Structural and functional analysis of NPA and FAR nematode proteins

by

Claire Patricia Martin

UNIVERSITY
of
GLASGOW

A thesis submitted for the degree of Doctor of Philosophy in the
Department of Chemistry
University of Glasgow
July 2003

Declaration

This thesis has been written in accordance with the University of Glasgow regulations, has not been presented for a degree at any other university and is original except where indicated otherwise by reference in the text. The work contained within is the author's own except where work was done in collaboration as indicated.

Signed	Date
C P Martin	21/7/03
Claire Patricia Martin	July21st 2003

Acknowledgements

Firstly, I would like to thank my supervisor, Dr Andy Freer, for all his help and support over the last few years. Thanks for your friendship and patience.

Thanks to all the PX lab: Colin, Alan, Ally, Tall Paul, Rhona, Chris, Neil, Adrian, Dave, The boy, Derek, Kirsty and Isobel who have all played their own parts in many different ways. Not forgetting Dina, Paul, Dominic, Russell and Neil.

I would also like to thank Nicola Meenan, for her help and friendship. Prof. Alan Cooper and Prof. Malcolm Kennedy, for their knowledge and experience, it was much appreciated.

Finally, I wish to thank my Mum, Dad, Grandparents, Richard and Andrew for their support emotionally and financially over the last four years. I hope I have made you proud.

Dedicated to Papa Fred who always believed in me.

Summary

Parasitic nematodes cause a wide range of diseases in humans, including river blindness caused by *Onchocerca volvulus* and elephantiasis caused by *Ascaris lumbricoides* (www.medic-planet.com). The greatest proportion of agricultural loss is in potato crops and is caused by potato cyst parasitic nematodes including *Globodera pallida.*

For survival, nematodes need to sequester fatty acids (usually from the host) and therefore must produce proteins capable of binding essential fatty acids (Kennedy *et al*, 1995). The nematode proteins studied in this project have been classified into two new types of fatty acid and retinoid binding proteins. Firstly, nematode polyprotein allergens/antigens (NPA), an example of this class of protein is the *Ascaris* body-fluid allergen (ABA-1) of the parasitic nematode *Ascaris.* The second class are known as fatty acid and retinol binding proteins (FAR), an example of this class is OvFAR1 from the parasitic nematode *Onchocerca volvulus.*

The nematode proteins are believed to be a cause of severe allergic-type reactions in nematode infections (Christie *et al*, 1990). Since the pathology of these allergic reactions dictate the involvement of the immune response system, it is thought that these classes of protein are potential candidates for vaccines containing recombinant protein antigens (McDermott, 1999).

ABA-1 was purified from *Ascaris suum* (pABA-1) and produced recombinantly (rABA-1). Fluorescence studies using DAUDA showed that rABA-1 is similar in binding to the pABA-1. Secondary structure studied using circular dichroism indicated rABA-1 and pABA-1 have similar secondary structures. Crystallisation of pABA-1 and rABA-1 using a large range of conditions failed to produce crystals suitable for diffraction and was investigated using dynamic light scattering (DLS), which suggested a dimer structure.

GpFAR1, OvFAR1, CeFAR2 and CeFAR3 proteins from the FAR family of proteins were overexpressed and purified. Ligand binding properties of rOvFAR1, rCeFAR2 and rCeFAR3 established that the CeFARs from *Caenorhabditis elegans* were good functional homologues of rOvFAR1, a major antigen of *Onchocerca volvulus*. Binding DAUDA suggested that the binding site is in an apolar environment and isolated from solvent. Competition between retinol and oleic acid indicated that the binding site was interactive or congruent. The FAR proteins also showed a preference for fatty acids with chain length C14-C18, investigated by competition of the DAUDA-FAR protein complex. This would indicate that the recombinant proteins are behaving as they do *in vivo*.

The biophysical properties of rOvFAR1, rCeFAR2, rCeFAR3 and rGpFAR1 were investigated. Circular dichroism suggested that these FAR proteins have essentially the same secondary structure. MALDI analyses suggested that fatty acids are bound to the proteins and were persistently bound after a multistep purification, possibly sequestered from *E. coli* during overexpression. DLS measurements indicated that the proteins were heterogeneous.

The objective of the project was to characterize a number of these NPA and FAR proteins with the ultimate goal of determining their 3-D structure. Initially, seven proteins were investigated (ABA-1, GpFAR1, OvFAR1, CeFAR1, CeFAR2, CeFAR3, CeFAR5) and only the more viable and attainable were carried forward for further characterization and attempts at crystallisation. However, rGpFAR1 was the only FAR protein that grew protein crystals, diffraction of these needle crystals showed that they were twinned or hollow.

For future crystallisations to be successful, new purification methods are required for the FAR proteins.

Abbreviations

Crystallography Terms

Å	Angstrom
a, b, c	Real space unit cell dimensions
CCD	Charge coupled device

Chemical terms and species

C. elegans	*Caenorhabditis elegans*
E. coli	*Eschrichia coli*
IPTG	Isopropyl-β-D-thiogalactoside
LB	Luria broth
EDTA	Ethylenediaminetetraacetic acid
PAGE	Polyacrylamide Gel Electrophoresis
SDS	Sodium dodecyl sulphate
Tris-HCl	Tris(hydroxymethyl)methylamine
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
PIPES	Piperazine-N,N-bis(2-ethane sulfonic acid)
DAUDA	11-((5-dimethylaminonaphthalene-1-sulfonyl)amino)undecanoic acid
DACA	Dansyl-D,L-α-aminocaprylic acid
PBS	Phosphate-buffered saline
Ni-NTA	Nickel nitrilotriacetic acid
PEG	Polyethylene glycol
MPD	(±)-2-Methyl-2,4-pentanediol

Proteins

NPA	Nematode polyprotein allergen/antigen
ABA-1	*Ascaris* body-fluid allergen-1
rABA-1	Recombinant *Ascaris* body-fluid allergen-1
pABA-1	Parasitic *Ascaris* body-fluid allergen-1
ABF	*Ascaris* body fluid
gp15/400	Polyprotein of *Brugia malayi*
DvA-1	*Dictyoculus viviparus* allergen-1
TBA-1	*Toxocara* body-fluid allergen
FABP	Fatty acid binding protein
FAR	Fatty acid and retinol binding protein
OvFAR1	*Onchocerca volvulus* fatty acid and retinol binding protein-1
rOvFAR1	Recombinant *Onchocerca volvulus* fatty acid and retinol binding protein-1
Bm20	*Brugia malayi*-20
GpFAR1	*Globodera pallida* fatty acid and retinol binding protein-1

rGpFAR1	Recombinant *Globodera pallida* fatty acid and retinol binding protein-1
CeFAR1	*Caenorhabditis elegans* fatty acid and retinol binding protein-1
CeFAR2	*Caenorhabditis elegans* fatty acid and retinol binding protein-2
CeFAR3	*Caenorhabditis elegans* fatty acid and retinol binding protein-3
CeFAR5	*Caenorhabditis elegans* fatty acid and retinol binding protein-5
rCeFAR2	Recombinant *Caenorhabditis elegans* fatty acid and retinol binding protein-2
rCeFAR3	Recombinant *Caenorhabditis elegans* fatty acid and retinol binding protein-3
HNF4	Hepatocyte nuclear factor 4

Miscellaneous

IMAC	Immobilised-metal affinity chromatography
pI	Protein isoelectric point
K_d	Dissociation constant
ESRF	European Synchrotron Radiation Facility
SRS	Synchrotron Radiation Source
NMR	Nuclear magnetic resonance spectroscopy
GOR	Garnier, Osguthorpe and Robson secondary structure prediction
PHD	Protein secondary structure prediction program
CD	Circular dichroism
DLS	Dynamic light scattering
MALDI-TOF	Matrix assisted laser desorption ionisation-time of flight
MS	Mass spectrometry
DSC	Differential scanning calorimetry
IEF	Isoelectric focussing electrophoresis
UV	Ultra violet
VIS	Visible

Contents

List of figures

Chapter One: Introduction

Chapter Two: Materials and methods

Chapter Three: Comparison of parasite-derived ABA-1 and recombinant ABA-1 proteins

Chapter Four: OvFAR1, CeFAR2 and CeFAR3. Expression, purification and fluorescence studies

Chapter Five: GpFAR1 overexpression and purification

Chapter Six: Matrix Assisted Laser Desorption Ionisation- Time Of Flight, Circular Dichroism and Dynamic Light Scattering analysis of the FAR proteins

Chapter Seven: Crystallisation of FAR proteins

List of tables

Chapter 1 - Introduction

1.1 Parasitic nematodes

Parasitic nematodes are multicellular, eukaryotic organisms with highly specialised mechanisms, which allow them to evade the immune system of their host and survive for long periods in the host. They are predominantly found in the tropics and subtropics. Parasitic nematodes cause a wide range of diseases in humans, including river blindness caused by *Onchocerca volvulus* and elephantiasis caused by *Ascaris lumbricoides* and other parasitic nematodes (www.medic-planet.com). In developed countries they cause loss through chronic or fatal diseases and substantial economic losses of up to 60 billion pounds worth of crops each year. The greatest proportion of agricultural loss is in potato crops and is caused by potato cyst parasitic nematodes including *Globodera rostochiensis* and *Globodera pallida* (www.inra.fr). Parasitic nematodes also cause problems in livestock.

To survive, there is a requirement for these nematodes to sequester fatty acids (usually from the host) and therefore the nematodes must produce proteins capable of binding these essential fatty acids (Kennedy *et al*, 1995a). For the nematodes studied in this project the proteins have been classified into two new types of fatty acid and retinoid binding proteins called nematode polyprotein allergens/antigens (NPA) and fatty acid and retinol binding proteins (FAR) (Garofalo *et al*, 2002).

It is currently believed (McDermott, 1999) that both types of protein play major roles in the metabolism of the parasites and that they are involved in the transport and sequestration of hydrophobic anti-nematode ligands. The ligands that these proteins bind and presumably protect from degradation are of great importance to the parasitic nematodes because they cannot synthesise these complex lipids themselves and must therefore acquire them from their host.

The first of these newly classified proteins, known as nematode polyprotein allergens/antigens (NPAs) are synthesised as large polyproteins that are

subsequently cleaved at regularly spaced sites to form multiple copies of fatty acid binding protein having a molecular mass between 14 and 15 kDa (Spence *et al*, 1993). An example of this class of protein is the Ascaris body-fluid allergen (ABA-1) of the parasitic nematode *Ascaris*, which sequesters fatty acids, including palmitic acid and stearic acid. These NPAs are extracellular proteins, abundant in the pseudocoelomic fluid of nematodes. This fluid is analogous to human blood, in that it bathes all the internal organs of the nematode. Therefore, it is likely that NPAs act as carriers responsible for organising the distribution of hydrophobic ligands within the parasite.

The second class of protein known as fatty acid and retinol binding (FAR) proteins are slightly larger in size, approximately 20 kDa (Bradley & Kennedy, 1999). Like NPAs they bind fatty acids and retinol, but differ in that they possess a higher affinity for the latter. They also differ in that they are not synthesised as polyproteins. An example of this class is OvFAR1 from the parasitic nematode *Onchocerca volvulus*. OvFAR1 is also an extracellular protein, which is particularly abundant in the body wall of the worm and its larvae, which develop within the host uteri.

The major interest in both the NPAs and FAR proteins, with respect to infections, is that they are secreted by the parasites and so could be involved in controlling the tissue environment or acting to counter immune defence reactions of the host. They are believed to be a cause of severe allergic-type reactions, which occur in nematode infections (Christie *et al*, 1990). Since the pathology of these allergic reactions dictate the involvement of the immune response system, it is thought that these classes of protein are potential candidates for vaccines containing recombinant protein antigens (McDermott, 1999).

1.2 Nematode polyprotein allergens/antigens (NPAs)

Nematode allergens were first discovered due to the hypersensitive response of the immune system against them. There are a large number of nematode allergens that are responsible for conditions as diverse as elephantiasis, filariasis and asthma. Elephantiasis is a disorder of the lymphatic system and is characterised by enlargement of body areas including limbs, torso and head (figure 1.1).

Figure 1.1 Elephantiasis of legs due to filariasis (www.gsk.com/filariasis)

The swelling of body areas is caused by an accumulation of fluid as a result of obstruction of lymph or blood flow. This blockage can occur due to bacterial infection or nematodes living in ducts or glands (www.elephantiasis.freeyellow.com). Proteins synthesised by these parasitic nematodes are potent allergens that can cause many diseases and are abundant in both the somatic and excreted/secreted material of the parasites (Britton *et al*, 1995). Allergens have been characterised from several different species of nematode (figure 1.2), including ABA-1 from *Ascaris lumbricoides* and

Ascaris suum, gp15/400 from *Brugia pahangi*, DvA-1 from *Dirofilaria immitis* and TBA-1 from *Toxocara canis*.

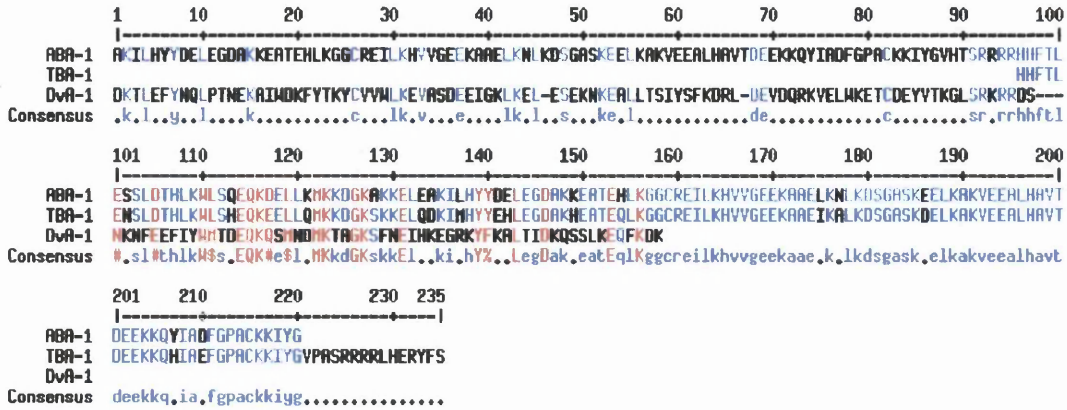

Figure 1.2 Sequence alignment of nematode polyprotein allergens. Accession numbers are Q06811, Q24702 and P49149 for ABA-1, DvA-1 and TBA-1 respectively. Red signifies 100% similarity and blue signifies 50% similarity.

Nematode allergens are synthesised as large precursors (200 - 400 kDa), which are proteolytically cleaved giving rise to multiple polypeptide units of molecular mass 14 - 15 kDa (Britton, 1995 *et al*; Poole *et al*, 1996). For this reason these "ladder proteins" are called "nematode polyprotein allergens/antigens" (NPAs). Production of polyproteins is a highly efficient means of producing proteins rapidly and economically. The large precursors are cleaved at proteolytic sites similar to the cleavage sites of endoproteinases of the subtilisin family (Barr, 1991), as the target sites are composed of basic amino acids, predominantly Arginine (figure 1.3). Repeating units within each NPA array also show variation in amino acid sequence. For example, *Dictyocaulus viviparus* polyprotein is composed of 12 different units that vary in size and may exhibit between 17 % and 98 % amino acid identity (Britton *et al*, 1995). The amino acid sequences of individual repeat units vary, although all NPAs contain conserved residues, Trp15, Gln20, Leu42, Cys64 and Cys120 (figure 1.7) (McDermott, 1999). Fluorescence studies suggest Trp15 is either buried internally in the structure or surrounded by several hydrophobic groups (McDermott, 1999). Cys64 and Cys120 are probable candidates for a disulphide bond which would stabilise the structure. Since individual parasites

can produce a different NPA, the structure-function of each NPA needs to be investigated individually.

D. immitis	...	Asn	Glu	**Arg**	**Arg**	**Lys**	**Arg**		Asn	Asp	His ...
B. malayi	...	Glu	Asp	**Arg**	**His**	**Lys**	**Arg**	Asp	Asn	His	Glu His ...
A. suum	...	Thr	Ser	**Arg**	**Arg**	**Arg**	**Arg**		His	His	Phe ...

Figure 1.3 Amino acid sequence of proposed cleavage site. Residues marked in bold indicate the cluster of 4 basic amino acids at which the polyprotein is cleaved. (Kennedy, 2000).

Structural evidence

The first fatty acid binding proteins were from a family of cytosolic proteins, which transport and protect hydrophobic compounds. They comprise 10 anti-parallel β-strands and 2 short α-helices surrounding the interior binding cavity in a β-barrel type structure (Banaszak *et al*, 1994). Multiple alignments of NPA sequences show no significant similarity to any protein of documented function on current databases and only match other NPAs (Kennedy *et al*, 1995b). Throughout this project no X-ray or NMR structures were available for any NPA protein. However, in November 2002 an initial structure from NMR data became available (N. Meenan, A. Cooper, personal communication).

Single gel filtration peaks correspond to dimers of DvA-1 and ABA-1 (Kennedy, 2000). Far UV circular dichroism studies analysed over 240-190 nm by the CONTIN procedure (Provencher & Glőckner, 1981) show the proteins to have a high α-helical signal (Kennedy, 2000). Secondary structure predictions of individual sequences of NPAs have been performed using GOR algorithms (Garnier *et al*, 1978), Chou and Fasman (Chou & Fasman, 1974) and PHD (Rost *et al*, 1994) and were found to agree with CD studies, as they predict that a single NPA unit is extensively helical, with little or no evidence of β-structure (Kennedy *et al*, 1995b).

Functional evidence

NPAs are fatty acid and retinoid binding proteins. The fluorescence emission of DAUDA is taken as a measure of the polarity of the binding site. Thus NPAs binding DAUDA with an increase in intensity and a blue shift in fluorescence indicates the ligands bind in an apolar environment (Kennedy *et al*, 1995a; McDermott, 1999). Table 1.1 shows lists of ligands found to bind or not bind to NPAs. There is only one binding site for each NPA molecule per ligand molecule. This was investigated by titrating the NPA molecule with DAUDA, with a dissociation constant (K_d) estimated to be $(8.8 \pm 0.5) \times 10^{-8}$ M (Kennedy *et al*, 1995a). Ligand competition experiments indicate that the binding site for fatty acids is the same as that for retinol and retinoic acid (Kennedy *et al*, 1995(a), 1995(b), 2000).

Binding	**No binding**
Oleic acid	Cholesterol
Palmitic acid	Tryptophan
Stearic acid	Caproic acid
Retinoic acid	Squalene
Retinol	Tocopherol
Oleoyl-CoA	Tocopherol acetate
Bilirubin	Succinyl CoA
Cis parinaric acid	2-methylbutyric acid
Trans parinaric acid	2-methylvaleric acid
Arachidonic acid	Bilverdin
Lysophosphatidic acid	Mebendazole
Lysophosphatidyl ethanolamine	Albendazole
Lysophosphatidyl choline	Thiabendazole
Platelet activating factor	Oxibendazole
Lysoplatelet activating factor	Piperazine
Leukotrienes B_4, C_4, D_4, E_4	Tetramisole
	Pyrantel
	DEC
	Levamisole

Table 1.1 Ligand binding by nematode polyprotein allergens/antigens.

NPAs are abundant in the pseudocoleomic fluid of nematodes, analogous to our blood as it bathes the organs of the body, and therefore it is likely that NPAs act as carriers of hydrophobic ligands within the parasite. The ligands they bind, and presumably protect from degradation, are essential to parasitic nematodes. This is because nematodes cannot synthesise these complex lipids themselves, and hence must acquire them from their host. Since NPA function has not yet been determined, their allergenicity remains a focus of attention (Kennedy, 2000). The NPAs could be potential drug targets, given that they appear to have no counterparts in vertebrates and their ligand binding sites have unusual properties.

1.2.1 NPAs specific to this project - *Ascaris* and ABA-1

The parasitic nematodes *Ascaris suum* and *Ascaris lumbricoides* are responsible for Ascariasis or roundworm infection, which is common in tropical and temperate climates (figure 1.4) where there is high humidity and low standards of hygiene and sanitation. About 26 % of the world population particularly children, are infected at any one time.

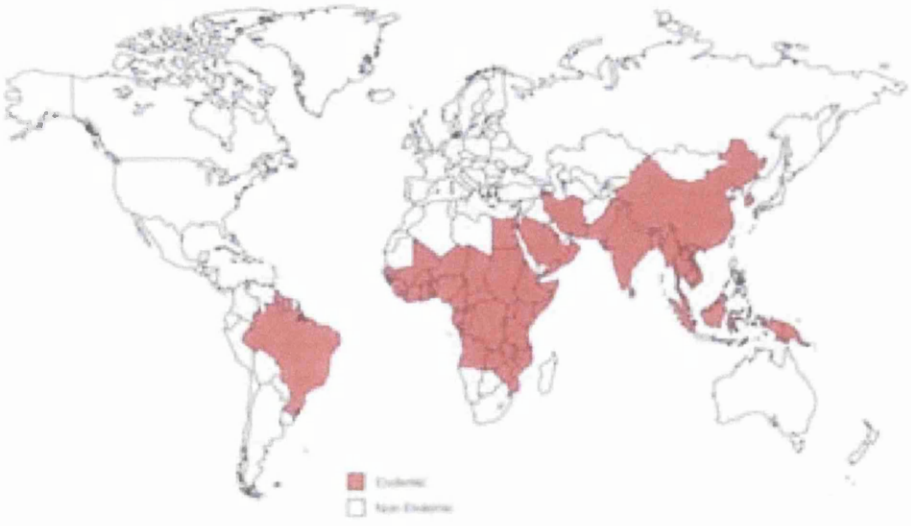

Figure 1.4 Map showing in red the countries affected by elephantiasis (www.gsk.com/filariasis)

Symptoms of infection include inflammation and haemorrhage of the liver and lungs due to an accumulation of dead larvae and tissue, which contribute to disease manifestations, including elephantiasis (see figure 1.1). The majority of infections show no symptoms, but the presence of even a few worms can be fatal. Although the adult nematodes are not normally long-lived, there is an annual cycle (figure 1.5) of infection coinciding with the lifecycle of Ascaris (Muller *et al*, 1975).

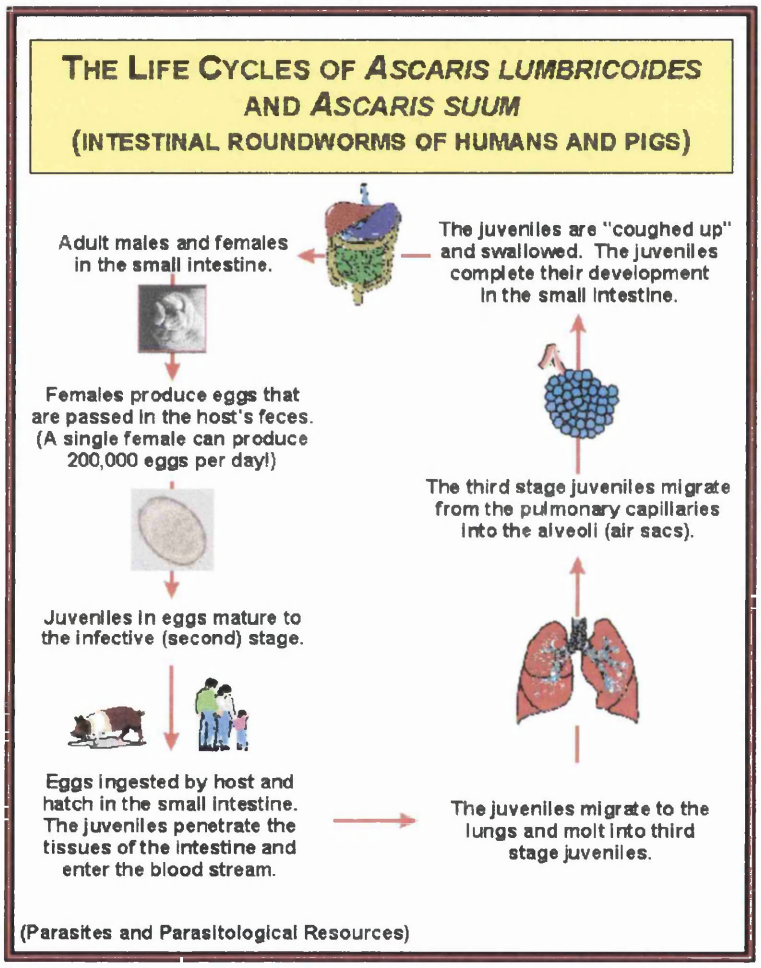

Figure 1.5 Life cycle of Ascaris (www.biosci.ohio-state.edu/~parasite/lifecycles/ascaris_lifecycle.html).

Fertilised eggs are passed in faeces. In the soil the first-stage larva in the egg moult and an egg containing second-stage larva develops, which is infective and can survive for up to 7 years. When the eggs are swallowed the larvae hatch in the duodenum or small intestine which are of ideal conditions, 37 °C and pH 7. The larvae then burrow into intestinal wall and are carried in lymph

or blood through the lungs where they moult again. Finally they travel through the oesophagus to the ileum. Mature female nematodes can survive for 1 – 2 years and lay more eggs for the cycle to repeat and with it the cycle of infection.

Within the context of infection, Ascaris Body-fluid Allergen (ABA-1) is the potent allergen protein of *Ascaris Lumbricoides* and *Ascaris suum.* The relationship of *Ascaris Lumbricoides* in man compared to the parasite *Ascaris Suum* in the pig is still unresolved. Though, the latter is usually regarded as a separate species (Muller *et al*, 1975). Purified ABA-1 from *Ascaris suum* was found to be potently allergenic in experimental animals (Tomlinson *et al*, 1989), and could therefore represent an important stimulus to the potentially lethal pulmonary and intestinal hypersensitivity reactions that arise in Ascariasis (Kennedy *et al*, 1995b). ABA-1 has been characterised as the most abundant protein in the pseudocoelomic fluid and in the somatic tissues of the adult Ascaris and it is released into the host's tissues by the invasive stages of the parasite.

ABA-1 structural predictions

ABA-1 behaves like all other NPAs and shows a single peak with a molecular weight of 27400 Da in gel filtration, yet SDS PAGE gives a single band at 14400 Da suggesting that ABA-1 may exist as a dimer (Kennedy *et al*, 1995b). DSC analysis revealed two major transitions, which is consistent with the existence of two distinct structural domains within the ABA-1 (Kennedy *et al*, 1995b; Christie *et al*, 1993). It is not known whether biochemical activity depends on the formation of a dimer. The non-glycoslated protein has a mass of 15.1 kDa, consists of 129 amino acids and has an isoelectric point of 5.2 (Christie *et al*, 1990). The polyprotein amino acid sequence consists of 395 amino acids (figure 1.6).

AKILHYYDEL EGDAKKEATE HLKGGCREIL KHVVGEEKAA
ELKNLKDSGA SKEELKAKVE EALHAVTDEE KKQYIADFGP
ACKKIYGVHT S**RRRR**HHFTL ESSLDTHLKW LSQEQKDELL
KMKKDGKAKK ELEAKILHYY DELEGDAKKE ATEHLKGGCR
EILKHVVGEE KAAELKNLKD SGASKEELKA KVEEALHAVT
DEEKKQYIAD FGPACKKIYG VHTS**RRRR**HH FTLESSLDTH
LKWLSQEQKD ELLKMKKDGK AKKELEAKIL HYYDELEGDA
KKEATEHLKG GCREILKHVV GEEKAAELKN LKDSGASKEE
LKAKVEEALH AVTDEEKKQY IADFGPACKK IYGVHTS**RRR**
RYHAEDGTDD IDGLAQSRQR RSGFFEKLID VFAFF

Figure 1.6 Amino acid sequence of polyprotein ABA, with the arginine tetrad where cleavage occurs highlighted in bold (www.ca.expasy.org/sprot/). Accession number for ABA-1 is Q06811.

The polyprotein ABA is cleaved into repeat units, which are not identical; adjacent repeats can differ by as much as 51 % of their amino acid residues (Kennedy, 2000).

HHFTL ESSLDTHLK**W** LSQE**Q**KDELL KMKKDGKAKK ELEAKI**L**HYY
DELEGDAKKE ATEHLKGG**C**R EILKHVVGEE KAAELKNLKD
SGASKEELKA KVEEALHAVT DEEKKQYIAD FGPA**C**KKIYG
VHTSRRRR

Figure 1.7 Amino acid sequence of ABA-1 (http://ca.expasy.org/sprot) cleaved from polyprotein residues 96 – 228, with conserved residues in bold.

CD analysis of the far UV spectrum analysed by the CONTIN procedure predicted a high degree of α-helix in the secondary structure of ABA-1. Secondary structure predictions of individual sequences of NPAs have been performed using GOR algorithms (Garnier *et al*, 1978), Chou and Fasman (Chou & Fasman, 1974) and PHD (Rost *et al*, 1994) and were found to agree with CD studies, as they predict that a single NPA unit is extensively helical, with no evidence of β-structure (Kennedy *et al*, 1995b). The amino acid

residues fell into four regions of predicted helix, so it might therefore form a four-helix bundle similar to carrier proteins such as cytochromes. The predicted four helix structural motif (primarily amphipathic helices) shows the amino acids involved in the helices have well-defined hydrophobic faces and they orient towards a hydrophobic protein core (Kennedy *et al*, 1995b).

The proposed four-helix structure for ABA-1 is conserved across the whole family of NPAs. The hydrophilic and hydrophobic nature of the residues are well conserved across all the NPAs (Kennedy *et al*, 1995b), strongly suggesting that this pattern is structurally or functionally significant.

ABA-1 provides a highly manipulative model for the investigation of the interaction between hydrophobic ligands and α-helical proteins. If structural predictions of ABA-1 are correct then it is distinct from the family of small lipid binding proteins, which include intestinal fatty acid binding proteins, cellular retinoic acid and retinol binding proteins, all of which are predominantly β-barrel with little α-helical content (Banaszak *et al*, 1994).

ABA-1 binding characteristics

Binding studies using the fluorescent fatty acid analogs DAUDA and DACA, have demonstrated that ABA-1 binds fatty acids, with a single binding site per monomer unit, with a dissociation constant (K_d) estimated to be $(8.8 \pm 0.5) \times 10^{-8}$ M (Kennedy *et al*, 1995b; Kennedy, 2000). The fluorescence emission of DAUDA and DACA increases and exhibits a blue shift in emission wavelength when bound to ABA-1. This is distinctive from β-barrel type fatty acid-binding proteins (FABPs), which show less of a blue shift in emission wavelength. The blue shift of DAUDA in ABA-1 has also been observed with cyclohexane suggesting that either the environment of the protein's binding site is apolar or that there are unusual specific interactions in the binding site. Similar blue shifts in fluorescence when bound to DACA could indicate that

the ligand is bound in a non-polar environment and in isolation from solvent water (Kennedy *et al*, 1995b; McDermott, 1999).

ABA-1 has a single conserved tryptophan (Trp 15). Trp 15 emission at a short wavelength indicates that it is buried. The accessibility of the Trp 15 to the water environment examined using quenching of fluorescence by succinimide showed minimal quenching confirming that the Trp 15 is buried within the hydrophobic environment of the structure. Other fluorescent experiments show it to be unaltered when oleic acid was added to ABA-1 (Kennedy *et al*, 1995b), therefore unlikely to constitute a binding site.

ABA-1 also bound retinal and retinoic acid (Kennedy *et al*, 1995a). The change in retinol and retinoic acid fluorescence in ABA-1 was reversed by addition of oleic acid, indicating that the binding site for retinol and oleic acid was competitive (Kennedy *et al*, 1995b).

1.3 Fatty acid and retinol (FAR) binding proteins

Onchocerca volvulus FAR and its homologues represent another novel family of fatty acid and retinoid binding protein from nematodes that are manifest by the severe allergic response of the immune system against them. They are responsible for conditions as diverse as river blindness, lymphatic filariasis and agricultural damage. River blindness has a variety of symptoms including, serious visual impairment, including blindness (figure 1.7), rashes, lesions, intense itching and depigmentation of the skin. The symptoms are due to a parasitic worm that lives and migrates throughout the human body.

Figure 1.8 Image of *Onchocerca volvulus* infection of the eye (www.trachomahki.org/imageye3).

Proteins synthesised by these parasitic nematodes are potent allergens from several different species of nematode and include, OvFAR1 a major antigen of the parasitic nematode *Onchocerca volvulus*, Bm20 from the human lymphatic filariasis parasite, *Brugia malayi*, GpFAR1 from the potato cyst nematode, *Globodera pallida* and CeFAR2 and CeFAR3 from *Caenorhabditis elegans* (figure 1.9).

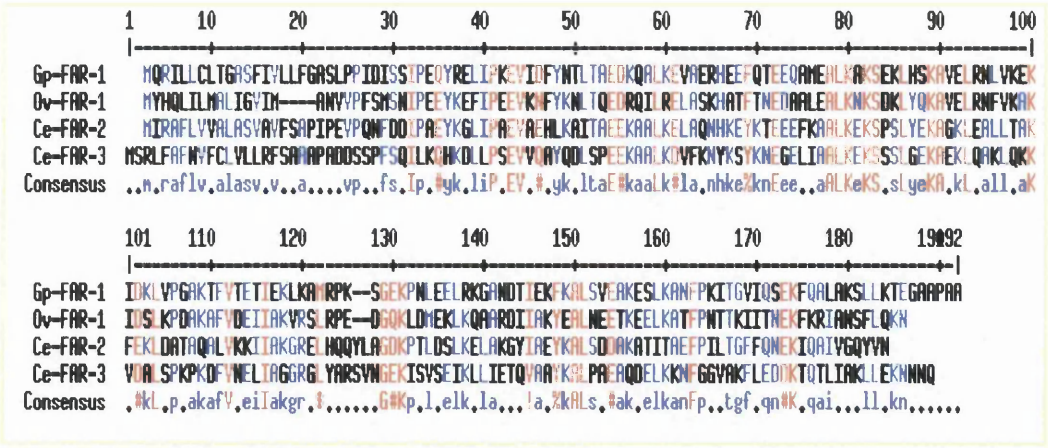

Figure 1.9 Sequence alignment of the FAR proteins. Accession numbers are CAA70477, AAC32662, NM_073850 and NP_506251 for GpFAR1, OvFAR1, CeFAR2 and CeFRA3 respectively. Red signifies 100% similarity and blue signifies 50% similarity.

Similar to the nematode polyprotein allergens (NPAs), these proteins are helix-rich, but are slightly larger, having a molecular mass in the range of 20 kDa. They are distinct from the NPAs in that they are not produced as polyproteins and contain none of the conserved residues of NPAs (McDermott, 1999). In particular, OvFAR1 and its homologues do not contain a trytophan residue.

Structural evidence

The first fatty acid binding proteins were from a family of cytosolic proteins, which comprised of 10 anti-parallel β-strands and 2 short α-helices surrounding the interior-binding cavity in a β-barrel type structure (Banaszak *et al*,1994). Multiple alignments of FAR sequences show no significant similarity to any protein of documented function on current databases (Kennedy *et al*, 1995b). No X-ray or NMR structures yet available for any FAR protein.

Secondary structure sequence based predictions using GOR (Garnier *et al*, 1978) and PredictProtein (http://ca.expasy.org/tools) indicate that the FAR proteins are helix-rich with no predictions for beta structures (Garofalo *et al*, 2002). Circular Dichroism (CD) studies analysed over 240-190 nm by the CONTIN procedure were found to agree with secondary structure sequence based predictions, as CD shows the proteins to have a high α-helical signal

(Kennedy *et al*, 1997).

Analytical ultracentrifugation shows that the dimerisation states of the FAR proteins differ with each protein, for example, rOvFAR1 forms a tight dimer whilst rCeFAR-5 is a monomer (Solovyova *et al*, 2003). These oligomerisation states are unaffected by ligand binding and protein concentration.

Functional evidence

FAR proteins have been shown to bind retinoids with a high affinity, fatty acids with a lesser affinity and do not bind cholesterol. The former is of particular importance considering that infection with *Onchocerca volvulus* causes eye pathology (figure 1.7) and disfiguring skin conditions. Similar symptoms are caused by a deficiency in available retinoids (Bradley, 1999).

The fluorescence emission of retinol upon mixing with a FAR protein undergoes a dramatic increase, indicative of entry into a protein binding site (Garofalo *et al*, 2002). There is competition for the binding site of the FAR proteins demonstrated with oleic acid competing effectively with DAUDA for binding, indicating congruence or interaction between binding sites for lipids. The fluorescence emission of DAUDA is taken as a measure of the polarity of the binding site. FAR proteins bind DAUDA with an increase in intensity and a blue shift in fluorescence, which indicates the ligands bind in an apolar environment, isolated from solvent water (Kennedy *et al*, 1997; Garofalo *et al*, 2002).

Although the function and true *in vivo* binding propensities of the FAR proteins remains to be established (as with NPAs), their secretion into tissues of infected hosts could be important to the survival of the parasite. With no structural counterparts in mammals, OvFAR1 could provide a novel therapeutic target for onchocerciasis therapy.

The main reason for studying *Caenorhabditis elegans* FAR proteins is that

they are homologues of the OvFAR1 protein. *Caenorhabditis elegans* is selected as a model to investigate the function of these proteins since *Caenorhabditis elegans* is much more amenable to experimental analysis than the pathogenic *Onchocerca volvulus*, which requires an insect vector as well as a human host in order to complete its life-cycle.

1.3.1 *Onchocerca volvulus* and OvFAR1

The parasitic nematode *Onchocerca volvulus* is a major causative agent of river blindness, also known as onchocerciasis. It is most commonly found in West Africa and Central America (figure 1.10) where it is spread to humans through the black fly vector. It has been estimated that 120 million people worldwide are at risk of onchocerciasis and that 96 % of these are in Africa, with a total of 18 million people infected with the disease (www.who.int/tdr/diseases/oncho).

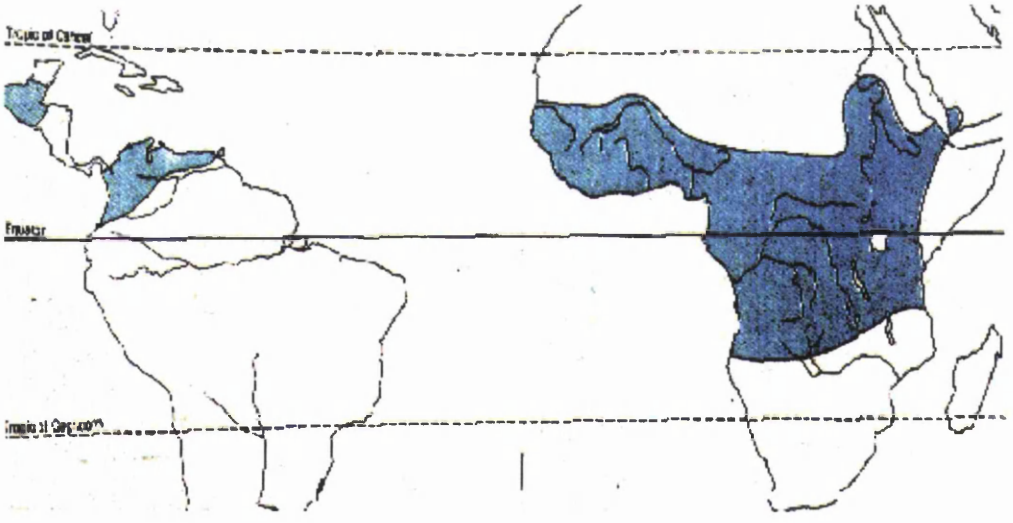

Figure 1.10 Geographical distribution of *Onchocerca Volvulus* (www.biosci.ohio-state.edu/~parasite/onchocerca).

Symptoms of infection include serious visual impairment, including blindness, rashes, lesions, intense itching and depigmentation of the skin. The parasite *Onchocerca volvulus* is transmitted by black fly, which breeds in fast-flowing rivers, in contrast to the stagnant-water breeding sites of mosquitoes. The

infective larvae of *Onchocerca volvulus* are transmitted when the fly takes a blood meal from an infected person, whereupon the larvae develop into adult parasites and gather in palpable, collagenous, subcutaneous nodules where they develop into sexually reproducing adult worms. Female adults release several hundred larvae daily (termed microfilariae), which migrate from the nodule to the skin where they may be taken up by the vector completing the lifecycle (figure 1.11). It is at the microfilarial stage of the parasite that causes the pathology of onchocerciasis (Muller *et al*, 1975; www.who.int).

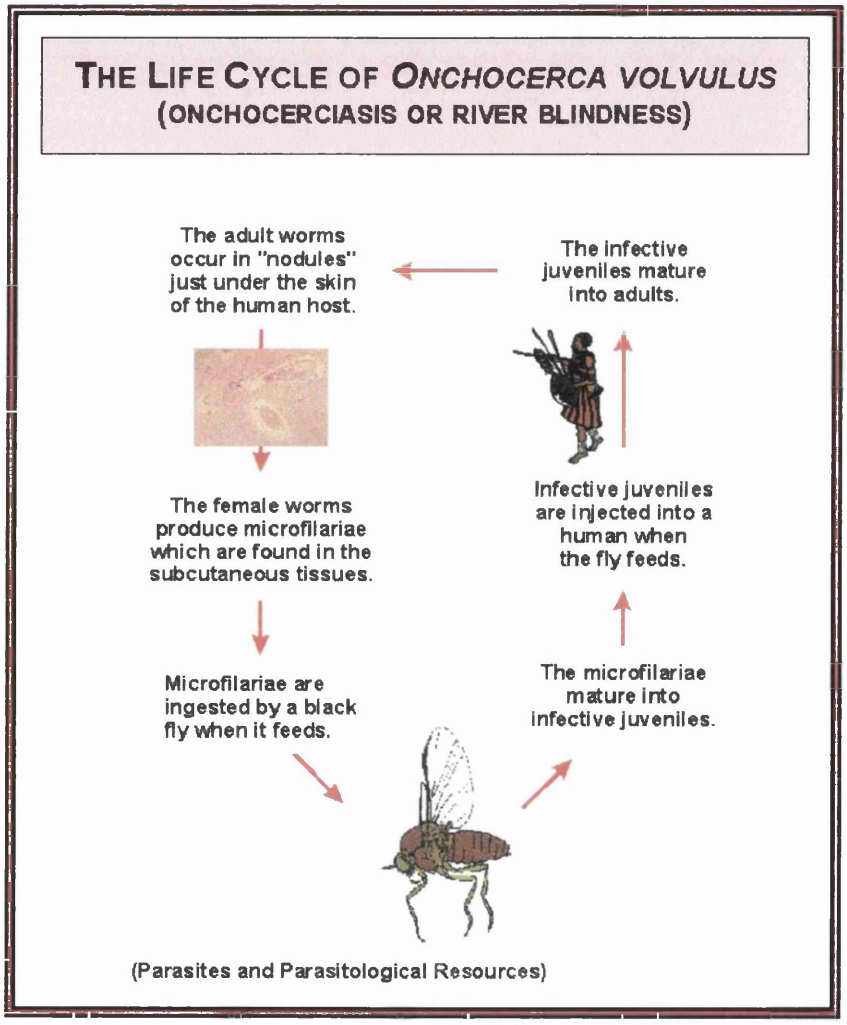

Figure 1.11 Lifecycle of *Onchocerca volvulus* (www.biosci.ohio-state.edu/~parasite/lifecycles/onchocerca_lifecycle.html).

