 mL deionized H_2O, sterile filtered (0.2 µm syringe filter), and added to the SelenoMet™ base solution.

Buffers

1 X TBS (tris-buffered saline) – To 6.06 g of tris base (50 mmol) and 8.77 g of NaCl (150 mmol) was added deionized H_2O to a final volume of 1 L. The pH was adjusted to 7.5 and

the solution was vacuum filtered through a 0.2 µm membrane filter using a bottle top filter unit (NALGENE™).

1 X TBS plus 2 mM EDTA (TBS + EDTA) – To 6.06 g of tris base (50 mmol), 8.77 g of NaCl (150 mmol), and 0.585 g EDTA (2 mmol) was added deionized H_2O to a final volume of 1 L. The pH was adjusted to 7.5 and the solution was vacuum filtered through a 0.2 µm membrane filter using a bottle top filter unit (NALGENE™).

1 X TAE – 100 mL 10 X TAE buffer (Bio-Rad) was diluted with 900 mL MiliQ H_2O.

SDS-PAGE electrophoresis

SDS-PAGE (gel electrophoresis) analysis used 10% Bis-Tris or 12% Bis-Tris NuPAGE® Novex® Mini Gels (Invitrogen) which were run in a XCell *SureLock*™ Mini-Cell (Invitrogen). MES NuPAGE® running buffer (20X) was used and diluted with Mili-Q water accordingly (final volume 800 mL). Gel loading buffer was prepared according to a published recipe: 100 mM Tris-HCl (pH 6.8), 4% w/w SDS (electrophoresis grade), 2% v/v β-mercaptoethanol, 0.2% w/v bromophenol blue, 20% v/v glycerol.[38] After adding sample loading buffer (2 µL) to the sample (10 µL), the sample was heated at 90 ^{o}C for 10 minutes. Gels were run at 200 V either for 40 minutes or 60 minutes. Novagen® Perfect Protein™ Marker, 15-150 kDa or 10-225 kDa, were used to approximate the molecular weights of the proteins in the samples by comparison. Proteins were visualized on the gels by staining with coomassie blue solution (Instant*Blue*™, Expedeon). Gels were destained with deionized H_2O to give a suitable contrast.

Agarose gel electrophoresis

50 mL 1 X TAE buffer was added to 0.5 g of agarose (electrophoresis grade) in a 250 mL Pyrex conical flask and microwaved for ≈1.5 minutes. The solution was poured into a gel try

and a well template was placed in the solution. When the gel had set, samples of 10 µL were loaded and the gel was run at 60 V for 1 hour. Proteins were visualized on the gels by staining with coomassie blue solution (Instant*Blue*™, Expedeon). Gels were destained with deionized H_2O to give a suitable contrast.

Determination of protein concentrations

Protein concentrations were determined using the Bradford assay.[25,26] Coomassie Plus (Bradford) reagent (Thermo Scientific) was added to 10 µL protein samples. Samples of bovine serum albumin (BSA) protein of known concentrations were measured as standards. Absorbance at 595 nm was measured using a Spectrophotometer. Protein samples and BSA standards were measured in triplicates and a mean average value was used to plot a graph of absorbance/concentration for the BSA standards. The plot was fitted to a quadratic equation and the concentrations of the samples were determined graphically.

Isopropyl β-D-1-thiogalactopyranoside (IPTG)

A stock solution of 1 M IPTG was made using Mili-Q water and filtering through a 0.2 µm filter. This was stored at 4 ^{o}C. Protein expression was induced by addition of this stock solution to media to give a final concentration of 1 mM IPTG.

2.6.4Protein Production Procedures

2.6.4.1 Initial production and purification of Qβ K16M in the presence of DL-TFM

Expression: Six 10 mL starter cultures of SelenoMet™ media containing 0.1 mg/mL ampicillin, supplemented with 40 mg/L (0.27 mM) L-Met were inoculated with 1 µL of a glycerol stock of *E. coli* B834(DE3) cells containing the p75M plasmid carrying the K16M Qβ gene. The starter cultures were incubated at 37 ^{o}C for 16 hours at 250 rpm. Flasks containing 625 mL SelenoMet™ medium supplemented with 0.1 mg/mL ampicillin and 40 mg/L (0.27 mM) L-Met were inoculated with 1 mL of the same overnight culture. The cells

were incubated at 37 ^{o}C at 250 rpm until the OD_{600} reached 0.6 – 0.8. The cells were then pelleted by centrifugation (20 minutes, 8000 rpm, Beckman JLA 9.1 rotor, 4 ^{o}C). For each flask, the supernatant was discarded and the cell pellet was resuspended in 400 mL of SelenoMet™ (base only, no nutrient mix added, no L-Met added) supplemented with 0.1 mg/mL ampicillin. The cells were pelleted again and washed in the same way. The cells were washed a total of three times. After harvesting the cells, each of the final cell pellets were resuspended in 20 mL of SelenoMet™ base and transferred to flasks of 625 mL of fresh SelenoMet™ media containing 0.1 mg/mL ampicillin, supplemented with 1.3 mM DL-TFM and pre-warmed to 37 ^{o}C. The flasks were incubated at 37 ^{o}C for 30 minutes, followed by 30 ^{o}C for 30 minutes (with 250 rpm shaking). Protein expression was then induced, by addition of 1 mM IPTG, and the flasks were incubated at 37 ^{o}C for 6 hours (with 250 rpm shaking). Cells were pelleted by centrifugation 20 minutes, 8000 rpm, Beckman JLA 9.1 rotor, 4 ^{o}C). Each cell pellet was resuspended in 20 mL TBS + EDTA. Each cell/buffer suspension was sonicated on ice (1 minute on, 1 minute off, for a total of 10 minutes). The cell debris was collected by centrifugation (30 minutes, 20000 rpm, Beckman JA-25.50 rotor, 4 ^{o}C) and the clear lysate from each was added to a 50 mL Falcon tube.

Extraction of Qβ virus-like particles (VLPs): The cleared lysates were added to an equal volume of 1:1 v/v *n*-butanol:chloroform and mixed by inversion and a vortex mixer. The layers were separated by centrifugation (20 minutes, 6500 rpm). The aqueous layer (top) was removed and the VLPs were precipitated by addition of PEG 8000 (10% w/v) and NaCl (100 mM). The mixture was vortexed and then incubated on a rocker at 4 ^{o}C for 30 minutes. The precipitated VLPs were collected by centrifugation (20 minutes, 6500 rpm) and dissolved in 10 mL of TBS. The extraction step was repeated once more.

Sucrose density gradient purification of VLPs: After the second extraction, a sucrose density gradient was used to purify the VLPs. Solutions of 10-40 sucrose w/w % in TBS buffer, with 5% increments, were prepared (solutions were filtered through 0.2 µm membrane). Tubes: Ultra-Clear, 14 X 89 mm. Ultracentrifuge: Beckman Ultracentrifuge L-90K. Rotor: SW 40 Ti, swing bucket. The gradients were spun at 37600 rpm, 3 hrs, 4 ^{o}C with maximum acceleration and the deceleration profile set to "No brake". The gradient was fractionated into 1 mL aliquots by pipetting from the meniscus downwards. Samples containing the VLP (as determined by SDS-PAGE) were pooled and precipitated by addition of 2 M ammonium sulphate and allowed to rock at 4 ^{o}C for 2 hours. The precipitate was collected by centrifugation (20 minutes, 8000 rpm). The precipitated protein was resuspended in 2 mL TBS.

ESI-MS: 3.3 µL of the protein sample was added to 46.7 µL TBS, 75 µL 8 M urea and 7.5 µL 1 M DTT. The sample was incubated for 37 ^{o}C for 1 hour before injecting into the mass spectrometer. The resulting mass spectrum is that shown in Fig. 2.8 B.

2.6.4.2 Production and purification of Qβ K16M in the presence of L-Met

Expression: Two 10 mL starter cultures of LB medium containing 0.1 mg/mL ampicillin were inoculated with 1 µL of a glycerol stock of *E. coli* B834(DE3) cells containing the p75M plasmid carrying the K16M Qβ gene. The starter cultures were incubated at 37 ^{o}C for 16 hours at 250 rpm. Flasks containing 800 mL LB medium, supplemented with 0.1 mg/mL ampicillin, were inoculated with 3 mL of overnight culture. The flasks were incubated at 37 ^{o}C at 250 rpm until the OD_{600} reached 0.6 – 0.8. Protein expression was then induced, by addition of 1 mM IPTG, and the flasks were incubated at 37 ^{o}C for 6 hours (with 250 rpm shaking). Cells were pelleted by centrifugation (10 minutes, 8000 rpm, Beckman JLA 9.1 rotor, 4 ^{o}C). Each cell pellet was resuspended in 40 mL TBS + EDTA. Each cell/buffer

suspension was sonicated on ice (1 minute on, 1 minute off, for a total of 10 minutes). The cell debris was collected by centrifugation (30 minutes, 20000 rpm, Beckman JA-25.50 rotor, 4 ^{o}C) and the clear lysate from each was added to a 50 mL Falcon tube.

Extraction of Qβ virus-like particles (VLPs): The procedure used for solvent extraction of the Qβ VLPs was the same as that described in Section 2.6.4.1 (page 88).

Sucrose density gradient purification of VLPs: The procedure used for sucrose density gradient purification was the same as that described in Section 2.6.4.1 (page 89).

ESI-MS: 6.6 μL 0.5 M TCEP was added to 20 μL of the protein sample (0.08 mg/mL). The resulting mass spectrum is that shown in Fig. 2.8 A.

2.6.4.3 Variation of the concentration of L-Met supplemented into growth medium

Six 10 mL starter cultures of SelenoMet™ medium containing 0.1 mg/mL ampicillin were prepared and supplemented with L-Met to give the following final concentrations of L-Met: 0.05 mM, 0.10 mM, and 0.27 mM (two cultures of each concentration). The started cultures were inoculated with 1 μL of a glycerol stock of *E. coli* B834(DE3) cells containing the p75M plasmid carrying the K16M Qβ gene and incubated at 37 ^{o}C at 250 rpm overnight. The next morning 4.2 mL of the starter cultures were each used to inoculate 420 mL of prewarmed (to 37 ^{o}C) SelenoMet™ medium containing 0.1 mg/mL ampicillin supplemented with the same concentration of L-Met as in the starter culture of the inoculate (i.e. 0.05 mM, 0.10 mM, or 0.27 mM). The absorbance at 600 nm (OD_{600}) was measured for each of the flasks at the following number of minutes after inoculation: 60, 120, 180, 240, 320, 360, 420, 480. The results were plotted to give the graph shown in Fig. 2.9.

2.6.4.4 Production of Qβ K16M with 2.08 mM and 4.11 mM DL-TFM

Expression: Three 10 mL starter cultures of SelenoMet™ medium containing 0.1 mg/mL ampicillin and supplemented with 14.9 mg/L (0.1 mM) L-Met were inoculated with 1 µL of a glycerol stock of *E. coli* B834(DE3) cells containing the p75M plasmid carrying the K16M Qβ gene. The starter cultures were incubated at 37 ^{o}C for 15.25 hours at 250 rpm. Each of three 500 mL flasks containing 157.5 mL SelenoMet™ medium, supplemented with 0.1 mg/mL ampicillin and 14.9 mg/L (0.1 mM) L-Met, were inoculated with 1.5 mL of the same overnight culture (final volume = 159.2 mL). The cells were incubated at 37 ^{o}C at 250 rpm until the OD_{600} reached 0.65. The cells were then pelleted by centrifugation (10 minutes, 9340 rpm, Beckman JA-10 rotor, 4 ^{o}C). For each flask, the supernatant was discarded and the cell pellet was resuspended in 100 mL of SelenoMet™ (base only, no nutrient mix added, no L-Met added) supplemented with 0.1 mg/mL ampicillin. The suspension was shaken and inverted. The cells were pelleted again and washed in the same way. The cells were washed a total of four times. Whilst the cells were being harvested, two 500 mL flasks containing 157.5 mL of fresh SelenoMet™ medium, 0.1 mg/mL ampicillin and supplemented with DL-TFM, were pre-warmed to 37 ^{o}C. The final concentrations of DL-TFM were 2.08 mM and 4.11 mM (final culture volumes after adding DL-TFM solution, 160.2 mL and 162.8 mL, respectively). After harvesting the cells, each of the final cell pellets was resuspended in 5 mL of the pre-warmed medium and then added to the bulk. The flasks were incubated at 37 ^{o}C for 30 minutes, followed but 30 ^{o}C for 30 minutes (with 250 rpm shaking). Protein expression was then induced, by addition of 1 mM IPTG at 37 ^{o}C for 6 hours (with 250 rpm shaking). Cells were pelleted by centrifugation 10 minutes, 9340 rpm, Beckman JA 10 rotor, 4 ^{o}C). Each cell pellet was resuspended in 15 mL TBS + EDTA. Each cell/buffer suspension was sonicated on ice (5 seconds on, 5 seconds off, for a total of 10 minutes). The cell debris was collected by

centrifugation (30 minutes, 20000 rpm, Beckman JA-25.50 rotor, 4 ^{o}C) and the clear lysate from each was added to a 50 mL Falcon tube.

Extraction of Qβ virus-like particles (VLPs): The procedure used for solvent extraction of the Qβ VLPs was the same as that described in Section 2.6.4.1 (page 88).

Sucrose density gradient purification of VLPs: The procedure used for sucrose density gradient purification was the same as that described in Section 2.6.4.1 (page 89). Fractions of the sucrose density gradient were analysed by SDS-PAGE (Fig. 2.15). Fractions 5 to 9 of each sample were pooled and the sucrose was removed by dialysis against 4 L TBS (Slide-A-Lyzer, 10 K MWCO, 3-12 mL sample volume). Samples were concentrated using VivaSpin 6 units.

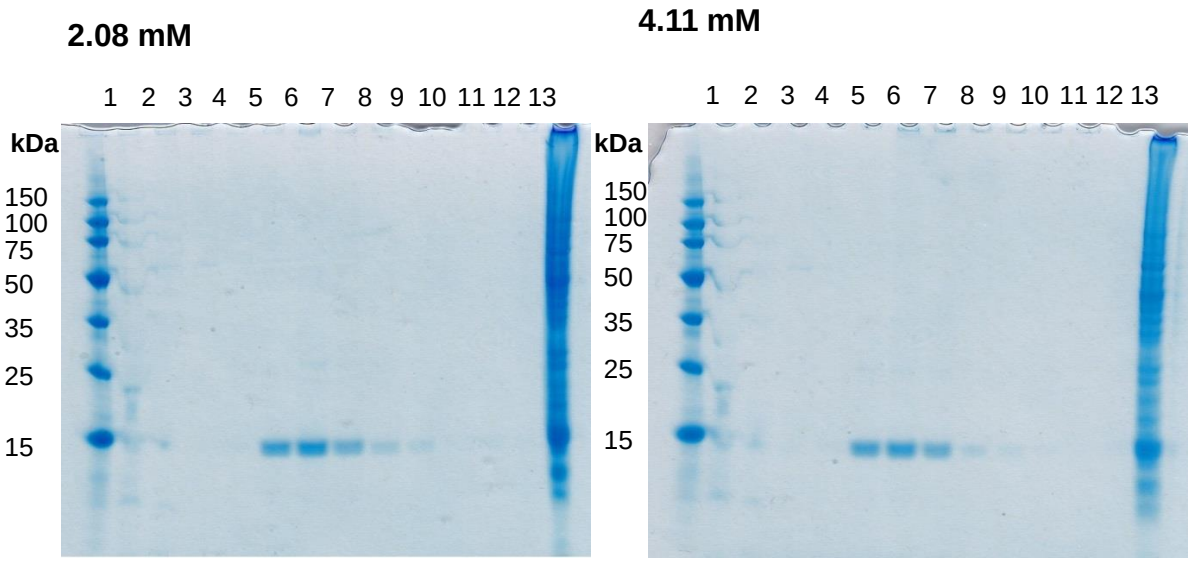


Figure 2.15 *SDS-PAGE of sucrose density gradient fractions of Qβ K16M expressed in SelenoMet™ medium supplemented with a concentration of* DL*-TFM as labelled. A 15-150 kDa ladder was used for each gel.*

ESI-MS: 6.6 µL of TBS and 6 µL (0.5 M) of TCEP were added to 13.4 µL (2.9 mg/mL) of Qβ-F sample. The resulting mass spectra for the purified protein expression in the presence of 2.08 mM and 4.11 mM DL-TFM are shown in Fig. 2.10 A and B, respectively.

2.6.4.5 Production of Qβ K16M with varying concentrations of DL-TFM

Expression trials: Three 10 mL starter cultures of SelenoMet™ medium containing 0.1 mg/mL ampicillin and supplemented with 14.9 mg/L (0.1 mM) L-Met was inoculated with 1 μL of a glycerol stock of *E. coli* B834(DE3) cells containing the p75M plasmid carrying the K16M Qβ gene. The starter cultures were incubated at 37 ^{o}C for 14.4 hours at 250 rpm. 500 mL flasks containing 157.5 mL SelenoMet™ medium supplemented with 0.1 mg/mL ampicillin and 14.9 mg/L (0.1 mM) L-Met were inoculated with 1.5 mL of the same overnight culture (final volume = 159.2 mL). The cells were incubated at 37 ^{o}C at 250 rpm until the OD_{600} reached 0.65 (≈4 hours). The cells were then pelleted by centrifugation (10 minutes, 8000 rpm, Beckman JA-10 rotor, 4 ^{o}C). For each flask, the supernatant was discarded and the cell pellet was resuspended in 100 mL of SelenoMet™ (base only, no nutrient mix added, no L-Met added) supplemented with 0.1 mg/mL ampicillin. The suspension was shaken and inverted. The cells were pelleted again and washed in the same way. The cells were washed a total of four times. Whilst the cells were being harvested, four 500 mL flasks containing 157.5 mL of fresh SelenoMet™ medium containing 0.1 mg/mL ampicillin, supplemented with DL-TFM, were pre-warmed to 37 ^{o}C. The final concentrations of DL-TFM were 6.48 mM, 8.51 mM, 10.46 mM and 15.06 mM. After harvesting the cells after the fourth wash, each of the final cell pellets was resuspended in 5-10 mL of the pre-warmed medium and then added to the bulk. The flasks were incubated at 37 ^{o}C for 30 minutes, followed but 30 ^{o}C for 1 hour (with 250 rpm shaking). Protein expression was then induced, by addition of 1 mM IPTG at 37 ^{o}C for 6 hours (with 250 rpm shaking). Cells were pelleted by centrifugation (10 minutes, 9340 rpm, Beckman JA 10 rotor, 4 ^{o}C). Each cell pellet was resuspended in 25 mL TBS + EDTA. Each cell/buffer suspension was sonicated on ice (5 seconds on, 5 seconds off, for a total of 10 minutes). The cell debris was collected by

centrifugation (30 minutes, 20000 rpm, Beckman JA-25.50 rotor, 4 ^{o}C) and the clear lysate from each was added to a 50 mL Falcon tube.

Extraction of Qβ virus-like particles (VLPs): The procedure used for solvent extraction of the Qβ VLPs was the same as that described in Section 2.6.4.1 (page 88).

Sucrose density gradient purification of VLPs: The procedure used for sucrose density gradient purification of the Qβ VLPs was the same as that described in Section 2.6.4.1 (page 89). Fractions of the sucrose density gradient were analysed by SDS-PAGE (Fig. 2.11). Fractions 5 to 7 of each sample were pooled and the sucrose was removed using PD-10 desalting columns.

ESI-MS: 6.48 mM DL-TFM sample: 2 µL of TCEP (0.5 M) were added to 10 µL of protein sample (0.16 mg/mL). 8.51 mM DL-TFM sample: 2.3 µL of TBS and 2 µL of TCEP (0.5 M) were added to 7.7 µL of protein sample (0.39 mg/mL). The resulting mass spectra for the purified protein expressed in the presence of 6.48 mM and 8.51 mM DL-TFM are shown in Fig. 2.12 A and B, respectively.

2.3.4.6 Production of Qβ-K16M with varying expression times

Expression: Three 10 mL starter cultures of SelenoMet™ medium containing 0.1 mg/mL ampicillin and supplemented with 14.9 mg/L (0.1 mM) L-Met were inoculated with 1 µL of a glycerol stock of *E. coli* B834(DE3) cells containing the p75M plasmid carrying the K16M Qβ gene. The starter cultures were incubated at 37 ^{o}C for 14.7 hours at 250 rpm. Each of four 500 mL flasks containing 157.5 mL SelenoMet™ medium supplemented with 0.1 mg/mL ampicillin and 14.9 mg/L (0.1 mM) L-Met were inoculated with 1.5 mL of the same overnight culture. The cells were incubated at 37 ^{o}C at 250 rpm until the OD_{600} reached 0.65. The cells were then pelleted by centrifugation (10 minutes, 8000 rpm, Beckman JA-10 rotor, 4 ^{o}C). For each flask, the supernatant was discarded and the cell pellet was resuspended in

100 mL of SelenoMet™ (base only, no nutrient mix added, no L-Met added) supplemented with 0.1 mg/mL ampicillin. The suspension was shaken and inverted. The cells were pelleted again and washed in the same way. The cells were washed a total of four times. Whilst the cells were being harvested, four 500 mL flasks containing 157.5 mL of fresh SelenoMet™ media containing 0.1 mg/mL ampicillin were pre-warmed to 37 ^{o}C. After harvesting the cells after the fourth wash, each of the final cell pellets was resuspended in 5-10 mL of the pre-warmed media and then added to the bulk. The flasks were incubated at 37 ^{o}C for 30 minutes, followed but 30 ^{o}C for 30 minutes (with 250 rpm shaking). DL-TFM was added to each flask to a concentration of 6.1 mM. The flasks were then incubated at 37 ^{o}C at 250 rpm for 15 minutes before adding IPTG to a concentration of 1 mM. The four flasks were then incubated at 37 ^{o}C at 250 rpm for different periods of time: 10 hours, 12 hours, 14 hours and 16 hours.

The cells were pelleted by centrifugation (10 minutes, 8000 rpm, Beckman JA-10 rotor, 4 ^{o}C) and resuspended in 25 mL TBS + EDTA. Each cell/buffer suspension was sonicated on ice in a glass beaker (5 seconds on, 5 seconds off, for a total of 10 minutes). The cell debris was collected by centrifugation (30 minutes, 20000 rpm, Beckman JA-25.50 rotor, 4 ^{o}C) and the clear lysate from each was added to a 50 mL Falcon tube.

Extraction of Qβ virus-like particles (VLPs): The procedure used for solvent extraction of the Qβ VLPs was the same as that described in Section 2.6.4.1 (page 88).

Sucrose density gradient purification of VLPs: The procedure used for sucrose density gradient purification was the same as that described in Section 2.6.4.1 (page 89). Fractions of the sucrose density gradient were analysed by SDS-PAGE (Fig. 2.13). Fractions 5 to 9 of each sample were pooled and the sucrose was removed by dialysis against 4 L deionized water (Slide-A-Lyzer, 10K MWCO, 12-30 mL sample volume). Samples were freeze-dried and the resulting white solid was dissolved in 500 µL of TBS.

ESI-MS: 10 hours expression time sample: 1 µL of DTT (1 M) and 2 µL of urea (8 M) were added to 10 µL of protein sample (0.08 mg/mL). 12 hours expression time sample: 1 µL of DTT (1 M) and 2 µL of urea (8 M) were added to 10 µL of protein sample (0.05 mg/mL). The resulting mass spectra for the purified proteins expressed for 10 hours and 12 hours are shown in Fig. 2.14 A and B, respectively.

2.6.5 Validation of quantification of incorporation level by mass spectrometry

Mass spectrometry was used to determine the percentage of Qβ-F monomers present in the Qβ samples purified after each attempted production of Qβ-F. It was proposed that the ratios of the peak areas of the Qβ-Met and Qβ-F peaks in the mass spectrum would be proportional to the relative amounts of the two proteins in the sample. This is based on the assumption that the Qβ-Met and Qβ-F ionize with the same efficiency. To validate this method and assumption, a sample of Qβ-F, of known concentration, was doped with Qβ-Met, of known concentration, and the concentration ratio was correlated with the ratio measured using peak areas in the mass spectrum. The procedure was carried out as follows:

The sample of Qβ-F and the sample of Qβ-Met were diluted and their concentrations were measured using the Bradford assay and calculated to be 0.12 mg/mL and 0.11 mg/mL, respectively. Eleven protein samples were prepared by mixing the two proteins in the volume fractions shown in the table below. The proteins in the mixtures were reduced by addition of 1 µL of 0.5 M TCEP per 10 µL of protein mixture prior to injection into the mass spectrometer. The resulting mass spectra are shown in Fig. 2.16.

Volume fraction of Qβ-F sample	Volume fraction of Qβ-Met sample
5	0
5	1
5	2

5	3
5	4
5	5
4	5
3	5
2	5
1	5
0	5

The mass spectra of the mixtures of Qβ-F and Qβ-Met proteins were deconvoluted using the MaxEnt algorithm (Mass Lynx software) to a resolution of 0.05 Da per channel. The mass spectrometry data was processed as follows. A mean value of the noise was calculated using the ion count data in the regions of the mass spectrum which are absent of the protein peaks. This method is valid on the assumption that the only mass spectrum peaks corresponding to protein are those of reduced Qβ-F or Qβ-Met, that is, that the protein sample is pure. The regions were taken to be in the intervals 10 000 Da – 13 4999.9 Da and 15 000 Da – 20 000 Da. The ion count data for the whole spectrum was then decreased by this mean noise value. If the ion count was less than the mean noise value, it was set to zero and unaltered otherwise. The area under the peaks was calculated using the trapezium rule. The fraction of Qβ-Met present in the mixed samples calculated using the concentrations determined by the Bradford Assay and known volume dilutions was plotted against the

fraction of Qβ-Met present in the mixed samples calculated using peak areas in the mass spectrum (Fig. 2.17). A strong positive correlation was observed and so all incorporation levels in this chapter were calculated using the fraction of the peak areas. That is:

$$\%\ \text{incorporation of TFM} = \frac{\text{area of Qβ-F mass peak}}{(\text{area of Qβ-F mass peak})+(\text{area of Qβ-Met mass peak})} \times 100$$

3. Analysis of the particle structure of Qβ-F and comparison with its Met analogue

3.1 Introduction

Chapter 2 described the optimisation of the methodology for the incorporation of trifluoromethionine (TFM) into the coat protein of Qβ bacteriophage. It was next investigated whether the incorporation of TFM hinders the ability of the coat protein to form intact icosahedral particles, as occurs for the wild-type Qβ particle (Qβ-WT). This chapter describes work on the analysis of the particle structure of Qβ-F using established analytical techniques and compares these results with those for Qβ-WT and Qβ-Met.

3.2 Electrophoresis

Firstly, protein samples Qβ-F, Qβ-WT, and Qβ-Met were analysed by reducing SDS-PAGE (Fig. 3.1). The results indicate that each Qβ protein sample is highly purified with respect to other proteins since only a single band corresponding to the mass of a Qβ protein monomer (~ 14 kDa) is observed for each sample.

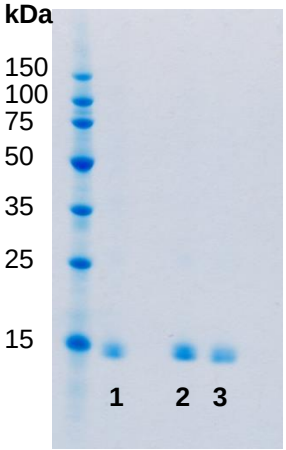


Figure 3.1 *Reducing SDS-PAGE (10% Bis-Tris gel (Invitrogen), run in MES buffer. The gel was stained with Instant*Blue™ *(coomassie blue solution)). Lane 1: Qβ-F (0.03 mg/mL); Lane 2: Qβ-WT (0.04 mg/mL); Lane 3: Qβ-Met (0.04 mg/mL). Protein markers: 15-150 kDa.*

In order to investigate whether the Qβ-F, Qβ-WT and Qβ-Met proteins were present as intact particles or not, the samples were analysed by agarose gel electrophoresis. Agarose

gel electrophoresis is commonly used for analysing and characterizing intact viruses and virus-like particles since the pore size of agarose matrices can be much greater than that of polyacrylamide matrices whilst still maintaining structural stability.[1–5] The larger pore size of agarose gels compared to polyacrylamide gels allows the migration of much larger species. The electrophoretic mobility of a species is known to be (normally) proportional to both its size and its overall charge.[6,7]

For each of the three protein samples (Qβ-F, Qβ-WT and Qβ-Met), the non-denatured proteins sample were run alongside the reduced species (generated by the addition of β-mercaptoethanol-containing loading dye to the protein sample, and heating at 90 ^{o}C for 5 minutes). The non-denatured Qβ-F protein sample and the non-denatured Qβ-Met protein sample both show the same electrophoretic mobility, which is significantly greater than that for the non-denatured Qβ-WT protein (Fig. 3.2). All three protein samples show the same electrophoretic mobility after reducing treatment with β-mercaptoethanol-containing loading dye.