The use of vitamin A (retinol) in treatment of ocular onchocerciasis was introduced because it was believed that infection induced vitamin A

deficiency. The symptoms of vitamin A deficiency are similar to onchocerciasis and include night blindness, destruction of the cornea and also similar skin pathology (Bradley *et al*, 1999). A study (Mustafa *et al*, 1979) on onchocerciasis patients found that their level of Vitamin A was reduced. *Onchocerca volvulus* infection may contribute to the host being deficient in vitamin A by simple sequestration of retinol. The retinol concentration in the nodules formed by *Onchocerca volvulus* has been reported to be eight times that of the surrounding host tissue, which is likely to be particularly harmful for infected patients with a poor dietary intake of vitamin A (Lal & James, 1996). It has been suggested that *Onchocerca volvulus* requires retinol for a variety of metabolic and developmental purposes, including growth, differentiation, embryogensis, glycoprotein synthesis and anti-oxidants (Sani & Vaid, 1985). A role in growth and development was suggested by experiments with radio labelled retinoic acid in *Brugia malayi*, also a filarial parasite, which showed that it was taken up by worms and localised to high concentrations in early and late embryonic forms (Bradley & Kennedy, 1999).

The treatment of river blindness includes Ivermectin and Diethylcarbamazine drugs which are capable of reducing the numbers of skin microfilariae in infected people. Both drugs will kill the adult nematodes, however, symptoms such as blindness are irreversible (www.who.int).

OvFAR1 is the allergen protein secreted by *Onchocerca volvulus* and the causative agent of river blindness. It is hoped that by determining the 3-D structure by X-ray crystallography the structure/function relationship of OvFAR1 can be elucidated thereby creating a route to designing inhibitors to prevent river blindness.

OvFAR1 structural predictions

Predictions of secondary and biophysical analysis have shown that the molecule is predominantly α-helical, with no evidence of β structure. This immediately discriminates it from retinol-binding proteins of similar sizes in

humans, in which β structures predominate (FABPs). Recent work on *Onchocerca volvulus* showed that retinol binds to OvFAR1 protein. OvFAR1 has a molecular weight 19.7 kDa (Bradley & Kennedy, 1999), with 178 amino acids (figure 1.12) (http://ca.expasy.org/sprot/).

MYHQLILMALIGVIMANVVPFSMSNIPEEYKEFIPEEVKNFYKNLTQED
RQILRELASKHATFTNEDAALEALKNKSDKLYQKAVELRNFVKAKIDS
LKPDAKAFVDEIIAKVRSLRPEDGQKLDMEKLKQAARDIIAKYEALNE
ETKEELKATFPNTTKIITNEKFKRIANSFLQKN

Figure 1.12 Amino acid sequence of OvFAR1 (www.ca.expasy.org/sprot/).

Circular dichroism shows a strong α-helix signal. Analysis of the data over the range 190 nm to 240 nm by CONTIN procedure estimates secondary structure to be 60 % α-helix, 32 % β-sheet and a remainder of 8 % (Kennedy *et al*, 1997).

Searching databases for similar structures shows no significant matches. Secondary structure prediction performed using PHD (Rost, 1994) algorithms applied to the whole sequence alignment predicted an 86 % helical conformation. These helices are predicted to be strongly amphipathic (Kennedy *et al*, 1997).

OvFAR1 binding characteristics

OvFAR1 molecule binds a single molecule of retinol with an affinity similar to that of other serum retinol-binding proteins. The dissociation constant of retinol:rOvFAR1 binding has been estimated by fluorescence titrations, with increasing quantities of retinol added to rOvFAR1, the results were consistent with a single retinol-binding site per molecule of rOvFAR1, which provided a K_d of 8.5×10^{-8} M (Kennedy *et al*, 1997). OvFAR1 also binds fatty acids, but with a lesser affinity than ABA-1. The fluorescence emission intensity of DAUDA occurs at 541 nm, but moved to a shorter wavelength of 485 nm upon binding to OvFAR1. A blue shift of this magnitude can be taken as indicative

of entry of the fluorophore into an apolar environment and removed from contact with water. Changes similar to this are observed in ABA-1. The fluorescence enhancement of DAUDA by OvFAR1 was reversed upon addition of oleic acid, showing that the binding is competitive. The binding is also competitive between retinol and oleic acid with OvFAR1 indicating that the binding sites of fatty acids and retinol with the protein are interactive or the same (Kennedy *et al*, 1997).

A further important finding was that parasite retinol-binding proteins also bind Ivermectin, which is currently the main drug in onchocerciasis control programme (Bradley & Kennedy, 1999). For that reason these proteins may participate in the delivery of the drug to the parasite, or diminish the effect, and it is possible that they may be involved in the development of drug resistance. Ivermectin competes efficiently with retinol for retinol-binding sites on parasite retinol-binding protein, but not for the host-tissue retinol-binding protein sites. The drug has no affinity for retinoic acid-binding proteins from either parasite or host tissues (Lal & James, 1996).

1.3.2 *Globodera Pallida* and GpFAR1

The potato cyst nematode or 'eelworm' as it is commonly known is a major pathogen of potato crops, including the commercial potato *Solanum tuberosum* (Duncan, 1997). There are two main species of potato cyst nematodes (figure 1.11), *Globodera pallida* and *Globodera rostochiensis*. Plant parasitic nematodes destroy up to 10 % of the world's agricultural output at a cost of up to 77×10^9 dollars each year in yield losses and controlling the pests (Sasser & Freckman, 1987). These potato nematodes are speculated to have originated from South America and have become established in temperate regions and a few other highland locations (www.inra.fr/hyppz/RAVAGEUR/bgloros.htm).

Figure 1.13 Potato cyst nematode

Potato cyst nematodes reduce yield by about 30 %. Tiny cysts can survive in the soil for 15 – 20 years, each containing several hundred young nematodes unhatched, until a potato plant is grown in that soil. Chemicals in the roots of the potato plant stimulate hatching and they then invade the roots of the plant. After the adult female nematode has developed and mated, her body forms a cyst containing the next generation of nematodes. These cysts then become detached from the roots and lie dormant in the soil to continue the cycle.

GpFAR1 is secreted by the plant parasitic nematode *Globodera pallida*. Secreted proteins have been implicated in many aspects of the host-parasite interaction, including migration of the nematode through the root before induction of the feeding site and in induction and maintenance of the syncytium. It is the nematode secretions that lead to modification of plant tissue with the result that growth of the potato is slowed or stopped, leading to poor yield (www.inra.fr/hyppz/RAVAGEUR/bgloros.htm).

GpFAR1 structural predictions

GpFAR1 has a molecular weight of 20.9 kDa, and consists of 188 amino acids (http://ca.expasy.org/sprot/). When the leader sequence, predicted by SignalP program (Nielson, 1997), is removed it has a molecular weight of 18.8 kDa. The predicted cleavage site for the leader sequence is between Gly-19 and Ala-

20. The GpFAR1 sequence (figure 1.14) shows close similarity to other nematode-specific FAR proteins (figure 1.9).

MQRILLCLTGASFIVLLFGASLPPIDISSIPEQYRELIPKEVIDFYNTLTAE
DKQALKEVAERHEEFQTEEQAMEALKAKSEKLHSKAVELRNLVKEKI
DKLVPGAKTFVTETIEKLKAMRPKSGEKPNLEELRKGANDTIEKFKAL
SVEAKESLKANFPKITGVIQSEKFQALAKSLLKTEGAAPAA

Figure 1.14 Amino acid sequence of GpFAR1 (http://ca.expasy.org/sprot/)

Secondary structure predictions of individual sequences of FAR proteins have been performed using PHD (Rost *et al*, 1994) algorithms and predict α-helix, with no evidence of β-structure (Prior *et al*, 2001). COILS algorithm (Lupas *et al*, 1991) predicted several strongly amphipathic stretches, with the strongest predicted between 60-100 and 130-168 (Prior *et al*, 2001). Circular dichroism analysis has not been carried out on GpFAR1, but other FAR proteins have been predicted to have a high α-helical signal (Kennedy *et al*, 1997; Garafalo *et al*, 2002).

There are no known similar structures and no X-ray or NMR structures yet available.

GpFAR1 binding characteristics

GpFAR1 binds fatty acids and retinol. It has been shown to bind the fluorophore tagged fatty acid DAUDA and DACA with an increase in intensity and with a blue shift in fluorescence indicating that it is bound in an apolar environment. The blue shift in fluorescence is greater than that of β-structure fatty acid binding proteins. Other fatty acids bind to GpFAR1 in competition with DAUDA and the fatty acids were found to displace retinol indicating the binding site is congruent or interactive. GpFAR1 also binds cis-parinaric acid suggesting that it is intrinsically fluorescent (Prior *et al*, 2001).

GpFAR1 fails to bind dihydroergosterol or a range of hydrophobic compounds related to plant defence responses, including glycoalkaloids (Prior *et al*, 2001). Therefore GpFAR1 is unlikely to be utilized in plant defences.

The preference of GpFAR1 for fatty acid chain length has been investigated (Prior *et al*, 2001) and was found to prefer saturated or unsaturated fatty acids with a chain length in the range C12 to C17. This correlates with the most abundant fatty acids found in the organism (Holz *et al*, 1997), which are fatty acids with chain length C14-C22.

1.3.3 *Caenorhabditis elegans* and CeFAR2 and CeFAR3

Caenorhabditis Elegans (*C. elegans*) is a small free-living nematode (figure 1.15), which lives in the soil, especially in rotting vegetation. It is found in many parts of the world, where they survive by feeding on microbes such as bacteria.

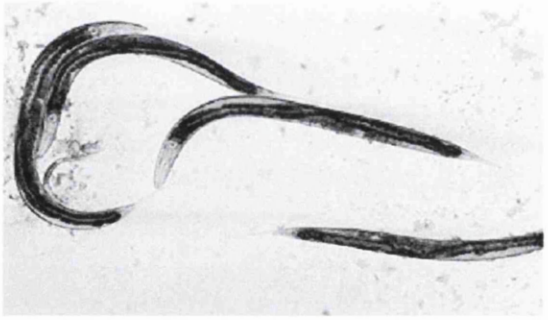

Figure 1.15 Diagram of *C. elegans*.

C. elegans is conceived as a single cell, which develops, starting with embryonic cleavage, proceeding through morphogenesis and growth into adult (figure 1.15). Its average life span is 2 – 3 weeks (http://elegans.swmed.edu) with a lifecycle that is a simple 3 day generation time. It can be grown quickly and with ease on agar plates using *Eschericha coli* as a food source (Hodgkin *et al*, 1995).

Though *C. elegans* is not an important nematode in terms of infection, it is likely to be an excellent model for the study of important parasitic nematodes

in its phylum (Blaxter, 1998), including *Ascaris lumbricodes, Wuchereria bancrofti* and *Brugia malayi.*

Since genetic resistance to current anti-nematode drugs is increasing novel control strategies are needed. By studying *C. elegans* these could be developed, possibly involving nematode-specific neurotropic agents, disrupting sex determination or embryogenic pathways of nematodes (Blaxter, 1998). The *C. elegans* genome project has identified 6 protein homologues of proteins from *Onchocerca volvulus* and other proteins from *Globodera pallida.* Therefore, there is growing interest in its potential as a vaccine component and as a marker of immune status in onchocerciasis (Blaxter, 1998). By studying the structure and function of the homologues of *C. elegans*, the results can be transferred to proteins of other parasites (*C. elegans* sequencing consortium, 1998).

CeFAR structural predictions

C. elegans contains protein homologues of NPA and FAR proteins. The protein sequences of CeFAR2 and CeFAR3 have been compared using the BLAST algorithm (Altschul *et al*, 1990) and show significant similarity of 34.9 % and 36 % respectively (Tree *et al*, 1995). The main reason for working on the *C. elegans* fatty acid and retinol (FAR) binding proteins is that they are homologues of the *Onchocerca volvulus* FAR protein (figure 1.9). *C. elegans* has been selected as a model to investigate the function of these proteins since it is much more amenable to experimental analysis than the proteins from the pathogenic *Onchocerca volvulus* protein.

CeFAR2 is found in chromosome III in *C elegans*, expression of this protein was observed in the body wall muscle cells of the nematode from early embryogenesis to adulthood (Tree *et al*, 1995). It has a molecular weight of 20035 Da and consists of 186 amino acids (figure 1.16) (http://ca.expasy.org/sprot/).

MIRAFLVVALASVAVFSAPIPEVPQNFDDIPAEYKGLIPAEVAEHLKAITAEEKAALKE
LAQNHKEYKTEEEFKAALKEKSPSLYEKAGKLEALLTAKFEKLDATAQALVKKIIAK
GRELHQQYLAGDKPTLDSLKELAKGYIAEYKALSDDAKATITAEFPILTGFFQNEKIQA
IVGQY VN

Figure 1.16 Amino acid sequence of CeFAR2 (http://ca.expasy.org/sprot/)

CeFAR3 is also found in chromosome III in *C elegans* has 189 amino acids and has a molecular weight of 20905 Da (figure 1.17) (http://ca.expasy.org/sprot/).

MSRLFAFNVFCLVLLRFSAAAPADDSSPFSQILKQHKDLLPSEVVQAYQDLSPEEKAA
LKDVFKNYKSYKNEGELIAALKEKSSSLGEKAEKLQAKLQKKVDALSPKPKDFVNELI
AGGRGLYARSVNGEKISVSEIKLLIETQVAAYKALPAEAQDELKKNFGGVAKFLEDD
KTQTLIAKLLEKNNNQ

Figure 1.17 Amino acid sequence of CeFAR3 (http://ca.expasy.org/sprot/)

They are from a new class of helix rich proteins from nematodes, which bind fatty acids and retinol and hence fit into the overall project. There are no known similar structures and no X-ray or NMR structures are yet available. Circular dichroism revealed that CeFAR2 and CeFAR3 have high content of α-helix (54 and 61% respectively) and little β-structure. They exhibit high probabilities of coiled coil structure by the Coils algorithm (Lupus *et al*, 1991).

CeFAR binding characteristics

All CeFAR proteins studied to date bind DAUDA with an increase in intensity and blue shift in fluorescence from 543 nm to 482 - 490 nm upon addition of protein, which is indicative of an apolar binding site. Dissociation constants ranged from 8×10^{-7} to 1×10^{-8} M. Binding was observed with cis-parinaric acid and retinol, which are intrinsically fluorescent and exhibit environment-sensitive fluorescence emission. Addition of oleic acid reversed retinol binding, suggesting binding sites for the fatty acid and retinol are either congruent or interactive (Garofalo *et al*, 2003).

This project looks at proteins from these classes to get a better understanding of their structure and function by biophysical techniques, including CD, DLS, fluorescence, MALDI-TOF and crystallography.

Chapter 2 - Materials and methods

All materials used in this project were purchased from Sigma Aldrich unless otherwise stated.

2.1 Protein analysis

It is important to determine the purity, concentration and activity of the proteins for biophysical techniques and crystallisation studies.

2.1.1 Molecular weight determination

Inputting the amino acid sequence for each protein into the ProtParam program available at http://ca.expasy.org/tools/protparam.html, the molecular weight can be estimated. The program counts the number of amino acids and amino acid composition, number of charged residues, calculates molecular weight, gives theoretical pI, estimates the extinction coefficient and estimates the stability of the proteins (The results of ProtParam for each of the proteins studied are given in Appendix 2).

2.1.2 Concentration determination

The concentration of protein solutions was determined using a Perkin Elmer UV/VIS spectrometer Lambda 40 and applying Beer-Lambert law $A = \varepsilon cl$. Absorbance, A, was measured at 280 nm with the protein in 1 cm path length cuvettes. Using the protein extinction coefficient, ε ($M^{-1}cm^{-1}$) at 280 nm with path length value 1 (cm), the protein concentration was calculated. Appendix 2 shows protein ε_{280} values, which were estimated using ProtParam available at http://ca.expasy.org/tools/protparam.html.

2.1.3 Sodium Dodecyl Sulphate – Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE gels were used to estimate molecular mass and evaluate protein purity. 15 % SDS-PAGE gels 0.75 mm thick were prepared in the Hoefer® Mighty small SE245, dual gel caster (Amersham Pharmacia Biotech). Boiling for 5 mins denatured the proteins that had been mixed 1:1 with loading buffer. Electrophoreses took place under reducing conditions using a discontinuous buffering system at 200 V and 20 mA per gel. Low molecular weight marker proteins (phosphorylase (94 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (30 kDa), soybean trypsin inhibitor (20.1 kDa) and α-lactalbumin (14.4 kDa)) from Amersham Pharmacia Biotech were dissolved in 200 µl loading buffer.

Gels were stained with Coomassie Brilliant Blue R250 for 30 minutes, followed by destaining with distilled water until a clear background was obtained.

2.1.4 Isoelectric focussing electrophoresis (IEF)

IEF gels are used for isoelectric point (pI) determination and confirmation of the presence of charged isoforms.

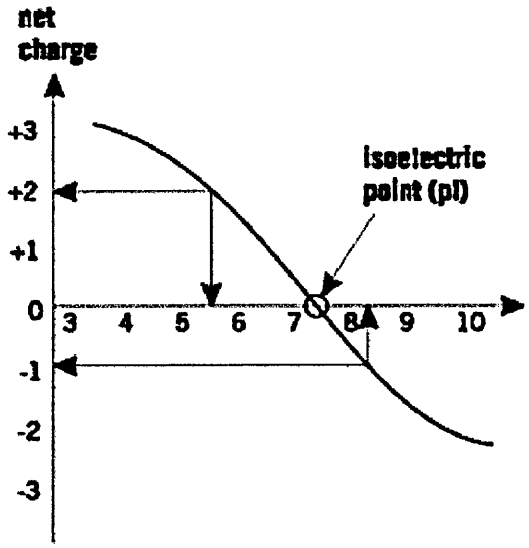

Figure 2.1 Graph showing what the pI of a protein is when the net charge is zero.

Novex® IEF gels are 5 % polyacrylamide, non-denaturing gels with a pH 3 – pH 10. These gels were used on the Xcell Surelock™ system from Invitrogen. A broad range of pI (pI 3.5 - 9.3) markers (amylglucosidase (pI 3.5), methyl red dye (pI 3.75), trypsin inhibitor (pI 4.55), b-lactoglobulin A (pI 5.2), bovine carbonic anhydrase B (pI 5.85), human carbonic anhydrase B (pI 6.55), myoglobin acidic band (pI 6.85), myoglobin basic band (pI 7.35), lentil lectin acidic (pI 8.15), lentil lectin middle (pI 8.45), lentil lectin basic (pI 8.65) and trypsinogen (pI 9.3)) from Amersham Pharmacia biotech were dissolved in 100 µl loading buffer. Cathode and anode buffers were purchased from Invitrogen. Protein samples were mixed 1:1 with sample buffer and electrophoresed at 100 V for 1 hour, 200 V for 1 hour, then 500 V for 30 minutes, at 5 mA per gel.

Gels were placed in fixing solution before staining with Coomassie Brilliant Blue R250 for 5 minutes. Destaining with water until the required background was obtained.

2.2 Biophysical techniques

2.2.1 Circular dichroism

Sharon M. Kelly and Nicholas C. Price (Kelly & Price, 1997) have reviewed the application of circular dichroism. Circular dichroism (CD) reports on the general secondary and tertiary structures of proteins by measuring the differential absorption of the left and right circularly polarised components of plane-polarised radiation. This effect occurs when a chromophore is chiral either, intrinsically by reason of its structure, by being covalently linked to a chiral centre or by being placed in an asymmetric environment. Chiral molecules differentially absorb beams of left and right handed circularly polarised light, causing two beams to travel at different speeds through these molecules, rotating the polarised light. A CD spectrum is obtained when the dichroism is measured as a function of wavelength. Studies of the far-UV region (190 – 210 nm) can be used to assess quantitatively the overall secondary structure content of the protein. In this region the peptide bond is the principal absorbing group.

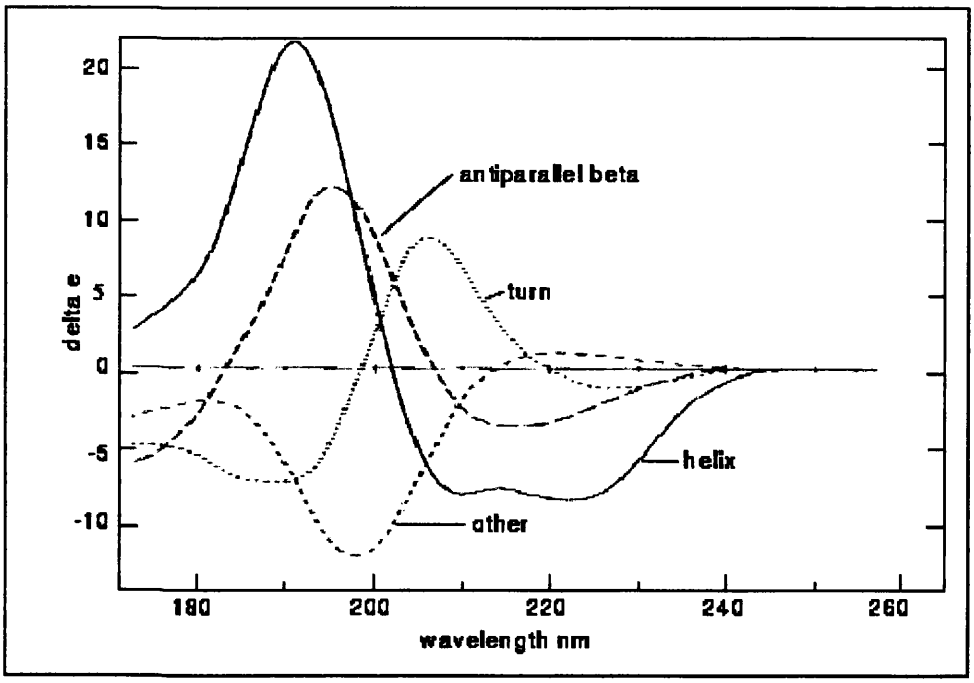

Figure 2.2 Circular dichroism spectra of "pure" secondary structures (Brahms & Brahms, 1980).

Provencher and Glöckner (Provencher & Glöckner, 1981) introduced a procedure of deducing the contributions of the different structural forms. It is called the CONTIN procedure and involves the direct analysis of a CD spectrum over the range from 240 nm to 190 nm as a linear combination of the CD spectra of 16 proteins whose structures have been determined to high resolution by X-ray crystallography. Several methods have been developed which analyse the experimental CD spectra using a database of reference protein CD spectra containing known amounts of secondary structure (Provencher & Glöckner, 1981; Hennessey & Johnson, 1981; Sreerama & Woody, 1994). These methods are, in general, more accurate and reliable than the Provencher and Glöckner approach.

In the near UV (250 –290 nm), the aromatic amino acid side chains (phenylalanine, tyrosine and trypothan) absorb. The tertiary folding of the polypeptide chain can place these side chains in chiral environments, thus giving rise to CD spectra that can serve as characteristic "fingerprints" of the native structure. These "fingerprints" can be useful for comparisons of tertiary structures between related proteins.

CD experimental method

All proteins were dialysed into 20 mM phosphate, pH 7.0 and concentrated to 1 mg/ml as estimated by absorbance at 280 nm (see chapter 2.1). Mean residue weights for each recombinant protein were calculated from their amino acid sequence, then their molar elliptically values were obtained. CD spectra were recorded at room temperature in a JASCO J-600 spectropolarimeter under PC control. The proteins were placed in quartz cells of path length 0.02 or 0.05 cm. Analysis of secondary structure of proteins was performed using the CONTIN procedure (Provencher, 1981) and SELCON procedure (Sreerama, 1994) in the far UV region (190 – 210 nm) and also the near UV region (250 – 290 nm).

2.2.2 Fluorescence

Fluorescence spectroscopy is one of the most widely used spectroscopic techniques in biochemical and biophysical fields due to its sensitivity to changes in structural and dynamic properties of biomolecules and biomolecular complexes. With macromolecules, fluorescence measurements can give information about conformation, binding sites, solvent interactions, degree of structural flexibility, intermolecular distances and the rotational diffusion coefficient.

When a molecule absorbs a photon of light the electrons in the molecule gain energy and are excited from their ground state energy level to a higher energy level. If the molecule contains no fluorescent chromophore the excess energy is lost as heat by collision with solvent molecules, as the electron returns to ground state. If the molecule contains fluorescent chromophores, reradiation of the absorbed energy as fluorescence occurs on return to ground state. The emitted light will have less energy than the absorbed light, therefore fluorescent light has a longer wavelength than the exciting light.

The Jablonski diagram (figure 2.3) is a schematic portrayal of molecular electronic and vibrational energy levels, which shows the sequence of steps involved in fluorescence.

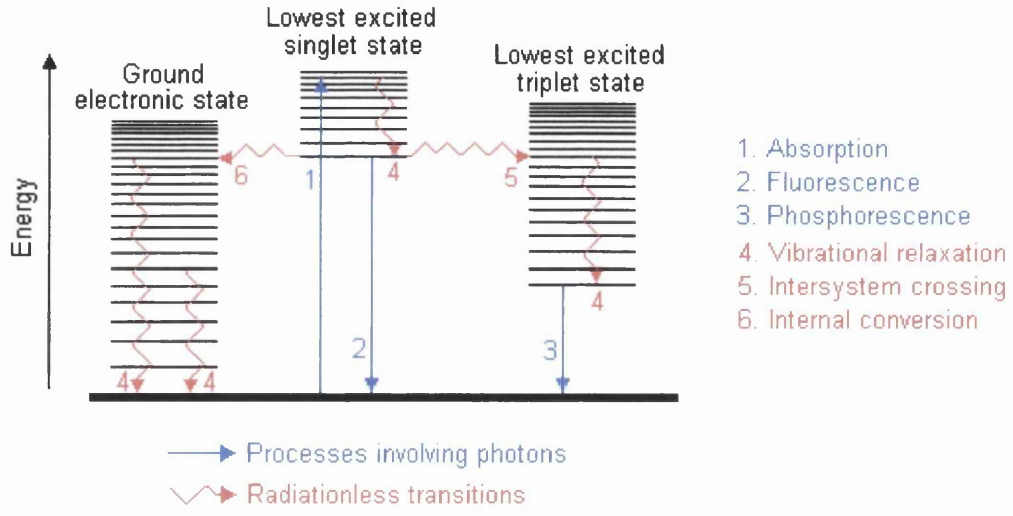

Figure 2.3 Jablonski diagram

Substances that normally exhibit fluorescence are those structures that contain delocalised electrons present in conjugated double bonds. Two types of fluorophores are used in fluorescence analysis of macromolecules: intrinsic fluorophores contained in proteins themselves and extrinsic fluorophores added to proteins and usually binding to them.

For proteins there are three intrinsic fluorophores – tryptophan, tyrosine and phenylalanine. The fluorescence of tryptophan is most commonly studied, as it is extremely sensitive to the hydrophobicity of its surroundings and so can be used as a probe to assess its environment.

Extrinsic fluorophores used for protein investigation include, dansyl chloride, fluorescein, dansyl amide and ANS.

Measurement of a fluorescence lifetime, the average time the fluorophore spends in the excited state, is a useful probe of the dynamics in a system. Fluorescence anisotropy or polarisation is the degree to which fluorescence emission is depolarised relative to the polarising exciting light. These experiments yield information on the degree of rotational freedom of the fluorophore on the lifetime time scale. Fluorescence quenching is the ability that compounds have to decrease the fluorescence of some fluorophores. Quenching of tryptophan fluorescence can reveal the extent of tryptophan exposure within proteins. Energy transfer is the effect observed when an electronically excited fluorophore (donor) can transfer its energy to another fluorophore (acceptor), which then emits the energy as fluorescence. Energy transfer experiments provide information on the distance between the donor and the acceptor.

Fluorescence experimental method

Fluorescence experiments were performed with a SPEX FluorMax Spectrofluorimeter (SPEX industries, Edison, NJ) at 25 °C using 2 ml samples in a silica cuvette, and the emitted light was detected at 90° relative to the incident beam.

Proteins were concentrated to 1 mg/ml in PBS buffer. PBS blanks were used to correct for Raman and background scattering where necessary.

The fluorescent fatty acid analogue 11-((5-dimethylaminonaphthalene- 1-sulfonyl)amino)undecanoic acid (DAUDA) (Wilkinson & Wilton, 1986) (figure 2.4) was obtained from Molecular Probes. A 10 mM stock solution was made in ethanol and stored at –20 °C, this was diluted 1:1000 in PBS for use in assays. A 1:10 DAUDA stock solution was diluted in methanol and the concentration was checked by absorbance at 335 nm, using an extinction coefficient ε_{335} of 4400 $M^{-1}cm^{-1}$. The excitation wavelength of DAUDA is 345 nm.

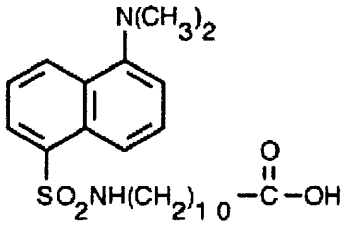

Figure 2.4 Structure of 11-((5-dimethylaminonaphthalene- 1-sulfonyl)amino)undecanoic acid (DAUDA)

A 10 mM stock solution of the fluorescent fatty acid retinol (figure 2.5) was made in ethanol and was stored in the dark at –20 °C. Free retinol is poorly soluble in aqueous solution, so was dissolved and diluted in ethanol immediately before use. The concentration of retinol was estimated by absorbance of a solution of retinol in ethanol at 325 nm, using an extinction coefficient of ε_{325} of 52480 $M^{-1}cm^{-1}$. The excitation wavelength of retinol is 350 nm.

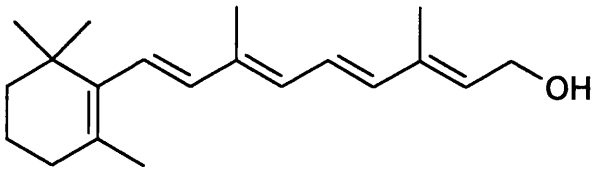

Figure 2.5 Structure of Retinol

Stock solutions of all non-fluorescent competitors (figure 2.6) were made to approximately 10 mM in ethanol, then diluted 1:10, 1:100 and 1:1000 in PBS for use in assays.

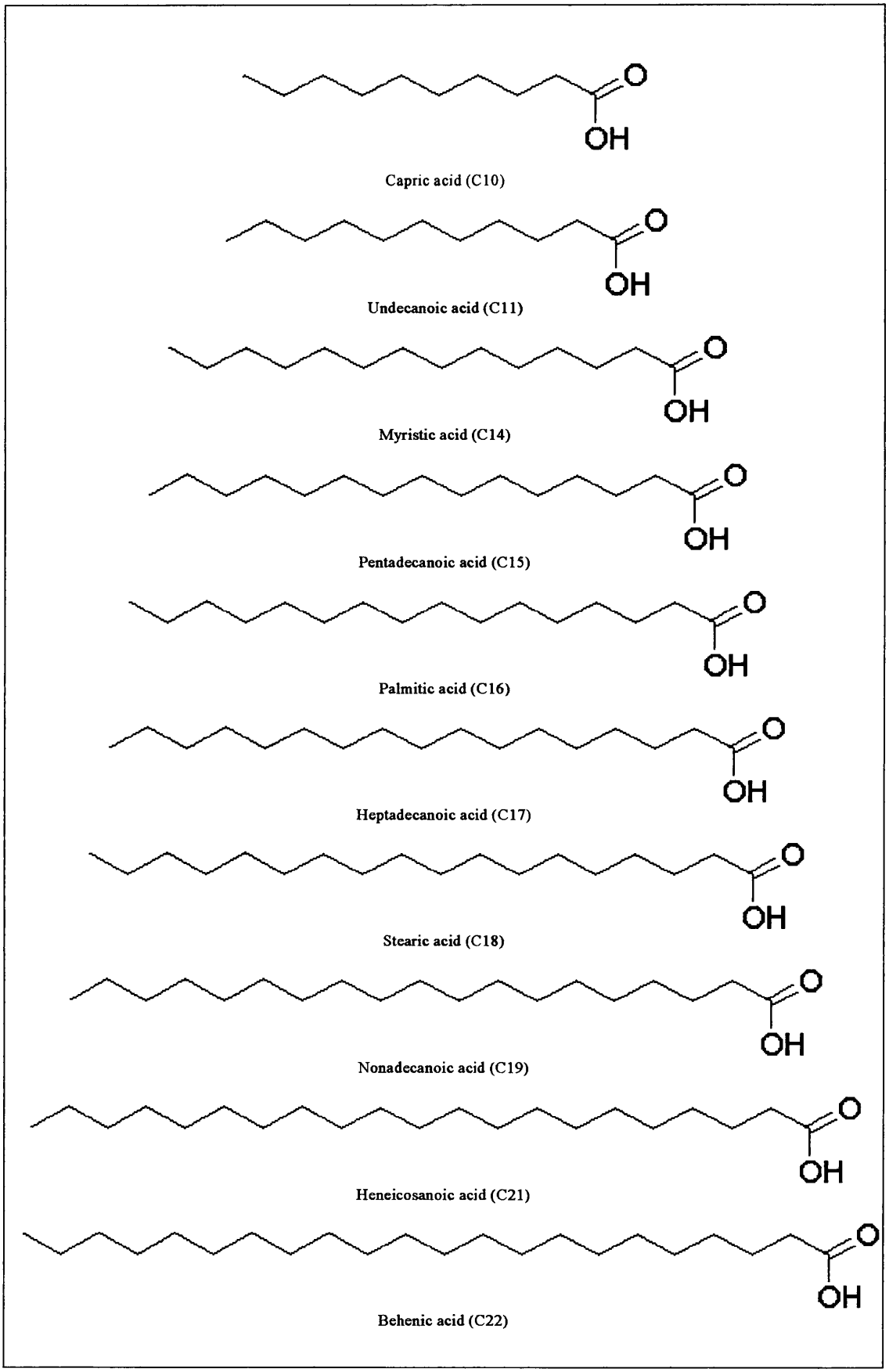

Figure 2.6 Structures of fatty acids with different carbon chain length

As the proteins studied are fatty acid and retinol binding proteins, fluorescence was used to determine if they had retained their activity. The ligand binding capacity of the proteins was investigated using DAUDA. 10 µl samples of protein at 1 mg/ml in PBS were added successively to 2 ml of DAUDA and the maximal fluorescence intensity was noted. Fluorescence data were corrected for dilution where necessary and fitted by standard non-linear regression techniques using the standard hyperbolic/non-cooperative binding equation (shown below) (Microcal, ORIGIN 5.0) to a single non-competitive binding model to give estimates of the dissociation constants (Kennedy *et al*, 1997).

$$F_{max}/F = 1+K_d/ (C_L - nC_PF/F_{max})$$

C_L = total ligand (DAUDA) concentration
C_P = total protein concentration
F_{max} = estimated extrapolated maximal fluorescence intensity
K_d = dissociation constant
n = number of binding sites per protein monomer unit.

The dissociation constant for retinol was calculated similarly. To correct for the fluorescence of free retinol, successive additions of free retinol were made to a cuvette containing only PBS. Free retinol intensities were subtracted from protein-retinol intensities before data analysis.

For retinol titration experiments, 10 mM retinol in ethanol was added in five 2 µl aliquots to 2 ml of a 1 mg/ml solution of protein in PBS, mixed immediately and the change in emission intensity noted at 350 nm. Addition of 10 µl samples of oleic acid (figure 2.7) of 1:1000, 1:100 and 1:10 dilutions and 10 µl of 10 mM oleic acid stock were made to the retinol-protein complex to determine if bound retinol was competitively displaced by oleic acid.

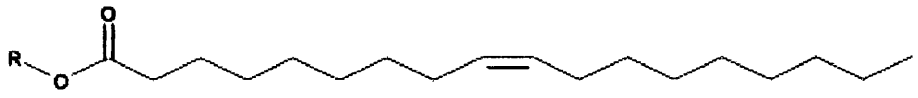

Figure 2.7 Structure of Oleic acid

To investigate the preference of the FAR proteins for fatty acid chain length, the same molarity of non-fluorescent competitor fatty acid (figure 2.6) was added to a DAUDA-protein complex and the percentage decrease in DAUDA fluorescence was recorded at the peak fluorescence emission wavelength of DAUDA in the protein (488 nm). This value was recorded for successive dilutions of each fatty acid.

2.2.3 Dynamic light scattering

(http://cbr.med.harvard.edu/investigators/springer/lab/protocols/sara_light_scattering)

Scattered light from a protein sample contains information about the particles in solution. This information is a valuable asset in characterizing hydrodynamic size, conformation, interaction or aggregation of the macromolecule. The hydrodynamic radius of particles e.g. proteins in solution, estimates the molecular weight if the protein is perfectly globular. If not, then this measurement roughly corresponds to half the maximal dimension of the protein. Therefore this information may infer the oligomerisation state of the protein.

The method can only determine if a solution contains protein that is monomeric, dimeric or a mixture of oligomers. Therefore DLS is a good test for homogenous, pure, non-aggregated and monodispersed protein: pre-requisites for successful crystallisation trials.

Experiments were carried out on a DynaPro-801 TC in which the coherent light source is a laser. The fluctuations in the scattered intensity of the protein particles are detected over short time-scales, due to varying constructive and

destructive interference of light scattered from these particles travelling with different speeds and directions (Brownian motion). This fluctuation has a time-scale related to the translational diffusion coefficient, which is related to the hydrodynamic radius of the particle by the Stokes-Einstein equation (shown below).

$$R_H = k_b T / 6\pi\eta D_T$$

R_H = Hydrodynamic Radius (nm)

k_b = Boltzman's Constant (JK^{-1})

T = Absolute Temperature (K)

η = Solvent Viscosity ($gm\ cm^{-1}\ s^{-1}$)

D_T = Translational diffusion coefficient (m^2s^{-1})

DLS experimental method

2 mg/ml of each protein (rABA-1, rGpFAR-1, rOvFAR-1, rCeFAR2, rCeFAR3 and lysozyme as a standard) were filtered through 0.2 μm filter and injected into the DynaPro-801 TC instrument. For each protein a range of buffers at specific pH values were used:

20 mM ACETATE pH 4 and 150 mM NaCl

20 mM PIPES pH 6.6 and 150 mM NaCl

50 mM HEPES pH 7.5 and 150 mM NaCl

20 mM TRIS pH 8.5 and 150 mM NaCl

For each protein in each buffer DLS was performed over the range of temperature 5 - 25 °C in increments of 5°. Table 2.1 gives an explanation of DLS parameters that were recorded.

Ampl ( )	Amplitude of the scattered light fluctuations
Diffn Coeff	Translational diffusion coefficient
Radius (nm)	Hydrodynamic radius calculated from Stokes-Einstein equation
Polyd (nm)	Width of the Gaussian distribution of radii fitting data, which has its centre at radius
Estd MW	Estimated molecular weight in kDa, assuming protein is spherical
Temp (°C)	Instrument is temperature controlled
Count (cnts/s)	Number of photons scattered per millisecond at 90°
Base line	Quality of fit of the sample monodisperse model to the data. Above 1.004 indicates a bad model
SOS Error	Second quality control. Needs to be less than 5

Table 2.1 Table explaining the parameters obtained from a DLS experiment *(http://cbr.med.harvard.edu/investigators/springer/lab/protocols/sara_light_scattering)*

2.2.4 Matrix Assisted Laser Desorption Ionisation – Time Of Flight (MALDI-TOF)

MALDI (Hillenkamp *et al*, 1991) has been used successfully in determining molecular weights of biomolecules, monitoring bioreactions, detecting post-translational modifications, performing protein and oligonucleotide sequencing analysis.

MALDI is based on the bombardment of sample molecules with laser light to bring about sample ionisation. The sample (analyte) is pre-mixed with a highly absorbing matrix compound. Matrices are small organic compounds that spare the analyte from degradation resulting in the detection of intact molecules.

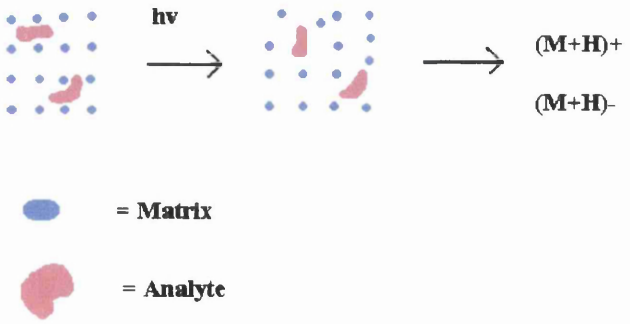

Figure 2.8 – Matrix Assisted Laser Desorption Ionisation (http://www.astbury.leeds.ac.uk/Facil/MStut/mstutorial.htm)

The matrix absorbs the laser energy, which in turn excites the sample. In this way energy transfer is efficient and also the analyte molecules are spared excessive direct energy. Once the sample molecules are vaporised and ionised they are transferred electrostatically into a time-of-flight mass spectrometer, where they are separated from the matrix ions, and individually detected and analysed according to their mass-to-charge ratios.

MALDI is a "soft" ionisation method and so results predominantly in the generation of singly charged molecular-related ions regardless of the molecular weight.

Positive ionisation mode is used for protein and peptide analyses. Protonated molecular ions ($M+H^+$) are usually the dominant species. Although they can be accompanied by salt adducts, a trace of the doubly charged molecular ion, and/or a trace of a dimeric species.

MALDI-TOF experimental method

10 pmol/ml of each protein in 10 mM TRIS pH 7.5 was prepared for MALDI-TOF. The protein was dissolved in an appropriate volatile solvent (sinapinic acid is common for protein analyses), with a trace of trifluoroacetic acid for positive ionisation. An aliquot (1-2 μL) of this was removed and mixed with an equal volume of a solution containing a vast excess of the matrix. 2 μL of the final mixture was dried on a sample plate and inserted into the vacuum of a Voyager-DE™ PRO Biospectrometry™ Workstation mass spectrometer. Protein standards of known molecular weights greater than, and less than, the molecular weight of the protein being analysed were also run to calibrate the instrument. The proteins used were Chymotrypsinogen (25657.1 Da) and Apomyoglobin (16952.5 Da). MALDI-TOF is a relatively simple instrument to use compared to Electrospray. However, the downside is that the accuracy of the technique is compromised (especially if calibration is not performed carefully) and errors of approximately 50 Da in the mass measurement can be expected.

Chapter 3 - Comparison of parasite-derived ABA-1 and recombinant ABA-1 proteins.

3.1 Introduction

The nematode polyprotein allergens (NPAs) are a family of helical proteins, which bind fatty acids and retinoids. The human parasite roundworm, *Ascaris lumbricoides* and the pig parasitic roundworm *Ascaris suum* secrete *a*scaris *b*ody-fluid *a*llergen (ABA-1). ABA-1 is the most abundant NPA in the pseudocoelomic fluid of the nematode *Ascaris*. Since *Ascaris suum* can be obtained from pigs in a local abattoir it is the easiest to acquire. Other parasites producing NPAs do not produce the allergens in sufficient quantity that would provide adequate amounts of proteins for crystallisation trials.

In this chapter the protein purification protocol and subsequent purity of ABA-1 obtained from the parasite *Ascaris* (hereafter denoted pABA-1) is compared to the recombinant ABA-1 (rABA-1). Crystallisation of ABA-1 from both sources is attempted using a large range of conditions. Failure to obtain crystals suitable for diffraction is investigated using dynamic light scattering (DLS).

3.2 Purification of parasite derived wild-type ABA-1

Ascaris body fluid (ABF) was collected from Prof. M. W. Kennedy's laboratory, University of Glasgow, were it had been removed from adult *Ascaris suum* obtained from a local abattoir. The large fat females were separated from the small curly tailed males. In a fumehood using a scalpel blade, the end of the female worm was removed and the body fluid allowed to drip into an Eppendorf. The body fluid was subsequently spun at 13000 rpm for 10 minutes then filtered through a 0.45 μm sterile filter to remove any fine particulate matter. The resulting fluid was stored at –70 °C.

ABF was purified to yield parasite-derived ABA-1 protein (pABA-1). A Superose 12 gel-filtration column mounted on a Pharmacia FPLC system was equilibrated with 0.15 M NaCl and 0.05 M K_2PO_4, pH 7.0. ABF was centrifuged at 13000 rpm for 10 minutes to remove fine particulate matter then loaded onto the Superose 12 column at 0.3 ml/min, 1 ml at a time to achieve the best separation of the different proteins present (figure 3.1). Protein was detected at 280 nm and by SDS-PAGE analysis (figure 3.1).

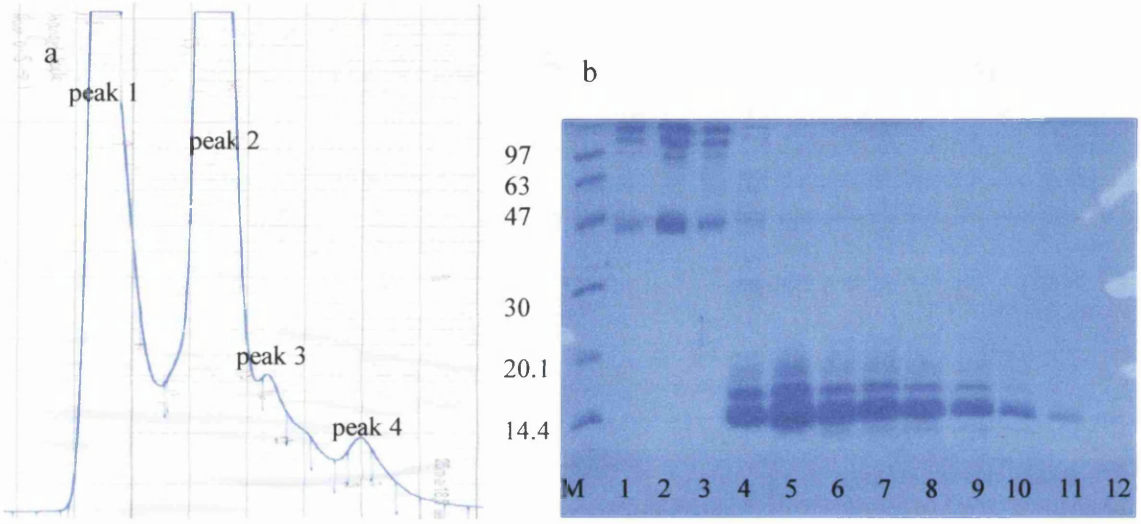

Figure 3.1 a) chart showing the 4 peaks eluted from ABF
b) M = markers, 1-3 = eluted in peak 1, 4-8 = eluted in peak 2 containing pABA-1,
9-10 = eluted in peak 3 containing pABA-1 and 11-12 = eluted in peak 4 containing pABA-1.

The fractions containing pABA-1, lanes 4-12 on SDS-PAGE (figure 3.1b) were pooled and dialysed into 50 mM HEPES pH 7.5 in preparation for loading onto an ion exchange column on a BioCAD Sprint system (Perfusion chromatography systems). A Hitrap Q anion exchange column was used, as homologues of ABA-1 have isoelectric points (pI) around 5.2. It was therefore expected that at pH 7.5 the proteins would have a negative charge and bind to the anion exchanger and subsequently be eluted individually with a salt gradient. The anion exchange column was washed with 2 M NaCl, then 50 mM HEPES pH 7.5 before the protein sample was loaded to remove any previously bound proteins and then given a final wash with 3 column volumes

of 50 mM HEPES pH 7.5. The protein was detected at 280 nm (figure 3.2). 5 ml of protein sample was loaded onto the column at 2 ml/min. Unbound protein was collected in flow through. A continuous gradient from 0 to 100 % in 50 mls was introduced using 2 M NaCl, to gradually elute off protein that had bound to the column.

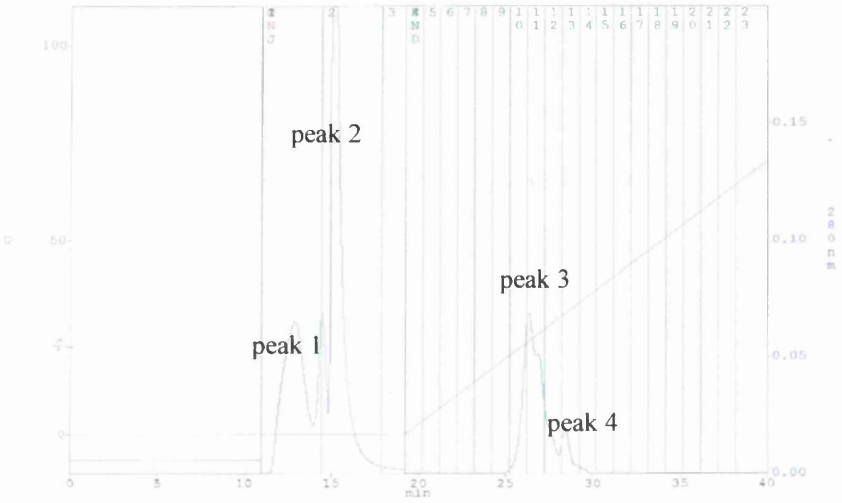

Figure 3.2 Chart of pABA-1 run on HQ anion exchange column using the Sprint system

SDS-PAGE (figure 3.3) was run on all fractions that eluted from the HQ column.

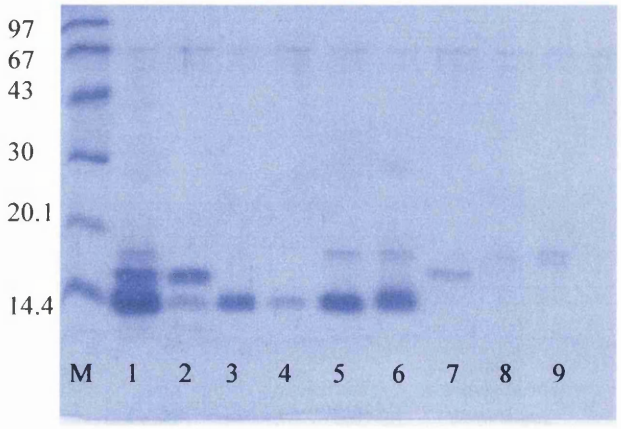

Figure 3.3 M = markers, 1 = sample loaded, 2 = eluted in peak 1 of flow through, 3-4 = eluted in peak 2 of flow through, 5-6 = eluted in peak 3, 7 = eluted in shoulder of peak 3 and 8-9 = eluted in peak 4

After anion exchange the main peak 3 (figure 3.2) has two bands present shown in lanes 5 and 6 of SDS-PAGE (figure 3.3), one close to the molecular weight of ABA-1 (15.1 kDa) and another of higher molecular weight, approximately 17 kDa, which is likely to be an isomer of ABA-1 (chapter 1.3) and could not be separated from the 15.1 kDa protein, despite further purification attempts. A single band is present in lanes 3 and 4 of SDS-PAGE at approximately 14 kDa, which is close to the calculated molecular weight of ABA-1. Other ion exchange columns (HS, mono S and mono Q) were also tried in an attempt to separate the 17 kDa protein and the charged isoforms (figure 3.8) of pABA-1 without success.

It has been shown (Kennedy, 1995b) that ABA-1 exhibits binding to fatty acids with a blue shift in fluorescence. Therefore, pABA-1 eluted in the second flow through peak (lanes 3 and 4 on gel) was assayed using DAUDA, a fluorescent labelled fatty acid known to bind to NPA proteins giving a blue shift. The fluorescence spectrum shows clearly the expected blue shift confirming DAUDA has bound to pABA-1 eluted in the flow through and that the protein has retained activity.

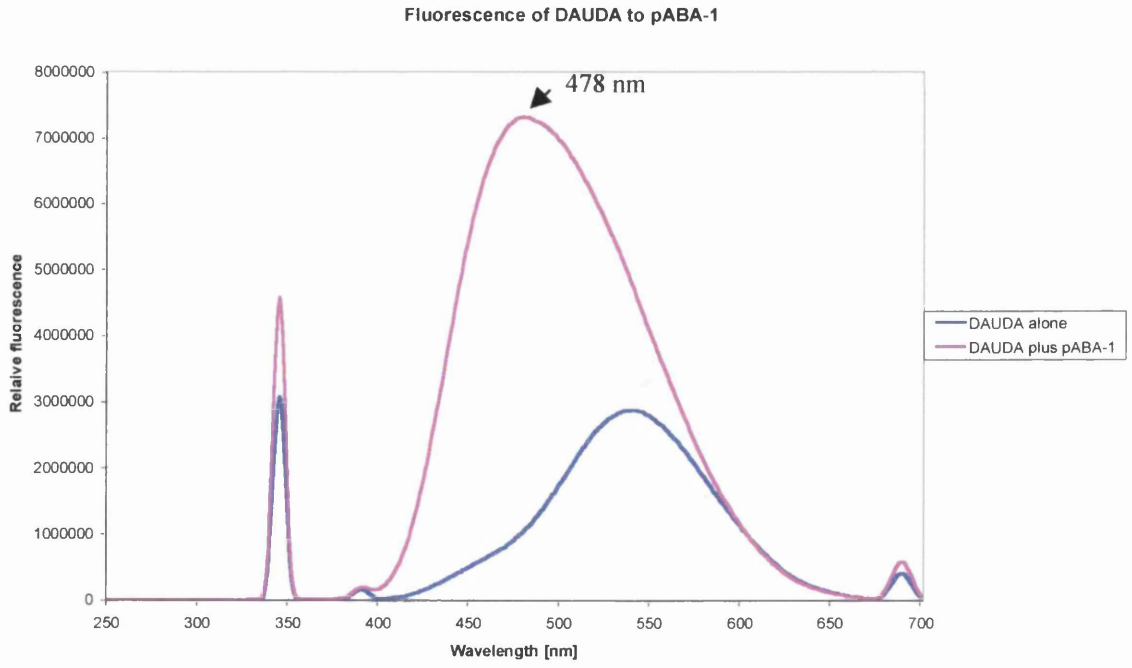

Figure 3.4 Binding of the fluorescent labelled fatty acid DAUDA to pABA-1

3.3 Overexpression and purification of recombinant ABA-1

N. Meenan of Prof. Cooper's and Prof. Kennedy's lab, University of Glasgow, supplied a pGEX-2T plasmid encoding the antigen of ABA-1.

The plasmid was extracted and transformed into BL21 host cells. The plasmid was inserted by heat shock. 100 µl of host strain BL21 was placed in a sterile Eppendorf tube on ice. 1 µl β-mercaptoethanol was added and left on ice for 2 minutes, then 5 µl of plasmid was added to the tube and left on ice for 1 hour. A water bath was heated to 42 °C and the cells placed in it for 30 seconds and then on ice to cool back down to 4 °C. 400 µl of LB (Millipore) was then added and the tube was incubated at 37 °C for 1 hour. The cells were plated on agar with ampicillin and incubated overnight at 37 °C. After incubation 10 ml of LB 200 µg/ml ampicillin was inoculated with a single colony from the plate and incubated overnight at 37 °C with rotary shaking at 200 rpm. Glycerol stocks were made using 850 µl of culture and 150 µl of 80 % glycerol and stored at –80 °C.

A 1 litre flask containing 500 ml of 2-YT broth (Millipore) and 200 µg/ml ampicillin was inoculated with a glycerol stock and grown overnight at 37 °C with shaking at 200 rpm. This overnight broth was added to 2 more litre flasks containing 1 litre 2-YT broth and 200 µg/ml ampicillin each and grown at 37 °C at 200 rpm for approximately 3 hours until cultures had grown to an A_{600nm} between 0.6–1. 1 mM dioxane free Isopropyl β-D-thiogalactoside (IPTG) was added to induce gene expression. Cells were grown for a further 3 hours at 37 °C with shaking at 200 rpm. The culture was transferred to appropriate centrifuge bottles and centrifuged at 8000 rpm for 15 minutes to sediment the cells and the supernatant was removed.

Each pellet was then resuspended in 100 ml PBS with a Roche EDTA free protease inhibitor tablet. The cells were disrupted by sonication on ice using a

Status US200 sonicator (Philip Harris Scientific) for six, 30 second blasts at half power. Cell disruption was evidenced by partial clearing of the suspension. The sonicate was centrifuged at 10000 rpm for 10 minutes and the supernatant transferred to a fresh container. 1 ml of Glutathione Sepharose 4B was added to 50 ml of supernatant and incubated at room temperature with gentle rocking for an hour. After an hour the supernatant was gently spun at 500 rpm for 10 minutes and the resin was washed 3 times with PBS. Finally, the protein was released from the GST by cleavage with thrombin protease, 50 units of thrombin per 1 ml of resin. After 16 hours at room temperature the resin was spun down at 200 rpm for 5 minutes. To remove the thrombin protease, the protein was run through a Benzamidine FFTrap column (Amersham Pharmacia Biotech) at 0.5 ml/min on an AKTA purifier 100 system (Amersham Pharmacia Biotech) followed by gel filtration using a Superose 12 column with elution buffer 50 mM HEPES pH 7.5 run at 0.3 ml/min. Protein was detected in the supernatant at 280 nm (figure 3.5) and by SDS-PAGE (figure 3.6).

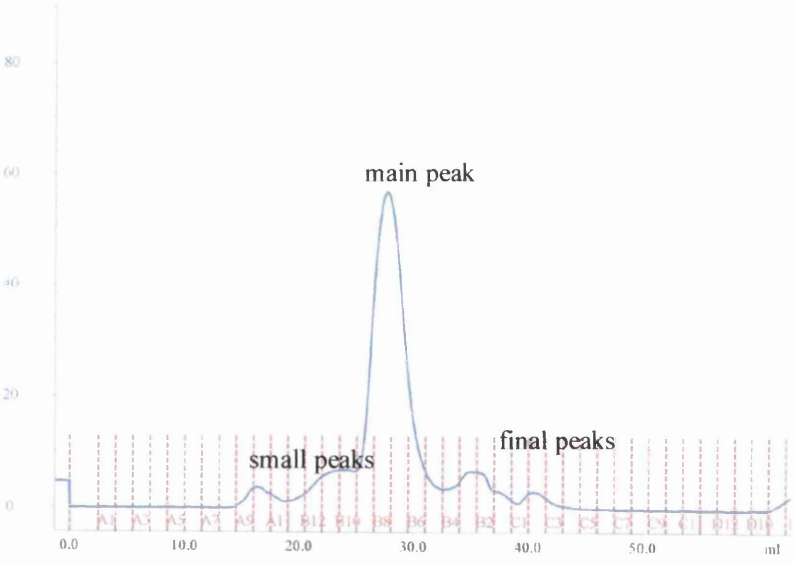

Figure 3.5 Gel filtration of rABA-1 on the superose 12 column on the AKTA

SDS-PAGE (figure 3.6) was run on all fractions that eluted from the Superose 12 gel filtration column.

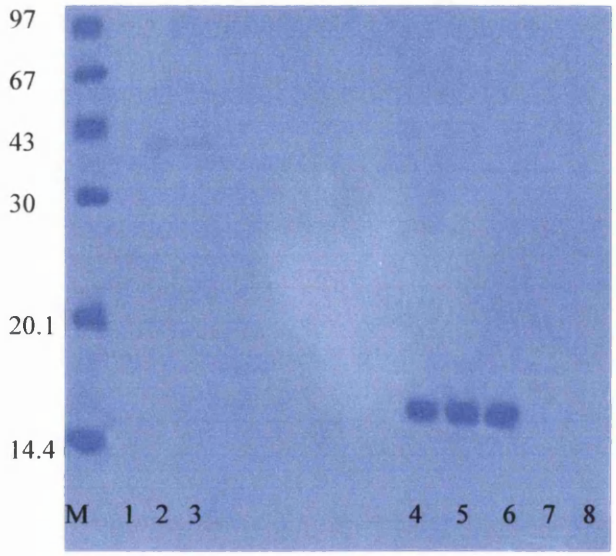

Figure 3. 6 M = markers, 1-3 = proteins eluted from initial small peaks, 4-6 = proteins eluted from main peak, including rABA-1 and 7-8 = protein eluted from final peaks

The resulting SDS-PAGE showed bands of approximately 15 kDa in lanes 4-6, which is close to the predicted molecular weight of rABA-1 (15.1 kDa). These fractions were assayed for DAUDA binding (figure 3.7) and were found to exhibit fatty acid binding activity confirming functional ABA-1 to be present.

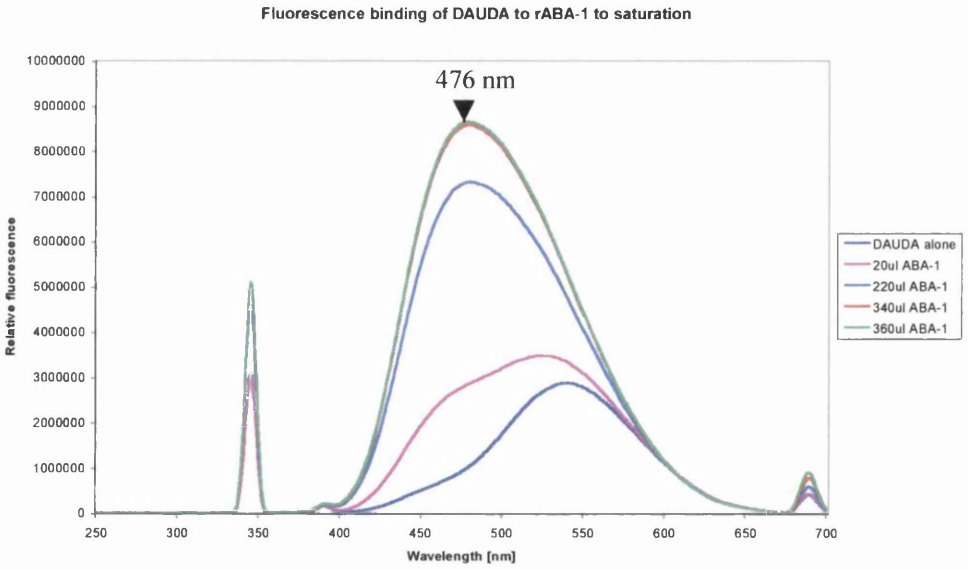

Figure 3.7 Binding of fluorescent labelled fatty acid DAUDA to rABA-1

Having successfully purified rABA-1 and parasite-derived ABA-1 an isoelectric focusing gel (figure 3.8) was run to determine if charged isoforms, which are not readily seen on SDS-PAGE were present. Other proteins studied in the project are also shown and will be referred to in subsequent chapters.

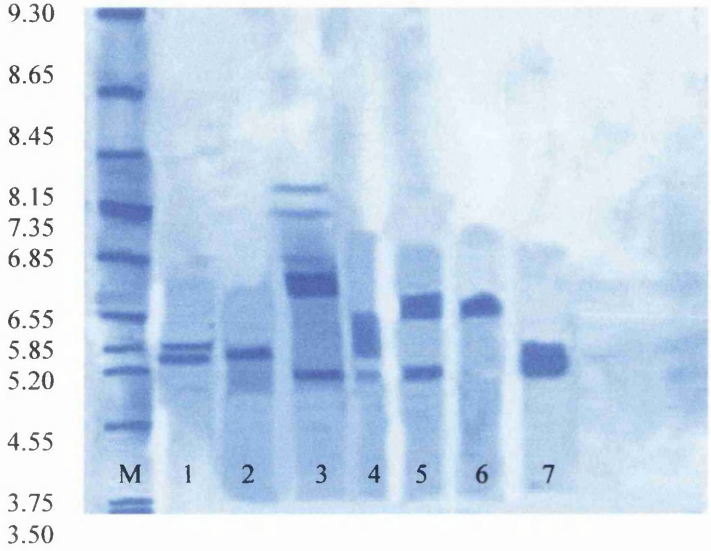

Figure 3.8 Novex® IEF gel. M = markers, 1 = rABA-1, 2 = rGpFAR1, 3 = rCeFAR2, 4 = rOvFAR1, 5 = rCeFAR3 before ion exchange, 6 = rCeFAR3 after ion exchange and 7 = pABA-1.

The resulting IEF gel (figure 3.8) indicates that parasite-derived ABA-1 in lane 7 (samples after ion exchange on HQ column) consists of several, differently charged isoforms of ABA-1 grouped together in a broad band that have similar pIs. Having tried ion exchange (chapter 3.2) it is difficult, if not impossible, to achieve separation into homogenous protein. Recombinant ABA-1 (lane 1), on the other hand, consists of two isoforms and appears as a single band on SDS-PAGE. Compared to pABA-1, rABA-1 is easier to produce and purify in large quantities for structural and biophysical studies. Upon interaction with DAUDA (figures 3.4 and 3.7) both ABA-1 proteins undergo dramatic blue shifts in the wavelength of maximum fluorescence emission from 542 nm for DAUDA in buffer to 478 nm and 476 nm for DAUDA interacting with pABA-1 and rABA-1 respectively, showing that the rABA-1 is similar, if not identical, in binding to the pABA-1.

3.4 Circular dichroism of pABA-1 and rABA-1

Circular dichroism (CD) analysis was carried out as described in chapter 2. The secondary structure, derived from the far UV spectrum of both pABA-1 and rABA-1 (figure 3.9), show high α-helical content.

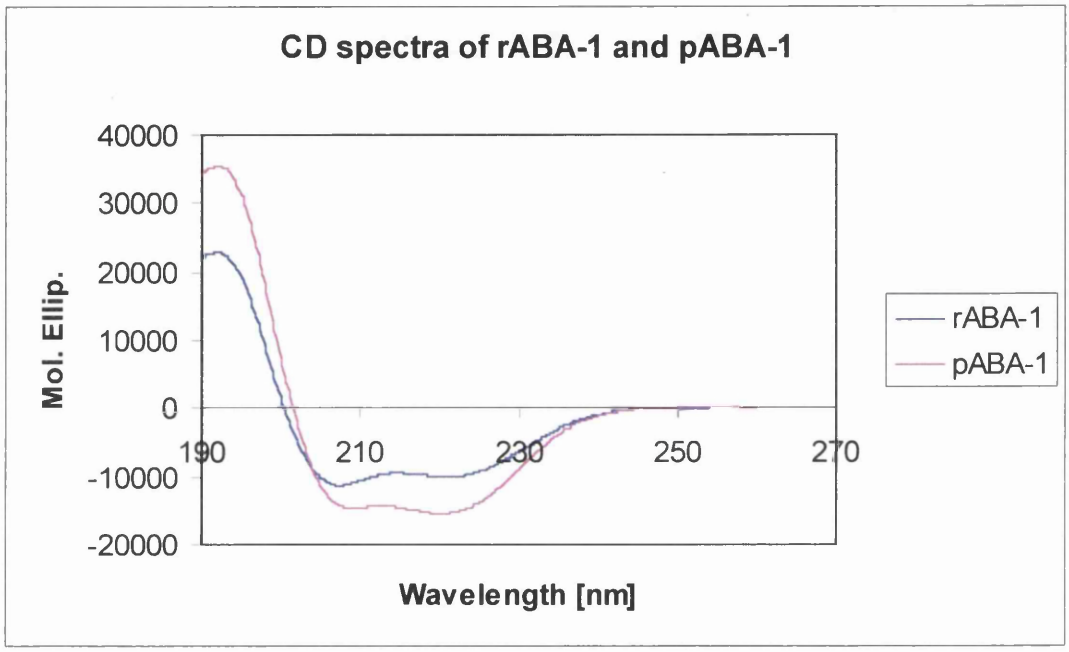

Figure 3.9 CD of the far UV spectra of rABA-1 and pABA-1

The far UV spectra of the ABA-1 proteins show that they have similar shape and that the secondary structures of pABA-1 and rABA-1 are very similar. The differences in the spectra may be due to concentration errors or the presence of the number of charged isoforms in pABA-1.

The secondary structure content of pABA-1 and rABA-1 was estimated using the CONTIN procedure (table 3.1) of Provencher and Glöckner (1981).

Protein	Secondary structure predictions using CONTIN			
	α-helix	β-sheet	β-turn	remainder
pABA-1	53 %	29 %	6 %	12 %
rABA-1	46 %	37 %	5 %	12 %

Table 3.1 Secondary structure contents of pABA-1 and rABA-1 estimated using the CONTIN procedure

The CD spectra show that both pABA-1 and rABA-1 have folded secondary structure with a higher percentage of α-helix, 53 and 46 % than β-sheet, 29 and 37 % for pABA-1 and rABA-1 respectively. This agrees with previous CD studies carried out on wild-type ABA-1, purified from the pseudocoelomic fluid of *Ascaris suum* by chromatofocussing and size-exclusion chromatography, which gave 59 % α-helix, 35 % β-sheet and 6 % remainder (Kennedy *et al*, 1995b). The main reason for the discrepancies between this study and that of Kennedy (1995b) is due to the difficulty in obtaining values for protein concentration, which in turn affects the secondary structure calculations. Other workers in this field (Kennedy and Garofalo) have also experienced this problem, which is possibly due to inaccurate protein molecular weight and the presence of DNA. However this data shows that ABA-1 is structurally different from the more ubiquitous lipid binding proteins from the family of cytosolic proteins, which are predominantly 10 anti-parallel β-stranded with 2 short α-helices surrounding the interior binding cavity in a β-barrel type structure (Banaszak *et al*, 1994).

3.5 Dynamic light scattering (DLS) on rABA-1

DLS analysis was only carried out on rABA-1 since IEF studies already indicated that pABA-1 was a highly heterogeneous protein. The DLS experimental was carried out for rABA-1 as described in chapter 2. 2 mg/ml rABA-1 was

investigated in 20 mM ACETATE pH 4.5, 20 mM PIPES pH 6.5, 20 mM HEPES pH 7.5 and 20 mM TRIS pH 8.5, with 150 mM NaCl present in all. Each of the above buffers was also studied at temperatures between 5 and 25 °C, in 5° increments. Lysozyme was used as standard. (Appendix 1, DLS results)

rABA-1 aggregated in acetate, therefore no results were obtained in this buffer. The bimodal distribution of rABA-1 in 20 mM HEPES pH 7.5 from DLS was analysed in a bimodal scatter diagram (figure 3.10). A gaussian graphical presentation (figure 3.11) for rABA-1 in 20 mM HEPES pH 7.5 is also shown. These diagrams give an indication of the aggregation and oligomerisation state of the protein (www.cbrweb.med.harvard.edu/investigators/protocols/sara_light_scattering).

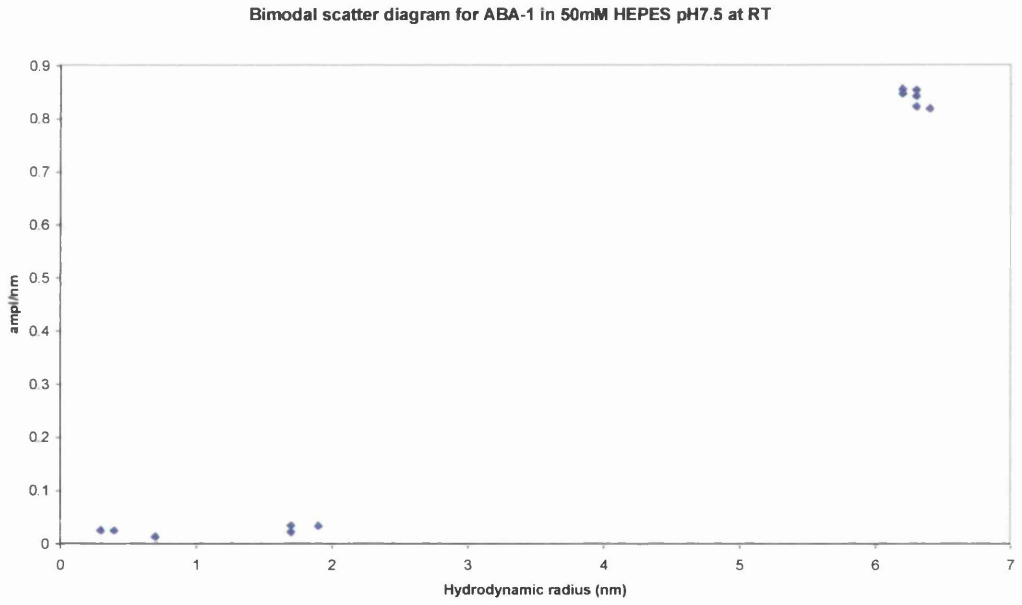

Figure 3.10 Bimodal scatter diagram of rABA-1 in 20 mM HEPES pH 7.5.

The resulting bimodal scatter diagram indicates that more than one species is present in the rABA-1 protein solution.

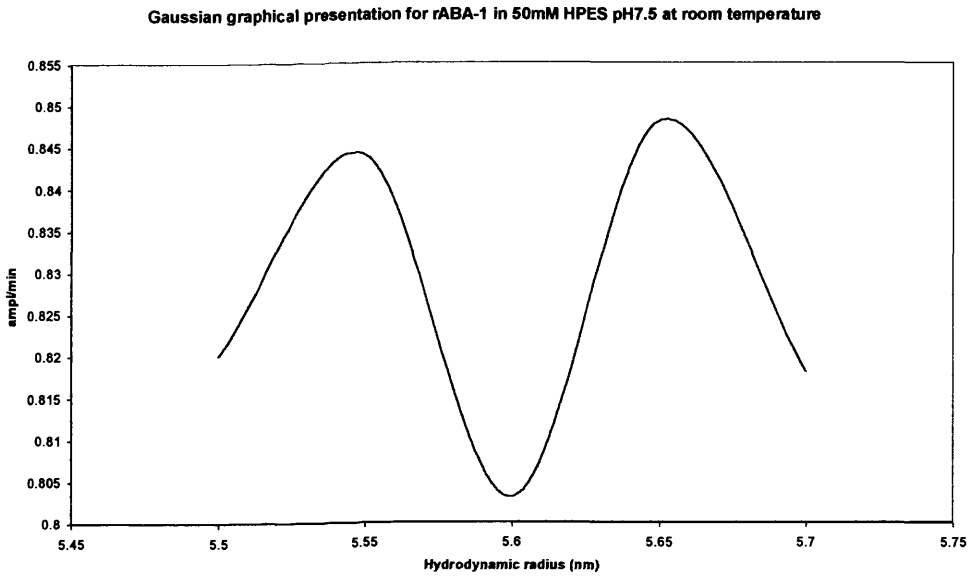

Figure 3.11 Gaussian graphical presentation for rABA-1 in 50 mM HEPES pH 7.5.

The gaussian graph agrees with the bimodal scatter diagram (figure 3.9), again indicating rABA-1 consists of two species possibly monomer or dimer.

The extent of polydispersity of a protein solution is often regarded as a good indicator as to the likelihood that crystallisation may occur (D'Arcy, 1994).

Polydispersity (%) = polydynamic radius (nm) / radius (nm) X 100

A polydispersity value of below 15 % indicates low or no polydispersity and a high likelihood of crystallisation of the protein occurring. Polydispersity of between 15 – 30 % indicates a moderate amount of polydispersity in which crystallisation is still possible and above 30 % indicates a largely heterogeneous solution with the likelihood that crystallisation may not take place.

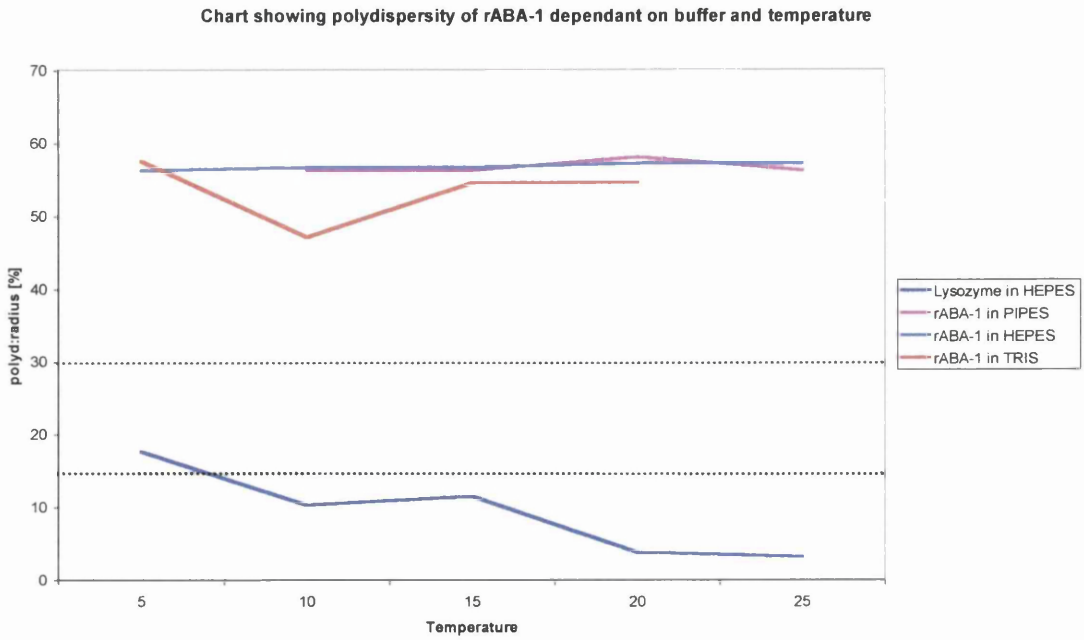

Figure 3.12 Polydispersity of lysozyme in HEPES and rABA-1 in PIPES, HEPES and TRIS

The polydispersity of rABA-1 (figure 3.12) is uniformly greater than 30 % between 5 and 25 °C in all buffers, which indicates a significant amount of polydispersity, or heterogeneity in the protein making crystallization an unlikely event. Baseline errors (appendix 1) indicate rABA-1 has multimodal size distribution, SOS errors (appendix 1) show high background noise in the sample, again indicating the protein is not homogenous and unlikely to crystallise. The results for lysozyme in HEPES pH 7.5 show moderate to negligible polydispersity and narrow monomodal size distribution with low noise errors, indicating that the lysozyme standard is homogenous and likely to crystallise (which it does).

The molecular weight of rABA-1 (figure 3.13) in each solution may be estimated if we assume the protein to be perfectly globular. If the protein is not globular, then this measurement roughly corresponds to half the maximal dimension of the protein and may infer the oligomerisation state of the protein.

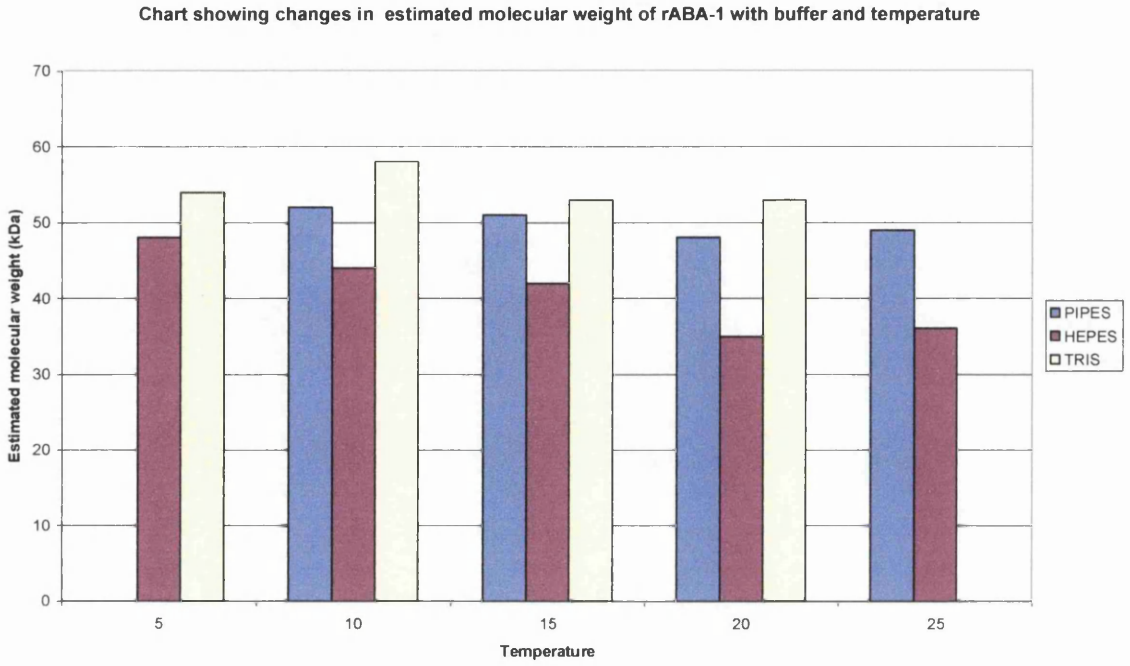

Figure 3.13 Estimated molecular weight of rABA-1 in PIPES, HEPES and TRIS between 5 and 25 °C

The molecular weight estimates derived from DLS suggests the protein has not aggregated. The lowest estimated molecular weight of 35 kDa may suggest one of two things, firstly, there may be a contribution from a dimer moiety (30 kDa) or secondly, rABA-1 is not a globular protein. That it is not a globular protein has been confirmed by the initial NMR studies, which suggest that ABA-1 is a two domain structure that is flat in one direction (N. Meenan, A. Cooper, private communication).

From these DLS experiments it was decided that probably the best elution buffer to be used in the final gel filtration step should be 20 mM HEPES pH 7.5.

3.6 Crystallisation trials on pABA-1 and rABA-1

Crystallisation trials were undertaken on both pABA-1 and rABA-1 since protein was available for both. The standard sitting drop vapour diffusion technique was used by mixing 4 µl of the protein solution with an equal volume of crystallisation solution in the drop and equilibrating against 1 ml of crystallising solution in the well. Initial trials were set up using the footprint screen (www.prg/stura/cryst/foot). Trials were set up at concentrations of pABA-1 and rABA-1 at 2, 5, 8, 10, 12, 15, 20 and 25 mg/ml. 21 mg/ml was found to be the optimum concentration for ABA-1 proteins as it resulted in fine precipitation in some of the drops. This high protein concentration suggests that the protein is fairly soluble and this is borne out by the fact that it is found in high concentrations in the pseudocoleomic fluid of the nematode. Trials were therefore set up with 4 µl of 21 mg/ml pABA-1 and rABA-1 using the standard sitting drop vapour diffusion technique and crystallising solutions from a large number of sparse matrix conditions (Brzozowski & Walton, 2001; Jancarik & Kim, 1991; Cryo I and II, Wizard I and II, www.emeraldbiostructures.com; Crystal Clear Strategy Screen I and II, www.moleculardimensions.com; www.hamptonresearch.com). Cryo screen 2, condition 9 (2 M ammonium sulphate and phosphate-citrate pH 4.2) gave small (20 µm) needle crystals, but optimisation failed to produced larger crystals. Frequently extensive precipitation and crystalline aggregates resulted and even though these conditions were followed up, no protein crystals suitable for diffraction were obtained. On several occasions microseeding techniques were used, but again suitable crystals were not obtained.

3.7 Discussion

Parasite derived ABA-1 shows multiple isoforms on SDS-PAGE and IEF gels. Despite attempts to separate the isoforms using ion exchange and gel filtration

chromatography, it was not possible to obtain homogenous protein. Although recombinant ABA-1 showed one band on SDS-PAGE there are two charged isoforms present on the IEF gel. Both pABA-1 and rABA-1 show an increase and blue shift in fluorescence emission when bound to DAUDA. With equivalent amounts of pABA-1 and rABA-1 shifts from 542 nm to 478 nm and 476 nm respectively were observed. Blue shifting of a dansyl fluorophore emission wavelength is indicative of a hydrophobic binding site (MacGregor *et al*, 1986). The DAUDA shifts obtained for both suggest that the binding site for rABA-1 is similar, if not identical, to that of pABA-1. Since rABA-1 is easier to obtain and can be produced in larger quantities, with better purity, it was decided to use rABA-1 for biophysical studies in preference to pABA-1.

Circular dichroism shows that both pABA-1 and rABA-1 have folded secondary structure with a higher percentage of α-helix, 53 and 46 % than β-sheet, 29 and 37 % for pABA-1 and rABA-1 respectively, which agrees with previous CD studies carried out on pABA-1, which calculated 59 % α-helix, 35 % β-sheet and 6 % remainder (Kennedy *et al*, 1995b).

Dynamic light scattering is a good indication of protein homogeneity, purity, aggregation state and monodispersity - all pre-requisites for successful crystallisation. DLS results on lysozyme indicate the technique can show when a protein is in the optimum buffer for crystallisation. DLS studies on rABA-1 suggest it is not optimized for crystallisation, as baseline errors indicate rABA-1 has multimodal size distribution and SOS errors show high background noise in the sample. There is also a significant amount of polydispersity of rABA-1 in all buffers indicating that the sample is not homogenous. The estimated molecular weight for rABA-1 (35 kDa), derived by DLS, was just over twice the actual molecular weight of rABA-1 (15.1 kDa) and suggests that the protein is not globular, confirmed by initial NMR studies suggesting that ABA-1 is a two domain structure (private communication, N. Meenan, A. Cooper). However, HEPES buffer seemed the best buffer to use at ambient temperatures. The

bimodal analysis showed that there may be two species present in the rABA-1 solution. This could indicate that rABA-1 may exist in a dimer conformation. The gaussian graph agrees with the bimodal scatter diagram information confirming two species present in rABA-1.