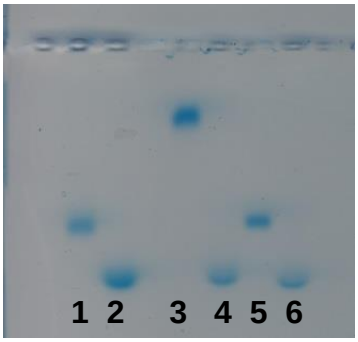


Figure 3.2 *Agarose gel electrophoresis (1% w/v agarose in TAE buffer solution). Lane 1: Qβ-F (0.04 mg/mL) native conditions; Lane 2: Qβ-F (0.04 mg/mL) after treatment with reducing dye and heating at 90 ^{o}C for 5 minutes; Lane 3: Qβ-WT (0.04 mg/mL) native conditions; Lane 4: Qβ-WT (0.04 mg/mL) after treatment with reducing dye and heating at 90 ^{o}C for 5 minutes; Lane 5: Qβ-Met (0.04 mg/mL) native conditions; Lane 6: Qβ-Met (0.04 mg/mL) after treatment with reducing dye and heating at 90 ^{o}C for 5 minutes. The gel was stained with Instant*Blue™ *(coomassie blue solution).*

Previous studies have shown that treatment of Qβ particles with reducing agent gives the monomeric form (M_r ~14 kDa).[8] Hence, the fact that the non-denatured samples have a

different electrophoretic mobility to the samples treated with reducing agent is evidence suggesting that the non-denatured form is indeed the intact VLP.

The differences in electrophoretic mobility between the Qβ-F and Qβ-Met particles and that of the Qβ-WT particles can be rationalised in terms of the different overall surface charge of the Qβ-F and Qβ-Met particles compared to that of the Qβ-WT particle. The latter has a lysine residue at position 16 (on the surface of the particle) which should be protonated at the pH of the buffer used (lysine ε-NH_3^+: pK_a 10.54[9]; TAE buffer pH 8.2), whereas Qβ-F and Qβ-Met possess a methionine residue (or analogue thereof) at position 16, which is not charged. Hence the Qβ-F and Qβ-Met particles likely have a more negative overall charge compared to the Qβ-WT particle and so migrate through the agarose gel further.[10]

3.3 Dynamic Light Scattering

Further evidence for the VLP structure of Qβ-F was sought using dynamic light scattering (DLS). DLS is a technique which uses the scattering of laser light by nanoparticles in solution to calculate the hydrodynamic diameter of the nanoparticles. Light incident on particles with a diameter much smaller than the wavelength of the light is scattered in all directions (Rayleigh scattering). The intensity of the light scattered by a particle is proportional to the sixth power of its diameter.[11] For a collection of particles the light scattered from individual particles interferes constructively and destructively forming a pattern of bright and dark areas, respectively. In a solution, the velocities of particles are constantly changing due to Brownian motion. Hence, the pattern formed changes with time. The Stokes-Einstein equation relates the rate of diffusion of a particle, D, to its hydrodynamic radius, r_s, (assuming a spherical particle):

$$D = \frac{k_B T}{6\pi\eta r_s}$$

where k_B is the Boltzmann constant, T is the absolute temperature, and η is the solvent viscosity.[12] Hence the larger the particle, the more slowly it moves in solution and so the more slowly the light pattern formed changes. The rate of change of the pattern is measured using a correlation function. The detector measures the intensity of the scattered light *I(t)* and *I(t + τ)*, at time *t* and after a time delay τ, respectively. The correlation between the two intensities is measured using the autocorrelation function:

$$g^2(\tau) = \frac{\langle I(t)I(t+\tau)\rangle}{\langle I(t)\rangle^2}$$

where the angular brackets denote the time averaged value of the enclosed function over a time range which is large compared to τ.[13] When *τ* is zero there is no difference between *I(t)* and *I(t + τ)*, and so $g^2(\tau)$ takes its maximum value - one. As *τ* increases, particles move through the solution, the difference in intensity at time *t* and time *t* + τ increases, and so $g^2(\tau)$ decreases. Since larger particles move more slowly, the decrease in $g^2(\tau)$ will be slower than for smaller particles. The diffusion coefficient *D* can be calculated from the rate of decay of $g^2(\tau)$ and using the Stokes-Einstein equation the diameter of the particle is calculated. DLS is regularly used to characterize virus-like particles[14,15] and to verify surface modifications by the resulting increase in the hydrodynamic diameter.[16,17]

Before analysing the Qβ-Met sample by DLS, the sample was centrifuged (15.7×10^3 *g* for 23 minutes) to remove large aggregates. The size distribution by intensity for the Qβ-Met sample is shown in Fig. 3.3. The data show a single population in the size distribution with a Z-Average of 30.2 nm and a polydispersity index of 0.085. This value is comparable to, but slightly greater than, previously published DLS data for Qβ-WT (28.4 nm[18]). This result shows that Qβ-Met successfully self-assembles into particles as with Qβ-WT; the slight discrepancy in value may be due to changes in solvation of the Qβ-Met particle due to its more hydrophobic surface compared to Qβ-WT.

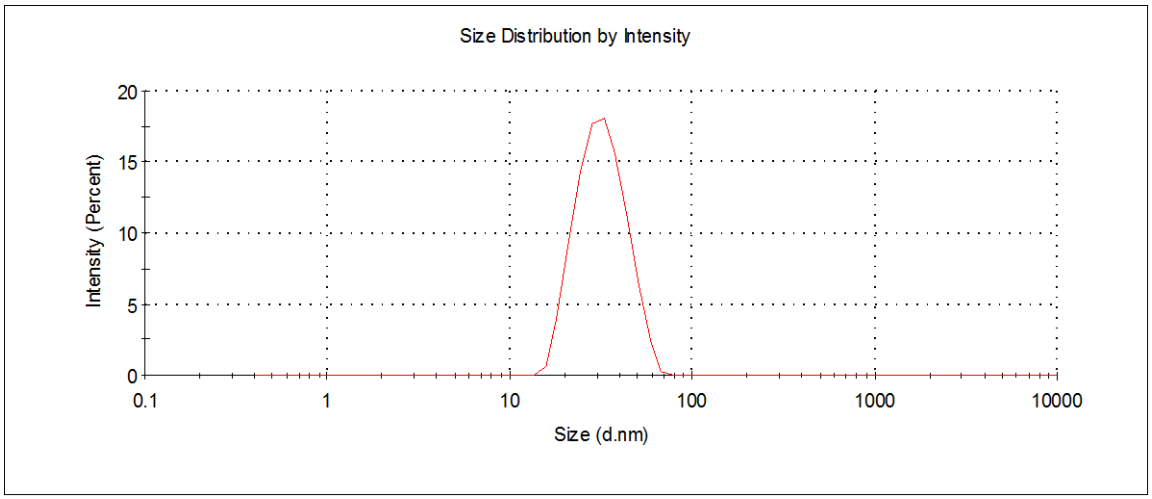


Figure 3.3 *DLS size distribution by intensity of Qβ-Met VLP sample (after centrifugation).*

The size distribution by intensity of the Qβ-F sample was then measured (Fig. 3.4). At least three populations can be seen from these data. The first is at 33.0 nm, of the order expected for a Qβ particle but 2.8 nm greater than that measured for Qβ-Met; there is a shoulder to this peak at ~120 nm, and finally a distinct peak at ~1000 nm. It was proposed that these latter two, high size peaks were due to aggregates of the VLP formed in the solution.

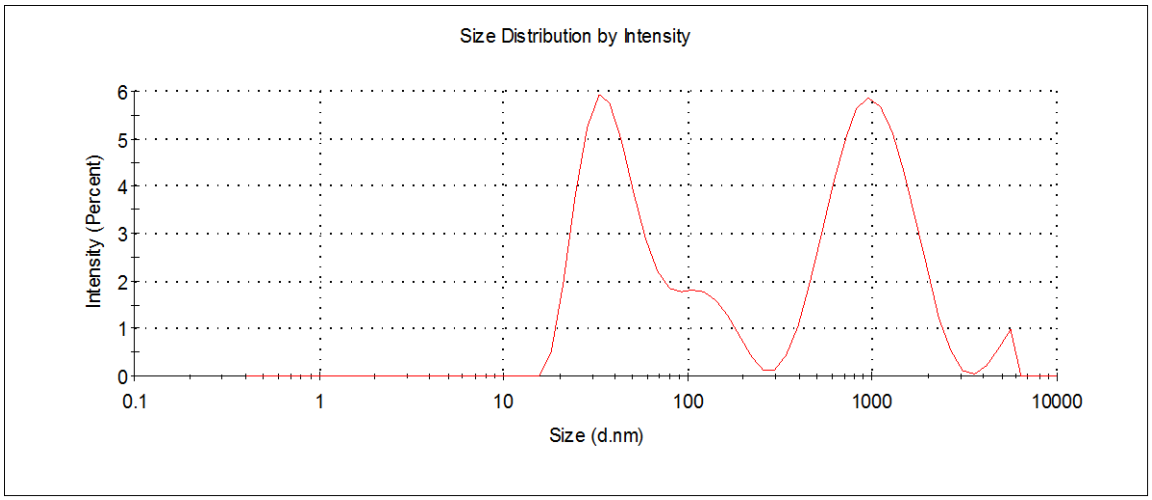


Figure 3.4 *DLS size distribution by intensity of Qβ-F VLP sample.*

In an effort to separate any aggregates present, the sample was sonicated for 3 minutes and then centrifuged (15.7 x 10^3 *g* for 30 minutes). The supernatant was removed

carefully, to prevent disturbing the pelleted aggregates, and was then immediately analysed by DLS. The resulting intensity distribution is shown in Fig. 3.5. It can be seen that the shoulder of the 33.0 nm peak is no longer present and that the high aggregate peak has reduced in intensity with respect to the 33.0 nm peak.

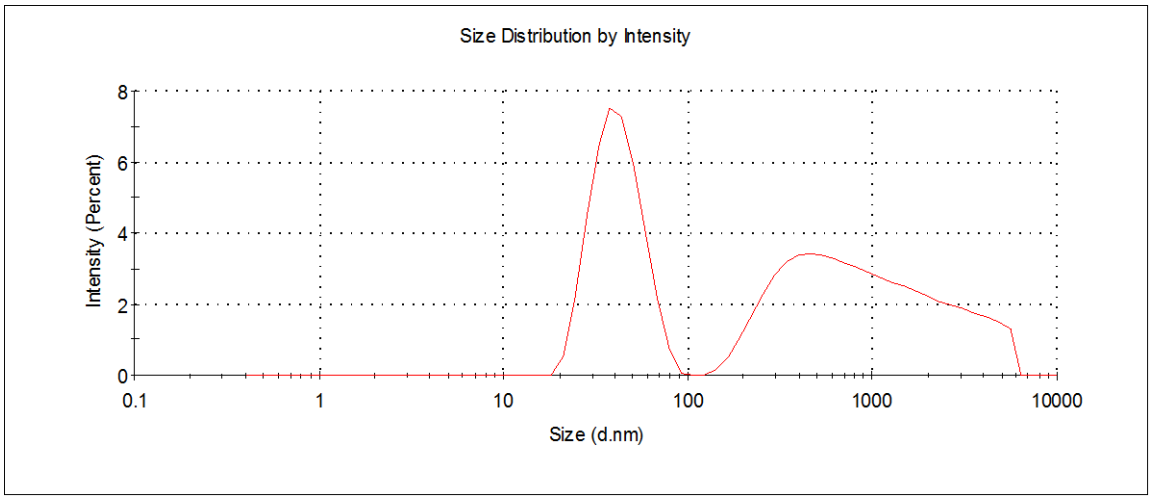


Figure 3.5 *DLS size distribution by intensity of Qβ-F VLP sample (after vortex and sonication treatment).*

Since larger particles scatter light more intensely (Rayleigh scattering, see above) the intensity distribution from DLS is biased towards the larger species in a given sample. The intensity distribution can be converted into a volume distribution (using Mie theory[19]). This gives a distribution such that the area under the peaks of different populations is related to the volume taken up by that kind of particle, and hence the distribution gives a more realistic sense of the differences in the different size populations. The size distribution by volume (Fig. 3.6) is dominated by a single symmetric peak at 30.0 nm. The tailing high-sized peak was not present in the size distribution by particle number, however the number distribution is not used here since large errors are known to be introduced when using the algorithm which converts the data from volume distribution to number distribution. As with Qβ-Met, the measured hydrodynamic diameter for Qβ-F is slightly greater than that reported for Qβ-WT (by 1.6 nm), and could be due to the difference in solvation as a result of the more hydrophobic surface of Qβ-F.

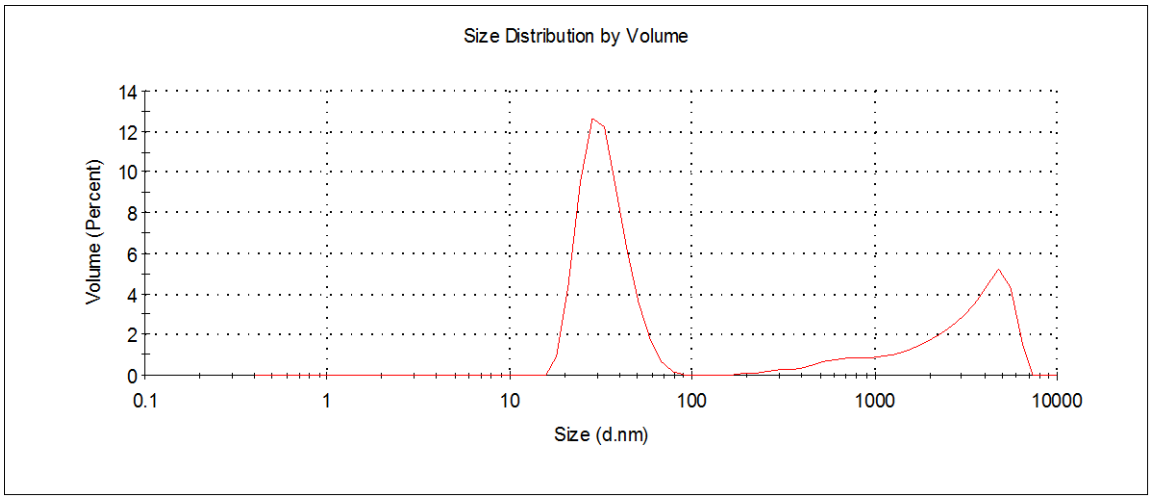


Figure 3.6 *DLS size distribution by volume of Qβ-F VLP sample (after vortex and sonication treatment).*

Overall, these data provide strong evidence that the Qβ-F protein has indeed formed intact VLPs of the expected diameter for the icosahedral bacteriophage Qβ particle.

3.4 Transmission Electron Microscopy (TEM)

Next, the protein samples Qβ-F and Qβ-WT were analysed by transmission electron microscopy (TEM). Qβ has previously been analysed by TEM and yields good quality micrographs in which the particle form can be clearly visualized.[8,18,20–23] The electron micrographs presented in this work were taken by Dr Alaa Aljibali. Samples of VLPs of concentration 0.1 mg/mL, purified by sucrose density gradient, were adsorbed onto graphite grids and stained with uranyl acetate solution. Fig. 3.7 shows a micrograph taken of Qβ-WT VLPs. The micrograph clearly shows intact particles of diameter ≈30 nm and is in agreement with previously published TEM data of Qβ-WT, in terms of both size and morphology.[20,24]

TEM images were then taken of a sample of Qβ-F. Again, the sample concentration was 0.1 mg/mL and the sample had been purified by sucrose density gradient. The protein sample was adsorbed onto graphite grids and stained with uranyl acetate solution. Fig. 3.8 shows a micrograph taken of the Qβ-F sample. The micrograph clearly shows intact particles of diameter ≈30 nm and are indistinguishable from those of Qβ-WT. Interestingly, however,

these preliminary results indicate that there is apparently a small fraction of incompletely formed shells (highlighted in Fig. 3.8 with a red box).

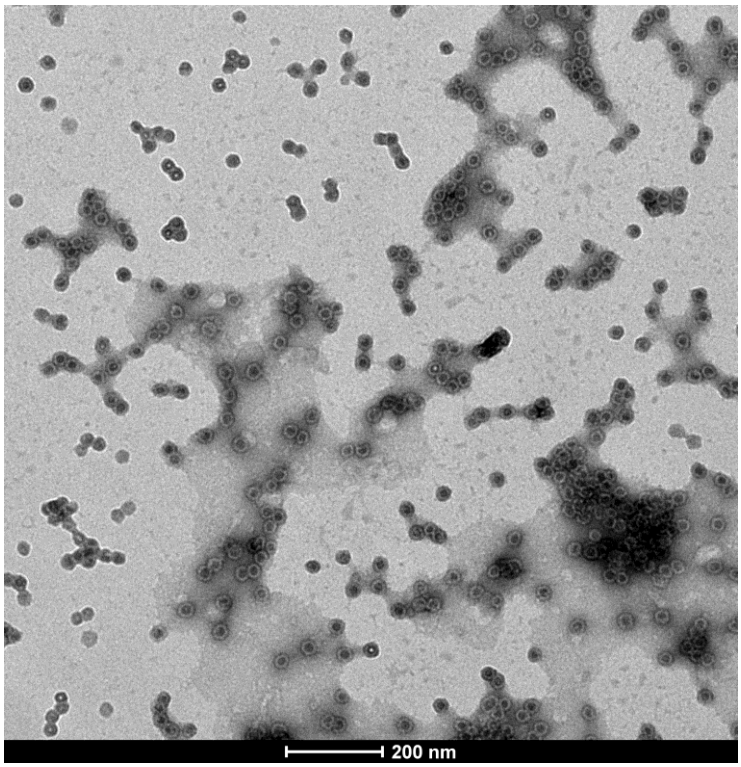


Figure 3.7 *TEM of Qβ-WT VLPs.*

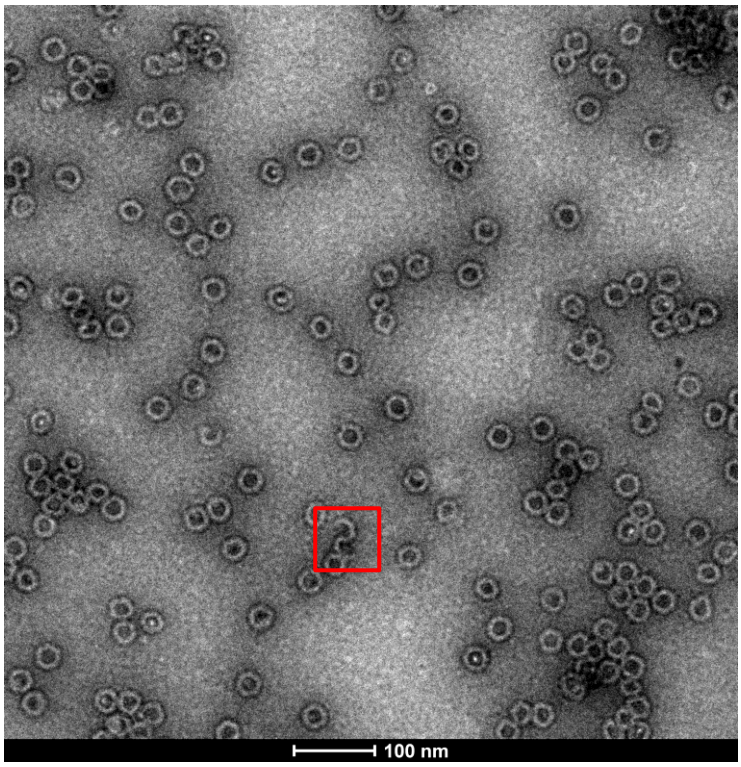


Figure 3.8 *TEM of Qβ-F VLPs. The red box highlights incompletely formed viral capsids.*

3.5 Conclusions

The samples of Qβ-F, Qβ-Met and Qβ-WT were analysed by agarose gel electrophoresis. It was concluded from the lower electrophoretic mobility of the non-denatured samples, with respect to the samples heated with reducing agent, that the proteins of each sample had formed intact particles. The electrophoretic mobility of the Qβ-WT particles was smaller than that of the Qβ-F and Qβ-Met particles. This was rationalised in terms of the more positive overall surface charge of the Qβ-WT particle due to its lysine residue at the surface position 16, which is protonated at the pH of the running buffer used. Qβ-F and Qβ-Met particles have methionine residues (or analogues thereof) at position 16, which are not protonated.

DLS was used to measure the hydrodynamic diameter of the particles Qβ-Met and Qβ-F, and their values were compared with a literature value for Qβ-WT of 28.4 nm.[18] The hydrodynamic diameters measured for Qβ-Met and Qβ-F were 30.2 nm and 30.0 nm,

respectively. It was proposed that the increase in hydrodynamic diameter for these particles with respect to Qβ-WT may be due to the differing degree of solvation.

Finally, transmission electron micrographs were taken of Qβ-WT and Qβ-F. These micrographs clearly showed intact particles in both samples which were comparable with previously reported micrographs of Qβ particles.

3.6 References

1. Steinmetz, N. & Manchester, M. PEGylated Viral Nanoparticles for Biomedicine: The Impact of PEG Chain Length on VNP Cell Interactions In Vitro and Ex Vivo. *Biomacromolecules* **10,** 784–792 (2009).

2. Steinmetz, N. F., Evans, D. J. & Lomonossoff, G. P. Chemical Introduction of Reactive Thiols Into a Viral Nanoscaffold: A Method that Avoids Virus Aggregation. *ChemBioChem* **8,** 1131–1136 (2007).

3. Rolfsson, O., Toropova, K., Ranson, N. A. & Stockley, P. G. Mutually-induced Conformational Switching of RNA and Coat Protein Underpins Efficient Assembly of a Viral Capsid. *J. Mol. Biol.* **401,** 309–322 (2010).

4. Taghavian, O. *et al.* A potential nanobiotechnology platform based on infectious bursal disease subviral particles. *RSC Adv.* **2,** 1970–1978 (2012).

5. Pokorski, J. K., Breitenkamp, K., Liepold, L. O., Qazi, S. & Finn, M. G. Functional Virus-Based Polymer À Protein Nanoparticles by Atom. 9242–9245 (2011).

6. Overbeek, J. T. G. & Bijsterbosch, B. H. *Electrokinetic Separation Methods*. 1–32 (Elsevier/North-Holland, 1979).

7. Serwer, P. & Allen, J. L. Agarose gel electrophoresis of bacteriophages and related particles IV. An improved procedure for determining the size of spherical particles. *Electrophoresis* **4,** 273–276 (1983).

8. Takamatsu, H. & Iso, K. Chemical evidence for the capsomeric structure of phage Qβ. *Nature* **298,** 819–824 (1982).

9. Dawson, R. M. C., Elliott, D. C., Elliott, W. H. & Jones, K. M. *Data for Biochemical Research*. 1–31 (Oxford University Press, 1986).

10. Shaw, D. J. *Electrophoresis*. (Academic Press, 1969).

11. Kato, H. *Nanomaterials: Processing and Characterization with Lasers*. 537 (Wiley-VCH Verlag CmbH & Co. KGaA, 2012).

12. Dill, K. A., Bromberg, S. & Dirk, S. *Molecular driving forces: statistical thermodynamics in chemistry and biology*. (Garland Science, 2003).

13. Kato, H. *Nanomaterials: Processing and Characterization with Lasers*. 535–554 (Wiley-VCH Verlag CmbH & Co. KGaA, 2012).

14. Pokorski, J. K., Hovlid, M. L. & Finn, M. G. Cell Targeting with Hybrid Qβ Virus-Like Particles Displaying Epidermal Growth Factor. *ChemBioChem* **12,** 2441–2447 (2011).

15. Hovlid, M. L. *et al.* Guiding plant virus particles to integrin-displaying cells. *Nanoscale* **4,** 3698–3705 (2012).

16. Sen Gupta, S., Raja, K. S., Kaltgrad, E., Strable, E. & Finn, M. G. Virus-glycopolymer conjugates by copper(I) catalysis of atom transfer radical polymerization and azide-alkyne cycloaddition. *Chem. Comm.* 4315–4317 (2005).

17. Ribeiro-Viana, R. *et al.* Virus-like glycodendrinanoparticles displaying quasi-equivalent nested polyvalency upon glycoprotein platforms potently block viral infection. *Nat. Commun.* **3,** 1–8 (2012).

18. 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).

19. Van de Hulst, H. C. *Light Scattering by Small Particles*. Chapters 9 and 10 (Dover Publications, 1981).

20. Overby, L. R., Barlow, G. H., Doi, R. H., Jacob, M. & Spiegelman, S. Comparison of Two Serologically Distinct Ribonucleic Acid Bacteriophages I. Properties of the Viral Particles. *J. Bacteriol.* **91,** 442–448 (1966).

21. Cielens, I. *et al.* Mutilation of RNA phage Qβ virus-like particles: from icosahedrons to rods. *FEBS Lett.* **482,** 261–264 (2000).

22. Fiedler, J. D. *et al.* Engineered Mutations Change the Structure and Stability of a Virus-Like Particle. *Biomacromolecules* **13,** 2339–2348 (2012).

23. Rhee, J.-K. *et al.* Colorful Virus-like Particles: Fluorescent Protein Packaging by the Qβ Capsid. *Biomacromolecules* **12,** 3977–3981 (2011).

24. Engelberg-Kulka, H., Israeli-Reches, M., Dekel, L. & Friedmann, A. Qβ-Defective Particles Produced in a Streptomycin-Resistant Escherichia coli Mutant. *J. Virol.* **29,** 1107–1117 (1979).

3.7 Experimental procedures

3.7.1 Electrophoresis

SDS-PAGE: 10% Bis-Tris, MES buffer; Lane 1: 15-150 kDa protein ladder. Lane 2: 0.03 mg/mL Qβ-F sample; Lane 3: left blank; Lane 4: 0.04 mg/mL Qβ-WT; Lane 5: 0.04 mg/mL Qβ-Met. The gel was stained with coomassie blue solution and is shown in Fig. 3.1.

Agarose gel Fig. 3.2: An agarose gel was set by making a 1% solution of agarose in 1 X TAE buffer (500 mg into 50 mL buffer), heating in a microwave before pouring into the gel box and allowing to cool. Lane 1: 1 µL Qβ-F sample, 7 µL TBS, and 2 µL DNA loading dye; Lane 2: 1 µL Qβ-F sample, 7 µL TBS, and 2 µL SDS-PAGE reducing dye, heated at 90 ^{o}C for 5 minutes; Lane 3: Blank; Lane 4: 1 µL Qβ-WT sample, 9 µL TBS, and 2 µL DNA loading dye; Lane 5: 1 µL Qβ-WT sample, 9 µL TBS, and 2 µL SDS-PAGE reducing dye, heated at 90 ^{o}C for 5 minutes; Lane 6: 1 µL Qβ-Met sample, 9 µL TBS, and 2 µL DNA loading dye; Lane 7: 1 µL Qβ-WT sample, 9 µL TBS, and 2 µL SDS-PAGE reducing dye, heated at 90 ^{o}C for 5 minutes. The final protein concentration of each sample was 0.04 mg/mL.

The gel was run at 60 volts for 1 hour and subsequently stained with coomassie blue dye and is shown in Fig. 3.2.

3.7.2 Dynamic Light Scattering

Dynamic light scattering measurements were taken using a Malvern Zetasizer Nano ZS instrument. Samples were 0.4–0.5 mg/mL in TBS buffer. Measurements were taken by the instrument every 10 seconds; 10 acquisitions were taken for each of 3 runs. The final value is the average of the 3 runs.

3.7.3Transmission Electron Microscopy

Transmission electron micrographs were taken using a FEI Tecnai Spirit F12 instrument at the Wellcome Trust, Old Road Campus, Oxford. Micrographs were taken by Dr Alaa Aljibali. For each sample, 30 µL of Qβ solution (TBS) was pipetted onto a carbon-coated copper grid. The grid was rinsed with MiliQ water to remove buffer salts. 2 µL of 2% uranyl acetate was pipetted onto the grid and stained for several minutes. Grids were viewed at 80 kV.

4. ^{19}F NMR of Qβ-F

4.1 Introduction

An ideal ^{19}F imaging probe should have the following characteristics: 1) a high number of ^{19}F atoms per molecule in order to generate a high signal intensity, 2) the ^{19}F NMR spectrum should have a single peak to avoid chemical shift artefacts.[1] Using bacteriophage Qβ as a construct for loading ^{19}F nuclei fulfils these criteria since: 1) each Qβ particle comprises 180 monomers, hence a very large number of ^{19}F nuclei can be introduced, and 2) the icosahedral symmetry of Qβ ensures the ^{19}F-groups are symmetrically related, hence each ^{19}F-group would be equivalent and give rise to a single NMR peak. Both of these features would contribute to a high NMR sensitivity.