Crystallisation trials on both pABA-1 and rABA-1 showed that a high concentration of 21 mg/ml was needed to be close to supersaturation. This is a fairly high protein concentration and made it labour intensive to obtain enough protein, recombinant or parasite-derived for crystallisation trials. In the nematode, the concentration of pABA-1 in the pseudocoelomic fluid is high indicating a very soluble protein. This is confirmed by the 21 mg/ml required before supersaturation is approached. Having set up a large number of crystallisation trials without any great measure of success the protein solution was investigated using DLS, which indicated that rABA-1 was not homogenous. Two possible reasons for this are heterogeneity due to the isoforms of ABA-1 as seen on the IEF gel (figure 3.7) or heterogeneity due to a mixture of fatty acids present from *E. coli*, which may be in competition for the binding site. The current NMR studies (N Meenan, A Cooper, private communication) also show this ambiguity in fatty acid present in the binding site. The presence of these *E. coli* derived fatty acids is further investigated for the FAR proteins.

Chapter 4 – OvFAR1, CeFAR2 and CeFAR3. Expression, purification and fluorescence studies.

4.1 Introduction

The fatty acid and retinol-binding (FAR) proteins are a family of 20 kDa helical proteins, which bind fatty acids and retinoids with a high affinity. The proteins from this family are not available in sufficient quantity to provide adequate amounts of protein directly from the organisms, thus necessitating the use of recombinant protein in order to study their functional properties. The use of recombinant protein is also advantageous in that it enables easier purification compared to purifying the proteins from the nematodes themselves.

GpFAR1, the 18.8 kDa protein secreted by the potato cyst nematode *Globodera pallida* (see chapter 5), is the FAR family protein which has been characterised most by ligand binding studies (Prior, 2001). This chapter studies the ligand binding properties of other proteins in the FAR family, which have been overexpressed and purified. These are rCeFAR2, rCeFAR3 and rOvFAR1. This is to establish if the proteins rCeFAR2 and rCeFAR3 from the nematode *Caenorhabditis elegans* are good functional homologues of the rOvFAR1 protein, a major antigen of the causative agent of river blindness, *Onchocerca volvulus*.

4.2 Overexpression and purification of recombinant OvFAR1

Dr Antonio Garofalo of Prof. Bradley's lab in Nottingham University supplied a pET 15b plasmid encoding antigen of OvFAR1. For expression the plasmid was cloned into c41 cells. Transformation is the process of inserting the plasmid into a host strain. The plasmid was inserted into host strain c41 by heat shock. 100 µl of host strain c41 was placed in a sterile Eppendorf tube on ice. 1 µl β-mercaptoethanol was added and left on ice for 2 minutes, then 5 µl of plasmid was added to the tube and left on ice for 1 hour. A water bath was heated to 42 °C and

the cells were placed in it for 30 seconds and then replaced on ice to cool back down to 4 °C. 400 µl of LB (Millipore) was then added and the tube was incubated at 37 °C for 1 hour. The cells were then plated on agar plus ampicillin and incubated overnight at 37 °C. 12 ml of LB 200 µg/ml ampicillin was inoculated with a single colony from the plate and incubated overnight at 37 °C with rotary shaking at 200 rpm. Glycerol stocks were made with this culture. The next day two 2-litre flasks containing 500 ml of 2-YT broth 200 µg/ml ampicillin were inoculated with 5 ml of the overnight culture and incubated at 37 °C with rotary shaking at 200 rpm. When the cultures had grown to an A_{600nm} between 0.6–1 (about 3 hours) they were induced with 1 mM dioxane free Isopropyl β-D-thiogalactoside (IPTG) from Roche and incubated for a further 3 hours at 37 °C with rotary shaking at 200 rpm. The cells were then pelleted by centrifugation at 5000 rpm for 15 minutes at 4 °C. The supernatant was discarded and the cell pellets were stored at –70 °C until required. The pellets were resuspended in lysis buffer (5 mM imidazole, 50 mM TRIS and 300 mM NaCl) with a Roche EDTA free protease inhibitor tablet. The cells were lysed by sonication (Status US200, Philip Harris) on ice. Cells were sonicated 6 times at half power for 30 seconds. Cell disruption is observed by a partial clearing of the suspension. Sonicate was then centrifuged at 10000 rpm for 10 minutes at 4°C. The supernatant was transferred to a fresh beaker. This lysate was loaded onto an immobilised-metal affinity chromatography (IMAC) column on the Pharmacia FPLC. The column uses Nitrilotriacetic acid (Ni-NTA) to bind the 6 Histidine tag residue incorporated into the FAR protein, which was pre-equilibrated with the lysis buffer. The column was washed with lysis buffer until the A_{280} was stable (about 5 column volumes), followed by two different wash steps with wash buffer 1 (10 mM imidazole, 50 mM TRIS and 300 mM NaCl) and wash buffer 2 (75 mM imidazole, 50 mM TRIS and 300 mM NaCl), until the A_{280} was stable (figure 4.1a). Finally an elution buffer (250 mM imidazole, 50 mM TRIS and 300 mM NaCl) was washed onto the column to remove the bound protein (figure 4.1b). Protein was detected at 280 nm and by SDS-PAGE analysis (figure 4.2).

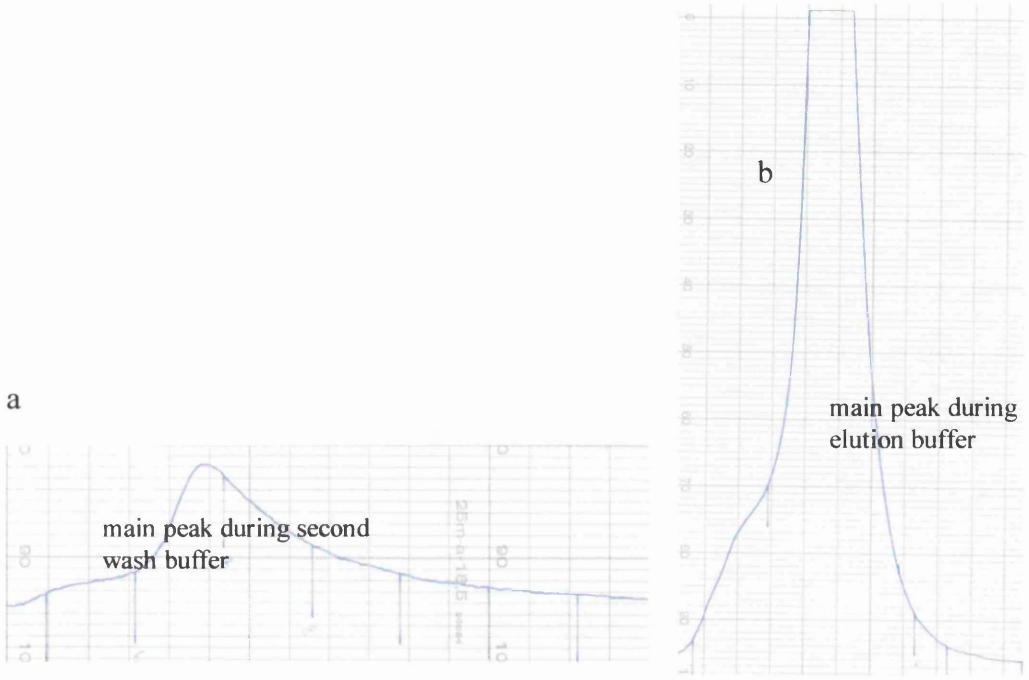

Figure 4.1 a) chart of the second wash buffer (containing 75 mM imidazole) step of Ni-NTA column during purification of rOvFAR1
b) Chart of the elution buffer (containing 250 mM imidazole) of Ni-NTA column during purification of rOvFAR1

SDS-PAGE (figure 4.2) was run on all fractions that eluted from Ni-NTA column.

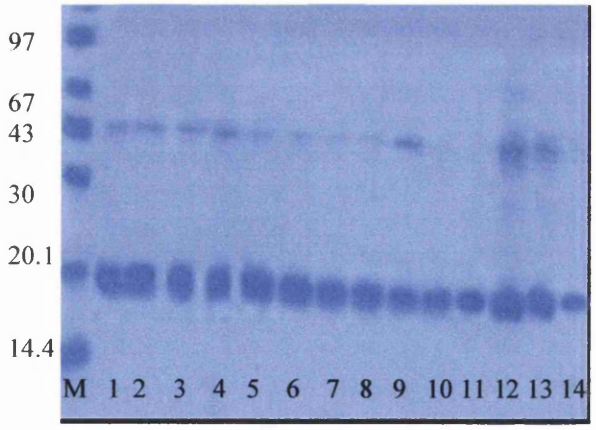

Figure 4.2 M = markers, 1-4 = rOvFAR1 protein eluted during first wash step of Ni-NTA column, 5-10 = rOvFAR1 protein eluted during second wash step of Ni-NTA column (figure 4.1a) and 11-14 = rOvFAR1 protein eluted during elution step of Ni-NTA column (figure 4.1b)

The resulting gel showed rOvFAR1 was present in all fractions, but as a smear of bands of approximately 20 kDa which is close to the predicted molecular weight for rOvFAR1. As a final step, protein was loaded onto a pre-equilibrated Superose 12 gel filtration column (figure 4.3) attached to an AKTA purifier 100 (Amersham Pharmacia Biotech) and loaded at 0.3 ml/min and eluted in 50 mM HEPES pH 7.5.

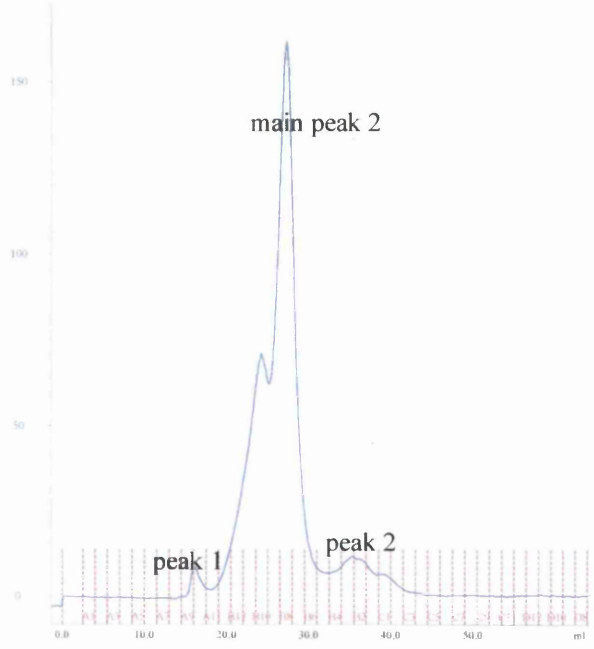

Figure 4.3 Chart showing peaks obtained from the gel filtration of rOvFAR1

The gel filtration yielded 3 peaks with the main peak 2 having a shoulder peak. Samples of each peak were run on SDS-PAGE (figure 4.4).

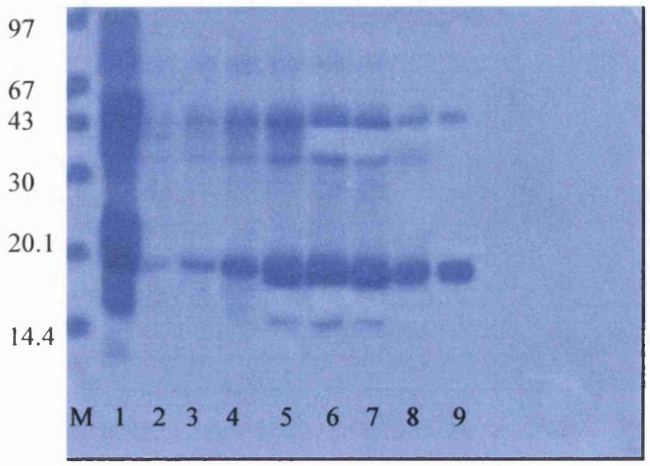

Figure 4.4 M = markers, 1 = sample loaded, 2-3 = protein eluted in peak 1 of gel filtration, 4-7 = protein including rOvFAR1 eluted from main peak 2 of gel filtration and 8-9 = protein including rOvFAR1 eluted in final peak 3 of gel filtration

rOvFAR1 appears purer in lanes 8 and 9 on the SDS-PAGE, with a band of 20 kDa, which is close to the molecular weight of rOvFAR1. A higher molecular weight band is present at about 43 kDa. Although this was a denaturing gel analysis by MALDI-TOF showed (see chapter 6) a dimer to be present and also the presence of higher oligomeric states. This higher molecular weight band could not be removed by gel filtration.

MALDI-TOF analysis also indicates that possible lipid groups are present, bound to rOvFAR1 (chapter 6). Subsequently (chapter 6), rOvFAR1 has been shown by DLS studies to be heterogeneous.

Further purification of rOvFAR1 was attempted by loading onto a Hitrap Q anion exchange column on a Sprint (Biocad) (figure 4.5) in 50 mM HEPES pH 7.5. It was hoped that at pH 7.5 the protein would have a negative charge and bind to an anion exchanger and subsequently eluted by a salt gradient. The anion exchange column was washed with 2 M sodium chloride before the protein sample was loaded to remove any previously bound material and then equilibrated with 3

column volumes of 50 mM HEPES pH 7.5. The protein was detected at 280 nm. 5 ml of protein sample, from lane 8 of figure 4.4, was loaded onto the column at 2 ml/min. Unbound protein was collected in the flow through. Bound protein was eluted using a salt gradient of 0 – 100 % 50 mM HEPES pH 7.5, 2 M NaCl.

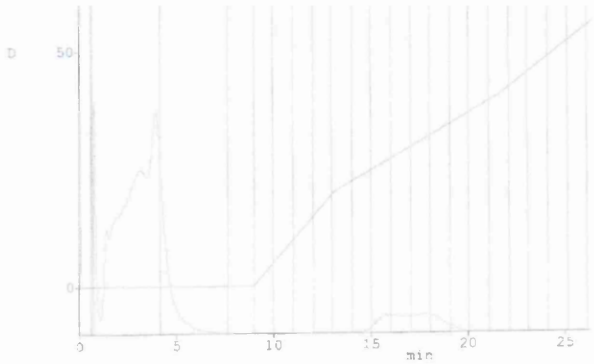

Figure 4.5 Chart of rOvFAR1 run on HQ anion exchange column using the Sprint system

SDS-PAGE (figure 4.6) was run on all fractions that eluted from the HQ column.

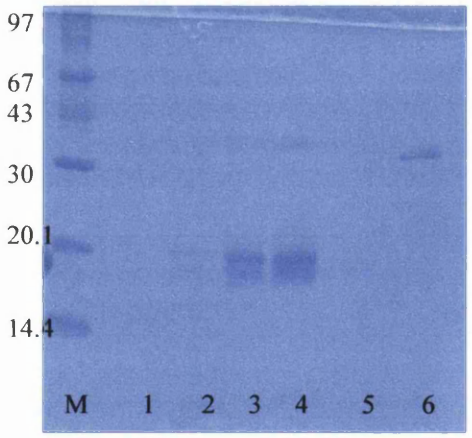

Figure 4.6 M = markers, 1-4 = flow through and 5-6 = proteins eluted in 25-40 % NaCl

The resulting gel after anion exchange of rOvFAR1 shows 2 bands in lanes 3-4 of 20 kDa which are close in molecular weight to the protein OvFAR1. A higher molecular weight band at approximately 40 kDa is also present, possibly a non-

denatured dimer. Ion exchange of rOvFAR1 did not significantly improve the purification of rOvFAR1 and subsequently was not used as a purification step.

4.3 Overexpression and purification of recombinant CeFAR2

Dr Antonio Garofalo of Prof. Bradley's lab in Nottingham University supplied a pET 15b plasmid encoding antigen of CeFAR2. For expression the plasmid was cloned into c41 cells. The plasmid was inserted into host strain c41 by heat shock. 100 µl of host strain c41 was placed in a sterile Eppendorf tube on ice. 1 µl β-mercaptoethanol was added and left on ice for 2 minutes, then 5 µl of plasmid was added to the tube and left on ice for 1 hour. A water bath was heated to 42 °C and the cells were placed in it for 30 seconds and then replaced on ice to cool back down to 4 °C. 400 µl of LB (Millipore) was then added and the tube was incubated at 37 °C for 1 hour. The cells were then plated on agar plus ampicillin and incubated overnight at 37 °C. 12 ml of LB 200 µg/ml ampicillin was inoculated with a single colony from the plate and incubated overnight at 37 °C with rotary shaking at 200 rpm. Glycerol stocks were made with this culture. The next day two 2-litre flasks containing 500 ml of 2-YT broth 200 µg/ml ampicillin were inoculated with 5 ml of the overnight culture and incubated at 37 °C with rotary shaking at 200 rpm. When the cultures had grown to an A_{600nm} between 0.6–1 (about 3 hours) they were induced with 1 mM dioxane free Isopropyl β-D-thiogalactoside (IPTG) from Roche and incubated for a further 3 hours at 37 °C with rotary shaking at 200 rpm. The cells were then pelleted by centrifugation at 5000 g for 15 minutes at 4 °C. The supernatant was discarded and the cell pellets were stored at –70 °C until required. The pellets were resuspended in lysis buffer (5 mM imidazole, 50 mM TRIS and 300 mM NaCl) with a Roche EDTA free protease inhibitor tablet. The cells were lysed by sonication (Status US200, Philip Harris) on ice. Cells were sonicated 6 times at half power for 30 seconds. Cell disruption is observed by a partial clearing of the suspension. Sonicate was then

centrifuged at 10000 rpm for 10 minutes at 4°C. The supernatant was transferred to a fresh beaker. This lysate was loaded onto an immobilised-metal affinity chromatography (IMAC) column on the Pharmacia FPLC. The column uses Nitrilotriacetic acid (Ni-NTA) to bind the 6 Histidine tag group incorporated onto the FAR protein. The column was pre-equilibrated with the lysis buffer. The column was washed with lysis buffer until the A_{280} was stable (about 5 column volumes), followed by two different wash steps with wash buffer 1 (10 mM imidazole, 50 mM TRIS and 300 mM NaCl) and wash buffer 2 (75 mM imidazole, 50 mM TRIS and 300 mM NaCl), until the A_{280} was stable. Finally an elution buffer (250 mM imidazole, 50 mM TRIS and 300 mM NaCl) was washed onto the column to remove the bound protein (figure 4.7a). Protein was detected at 280 nm and by SDS-PAGE analysis (figure 4.7).

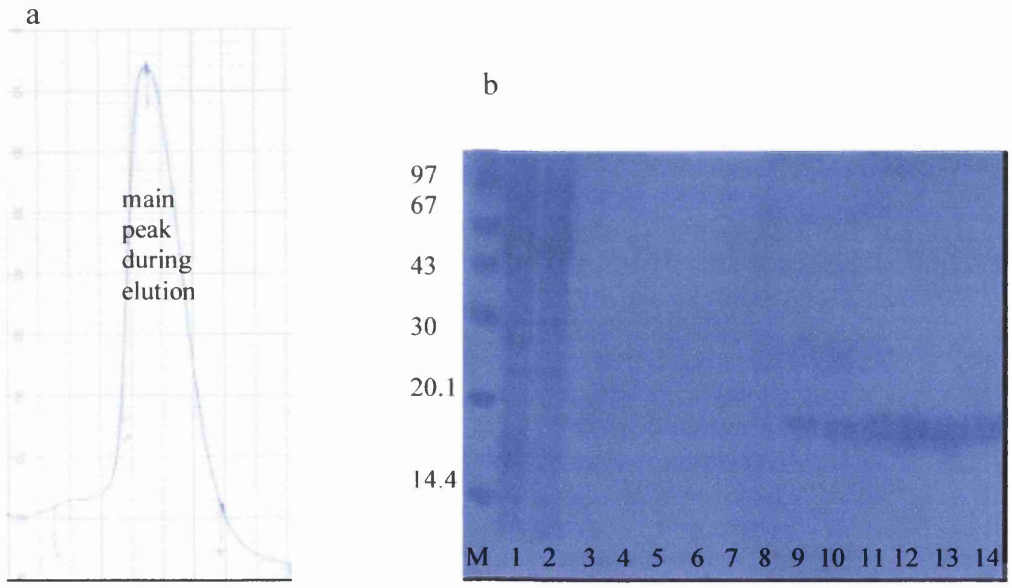

Figure 4.7 a) chart showing the main peak during elution (buffer containing 250 mM imidazole) from the Ni-NTA column

b) M = markers, 1-2 = eluted during wash buffer 1 (containing 10 mM imidazole) on the Ni-NTA column, 3-11 = eluted during wash buffer 2 (containing 75 mM imidazole) on Ni-NTA column and 12-14 = proteins including rCeFAR2 eluted during elution buffer step (containing 250 mM imidazole) on Ni-NTA column

Lanes 9 – 14 on SDS-PAGE show smears of bands at 20 kDa, which is close to the predicted molecular weight of rCeFAR2. These fractions were run on the Superose 12 gel filtration column as a final purification step. Protein was loaded onto a pre-equilibrated Superose 12 gel filtration column (figure 4.8a) attached to the Pharmacia FPLC system, loaded at 0.3 ml/min and eluted in 50 mM HEPES pH 7.5. One main peak eluted and the fractions were run on SDS-PAGE (figure 4.8b).

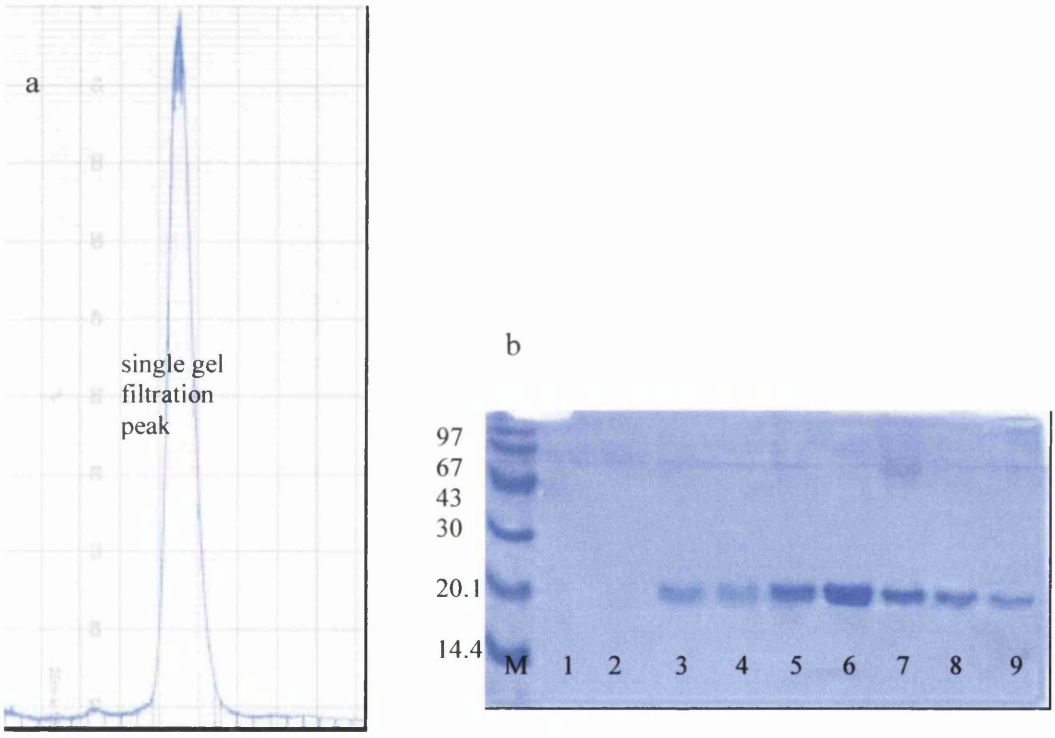

Figure 4.8 a) chart of rCeFAR2 eluting from gel filtration column as a single peak, b) M = markers, 1-9 = fractions eluted in main peak during gel filtration of rCeFAR2

Lanes 3-9 show bands of 20 kDa, which are close to the predicted molecular weight of rCeFAR2. MALDI-TOF indicates the protein to be close to its molecular weight of 20 kDa (chapter 6). Dynamic light scattering results indicate protein is not homogenous (chapter 6). Isoelectric focusing gel shows at least 2 distinct bands (figure 3.8, lane 3) indicating that rCeFAR2 is probably

heterogeneous. This could be due to degradation, loss of the end terminal amino acids or as is more likely, two species present in the sample, one species exists with lipid bound to rCeFAR2 and one without lipid bound to rCeFAR2 (chapter 6).

Further purification was attempted by ion exchange. rCeFAR2 in 50 mM HEPES pH 7.5 was loaded a Hitrap Q anion exchange column on an AKTA purifier 100 chromatography system (Amersham Pharmacia Biotech). Anion exchange column was washed with 2 M sodium chloride before the protein sample was loaded to remove any previously bound material and then equilibrated with 3 column volumes of 50 mM HEPES pH 7.5. Protein was detected at 280 nm. 5 ml of protein sample was loaded onto the column at 2 ml/min. Unbound protein was collected in the flow through. Bound protein was eluted using a 0 – 100 % salt gradient with 50 mM HEPES pH 7.5, 2 M NaCl. SDS-PAGE (figure 4.9) was run on all fractions that eluted from the column.

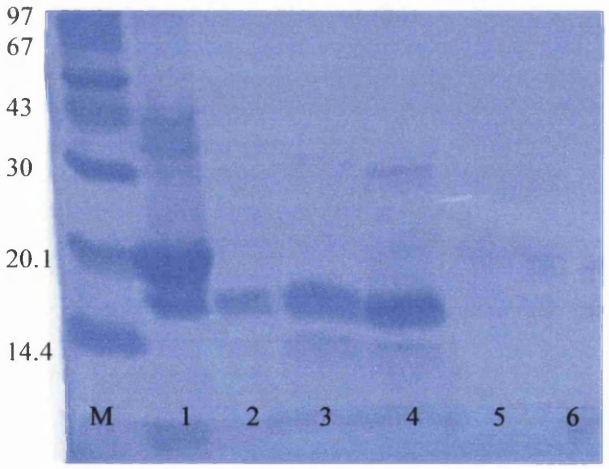

Figure 4.9 M= markers, 1 = sample loaded, 2-4 including CeFAR2 eluted from HQ column in 20-25% NaCl and 5-6 eluted from HQ column in 25-30% NaCl.

The resulting SDS-PAGE from anion exchange indicates that further attempts at purification were less than successful. Smeared bands in lanes 3 and 4 in figure 4.9 show that the protein has degraded. This step was subsequently not used in the purification of rCeFAR2.

4.4 Overexpression and purification of recombinant CeFAR3

Dr Antonio Garafalo of Prof. Bradley's lab in Nottingham University supplied a pET 15b plasmid encoding antigen of CeFar3. For expression the plasmid was cloned into c41 cells. The plasmid was inserted into host strain c41 by heat shock. 100 µl of host strain c41 was placed in a sterile Eppendorf tube on ice. 1 µl β-mercaptoethanol was added and left on ice for 2 minutes, then 5 µl of plasmid was added to the tube and it was left on ice for 1 hour. A water bath was heated to 42 °C and the cells were placed in it for 30 seconds and then replaced on ice to cool back down to 4 °C. 400 µl of LB (Millipore) was then added and the tube was incubated at 37 °C for 1 hour. The cells were then plated on agar plus ampicillin and incubated overnight at 37 °C. 12 ml of LB 200 µg/ml ampicillin was inoculated with a single colony from the plate and incubated overnight at 37 °C with rotary shaking at 200 rpm. Glycerol stocks were made with this culture. The next day two 2-litre flasks containing 500 ml of 2-YT broth 200 µg/ml ampicillin were inoculated with 5 ml of the overnight culture and incubated at 37 °C with rotary shaking at 200 rpm. When the cultures had grown to an A_{600nm} between 0.6–1 (about 3 hours) they were induced with 1 mM dioxane free Isopropyl β-D-thiogalactoside (IPTG) from Roche and incubated for a further 3 hours at 37 °C with rotary shaking at 200 rpm. The cells were then pelleted by centrifugation at 5000 rpm for 15 minutes at 4 °C. The supernatant was discarded and the cell pellets were stored at –70 °C until required. The pellets were resuspended in lysis buffer (5 mM imidazole, 50 mM TRIS and 300 mM NaCl) with a Roche EDTA free protease inhibitor tablet. The cells were lysed by sonication (Status US200, Philip Harris) on ice. Cells were sonicated 6 times at half power for 30 seconds. Cell disruption is observed by a partial clearing of the suspension. Sonicate was then centrifuged at 10000 rpm for 10 minutes at 4°C. The supernatant was transferred to a fresh beaker. This lysate was loaded onto an immobilised-metal affinity chromatography (IMAC) column on the Pharmacia FPLC. The column uses Nitrilotriacetic acid (Ni-NTA) to bind the 6 Histidine tag residue on FAR

proteins and was pre-equilibrated with the lysis buffer. The column was washed with lysis buffer until the A_{280} was stable (about 5 column volumes), followed by two different wash steps with wash buffer 1 (10 mM imidazole, 50 mM TRIS and 300 mM NaCl) and wash buffer 2 (75 mM imidazole, 50 mM TRIS and 300 mM NaCl), until the A_{280} was stable (figure 4.10a). Finally, an elution buffer (250 mM imidazole, 50 mM TRIS and 300 mM NaCl) was washed onto the column to remove the bound protein (figure 4.10b). Protein was detected by SDS-PAGE analysis (figure 4.11).

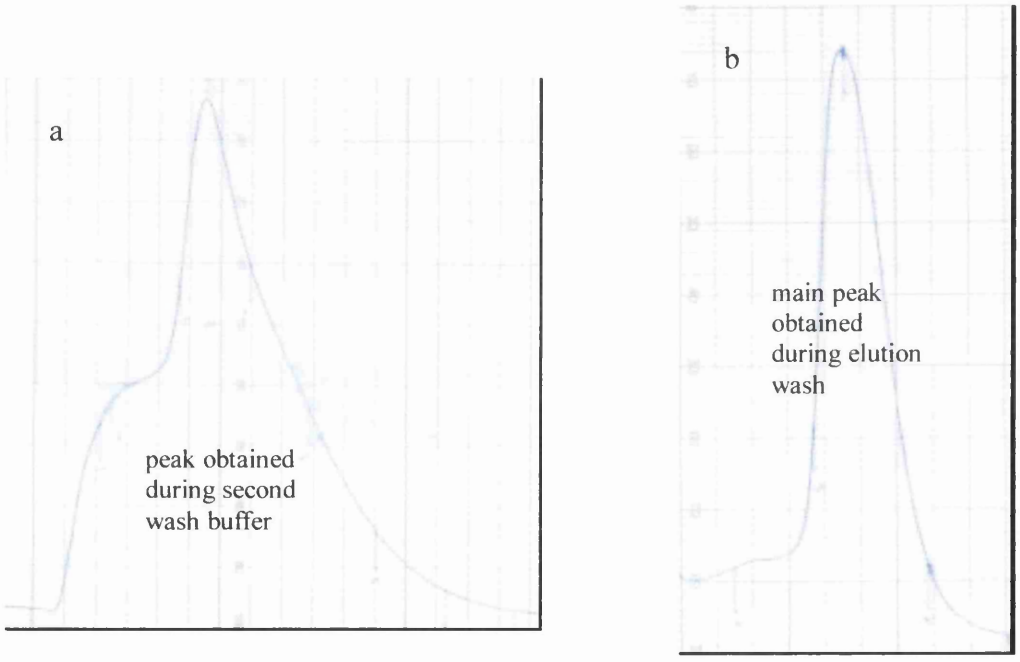

Figure 4.10 a) chart from wash buffer 2 (containing 75 mM imidazole) step of rCeFAR3 eluting from Ni-NTA column,
b) Chart from elution buffer (containing 250 mM imidazole) step of rCeFAR3 eluting from Ni-NTA column.

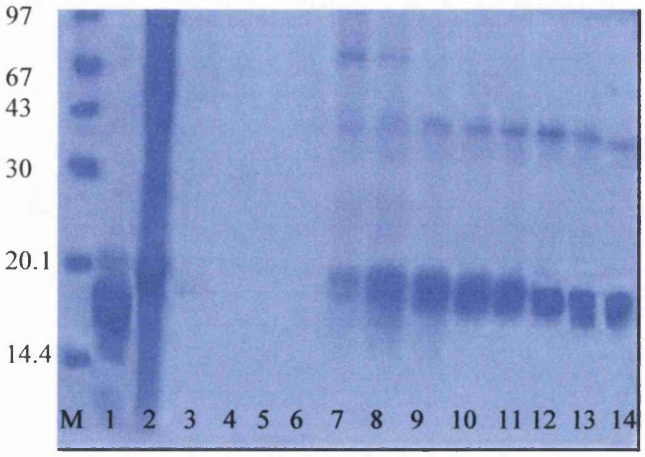

Figure 4.11 M = markers, 1-2 = eluted from wash buffer 1 (containing 10 mM imidazole), 3-11 = proteins eluted during wash buffer 2 (containing 75 mM imidazole) of rCeFAR3 from Ni-NTA column and 12-14 = proteins eluted during elution buffer (containing 250 mM imidazole) step of rCeFAR3 from Ni-NTA column

SDS-PAGE lanes 7 – 14 contain smear bands of protein at 20 kDa, which is close to the molecular weight of rCeFAR3.

Protein was dialysed into 50 mM HEPES pH 7.5 for loading a Hitrap Q anion exchange column on the AKTA purifier 100 chromatography system (figure 4.12a). Ion exchange was tried instead of gel filtration since gel filtration on previous FAR proteins, rOvFAR1 and rCeFAR2, indicated protein degradation was occurring while on the column. The anion exchange column was washed with 2 M sodium chloride before the protein sample was loaded to remove any previously bound proteins and then given a final wash with 3 column volumes of 50 mM HEPES pH 7.5. Protein was detected at 280 nm. 5 ml of protein sample was loaded onto the column at 2 ml/min. Unbound protein was collected in the flow through. Bound protein was eluted using a salt gradient (50 mM HEPES pH 7.5, 2 M NaCl). SDS-PAGE (figure 4.12b) was run on all fractions that eluted from the column.

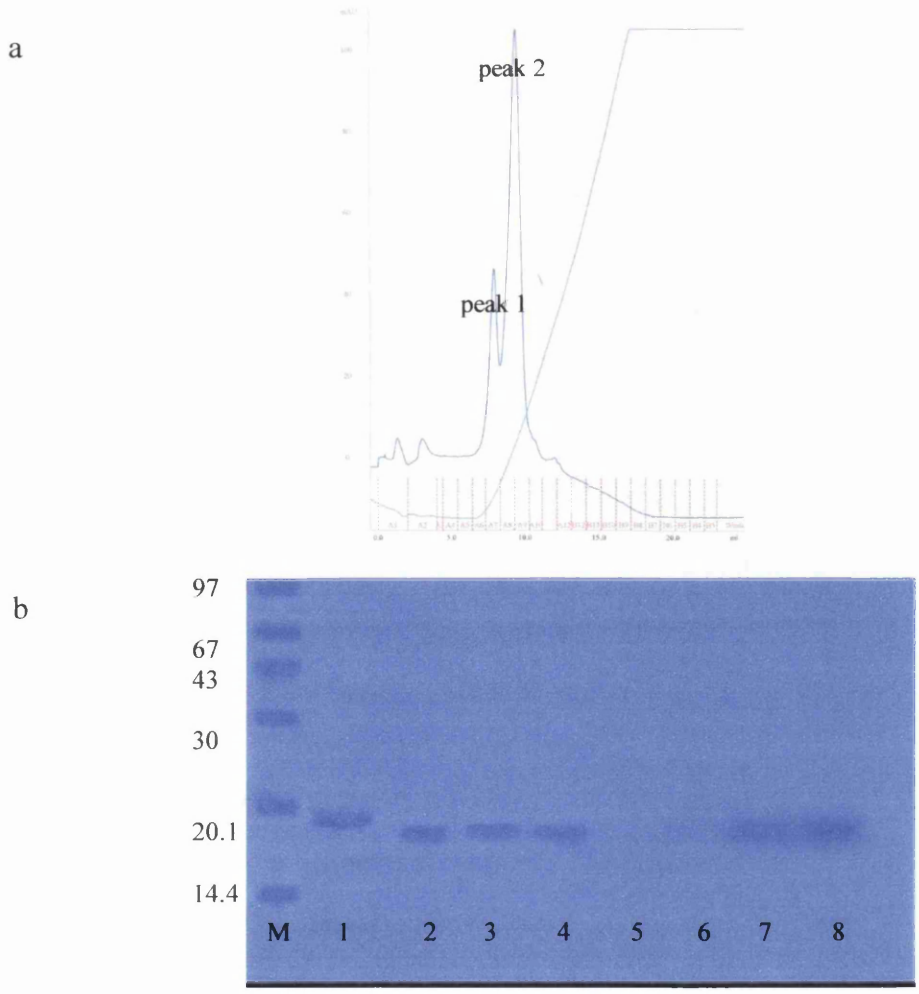

Figure 4.12 a) chart of HQ anion exchange run on AKTA system,
b) M= markers, 1= flow through, 2-4 = proteins eluted in peak one from HQ column in 0-5 % NaCl and 7-8 = proteins eluted in peak two of HQ column in 5-15 % NaCl

The resulting gel showed bands of 20 kDa in lanes 2-4 and 7-8 and lane 1, which is close to the molecular weight of rCeFAR3. Protein in lane 1 did not bind fluorescent tagged fatty acid DAUDA, therefore was not rCeFAR3. The two peaks that resulted from ion exchange, both contain rCeFAR3. These two peaks may well be isoforms of the protein since an isoelectric focusing gel (figure 3.8, lane 5) shows at least 2 bands (charged isoforms) present. MALDI-TOF analysis indicates other higher molecular weight oligomers may also be present (chapter 6), although dynamic light scattering of rCeFAR3 in HEPES and PIPES buffer suggest only moderate polydispersity of rCeFAR3 (chapter 6).

4.5 Results of ligand binding studies of rOvFAR1, rCeFAR2 and rCeFAR3

4.5.1 DAUDA binding

Environment sensitive probes, like DAUDA can explore even subtle difference in the nature of a fatty acid binding site (Kennedy *et al*, 1995b). These binding studies were undertaken to determine the strength of ligand binding between the different FAR proteins. Upon interaction with DAUDA the FAR proteins undergo dramatic blue shifts in the wavelength of maximum fluorescence emission from 542 nm for DAUDA in buffer to 492 nm, 489 nm and 486 nm for DAUDA interacting with rOvFAR1, rCeFAR2 and rCeFAR3 respectively (figures 4.13, 4.14 and 4.15 respectively). This indicates that the FAR proteins have bound the fluorescent fatty acid. Greater shift enhancement of DAUDA fluorescence was observed in rCeFAR3 than the other FAR proteins, as a greater blue shift in fluorescence occurred.

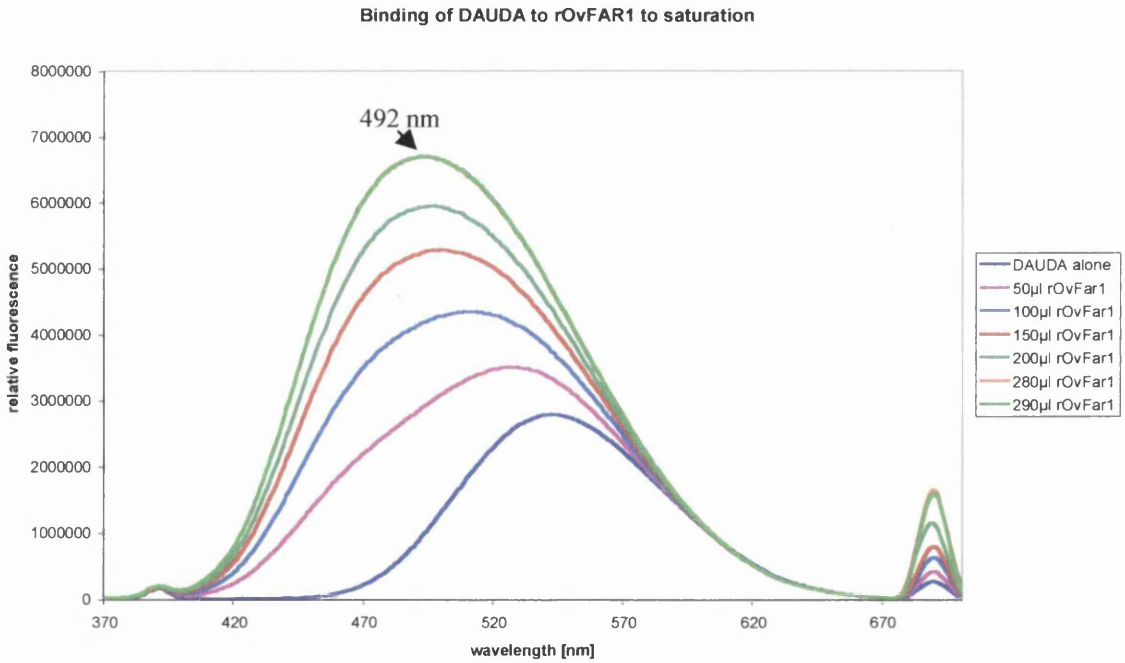

Figure 4.13 Binding of the fluorescent labelled fatty acid DAUDA to rOvFAR1

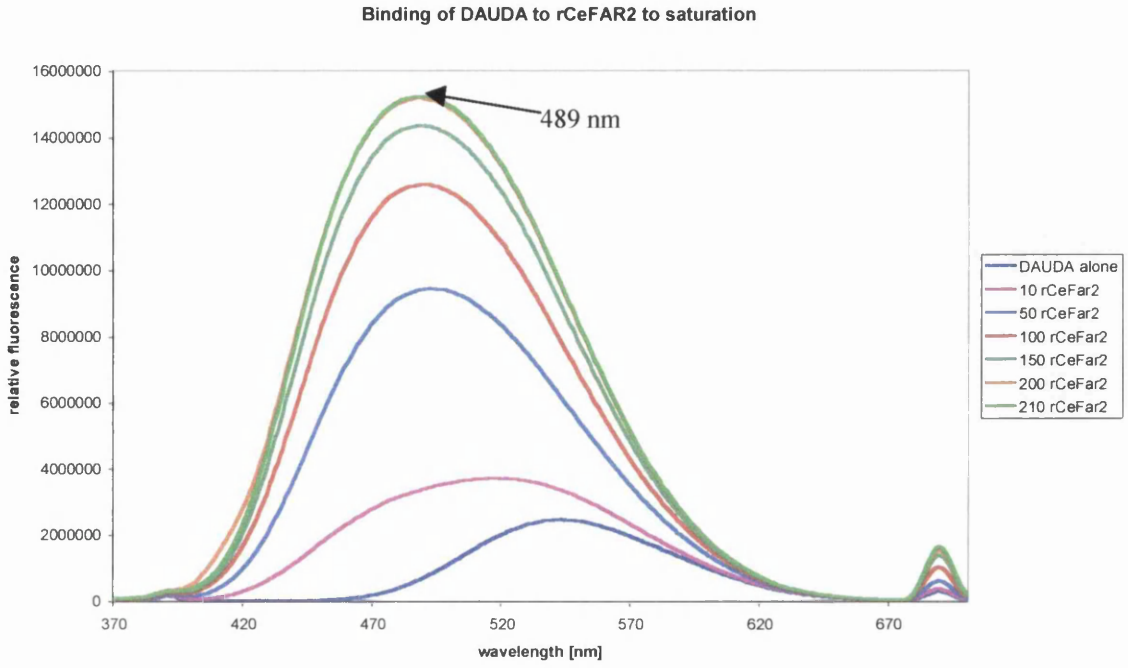

Figure 4.14 Binding of the fluorescent labelled fatty acid DAUDA to rCeFAR2

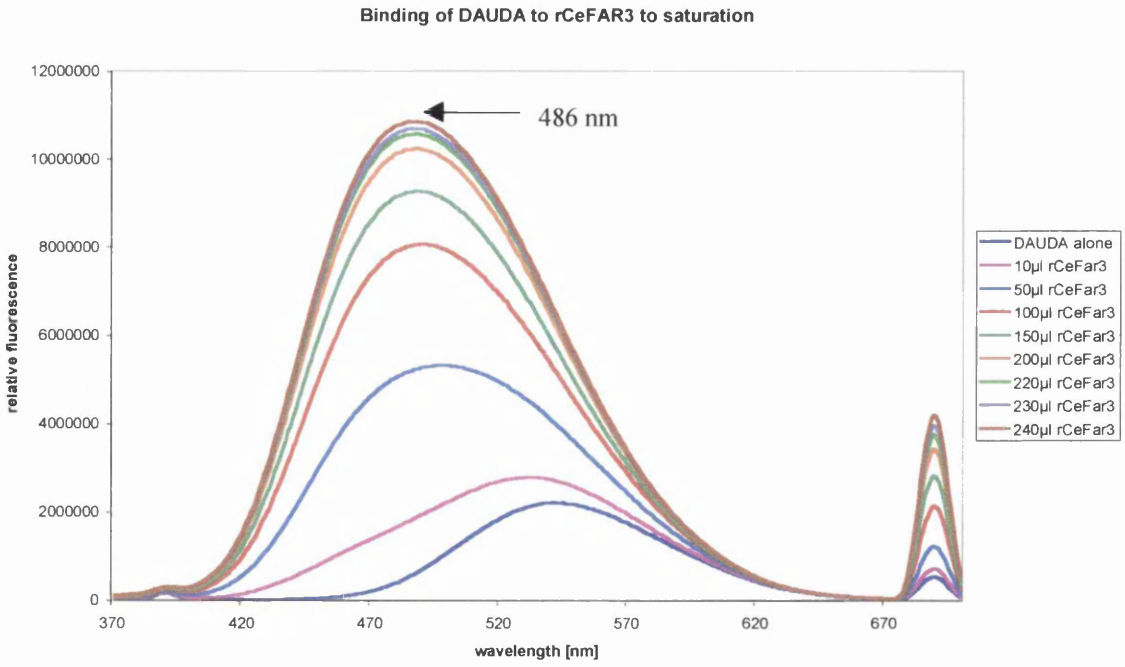

Figure 4.15 Binding of the fluorescent labelled fatty acid DAUDA to rCeFAR3

Experiments involving titrations of the FAR proteins with DAUDA (figures 4.16, 4.17 and 4.18) showed that the proteins have similar dissociation constants (Kd) (Table 4.1) and emphasises the greater fluorescence emission produced in DAUDA binding to rCeFAR3. This is shown by rCeFAR3 having the highest Kd value of 0.09(±3)mM.

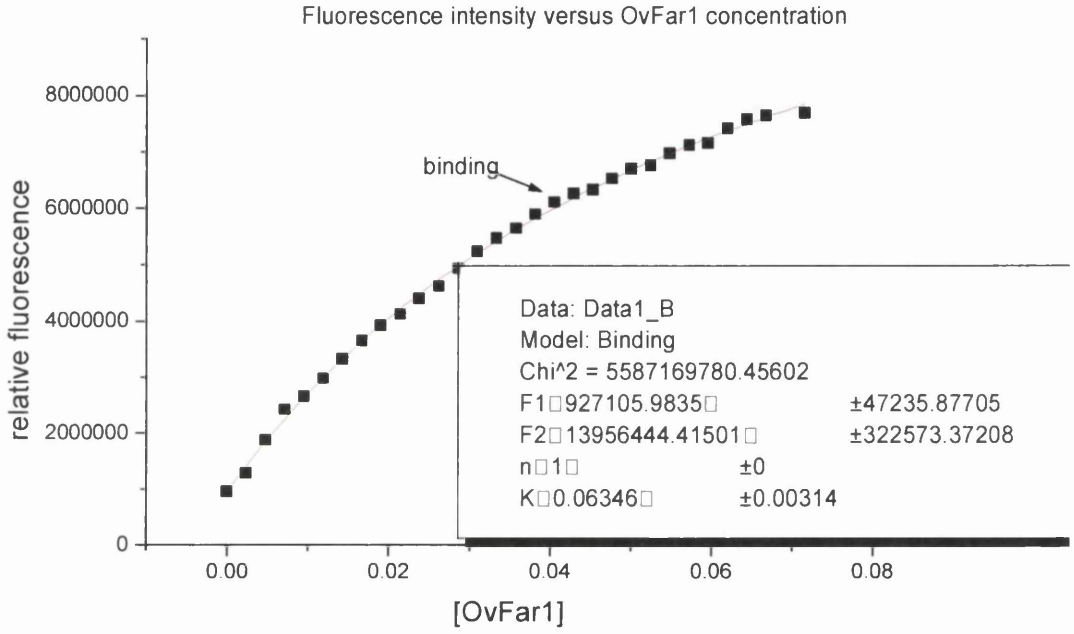

Figure 4.16 Titration curve for the binding of DAUDA to rOvFAR1. Relative intensity changes (recorded at the wavelength of maximum fluorescence emission of 492 nm for rOvFAR1) of 0.000406 mg/ml DAUDA in PBS buffer, upon addition of increasing concentration of rOvFAR1. The data is corrected for the effects of dilution. The solid line is the theoretical binding curve for complex formation with dissociation constant, Kd = 0.063(±3)mM.

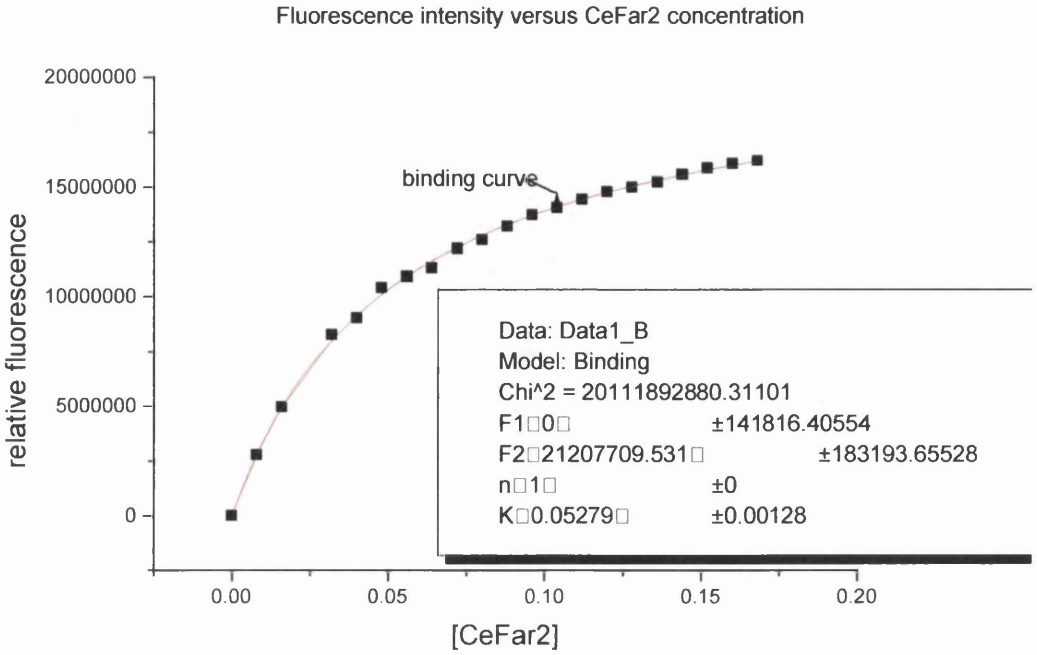

Figure 4.17 Titration curve for the binding of DAUDA to rCeFAR2. Relative intensity changes (recorded at the wavelength of maximum fluorescence emission of 489 nm for rCeFAR2) of 0.000406 mg/ml DAUDA in PBS buffer, upon addition of increasing concentration of rCeFAR2. The data is corrected for the effects of dilution. The solid line is the theoretical binding curve for complex formation with dissociation constant, Kd = 0.053(±1)mM.

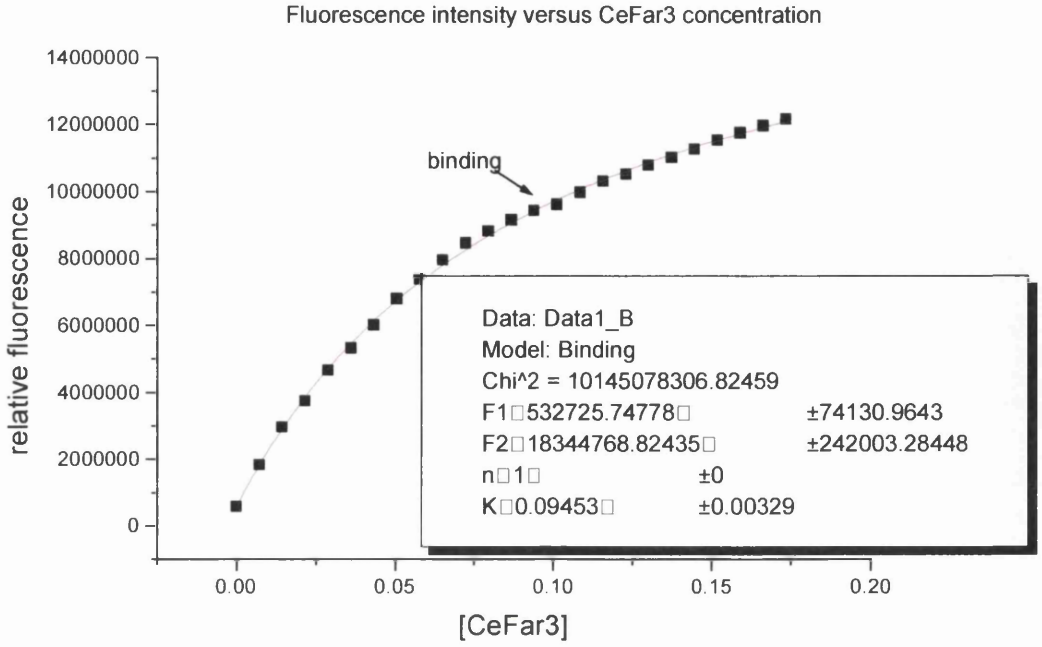

Figure 4.18 Titration curve for the binding of DAUDA to rCeFAR3. Relative intensity changes (recorded at the wavelength of maximum fluorescence emission of 486 nm for rCeFAR3) of 0.000406 mg/ml DAUDA in PBS buffer, upon addition of increasing concentration of rCeFAR3. The data is corrected for the effects of dilution. The solid line is the theoretical binding curve for complex formation with dissociation constant, Kd = 0.095(±3)mM.

4.5.2 Retinol binding

Since retinol is the major ligand sequestered by OvFAR1 and perhaps CeFAR2 and CeFAR3, the interaction of retinol with FAR proteins was monitored using fluorescence. Retinol (all-trans-retinol), figure 2.5, contains 5 conjugated double bonds and is therefore prone to oxidation by light and oxygen. It is highly unstable in aqueous solution, although the ligand appears stable to degradation upon binding to lipid binding proteins (Hemley *et al*, 1979). The fluorescence emission of all-trans-retinol is minimal in buffer alone, but it is dramatically enhanced on addition of retinol-binding proteins (Hemley *et al*, 1979; Narayan & Berliner, 1997; Papiz *et al*, 1986). To minimise retinol deterioration, the

dissociation constant of retinol:protein binding was estimated by fluorescence titration in which increasing concentrations of retinol were added to a protein solution in a cuvette. This means that any retinol not complexed with protein is prone to deterioration. Dissociation curves (figures 4.19) were corrected for free retinol. rCeFAR2 protein produces a slightly greater increase in the fluorescence emission intensity of bound fluorescent ligand than the other FAR proteins, but not enough to be significant. Unfortunately, the Kd values for retinol bound to these FAR proteins are unreliable as the errors for each are greater than the Kd value obtained (table 4.1). The blue shift in wavelength from 542 nm to 477 nm, 476 nm and 477 nm for rOvFAR1, rCeFAR2 and rCeFAR3 respectively indicates that the FAR proteins have a similar affinity for retinol.

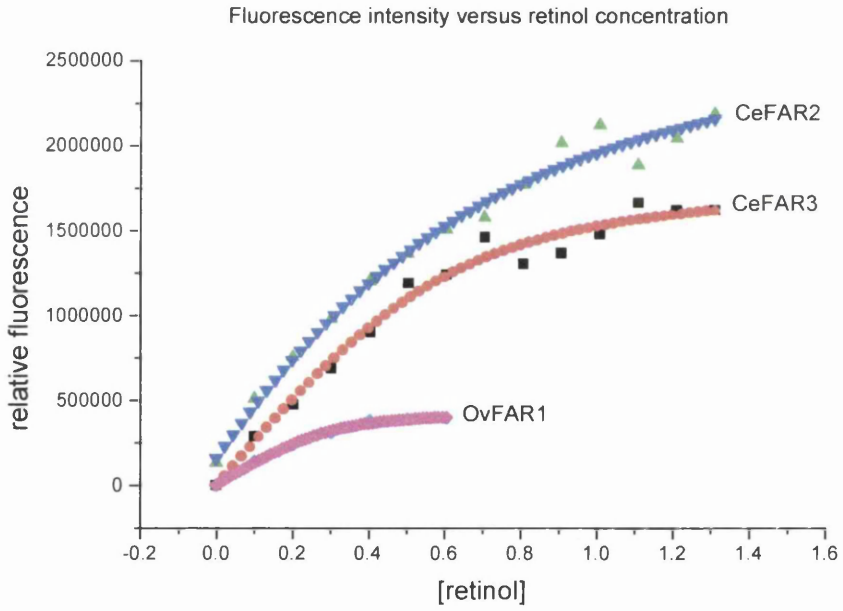

Figure 4.19 Fluorescence titration curves showing the binding of all-trans-retinol to rOvFAR1, rCeFAR2 and rCeFAR3. Consecutive additions of retinol were added to a cuvette containing 200µl of rOvFAR1, rCeFAR2 and rCeFAR3 and changes in the relative fluorescence intensity at 477 nm, 476 nm and 477 nm respectively. The data were corrected for the effects of dilution and retinol fluorescence for the curves shown. Theoretical binding curves for complex formation are represented by the solid lines with dissociation constants, Kd = 0.3 mM for rOvFAR1, Kd = 0.34 mM for rCeFAR2 and Kd = 0.09 mM for rCeFAR3 (not enough to give a reliable standard deviation for rOvFAR1).

	rOvFAR1	rCeFAR2	rCeFAR3
ligand binding	λ_{max} Kd	λ_{max} Kd	λ_{max} Kd
DAUDA	492nm 0.06mM	489nm 0.05mM	486nm 0.09mM
Retinol	477nm 0.3mM	476nm 0.34mM	477nm 0.09mM

λ_{max} : Wavelength of maximal fluorescence emission of ligand when bound to protein.

Kd : All dissociation constant estimations obtained by titration with fluorescent ligands. Retinol Kds have large errors and the data has to be judged accordingly.

Table 4.1 - Functional properties of FAR proteins

4.5.3 Competitive binding studies with oleic acid

Competition for the retinol binding site in the FAR proteins was investigated by adding oleic acid to the retinol-FAR protein complexes and observing a decrease in the fluorescence emission (figures 4.20, 4.21 and 4.22).

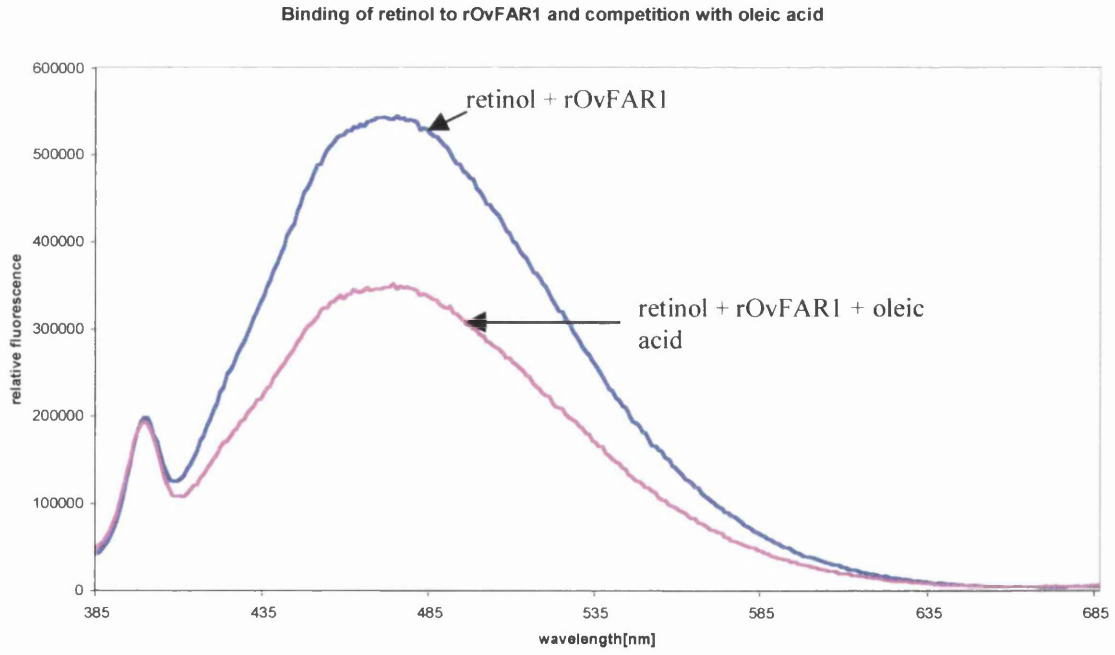

Figure 4.20 Competitive effects of oleic acid on rOvFAR1-retinol complex. The enhanced emission of retinol when bound to rOvFAR1 was reversed by addition of oleic acid.

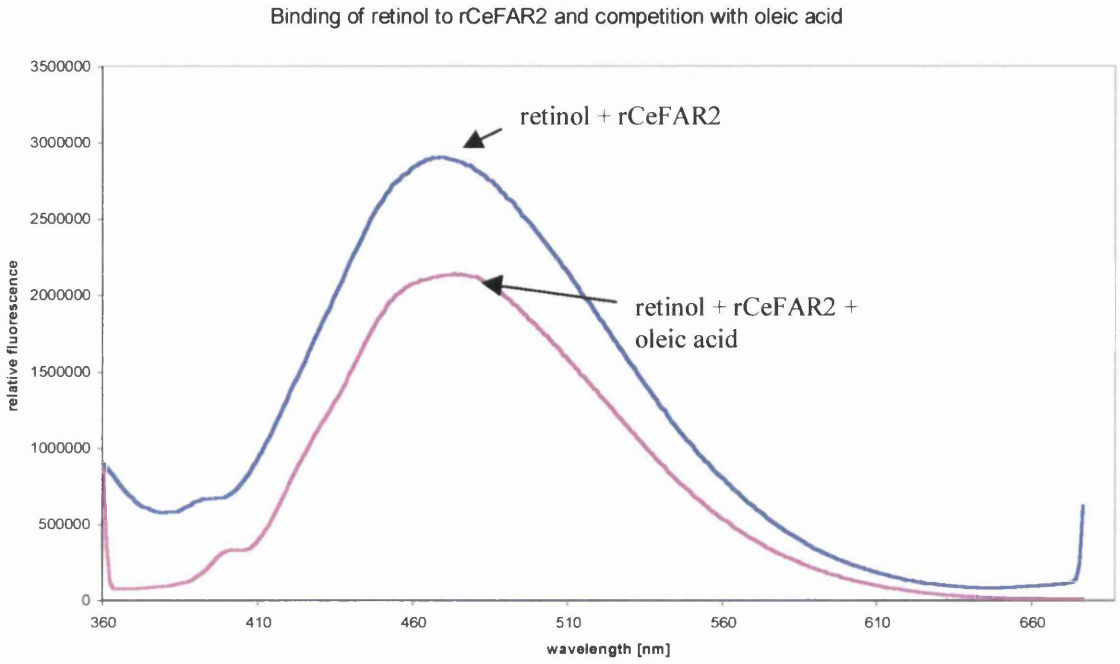

Figure 4.21 Competitive effects of oleic acid on rCeFAR2-retinol complex. The enhanced emission of retinol when bound to rCeFAR2 was reversed by addition of oleic acid.

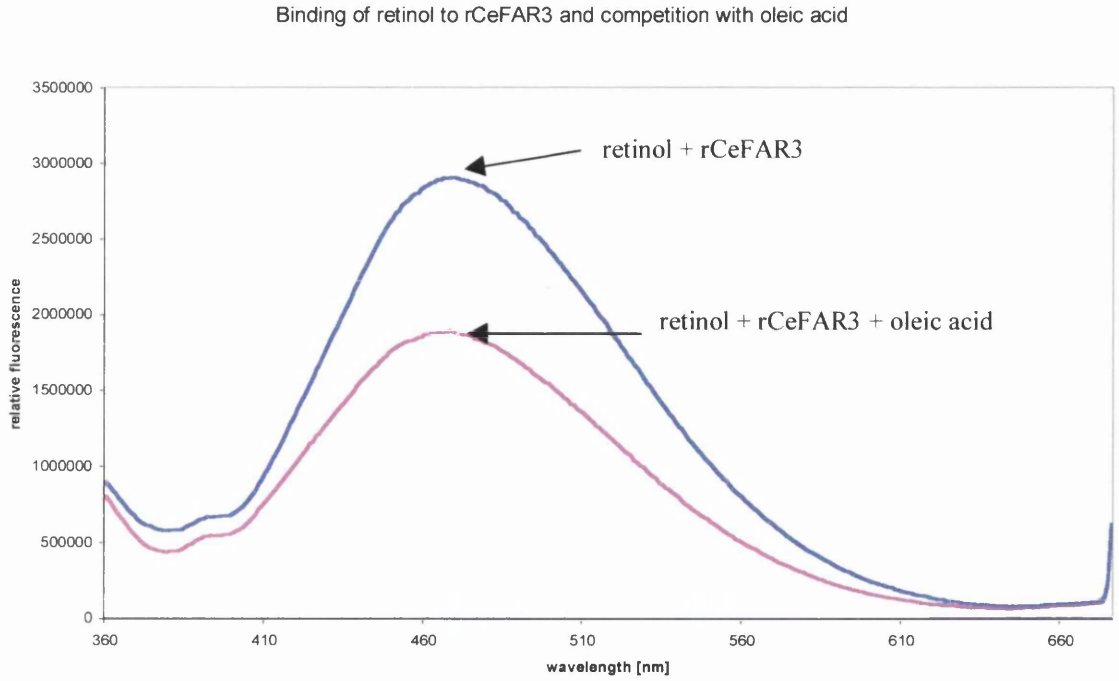

Figure 4.22 Competitive effects of oleic acid on rCeFAR3-retinol complex. The enhanced emission of retinol when bound to rCeFAR3 was reversed by addition of oleic acid.

These competitive binding experiments show that the enhanced emission of retinol when bound to the FAR proteins was reversed by addition of oleic acid. This suggests that the retinol and oleic fatty acid do compete for the binding site, indicating that the binding sites are either the same (the more probable), or interactive because the oleic acid affected the retinol bound to the FAR proteins.

4.5.4 Preference of FAR proteins for fatty acid chain length

Experimental work to determine preference of fatty acid chain length has previously been carried out on rGpFAR1 (Prior *et al*, 2001) where the results indicated no preference for saturated or unsaturated fatty acids, but showed a preference for chain length C14-C18, with the greatest displacement of DAUDA occurring at C15. In this section we investigate the preference of rOvFAR1, rCeFAR2 and rCeFAR3 proteins for fatty acid chain length (figures 4.23, 4.24 and 4.25). The same molarity of test fatty acid (chapter 2, figure 2.6 - lists fatty acids used) was added to a DAUDA-FAR protein complex. The percentage decrease in DAUDA fluorescence was recorded at the maximum fluorescence emission of 492 nm, 489 nm and 486 nm for rOvFAR1, rCeFAR2 and rCeFAR3 respectively.

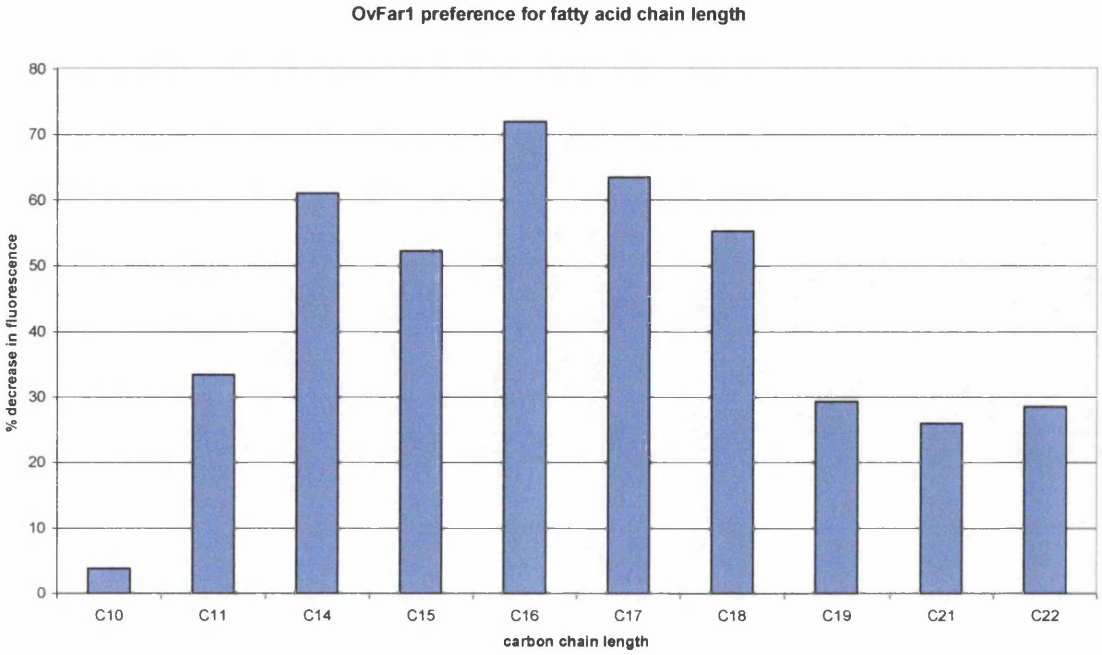

Figure 4.23 Chart showing the preference of rOvFAR1 for fatty acid chain length

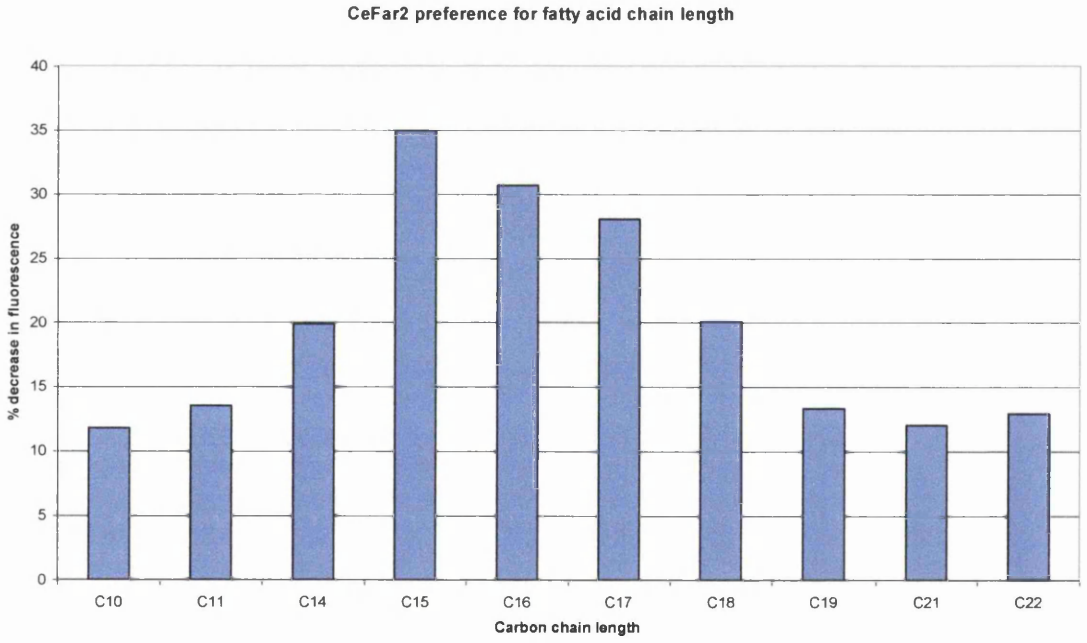

Figure 4.24 Chart showing the preference of rCeFAR2 for fatty acid chain length

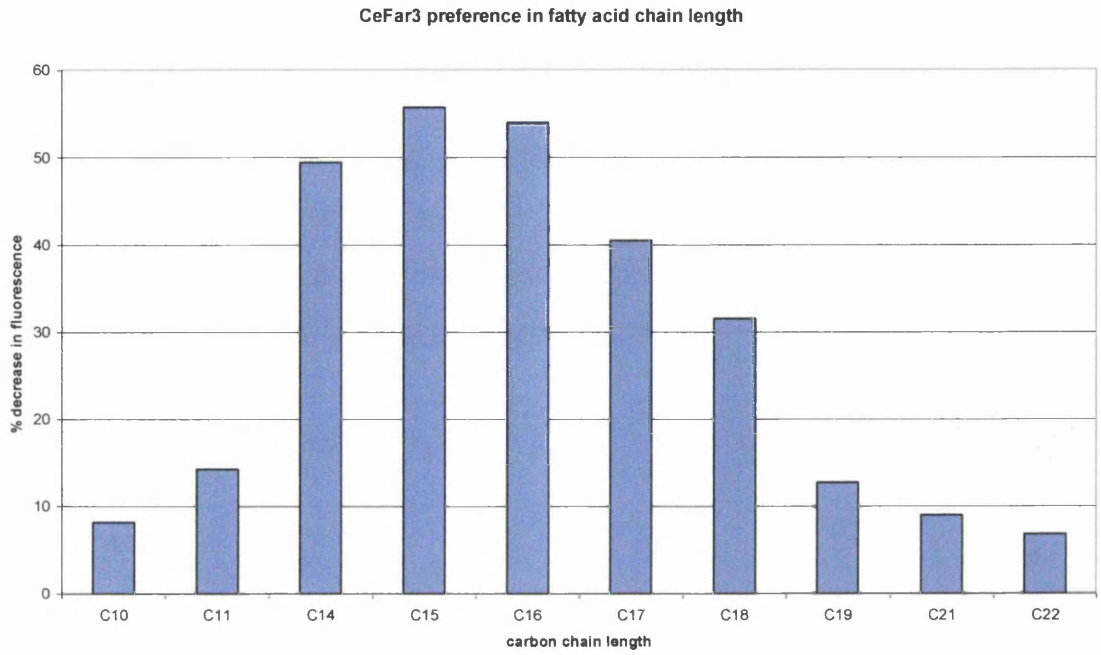

Figure 4.25 Chart showing the preference of rCeFAR3 for fatty acid chain length

Significant binding occurred in the range C14 - C18 for all the FAR proteins investigated, with the maximum DAUDA displacement occurring for fatty acids, pentadecanoic acid and palmitic acid, which have a C15 or C16 chain length.

4.5.5 Lipid binding

Lipids are naturally present in the *E. coli* expression system used (Wisely *et al*, 2002). Because of this, there may be competition between these lipids and the fatty acids for the FAR protein binding site. During normal overexpression, there are no fatty acids present and therefore these lipids may occupy the binding site. It is unlikely that this will be either complete or in a fixed ratio. This is likely to give rise to a heterogeneous sample with some proteins with lipid in their binding site and others with empty binding sites. MALDI-TOF analysis of rGpFAR1 before and after lipid removal columns suggests that the proteins do bind lipids or other molecules during their preparation (chapter 6). The competition for the

binding site in the FAR proteins suggests that the molecules bound during preparation may be removed by competition with other fatty acids. This removal is not a straight forward procedure (N. Meenan, private communication), since the protein requires to be bound to a media by reverse phase chromatography, denatured, washed with buffer, then refolded before the required fatty acid is added. There is also evidence (Wisely *et al*, 2002) that during normal protein folding in over-expression lipid may also get trapped in the protein interior.

4.6 Discussion

The aim of this chapter was to investigate and improve the purity of the overexpressed proteins rOvFAR1, rCeFAR2 and rCeFAR3, to compare their ligand binding properties with previously studied rGpFAR1 and to establish if rCeFAR2 and rCeFAR3 are good functional homologues of rOvFAR1.

The results indicate that the three recombinant FAR proteins investigated in this chapter are indeed functional lipid binding proteins. Fluorescence results showing that the proteins have similar activity despite sequence similarities of only 32 %. The purity of the three FAR proteins as depicted on SDS-PAGE show that in each case two very close bands of approximately 20 kDa at the expected molecular weight are present, indicating that the proteins either degrade, the lower band has been cleaved of a small number of amino acids or as is more likely the presence of loosely bound lipid. MALDI-TOF analysis confirms that the proteins are of a molecular weight close to that expected, also shown is that the proteins may have bound lipids or endotoxins derived from E. coli overexpression (chapter 6).

The FAR proteins investigated undergo dramatic shifts in wavelength of maximum fluorescence emission from 542 nm for DAUDA in buffer to 492 nm, 489 nm and 486 nm for DAUDA interacting with rOvFAR1, rCeFAR2 and rCeFAR3 respectively. This shows that the FAR proteins have bound to the

fluorescent fatty acid. The greatest enhancement was observed in DAUDA binding rCeFAR3 with a blue shift in wavelength intensity from 542 nm for DAUDA in buffer to 486 nm. The dissociation constants (table 4.1) calculated for DAUDA binding the FAR proteins emphasises the greater fluorescence emission produced in DAUDA binding rCeFAR3 with a Kd of 0.095(±3) mM compared to Kd of 0.053(±) mM for rCeFAR2 and Kd 0.063(±3) mM for rOvFAR1. Binding DAUDA with enhanced fluorescent emission and a large blue shift suggests that the binding site for the FAR proteins is in an apolar environment and isolated from solvent. This has already been shown for rOvFAR1 (Kennedy *et al*, 1997) and GpFAR1 (Prior, 2001) but shows that rCeFAR2 and rCeFAR3 of the FAR family interact in a similar way and are therefore in an apolar environment and isolated from solvent. This isolation of ligands from solvent water is useful in that it would protect oxidation-sensitive ligands, such as retinol and other hydrophobic ligands.

The FAR proteins have also been shown to bind retinol and rCeFAR2 was shown to produce a slightly greater increase in the fluorescence emission intensity of the bound fluorescent ligand than the other FAR proteins (figure 4.19). Blue shifts in wavelength of maximum fluorescence emission from 542 nm for retinol in buffer to 477 nm, 476 nm and 477 nm for retinol interacting with rOvFAR1, rCeFAR2 and rCeFAR3 respectively. Competitive effects with the retinol-FAR protein complex were observed when 10 mM oleic acid was added. This indicates that the binding site for retinol and fatty acids is interactive or congruent in the case of all the FAR proteins. This has also been shown for GpFAR1 (Prior *et al*, 2001) another member of the FAR family, which will be discussed in a later chapter. Binding retinol is important for OvFAR1 as it is thought that the nematode *Onchocerca volvulus* that secretes the protein OvFAR1 requires retinol for a variety of metabolic and developmental purposes including, growth, differentiation, embryogensis, glycoprotein synthesis and anti-oxidants. This suggests that the FAR proteins role in the nematodes is related to retinol acquisition or transport.

The preference of the FAR proteins for fatty acid chain length was investigated by competition of the DAUDA-FAR protein complex. The maximal binding occurred in the range C14-C18 for all the FAR proteins, with the greatest displacement occurring between C15-C16. This is also the case for GpFAR1 (Prior *et al*, 2001). This preference for a C15-C16 fatty acid chain length correlates favorably with the most abundant fatty acids found in the *Globodera* organisms which are fatty acid chain C14 – C22 (Holz *et al*, 1998a; Holz *et al*, 1998b), with the main fatty acids found in potato roots being of chain length C16 and C18 (Holz *et al*, 1998a; Holz *et al*, 1998b). This would indicate that the recombinant proteins are behaving as they do *in vivo*, by binding similar fatty acid chain lengths.

The results from this chapter show that the FAR proteins investigated here have similar ligand binding activities, indicating that the binding sites for these proteins are most probably conserved. The sequence alignment of FAR proteins (figure 1.9), OvFAR1, CeFAR2, CeFAR3 and GpFAR1 shows areas of sequence similarity between amino acid residues 41-46, 129-131 and a large region of similarity between amino acid residues 78-90. These conserved regions between the FAR proteins indicate probable binding sites.

Chapter 5 – GpFAR1 overexpression and purification.

5.1 Introduction

GpFAR1 protein is secreted by the potato cyst nematode *Globodera pallida*, one of the major pathogens of the commercial potato *Solanum tuberosum*. GpFAR1 is an 18.8 kDa protein that has been characterised by ligand binding studies as a fatty acid and retinol binding protein. This chapter studies GpFAR1 separately from the other FAR proteins as it has been already been characterised by ligand binding (Prior *et al*, 2001) and was the FAR protein that grew protein crystals. Recombinant protein is used to study the functional and structural properties as protein from the nematode is not available in sufficient quantity.

5.2 Overexpression and purification of rGpFAR1

Dr A. Prior (Scottish Crop Research Institute, Dundee) supplied a plasmid encoding the antigen of GpFAR1 as a pGEX-2T plasmid containing the GpFAR1 gene.

The plasmid was extracted and transformed into BL21 host cells. The plasmids were inserted by heat shock. 100 µl of host strain BL21 was placed in a sterile Eppendorf tube on ice. 1 µl β-mercaptoethanol was added and left on ice for 2 minutes, then 5 µl of plasmid was added to the tube and left on ice for 1 hour. A water bath was heated to 42 °C and the cells placed in it for 30 seconds and then replaced on ice to cool back down to 4 °C. 400 µl of LB (Millipore) was then added and the tube was incubated at 37 °C for 1 hour. The cells were then plated on agar with ampicillin and incubated overnight at 37 °C. After incubation 10 ml of LB 200 µg/ml ampicillin was inoculated with a single colony from the plate and incubated overnight at 37 °C with rotary shaking at 200 rpm. Glycerol stocks

were made using 850 µl of culture and 150 µl of 80 % glycerol and stored at –80 °C.

A 1 L flask containing 500 ml of 2-YT broth and 200 µg/ml ampicillin was inoculated with a glycerol stock and grown overnight at 37 °C with shaking at 200 rpm. This overnight broth was added to 2 more litre flasks containing 1 litre 2-YT broth 200 µg/ml ampicillin each and grown at 37 °C and 200 rpm for approximately 3 hours until cultures had grown to an A_{600nm} between 0.6–1. 1 mM dioxane free Isopropyl β-D-thiogalactoside (IPTG) was added to induce gene expression. Cells were grown for a further 3 hours at 37 °C with shaking at 200 rpm. The culture was transferred to appropriate centrifuge bottles and centrifuged at 8000 rpm for 15 minutes to sediment the cells and the supernatant was removed.

Each pellet was then resuspended in 100 ml PBS with a Roche EDTA free protease inhibitor tablet. The cells were disrupted by sonication on ice, using a Status US200 sonicator from Philip Harris Scientific for six, 30 second blasts at half power. Cell disruption is evidenced by partial clearing of the suspension. A small aliquot of sonicate was saved for analysis by SDS-PAGE. Sonicate was then centrifuged at 10000 rpm for 10 minutes and the supernatant transferred to a fresh container. 1 ml of Glutathione Sepharose 4B was then added to 50 ml of supernatant and this was incubated at room temperature with gentle rocking for an hour. After an hour this was spun down gently in the centrifuge at 500 rpm for 10 minutes and the resin was washed 3 times with PBS. Finally the proteins were released from the GST by cleavage with thrombin protease, 50 units of thrombin per 1 ml of resin. After 16 hours at room temperature the resin was spun down at 200 rpm for 5 minutes and the supernatant saved for SDS-PAGE analysis. The protein was detected in the supernatant on SDS-PAGE. To remove the thrombin protease the protein was run through a Benzamidine FFTrap column from Amersham Pharmacia Biotech at 0.5 ml/min using AKTA purifier 100 (Amersham Pharmacia Biotech).

SDS-PAGE (figure 5.1) shows the thrombin cleavage steps of rGpFAR1 preparation.

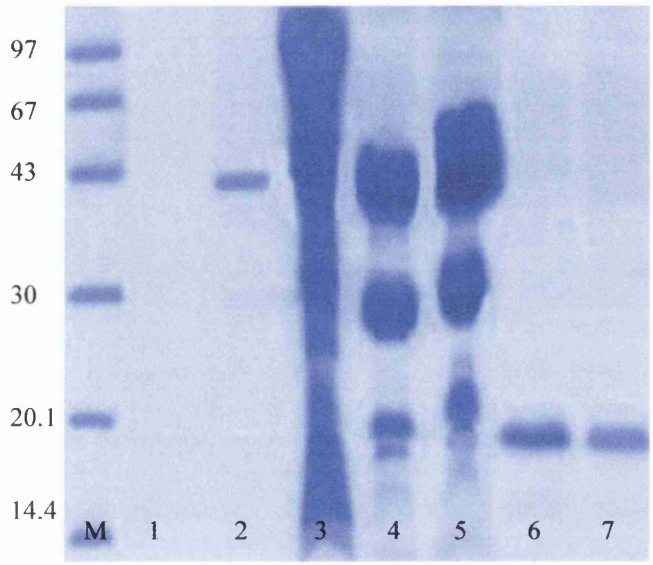

Figure 5.1 M = marker, 2 = resin after rGpFAR1 removal, 3 = pellet before resin added, 4-5 = washes of resin and 6-7 = rGpFAR1 removed from resin.