The first goal was to measure a ^{19}F NMR signal from the fluorinated Qβ virus-like particle (VLP) developed in the previous chapters.

4.2 ^{19}F NMR of Qβ-F virus-like particle

Firstly, a ^{19}F NMR spectrum of DL-trifluoromethionine (DL-TFM) was measured in order to investigate the ^{19}F chemical shift range of the SCF_3 group. The spectrum was referenced to an internal standard (1,1,1-trifluoroacetone δ_F -84.60 ppm), and gave a peak at -39.12 ppm (Fig 4.1).

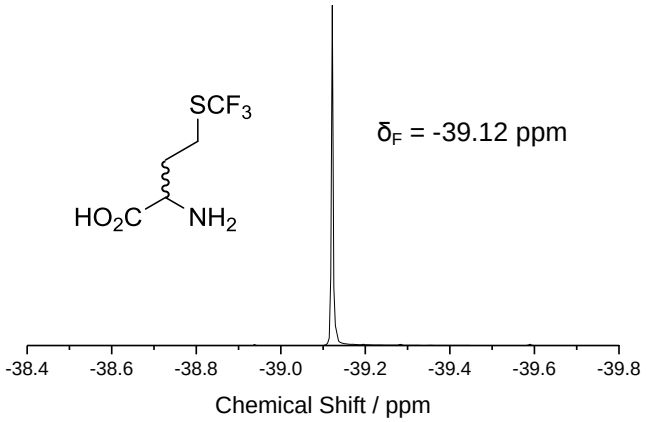
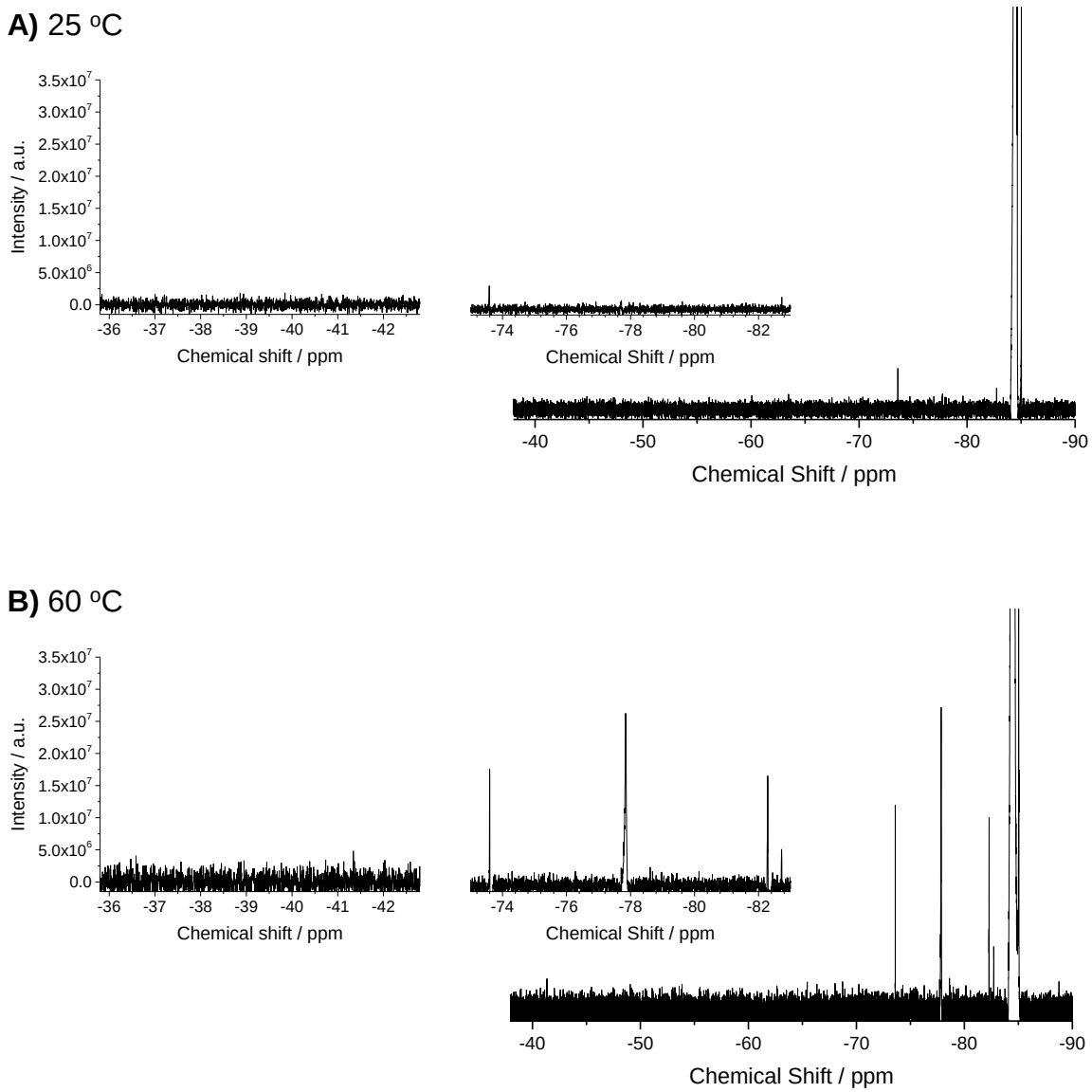


Figure 4.1 *^{19}F NMR of DL-TFM. The spectrum is referenced to 1,1,1-trifluoroacetone δ_F = -84.60 ppm.*

Next, a ^{19}F NMR spectrum of a purified sample of Qβ-F VLPs (0.4 mg/mL, 0.03 mM in monomer) was measured at room temperature. Disappointingly, no ^{19}F NMR signal was observed at the expected chemical shift value of the SCF_3 group (Fig. 4.2 A). The peaks observed at -73.58 ppm and -82.73ppm both have FWHH values of ~3 Hz, and so were suspected to originate from small organic molecules as opposed to from a protein. On further inspection of the ^{19}F NMR spectrum of DL-TFM, peaks were observed at -73.59 ppm and -82.41 ppm, both with FWHH values of 2.0 Hz (peak areas: 0.08% and 0.26% that of the CF_3 peak at -39.12 ppm, respectively). Hence the peaks observed in the ^{19}F NMR spectrum of Qβ-F VLP in Fig. 4.2 A were taken to be residual small molecule impurities from the DL-TFM used in the production of Qβ-F VLP.

The absence of a protein signal in the ^{19}F NMR spectrum of Qβ-F VLP at 25 ^{o}C is most likely due to the large size of the Qβ-F VLP (molecular weight, M_r = 2.55 MDa, hydrodynamic diameter, R_H = 30 nm; see Chapter 3) as will now be explained.



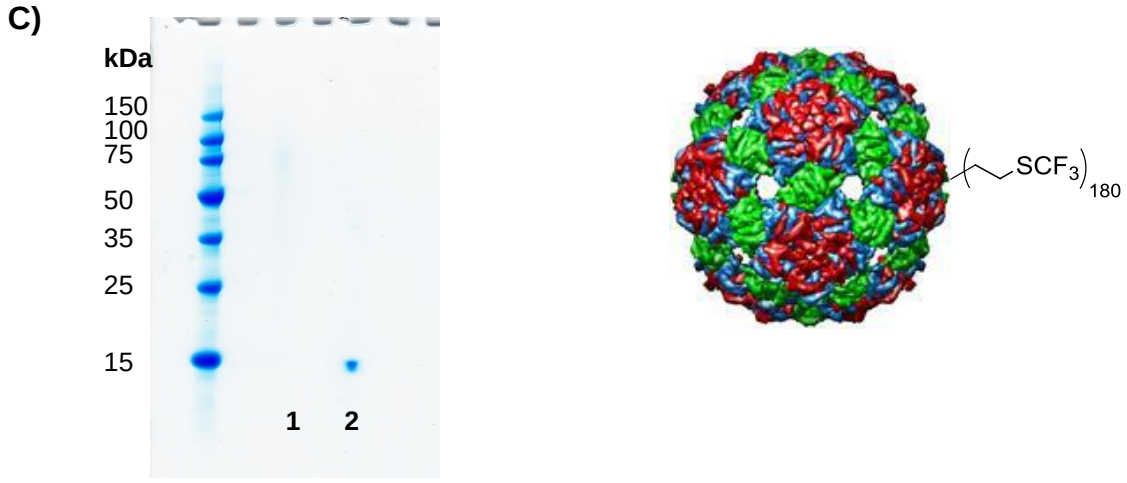


Figure 4.2 *^{19}F NMR of Qβ-F VLP (0.40 mg/mL) at A: 25 ^{o}C; and B: 60 ^{o}C. The spectra are referenced to 1,1,1-trifluoroacetone δ_F = -84.60 ppm. C: SDS-PAGE of the NMR sample (diluted six-fold, final VLP concentration 0.07 mg/mL). Lane 1: DNA native loading dye, no heat treatment; Lane 2: reducing loading dye, sample heated at 90 ^{o}C for 10 minutes. A 15-150 kDa protein marker is shown. 10% Bis-Tris gel (Invitrogen), run in MES buffer.*

The relationship between molecular tumbling, transverse relaxation (characterized by the T_2 time constant), and NMR peak broadening is explained as follows. The magnetic moments of neighbouring magnetic nuclei interact; termed dipolar coupling. As a result, the local magnetic field of a given nucleus (dominated by the applied magnetic field, B_0) will be altered by the magnetic fields produced by its neighbouring nuclei. For nucleus A at a point (r, θ) experiencing a magnetic field from the magnetic moment μ of neighbouring nucleus B, the z component of the magnetic field, is given by:

$$B_{\mu z} = \left(\frac{\mu_0}{4\pi}\right)\left(\frac{\mu}{r^3}\right)(3cos^2\theta - 1)$$

where μ_0 is the permeability of vacuum (4π x 10^{-7} Hm^{-1}), and r and θ are as defined in Fig. 4.3.[2]

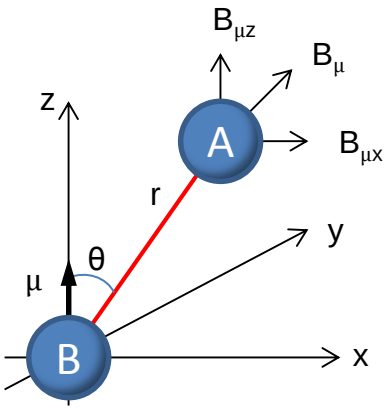


Figure 4.3 *The magnetic field from the magnetic moment μ of nucleus B, experienced by nucleus A positioned at a point (r, θ), with nucleus B at the origin.*

As such, as a molecule rotates in solution, r and θ are constantly varied, and the field experienced by a given nucleus will be modulated by this motion. Differences in local magnetic field cause differences in precessional frequency (since $\nu=\gamma B/2\pi$).[2] Large proteins tumble slowly in solution and so exhibit a broad range of precessional frequencies. The transverse magnetization is therefore lost quickly, hence T_2 is short. The broader range of precessional frequencies gives rise to a broader NMR peak.[3]

Short T_2 and consequent line broadening is a well-known NMR phenomenon, and is the principle barrier to NMR analysis of large molecular weight proteins.[4] Here, it seems likely that the ^{19}F CF_3 peak has broadened so extensively that it cannot be observed.

Following this reasoning, it was thought that increasing the temperature could result in the development of a peak due to the resulting increased molecular tumbling (and hence narrowing of the NMR peak). A ^{19}F NMR spectrum was taken at 60 ^oC which also gave no signal at the expected region for the SCF_3 group (Fig. 4.2 B). However, the development of a peak at -77.85 ppm, with a FWHH of 20.0 Hz, was observed. The development of the peak at -77.85 ppm on heating to 60 ^oC is likely due to a shift in the VLP/multimer equilibrium generating at least a small proportion of multimers of the Qβ-F protein in the sample which have a low enough molecular weight to be detected by NMR. The change in chemical shift from that expected for the SCF_3 group (δ_F ~ -39 ppm, see Fig. 4.1) to -77.85 ppm was proposed to be due to oxidation of the thiol ether of the trifluoromethionine residue to its sulfone ($S(O)_2CF_3$); the rate of which is likely to be enhanced at the higher temperature. Evidence to support this proposal can be found from a text book value of the chemical shift of $C\underline{F}_3SO_3CF_3$ quoted as -74 ppm.[5] Additionally, a sample of DL-TFM (10 mg) was incubated at room temperature with nitric acid (70%, double-distilled), a strong oxidising agent. The ^{19}F NMR spectrum of the resulting mixture is shown in Fig. 4.4 B, below the spectrum of DL-TFM in D_2O (Fig. 4.4 A). The spectra show that after incubation in HNO_3 the ^{19}F NMR peak shifts to ~ 76 ppm. Hence, the peak observed at -77.85 ppm on heating to Qβ-F VLP to 60 ^oC was taken to be from Qβ-F multimers with oxidised trifluoromethionine residues.

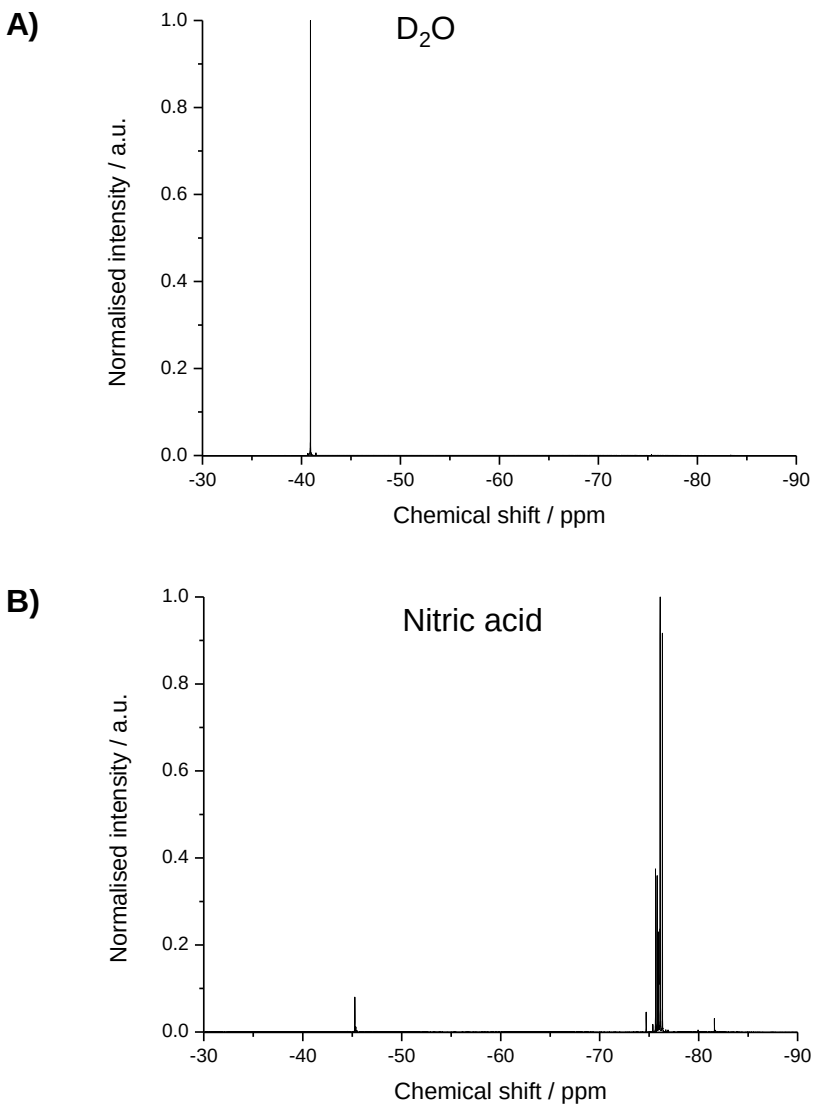


Figure 4.4 *^{19}F NMR spectra of A: DL-TFM in D_2O-TBS, B: DL-TFM after incubation in 70% nitric acid.*

To confirm the presence of protein in the NMR sample used to give the spectra shown in Fig. 4.2, the NMR sample was subjected to SDS-PAGE analysis, under both non-reducing and reducing conditions after NMR measurements were taken (Fig. 4.2 C). No band corresponding to the monomeric mass (~14 kDa) was observed under non-reducing conditions. However streaking can be observed around 75 kDa on the SDS-PAGE gel. The mass of the Qβ-F pentamer and hexamer is 70.9 kDa and 85.1 kDa, respectively. Hence, the streaking observed may be from multimers as a result of heating the Qβ-F VLP sample to 60 ^{o}C. This conclusion would be in agreement with the development of the broad signal

observed in the ^{19}F NMR spectrum at -77.85 ppm after heating the sample to 60 ^{o}C (Fig. 4.2 B). A clear band is seen at the monomer mass when the NMR sample was heated to 90 ^{o}C in the presence of a reducing agent.

4.3 ^{19}F NMR of disassembled Qβ-F

Since no ^{19}F NMR peak at δ_F ~ -39 ppm was obtained from the intact Qβ-F particle, a signal from a lower oligomeric state of the particle, and so a lower molecular weight and longer T_2, was sought. Takamatsu and Iso showed empirical evidence for the capsomeric structure of Qβ phage;[6] each particle comprising ~12 pentamers and ~ 20 hexamers of the coat protein. It was shown that neighbouring multimers, held together by hydrophobic forces,[7] could be separated by treatment of the denaturant SDS; and that monomers within an individual multimer, held together by disulphide bonds, could be separated by treatment of the reducing agent β-mercaptoethanol (0.43 M) and heating to 90 ^{o}C for 30 minutes. Takamatsu and Iso separated the hexamers, pentamers and monomers by gel electrophoresis.[6]

A ^{19}F NMR spectrum of the sample of Qβ-F VLP used in Fig. 4.2 (0.4 mg/mL, 0.03 mM monomers) was taken after 14 days storage at 4 ^{o}C and is shown in Fig. 4.5 A. The sample was then treated with SDS (final concentration 0.17 M) and a ^{19}F NMR spectrum was taken at temperatures 50 ^{o}C and 90 ^{o}C (Fig. 4.5 B and C). All spectra shown in Fig. 4.5 were taken using the same sample, so the protein and SDS concentrations are identical. The ^{19}F chemical shifts were calibrated using 1,1,1-trifluoroacetone as an internal standard set to δ_F -84.60 ppm. No signal was observed in the chemical shift range of the SCF_3 group (δ_F ~ -39 ppm, see Fig. 4.1) at 27 ^{o}C, which might be an indication that the Qβ-F VLP was still intact. (The signal at -77.83 ppm was of negligible intensity). At 50 ^{o}C a signal at -39.17 ppm was observed, showing that the particle had at least partially disassembled to a lower oligomeric state(s). Furthermore, a signal at -77.93 ppm was observed, presumably from the multimers

bearing oxidized trifluoromethionine residues (as discussed above). When heated to 90 ^{o}C, the peak at -39.17 ppm shifted to -39.32 ppm and appeared to have broadened slightly. It is tempting to take this peak broadening as evidence for further disassembly into two species, hexamers and pentamers, with different chemical shifts which are not fully resolved. However, firstly the differences in peak widths are within the errors for the level of digital resolution used here, and so such a conclusion cannot be confidently made. Secondly, at high temperatures, particularly at 90 ^{o}C, close to the solvent boiling point, shimming becomes difficult and peak broadening could likely be due to the inhomogeneity of the sample as a result of the dynamic physical changes occurring in the solution at this temperature. At 90 ^{o}C, the peak at -77.93 ppm shifted slightly, to -78.10 ppm, and, more significantly, had increased in intensity by a factor of 7.3, when compared to at 50 ^{o}C. From the data shown in Fig. 4.5 C it seems that the majority of the Qβ protein contains the oxidised trifluoromethionine residue.

A ^{19}F NMR spectrum was taken after leaving the sample at 4 ^{o}C for 20 days (Fig. 4.6). The ^{19}F NMR spectrum shows a peak -39.00 ppm, corresponding to the SCF_3 group. This shows that the Qβ-F protein had remained in the lower oligomeric state and had not reassembled to the VLP.

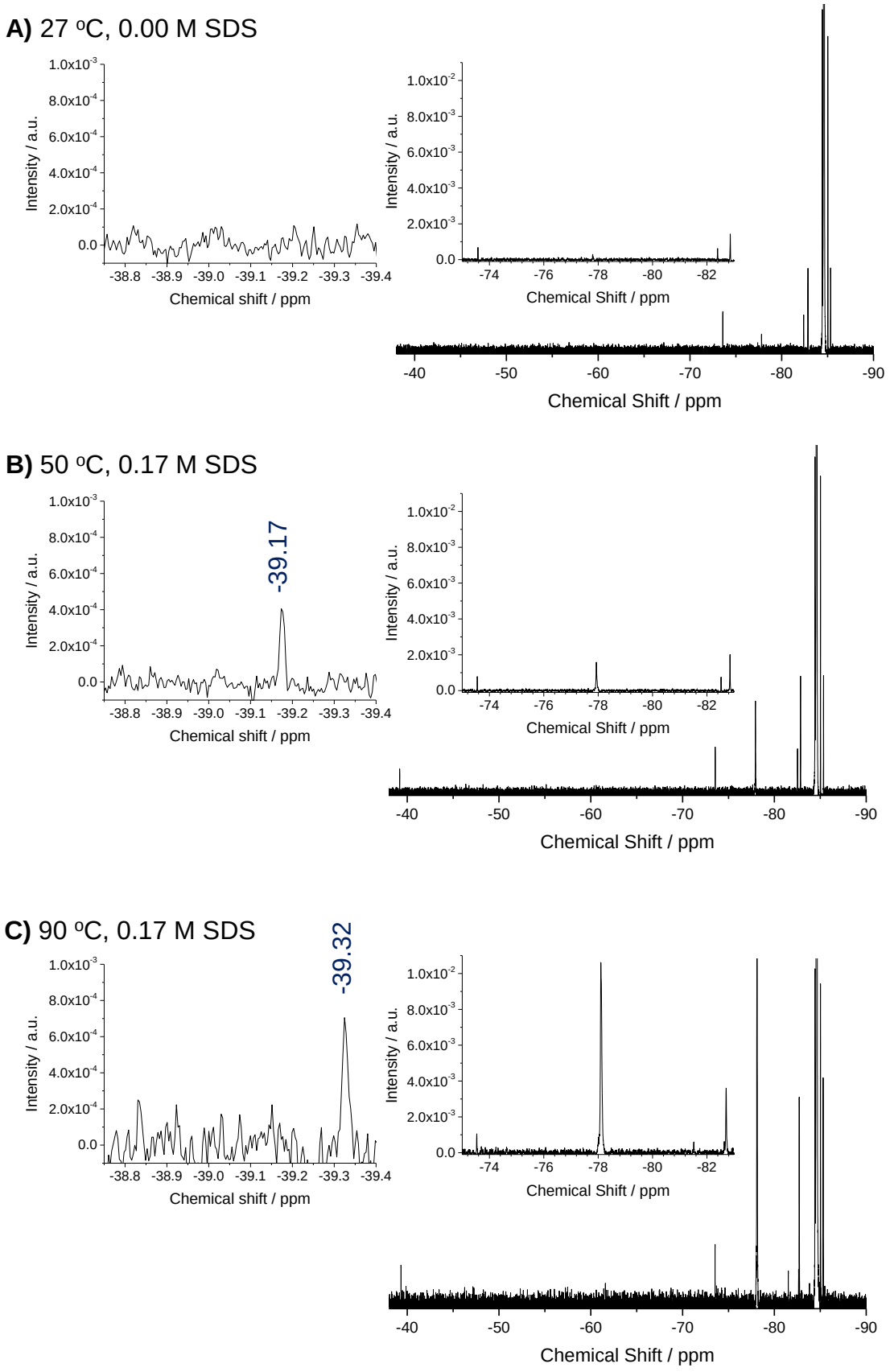


Figure 4.5 *^{19}F NMR of Qβ-F VLP (0.40 mg/mL) treated with 0.17 M SDS and heated to the temperatures as labelled. The spectra are referenced to 1,1,1-trifluoroacetone δ_F = -84.60 ppm.*

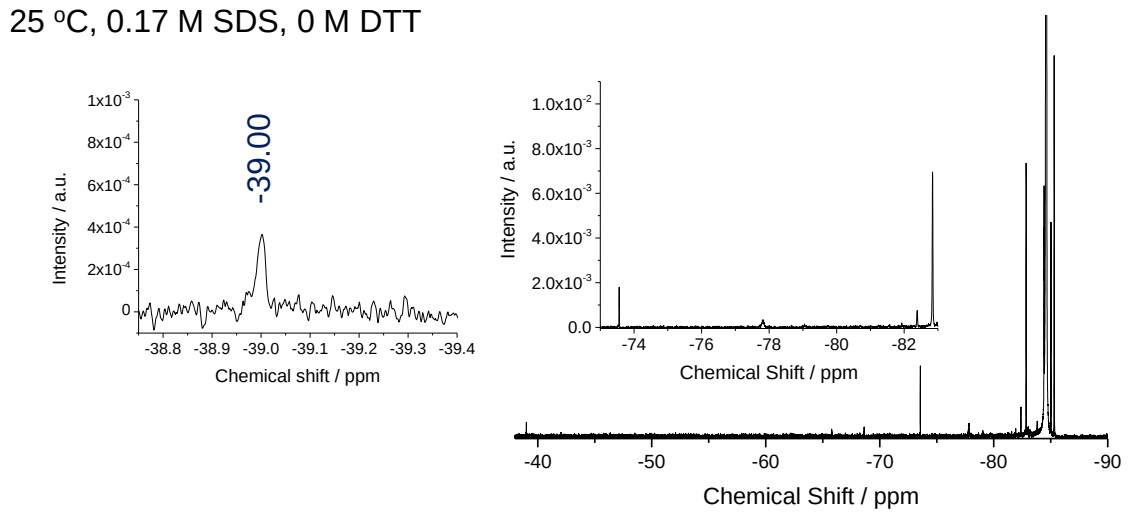


Figure 4.6 *^{19}F NMR of Qβ-F VLP (0.40 mg/mL) at 25 °C in 0.17 M SDS after storing at 4 °C for 20 days after the experiments shown in Fig. 4.5. The spectrum is referenced to 1,1,1-trifluoroacetone δ_F = -84.60 ppm.*

Since Takamatsu and Iso had shown that the pentamers and hexamers obtained after treatment of Qβ VLP with 0.17 M SDS could be reduced to monomers by addition of a reducing agent,[6] this procedure was next investigated to see if the change could be tracked by ^{19}F NMR. Dithiothreitol (DTT) was used as the reducing agent.[8] It was envisaged that if a spectrum was measured at increasing temperatures after addition of DTT, progression of the chemical reduction of the multimers present may be measured by differences in peak width (FWHH, full width at half height) and peak height. After recording a ^{19}F NMR spectrum at room temperature (Fig. 4.7 A), DTT was added to the sample (final concentration 0.82 M) and heated to 50 °C. This was the same sample as that used in Fig. 4.5. A spectrum was recorded at 50 °C before further heating to 90 °C and recording a spectrum at 90 °C (Fig. 4.7 C). Finally, the sample was left to stand at room temperature overnight and a spectrum was taken the next morning (Fig. 4.7 D).