A one step purification is affected by loading the supernatant (figure 5.1, lanes 6 and 7) onto a pre-equilibrated Superose 12 gel filtration column (figure 5.2a) at 0.3 ml/min and eluted in 50 mM HEPES pH 7.5. SDS-PAGE (figure 5.2b) was run on all fractions that were eluted.

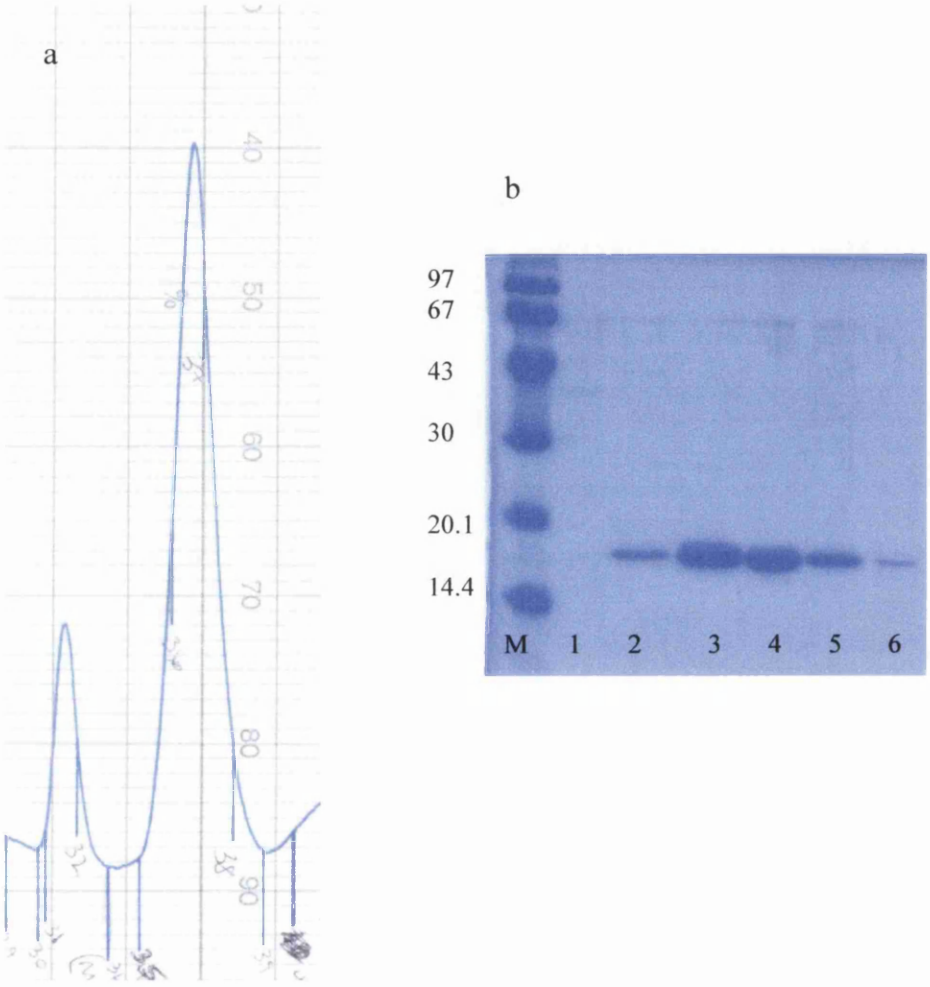

Figure 5.2 a) gel filtration of rGpFAR1 showing 2 peaks.
b) M = markers, 1 = first peak from gel filtration containing high molecular weight protein of approximately 42 kDa, 2-6 = second peak from gel filtration containing rGpFAR1.

rGpFAR1 was judged to be homogenous by the single band in lanes 3-6 with an apparent molecular weight of 18 kDa, which approximates to the predicted molecular weight of rGpFAR1 (18.8 kDa).

These rGpFAR1 fractions were assayed for DAUDA binding and were shown to exhibit fatty acid binding activity confirming functional rGpFAR1 to be present. 9 mg/ml of rGpFAR1 was produced from an 8 L culture.

5.3 Discussion

The overexpression and purification of rGpFAR1 resulted in a single band on SDS-PAGE analysis. From isoelectric focusing gel (chapter 3, figure 3.8, lane 2), rGpFAR1 may be judged to be homogenous as only a single band is shown. MALDI-TOF analysis (chapter 6) shows no dimeric species present, though it is likely that lipid is bound to rGpFAR1. Dynamic light scattering suggests that the protein may be polydispersed. Since SDS-PAGE and IEF gel analysis suggest a monomeric protein crystallisation trials were attempted (chapter 7).

Chapter 6 – Matrix Assisted Laser Desorption Ionisation – Time Of Flight, Circular Dichroism and Dynamic Light Scattering analysis of the FAR proteins.

6.1 Introduction

This chapter studies the biophysical properties of all the FAR proteins to get a better understanding of their secondary structure and to determine the optimum conditions for crystallisation studies.

6.2 Matrix Assisted Laser Desorption Ionisation – Time Of Flight (MALDI-TOF) analysis of FAR proteins

MALDI (Hillenkamp *et al*, 1991) has been used successfully in determining molecular weights of proteins. In this chapter it is used not only to determine the molecular weight of the FAR proteins and to determine the presence of oligomers, but also in the case of rGpFAR1 it has been used to determine if lipid is present. Lipids and endotoxins from *E. coli* are thought to have bound to these fatty acid and retinol binding proteins during their expression. MALDI-TOF analysis was carried out as described in chapter 2.2. All spectra were baseline corrected (BC) and mass corrected (MC) using the standards described in chapter 2.2.

6.2.1 rOvFAR-1

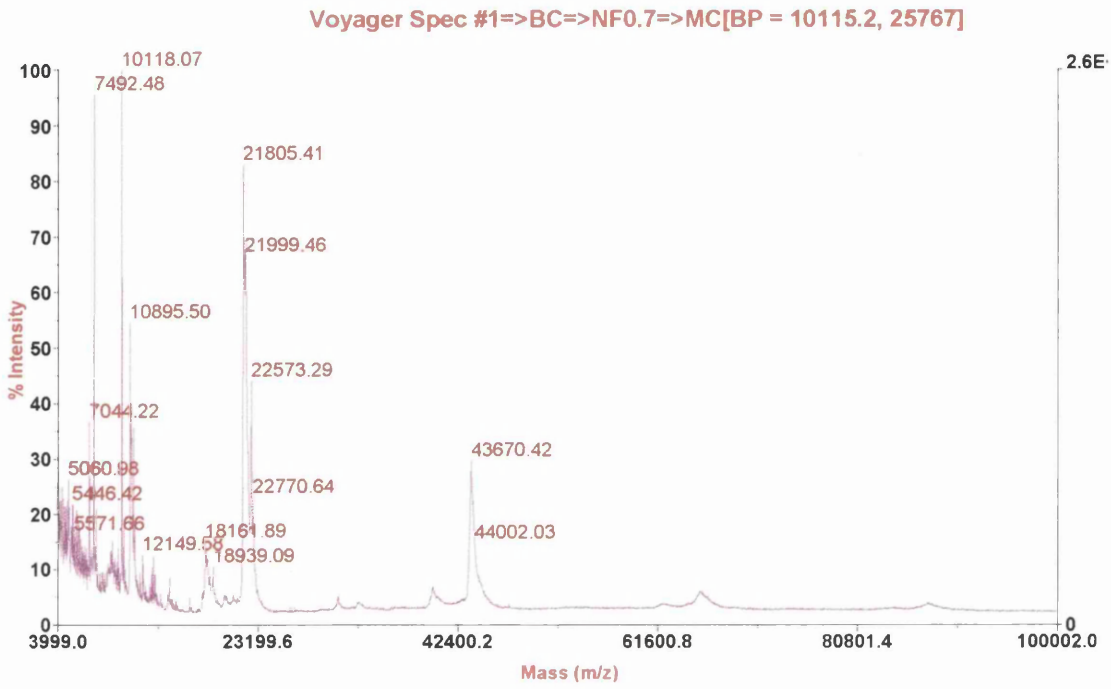

Figure 6.1 MALDI-TOF analysis of rOvFAR1

The expected molecular weight of rOvFAR1 from the amino acid sequence is 20580 Da. The actual mass from MALDI-TOF (figure 6.1) of the parent peak was 21805 Da, giving a difference between the calculated and observed molecular masses of approximately 1225 Da. The discrepancy between calculated and observed values is likely to be partially due to the sequestering of lipid into the binding site during overexpression, partially due to lipid non-covalently bound to the protein surface, and perhaps *E. coli* lipid(s) trapped within the protein interior during protein folding. This latter effect has already been evidenced for Hepatocyte nuclear factor 4 (HNF4), (Wisely *et al*, 2002). Since over-expression is in bacteria (*E. coli*), post-translational modification is unlikely in this instance. From experimental binding studies, described in chapter 4, it was found that there is a preference for fatty acids with a chain length between 14 – 18 carbon atoms. This would give an additional approximate mass value between 226 – 290 Da.

The difference between the calculated and observed molecular weights of approximately 1225 Da makes it likely that 4 fatty acids are associated with each rOvFAR1 molecule. These fatty acids may be non-covalently bound to the surface of rOvFAR1 and fly with it during MALDI-TOF. There may also be at least one fatty acid in the binding pocket. The spectrum (figure 6.1) also shows doubly, 10895 Da and (approximately) triply, 7492 Da charged peaks are present.

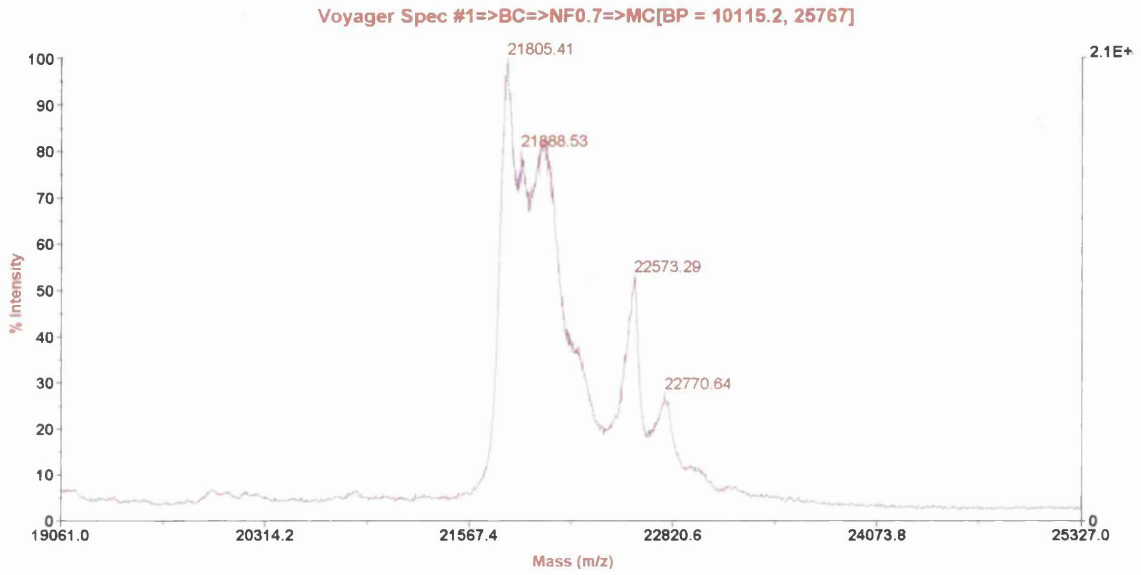

Figure 6.2 MALDI-TOF of rOvFAR1 expanded main peak

It is worth noting that the main peak at 21805 Da (figure 6.2) is really a mixture of several sub peaks of near identical mass. The reason for this could be the presence of loosely bound lipid, which may indicate heterogeneity. The peak at 43670 Da (figure 6.1) strongly suggests that rOvFAR1 may exist as a dimer in its native form. The dimer peak is twice the singly charge parent peak (21805 Da). Further low abundance peaks at approximately 66,000 and 88,000 show higher oligomeric states may have existed. SDS-PAGE analysis of rOvFAR1 (figure 4.4) shows a band of 20 kDa, which is close to the molecular weight of rOvFAR1 and a higher molecular weight band present at approximately 43 kDa, which again may support the native dimer conformation.

6.2.2 rCeFAR2

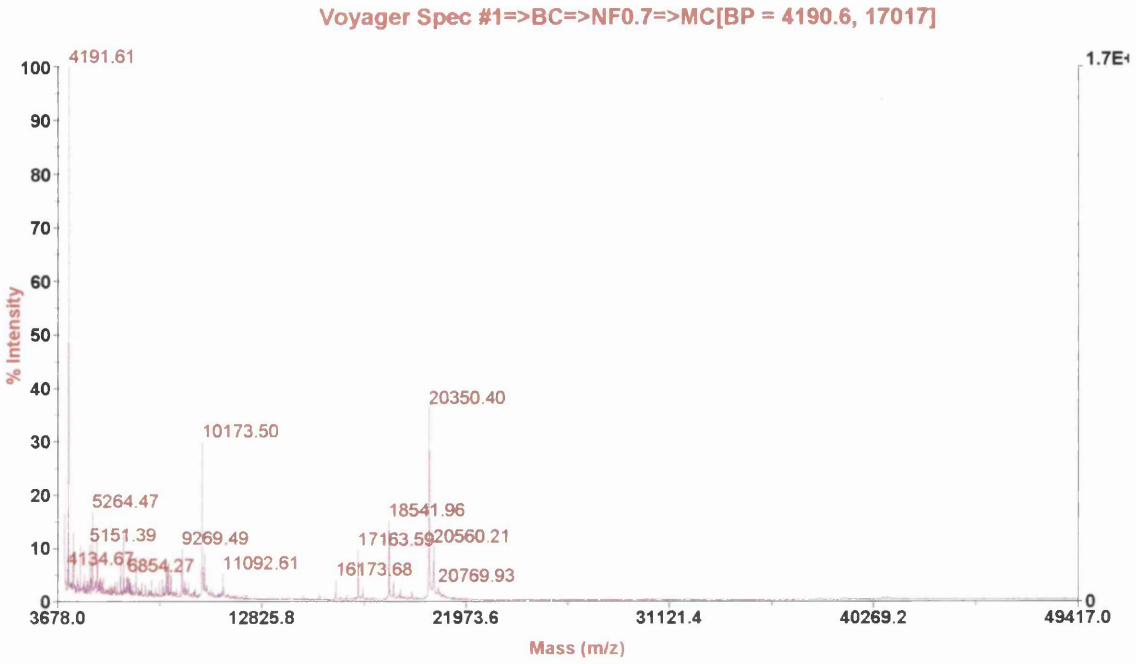

Figure 6.3 MALDI-TOF analysis of rCeFAR2

The expected molecular mass calculated from the rCeFAR2 sequence is 20035 Da, the molecular weight obtained from MALDI-TOF analysis (figure 6.3) was 20350 Da, a difference of 315 Da. The discrepancy is likely to be due to the partial sequestering of lipid, possibly fatty acid of chain length C20. The doubly charged peak of 10173 Da can be clearly observed and is half the parent peak. In this instance the spectrum shows no evidence to support the formation of a dimer, but a peak at 20560 Da adjacent to the 20350 Da peak was observed. This may indicate protein heterogeneity, with two species present in the sample. The most likely explanation is that one species exists with two lipid molecules bound to rCeFAR2 (20560 Da) and one with only one lipid molecule bound rCeFAR2 (20350 Da). The difference between the two species of 210 Da may well be the result of an extra fatty acid of C12 chain length, molecular mass 212 Da, being incorporated during folding. SDS-PAGE (figure 4.8b) shows two bands of approximately 20 kDa close together and dynamic light scattering results indicate

rCeFAR2 is not homogenous (see chapter 6.5) agreeing with the MALDI-TOF spectrum.

6.2.3 rCeFAR3

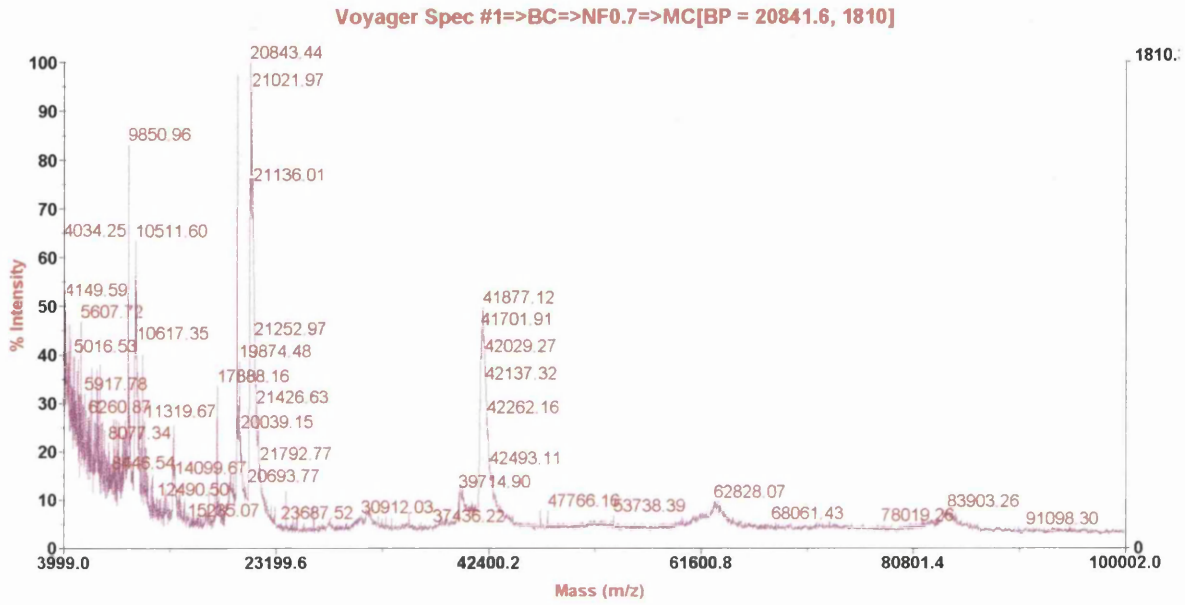

Figure 6.4 MALDI-TOF analysis of rCeFAR3

The rCeFAR3 spectrum is poor, indicating rCeFAR3 did not fly well. Additionally, several electronic noise peaks (spikes) can be seen that are due to traces of salt present in the solution. rCeFAR3 was purified (chapter 4.4) by ion exchange and eluted using a salt gradient. rCefar3 was dialysed to remove salt before MALDI-TOF analysis, possibly traces still remained. The expected molecular mass of rCeFAR3 from amino acid sequence is 20905 Da, the molecular weight obtained from MALDI-TOF analysis (figure 6.4) was 20843 Da, a difference of 63 Da. This discrepancy between calculated and observed values is too small to assign a fatty acid, but may be due to inaccuracy due to calibration errors in the MALDI-TOF experiment as well as the quality of the spectrum. It should also be noted that sub peaks are present around the main parent peak, which indicates heterogeneity of the protein sample due to the acquisition of *E.*

coli lipids. The spectrum also showed that the native structure consists of a dimer and several other oligomers (peaks at 41877, 62828 and 83904 Da). SDS-PAGE (figure 4.12) shows bands of 20 kDa, which is close to the molecular weight of rCeFAR3 (20.9 kDa), but also other higher molecular weight oligomers present, as demonstrated in the MALDI-TOF experiment.

6.2.4 rGpFAR1 and lipid analysis

FAR proteins are fatty acid and retinol binding proteins as observed in the fluorescence studies carried out in chapter 4. An abundance of lipids are present in *E. coli*, including palmitic acid (fatty acid chain length 16), myristic acid (fatty acid chain length 14) and cyclo heptadecanoic acid (fatty acid chain length 17 with a cyclopropyl group) (Dhe-Paganon, 2002). Since *E. coli* is used for overexpressing the FAR proteins, it is more than likely that during overexpression these *E. coli* fatty acids will be sequestered by the FAR proteins. To determine if the FAR proteins had sequestered *E. coli* lipids during their overexpression MALDI-TOF analysis of rGpFAR1 was carried out before and after lipid/endotoxin removal (figures 6.5 and 6.6). This is important to know since lipid uptake may produce heterogeneous protein consisting of rGpFAR1 with and without lipids, which will in turn affect the quality (homogeneity) of protein available for crystallisation. Not only would there be a mix of bound/unbound lipid proteins, but the size (chain length) of the lipids themselves could be varied in the binding site. There is also another important point in trying to determine which *E. coli* fatty acid(s) have been sequestered. It appears that there may be a strong affinity for these *E. coli* fatty acids and they may have a preference over the fatty acids normally sequestered by the FAR proteins in their natural environment. If this were the case, then a ready made inhibitor would be available to combat the parasite. Endotoxins and lipids were removed by passing purified rGpFAR1 down a gravity flow disposable column containing either polymyxn B resin (Sigma) column or a D-detergent removal column (Pierce).

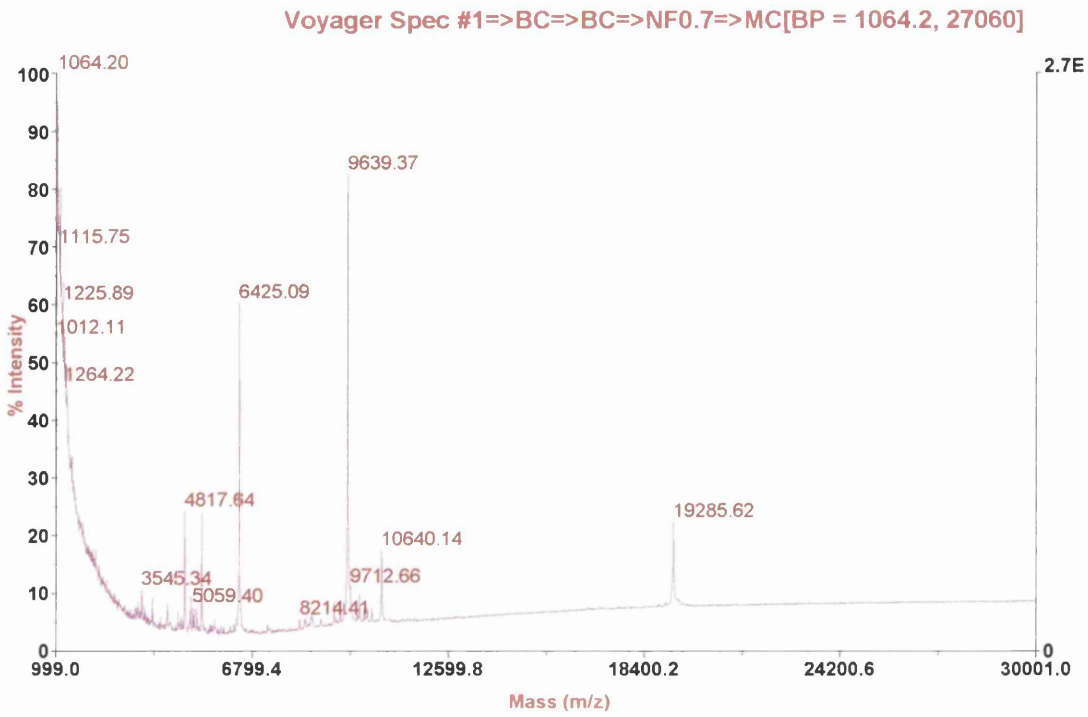

Figure 6.5 MALDI-TOF analysis of rGpFAR1 before lipid removal

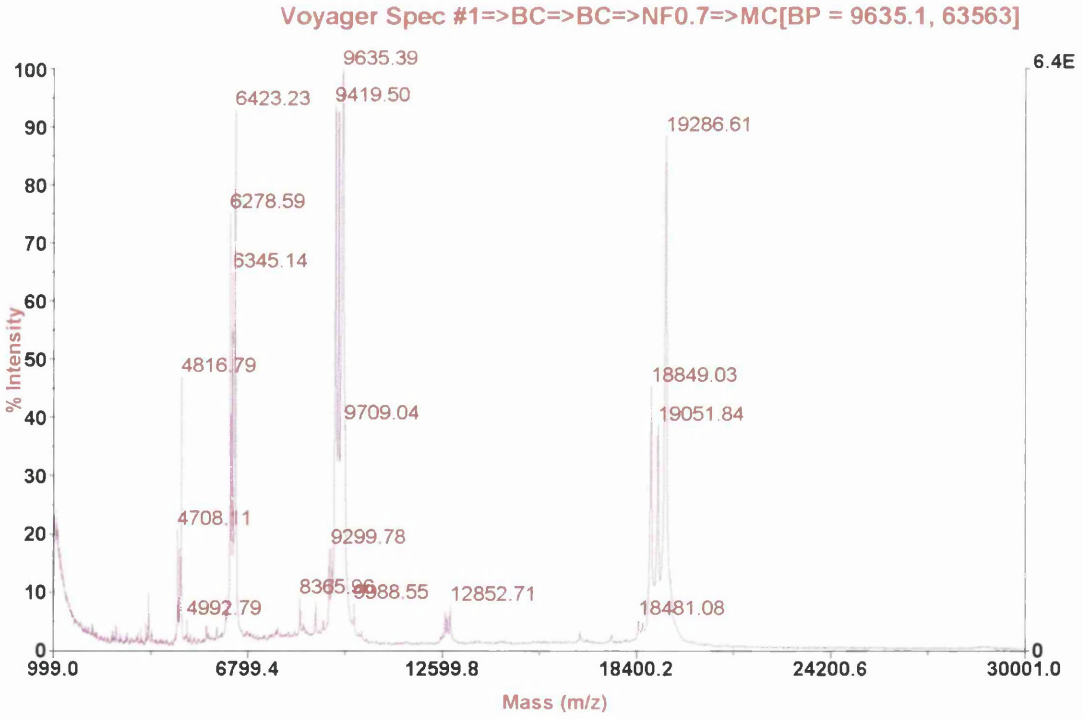

Figure 6.6 MALDI-TOF analysis of rGpFAR1 after lipid removal

The expected molecular mass of rGpFAR1 from the sequence is 18844 Da. The molecular mass from MALDI-TOF before lipid removal was 19286 Da, a difference of 442 Da indicating *E. coli* derived lipid may have been incorporated into the fatty acid binding site or non-covalently bound to the molecular surface. 442 Da could support two C12 fatty acids present in binding site. The spectrum (figure 6.5) clearly shows doubly (9639 Da), triply (6425 Da) and quadruply (4817 Da) charged species. The analysis also shows that no dimeric or higher order oligomers are present. The MALDI-TOF analysis of rGpFAR1 after lipid removal (figure 6.6) shows many more sub peaks compared to the analysis before lipid removal, indicating that the protein is less stable when lipid is removed. However, after lipid removal the molecular weight obtained for rGpFAR1 is still 19287 Da, almost identical to that before the protein is passed down the lipid removal columns. The near identical value (442Da) for before and after delipidating column usage perhaps more than anything, shows that the fatty acid (lipid) may be significantly bound and probably inserted at an early stage of expression. Although, after lipid removal, there are several sub peaks close to the parent peak 19287 Da, this may indicate that the extent of lipid binding has decreased.

Protein	**Difference between observed and calculated (Da)**	**Dimer**
OvFAR1	1225	Yes
CeFAR2	315	No
CeFAR3	63	Yes
GpFAR1 before lipid removal	442	No
GpFAR1 after lipid removal	442	No

Table 6.1 Showing the discrepancies between observed and calculated molecular weights from MALDI.

6.3 Circular dichroism of FAR proteins

Circular dichroism (CD) analysis was carried out as described in chapter 2. The secondary structure charts of the far UV spectra of the FAR proteins rOvFAR1, rCeFAR2, rCeFAR3 and rGpFAR1 are compared in figure 6.6.

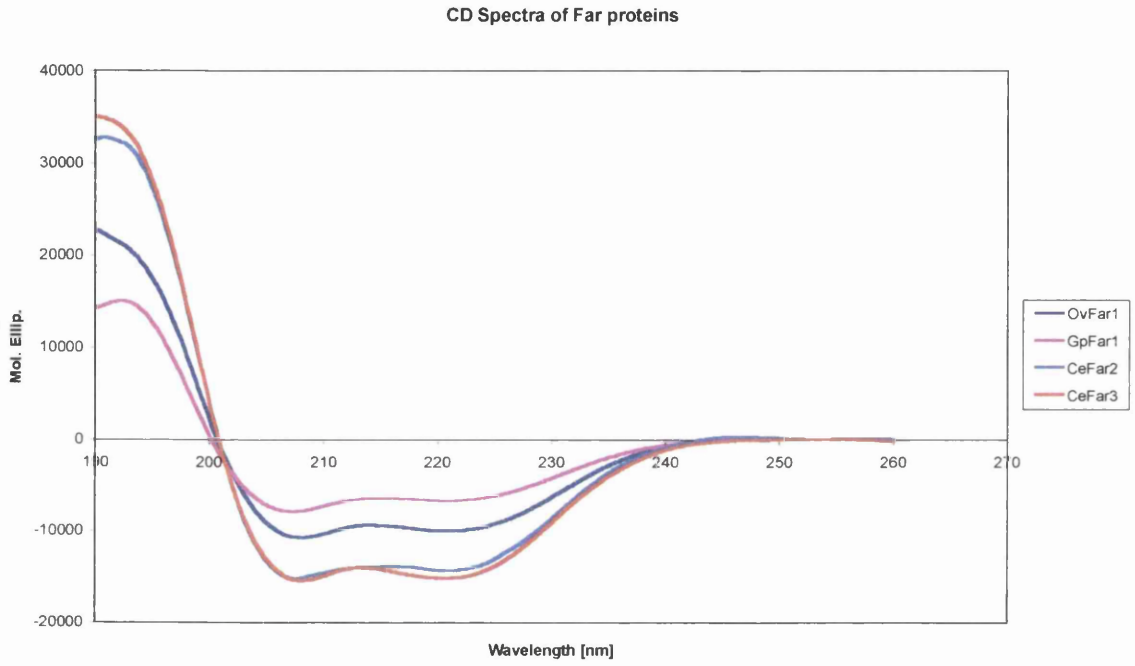

Figure 6.7 Comparison of the CD of the far UV spectra of the Far proteins

The far UV spectra of the FAR proteins shows that they have a similar shape and that the secondary structures of rCeFAR2 and rCeFAR3 are almost identical.

The secondary structure content of each protein was estimated using the CONTIN procedure (table 6.2) of Provencher and Glöckner (1981).

Protein	Secondary structure predictions using CONTIN			
	α-helix	β-sheet	β-turn	remainder
rGpFar1	35%	42%	15%	8%
rOvFar1	27%	32%	27%	15%
rCeFar2	44%	34%	9%	13%
rCeFar3	42%	29%	14%	15%

Table 6.2 Secondary structure content of each protein estimated using the CONTIN procedure.

In a number of studies on proteins it has been found (Kelly & Price, 1997) that the CONTIN procedure tends to overestimate the β -sheet content. The SELCON procedure (table 6.3) (Sreerama & Woody, 1993) has been attributed with a more accurate method of secondary structure assignment. This was used for rOvFAR1, rCeFAR2 and rCeFAR3.

Protein	Secondary structure predictions using SELCON				
	α-helix	anti-parallel β-sheet	parallel β-sheet	turn	other
rOvFar1	31%	12.6%	7.7%	17.7%	31.1%
rCeFar2	44.3%	8.9%	7.3%	14.7%	24.5%
rCeFar3	46.6%	7.7%	7.0%	14.1%	24.9%

Table 6.3 Secondary structure content of each protein estimated using the SELCON procedure.

Recombinant GpFAR1 was a poor match to other proteins using the SELCON procedure and hence its secondary structure was not estimated by this method. The accuracy of these results is directly dependent upon the accuracy of the protein concentrations used to calculate the mean residue ellipticity values. rGpFAR1 did not give any solutions using the SELCON procedure, which may reflect the possibility that the concentration had been overestimated.

The CD spectra show that the recombinant proteins have folded into a secondary structure. The SELCON results show that the rCeFARs and rOvFAR1 proteins contain a higher percentage of α-helix than β-sheet.

CONTIN results show rGpFAR1 contains approximately equal amounts of both α-helix and β-sheet. This shows that these FAR proteins are structurally different from the more ubiquitous lipid binding proteins, which are predominantly 10 anti-parallel β-stranded and 2 short α-helices surrounding the interior binding cavity in a β-barrel type structure (Banaszak *et al*, 1994). Previous CD studies carried out (Garofalo *et al*, 2003) show 54 and 61 % α-helical content for CeFAR2 and CeFAR3 respectively. The main reason for the difference in helical content is probably due to inaccurate measurement of protein concentration. Protein concentration may be inaccurate due to lipid being attached to the protein, as a result the molecular weight is incorrect as witnessed in the MALDI experiment. The presence of lipid can also affect the absorption of the protein sample and also the presence of DNA absorbing at 260 nm lead to errors in protein concentration.

6.4 Dynamic light scattering (DLS) on FAR proteins

The DLS experimental protocol is described in chapter 2. DLS was carried out to investigate the effects of buffer, pH and temperature on each FAR protein to determine the optimum conditions for crystallisation. Each 2 mg/ml protein was investigated in 20 mM ACETATE pH 4.5, 20 mM PIPES pH 6.5, 20 mM HEPES pH 7.5 and 20 mM TRIS pH 8.5, with 150 mM NaCl present in all. Each protein solution was investigated at temperatures between 5 – 25 °C in 5° increments. Lysozyme was used as a standard. (Appendix 1, DLS results)

Polydispersity of FAR proteins

Polydispersity (%) = polydynamic radius (nm) / radius (nm) X 100

The degree of polydispersity of each protein solution may indicate the conditions whereby there is a greater likelihood that crystallisation may occur.

(i) ACETATE

In 20 mM ACETATE pH 4.5 DLS results were obtained for lysozyme, rCeFAR2 and rCeFAR3. rOvFAR1 and rGpFAR1 aggregated in this buffer therefore no results were obtained for these proteins.

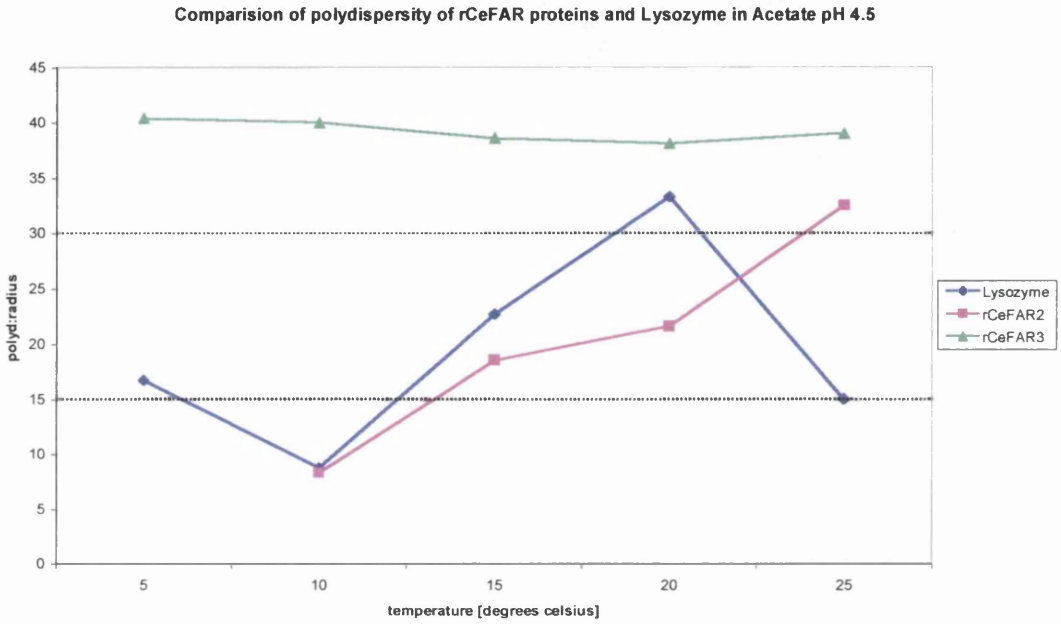

Figure 6.8 Polydispersity of lysozyme, rCeFAR2 and rCeFAR3 in ACETATE pH 4.5

The polydispersity of rCeFAR3 is uniformly greater than 30 % between 5 – 25 °C in ACETATE pH 4.5, which indicates a significant amount of polydispersity indicating that these conditions are unlikely to be optimum for crystallisation. The baseline parameters (appendix 1) for rCeFAR3 indicate narrow monomodal size distribution between 15 – 25 °C and broad monomodal size distribution below 10 °C. The SOS noise (appendix 1) indicates significant background noise between 5 – 25 °C, corroborating the polydispersity values.

The polydispersity for rCeFAR2 is below 15 % at 10 °C in ACETATE pH 4.5 suggesting negligible polydispersity. At 15 °C and 20 °C the polydispersity is

below 30 % indicating a moderate amount of polydispersity and above 25 °C there is a significant amount of polydispersity. The SOS errors (appendix 1) indicate low noise, however the baseline parameters show multimodal size distribution. The results suggest the protein is not homogenous, but that crystallisation in this buffer may be possible at temperatures below 10 °C.

The results for lysozyme in ACETATE pH 4.5 shows moderate to negligible polydispersity and narrow monomodal size distribution with low noise errors, indicating that lysozyme is homogenous and this buffer is suitable for crystallisation. Subsequent analysis on lysozyme with PIPES, HEPES and TRIS show, as expected, very similar results as to that found for ACETATE.

(ii) PIPES

In 20 mM PIPES pH 6.5 DLS results were obtained for lysozyme, rGpFAR1, rOvFAR1 and rCeFAR3. rCeFAR2 aggregated in this buffer therefore no results were obtained.

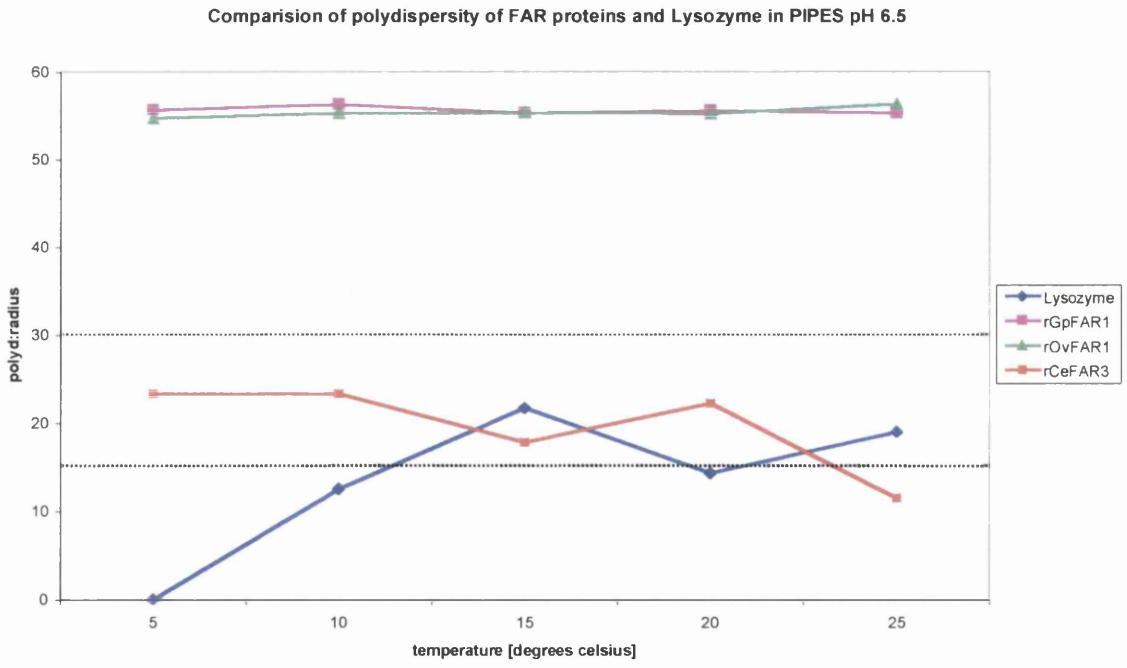

Figure 6.9 Polydispersity of lysozyme, rGpFAR1, rOvFAR1 and rCeFAR3 in PIPES pH 6.5

rGpFAR1 and rOvFAR1 show polydispersity greater than 30 % with baseline errors indicating multimodal size distribution and SOS noise (appendix 1) showing high error for rGpFAR1 and significant background noise for rOvFAR1. These results indicate that PIPES pH 6.5 is not an appropriate buffer for rGpFAR1 and rOvFAR1.

rCeFAR3 shows a moderate amount of polydispersity below 20 °C and negligible polydispersity above 25 °C. The baseline errors (appendix 1) suggest narrow monomodal size distribution and SOS errors (appendix 1) indicate low noise, agreeing with the polydispersity values. rCeFAR3 is monodispersed in 20 mM PIPES pH 6.5 above 25 °C and hence the most favourable buffer and temperature for crystallisation.

(iii) HEPES

In 20 mM HEPES pH 7.5 DLS results were obtained for lysozyme, rGpFAR1, rOvFAR1 and rCeFAR3. rCeFAR2 aggregated in this buffer therefore no results were obtained.

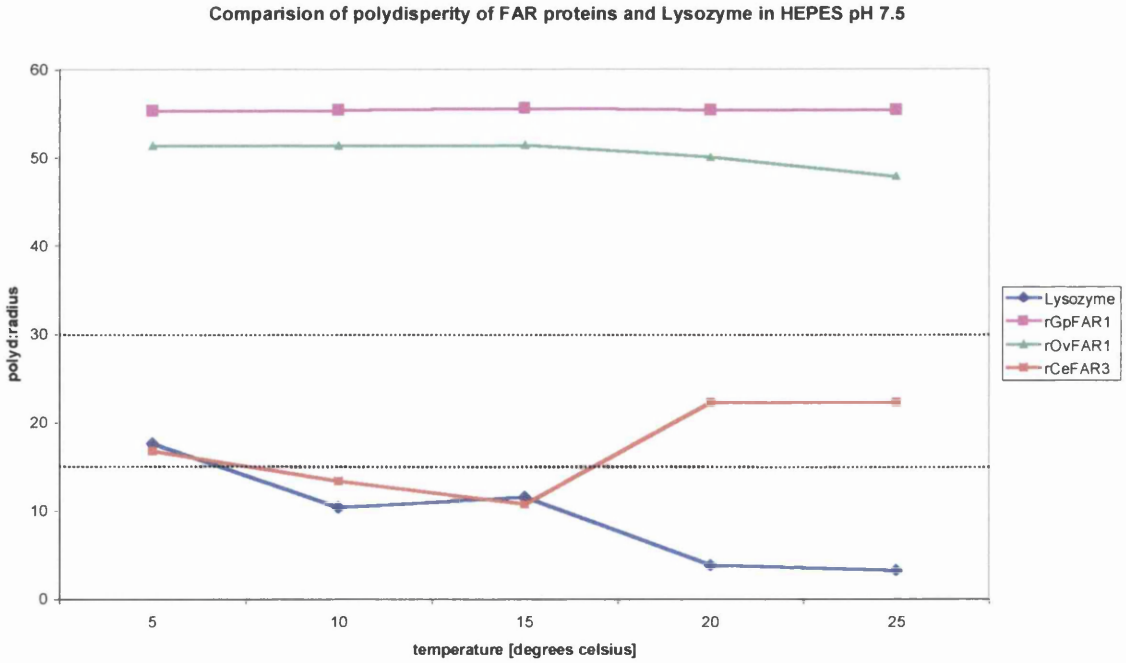

Figure 6.10 Polydispersity of lysozyme, rGpFAR1, rOvFAR1 and rCeFAR3 in HEPES pH 7.5

rGpFAR1 and rOvFAR1 show polydispersity greater than 30 % with baseline errors (appendix 1) indicating multimodal and broad monomodal size distribution and SOS noise (appendix 1) showing high error for rGpFAR1 and significant background noise for rOvFAR1. These results indicate that HEPES pH 7.5 is not an optimum buffer for rGpFAR1 or rOvFAR1.

rCeFAR3 shows a moderate amount of polydispersity above 20 °C and negligible polydispersity below 15 °C. The baseline errors (appendix 1) suggest multimodal and broad monomodal size distribution and SOS errors (appendix 1) indicate significant background noise. Crystallisation of rCeFAR3 could be attempted in 20 mM PIPES pH 6.5 below 15 °C, even though SOS errors are high.

(iv) TRIS

In 20 mM TRIS pH 8.5 DLS results were obtained for lysozyme, rOvFAR1, rCeFAR2 and rCeFAR3. rGpFAR1 aggregated in this buffer therefore no results were obtained.

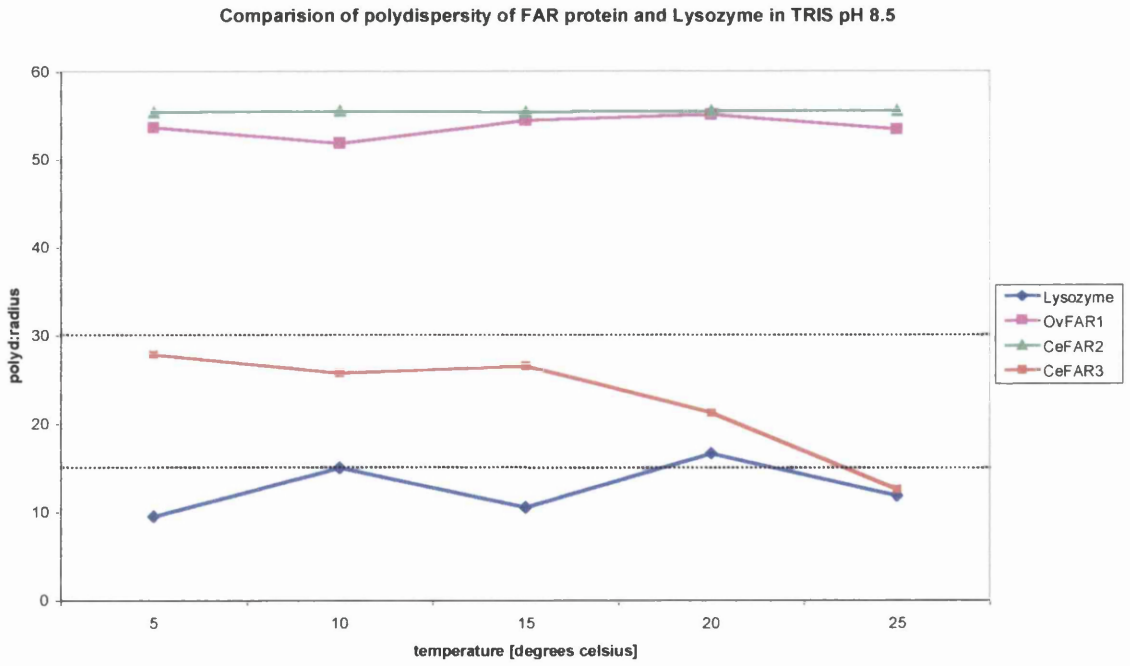

Figure 6.11 – Polydispersity of lysozyme, rOvFAR1, rCeFAR2 and rCeFAR3 in TRIS pH 8.5

rOvFAR1 and rCeFAR2 show polydispersity greater than 30 % with baseline errors (appendix 1) indicating multimodal size distribution and SOS noise (appendix 1) showing significant background noise for rOvFAR1 and high SOS error for rCeFAR2. These results indicate that TRIS pH 7.5 would not be a suitable buffer for rOvFAR1 and rCeFAR2.

rCeFAR3 shows a moderate amount of polydispersity below 20 °C and negligible polydispersity above 25 °C. The baseline errors (appendix 1) suggest narrow monomodal size distribution and SOS errors (appendix 1) indicate low background noise. rCeFAR3 is monodispersed in 20 mM TRIS pH 8.5 at 25 °C.

6.4.1 Molecular weight estimates of FAR proteins using DLS measurements.

The molecular weight (figures 6.12, 6.13, 6.14 and 6.15) of each protein in each solution can be estimated assuming the protein is perfectly globular. If not, then this measurement roughly corresponds to half the maximal dimension of the protein. Therefore this information may infer the state of oligomerisation of the protein. The molecular weight of each protein was only investigated in buffers in which the protein was stable.

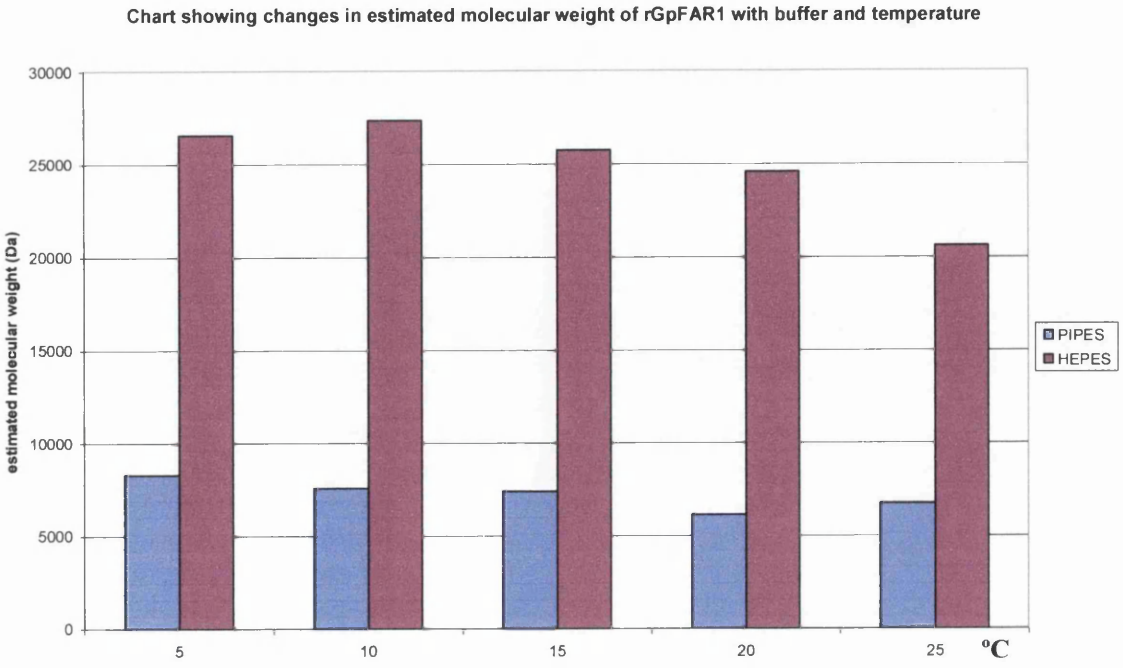

Figure 6.12 Estimated molecular weight of rGpFAR1 in PIPES and HEPES between 5 and 25 °C

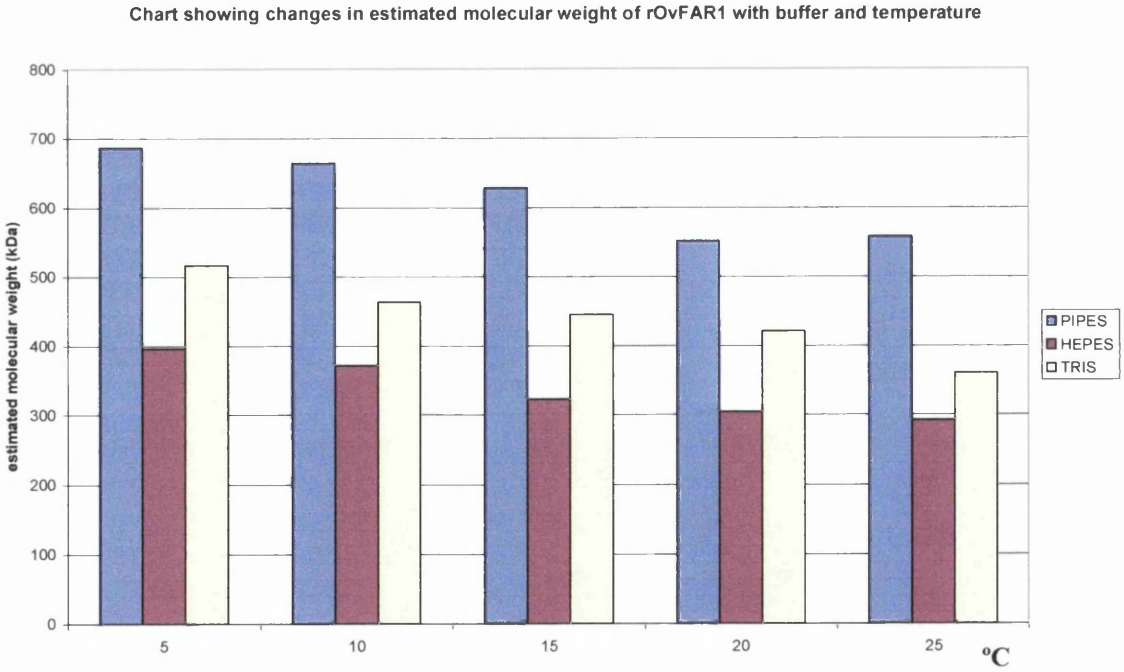

Figure 6.13 Estimated molecular weight of rOvFAR1 in PIPES, HEPES and TRIS between 5 and 25 °C

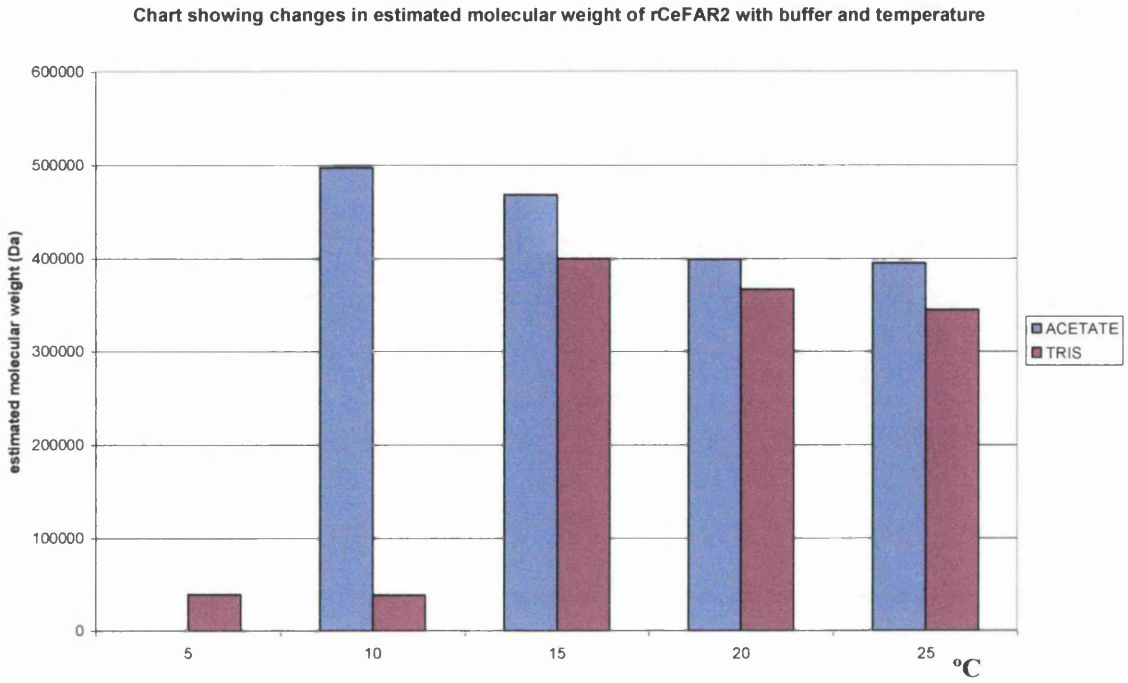

Figure 6.14 Estimated molecular weight of rCeFAR2 in ACETATE and TRIS between 5 and 25 °C

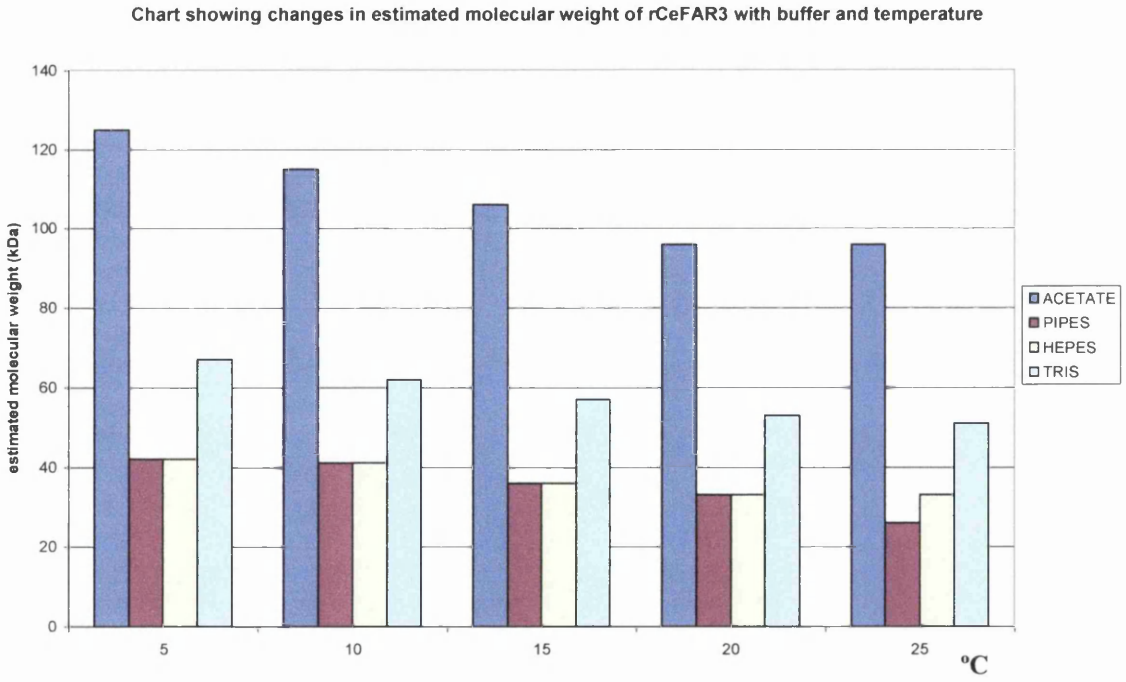

Figure 6.15 Estimated molecular weight of rCeFAR3 in ACETATE, PIPES, HEPES and TRIS between 5 and 25 °C

The graphs (figures 6.12, 6.13, 6.14 and 6.15) give an indication of which buffer and temperature would give a good starting point for crystallisation trials of individual proteins. For rGpFAR1 in HEPES (figure 6.12) the estimated molecular weight is in the range 21 – 27 kDa suggesting crystallisation could be carried out in HEPES buffer since this gives a result close to the actual molecular weight of rGpFAR1, 18.8 kDa. The estimated molecular weight for rOvFAR1 (figure 6.13) is too high being in the range 400 – 700 kDa and indicates that the protein is forming large oligomers and is probably close to aggregation. rOvFAR1 shows a lower estimated molecular weight in HEPES at ambient temperatures (300 kDa) rather than PIPES (600 kDa) and TRIS (400 kDa), suggesting crystallisation could be attempted in HEPES buffer at around room temperature, though all buffers seem unsuitable for crystallisation. rCeFAR2 (figure 6.14) in TRIS below 10 °C shows an estimated molecular weight of 40 kDa, which is closer to the actual molecular weight of rCeFAR2 (20 kDa) than at other temperatures and in other buffers. Crystallisation of rCeFAR2 could therefore be attempted in TRIS buffer at temperatures below 10 °C. MALDI-TOF analysis also showed that rCeFAR3 was present as a dimer in its native state. In PIPES and HEPES buffers the estimated molecular weight of rCeFAR3 (figure 6.15) from DLS is between 25 – 42 kDa, which is close to the real molecular weight of the monomer (20.9 kDa) and dimer (41.8 kDa). It is also encouraging that polydispersity is relatively low for these two buffers. In ACETATE and TRIS buffers the estimated molecular weight was too high. Crystallisation could be carried out in HEPES or PIPES buffer, as the molecular weights indicate they could be suitable buffers for crystallisation.

6.5 Discussion

MALDI-TOF analysis of rGpFAR1 suggested that the protein was a monomer, but that lipids or endotoxins sequestered from *E. coli* overexpression had bound irreversibly, possibly to the fatty acid/retinol binding site. It is probable that two C12 fatty acids may be present in the binding site or, at least strongly bound to the protein. Removal of these lipids or endotoxins resulted in several peaks around 19287 Da, possibly loosening the lipid bound. For hepatocyte nuclear factor 4 (HNF4), an orphan member of the nuclear hormone receptor family with prominent functions in liver, gut, kidney and pancreatic β cells, it has been shown by X-ray crystallography (Dhe-Paganon, 2002) and gas chromatography/mass spectrometry (Wisely *et al*, 2002) that a mixture of fatty acids occupy the binding site. The mixture was found to be saturated and monounsaturated C16-18 fatty acids also found in *E. coli*. These fatty acids were found to be persistently bound after a multistep purification indicating that they were tightly associated and failed to strip the lipids from the protein. Wisely *et al.* have clearly shown that fatty acids can be sequestered from *E. coli* and it is probable that a similar situation exists with the FAR proteins. They also speculate that the fatty acid goes through a selection process during the HNF4 translation and folding and is then trapped in the fully folded protein, therefore the fatty acid does not behave like an exchangeable ligand (Wisely *et al*, 2002). A similar situation appears to exist for rGpFAR1, and probably all the other FAR proteins studied in the project.

MALDI-TOF analysis of rOvFAR1 clearly showed (peak at 43670 Da) that the native structure is likely to exist as a dimer, while DLS analysis has shown the protein either is a large oligomer (400 – 700 kDa) or close to aggregation in all the buffers tried. Heterogeneity is probably due to a monomer-dimer equilibrium or variation in stoichiometry, consisting of one species existing with lipid bound to rOvFAR1 and one with no lipid bound. Unlike rCeFAR3, no higher order oligomerisation is seen from MALTI-TOF analysis of rCeFAR2. However, it is shown from analysis of rCeFAR2 that there are two species of rCeFAR2 present,

possibly one species exists with two lipid molecules bound to rCeFAR2 (20560 Da) and one with only one lipid molecule bound rCeFAR2 (20350 Da), possibly a fatty acid with chain length C20. The difference between the two species of 210 Da may well be the result of an extra fatty acid of C12 chain length, molecular mass 212 Da, being incorporated during folding. Finally, MALDI-TOF analysis of rCeFAR3 shows that oligomers are present and fatty acids may also be bound to the protein. The analysis of rCeFAR2 and rGpFAR1 suggests they are monomers and analysis of rOvFAR1 and rCeFAR3 suggest they are being expressed as dimers and possibly oligomers.

Circular dichroism shows that the recombinant proteins are folded and that they all have essentially the same secondary structure, with rCeFAR2 and rCeFAR3 being almost identical. Analysis by the CONTIN and SELCON procedures indicate that they have a high α-helical content of 35 %, 27 %, 44% and 42 % for rGpFAR1, rOvFAR1, rCeFAR2 and rCeFAR3 respectively by CONTIN and 31 %, 44.3 % and 46.6 % for rOvFAR1, rCeFAR2 and rCeFAR3 respectively by SELCON. Previous CD studies on CeFAR2 and CeFAR3 proteins that show 54 and 61 % α-helix respectively (Garofalo, 2003), which is a higher α-helical content prediction. The main reason for the difference in helical content is probably due to inaccurate protein concentration. Possibly due to lipid being attached to the protein, as a result the molecular weight is incorrect. The presence of lipid can also affect the absorption of the protein sample and also the presence of DNA absorbing at 260 nm lead to errors in protein concentration.

Dynamic light scattering analysis may give an indication of protein homogeneity, purity, aggregation state and monodispersity, all pre-requisites for successful crystallisation. DLS results on lysozyme suggest the technique can indicate the best temperature, buffer and pH for protein crystallisation. rGpFAR1 shows aggregation in ACETATE and TRIS buffer and was shown to be polydispersed in PIPES and HEPES, though the estimated molecular weight of rGpFAR1 in HEPES (20 – 27 kDa) was shown to be close to real molecular weight of

rGpFAR1 (18.8 kDa). This suggests that HEPES may be a better buffer to use in crystallisation trials of rGpFAR1, though polydispersity, baseline errors and SOS noise indicates these conditions are not ideal. rOvFAR1 aggregated in ACETATE buffer and was shown to be polydispersed in PIPES, HEPES and TRIS. From the estimated molecular weights of rOvFAR1 (300-600 kDa) in each condition, DLS suggests that rOvFAR1 may be unsuitable for crystallisation. rCeFAR2 aggregated in PIPES and HEPES buffers and was polydispersed in TRIS buffer. DLS analysis showed rCeFAR2 to have negligible polydispersity in ACETATE buffer at 10 °C, SOS errors show low noise, but the baseline parameter shows multimodal size distribution suggesting the protein is not homogenous, but crystallisation may be possible. The estimated molecular weight calculated by DLS for rCeFAR2 in TRIS buffer below 10 °C was 40 kDa, close to the molecular weight of a dimer of rCeFAR2 (41 kDa), though from MALDI measurements, there is no evidence of a dimer moiety. These results suggest that although rCeFAR2 is not optimised for crystallisation, it may be possible at temperatures below 10 °C. Finally rCeFAR3 was shown to be polydispersed in ACETATE, moderately polydispersed in PIPES and TRIS below 20 °C, but negligibly polydispersed at 25 °C in PIPES and TRIS. Negligible polydispersity was also observed in HEPES below 15 °C. The estimated molecular weight for rCeFAR3 was between 25 – 42 kDa in PIPES and HEPES buffer. Since MALDI-TOF analysis showed that rCeFAR3 is present as a dimer in its native state with a molecular weight of approximately 42 kDa, therefore the molecular weight in HEPES and PIPES buffers is close to the actual molecular weight of the rCeFAR3 monomer (20.9 kDa) and dimer (42 kDa). As a result of these measurements, crystallisation of rCeFAR3 may give better results in HEPES buffer below 15 °C. Although DLS analysis suggests that these proteins are not in the optimum conditions for crystallisation trials, there are many examples of proteins that has been successfully crystallised, even though DLS measurements suggest that conditions were far from optimal (Stura *et al*, 2002; Bergfors, 2003).

The objective of looking at a range of these novel fatty acid and retinol binding proteins was to find one that would crystallise. Since only rGpFAR1 gave crystals DLS was used to determine suitable conditions for crystallisation and possible reasons for the low success rate. However, from the data presented in this chapter it can be concluded that from DLS measurements rOvFAR1, rGpFAR1 and rCeFAR2 are polydispersed and rCeFAR3 has negligible polydispersity in HEPES. MALDI shows that lipids may be attached to rGpFAR1 and rOvFAR1. MALDI also confirmed that rCeFAR2 is a heterogeneous species and that rOvFAR1 and rCeFAR3 are present as oligomers. All of which would make crystallisation a difficult task.

Chapter 7 - Crystallisation studies of FAR proteins

7.1 Crystallisation theory

For crystal formation to occur, molecules have to be brought to a supersaturated, thermodynamically unstable state, which may develop into a crystalline or amorphous phase when it returns to equilibrium. Supersaturation (figure 7.1) can be achieved by slow evaporation of the solvent or by varying parameters, such as ionic strength, pH and temperature. Crystallisation starts by a nucleation phase, followed by a growth phase. Nucleation requires greater supersaturation than growth, and the rate of crystallisation increases when supersaturation increases. In general, it is desirable to avoid any amorphous precipitate and produce large, flawless crystals.

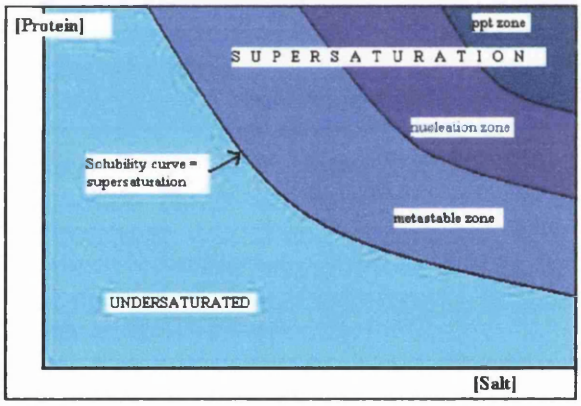

Figure 7.1 Schematic diagram of the states involved in crystallisation

Nucleation

The first step in crystallisation is nucleation. It is the process by which free molecules or non-crystalline aggregates (dimers, trimers, etc) in a supersaturated solution produce a stable aggregate repeating lattice. Crystals have ~3-6 kcal/mole lower free energy than the relative solution state (Drenth & Haas,

1992). The aggregate must exceed a critical size (between 10 – 200 molecules) to form a macroscopic crystal, otherwise spontaneous dissolution will occur and non-specific aggregates and noncrystalline precipitation will form. The degree of supersaturation determines the degree of nucleation that occurs. Therefore, higher degrees of supersaturation produces more stable aggregates and increases the likelihood of the formation of stable nuclei. Thus supersaturation must be approached slowly to achieve only a few nuclei and resultant larger crystals in preference to many nuclei and showers of small crystals.

Growth

Crystal growth occurs at solute concentrations sufficient for nucleation to occur and continues at concentrations beneath the nucleation threshold. Growth rate is determined by the growing crystal surface and the diffusional rate of molecules in the crystallising solution (McPherson, 1995). Ideally, crystals suitable for diffraction have dimensions around 0.2-0.3 mm in all dimensions.

Cessation of growth

There are many reasons for crystal growth to stop. The decrease in concentration of the crystallising solute to a point where the solid and solution phases reach equilibrium or cumulative lattice strain effects (Feher, 1985). Poisoning of growth surfaces (Sato *et al*, 1992) due to impurities or damaged crystallising protein resulting in defects on the crystal surface preventing further packing of molecules. McPherson *et al.* have used atomic force microscopy to study the crystals as they grow to see how a range of impurities, including foreign particles, influence the properties of the crystals (McPherson, 2000). Macroseeding or microseeding (chapter 7.1.3) can be attempted to further crystal growth.

7.1.1 Crystallisation parameters

There are a number of factors that affect nucleation and growth of macromolecules, including type of precipitant, precipitant concentration, protein concentration, temperature, pH and factors specific to the protein molecule, such as ligands and co factors.

Precipitants

The use of precipitants is the most widely used method of achieving protein supersaturation, thereby inducing crystallisation by competing for water of solvation. The range of precipitants includes salts, high molecular weight straight chain polymers (PEG), MPD and organic solvents. These precipitants are often used at a range of concentrations in sparse matrix trials (Jancarik & Kim, 1991). If a precipitant yields crystals, a fine linear grid screen surrounding this condition can be set up to optimise the initial condition. Different ratios of precipitant to protein concentration can influence the rate of nucleation and crystal growth.

pH

Proteins are charged species and the net charge often determines its solubility. Buffers included in the crystallisation solution control the pH of the solution. It is typical that crystallisation occurs over a narrow range of pH (<1 pH unit). Crystal morphology, including twinned and polynucleated growth forms are often related to pH (McPherson, 1995). As the proper pH is approached a decrease in crystal quality on either side of the optimal condition is observed.

Temperature

Crystallisation of proteins has been accomplished in a range of 0 – 60 °C. Temperature is an important parameter; for example, small rapidly formed

crystals grown at room temperature can yield fewer, larger crystals at a lower temperature.

Molecule specific factors

Ligands and small molecules have been shown to affect crystallisation of proteins (Carter & Yin, 1994). Interactions of other molecules with the protein may fix the protein into one conformation or may alter interactions between solvent, other molecules and the protein. Two other chemicals that can be used to initiate or enhance crystallisation are viscosity altering compounds and anti-twinning, solubilising compounds (detergents and ethers). By including these compounds in the mother liquor, they may alter the rate of crystal growth.

7.1.2 Vapour diffusion experiment

Several techniques have been used to attain supersaturation for crystallisation (Bergfors, 1999). The vapour diffusion technique (figure 7.2) utilises evaporation and diffusion of water between solutions of different concentrations as a means of approaching and achieving supersaturation of macromolecules. It is the optimal technique to use when screening a large number of conditions. In a vapour diffusion experiment the solution containing the protein is mixed 1:1 with the reservoir solution. The drop containing the mixture has both solutions diluted to half their original concentration and is then suspended and sealed over the well solution, which contains the precipitant at the target concentration. Differences in precipitant concentration between drop and well means a non-equilibrium exists that causes water to evaporate from the drop until the concentration of precipitant in the drop equals that of well solution and equilibrium is reached.

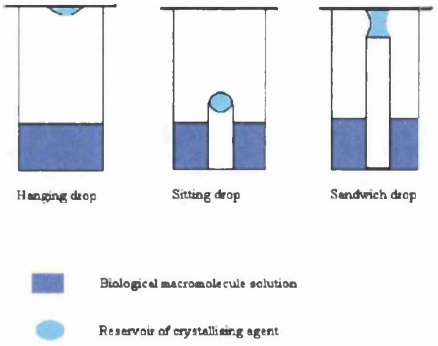

Figure 7.2 – Vapour diffusion techniques

7.1.3 Seeding to promote crystallisation

Seeding is a method used to provide nucleation sites for crystal growth. A protein crystal or crystalline aggregate is placed in a solution of slightly lower precipitant concentration, crushed and mixed to make a seed solution. Ratios of 1:10, 1:100 and 1:1000 of this seed solution are then added to the crystallisation drop, ensuring the presence of nuclei for crystal growth (Bergfors, 2003). There are two commonly used methods of seeding: microseeding involves the transfer of submicroscopic seeds, which are too small to be distinguished individually and macroseeding, which involves the transfer of a single crystal.

7.1.4 Gelled surface crystallisation

Gels promote the growth of a few larger crystals by reducing excessive nucleation and precipitation. They can change the morphology of the crystal or slow crystal nucleation. Silica gels (Hampton research) are stable over a wide range of temperatures and are compatible with a wide variety of precipitants and additives used for crystal growth. This experiment can be carried out in a sitting drop plate.

7.1.5 Crystallisation with oils

The use of oil as a sealant can play an important role in the crystallisation experiment (Chayen, 1999). The principal role of oil is to act as an inert sealant to prevent evaporation of small-volume trials and also to provide a controlled environment for the nucleation and growth stages of crystallisation. The rationale behind the use of oils is that water can evaporate at different rates through different oils. Paraffin oil can act as a good sealant allowing only negligible amounts of water evaporate through it. In contrast, water can diffuse freely through silicone oils. A mixture of paraffin and silicone oil permits partial diffusion, depending on the ratio at which they are mixed.

The crystallisation techniques outlined above have been used in attempts to grow crystals of sufficient diffraction quality for the FAR proteins used in this project.

7.2 Crystallisation of FAR proteins

7.2.1 rOvFAR1

MALDI and DLS analysis (chapter 6) indicated that rOvFAR1 was heterogeneous and polydispersed in PIPES, HEPES and TRIS. Though, DLS suggested that crystallisation could be set up in 20 mM HEPES pH 7.5 at room temperature since the polydispersity and aggregation state was slightly lower than in other buffers tried. Crystallisation was attempted using Cryschem plates (Hampton), using the standard sitting drop vapour diffusion technique by mixing 4 µl of the protein solution with an equal volume of crystallisation solution from the footprint screen (www.prg/stura/cryst/foot) and equilibrated against 1 ml of the same crystallising solution. The trials were set up at various protein concentrations of 2, 5, 8, 10 and 12 mg/ml to check for the best concentration. 9 mg/ml was found to be the optimum concentration for the rOvFAR1 protein as it resulted in light precipitation in several wells. More trials were set up with 4 µl of 9 mg/ml

rOvFAR1 using the standard sitting drop vapour diffusion technique and crystallising solutions from sparse matrix solutions (Brzozowski & Walton, 2001; Jancarik & Kim, 1991; Cryo I and II, Wizard I and II, www.emeraldbiostructures.com; Crystal Clear Strategy Screens, www.moleculardimensions.com; www.hamptonresearch.com). After 1 week small needle shaped crystals (>10μm) were formed in 0.1 M TRIS pH 8.5, 10 % PEG 8K and 0.5 M $MgCl_2$ at 25 °C. Optimisation of PEG concentration from 5 – 30 % w/v, pH from 6 – 9.5 of TRIS did not result in larger needles that would have been suitable for diffraction studies. Optimisation was also attempted at 4 and 15 °C, with no improvement to crystal size. Other trials resulted in precipitation and crystalline aggregates, but no protein crystals suitable for diffraction experiments.

7.2.2 rCeFAR2

DLS analysis of rCeFAR2 (chapter 6) suggested that at 10 °C in 20 mM ACETATE pH 4.5 the protein has negligible polydispersity and also in 20 mM TRIS pH 8.5 below 10 °C the estimated molecular weight of rCeFAR2 was that of a dimer of rCeFAR2 (41 kDa), although no dimer was observed from MALDI. Crystallisation trials were set up under these conditions. The Cryschem (Hampton) sitting drop vapour diffusion plate was used with rCeFAR2 at various concentrations between 2 – 10 mg/ml with the footprint screen (www.prg/stura/cryst/foot) as for rOvFAR1. Light precipitation resulted at 7 mg/ml whereas a higher concentration produced very heavy precipitation. Crystallisation trials were set up with 4 μl drops of rCeFAR2 at 7 mg/ml using the many available sparse matrix screens (Brzozowski & Walton, 2001; Jancarik & Kim, 1991; Cryo I and II, Wizard I and II, www.emeraldbiostructures.com; com; Crystal Clear Strategy Screens, www.moleculardimensions.com; www.hamptonresearch.com). rCeFAR2 in 20 mM ACETATE pH 4.5 was tried but at 7 mg/ml the protein precipitated out of solution and at lower concentrations of CeFAR2 no crystals or precipitation resulted. rCeFAR2 trials in 20 mM TRIS

pH 8.5 at 10 and 25 °C resulted in heavy precipitation in many drops, but no crystals resulted from any of the sparse matrix screens.

7.2.3 rCeFAR3

rCeFAR3 crystallisations were set up in 20 mM PIPES pH 6.5, 20 mM HEPES pH 7.5 and 20 mM TRIS pH 8.5 at 20 °C since DLS indicated negligible polydispersity (chapter 6) in these buffers. MALDI however, (chapter 6) suggested that rCeFAR3 had different oligomerisation states and therefore was heterogeneous. Crystallisation was attempted using the standard sitting drop vapour diffusion technique on Cryschem (Hampton) by mixing 4 µl of the protein solution with an equal volume of crystallisation solution from the footprint screen (www.prg/stura/cryst/foot) and equilibrated against 1 ml of the same crystallising solution at a variety of concentrations from 2 mg/ml to 12 mg/ml. The optimum concentration was 7 mg/ml, as this produced light precipitation. Trials were set up with 4 µl of 7 mg/ml CeFAR3 using sitting drop plates and crystallising solutions from sparse matrix solutions (Brzozowski & Walton, 2001; Jancarik & Kim, 1991; Cryo I and II, Wizard I and II, www.emeraldbiostructures.com; Crystal Clear Strategy Screens, www.moleculardimensions.com; www.hamptonresearch.com). These trials resulted in extensive precipitation and crystalline aggregates, but no protein crystals suitable for diffraction.

7.2.4 rGpFAR1

MALDI analysis (chapter 6) indicated rGpFAR1 had slight heterogeneity, probably due to presence of *E. coli* lipids as discussed in previous chapters. DLS results (chapter 6) agreed with MALDI analysis suggesting rGpFAR1 was not monodispersed, though in 20 mM HEPES pH 7.5 at 25 °C the estimated molecular weight of rGpFAR1 was approximately 20 kDa, which is close to the actual molecular weight of GpFAR1 (18.8 kDa). Therefore rGpFAR1 crystallisations were set up in 20 mM HEPES pH 7.5 at room temperature. The optimum

concentration of rGpFAR1 was established using the standard sitting drop vapour diffusion technique with Cryschem plates (Hampton) and mixing 4 µl of the protein solution with an equal volume of crystallisation solution from the footprint screen (www.prg/stura/cryst/foot) equilibrated against 1 ml of the same crystallising solution. The trials were set up at various concentrations between 2 - 15 mg/ml to check for the best concentration. 9.6 mg/ml was found to be the optimum concentration for the rGpFAR1 protein as it resulted in light precipitation in several wells whereas greater than 10 mg/ml resulted in heavy precipitation. Trials were set up with 4 µl of 9.6 mg/ml rGpFAR1 in 20 mM HEPES pH 7.5 with crystallising solutions from commercially available sparse matrix screens (Brzozowski & Walton, 2001; Jancarik & Kim, 1991; Cryo I and II, Wizard I and II, www.emeraldbiostructures.com; Crystal Clear Strategy Screens, www.moleculardimensions.com; www.hamptonresearch.com). Condition 3 M Na/K Phosphate pH 7.5, 0.1 M Na Tartrate and 0.1 M HEPES pH 7.5 gave small rod shaped crystals (figure 7.3).

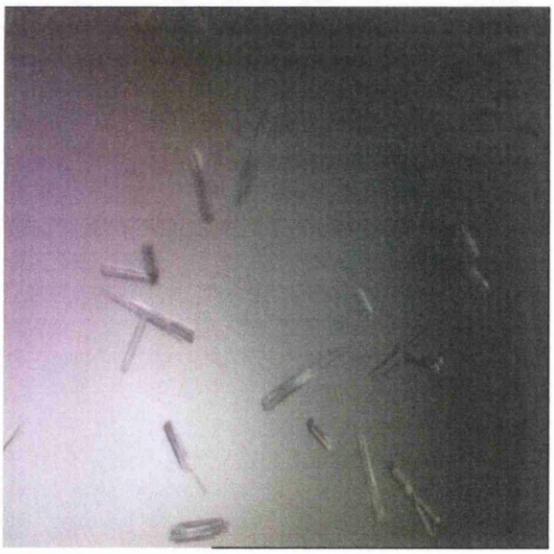

Figure 7.3 Initial rod shaped crystals in 3 M Na/K Phosphate pH 7.5, 0.1 M Na Tartrate and 0.1 M HEPES pH 7.5.

This condition for rGpFAR1, buffered with 20 mM HEPES pH 7.5, was optimised to give crystals (1 X 0.4 X 0.1 mm) large enough for diffraction (3 M Na/K Phosphate pH 7.5, 0.1 M Na Tartrate and 0.1 M HEPES pH 7.3 (figure 7.4)).

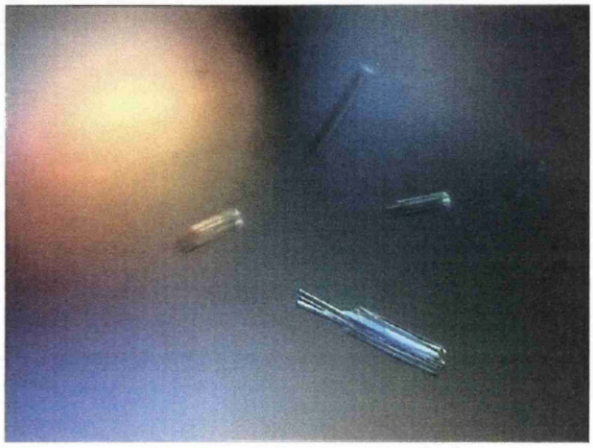

Figure 7.4 rGpFAR1 rod shaped crystals in 3 M Na/K Phosphate pH 7.5, 0.1 M Na Tartrate and 0.1 M HEPES pH 7.3.

On close inspection, these rods were found to be many rod shaped crystals bundled together. Even with extreme care it proved impossible to separate these bundles into individual crystals. It was then decided to try and optimise each component of the crystallising solution in turn. + and – tartrate were tried as a replacement for the racemic mixture of tartrate. A finer sampling of pH was tried varying the pH of the 3 M Na/K Phosphate present in the optimised condition (3 M Na/K Phosphate pH 7.5, 0.1 M Na Tartrate and 0.1 M HEPES pH 7.3) between pH 6 – 8 by 0.05 pH units - no larger crystals were obtained. In 20 mM PIPES pH 6.5 in the same condition (3 M Na/K Phosphate, 0.1 M Na Tartrate and 0.1 M HEPES pH 7.3) longer rod shaped crystals (figure 7.5) were obtained.

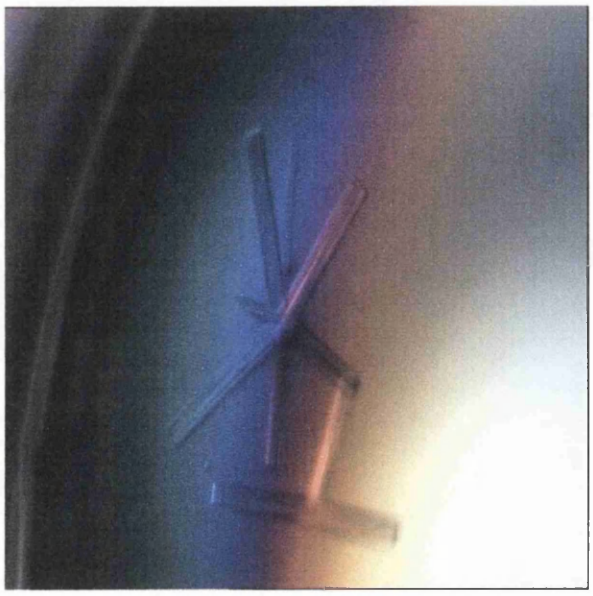

Figure 7.5 Optimised rGpFAR1 crystals in 3 M Na/K Phosphate pH 7.5, 0.1 M Na Tartrate and 0.1 M HEPES pH 7.3.