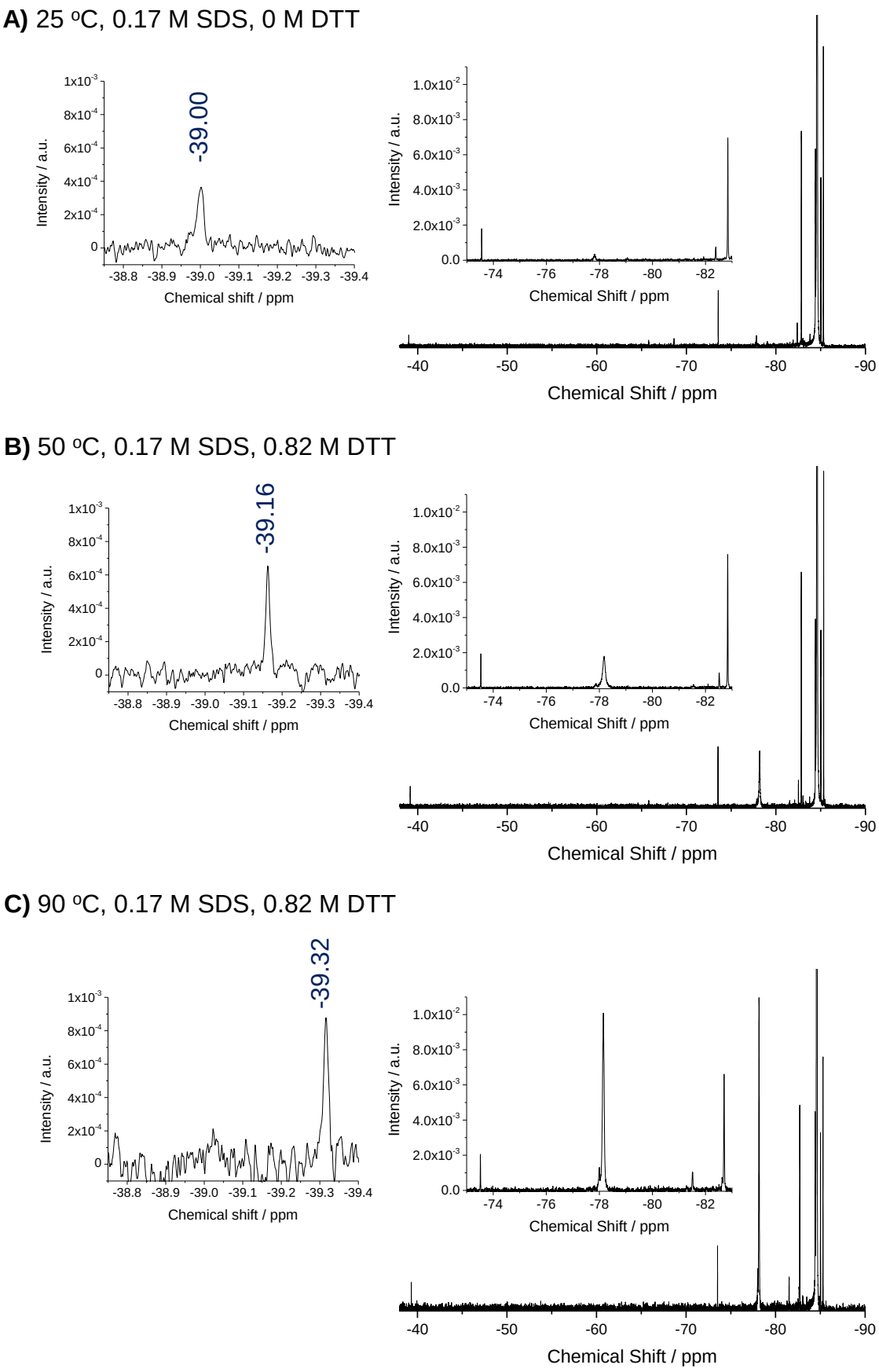


Figure 4.7 *^{19}F NMR of Qβ-F VLP(0.40 mg/mL) treated with 0.17 M SDS and DTT at concentrations as labelled and at the temperatures shown. All spectra are plotted with the same intensity scale. The spectra are referenced to 1,1,1-trifluoroacetone δ_F = -84.60 ppm.*

D) 25 ºC, 0.17 M SDS, 0.82 M DTT

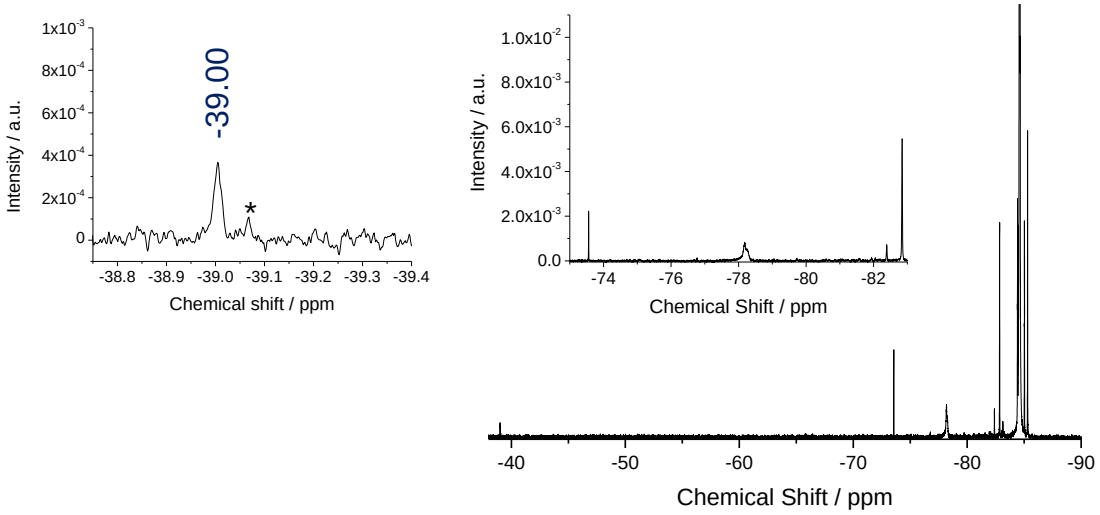


Figure 4.7 continued.

Comparing the spectra recorded at the same temperatures for Qβ-F VLP after treatment of SDS with those for Qβ-F VLP after treatment with SDS and DTT, the intensity of the peak at ~ -39 ppm increases as temperature increases form 25 ^{o}C, to 50 ^{o}C, to 90 ^{o}C. Furthermore, when comparing measurements at 25 ^{o}C and 50 ^{o}C, the FWHH is smaller for samples treated with DTT. This is clear evidence for a reduction of the initial multimeric Qβ-F species to a lower oligomer, most likely the monomer, after treatment with DTT. The data is summarised in Table 4.1. The slightly higher FWHH at 90 ^{o}C compared with at 50 ^{o}C is likely due to inhomogeneities in the solution at higher temperatures (see discussion above).

Treatment	Temperature	Normalized Intensity	FWHH
SDS	25 ^{o}C	3.6 x10^{-4}	9.5 Hz
SDS	50 ^{o}C	4.1 x 10^{-4}	6.1 Hz
SDS	90 ^{o}C	7.0 x 10^{-4}	6.9 Hz
SDS + DTT	25 ^{o}C	3.7 x 10^{-4}	8.6 Hz
SDS + DTT	50 ^{o}C	6.5 x 10^{-4}	5.7 Hz
SDS + DTT	90 ^{o}C	8.8 x 10^{-4}	7.0 Hz

Table 4.1 *^{19}F NMR data for the -39 ppm peak in the spectra shown in Fig. 4.5 and Fig. 4.7.*

Treatment	Temperature	Normalized Intensity	FWHH	Area
SDS	25 ^{o}C	3.0 x 10^{-4}	-	<0.00%
SDS	50 ^{o}C	1.6 x 10^{-3}	15 Hz	0.37%
SDS	90 ^{o}C	1.1 x 10^{-2}	22 Hz	2.38%
SDS + DTT	25 ^{o}C	4.0 x 10^{-4}	-	0.10%
SDS + DTT	50 ^{o}C	1.8 x 10^{-3}	55 Hz	1.06%
SDS + DTT	90 ^{o}C	1.1 x 10^{-2}	34 Hz	2.94%

Table 4.2 *^{19}F NMR data for the -78 ppm peak in the spectra shown in Fig. 4.5 and Fig. 4.7.*

The FWHH of DL-TFM was found to be 2.14 Hz (Fig. 4.1). This value is smaller than the FWHH values recorded for the disassembled Qβ-F protein samples (Fig. 4.5 – Fig. 4.7 and Table 4.1). This provides evidence that the ^{19}F signal observed in these protein NMR spectra does indeed originate from fluorine bound to the Qβ protein, as opposed to a false positive originating from residual DL-TFM in the sample solution from protein expression.

The ^{19}F NMR spectrum taken after allowing the sample and spectrometer bore to cool to 25 ^{o}C shows an additional peak at δ_F -39.08 (indicated with an asterisk in Fig. 4.7 D). I thought this might have been resolution of different oligomeric states. This lead me to repeat these experiments with a more concentrated protein sample to investigate this further. This will be discussed in Section 4.5.

4.4 A novel off-to-on ^{19}F NMR switch

A general comment in the light of the above results is that Qβ-F is NMR-silent in its intact VLP form, but on disassembly to lower constituent oligomers, presumably pentamers, hexamers and monomers, Qβ-F exhibits two distinct ^{19}F NMR signals. This opens up a potential use for this imaging construct as an off-to-on imaging probe, only giving a ^{19}F signal when it is activated by disassembly (Fig. 4.8). Furthermore, if the disassembly could be triggered by a specific biological event, this probe would be NMR-silent *in vivo* until it reaches a particular region or target of interest where activation would cause a strong ^{19}F signal to develop. Hence, instead of relying on accumulation of the imaging agent at the biological target to achieve image contrast (e.g. as in the enhanced permeability effect[9]), with "activatable Qβ-F", a ^{19}F NMR signal would only originate from Qβ-F at the site of interest.

Hamachi *et. al.* reported the first example of an off/on ^{19}F probe by means of disaggregation of a supramolecular organic nanoparticle.[10] (Note, the probe of Hamachi is termed off/on since the disassembly process was shown to be reversible). The monomer unit comprised a 3,5-bis(trifluoromethyl)benzene derivative attached to a benzenesulfonamide moiety, a ligand for human carbonic anhydrase I (hCAI), *via* a linker. A solution of the monomer, shown to be present as an aggregate of diameter 200 nm by TEM, gave no ^{19}F NMR signal. Addition of hCAI and concomitant ligand binding resulted in a strong signal at -62.2 ppm, suggesting disassembly of the aggregate.[10] Gao *et. al.* have also exploited this

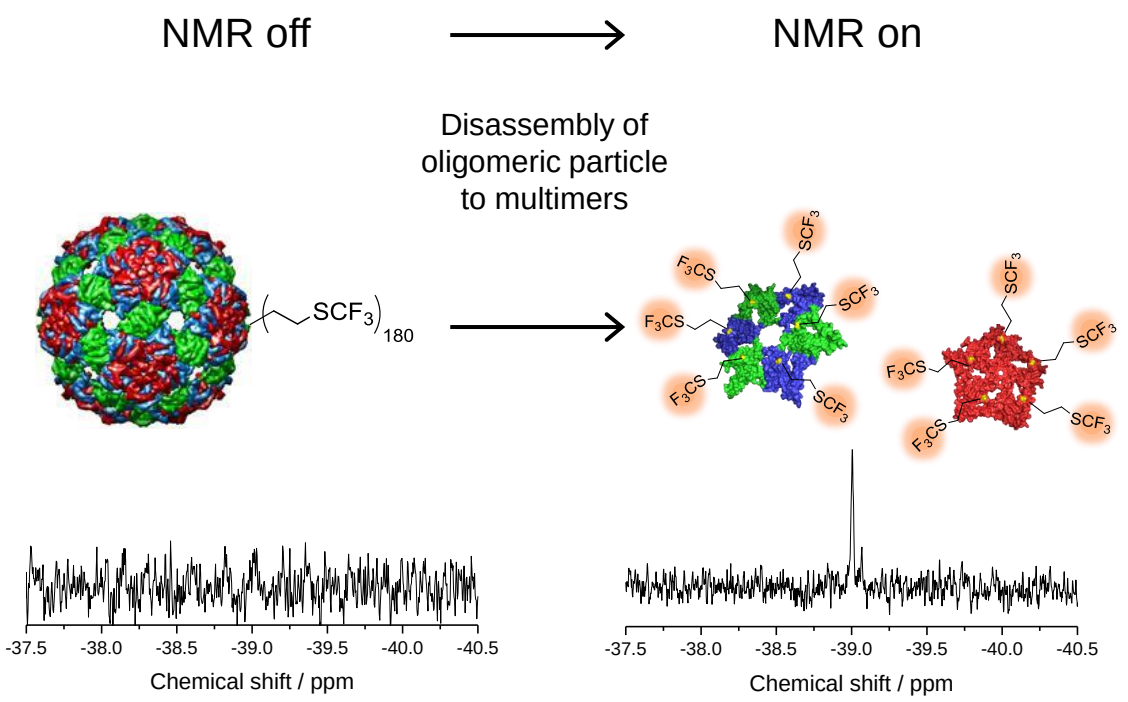


Figure 4.8 *Potential use of Qβ-F as an off-to-on imaging switch when triggered.*

off/on principle using the disassembly of a micelle of copolymers which occurred on lowering the pH of the solution.[11] Thus, the micelle can be used as a pH detector.

An advantage of the current approach is the unusually high payload of ^{19}F nuclei per imaging molecule achievable with Qβ-F VLP. Furthermore, the surface of the VLP provides an extremely versatile platform for loading of targeting ligands and fluorophores *via* chemical ligation of lysines, cysteines and/or incorporation of another unnatural amino acid presenting a chemical handle (see Chapter 1 section 1.3.1).

The application of Qβ-F as a probe for medical imaging is further investigated and discussed in Chapter 5.

4.5 Probing the structure of Qβ phage using ^{19}F NMR

Significant research efforts have focused on virus disassembly, the mechanism of which is still largely unknown. Many antiviral drugs are based upon inhibition of the replication process of viruses.[12–14] Hence, a more detailed understanding of the disassembly process would guide the design of new antiviral drugs. Furthermore, viral nanoparticles are

being used as gene-delivery vectors[15] and "Trojan Horses" to get drugs across the blood-brain barrier.[16]

I envisaged that the Qβ-F VLP could be used as a model for virus disassembly by exploiting the ^{19}F label to track the disassembly process of bacteriophage Qβ. Imaging of virus disassembly and internalization to date has predominantly used fluorescence microscopy[17–19] and transmission electron microscopy (TEM).[13,20–23] Although these techniques give data on the translational position and motion of viruses relative to cells, they do not provide information on the molecular changes of the virus during the binding and infecting process. Using NMR to follow a virus with an NMR-active label could allow a "virus eye-view" of the disassembly process.

Recently, ^{19}F NMR has been used to probe the dynamics of protein unfolding,[24–27] driven by the link between misfolded proteins and a number of human diseases, such as Alzheimer's.[28] Prosser *et. al.* successfully characterized three distinct oligomeric states of mammalian prion protein (PrP) by differences in ^{19}F chemical shift and FWHH.[29] Furthermore, Marsh *et. al.* used fluorine labelling to follow the time course of aggregation of $A\beta_{1\text{-}40}$ (a peptide derived from cleavage of the amyloid precursor protein) *via* chemical shift differences in the oligomeric states.[30]

Encouraged by these findings, it was envisaged that if disassembly experiments could be conducted on the fluorinated Qβ developed in this work whilst measuring ^{19}F NMR, the δ_F of the multimers and monomer could be measured. This would provide a chemical shift signature for the disassembly components of Qβ.

The ^{19}F NMR of Qβ-F VLP treated with SDS shows, i.e. by observation of the development of signals, that the particle disassembles on addition of SDS (Fig. 4.5). However, the NMR spectra were recorded at three temperatures and so only give a "snap-shot" of disassembly at these points (Fig. 4.5). It was envisaged that monitoring the

disassembly over a time course would be more informative of the mechanism of the disassembly and so a time course of 20-minute acquisitions taken each hour overnight was planned. Following previous ^{19}F NMR results in this chapter and the mentioned uncertainties in line broadening at high temperatures, it was decided that NMR spectra would not be measured at a temperature greater than 37 ^{o}C. Hence, I planned to incubate the sample overnight at 37 ^{o}C. This also provides a more biologically relevant model, since 37 ^{o}C is the normal human body temperature and incubation temperature for *E. coli*.

SDS (final concentration 0.17 M) was added to a sample of Qβ-F VLP (1.26 mg/mL, 0.09 mM with respect to monomer concentration). However, within the time required to take an initial ^{19}F NMR spectrum (~5 minutes after addition of SDS) a signal was observed at ~ -39 ppm (Fig. 4.9 B). A peak at -78 ppm was also observed, albeit at much lower intensity. Another spectrum was measured after a 20-minute period (incubation at 37 ^{o}C in NMR bore) but no change in the peak intensity or width was observed, suggesting that an endpoint had been reached (data not shown). As such, an overnight time course was not performed. A spectrum was then recorded at 25 ^{o}C to ensure a signal could also be detected at ambient temperature (Fig. 4.9 C). At 25 ^{o}C a signal was observed at -39 ppm, however no signal was observed at -78 ppm. It is interesting to note that when compared to the spectra of Fig. 4.5 B and C, the intensity of the peak at -78 ppm is significantly weaker. Since the difference between the conditions in Fig. 4.5 and Fig. 4.9 is the increased temperature in the former, it seems that the development of the peak at -78 ppm is induced by temperature.

To investigate if the particle had reassembled, another spectrum of the sample of Fig. 4.9 was also measured after overnight standing at room temperature (Fig. 4.9 D). A peak, of approximately the same peak height as in Fig. 4.9 C, was observed, showing that the protein remained in its multimeric state. Note that a more concentrated protein sample was used here than in previous experiments (Fig. 4.5 and Fig. 4.7), giving a higher signal-to-noise ratio.

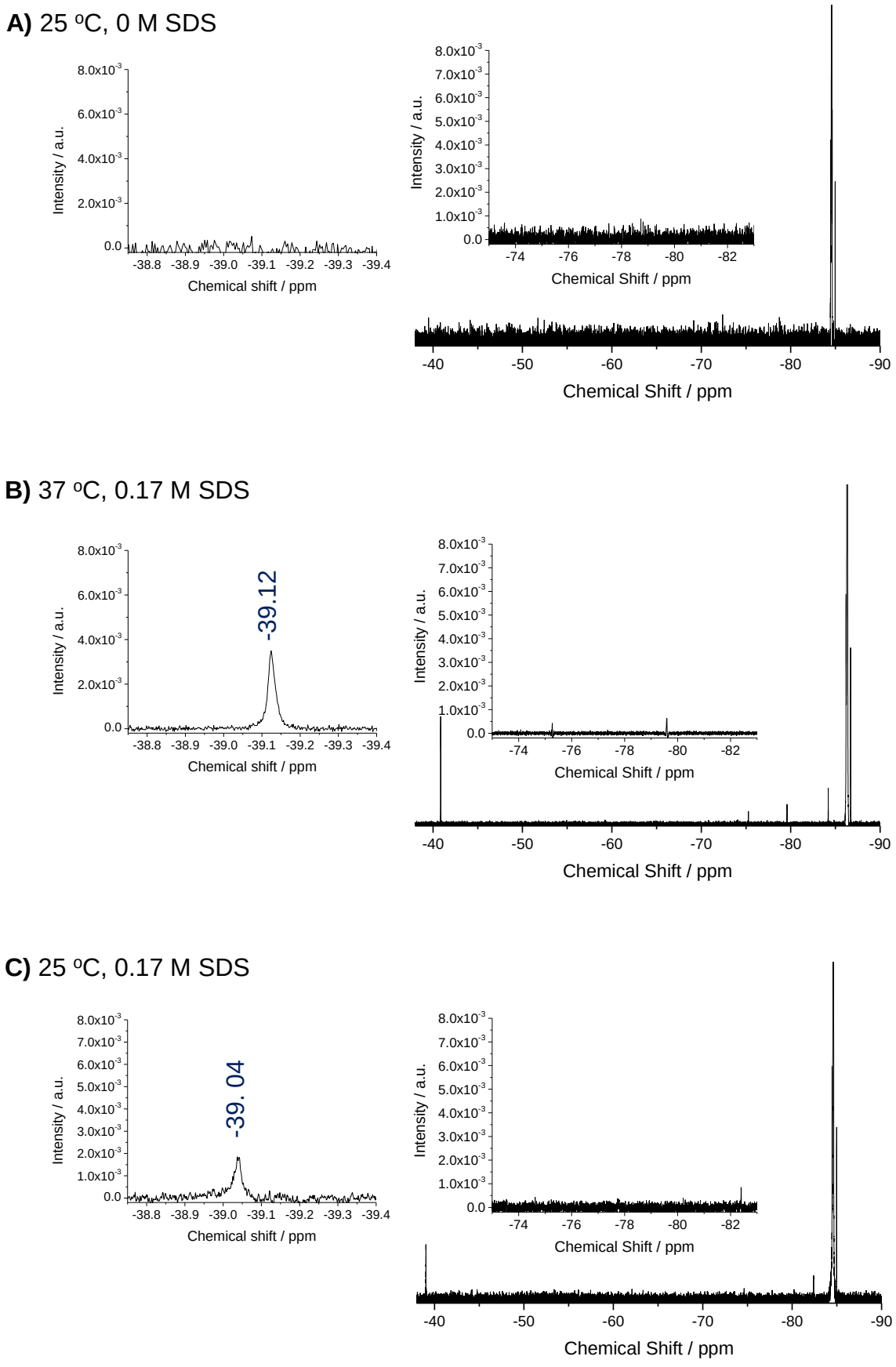


Figure 4.9 *^{19}F NMR of Qβ-F VLP (1.26 mg/mL) in 0.17 M SDS at the temperatures shown. The spectra are referenced to 1,1,1-trifluoroacetone δ_F = -84.60 ppm.*

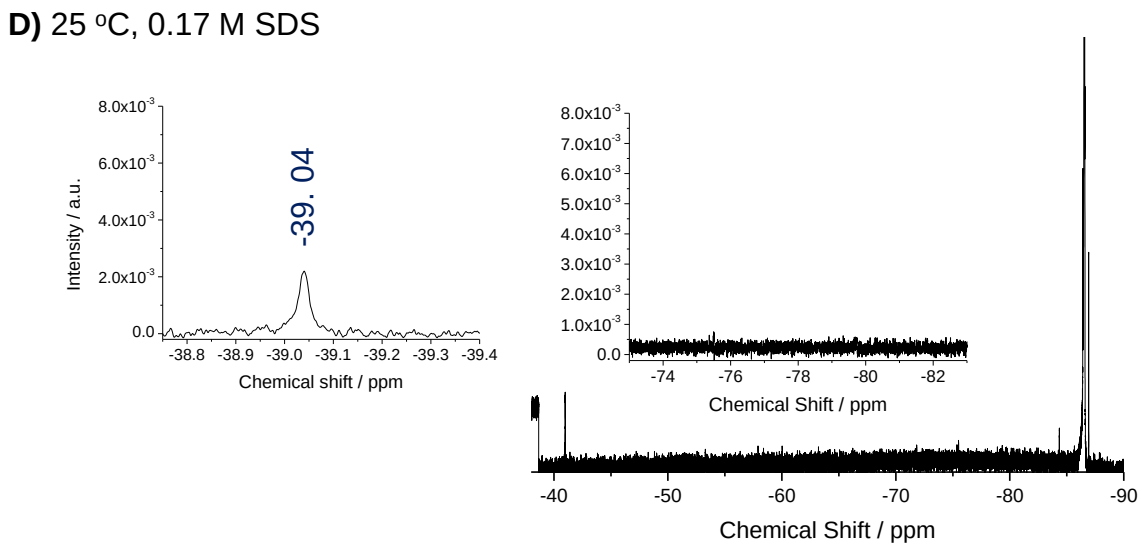


Figure 4.9 continued.

Diffusion ordered NMR spectroscopy (DOSY) was next employed as a more reliable method for comparing the sizes of the multimers by NMR than by using peak FWHH. DOSY uses the pulsed-gradient spin-echo method to differentiate species according to their translational diffusion coefficient.[31–33] An initial 90^o pulse tips the magnetization vector from the *z* axis onto the *xy* plane and the spin vectors begin to dephase (transverse magnetization decreases). After a time τ a second pulse rotates the vectors 180^o and the spins begin to rephrase along the *x* direction where an echo signal is detected (Fig. 4.10). In order for the spins to *completely* rephrase so that the echo signal is of the same intensity as the initial signal at time zero, the precession frequencies must remain constant throughout the experiment. However, since a gradient field is applied, the local magnetic field, and therefore the precession frequency of the nuclei depends on their position along the field gradient lines. The greater the diffusion rate of a molecule, the greater the difference in the local magnetic field experienced by, and therefore frequencies of, the nuclei before and after the 180^o pulse. This results in greater dephasing of the spins and ultimately a diminution of the signal.

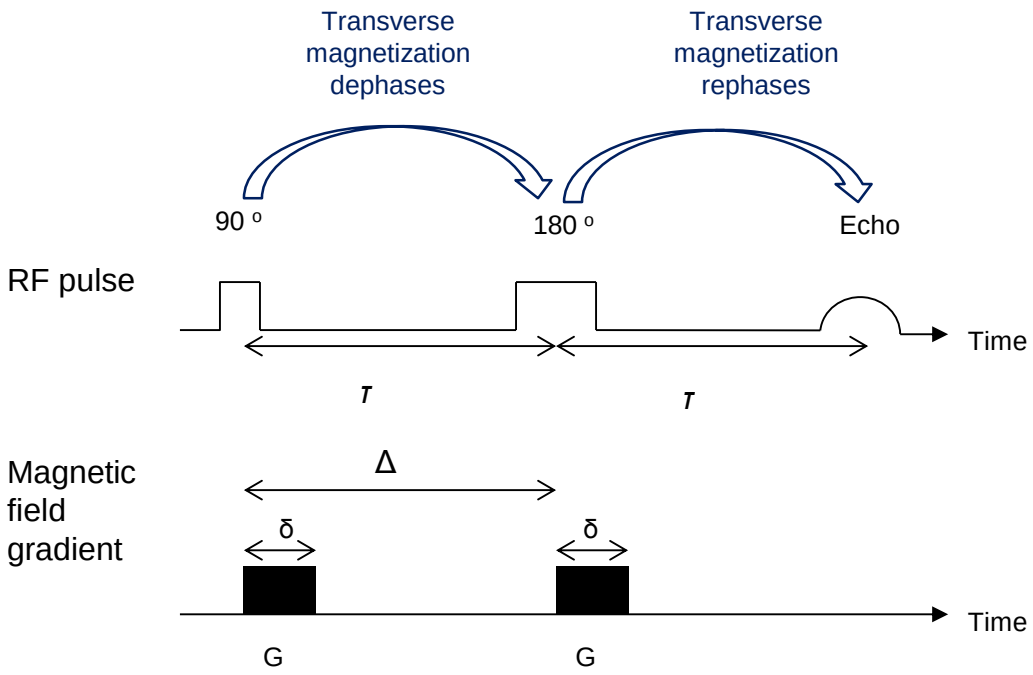


Figure 4.10 *Sequence for the pulse-gradient spin-echo experiment used in DOSY. (Adapted from Fig. 9.1 of reference [3]).*

By measuring the intensity of the echo, I, of an initial signal of intensity I_0, at different magnetic field strengths, G, the diffusion coefficient, D can be graphically determined:

$$\frac{I}{I_0} = -exp\left[D\gamma^2\delta^2G^2\left(\Delta - \frac{\delta}{3} - \frac{\tau}{2}\right)\right]$$

where γ is the magnetogyric ratio of the observed nucleus and δ and τ are time durations as defined in Fig. 4.10.[31] The diffusion coefficient, D, is related to the hydrodynamic radius of a sphere, r_s, by the Stokes-Einstein equation:

$$D = \frac{k_BT}{6\pi\eta r_s}$$

where k_B is the Boltzmann constant, T is the absolute temperature, and η is the solvent viscosity.[34]

^{19}F DOSY was measured for the peak obtained from treatment of Qβ-F VLP with SDS (Fig. 4.11 A). Calculation of the diffusion coefficient D gave a value of 2.1 x 10^{-11} m^2s^{-1}. To compare this diffusion coefficient with that of the monomer, a sample of the monomer was synthesized and purified by reducing Qβ-F VLP with DTT and subsequently reacting it

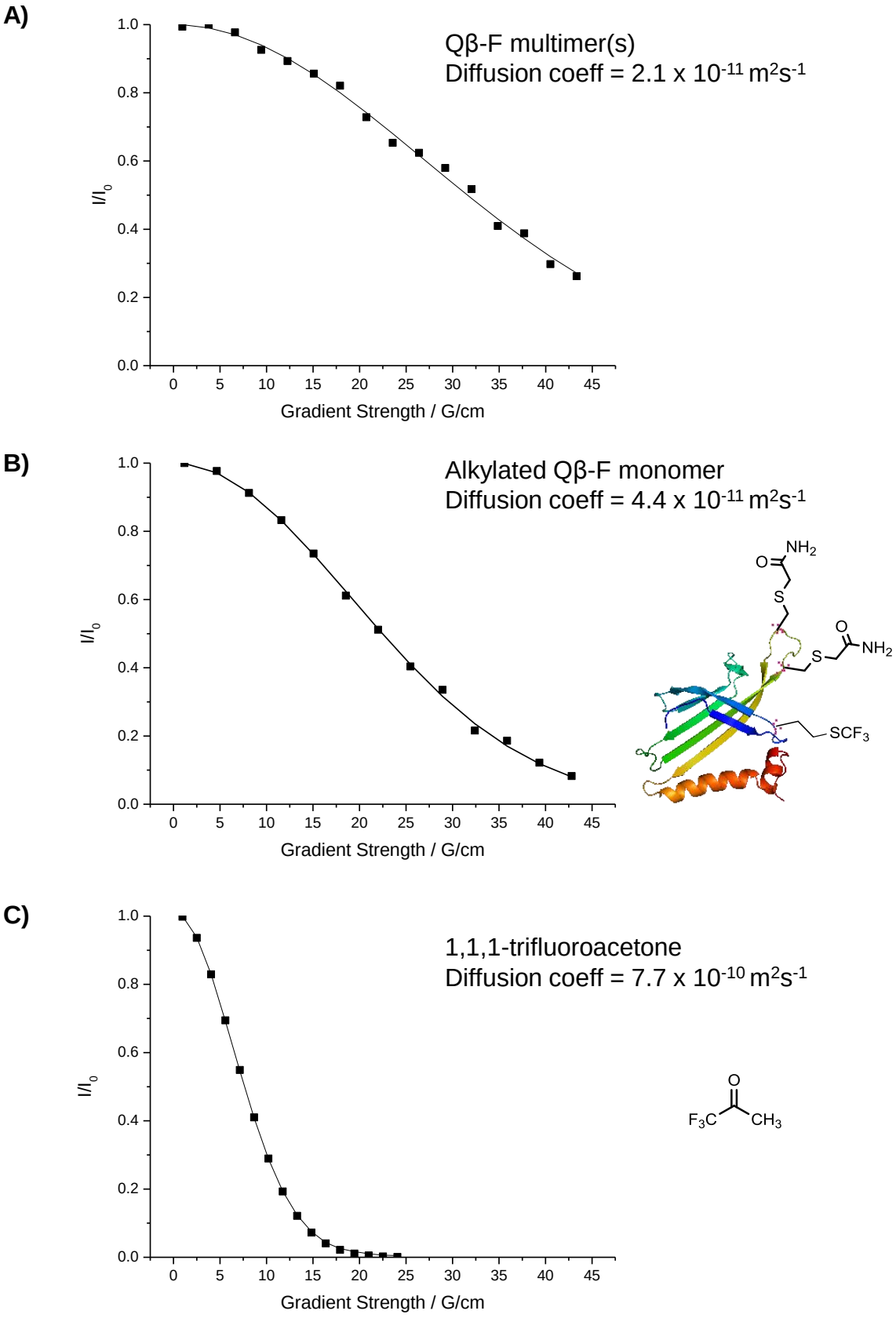


Figure 4.11 *^{19}F DOSY plots for A) Qβ-F multimer, B) Qβ-F alkylated monomer, C) 1,1,1-trifluoroacetone reference.For each experiment, Δ = 199.9 ms and δ = 5.0 ms.*

with iodoacetamide to alkylate the cysteines. This ensures that the protein remains in the monomeric state since assembly to the higher order oligomers requires formation of disulphides. Fig. 4.12 shows an SDS-PAGE gel of Qβ-Met VLP without (lane 1) and with (lane 2) reducing treatment, and Qβ-F alkylated monomer without (lane 3) and with (lane 4) reducing treatment. For the Qβ-Met in its VLP form, the gel shows the protein is not present as monomers when not treated with reducing loading buffer (lane 1), however is clearly present as monomers when treated with reducing agent and heating (lane 2). For the Qβ-F alkylated monomer, there is a distinct band for the Qβ monomer under both native conditions and reducing conditions (lanes 3 and 4, respectively), confirming that the sample consists of isolated monomers and not assembled particles.

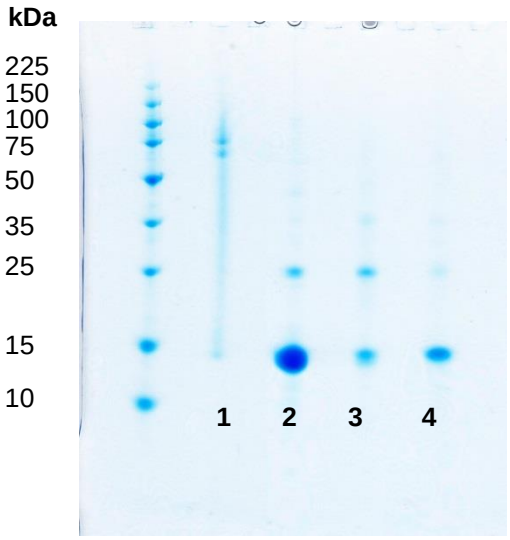


Figure 4.12 *SDS-PAGE: Lane 1: Qβ-Met VLP (0.04 mg/mL) after non-reducing treatment; Lane 2: Qβ-Met VLP (0.04 mg/mL) after treatment with reducingdye and heating at 90 ^{o}C for 10 minutes; Lane 3: Qβ-F alkylated monomer (0.18 mg/mL) after non-reducing treatment; Lane 4: Qβ-F alkylated monomer(0.18 mg/mL) after treatment with reducingdye and heating at 90 ^{o}C for 10 minutes. A 10-225 kDa protein marker is shown. 10% Bis-Tris gel (Invitrogen), run in MES buffer. The gel was stained with Instant*Blue ™ *(coomassie blue solution).*

SDS was added to the monomer sample at the same concentration as that previously used (i.e. a final concentration of 0.17 M) since a change in concentration is likely to change the solvent viscosity and hence the translational diffusion of the monomer. For the same reason the DOSY experiment was taken at the same temperature as that of the multimer. The

diffusion coefficient calculated from the signal of the monomer was 4.4 x 10^{-11} m^2s^{-1} (Fig. 4.11 B), approximately twice that as for the species generated on addition of SDS to Qβ-F VLP. From the Stokes-Einstein equation (above), it follows that the hydrodynamic radius, r_s, is greater for the species generated on addition of SDS to Qβ-F VLP, than for the monomer. This clearly shows that the species generated on addition of SDS to Qβ-F VLP is a multimeric state i.e. of higher molecular weight than the monomer. Finally, as a small molecule reference, the diffusion constant *D* of the reference compound 1,1,1-trifluoroacetone was measured and calculated to be 7.7 x 10^{-11} m^2s^{-1} (Fig. 4.11 C).

These multimeric states (Fig. 4.11 A) give a characteristic broad peak at -39.04 ppm, a higher chemical shift than that of the monomer, -39.10 ppm (Fig 4.13). Such information would be extremely useful to characterize the state of the Qβ bacteriophage during its infection cycle. This would allow *in vivo* virus disassembly mechanisms to be studied in much greater detail than has currently been possible.

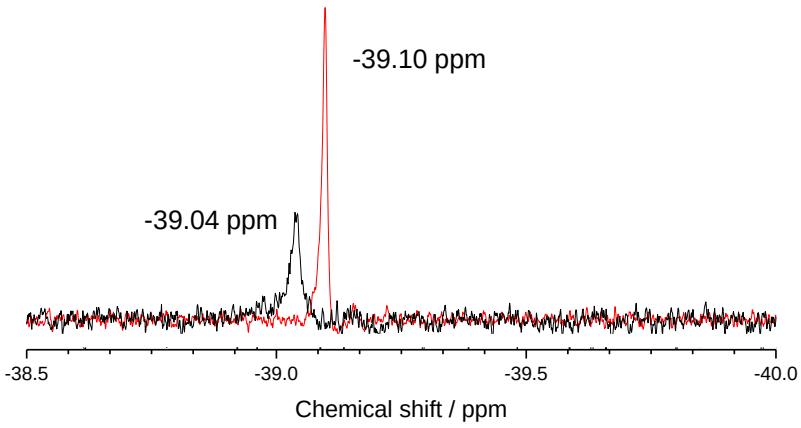


Figure 4.13 *^{19}F NMR Qβ-F VLP (1.26 mg/mL) treated with 0.17 M SDS (black) and Qβ-F alkylated monomer(1.07 mg/mL) (red). The spectra are referenced to 1,1,1-trifluoroacetone δ_F = -84.60 ppm.*

Having analysed the disassembly of the Qβ-F VLP to a lower multimeric state by ^{19}F NMR and ^{19}F DOSY NMR, the process of disassembly from this multimeric state to the monomeric state of Qβ-F was next investigated. As discussed above, the disassembly of

the multimer to monomers can be brought about by the addition of a reducing agent.[6] The sample used previously (after addition of SDS, Fig. 4.9) was used so that direct comparisons could be made between spectra before and after DTT treatment. DTT (100 equivalents with respect to Qβ-F disulfide bonds, 18 mM final concentration) was added to the sample before heating at 90 ^{o}C for 40 minutes. Fig. 4.14 A shows the resulting ^{19}F NMR spectrum. A prominent, additional peak at -39.07 ppm was observed. This was promising as it indicated the presence of a species distinct from the multimer; the multimer peak still being observable at -39.04 ppm. The peaks were fitted using ACD software and the FWHH of the peaks at -39.04 ppm and -39.07 ppm were 10.3 Hz and 7.0 Hz, respectively. Unfortunately, the peaks were not sufficiently resolved or high enough in intensity to measure the diffusion constant by DOSY. The peak at -39.04 ppm is taken to be the species observed on addition of 0.17 M SDS to Qβ-F VLP (see Fig. 4.9). Since the additional peak at -39.07 ppm resulted from heat treatment with DTT, and that the FWHH is smaller than that observed for the multimer peak, it seems reasonable to suspect that the peak at -39.07 ppm is from a reduced form of the multimer, that is, a lower oligomer. However, whether the peak at -39.07 ppm is from a monomer is not clear. In an effort to further extend the reduction of the multimer, and hence determine whether a separate peak corresponding to the monomer would be observed, the concentration of DTT was doubled to 36 mM and the sample was heated at 90 ^{o}C for 40 minutes. The resulting spectrum is shown in Fig. 4.14 B. Peaks corresponding to peak I and II in spectrum Fig. 4.14 A were observed at -39.04 and -39.06 with FWHH values of 10.4 and 7.1 Hz, respectively. The peak at -39.06 had increased in intensity from 4.7 x 10^{-3} to 6.0 x 10^{-3}, suggesting further reduction of the multimer to this species. In addition, a new peak at -39.02 was observed with a FWHH value of 10.3 Hz. The spectrum after overnight incubation at room temperature was unchanged (Fig. 4.15).

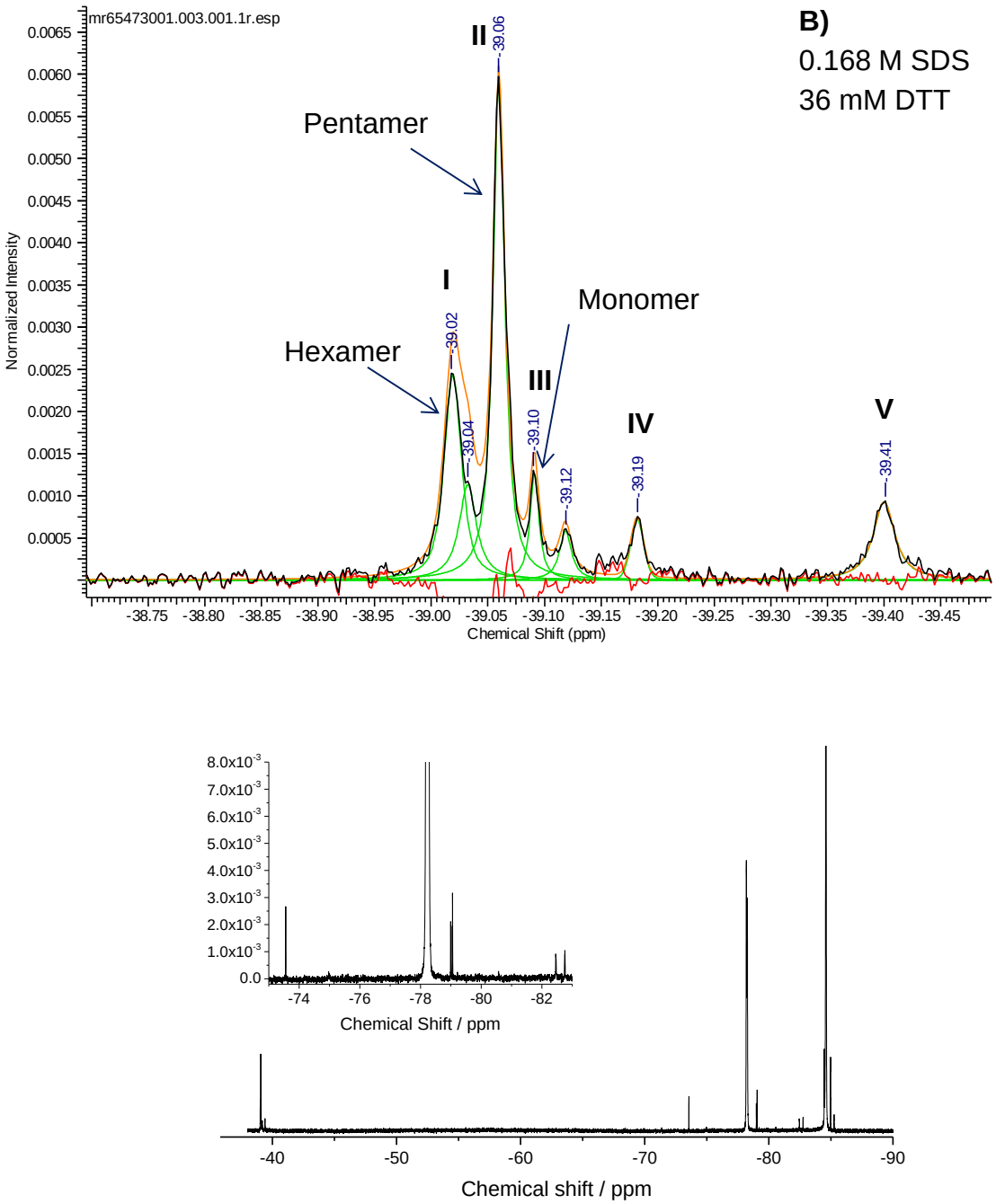


Figure 4.14 Continued.

25 °C
0.168 M SDS
36 mM DTT o/n

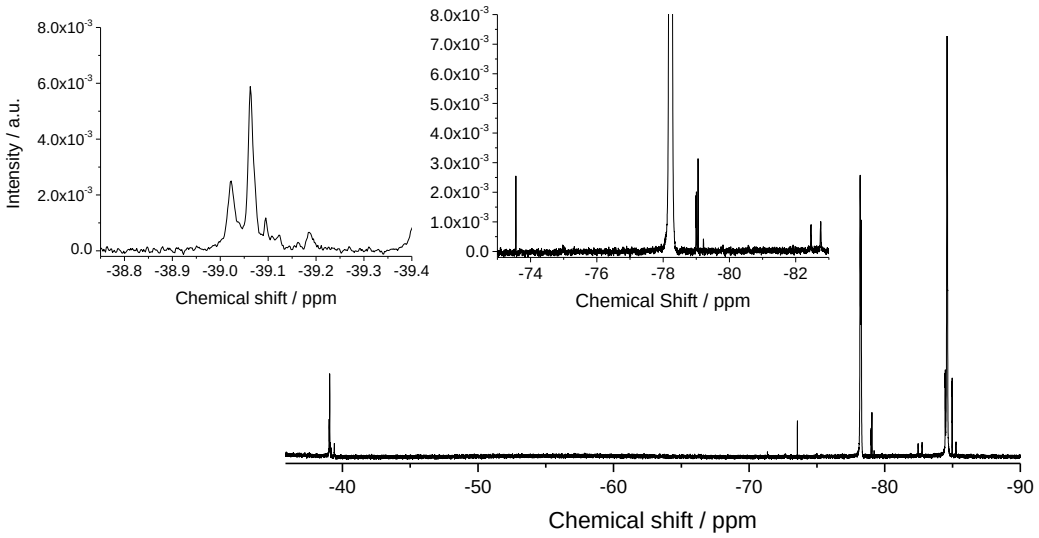


Figure 4.15 *^{19}F NMR spectrum of Qβ-F VLP (1.26 mg/mL) treated with 0.17 M SDS and DTT after incubation overnight at room temperature. The spectrum is referenced to 1,1,1-trifluoroacetone δ_F = -84.60 ppm.*

Finally, the sample was treated further with DTT as follows. The final concentration of DTT was raised to 102 mM and then heated at 90 °C for 165 minutes. The final concentration of DTT was then raised to 169 mM and the sample was heated at 90 °C for a further 60 minutes. The ^{19}F NMR spectrum of the resulting sample is shown in Fig. 4.16.

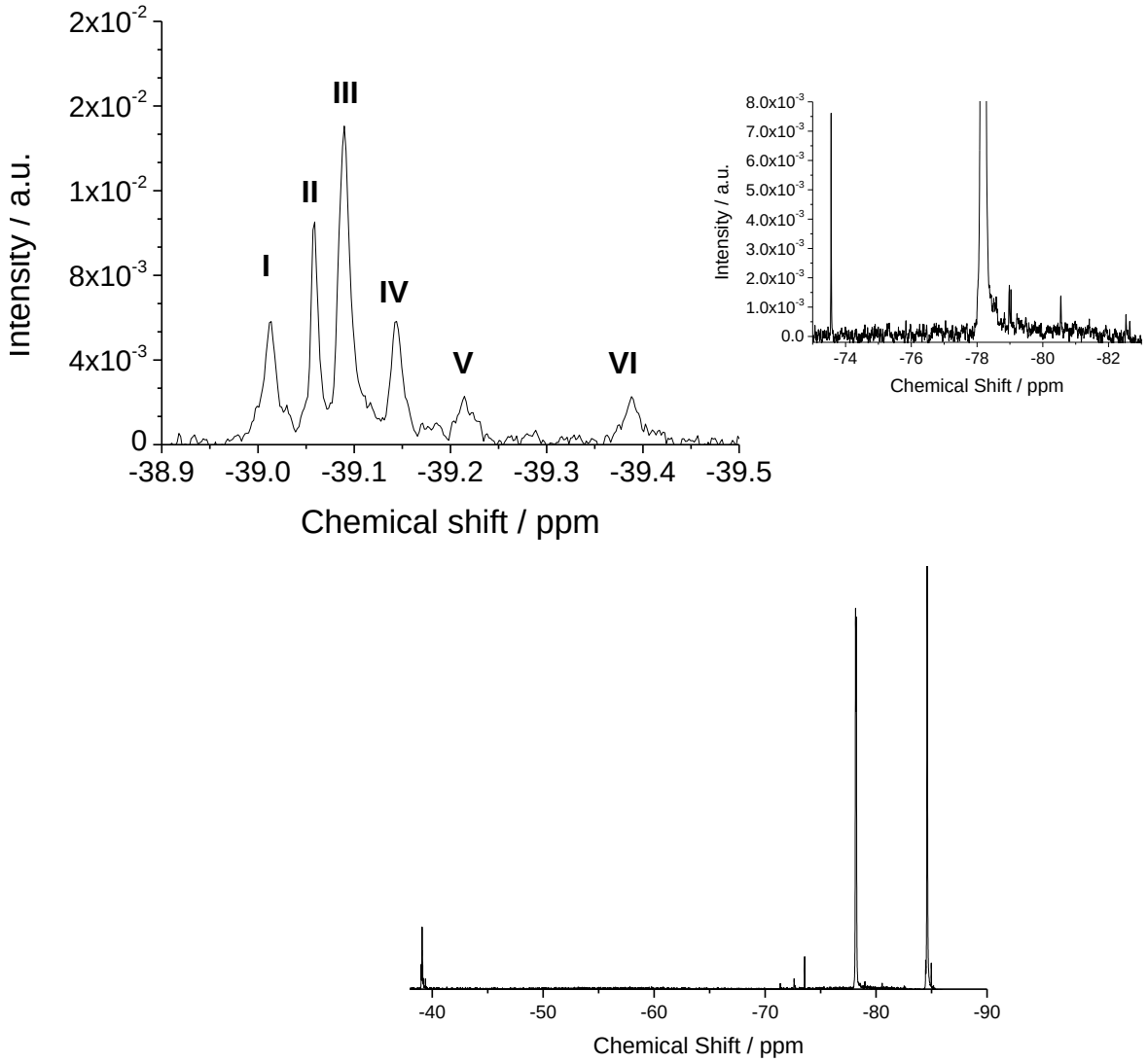


Figure 4.16 *^{19}F NMR of Qβ-F VLP (1.26 mg/mL) after further addition of DTT and heating at 90 ^{o}C for 225 minutes. The spectrum is referenced to 1,1,1-trifluoroacetone δ_F = -84.60 ppm.*

A ^{19}F NMR spectrum of purified Qβ-F capped monomer was taken, under same concentration of SDS as a chemical shift standard. Fig. 4.17 shows the different stages of disassembly from treatment with SDS to successive additions of DTT. The ^{19}F NMR spectra are aligned with that of the alkylated monomer (spectrum in red) and hence allows characterization of this peak in the disassembled samples.

The ^{19}F NMR spectrum A in Fig. 4.17 is of the intact Qβ-F VLP (0.09 mM). Spectrum B was taken after addition of SDS to a final concentration of 0.17 M and incubation at room temperature incubation for 5 minutes (taken from Fig. 4.9 C). The peak

observed at -39.04 ppm was attributed to a multimeric state (see above). Spectrum C is that shown in Fig. 4.14 A. Spectrum D is that shown in Fig. 4.14 B. Spectrum E is that shown in Fig. 4.16. Fig. 4.17 shows that the species present after heat treatment in 18 mM and 36 mM DTT at -39.06/-39.07 (shown in Fig. 4.14) was not in fact the monomer. Peak III of spectrum Fig. 4.16 is seen to have same chemical shift as the alkylated monomer, and required additional DTT and heating compared to the spectra Fig. 4.17 C and D. Based upon the FWHH values, the peaks labelled I and II in Fig. 4.14 can be tentatively assigned as hexamer and pentamer, respectively. However, further experiments are required to unequivocally characterize these peaks in the spectra. These experiments will be discussed at the end of this chapter.

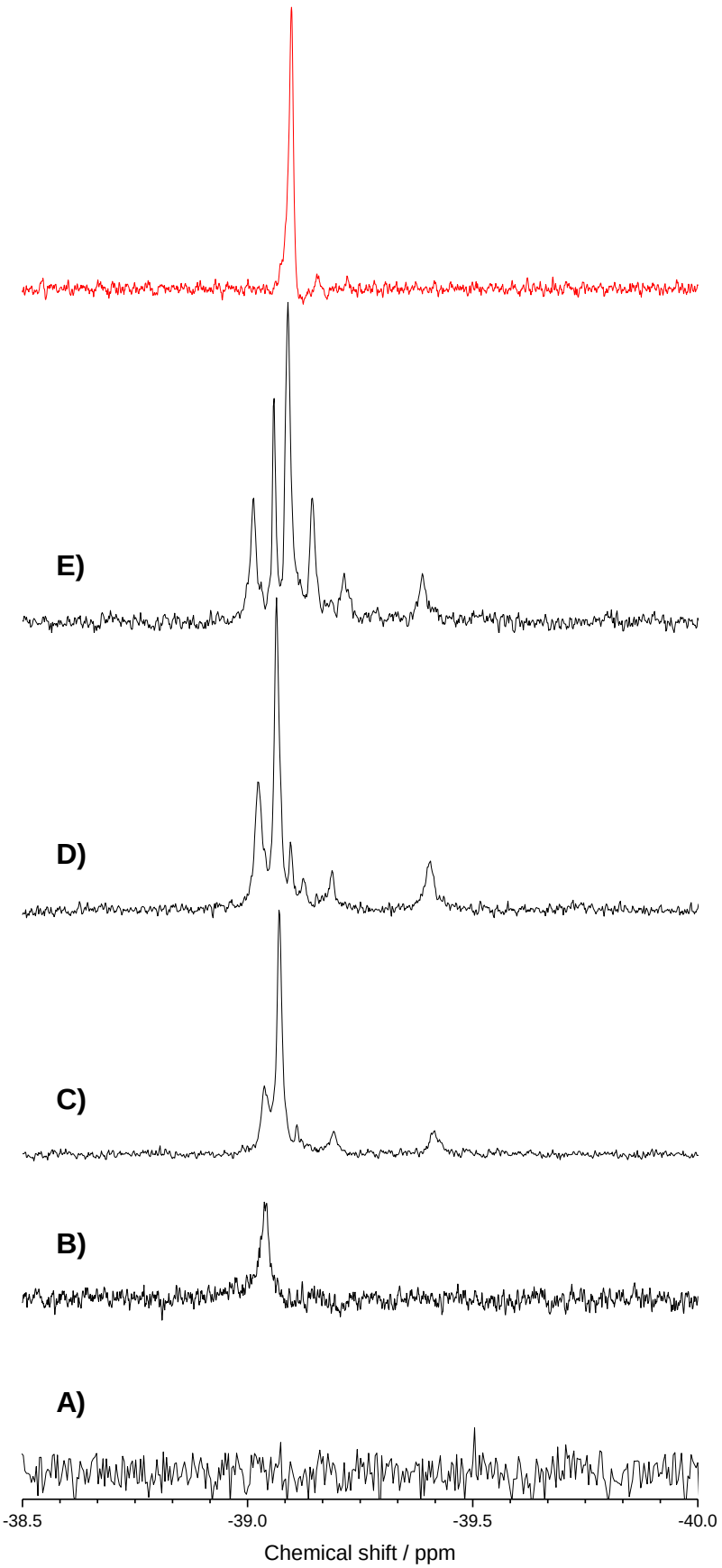


Figure 4.17 *^{19}F NMR of Qβ-F VLP (1.26 mg/mL) treated with: A: 0 M SDS; B: 0.17 M SDS; C: 0.17 M SDS, 18 mM DTT then 90 °C for 40 minutes; D: 0.17 M SDS, 36 mM DTT then 90 °C for 40 minutes; E: 0.17 M SDS, 102 mM DTT at 25 °C, then 90 °C for 165 minutes, then 169 mM DTT heated to 90 °C for 60 minutes. Red spectrum: Qβ-F alkylated monomer (1.07 mg/mL). The spectra are referenced to 1,1,1-trifluoroacetone δ_F = -84.60 ppm.*

Next, an attempt was made to track the disassembly of Qβ-F in a more dynamic fashion over a longer time period. It was proposed that addition of DTT followed by monitoring by ^{19}F NMR at set time points over an extended period of time (overnight) would allow changes in oligomeric states to be more closely followed. A fresh sample of Qβ-F VLP (0.15 mM with respect to monomer) was treated with 7.6 mM DTT and incubated for 20 minutes at room temperature before recording an initial ^{19}F NMR spectrum (Fig. 4.18 labelled 20 mins.). The sample was then heated to 37 ^{o}C in the NMR bore and a series of 20-minute ^{19}F NMR spectra were taken with a 40 minute wait between each measurement.

The ^{19}F NMR spectra at each of the time points are shown in Fig. 4.18. After room temperature incubation with DTT (7.6 mM) for 20 minutes, a broad peak centred at -38.9 ppm was observed. This was taken to be from an oligomeric species due to its large FWHH compared to that observed for the alkylated monomer in Fig. 4.13. No peak was observed at -78 ppm, the expected region for oxidised trifluoromethionine residue. However, a broad signal developed at -78 ppm after 132 minutes (Fig. 4.18), which was observed to increase in intensity as a function of time.

The spectra zoomed in the region of the -39 ppm peak are shown in Fig. 4.19 for a representative selection of time points. After incubation at 37 ^{o}C, a sharp peak at -39.14 ppm developed, approximately equal to the chemical shift measured for the alkylated monomer (see Fig. 4.13). The peak at -39.14 ppm is consistent throughout the timecourse. The ^{19}F NMR spectrum taken 532 minutes after addition of DTT, shows the main peak had split, which is more prominently seen after 1012 minutes. This could be due to individual multimers. Although it is not possible to fully characterize the individual multimers, these results do show a characteristic feature of a broad multimer peak centred at δ_F -38.9 and a sharp monomer peak at δ_F -39.14, consistent with the results shown in Fig. 4.17.

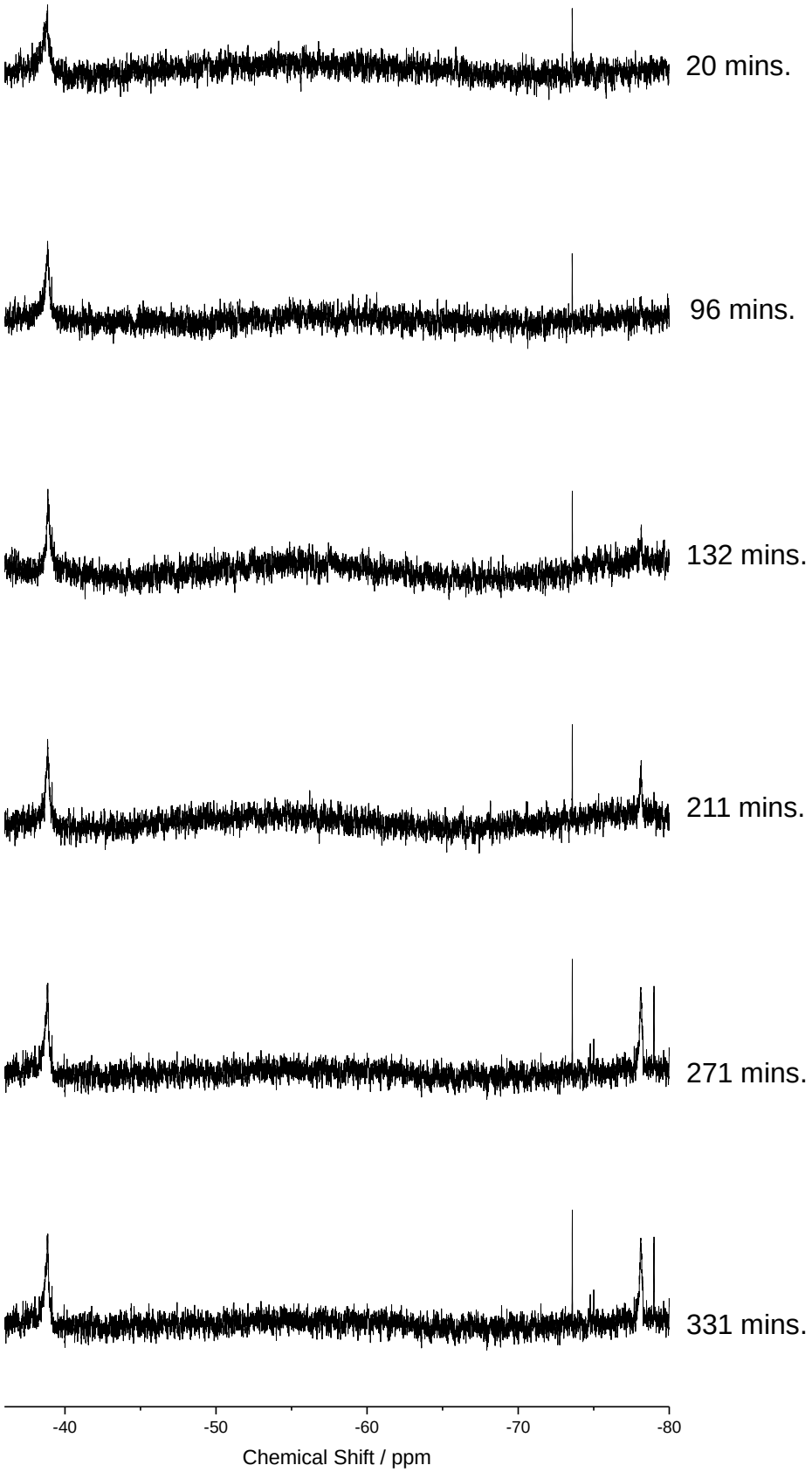


Figure 4.18 *^{19}F NMR of 0.15 mM Qβ-F VLP treated with 7.6 mM DTT taken at the labelled time after initial addition of DTT. The spectra were taken at 37 ^{o}C except for the spectrum at 20 mins., which was taken at 25 ^{o}C. The spectra are referenced to 1,1,1-trifluoroacetone δ_F = -84.60 ppm.*

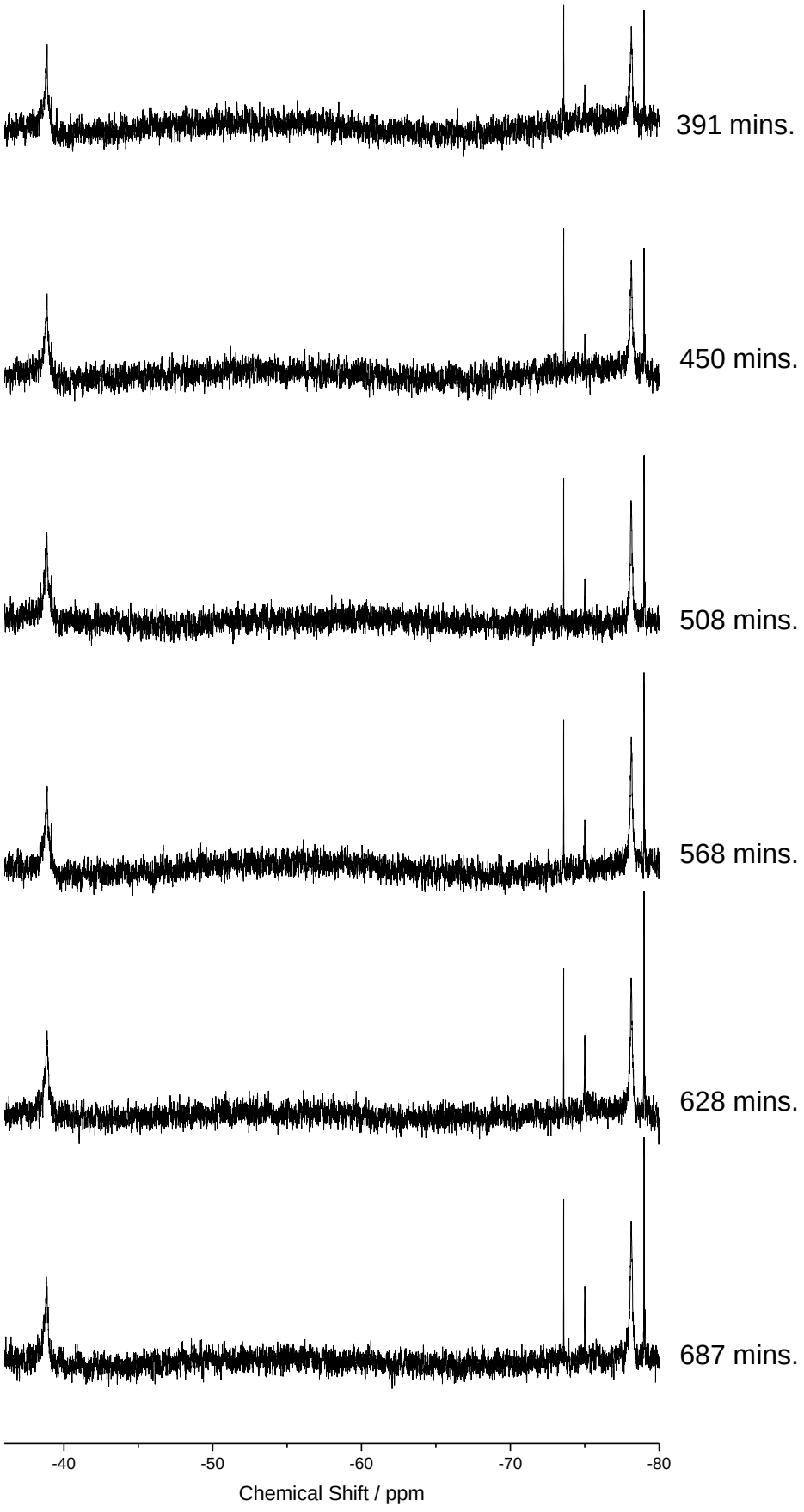


Figure 4.18 Continued.

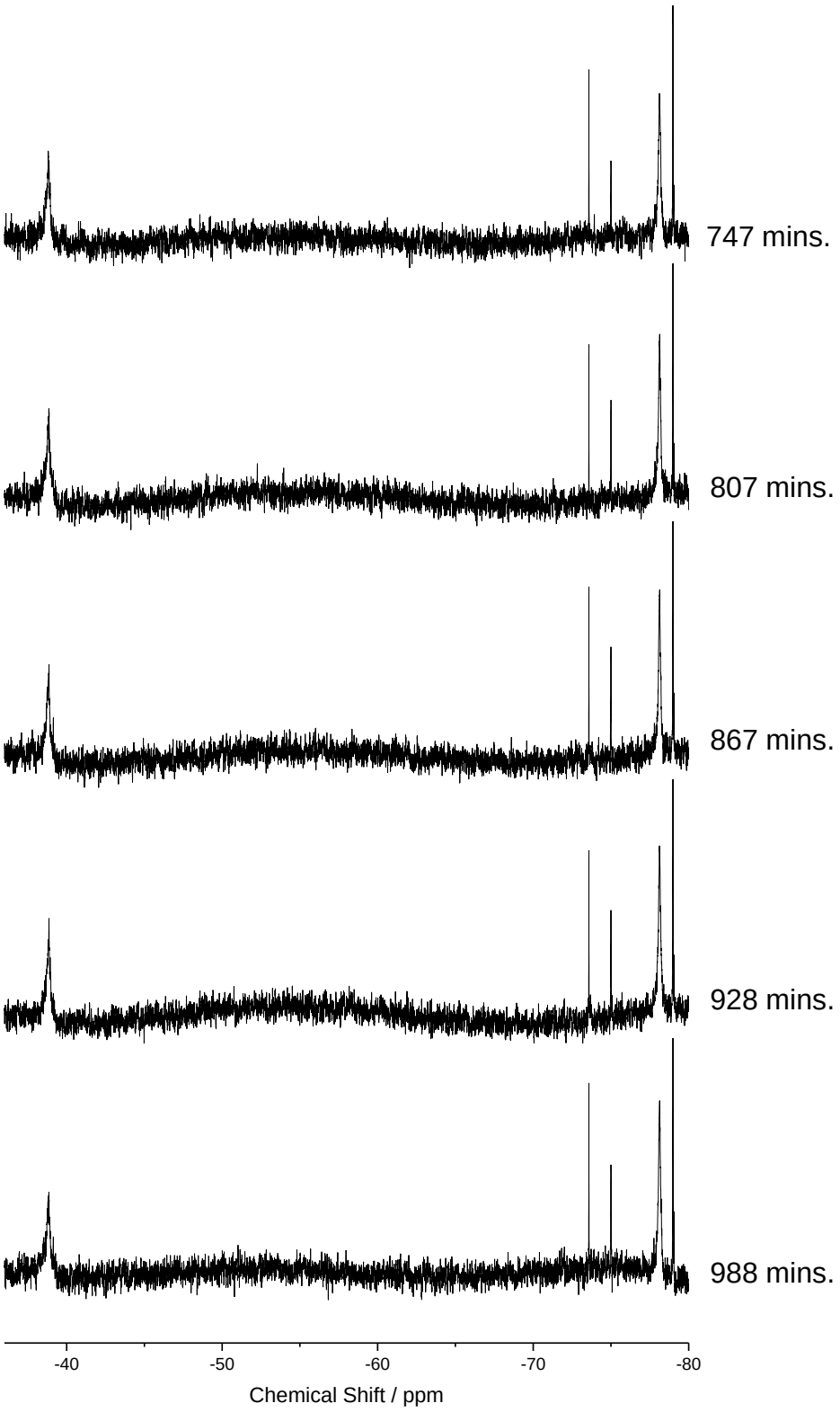


Figure 4.18 Continued.

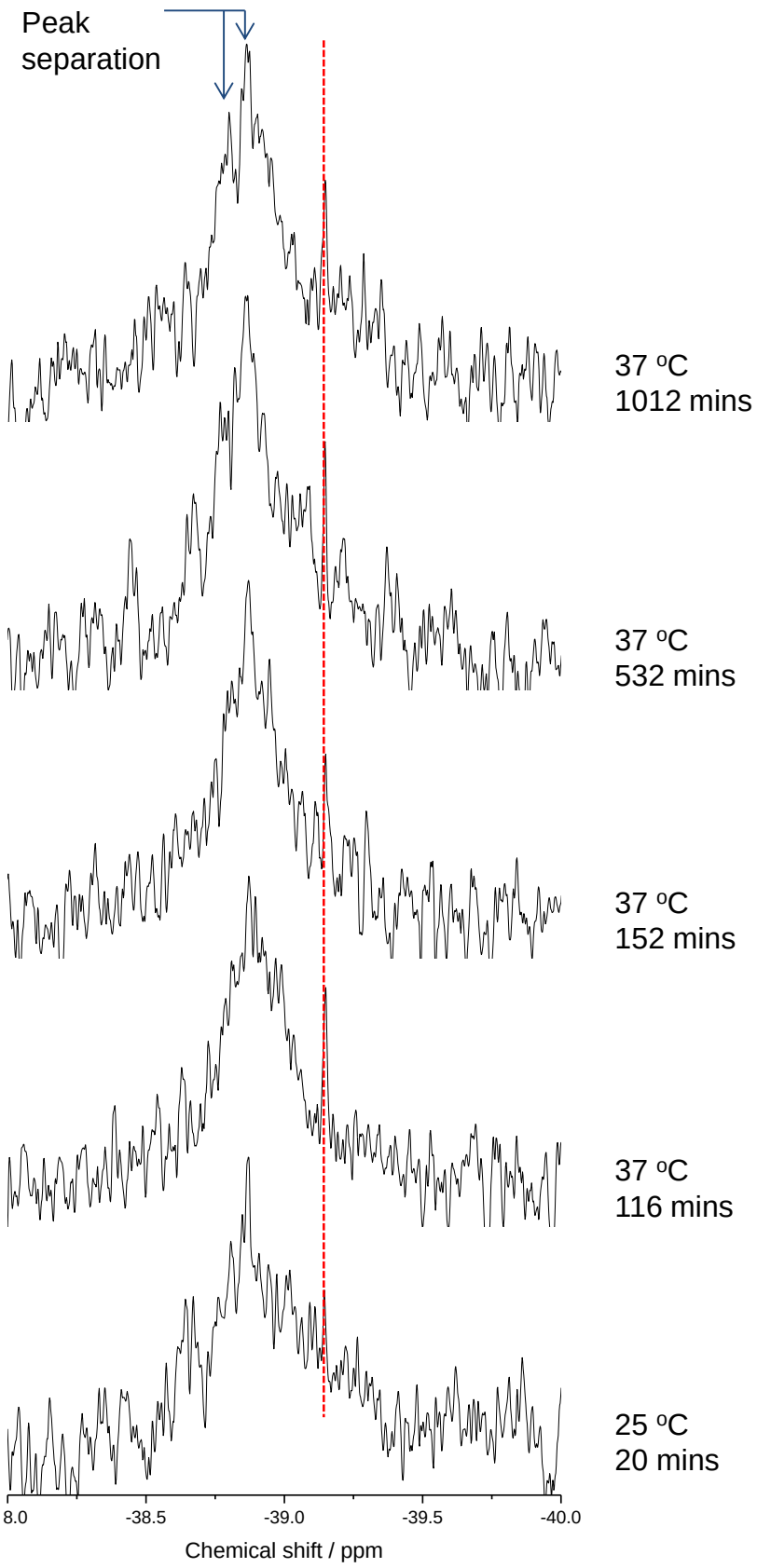


Figure 4.19 *^{19}F NMR of 0.15 mM Qβ-F VLP treated with 7.6 mM DTT taken at the temperatures shown and the time after initial addition of DTT. The red dotted line is at δ_F -39.14 ppm. The spectra are referenced to 1,1,1-trifluoroacetone δ_F = -84.60 ppm.*

In an effort to resolve the broad multimer peaks observed at -39 ppm and -78 ppm (Fig. 4.18), it was decided to titrate SDS into the sample. Fig. 4.20 shows the ^{19}F NMR spectra of the sample after addition of different amounts of SDS. The sample from Fig. 4.19 was stored at 4 ^{o}C for 15 days before carrying out the SDS titration experiments shown in Fig. 4.20. Unexpectedly the peak at -78 ppm had increased in intensity significantly whereas the peak at -39 ppm remained unchanged.

After the first addition of SDS (final concentration 1.68 mM), a sharp peak at -39.15 ppm was observed. This is in line with the peak observed in Fig. 4.19 which was tentatively assigned to the monomer. A broad multimer feature is still observed around -38.90 ppm. On subsequent additions of SDS, the intensity of this multimer feature is seen to diminish with simultaneous increase in intensity of the monomer peak. This is seen more clearly in the spectrum zoomed in the -39 ppm region, shown in Fig. 4.21. In contrast, the peak at -78 ppm appears to be unaffected by the addition of SDS (Fig. 4.20).

The results of the work leading to Fig. 4.19 and Fig. 4.21 show that significant intermolecular bonding interactions exist between Qβ monomers with trifluoromethionine residues other than disulphide bonding, most likely involving hydrophobic interactions. The data also imply that these intermolecular bonding interactions are not present between Qβ monomers with oxidised trifluoromethionine residues since a sharp resonance with a FWHH close to that seen for the monomer was observed *before* the addition of the denaturant, SDS (Fig. 4.20). This may be due to changes in the secondary structure of the Qβ monomer when the trifluoromethionine residue is oxidised which prevents effective dimerization through hydrophobic interactions.

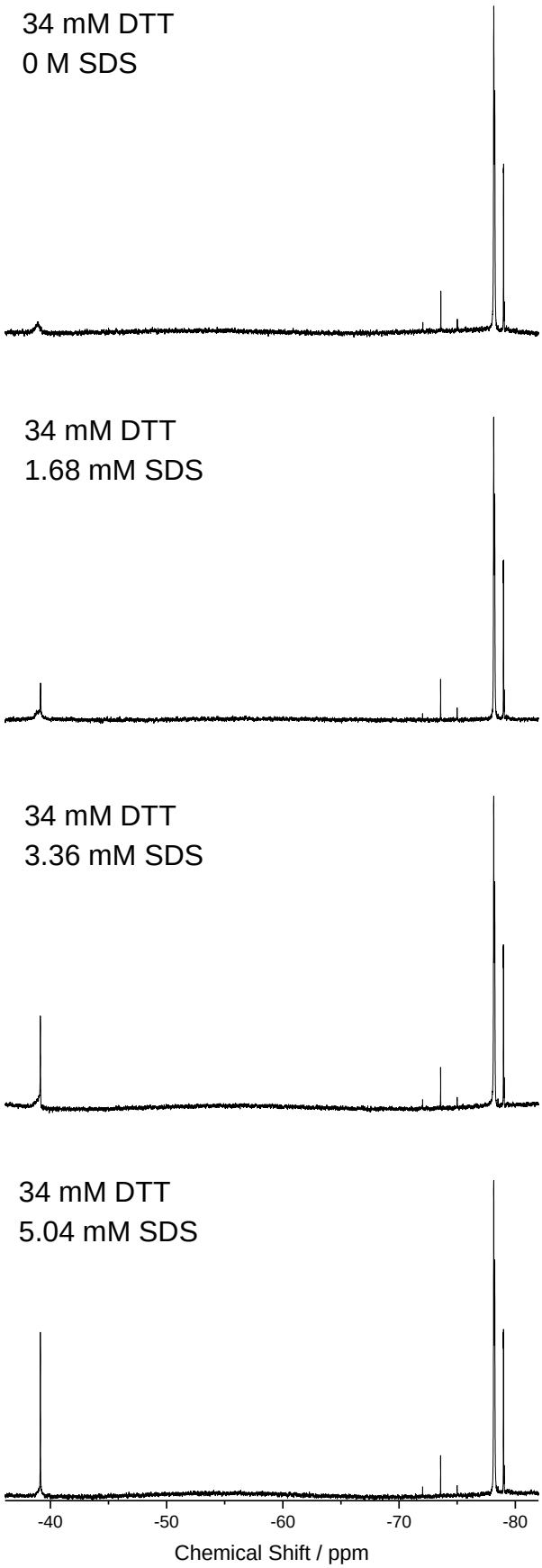


Figure 4.20 *^{19}F NMR of 0.15 mM Qβ-F VLP after treatment with 34 mM DTT and addition of SDS to a final concentration as indicated. The spectra are referenced to 1,1,1-trifluoroacetone δ_F = -84.60 ppm.*

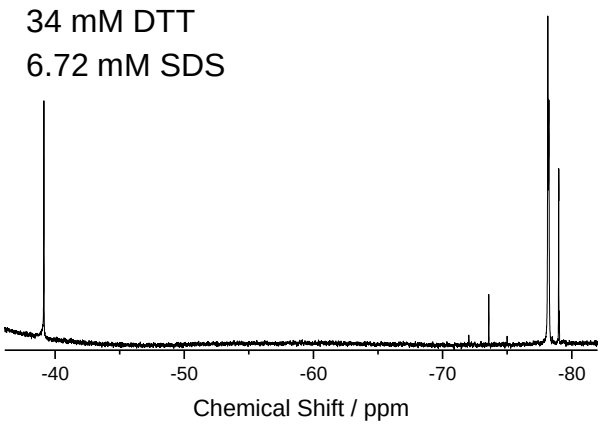


Figure 4.20 Continued.

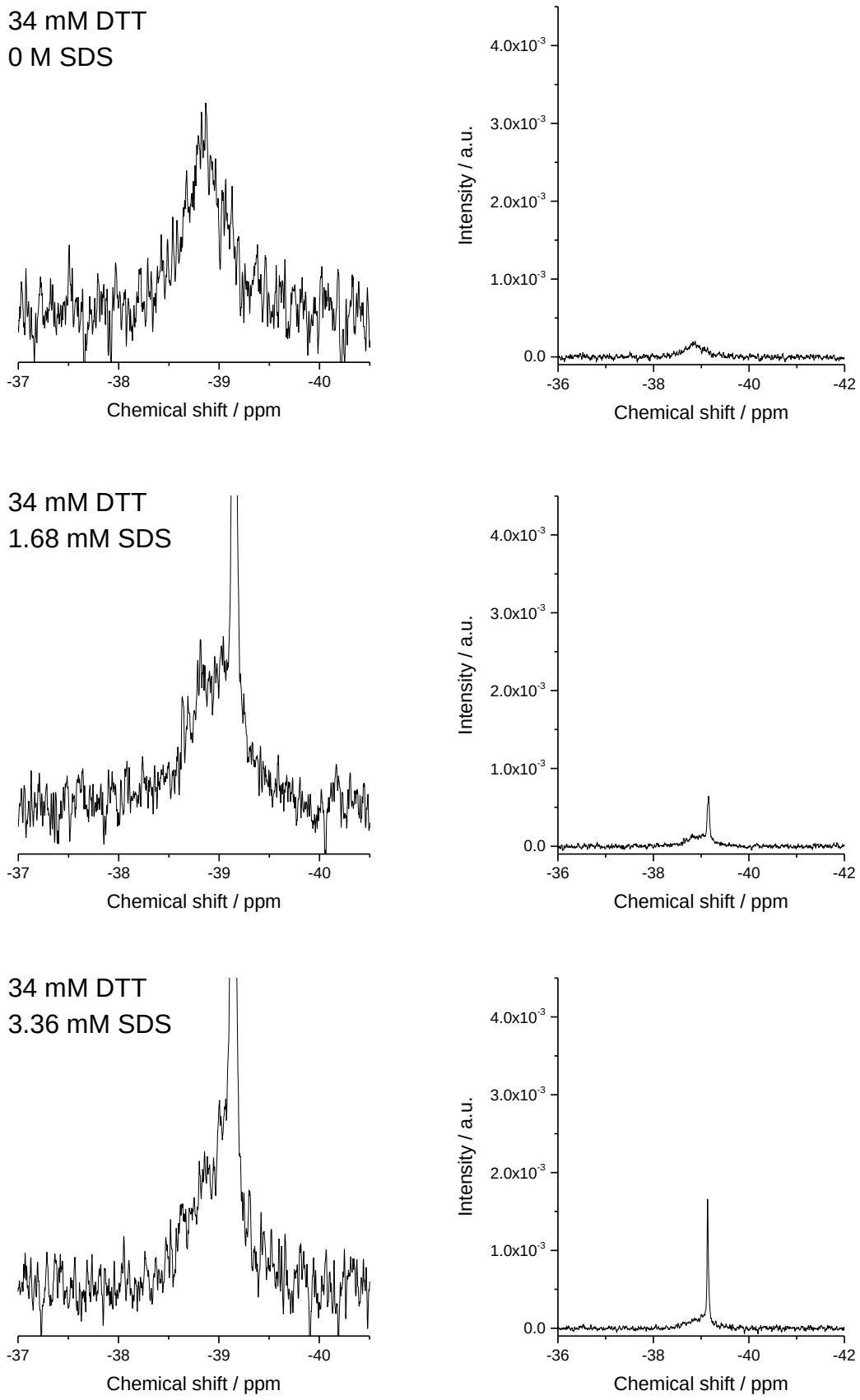


Figure 4.21 *^{19}F NMR of 0.15 mM Qβ-F VLP after treatment with 34 mM DTT and addition of SDS to a final concentration as indicated. The spectra are referenced to 1,1,1-trifluoroacetone δ_F = -84.60 ppm.*

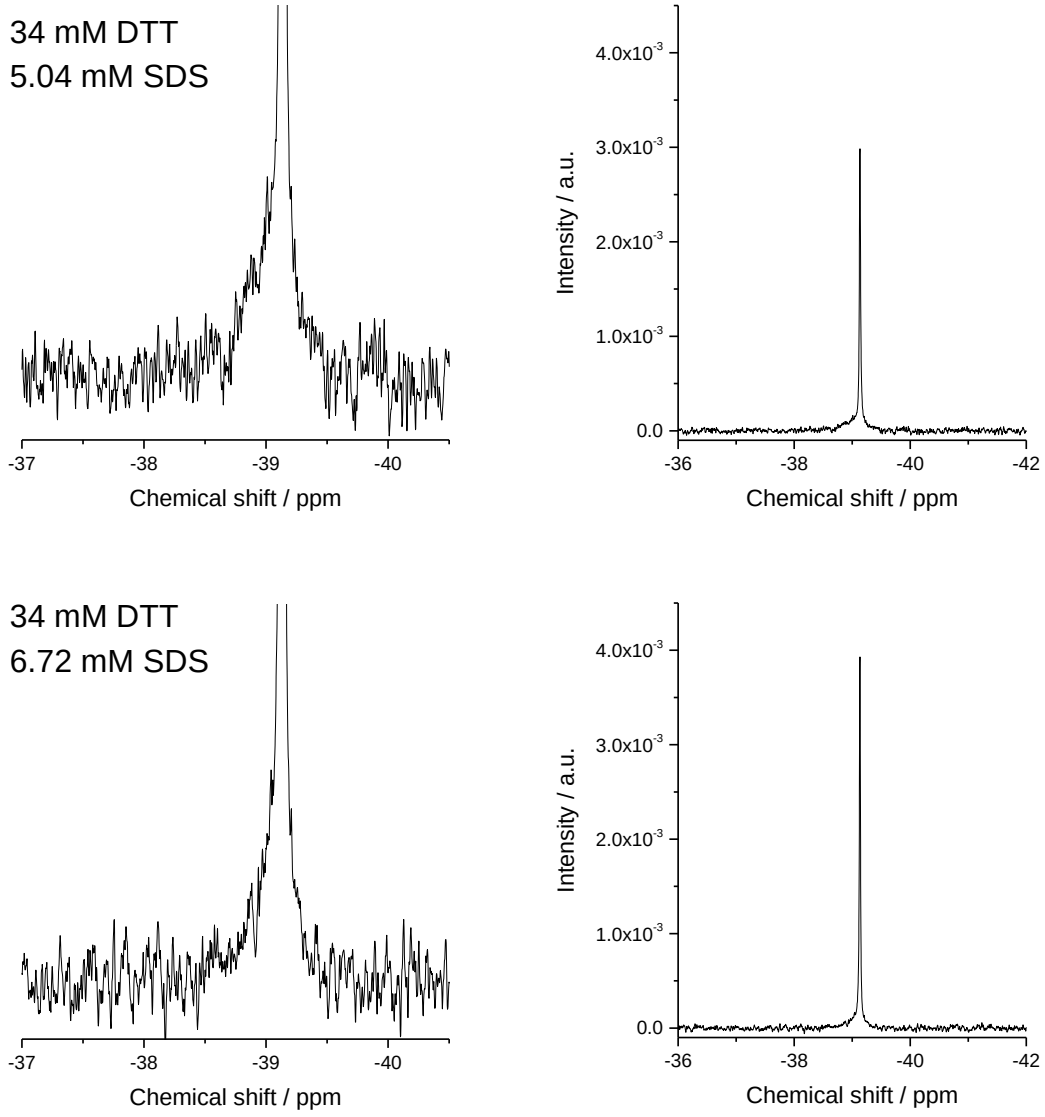


Figure 4.21 continued.

4.6 In-cell NMR

The results described above for disassembly of Qβ-F *in vitro*, demonstrate the potential for this construct to be used to probe disassembly of the phage during infection of *E. coli.* Moving from *in vitro* NMR analysis of proteins to *ex vivo* (isolated living cells) and *in vivo* (inside a living body) NMR studies presents a number of challenges, including that associated with molecular crowding in a cellular environment. Molecular crowding causes broadening of resonance peaks due to increased solution viscosity and interactions of the protein of

interest with intracellular proteins, thus making interpretation of spectra more difficult and ultimately reducing sensitivity.[35–38]

RNA phages only infect F+ (male) *E. coli* strains;[39] hence, for infection experiments, C3000 (F+) *E. coli* (LGC standards) were used. The cells were grown according to the product protocol.[40] Cultures were grown from glycerol stocks, originating from single colonies, to an OD_{600} of 0.5 – 0.6.[41] The cells were centrifuged and resuspended in tryptone medium containing D_2O, since D_2O is required to lock the magnetic field of the NMR spectrometer. Qβ-F VLPs were added to a final concentration of 0.5 mg/mL (~35 μM in monomer). A control containing deuterated tryptone medium instead of cell slurry was also prepared.

A ^{19}F NMR spectrum of the *E. coli*/Qβ-F VLPs mixture was taken at 37 ^{o}C; the sample was allowed to equilibrate in the bore for 10 minutes before taking the spectrum. No signal was seen after 1000 scans, implying that the Qβ-F VLPs had not disassembled for a lack of signal at -39 ppm. In the control sample, the same amount of Qβ-F added to buffer instead of *E. coli* was measured for the same number of scans; this was also observed not to give a signal. This result supports the proposal that the original sample of Qβ-F was in its VLP state. As a positive control, to ensure the lack of signal was not due to an insufficient number of scans, the sample of Qβ-F used to record spectrum Fig. 4.14 B (i.e. treated with SDS and DTT) was measured. The concentrations of the samples were similar (control sample: 0.07 mM; Fig. 4.14 B: 0.09 mM). A signal was observed in the positive control spectrum (run for 1000 scans) showing that the lack of signal in the *E. coli* samples is not due to an insufficient number of scans. The *E. coli*/Qβ-F VLP slurry was run overnight at 37 ^{o}C; 24 000 scans were recorded, after which no signal was present in the ^{19}F NMR spectrum. Together, these results provide evidence that no disassembly of the Qβ-F bacteriophage occurs when the bacteriophage is incubated with its host, male *E. coli* F+, at least under the tested conditions.

The results support work by Stockley and co-workers on the interaction of a closely related bacteriophage, MS2, with F-pili of its host, also F+ *E. coli*.[42] Stockley's work was based upon TEM images which show intact MS2 particles after incubation with F-pili. The benefit of the ^{19}F NMR data presented here is that ^{19}F NMR is acutely sensitive to the oligomeric state of the virus, and so can record the presence of lower oligomers which are undetectable by TEM.

Since it had been determined that the Qβ-F bacteriophage remains intact on incubation with its host, the next aspect to determine was whether the bacteriophage entered its host cell or not. After overnight incubation at 37 ^{o}C the sample of Qβ-F and *E. coli* was centrifuged to collect the cells before removing the supernatant, resuspending the cells in 500 µL of buffer and lysing by sonication. The cell debris was then collected by centrifugation and resuspended in buffer (500 µL).

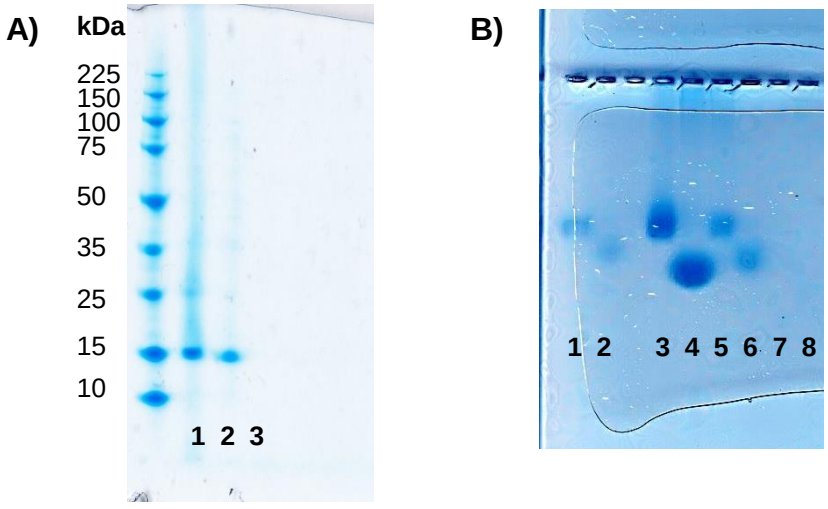


Figure 4.22 A: *SDS-PAGE. Lane 1: Supernatant after centrifuging E. coli NMR sample. Lane 2: Supernatant after lysing E. coli pellet from NMR sample resuspended in buffer (lysate). Lane 3: Cell debris after lysing cells, resuspended in buffer.* ***B:*** *Agarose gel. Lane 1: Qβ-Met VLP (non-reducing). Lane 2: Qβ-Met VLP (reduced). Lane 3: Supernatant after centrifuging E. coli NMR sample (non-reducing). Lane 4: Supernatant after centrifuging E. coli NMR sample (reduced). Lane 5: Supernatant after lysing E. coli pellet from NMR sample resuspended in buffer (lysate) (non-reducing). Lane 6: Supernatant after lysing E. coli pellet from NMR sample resuspended in buffer (lysate) (reduced). Lane 7: Cell debris after lysing cells, resuspended in buffer (non-reducing). Lane 8: Cell debris after lysing cells, resuspended in buffer (reduced).*

The samples were analysed by reducing SDS-PAGE and agarose gel electrophoresis. The samples analysed by agarose gel electrophoresis were run alongside a standard of Qβ-Met VLP as a marker for the intact VLP (Qβ-F VLP and Qβ-Met VLP have been shown to co-migrate on agarose gel in Chapter 3 Page 104). The SDS-PAGE shows the presence of Qβ-F protein in both the supernatant after centrifuging the cells and the cell lysate. Qβ-F protein observed in the lysate implies the Qβ-F protein had been internalised into the *E. coli* cells. On the agarose gel, each sample was run in duplicate with the first being the non-denatured form and the second being after treatment with reducing loading dye and heat treatment. The agarose gel also shows the presence of Qβ-F in both the supernatant after centrifuging the cells and the cell lysate. Furthermore, the agarose gel clearly shows the intact VLP form of the Qβ-F in the cell lysate. This is in agreement with the result that no ^{19}F NMR signal was observed in after incubation of Qβ-F with the *E. coli* cells (discussed above).

The ^{19}F NMR and electrophoretic results above together provide strong evidence to show that the Qβ-F particle does not disassemble when incubated with its host, F+ *E. coli*. However, these results do not conclusively determine whether the Qβ-F particle is internalised into the F+ *E. coli* cells since the Qβ-F detected in the cell lysates by the SDS-PAGE and agarose gels could have been associated with the cell outer surface, as opposed to actually being inside the cells. To investigate whether the Qβ-F particle is internalized into its host *E. coli*, the surface of the particle could be labelled with a fluorophore. The process could then be imaged by fluorescence microscopy.

4.7 Conclusions and future work

The above results demonstrate that fluorine labelling of the Qβ VLP and monitoring of ^{19}F NMR spectra during disassembly provides a new way of analysing virus disassembly. The constituent oligomeric states of the capsid can be characterized by their differing ^{19}F chemical shifts. Furthermore, I believe this ^{19}F-labelling technique could be generalised to

other viruses. Of particular interest will be application of the labelling procedure to mammalian viruses where ^{19}F NMR could be used to assign a more detailed mechanism of infection and thus guide rational design of potent antivirus molecular therapeutics. For example, the Hepatitis B virus (HBV) is being extensively researched due to its cause of viral infection killing 1 million people annually.[43–45] Thus, ^{19}F labelling of HBV would be of interest.

The peak in the ^{19}F NMR spectrum of disassembled Qβ-F corresponding to the monomer has been assigned by matching the chemical shift with that of an isolated sample of the purified fluorinated Qβ monomer. The other peaks in the spectrum of disassembled Qβ-F are still to be assigned. This could be done using a similar technique to that used here for the monomer. Pentamers and hexamers of the virus can be obtained by treatment of the VLP with SDS.[6] The pentamers and hexamers could then be separated either by size-exclusion chromatography or ion-exchange chromatography. If successful, ^{19}F NMR spectra of the separated hexamers and pentamers could be measured to obtain a fluorine chemical shift for these species. Possible difficulties with this approach could be the potential loss of protein sample to the chromatography system and column.

Finally, the potential for using Qβ-F VLP as an *in vivo* off-to-on imaging agent was proposed (Fig. 4.8). To this end, it will be important to analyse Qβ-F VLP incubated with mammalian cells. Finn and co-workers have shown that Qβ modified with the glycoprotein transferrin can be internalised into mammalian cells by the transferrin receptor.[18] This was achieved by labelling the surface of Qβ with AlexaFluor 568 and imaging stained cells by fluorescence microscopy. However, the reported results do not show whether Qβ disassembles when internalized into the cells. Qβ-F VLP would show this since it would produce a ^{19}F NMR signal if disassembly occurred. Hence, future work could be guided towards modification of the lysine residues of Qβ-F VLP with transferrin.

4.8 References

1. Srinivas, M. *et al.* Customizable, multi-functional fluorocarbon nanoparticles for quantitative in vivo imaging using 19F MRI and optical imaging. *Biomaterials* **31,** 7070–7077 (2010).

2. Hore, P. *Nuclear Magnetic Resonance*. (Oxford University Press, 2006).

3. Claridge, T. D. W. *High-Resolution NMR Techniques in Organic Chemistry*. (Elsevier, 2009).

4. Fiaux, J., Bertelsen, E. B., Horwich, A. L. & Wüthrich, K. NMR analysis of a 900K GroEL-GroES complex. *Nature* **418,** 207–211 (2002).

5. Dolbier, W. R. *Guide to Fluorine NMR for Organic Chemists*. (Wiley, 2009).

6. Takamatsu, H. & Iso, K. Chemical evidence for the capsomeric structure of phage Qβ. *Nature* **298,** 819–824 (1982).

7. Matthews, K. S. & Cole, R. D. Shell Formation by Capsid Protein of f2 Bacteriophage. *Virology* **65,** 1–15 (1972).

8. Floyd, N., Vijayakrishnan, B., Koeppe, A. R. & Davis, B. G. Thiyl Glycosylation of Olefinic Proteins: S-Linked Glycoconjugate Synthesis. *Angew. Chem. Int. Ed.* **48,** 7798–7802 (2009).

9. Iyer, A. K., Khaled, G., Fang, J. & Maeda, H. Exploiting the enhanced permeability and retention effect for tumor targeting. *Drug Discov. Today* **11,** 812–818 (2006).

10. Takaoka, Y. *et al.* Self-assembling nanoprobes that display off/on 19F nuclear magnetic resonance signals for protein detection and imaging. *Nat. Chem.* **1,** 557–561 (2009).

11. Huang, X. *et al.* Multi-Chromatic pH-Cctivatable 19F-MRI Nanoprobes with Binary ON/OFF pH Transitions and Chemical-Shift Barcodes. *Angew. Chem. Int. Ed. Engl.* **52,** 8074–8078 (2013).

12. Deres, K. *et al.* Inhibition of Hepatitis B Virus Replication by Drug-Induced Depletion of Nucleocapsids. *Science* **299,** 893–896 (2003).

13. Stray, S. J. *et al.* A heteroaryldihydropyrimidine activates and can misdirect hepatitis B virus capsid assembly. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 8138–8143 (2005).

14. Stray, S. J. & Zlotnick, A. BAY 41-4109 has multiple effects on Hepatitis B virus capsid assembly. *J. Mol. Recognit.* **19,** 542–548 (2006).

15. Bausinger, R. *et al.* The Transport of Nanosized Gene Carriers Unraveled by Live-Cell Imaging. *Angew. Chem. Int. Ed. Engl.* **45,** 1568–1572 (2006).

16. Kreuter, J. Mechanism of polymeric nanoparticle-based drug transport across the blood-brain barrier (BBB). *J. Microencapsul.* **30,** 49–54 (2013).

17. Lakadamyali, M., Rust, M. J., Babcock, H. P. & Zhuang, X. Visualizing infection of individual influenza viruses. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 9280–9285 (2003).

18. 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).

19. Ewers, H. *et al.* Single-particle tracking of murine polyoma virus-like particles on live cells and artificial membranes. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 15110–15115 (2005).

20. Lewis, J. D. *et al.* Viral nanoparticles as tools for intravital vascular imaging. *Nat. Med.* **12,** 354–360 (2006).

21. Mukherjee, S., Thorsteinsson, M. V, Johnston, L. B., DePhillips, P. A. & Zlotnick, A. A Quantitative Description of In Vitro Assembly of Human Papillomavirus 16 Virus-Like Particles. *J. Mol. Biol.* **381,** 229–237 (2008).

22. Borodavka, A., Tuma, R. & Stockley, P. G. Evidence that viral RNAs have evolved for efficient, two-stage packaging. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 15769–15774 (2012).

23. Wiley, D. T., Webster, P., Gale, A. & Davis, M. E. Transcytosis and brain uptake of transferrin-containing nanoparticles by tuning avidity to transferrin receptor. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 8662–8667 (2013).

24. Wang, G.-F., Li, C. & Pielak, G. J. Probing the Micelle-Bound Aggregation-Prone State of α-Synuclein with 19F NMR Spectroscopy. *ChemBioChem* **11,** 1993–1996 (2010).

25. Schlesinger, A. P., Wang, Y., Tadeo, X., Millet, O. & Pielak, G. J. Macromolecular Crowding Fails To Fold a Globular Protein in Cells. *J. Am. Chem. Soc.* **133,** 8082–8085 (2011).

26. Thomas, C. A., Talaty, E. R. & Bann, J. G. 3S-Fluoroproline as a probe to monitor proline isomerization during protein folding by 19F-NMR. *Chem. Comm.* 3366–3368 (2009).

27. Li, H. L. & Frieden, C. Observation of sequential steps in the folding of intestinal fatty acid binding protein using a slow folding mutant and F-19 NMR. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 11993–11998 (2007).

28. Chiti, F. & Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **75,** 333–366 (2006).

29. Larda, S., Simonetti, K., Al-Abdul-Wahid, M., Sharpe, S. & Prosser, R. S. Dynamic Equilibria Between Monomeric and Oligomeric Misfolded States of the Mammalian Prion Protein Measured by 19F NMR. *J. Am. Chem. Soc.* **135,** 10533–10541 (2013).

30. Suzuki, Y. *et al.* Resolution of Oligomeric Species During the Aggregation of Aβ1-40 Using 19F NMR. *Biochemistry* **52,** 1903–1912 (2013).

31. Stilbs, P. Fourier Transform Pulsed-Gradient Spin-Echo Studies of Molecular Diffusion. *Prog. Nucl. Magn. Reson. Spectrosc.* **19,** 1–45 (1987).

32. Morris, K. F. & Johnson, C. S. Diffusion-Ordered Two-Dimensional Nuclear Magnetic Resonance Spectroscopy. *J. Am. Chem. Soc.* **114,** 3139–3141 (1992).

33. Price, W. S. Pulsed-Field Gradient Nuclear Magnetic Resonance as a Tool for Studying Translational Diffusion: Part 1. Basic Theory. *Concepts Magn. Reson.* **9,** 299–336 (1997).

34. Dill, K. A., Bromberg, S. & Dirk, S. *Molecular driving forces: statistical thermodynamics in chemistry and biology*. (Garland Science, 2003).

35. Li, C. *et al.* 19F NMR in Escherichia coli. *J. Am. Chem. Soc.* 321–327 (2010).

36. Li, C. *et al.* Differential Dynamical Effects of Macromolecular Crowding on an Intrinsically Disordered Protein and a Globular Protein: Implications for In-Cell NMR Spectroscopy. *J. Am. Chem. Soc.* **130,** 6310–6311 (2008).

37. 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).

38. Serber, Z., Ledwidge, R., Miller, S. M. & Dötsch, V. Evaluation of Parameters Critical to Observing Proteins Inside Living Escherichia coli by In-Cell NMR Spectroscopy. *J. Am. Chem. Soc.* **123,** 8895–8901 (2001).

39. Loeb, T. & Zinder, N. D. A Bacteriophage Containing RNA. *Proc. Natl. Acad. Sci. U. S. A.* **47,** 282–289 (1961).

40. ATCC Product Sheet Escherichia coli (ATCC® 15597™). at <https://www.atcc.org/~/ps/15597.ashx>

41. Ikeya, T. *et al.* NMR protein structure determination in living E. coli cells using nonlinear sampling. *Nat. Protoc.* **5,** 1051–1060 (2010).

42. Dent, K. C. *et al.* The Asymmetric Structure of an Icosahedral Virus Bound to Its Receptor Suggests a Mechanism for Genome Release. *Structure* **21,** 1225–1234 (2013).

43. Dienstag, J. L. Hepatitis B Virus Infection. *N. Engl. J. Med.* **359,** 1486–1500 (2008).

44. Aspinall, E. J., Hawkins, G., Fraser, A., Hutchinson, S. J. & Goldberg, D. Hepatitis B prevention, diagnosis, treatment and care: a review. *Occup. Med. (Lond).* **61,** 531–450 (2011).

45. Zhang, F. & Wang, G. A review of non-nucleoside anti-hepatitis B virus agents. *Eur. J. Med. Chem.* **75,** 267–281 (2014).

46. Bradford, M. M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **72,** 248–254 (1976).

47. Thermoscientific. Chemistry of Protein Assays. at <http://www.piercenet.com/method/chemistry-protein-assays#dyeassays>

48. Jain, R., Knorr, A. L., Bernacki, J. & Srivastava, R. Investigation of Bacteriophage MS2 Viral Dynamics Using Model Discrimination Analysis and the Implications for Phage Therapy. *Biotechnol. Prog.* **22,** 1650–1658 (2006).

4.9 Experimental procedures

All ^{19}F NMR spectra were referenced to the internal standard 1,1,1-trifluoroacetone which was set to δ_F = -84.60 ppm.

4.9.1 ^{19}F NMR of DL-TFM

DL-Trifluoromethionine (10 mg) and 1,1,1-trifluoroacetone (5.8 mg) were added to a glass vial and dissolved in deuterated TBS buffer (500 µL). The solution was transferred to a NMR tube using a needle and syringe. The spectrum was measured on a Bruker AVB500 spectrometer and is shown in Fig. 4.1.

4.9.2 ^{19}F NMR of Qβ-F VLP with varying temperatures

A sample of Qβ-F in TBS was concentrated as follows. The solution was exchanged into H_2O using a PD-10 column (GE Healthcare). Fraction volumes of 8 drops ≈ 300 µL were taken. The fractions were analysed by Nanodrop Spectrophotometer (Thermo Scientific) at 280 nm and fractions containing protein were pooled. The sample was freeze-dried to remove H_2O and the resulting white powder was dissolved in 400 µL deuterated TBS. The solution was

added to an NMR tube. The concentration of the sample was measured using the Bradford assay (Pierce) giving a value of 0.4 mg/mL. A spectrum was measured on a Bruker AVB500 spectrometer at 25 ^{o}C and is shown in Fig. 4.2 A. The temperature of the NMR bore was then increased to 60 ^{o}C. The sample was left to equilibrate in the bore for approximately 15 minutes before a spectrum was taken and is shown in Fig. 4.2 B.

SDS-PAGE of the NMR sample was next taken. Lane 1: 2 µL NMR sample, 8 µL TBS buffer, 2 µL DNA native loading dye, no heat treatment. Lane 2: 2 µL NMR sample, 8 µL TBS buffer, 2 µL SDS-PAGE reducing dye, sample heated at 90 ^{o}C for 10 minutes. A 15-150 kDa protein marker was also loaded. A 10% Bis-Tris gel was used with MES buffer. The gel was run for 40 minutes at 200V. The gel was stained with coomassie solution (Instan*Blue*™, Expedeon) to afford the gel shown in Fig. 4.2 C.

4.9.3 ^{19}F NMR of Qβ-F VLP treated with SDS

To the NMR sample used in Fig. 4.2, SDS (19.35 mg) was added to give a final SDS concentration of 0.17 M. A spectrum was measured on a Bruker AVB500 spectrometer at 27 ^{o}C and is shown in Fig. 4.5 A. The temperature of the NMR bore was then increased to 50 ^{o}C. The sample was left to equilibrate in the bore for approximately 15 minutes before a spectrum was taken and is shown in Fig. 4.5 B. The temperature of the NMR bore was then increased to 90 ^{o}C. The sample was left to equilibrate in the bore for approximately 15 minutes before a spectrum was taken and is shown in Fig. 4.5 C.

This sample was stored at 4 ^{o}C, and after 20 days of storage a ^{19}F NMR spectrum was measured. A spectrum was measured on a Bruker AVB500 spectrometer at 25 ^{o}C and is shown in Fig. 4.6.

4.9.4 ^{19}F NMR of Qβ-F VLP treated with SDS and DTT

A spectrum of the NMR sample used in Fig. 4.6 was measured on a Bruker AVB500 spectrometer at 25 ^{o}C and is shown in Fig. 4.7 A. Then 63 mg of DTT was added to the NMR sample giving a final DTT concentration of 0.816 M. The temperature of the NMR bore was then increased to 50 ^{o}C. The sample was left to equilibrate in the bore for approximately 15 minutes before a spectrum was taken and is shown in Fig. 4.7 B. The temperature of the NMR bore was then increased to 90 ^{o}C. The sample was left to equilibrate in the bore for approximately 15 minutes before a spectrum was taken and is shown in Fig. 4.7 C. Finally, the sample was left to stand at room temperature overnight and a spectrum was taken at 25 ^{o}C the next morning (Fig. 4.7 D).

The right side ^{19}F NMR spectrum shown in Fig. 4.8 is that of Fig. 4.7 D.

4.9.5 ^{19}F NMR of Qβ-F VLP treated with SDS

A new sample of Qβ-F was prepared as follows. A solution of 1.87 mg Qβ-F in 6.6 mL TBS buffer was dialysed into H_2O using a Slide-A-Lyzer dialysis cassette with a 10 kDa molecular weight cut-off. The solution was extensively dialysed by repeated refreshing of H_2O to remove TBS salts from the sample. The sample was then freeze-dried to remove H_2O and the resulting white powder was dissolved in 1 mL deuterated TBS and 600 µL TBS. The final solution was 2 mL, and its concentration was 1.26 mg/mL as determined by the Bradford assay.[46,47] 450 µL of this solution (0.567 mg Qβ-F) was used for the ^{19}F NMR spectra shown in Fig. 4.9. The bore of a Bruker AVB500 spectrometer was heated to 37 ^{o}C, the NMR sample was left to equilibrate in the bore for approximately 15 minutes before a spectrum was taken (Fig 4.9 A). 22.8 mg of SDS was added to the sample to give a final concentration of 0.17 M and the sample was placed back in the NMR bore. A ^{19}F NMR spectrum was recorded immediately (approximately 5 minutes after the addition of SDS). Another ^{19}F

NMR spectrum was recorded 30 minutes later, no change was observed in the spectrum. The NMR bore and sample were allowed to cool to 25 ^{o}C. A ^{19}F NMR spectrum was recorded 15 minutes after the temperature had been reached. The sample was left to stand overnight at room temperature, after which another ^{19}F NMR spectrum was recorded (Fig. 4.9 D).

4.9.6 Preparation of the alkylated Qβ-F monomer

A solution of Qβ-F VLPs (9.5 mL, 44 µg/mL) was concentrated using a Viva Spin 6 tube with a 10 kDa molecular weight cut off. The final volume was 5 mL. 10 mg of DTT was added to this solution, and the sample was heated at 95 ^{o}C for 20 minutes. A solution of iodoacetamide (≈100 mM) was made up in a foil-wrapped Eppendorf tube: 526 µL of H_2O was added to 9.46 mg of iodoacetamide. Once the protein solution had been heated with DTT, 236 µL of the iodoacetamide solution (≈100 mM) was added, the reaction was allowed to proceed in the dark at room temperature for 20 minutes. The reaction mixture was then loaded onto a sucrose density gradient (5%-40% w/w sucrose in TBS). The gradients were ultracentrifuged: Beckman Coulter Optima™ L-90K, 37600 rpm, 4 ^{o}C, 3 hours, with maximum acceleration and deceleration set to "No brake". After ultracentrifugation, the sucrose gradients were fractionated into 1 mL fractions. The fractions were then analysed for protein content using a Nanodrop Spectrophotometer. Protein was detected in the first three fractions and none in the fractions where VLPs are expected, thus confirming the monomeric state of the Qβ-F protein. Fractions were pooled and concentrated to 5 mL using a Viva Spin with a 5 kDa cut-off. The sucrose was removed from the solution using a PD-10 column. The column was eluted with TBS so the final solution was in the buffer.

The Qβ-F alkylated monomer was analysed by SDS-PAGE. Lane 1: 1 µL Qβ-Met VLP, 9 µL TBS, 2 µL DNA native loading dye, no heat treatment. Lane 2: 1 µL Qβ-Met VLP, 9 µL 0.244 M SDS solution, 2 µL SDS-PAGE reducing dye, 90 ^{o}C for 10 minutes. Lane 3: 2 µL

Qβ-F alkylated monomer in TBS, 8 µL TBS, 2 µL DNA native loading dye, no heat treatment. Lane 4: 2 µL Qβ-F alkylated monomer in TBS, 8 µL 0.244 M SDS solution, 2 µL SDS-PAGE reducing dye, 90 ^{o}C for 10 minutes. A 10-225 kDa ladder was also loaded. A 10% Bis-Tris gel was used with MES buffer. The gel was run for 40 minutes at 200V. The gel was stained with coomassie solution (Instan*Blue*™, Expedeon) to afford the gel shown in Fig. 4.12.

4.9.7 ^{19}F NMR of Qβ-F VLP treated with SDS and DTT

To the NMR sample of Qβ-F from Fig. 4.9 (0.576 mg Qβ-F, 0.17 M SDS) was added DTT (to a final concentration of 18 mM). The NMR sample was then heated at 90 ^{o}C for 40 minutes. A ^{19}F NMR spectrum was then recorded and is shown in Fig. 4.14 A. A further addition of DTT was made to give a final DTT concentration of 36 mM (i.e. the concentration of DTT was doubled). The sample was heated at 90 ^{o}C for 40 minutes. A ^{19}F NMR spectrum was then recorded and is shown in Fig. 4.14 B.

A further addition of DTT was made to give a final DTT concentration of 102 mM. The sample was heated at 90 ^{o}C for 165 minutes. The final concentration of DTT was then raised to 169 mM and the sample was heated at 90 ^{o}C for 60 minutes. A ^{19}F NMR spectrum was then recorded and is shown in Fig. 4.16.

4.9.8 ^{19}F NMR of Qβ-F VLP with overnight incubation with DTT

A fresh sample of Qβ-F VLP was prepared (0.15 mM of the monomer). DTT was added to the sample to give a final concentration of 7.6 mM. The sample was incubated at room temperature for 20 minutes before recording a ^{19}F NMR spectrum which is shown in Fig. 4.18 and Fig. 4.19. The temperature of the NMR bore was then raised to 37 ^{o}C. The temperature was held at 37 ^{o}C overnight and a ^{19}F NMR spectrum (20 minute-scan) was recorded every 40 minutes.

4.9.9 ^{19}F NMR of Qβ-F with titration of SDS

The DTT concentration of the sample from Fig. 4.18/4.19 was then raised to 34 mM and incubated at 37 ^{o}C overnight. SDS was added to the sample to give a final concentration of 1.68 mM. A ^{19}F NMR spectrum was recorded. SDS was added to the NMR sample in increments and a ^{19}F NMR spectrum was measured for each increment. Resulting spectra are shown in Fig. 4.20 and Fig. 4.21.

4.9.10 Growth of *E. coli* strain C-3000

C-3000 is an F^+ (male) *E. coli* strain and has been used before for investigating the infection of *E. coli* by bacteriophages.[48] The *E. coli* C-3000 cells were purchased from ATCC (product number 15597) and prepared according to the product information sheet.[40]

4.9.10.1 Preparation of growth medium

Nutrient solution was made of as followed and filtered through a 0.2 µm syringe filter:

Reagent	Mass	Volume of H_2O
Glucose	1 g	9 mL
$CaCl_2 \cdot H_2O$	0.294 g	2 mL
Thiamine	10 mg	1 mL

Bulk medium was made up by adding the following reagents to 1 L of deionized water and then autoclaving:

Reagent	Mass
Tryptone	10 g
Yeast extract	1.0 g

NaCl	8.0 g
Agar	15 g

The nutrient solution was then added to the bulk medium to give the growth medium.

The resulting solution was poured into petri dishes (≈20 mL for each petri dish). The plates were allowed to cool.

The growth medium (6 mL) (without agar) was used to hydrate the *E. coli* cells as purchased. 100 µL of this solution was used to inoculate an agar plate. Three agar plates were prepared and were incubated at 37 ^{o}C for 16.5 hours. A single colony was used to inoculate a fresh solution of medium (10 mL). The culture was incubated at 37 ^{o}C for 16 hours (OD_{600} = 1.0). A glycerol stock of the culture was prepared by adding 1 mL of the culture to 1 mL 50% glycerol (aqueous), snap freezing in liquid nitrogen and then storing at -80 ^{o}C. Each glycerol stock was prepared from a single colony.

4.9.10.2 ^{19}F NMR of Qβ-F with E. coli C-3000

The growth media (15 mL) was inoculated with a glycerol stock of *E. coli* C-3000. The culture was incubated at 37 ^{o}C until OD_{600} = 0.52 (≈6 hours). Two 1 mL aliquots of the culture were removed and the cells were harvested by centrifugation (400 g, 20 minutes, room temperature). The cells were resuspended in tryptone medium (43% D_2O). 250 µL of the cell slurry was added to each of two NMR tubes. 250 µL of Qβ-F VLPs (1.0 mg/mL) was added to one NMR tube containing cells. As a control, to the other NMR tube containing cells was added 250 µL growth medium (43% D_2O). The temperature of the NMR bore was set to 37 ^{o}C and a ^{19}F NMR spectrum was taken after the sample was allowed to equilibrate in the bore for 10 minutes.

1000 scans were measured for each sample. No ^{19}F NMR signal was observed for either the Qβ-F VLPs incubated with *E. coli* nor the control of the Qβ-F VLP without *E. coli*. As a positive control, to ensure the lack of signal was not due to an insufficient number of scans, the sample of Qβ-F used to record spectrum Fig. 4.14 B (i.e. treated with SDS and DTT) was measured. The concentrations of the samples were similar (control sample: 0.07 mM; Fig. 4.14 B: 0.09 mM). A signal was observed in the positive control spectrum (run for 1000 scans).

The Qβ-F/*E. coli* sample was run overnight (24 000 scans) at 37 ^{o}C, no signal was observed.

5. *In vivo* studies of Qβ-F VLP

5.1 Introduction

Since the ultimate application proposed for the Qβ-F virus-like particles (VLPs) is *in vivo* imaging, initial studies to investigate its tissue biodistribution were required. Viral nanoparticles for use in medicine are currently extensively researched, and must undergo tests to determine whether they are safe to use in animals and what pharmacological properties they display.[1–3] Modifications are often made after these initial tests, for example PEGylation to reduce nonspecific interactions with endogenous biomolecules and increase the circulation time.[2,4,5]

The only published experiments on the biodistribution of Qβ involved Gd(DOTA)-labelled Qβ and used atomic emission spectroscopy to measure the concentration of labelled virion in tissues by measuring the gadolinium concentration.[6] Here, the ^{19}F tag on the Qβ nanoparticle was exploited to detect the Qβ-F protein by using ^{19}F NMR. This should allow determination of the biodistribution of Qβ-F since any ^{19}F signal detected from *ex vivo* samples must originate from the Qβ-F administered (since no ^{19}F occurs in the body).

5.2 *In vivo* tissue biodistribution of Qβ-F VLP

Male Wistar rats (200 g, Charles River) were injected with a solution of Qβ-F VLPs (200 μL, 90 ng). After 240 minutes, the rats were euthanized and dissected. These *in vivo* experiments were carried out by Dr Daniel Anthony. To detect the Qβ-F VLPs in the organs, two different analytical techniques which have been previously used in this work were used: sucrose density gradient with agarose gel electrophoresis, and ^{19}F NMR.

First, each organ was homogenized in 700 μL of TBS for 5 minutes before removing the debris by centrifugation and subsequently subjecting the supernatant to sucrose density gradient ultracentrifugation, to separate potentially present Qβ-F VLPs from organ proteins

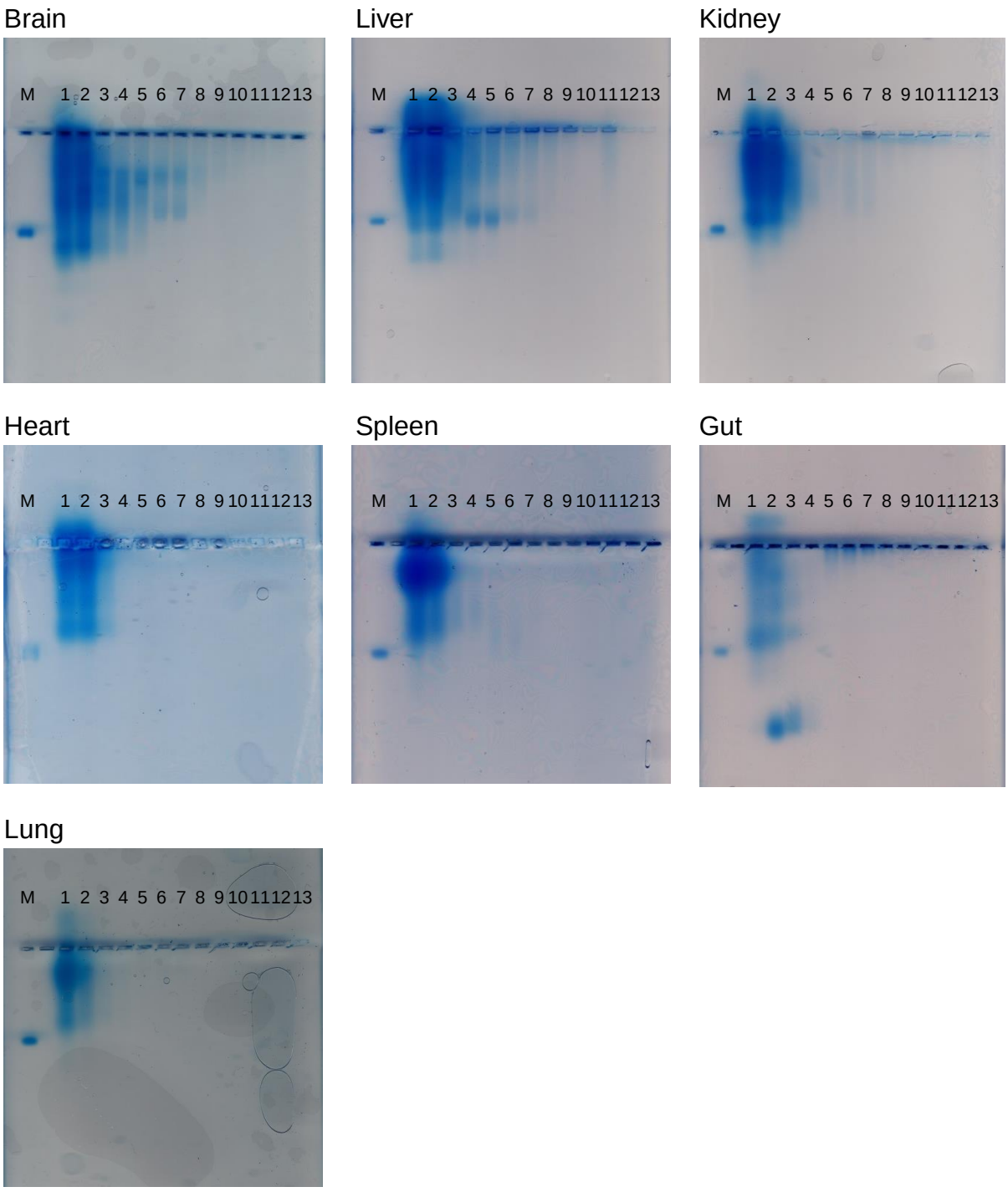


Figure 5.1 *Agarose gel electrophoresis (1% w/v agarose in TAE buffer solution). Sucrose density gradient fractions of homogenized rat organs as labelled. M = Qβ-Met VLP; 1 – 13 = sucrose gradient fractions, where 1 = top fraction and 13 = bottom fraction. The gels were stained with Instant*Blue™ *(coomassie blue solution).*

and other macromolecules. The gradients were fractionated and the fractions analysed by native agarose gel electrophoresis (Fig. 5.1). A sample of purified Qβ-Met VLPs, which have been shown to co-migrate with Qβ-F VLPs (see Chapter 3, Fig. 3.2), was run on each agarose gel as a marker for detection of Qβ-F VLPs. The agarose gel for the purified liver extract

clearly shows Qβ particles present in fractions 4 and 5 (and in 6 and 7, albeit at lower concentrations). No bands corresponding to the Qβ-F VLPs were observed in the agarose gels of the other organs.

In addition to its co-migration with Qβ-Met in the agarose gel, further evidence that Qβ was present in the liver is provided by the sedimentation coefficient of the species detected in fractions 4 and 5 (i.e. sucrose density gradient fraction number that the protein occurs in) of the homogenized liver, which is the same as that of Qβ particles purified after Qβ protein production by *E. coli* (see Chapter 2). If these bands are indeed Qβ-F VLP, they should show a gel shift after reducing treatment corresponding to reduction to the monomer (as observed in Chapter 3, Fig. 3.2). To test for this, the sucrose gradient fractions of the homogenized liver were treated with reducing SDS-PAGE loading buffer and heated at 90 ^{o}C for 10 minutes. The resulting agarose gel is shown in Fig. 5.2 and clearly shows no bands at the position expected for Qβ-F VLP in fractions 4, 5, 6, and 7 (proposed to contain Qβ-F VLPs) and instead bands corresponding to the Qβ monomer are observed.

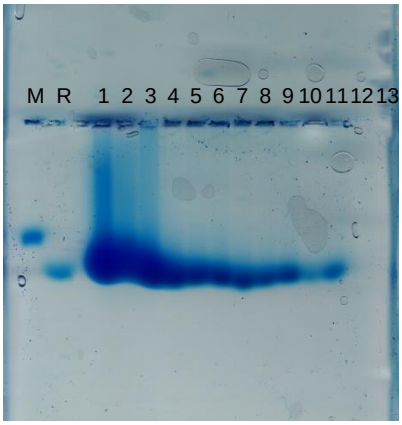


Figure 5.2 *Agarose gel electrophoresis (1% w/v agarose in TAE buffer solution): Sucrose density gradient fractions of homogenized liver after treatment with reducing SDS-PAGE loading dye and heating at 90 ^{o}C for 10 minutes. Qβ-Met VLP non-denatured (M) and after treatment with reducing SDS-PAGE loading dye and heating at 90 ^{o}C for 10 minutes (R) are also shown for reference. The gels were stained with Instant*Blue™ *(coomassie blue solution).*

To further confirm that the bands detected in fractions 4 to 7 by agarose gel electrophoresis of the homogenized liver did indeed correspond to Qβ-F VLP, these fractions

were analysed by ^{19}F NMR. Fractions 4 to 7 were pooled, sucrose was removed using a PD-10 column, and the sample was concentrated using a VivaSpin 500. The ^{19}F NMR of the resulting solution is show in Fig. 5.3 A. No peak was observed at the expected chemical shift of the SCF_3 group or the $S(O)_2CF_3$ group. SDS was then added to the sample (final concentration 0.17 M) and a ^{19}F NMR spectrum was recorded (Fig. 5.3 B). Finally, DTT was added to the sample (final concentration 10 mM) and the sample was incubated at 37 ^{o}C for 30 minutes before recording another ^{19}F NMR spectrum (Fig. 5.3 C).

No signal was observed at the chemical shift expected for the SCF_3 group after treatment with SDS or after subsequent incubation with DTT. However, after incubation of the sample with DTT, a peak was observed at -78.24 ppm. This chemical shift corresponds to that of the peak assigned to the sulfone of the trifluoromethionine residue (see Chapter 4). This provides robust evidence for the presence of Qβ-F in the liver sample. Furthermore, the fact that a peak was only observed after incubation with DTT at 37 ^{o}C shows that the Qβ-F VLP remained intact *in vivo*. This result shows that the Qβ-F VLP is stable *in vivo*, at least for the duration *in vivo* used in this investigation (240 minutes). *In vivo* stability of the Qβ-F VLP is important for its intended application as an imaging agent, since uncontrolled disassembly, before reaching the desired biological target, would lead to false positive signals.

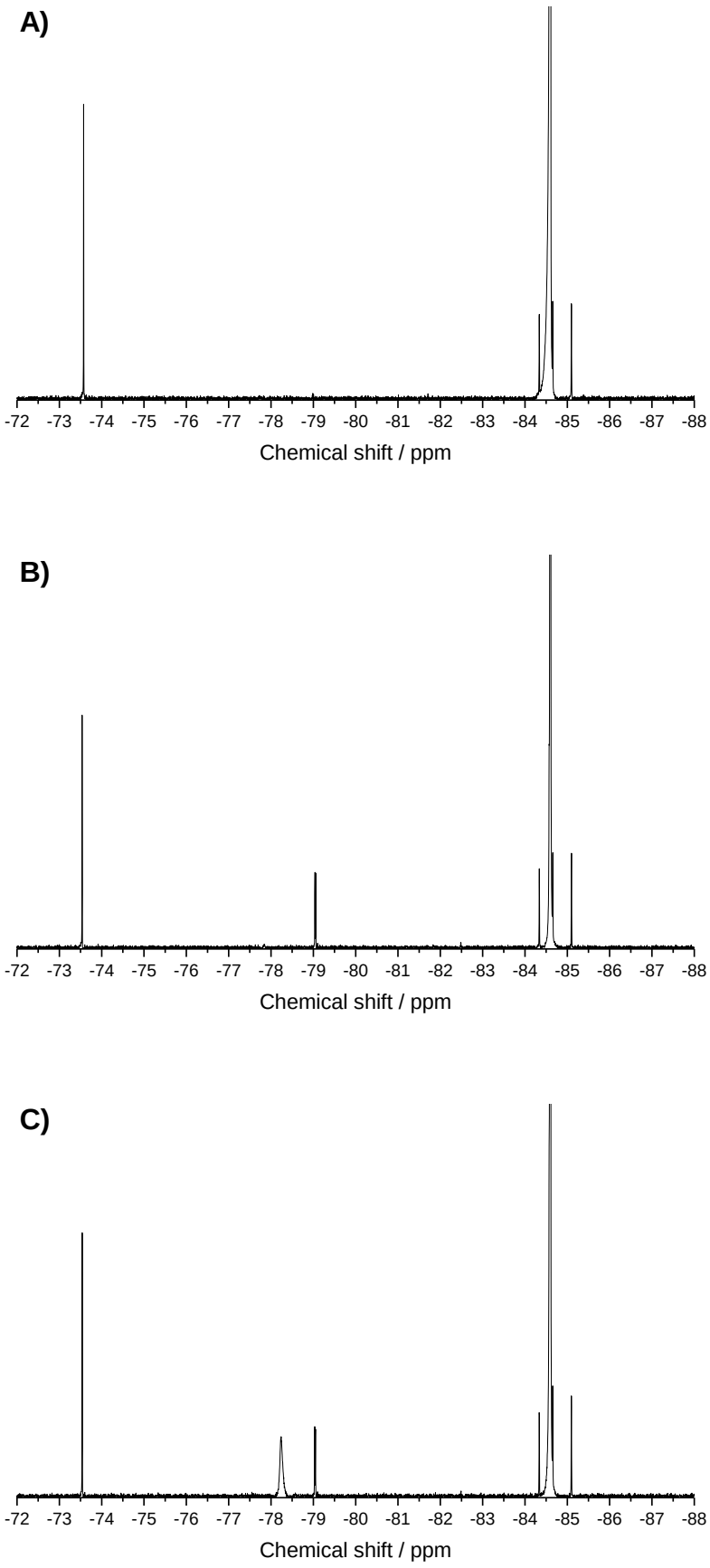


Figure 5.3 *^{19}F NMR spectra of collected sucrose density gradient fractions of homogenized liver A: in TBS, B: after treatment with 0.17 M SDS, C: after adding 10 mM DTT then heating at 37 ^{o}C for 30 minutes following treatment with 0.17 M SDS. The spectra are referenced to 1,1,1-trifluoroacetone δ_F = -84.60 ppm.*

5.3 Conclusions

Together, the sucrose density gradient and agarose gel electrophoresis, and ^{19}F NMR results show that after intravenous injection of Qβ-F VLP into rats, the Qβ-F VLPs accumulate in the liver. This conclusion agrees with data obtained by Prasuhn *et. al.* for Gd(DOTA)-labelled Qβ[6] and by Singh *et. al.* for the plant VLP, CPMV.[7]

Furthermore, the results from the sucrose gradient and agarose gel show that the Qβ-F VLP remained intact, i.e. it was not degraded *in vivo*. This is very promising when considering the utility of Qβ-F VLP as an *in vivo* off-to-on imaging agent as *in vivo* degradation would lead to false positive results. However, additional experiments should be carried out recovering the VLP at differing times post-injection. This would establish the *in vivo* stability half-life of Qβ-F VLP.

5.4 References

1. Wu, M. *et al.* Biobehavior in Normal and Tumor-Bearing Mice of Tobacco Mosaic Virus. *Biomacromolecules* **14,** 4032–4037 (2013).

2. Wen, A. M., Rambhia, P. H., French, R. H. & Steinmetz, N. F. Design rules for nanomedical engineering: from physical virology to the applications of virus-based materials in medicine. *J. Biol. Phys.* **39,** 301–325 (2013).

3. Bruckman, M. a *et al.* Biodistribution, pharmacokinetics, and blood compatibility of native and PEGylated tobacco mosaic virus nano-rods and -spheres in mice. *Virology* **449,** 163–173 (2014).

4. Koudelka, K. J. *et al.* Lysine Addressability and Mammalian Cell Interactions of Bacteriophage λ Procapsids. *Biomacromolecules* **14,** 4169–4176 (2013).

5. Steinmetz, N. & Manchester, M. PEGylated Viral Nanoparticles for Biomedicine: The Impact of PEG Chain Length on VNP Cell Interactions In Vitro and Ex Vivo. *Biomacromolecules* **10,** 784–792 (2009).

6. Prasuhn, D. E. *et al.* Plasma Clearance of Bacteriophage Qβ Particles as a Function of Surface Charge. *J. Am. Chem. Soc.* **130,** 1328–1334 (2008).

7. Singh, P. *et al.* Bio-distribution, toxicity and pathology of cowpea mosaic virus nanoparticles in vivo. *J. Control. Release* **120,** 41–50 (2007).

5.5 Experimental procedures

Animal experiments were carried out by Dr Daniel Anthony at the Department of Pharmacology, University of Oxford with UK Home Office Approval (licence 30/3076).

A solution of purified Qβ-F virus-like particle (VLP) of concentration 0.98 mg/mL (100 µL) was diluted with sterilised saline solution (100 µL). The resulting solution was intravenously injected into the tail veins of male rats (200 g, Wistar (Charles River)). After 240 minutes, the rats were euthanized and dissected. The following organs were isolated, placed in 1.5 mL Eppendord tubes and immediately stored at -80 ^{o}C: gut, liver, heart, brain, kidney, lung, spleen.

Preparation of the samples and sample analysis was carried out by Mr Matthew Robinson.

Each organ was thawed separately, combined with 0.7 mL of TBS buffer, and homogenized with a hand-held homogenizer for 5 minutes. The samples were then centrifuged to collect tissue debris (14 000 *g*, 20 minutes, 4 ^{o}C). The supernatant was carefully removed with a 1 mL pipette and transferred to a fresh 1.5 mL Eppendorf tube.

Each supernatant was loaded onto a sucrose density gradient (5% - 20% 1 mL layers, 25% - 40% 2 mL layers) and centrifuged for 3 hours at 37 600 rpm (178 530 *g*) at 4 ^{o}C in a Beckman ultracentrifuge. The acceleration was set to maximum and the deceleration was set to "No brake". The gradients were subsequently fractionated; 1 mL fractions were collected in Eppendorf tubes and analysed by native agarose gel electrophoresis: 500 mg agarose (electrophoresis grade) was dissolved in 50 mL 1 X TAE buffer. The solution was microwaved for ~ 2 minutes, poured into a gel tray and allowed to set. For each set of sucrose gradient fractions (i.e. each organ), a sample of intact Qβ-Met VLP was loaded. One lane was left empty before loading samples of consecutive fractions. The agarose gels were run at 60

volts for 1 hour. The resulting gels were stained with coomassie blue and are shown in Fig. 5.1

6. Summary and discussion

This thesis describes the design and development of a new ^{19}F MR imaging agent. ^{19}F MRI has a significant advantage over ^{1}H MRI since no covalently bound ^{19}F occurs naturally *in vivo* and so the only signal detected in a ^{19}F MR image is that from the administered ^{19}F-labelled agent. However, the principle disadvantage of MRI, that of low sensitivity, is still an issue with ^{19}F MRI as with conventional ^{1}H MRI. Hence, at the outset of this research, the primary goal was to develop a highly sensitive ^{19}F imaging agent.

The bacteriophage Qβ virus-like particle (VLP), a multimeric protein construct with iscosahedral symmetry, was used as the scaffold for the imaging agent. Qβ was chosen for two main reasons: firstly it comprises 180 coat proteins, thus allowing a large number of ^{19}F atoms to be loaded onto the VLP; and secondly, because the coat proteins are symmetrically related in the VLP and so the ^{19}F group of each modified Qβ monomer should have the same chemical shift. Using a protein scaffold such as Qβ could also allow further modification *via* other residues to introduce antibodies or glycans for cell targeting, and/or fluorescent labels for microscopy imaging.[1–7]

The Qβ VLP could have been modified chemically to bear ^{19}F atoms by chemically reacting either the lysine or cysteine residues on the surface of the VLP with fluorinated tags.[8–10] However, it was decided to use genetic incorporation of a fluorinated unnatural amino acid to introduce the ^{19}F-group since this would allow greater selectivity than chemical modification and so would lead to ^{19}F groups being more regularly disposed over the VLP surface. The sense codon reassignment method has been used extensively to incorporate unnatural amino acids, including fluorinated amino acids, into many different proteins (see Chapter 1 Section 1.4.2). Here, it was envisaged that this method could be used to incorporate a fluorinated amino acid into the Qβ coat protein thus producing the desired fluorinated imaging agent. The wild-type Qβ coat protein does not contain methionine residues, hence a

mutant gene sequence (K16M), containing a single methionine residue, was used so that the sense codon reassignment method would incorporate a single fluorinated methionine analogue, site-specifically, into each Qβ coat protein. Trifluoromethionine (TFM) was chosen as the fluorinated methionine analogue since it bears three chemically equivalent ^{19}F atoms, hence giving rise to a single ^{19}F NMR peak.

Previous reports of incorporation of TFM into proteins in the literature have made apparent the difficulty of achieving high levels of incorporation with this particular unnatural amino acid.[11,12] Indeed, the first attempt to incorporate TFM into the Qβ protein in this work gave an incorporation level of 72%. In an attempt to increase this incorporation level, the protocol for incorporation of TFM into Qβ K16M was optimised. The level of incorporation of TFM into Qβ for each set of conditions was determined by calculating the ratio of the areas of the Qβ-Met peak and the Qβ-F peak in the mass spectrum.

Firstly, the effect of different concentrations of L-methionine (L-Met) in the growth medium on the growth of auxotrophic *E. coli* cells was investigated. It was decided that 0.1 mM was the lowest concentration of L-Met which could be used whilst still amassing an acceptable quantity of viable *E. coli* cells for protein production. The concentration of DL-TFM in the growth medium was next investigated since it was expected that a higher concentration of DL-TFM in the growth medium would result in a higher incorporation level of the amino acid into the Qβ protein. Toxicity effects of DL-TFM on *E. coli* cells has been reported in the literature,[11] and so a maximum concentration of DL-TFM which could be tolerated by the *E. coli* cells was expected. From the concentrations tested in this work, 6.5 mM DL-TFM was selected as the maximum concentration of DL-TFM to be used. Finally, the duration for which protein production was carried out was varied based on the assumption that the pool of residual Met in the *E. coli* would be depleted before that of TFM and so any protein production beyond this point would incorporate TFM. Protein production times of 10

hours, 12 hours, 14 hours and 16 hours were tested. Interestingly, when protein production was allowed to proceed for 14 hours and 16 hours no protein was isolated. It was proposed that the lack of protein produced was due to degradation of the protein at these extended times. An incorporation level of 96% was achieved using the optimised protocol, which shows a substantial improvement on previously reported incorporations of TFM.[11,12] Furthermore, this is first time a fluorinated amino acid has been incorporated into a virus assembly.

As explained above, it was initially envisaged that a single ^{19}F NMR peak be observed from Qβ-F VLP, with each TFM residue contributing to the same peak due to the equivalence of the TFM residues through symmetry. However, initial ^{19}F NMR measurements of the Qβ-F VLP showed no signals. Guided by the work of Takamatsu and Iso, who worked on wild-type Qβ,[13] Qβ-F VLP was disassembled using the denaturant sodium dodecyl sulphate (SDS), and this disassembled form was shown to give ^{19}F NMR signals. A signal was observed at -39.04 ppm, as expected for the TFM residue, and another signal also observed at ~ -78 ppm, proposed to originate from protein assemblies of Qβ-F in which the TFM residue had oxidised to its corresponding sulfone. After addition of SDS, the Qβ-F sample was heat treated with dithiothreitol (DTT), which yielded an additional ^{19}F NMR peak at -39.10 ppm, arising from the monomer of Qβ-F. ^{19}F DOSY was used to confirm that the broad peak at -39.04 ppm and the sharper peak at -39.10 ppm corresponded to Qβ-F multimers and Qβ-F monomers, respectively. Additional peaks were observed in the region of -39 ppm at a disassembly stage between partial disassembly and complete reduction to the monomer. These are likely to correspond to different multimeric states of the Qβ-F VLP (hexamer, pentamer, and dimer). Hexamer and pentamer peaks were tentatively assigned based upon their respective FWHH values. These results show the potential of fluorinating virus monomers to follow the disassembly of a virus by ^{19}F NMR. Future work

should focus on characterizing the peaks other than that of the monomer in terms of the multimeric state of the protein species.

Finally, preliminary *in vivo* experiments were carried out to investigate the stability and biodistribution of Qβ-F VLP with a view to establish its potential as a medical imaging agent. "Spike-recovery" experiments were performed by injecting rats with Qβ-F VLP and analysing the homogenized organs. Qβ-F VLP was found to accumulate predominantly in the liver. Importantly, it was shown that they remained intact *in vivo*. This is an important result in terms of using Qβ-F VLP as an off-to-on imaging agent, since non-selective degradation of Qβ-F VLP would lead to false positive ^{19}F MRI signals.

The ^{19}F NMR results presented in Chapter 4, combined with the preliminary *in vivo* experiments of Chapter 5 show that there is great potential in the Qβ-F VLP construct developed in this thesis for a novel medical imaging agent. The next key questions to investigate are: 1) Can Qβ-F VLP be targeted to cells presenting specific biomarkers for diseases? 2) On reaching its target site, can Qβ-F VLP be made to disassemble *in vivo* and hence give a measurable signal?

M. G. Finn and co-workers have shown that Qβ can be targeted to the transferrin receptor by conjugating the glycoprotein transferrin onto the surface of Qβ.[14] Furthermore, it was shown, by fluorescence microscopy, that the Qβ was internalised into monkey kidney epithelial cells. However, these results do not show whether Qβ remained intact once internalized, or became disassembled. Hence, one possible starting point for work following the current thesis would be to modify Qβ-F VLP with transferrin, incubate the resulting construct with cells bearing the transferrin receptor and measure the ^{19}F NMR of the resulting mixture. The development of a ^{19}F NMR signal would indicate the disassembly of the particle. If successful, *in vivo* ^{19}F MRI with mouse models with and without transferrin receptors could be carried out.

References

1. Wang, Q., Lin, T., Tang, L., Johnson, J. E. & Finn, M. G. Icosahedral Virus Particles as Addressable Nanoscale Building Blocks. *Angew. Chem. Int. Ed. Engl.* **41,** 459–462 (2002).

2. Wang, Q., Kaltgrad, E., Lin, T., Johnson, J. E. & Finn, M. G. Natural Supramolecular Building Blocks: Wild-Type Cowpea Mosaic Virus. *Chem. Biol.* **9,** 805–811 (2002).

3. Lewis, J. D. *et al.* Viral nanoparticles as tools for intravital vascular imaging. *Nat. Med.* **12,** 354–360 (2006).

4. Shriver, L. P., Koudelka, K. J. & Manchester, M. Viral nanoparticles associate with regions of inflammation and blood brain barrier disruption during CNS infection. *J. Neuroimmunol.* **211,** 66–72 (2009).

5. Kaltgrad, E. *et al.* On-virus construction of polyvalent glycan ligands for cell-surface receptors. *J. Am. Chem. Soc.* **130,** 4578–4579 (2008).

6. Yin, Z. *et al.* Boosting Immunity to Small Tumor-Associated Carbohydrates with Bacteriophage Qβ Capsids. *ACS Chem. Biol.* **8,** 1253–1262 (2013).

7. Pokorski, J. K., Hovlid, M. L. & Finn, M. G. Cell Targeting with Hybrid Qβ Virus-Like Particles Displaying Epidermal Growth Factor. *ChemBioChem* **12,** 2441–2447 (2011).

8. Luchette, P. A., Prosser, R. S. & Sanders, C. R. Oxygen as a Paramagnetic Probe of Membrane Protein Structure by Cysteine Mutagenesis and 19F NMR Spectroscopy. *J. Am. Chem. Soc.* **124,** 1778–1781 (2002).

9. Oxenoid, K., Sönnichsen, F. D. & Sanders, C. R. Topology and Secondary Structure of the N-Terminal Domain of Diacylglycerol Kinase. *Biochemistry* **41,** 12876–12882 (2002).

10. Gerig, J. T. Fluorine NMR of Proteins. *Prog. Nucl. Magn. Reson. Spectrosc.* **26,** 293–370 (1994).

11. 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 19F NMR. *Biochemistry* **36,** 3404–3416 (1997).

12. Holzberger, B., Rubini, M., Möller, H. M. & Marx, A. A Highly Active DNAPolymerase with a Fluorous Core. *Angew. Chem. Int. Ed.* **49,** 1324–1327 (2010).

13. Takamatsu, H. & Iso, K. Chemical evidence for the capsomeric structure of phage Qβ. *Nature* **298,** 819–824 (1982).

14. 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).