Additive screens 1 – 3 from Hampton were added to the optimised conditions to try and grow single crystals; 5 – 20 % glycerol was added to try and improve the crystals - neither resulted in single crystals. Further techniques were attempted to grow single crystals of rGpFAR1: these included microseeding and streak seeding, growth of rGpFAR1 crystals in gels, growth of rGpFAR1 crystals under oil with varying percentages of silicon and paraffin oils all as described earlier. Seeding of rGpFAR1 into the optimised conditions resulted in smaller needle crystals being grown in bundles. Crystal growth in silica hydrogel (Hampton) was tried in an attempt to reduce excessive nucleation and grow fewer, larger crystals. 5 µl of gel was pipetted onto the post of the Cryschem plate and allowed to polymerise, then 5µl of precipitant was added. Once the precipitant was completely absorbed 10 µl rGpFAR1 at 9.6 mg/ml was placed onto the gel and once absorbed, the reservoir was filled with 1 ml of precipitant. The precipitant used for rGpFAR1 trials was 3 M Na/K Phosphate pH 7.5, 0.1 M Na Tartrate and 0.1 M HEPES pH 7.3. The pH of precipitant was also varied from pH 6 – 9, as was the concentration of Na Tartrate from 0.05 – 0.25 M. Crystal trials in gels

resulted in no crystals being grown. Growth of crystals under mixtures of 10, 25, 50 and 75 % silicon oil with 80, 75, 50 and 25 % paraffin oil respectively, resulted in smaller, needle crystals but again in bundles.

Crystals were also grown using the optimised conditions (3 M Na/K Phosphate pH 7.5, 0.1 M Na Tartrate and 0.1 M HEPES pH 7.3) at 4 °C and 15 °C as well as room temperature (25 °C). It was found that larger crystals were grown at room temperature (25 °C). Temperature shifting was tried by moving crystals from room temperature (25 °C) to 15 °C after 24 hours and vise versa and shifting the temperature between room temperature and 4 °C. All without improvement under visual inspection.

The hanging drop method was used instead of sitting drop, but again needle crystals in bundles were grown and were not of sufficient quality to use in diffraction experiments.

In chapter 6 it was demonstrated that fatty acids or lipids may have bound to the protein during expression in *E. coli*. In chapter 4 it was also shown that fatty acids, which are bound in the active site can be competitively displaced by other fatty acids. This poor crystal quality we were experiencing, despite exhaustive attempts to improve by rational optimisation, may be a result of heterogeneous binding of fatty acids. To try to minimise this heterogeneity 1 mM of fatty acids were added to 9.6 mg/ml rGpFAR1 before crystallisation. Palmitic acid was used as rGpFAR1 was shown to bind fatty acid chain length C15 with the highest affinity (Prior *et al*, 2001). Oleic acid was also used as it has been shown (chapter 4) to compete with other fatty acids and displace them from the binding site. Trials were set up using vapour diffusion experiments against the optimised condition, 3 M Na/K Phosphate pH 7.5, 0.1 M Na Tartrate and 0.1 M HEPES pH 7.3 and varying the pH of Na/K Phosphate from pH 6 – 9, varying concentration of Na Tartrate from 0.05 – 0.25 M, varying the protein concentration from 8.5 –

10.5 mg/ml and grown at 25 and 4 °C. These experiments resulted in small needle crystals forming, which were not large enough for diffraction experiments.

All the methods tried did not produce single crystals of rGpFAR1. Therefore the needle crystals of rGpFAR1 (figure 7.5) were used for diffraction at Daresbury SRS and ESRF at Grenoble by positioning the crystals so that a single rod was in the beam.

7.3 Data collection

7.3.1 Crystal Mounting

Once crystallisation conditions have been optimised and a crystal of suitable size and quality has been produced, the crystal needs to be stabilized before being aligned in the X-ray beam to allow the diffraction experiment to be performed. Removal of the crystal from the precipitant solution may damage it, as solvent makes up a large proportion of the volume of the crystal. Hence, the crystal needs to be kept in a stabilised state in the precipitant solution or in a saturated vapour of the mother liquor. Two methods commonly used to mount crystals are mounting in a glass capillary and frozen mounting in a loop (Garman, 1999).

Capillary mounting

The crystal is lifted from the drop and placed in a glass capillary tube of suitable diameter and sealed at one end. The thin walls of the glass capillary tube reduce the amount of absorption of X-rays. To prevent the crystal drying out, a reservoir of mother liquor solution is placed either side of the crystal and the tube is sealed.

Loop mounting

Crystals are lifted from the drop into a loop (from Hampton) of a similar size to the crystals. To prevent the crystal from being destroyed the crystals are dipped, or transferred, into a cryoprotectant. The cryoprotectant protects the crystal from shattering when introduced into the freezing liquid nitrogen stream. Successful cryoprotection should result in the formation of a transparent glass, which will greatly extend the lifetime of the crystal in the intense synchrotron X-ray beam. Cryoprotectants can also be introduced slowly into the mother liquor to lessen crystal damage. Cryoprotectants include dried paraffin oil or reservoir solution with added glycerol, MPD or sodium formate.

7.3.2 Synchrotron Radiation

Synchrotron radiation is generated when electrons (or positrons) are accelerated to near light speed and then forced to change course by the use of powerful magnets. The radiation emitted is composed of all wavelengths (white light) within the electromagnetic spectrum ranging from ultra-violet to short wavelength X-rays. A synchrotron radiation source is very intense making the collection of data not only viable, but faster than would be possible from weakly diffracting crystals and very small crystals. The wavelength of the radiation from a synchrotron can be tuned (to the absorption edge of heavy metals bound to the protein) so that multiple wavelength anomalous dispersion (MAD) experiments can be performed.

7.4 Data collection of rGpFAR1 crystals

Needle crystals of rGpFAR1 were lifted from the drop using a cryoloop of suitable size. Paraffin oil was used as a cryoprotectant, which formed a glass around the crystal on freezing in a nitrogen gas stream at 100K. Data were

collected at Daresbury, station 14.1 on an ADSC Quantum 4R CCD diffractometer at a wavelength of 1.488 Å and a distance of 350 mm between the detector and the crystal. Data were collected to 4 Å (figure 7.6).

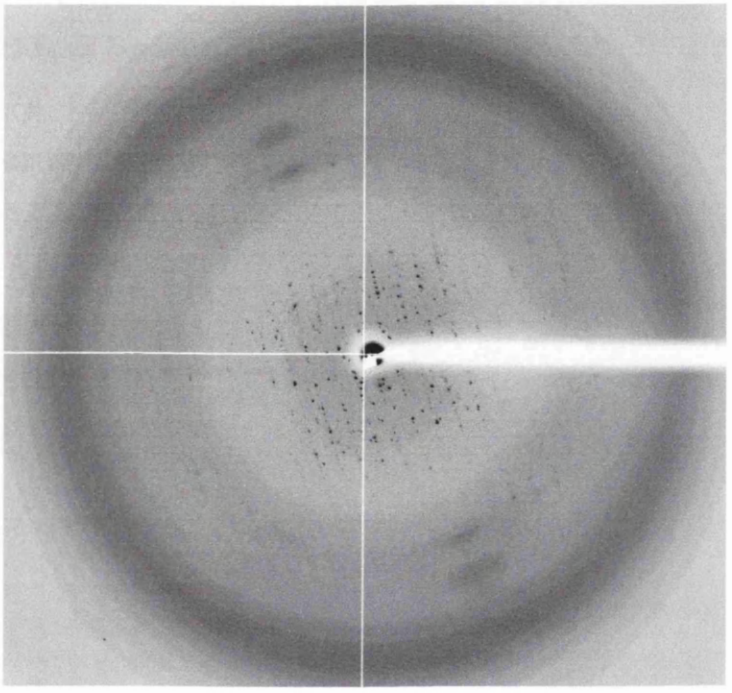

Figure 7.6 Diffraction of rGpFAR1 crystal to 4 Å on station 14.1 at Daresbury SRS.

20 % glycerol equilibrated in the crystallisation solution as well as 8 M Na formate were also trialed as cryoprotectants by immersing a rGpFAR1 crystal in each solution for varied lengths of time. Unfortunately, the diffraction did not improve.

From the diffraction in figure 7.6 it can be seen that more than one crystal lattice is present, showing that each rod broken from the rGpFAR1 crystal contains more than one crystal or else the needles may be hollow. Auto indexing using DENZO (Otwinowski & Minor, 1997) and MOSFLM (Powell, 1999) could not be accomplished, as the programs could not distinguish between the different crystal lattices.

Crystals were mounted in capillaries to ascertain if damage was a result of the cryoprotectants. Data were also collected at SRS Daresbury station 9.5 on an ADSC Quantum 4R CCD diffractometer. The crystal was mounted in a capillary and data were collected at room temperature. At a wavelength of 1.22 Å and a distance of 250 mm between the detector and the crystal data were collected to 2.8 Å (figure 7.7). Diffraction also showed that the lattice points were very close together indicating the presence of more than one lattice.

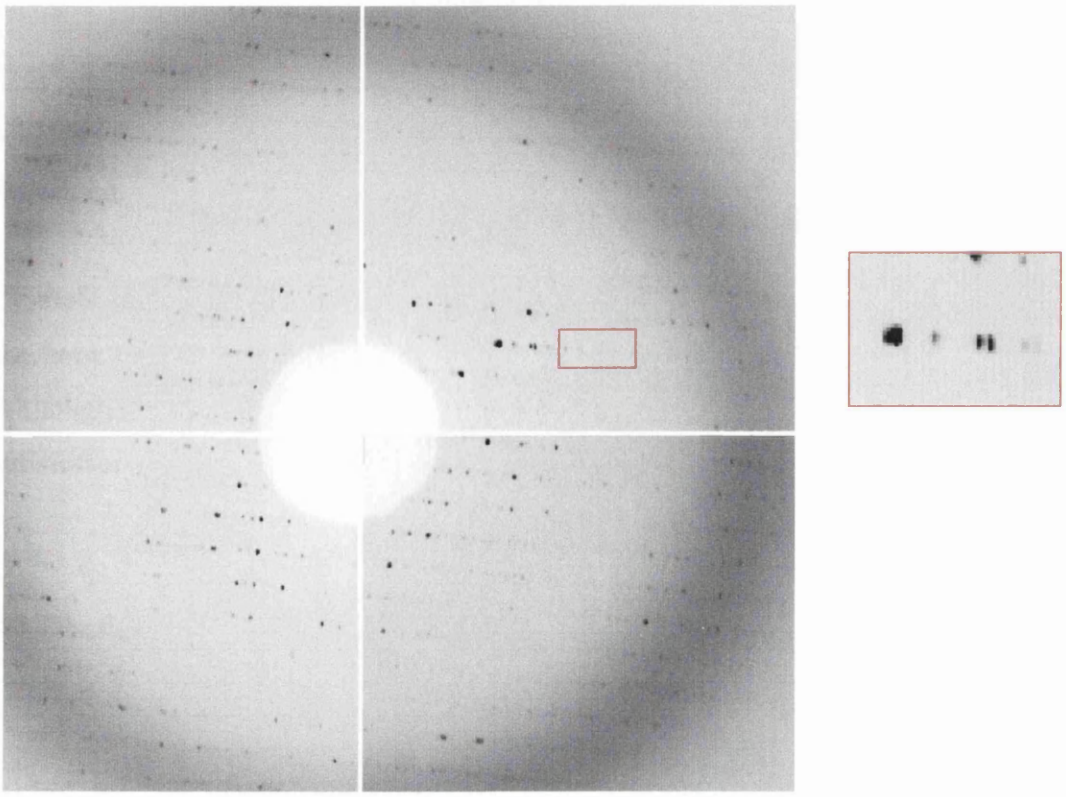

Figure 7.7 – Diffraction to 2.8 Å of rGpFAR1 crystals mounted in a capillary tube at station 9.5 at Daresbury SRS

The program DENZO was used to search for the peaks by hand and different unit cells were tried to determine which fitted the diffraction pattern best. Individual diffraction patterns for 0, 90 and 120° ϕ rotation positions were collected. 1° oscillation was used for each exposure.

For 0 degree rotation the crystal is probably orthorhombic with unit cell dimensions: 93.07, 47.88, 155.19 Å.

However, at 90 and 120 degrees rotation, the crystal lattice is possibly C centered monoclinic, with unit cell dimensions: 181.44, 47.83, 179.91 Å, $\beta = 115.15$.

This suggests that there is more than one lattice present in the crystal (figure 7.7), with the result that a complete data set could not be collected and successfully processed.

Crystals were also taken to the microfocus beamline at the ESRF (Grenoble) in the hope that it would be possible to find part of a rod that was a single crystal. Although several crystals were tried and different cryoprotectants used, satisfactory diffraction could not be attained.

7.5 Discussion

Crystallisation trials of rOvFAR1 resulted in no crystals suitable for diffraction. 9 mg/ml was found to be the optimum concentration for crystallisation trials. Unfortunately, the protein may not be pure enough for crystals to be grown as two bands were shown by SDS-PAGE (chapter 4) and possibly the protein has bound lipids acquired during expression in *E. coli* (chapter 4 and 6).

Crystallisation trials of rCeFAR2 and rCeFAR3 resulted in no crystals suitable for diffraction. 7 mg/ml was found to be the optimum concentration for

crystallisation trials. Unfortunately, like other FAR proteins it is likely that the rCeFAR proteins have bound fatty acids during overexpression in *E. coli* (chapter 6).

rGpFAR1 crystallisation trials resulted in needle crystals grown in condition 3 M Na/K Phosphate pH 7.5, 0.1 M Na Tartrate and 0.1 M HEPES pH 7.3. The crystals could not be improved to provide single crystals, though many methods were tried, including seeding, oils, growth in gels, pH variation and temperature shifting. Diffraction of these needle crystals showed that they were twinned or hollow. Twinning of crystals occurs during crystal growth as nuclei sometimes come together in entirely different orientations in ways that achieve good molecular packing but violate crystal symmetry. The twinning or hollow crystals of rGpFAR1 resulted in unsolvable diffraction. This could be due to the protein not being pure enough because lipids have bound during expression in *E. coli* (chapter 6). Adding fatty acids to displace the lipids bound to rGpFAR1 did not improve the crystals.

New purification methods are required for the FAR proteins or ways to stop the proteins binding lipids during their overexpression. Other crystallisation conditions could also be attempted to give better crystals that give diffraction to a high resolution in a single space group allowing the data to be processed.

Chapter 8 – Conclusions

8.1 Nematode polyprotein allergens

ABA-1 purified from *Ascaris suum* and purified recombinantly was found to have similar, if not identical activity, as both show an increase and blue shift in fluorescence emission when bound to DAUDA. With equivalent amounts of pABA-1 and rABA-1 shifts from 542 nm to 478 nm and 476 nm respectively were observed (chapter 3.2 and 3.3). Circular dichroism (figure 3.9) shows that both pABA-1 and rABA-1 have folded secondary structure with high percentages of α-helix, 53 and 46 % for pABA-1 and rABA-1 respectively. The bimodal scatter diagram (figure 3.10) and gaussian graphical presentation (figure 3.11) from DLS studies suggest that rABA-1 is possible a dimer. Estimated molecular weight indicates rABA-1 is not globular, confirmed by initial NMR studies suggesting that ABA-1 is a two domain structure (N. Meenan, A. Cooper, private communication). However, rABA-1 was shown to be heterogeneous from DLS analysis and both were confirmed to be heterogeneous from IEF gel analysis (figure 3.8). The species present could be due to isoforms of ABA-1 or heterogeneity due to a mixture of fatty acids present from *E. coli*, which may be in competition for the binding site. The current NMR studies (N Meenan, A Cooper, private communication) also show this ambiguity in fatty acid present in the binding site. Crystallisation trials were unsuccessful due to the heterogeneity of ABA-1. This would indicate that new purification methods are required or ways to stop the proteins binding lipids during their overexpression before it will be possible to crystallise ABA-1 and possibly other NPAs.

8.2 Fatty acid and retinol binding proteins

GpFAR1, OvFAR1, CeFAR2 and CeFAR3 were all produced recombinantly in this project. The overexpression was in *E. coli*, which has fatty acids present, including palmitic acid (fatty acid chain length 16), myristic acid (fatty acid chain

length 14) and cyclo heptadecanoic acid (fatty acid chain length 17 with a cyclopropyl group) (Dhe-Paganon *et al*, 2002). Fluorescence studies carried out in chapter 4 have shown that these proteins have an affinity for fatty acids with a chain length C14-C18. This is significant since not only are these recombinant proteins behaving as they do *in vivo*, they also bind the fatty acids which are similar in length to those found in *E. coli*. MALDI-TOF analysis of OvFAR1, CeFAR2 and CeFAR3 suggested that lipid may have bound to the proteins, giving a heterogeneous mixture of, possibly, lipid bound protein and unbound protein (chapter 6), the heterogeneous mixture was also confirmed by DLS. MALDI-TOF analysis of rGpFAR1 suggested lipids had bound irreversibly, possibly to the fatty acid/retinol binding site. Hepatocyte nuclear factor 4 (HNF4), discussed in chapter 6.5, has a mixture of fatty acids occupying the binding site and it is also speculated that the fatty acid goes through a selection process during the HNF4 translation and folding and is then trapped in the fully folded protein, therefore the fatty acid does not behave like an exchangeable ligand (Wisely *et al*, 2002). It is possible that this is the case for the FAR proteins and also ABA-1. Hence future overexpression of the FAR proteins should be carried out using a different expression system, possibly in Pichia (yeast), this is currently being attempted at Prof. Bradley's lab, Nottingham University.

Fluorescence studies showed competitive effects with the retinol-FAR protein complex when 10 mM oleic acid was added. This indicates that the binding site for retinol and fatty acids is interactive or congruent in the case of all the FAR proteins. Binding retinol is important for OvFAR1 as it is thought that the nematode *Onchocerca volvulus* that secretes the protein OvFAR1 requires retinol for a variety of metabolic and developmental purposes including, growth, differentiation, embryogensis, glycoprotein synthesis and anti-oxidants. This suggests that the FAR proteins role in the nematodes is related to retinol acquisition or transport.

Circular dichroism shows that the recombinant proteins are folded and that they all have essentially the same secondary structure, with rCeFAR2 and rCeFAR3 being almost identical. Analysis by the CONTIN and SELCON procedures indicate that they have a high α-helical content of 35 %, 27 %, 44% and 42 % for rGpFAR1, rOvFAR1, rCeFAR2 and rCeFAR3 respectively by CONTIN and 31 %, 44.3 % and 46.6 % for rOvFAR1, rCeFAR2 and rCeFAR3 respectively by SELCON, though accurate protein concentration was a problem. This indicates that they are a structurally different from the more ubiquitous lipid binding proteins from the family of cytosolic proteins, which are predominantly 10 anti-parallel β-stranded with 2 short α-helices surrounding the interior binding cavity in a β-barrel type structure (Banaszak *et al*, 1994).

Finally, rGpFAR1 was the only protein in this study to crystallize. However, diffraction of these needle crystals showed that they were twinned or hollow and resulted in unsolvable diffraction. A new purification method possibly involving a different expression system or ways to stop the protein binding lipids during overexpression could lead to crystals that diffract. Other crystallisation conditions could also be attempted to give better crystals that diffract to high resolution in a single space group allowing quality data to be collected and processed.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J. (1990). Basic local alignment search tool. J. Mol. Biol. **215**: 403-410.

Banaszak, L., Winter, N., Xu, Z., Bernlohr, D. A., Cowan, S., Jones, T. A. (1994). Lipid-binding proteins: a family of fatty acid and retinoid transport proteins. Adv. Protein Chem. **45**: 89-151.

Barr, P. J. (1991). Mammalian subtilisins: the long sought dibasic processing endoproteases. Cell. **66**: 1-3.

Bergfors, T. (2003). Seeds to crystals. J. Struct. Biol. **142**: 66-76.

Blaxter, M. (1998). *Caenorhabditis elegans* is a nematode. Science. **282**(5396): 2041-2046.

Bradley, J. E., Kennedy, M. (1999). Retinoids, novel retinoid-binding proteins and river blindness. Retinoids. **15**(1).

Brahms, S., Brahms, J. (1980). Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. J. Mol. Biol. **138**(2): 149-178.

Britton, C., Moore, J., Gilleard, J.S., Kennedy, M.W. (1995). Extensive diversity in repeat unit sequences of the cDNA encoding the polyprotein antigen/allergen from the bovine lungworm *Dictyocaulus viviparus*. Mol. Biochem. Parasitol. **72**(1-2): 77-88.

Brzozowski, A. M., Walton, J. (2001). Clear strategy screens for macromolecular crystallization. J. Appl. Cryst. **34**: 1-4.

Carter, C. W., Yin, Y. (1994). Quantitative analysis in the characterisation and optimization of protein crystal growth. Acta. Cryst. **D50**: 572-590.

Chayen, N. E. (1999). Crystallization with oils: a new dimension in macromolecular crystal growth. J. Cryst. Growth. **196**: 434-441.

Chou, P.Y., Fasman, G. (1974). Prediction of protein conformation. Biochemistry. **13**(2): 22-245.

Christie, J.F., Dunbar, B., Davidson, I., Kennedy, M.W. (1990). N-terminal amino acid sequence identity between a major allergen of *Ascaris lumbricoides* and *Ascaris suum*, and MHC-restricted IgE responses to it. Immunology. **69**(4): 596-602.

Christie, J, F., Dunbar, B., Kennedy, M.W. (1993). The ABA-1 allergen of the

nematode *Ascaris suum*: epitope stability, mass spectrometry, and N-terminal sequence comparison with its homologue in *Toxocara canis*. Clin. Exp. Immunol. **92**(1): 125-132.

Cooper, A., McReynolds, L., Kennedy, M.W. (1995). *C. elegans* makes a fatty acid/retinoid-binding polyprotein. Wormy serum albumin? Worm Breeder's Gazette. **14**(1): 76.

Consortium, *C. elegans*. sequencing. (1998). Genome sequence of the nematode *Caenorhabditis elegans*: A platform for investigating biology. Science. **282**: 2012-2018.

Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y., McPherson, A. (1994). Screening and optimization strategies for macromolecular crystal growth. Acta. Cryst. **D50**: 414-423.

D'Arcy, A. (1994). Crystallizing proteins – a rational approach? Acta. Cryst. **D50**: 469-471.

Dhe-Paganon, S., Duda, K., Iwamoto, M., Chi, Y., Shoelson, S. E.(2002). Crystal structure of the HNF4α ligand binding domain in complex with endogenous fatty acid ligand. J. Biol. Chem. **277**(41): 37973-37976.

Drenth, J., Haas, C. (1992). Protein crystals and their stability. J. Cryst. Growth. **122**: 107-109.

Duncan, L. H., Robertson, L. Robertson, W. M., Kusel, J. R. (1997). Isolation and characterisation of secretions from the plant-parasitic nematode *Globodera pallida*. Parasitology. **115**: 429-438.

Feher, G., Kam, Z. (1985). Nucleation and growth of protein crystals: general principles and assays. Meth. Enz. **114**: 77-111.

Garman, E. (1999). Cool data: quantity and quality. Acta. Cryst. **D55**: 1641-1653.

Garnier, J. Osguthorpe, D., Robson, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. **120**(1): 97-120.

Garofalo, A, Klager, S., Rowlinson, M., Nirmalan, N., Klion, A., Allen, J.E., Kennedy, M.W., Bradley, J.E. (2002). The FAR proteins of filarial nematodes: secretion, glycosylation and lipid binding characteristics. Mol. Biochem. Parasitol. **122**: 161-170.

Garofalo, A., Rowlinson, M., Amambua, N. A., Hughes, J. M., Kelly, S. M., Price, N. C., Cooper, A., Watson, D. G., Kennedy, M. W., Bradley, J. E. (2003).

The FAR protein family of the nematode *Caenorhabditis elegans*. J. Biol. Chem. **278**(10): 8065-8074.

Hemley, R., Kohler, B. E., Siviski, P. (1979). Absorption spectra for the complexes formed from vitamin A and beta-lactoglobulin. Biophys. J. **28**(3): 447-455.

Hennessey, J.P. Jr., Johnson, W. J. (1981). Information content in the circular dichroism of proteins. Biochemistry. **20**(5): 1088-1094.

Hillenkamp, F., Karas, M., Bevis, R.C., Chait, B. T. (1991). Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. Anal. Chem. **63**(24): 1193A-1203A.

Hodgkins, J., Plasterk, R. H. A., Waterson, R. H. (1995). The nematode *Caenorhabditis elegans* and its genome. Science. **270**: 410-414.

Holz, R. A., Wright, D. J., Perry, R. N. (1997). The lipid content and fatty acid composition of hatched second stage juveniles of *Globodera rostochiensis* and *G. pallida*. Fundam. appl. Nematol. **20**(3): 291-298.

Holz, R. A., Wright, D. J., Perry, R. N. (1998a). The influence of the host plant on lipid reserves of *Globodera rostochiensis*. Nematologica. **44**: 153-169.

Holz, R. A., Wright, D. J., Perry, R. N. (1998b). Changes in the lipid content and fatty acid composition of 2nd-stage juveniles of *Globodera rostochiensis* after rehydration, exposure to the hatching stimulus and hatch. Parasitology. **116**: 183-190.

Jancarik, J., Kim, S. H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. J. Appl. Cryst. **24**: 409-411.

Kelly, S. M., Price, N. C. (1997). Applications of circular dichroism to studies of protein folding and unfolding. Biochim. Biophys. Acta. **1338**: 161-185.

Kennedy. M.W., Fraser, E., Christie, J.F. (1991). MHC class II (I-A) region control of the IgE antibody repertoire to the ABA-1 allergen of the nematode *Ascaris*. Immunology. **72**(4): 577-579.

Kennedy, M.W., Allen, J., Wright, A.S., McCruden, A.B., Cooper, A. (1995a). The gp15/400 polyprotein antigen of *Brugia malayi* binds fatty acids and retinoids. Mol. Biochem. Parasitol. **71**: 41-50.

Kennedy, M.W., Brass, A., McCruden, A.B., Price, N.C., Kelly, S.M., Cooper, A. (1995b). The ABA-1 allergen of the parasitic nematode *Ascaris suum*: fatty acid

and retinoid binding function and structural characterization. Biochemistry. **34**(20): 6700-6710.

Kennedy, M.W., Garside, L., Goodrick, L.E., McDermott, L., Brass, A., Price, N.C., Kelly, S.M., Cooper, A., Bradley, J.E. (1997). The Ov20 protein of the parasitic nematode *Onchocerca volvulus*. A structurally novel class of small helix-rich retinol-binding proteins. J. Biol. Chem. **272**(47): 29442-29448.

Kennedy, M. W. (2000). The nematode polyprotein allergens/antigens. Parasitology today. **16**(9).

Lal, P. G., James, E. R. (1996). *Onchocerca* retinol- and invermectin-binding protein activity. Parasitology.**112**: 221-225.

Lupas, A., Van Dyke, M., Stock, J. (1991). Predicting coiled coils of the protein sequences. Science. **252**: 1162-1164.

MacGregor, R.B., Weber, G. (1986). Estimation of the polarity of the protein interior by optical spectroscopy. Nature. **319**(6048): 70-73.

McDermott, L., Cooper, A., Kennedy, M. W. (1999). Novel classes of fatty acid and retinol binding protein from nematodes. Molecular and Cellular Biochemistry. **192**: 69-75.

McDermott, L. (1999). Structural and functional analysis of the ABA-1 allergen of the nematode *Ascaris*. Ph.D. thesis, University of Glasgow, UK.

McDermott, L, Moore, J., Brass, A., Price, N.C., Kelly, S.M., Cooper, A., Kennedy, M.W. (2001). Mutagenic and chemical modification of the ABA-1 allergen of the nematode *Ascaris*: consequences for structure and lipid binding properties. Biochemistry. **40**(33): 9918-9926.

McDermott, L, Kennedy, M. W., McManus, D.P., Bradley, J.E., Cooper, A., Storch, J. (2002). How helminth lipid-binding proteins offload their ligands to membranes: differential mechanisms of fatty acid transfer by the ABA-1 polyprotein allergen and Ov-FAR-1 proteins of nematodes and Sj-FABPc of schistosomes. Biochemistry. **41**(21): 6706-6713.

McGibbon, A.M., Christie, J.F., Kennedy, M.W., Lee, T.D. (1990). Identification of the major *Ascaris* allergen and its purification to homogeneity by high-performance liquid chromatography. Mol. Biochem. Parasitol. **39**(2): 163-171.

McPherson, A. (1995). Increasing the size of microcrystals by fine sampling of pH limits. J. Appl. Cryst. **28**: 362-365.

Moore, J., McDermott, L., Price, N.C., Kelly, S.M., Cooper, A., Kennedy, M.W. (1999). Sequence-divergent units of the ABA-1 polyprotein array of the nematode *Ascaris suum* have similar fatty-acid- and retinol-binding properties but different binding-site environments. J. Biochem. **340**(1): 337-343.

Muller, P. (1975). Worms and disease: A manual of medical helminthology. Heinemann.

Mustafa, K. Y., Turumen, U., Gumaa, K. A., (1979). Serum vitamin A levels of patients with onchocerciasis from two areas of Sudan. Tropical medicine and hygiene. **82** : 122-127.

Narayan, M., Berliner, L. J. (1997). Fatty acids and retinoids bind independently and simultaneously to beta-lactoglobulin. Biochemistry. **36**(7): 1903-1911.

Neilson, H., Engelbrecht, J., Brunak, S., Heijne, G. V. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engineering. **10**: 1-6.

Otwinowski, Z., Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Meth. Enz. **276**: 307-326.

Papiz, M. Z., Sawyer, L., Eliopoulos, E. E., North, A. C., Findlay, J. B., Sivapradarao, R., Jones, T. A., Newcomer, M. E., Kraulis, P. J. (1986). The structure of beta-lactoglobulin and its similarity to plasma retinol-binding protein. Nature. **324**(6095): 383-385.

Poole, C.B., Hornstra, L., Benner, J.S., Fink, J.R., McReynolds, L.A. (1996). Carboxy-terminal sequence divergence and processing of the polyprotein antigen from *Dirofilaria immitis*. Mol. Biochem. Parasitol. **82**(1): 51-65.

Powell, H. R. (1999). The Rossman Fourier autoindexing algorithm in MOSFLM. Acta. Cryst. D Biol. Cryst. **55**(10): 1690-1695.

Prior, A, Jones, J., Blok, V.C., Beauchamp, J., McDermott, L., Cooper, A., Kennedy, M.W. (2001). A surface-associated retinol- and fatty acid-binding protein (Gp-FAR-1) from the potato cyst nematode *Globodera pallida*: lipid binding activities, structural analysis and expression pattern. J. Biochem. **356**(2): 387-394.

Provencher, S.W., Glockner, J. (1981). Estimation of globular protein secondary structure from circular dichroism. Biochemistry. **20**(1): 33-37.

Rost, B., Sander, C., Schneider, R .(1994). PHD an automatic mail server for protein secondary structure prediction. Comput. Appl. Biosci. **10**(1): 53-60.

Sani, B. P., Vaid, A. (1985). The Ov20 protein of the parasitic nematode *Onchocerca volvulus*. J. Biol. Chem. **232**: 557-583.

Sasser, J. N., Freckman, D. W. (1987) A commemoration of the 25th anniversary of the Society of Nematologists. In Vistas on Nematology. Society of Nematologists. : 7-15.

Sato, K., Fukuba, Y., Mitsuda, T., Hirai, K., Moriya, K. (1992). Observation of lattice defects in orthorhombic hen-egg white lysozyme crystal with laser scattering tomography. J. Cryst. growth. **122**: 87-94.

Solovyova. A.S., Meenan, N., McDermott, L., Garofalo, A., Bradley, J.E., Kennedy, M.W., Byron, O. (2003). The polyprotein and FAR lipid binding proteins of nematodes: shape and monomer/dimer states in ligand-free and bound forms. European Biophysics Journal.

Spence, H.J., Moore, J., Brass, A., Kennedy, M.W. (1993). A cDNA encoding repeating units of the ABA-1 allergen of *Ascaris*. Mol. Biochem. Parasitol. **57**(2): 339-343.

Sreerama, N., Woody, R. W. (1994). Protein secondary structure from circular dichroism spectroscopy. Combining variable selection principle and cluster analysis with neural network, ridge regression and self-consistent methods. J. Mol. Biol. **242**(4): 497-507.

Stura, E.A., Martin, L., Lortat-Jacob, H., Vives, R., Vita, C. (2002). Heparin-aggregated RANTES can be crystallised. Acta Cryst. D Biol. Cryst. **58**(10 and 1): 1670-1673.

Thumser, A.E., Evans, C., Worrall, A.F., Wilton, D.C. (1994). Effect on ligand binding of arginine mutations in recombinant rat liver fatty acid-binding protein. Biochem. J. **297**(1): 103-107

Tomlinson, L.A., Christie, J. F., Fraser, E. M., McLaughlin, D., McIntosh, A. E., Kennedy, M. W. (1989). MHC restriction of the antibody repertoire to secretory antigens, and a major allergen of the parasitic nematode *Ascaris*. J. Immunolgy. **143**: 2349-2356.

Tree, T. I., Gillespie, A. J., Shepley, K. J., Blaxter, M. L., Tuan, R. S., Bradley, J. E. (1995). Characterisation of an immunodominant glycoprotein antigen of *Onchocerca volvulus* with homologues in other filarial nematodes and *Caenorhabditis elegans*. Mol. Biochem. Parasitol. **69**: 185-195.

Wilkinson, T. C. I., Wilton, D. C. (1986). Studies on fatty-acid binding proteins – the detection and quantitation of the protein from rat liver by using a fluorescent fatty acid analog. J. Biochem. **238**: 419-424.

Wisely, G. B., Miller, A. B., Davis, R. G., Thornquest, Jr., A. D., Johnson, R., Spitzer, T., Sefler, A., Shearer, B., Moore, J. T., Miller, A. B., Willson, M., Williams, S. P. (2002). Hepatocyte Nuclear Factor 4 is a transcription factor that constitutively binds fatty acids. Structure. **10**: 1225-1234.

Wood, W. B. (1988). The nematode *Caenorhabditis elegans*, Cold Spring Harbor Laboratory: 1-867.

Xia, Y., Spence, H., Moore, J., Heaney, N., McDermott, L., Cooper, A., Watson, D.G., Mei, B., Komuniecki, R., Kennedy, M.W. (2000). The ABA-1 allergen of *Ascaris lumbricoides*: sequence polymorphism, stage and tissue-specific expression, lipid binding function, and protein biophysical properties. Parasitology. **120**(2): 211-224.

http://ca.expasy.org/sprot/

http://ca.expasy.org/tools/protparam.html

http://cbrweb.med.harvard.edu/investigators/springer/lab/protocols/sara_light_scattering

http://elegans.swmed.edu

www.elephantiasis.freeyellow.com

www.trachomahki.org/imageye3

www.gsk.com/filariasis

www.inra.fr/hyppz/RAVAGEUR/bgloros.htm

www.who.int/tdr/diseases/oncho/lifecycle.htm

www.biosci.ohio-state.edu/~parasite/lifecycles/onchocerca_lifecycle.html

www.biosci.ohio-state.edu/~parasite/lifecycles/ascaris_lifecycle.html

www.emeraldbiostructures.com

www.moleculardimensions.com

www.hamptonresearch.com

Appendix 1

Experiment: rABA-1 in PIPES at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.6	401.0	3.4	1.9	59.0	5.1	21994.0	1.1	37.6
2	20.0	0.6	419.0	3.3	1.8	52.0	5.1	21662.0	1.1	23.2
3	39.0	0.6	412.0	3.3	1.9	54.0	4.9	21755.0	1.0	38.3
4	59.0	0.6	414.0	3.3	1.9	54.0	5.1	21122.0	1.0	18.3
5	80.0	0.6	403.0	3.4	1.9	58.0	5.2	21029.0	1.0	25.6
6	101.0	0.6	443.0	3.1	1.7	46.0	5.1	20407.0	1.0	28.3
7	123.0	0.6	454.0	3.0	1.7	43.0	5.0	20625.0	1.0	39.4

Experiment: rABA-1 in PIPES at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.5	441.0	3.5	2.0	62.0	9.1	16701.0	1.0	27.8
2	25.0	0.5	492.0	3.2	1.8	50.0	9.9	16470.0	1.0	31.3
3	51.0	0.5	499.0	3.1	1.8	47.0	9.4	16785.0	1.0	33.7
4	77.0	0.5	420.0	3.8	2.1	73.0	9.8	18631.0	1.0	36.8
5	102.0	0.5	473.0	3.3	1.9	56.0	10.0	17580.0	1.0	28.3
6	127.0	0.5	439.0	3.6	2.0	67.0	10.1	19402.0	1.0	46.8
7	150.0	0.5	508.0	3.1	1.8	47.0	10.0	18777.0	1.0	29.4

Experiment: rABA-1 in PIPES at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.5	575.0	3.1	1.8	47.0	14.6	18348.0	1.0	26.6
2	24.0	0.5	569.0	3.2	1.8	49.0	14.8	17975.0	1.0	25.8
3	49.0	0.5	550.0	3.3	1.9	53.0	14.7	19134.0	1.0	30.1
4	71.0	0.5	535.0	3.4	1.9	57.0	14.8	20262.0	1.0	50.4
5	95.0	0.5	545.0	3.3	1.8	55.0	14.8	20241.0	1.0	18.4

Experiment: rABA-1 in PIPES at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.6	590.0	3.5	2.0	61.0	19.4	22269.0	1.0	47.9
2	21.0	0.6	678.0	3.1	1.7	45.0	19.7	21497.0	1.0	25.4
3	42.0	0.6	595.0	3.5	2.0	61.0	19.7	23528.0	1.0	41.4
4	62.0	0.6	630.0	3.3	1.9	54.0	19.9	22799.0	1.0	38.1
5	82.0	0.6	681.0	3.0	1.7	44.0	19.8	22805.0	1.0	23.7
6	101.0	0.6	668.0	3.1	1.8	47.0	19.8	23461.0	1.0	38.4

Experiment: rABA-1 in PIPES at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.6	704.0	3.3	1.9	55.0	24.4	25729.0	1.0	41.8
2	18.0	0.6	769.0	3.1	1.7	45.0	24.5	24879.0	1.0	20.8
3	37.0	0.6	737.0	3.2	1.8	50.0	24.6	26396.0	1.0	40.5
4	54.0	0.6	742.0	3.2	1.8	49.0	24.7	27343.0	1.0	44.6
5	88.0	0.6	760.0	3.1	1.8	47.0	24.7	27050.0	1.0	29.6
6	105.0	0.6	683.0	3.5	1.9	60.0	24.6	28125.0	1.0	30.2
7	122.0	0.6	738.0	3.2	1.8	50.0	24.8	27554.0	1.0	26.9
8	139.0	0.6	770.0	3.1	1.7	45.0	24.6	27620.0	1.0	22.7

Experiment: rABA-1 in HEPES at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.6	458.0	3.0	1.7	41.0	4.7	39000.0	1.0	19.9
2	33.0	0.6	448.0	3.0	1.7	44.0	4.8	39497.0	1.0	29.6
3	65.0	0.6	448.0	3.0	1.7	43.0	4.6	38726.0	1.0	22.6
4	98.0	0.6	411.0	3.3	1.9	53.0	4.6	40930.0	1.0	29.6
5	130.0	0.5	416.0	3.3	1.8	52.0	4.7	43591.0	1.0	22.3
6	161.0	0.6	429.0	3.2	1.8	49.0	4.9	40077.0	1.0	27.6
7	191.0	0.6	392.0	3.5	1.9	60.0	4.7	41947.0	1.0	30.6
8	222.0	0.6	410.0	3.3	1.9	54.0	4.6	40602.0	1.0	34.1
9	254.0	0.6	426.0	3.2	1.8	49.0	4.6	39562.0	1.0	25.3
10	285.0	0.6	443.0	3.1	1.7	45.0	4.7	38458.0	1.0	25.7

Experiment: rABA-1 in HEPES at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.6	494.0	3.2	1.8	48.0	9.5	34557.0	1.0	33.0
2	14.0	0.6	535.0	2.9	1.6	40.0	9.5	33672.0	1.0	20.9
3	29.0	0.6	494.0	3.1	1.8	48.0	9.4	33366.0	1.0	29.3
4	43.0	0.6	496.0	3.1	1.8	47.0	9.3	34171.0	1.0	27.3
5	57.0	0.6	522.0	3.0	1.7	42.0	9.4	34912.0	1.0	33.4
6	72.0	0.6	481.0	3.2	1.8	51.0	9.4	35405.0	1.0	32.1
7	87.0	0.6	516.0	3.0	1.7	43.0	9.4	34225.0	1.0	25.3
8	102.0	0.6	491.0	3.2	1.8	49.0	9.4	35036.0	1.0	24.9
9	117.0	0.6	496.0	3.1	1.8	48.0	9.5	35105.0	1.0	32.8
10	131.0	0.6	526.0	3.0	1.7	41.0	9.4	33869.0	1.0	21.3

Experiment: rABA-1 in HEPES at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.6	514.0	3.5	1.9	60.0	14.1	39150.0	1.0	47.9
2	31.0	0.6	625.0	2.9	1.6	38.0	14.3	36185.0	1.0	29.8
3	46.0	0.6	579.0	3.1	1.7	46.0	14.3	36349.0	1.0	30.2
4	77.0	0.6	585.0	3.0	1.7	44.0	14.1	36956.0	1.0	35.1
5	91.0	0.6	586.0	3.0	1.7	44.0	14.1	36688.0	1.0	32.3
6	126.0	0.6	621.0	2.9	1.6	38.0	14.1	35130.0	1.0	13.4
7	159.0	0.6	636.0	2.8	1.6	36.0	14.2	35116.0	1.0	23.1
8	174.0	0.6	597.0	3.0	1.7	42.0	14.2	36685.0	1.0	35.3
9	187.0	0.6	598.0	3.0	1.7	42.0	14.3	36418.0	1.0	28.7
10	200.0	0.6	644.0	2.8	1.6	35.0	14.1	34696.0	1.0	18.4

Experiment: rABA-1 in HEPES at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.6	755.0	2.7	1.5	32.0	18.8	34422.0	1.0	17.2
2	14.0	0.6	727.0	2.8	1.6	36.0	18.9	35299.0	1.0	38.5
3	28.0	0.7	745.0	2.7	1.5	34.0	19.2	33865.0	1.0	22.6
4	43.0	0.7	717.0	2.8	1.6	38.0	19.2	35837.0	1.0	31.8
5	57.0	0.6	725.0	2.8	1.6	36.0	19.0	34332.0	1.0	29.1
6	72.0	0.6	728.0	2.8	1.6	36.0	19.0	35972.0	1.0	30.5
7	86.0	0.6	700.0	2.9	1.6	39.0	19.0	35343.0	1.0	33.6
8	101.0	0.6	719.0	2.8	1.6	37.0	18.9	35098.0	1.0	23.7
9	133.0	0.7	778.0	2.6	1.5	31.0	19.2	34750.0	1.0	19.8
10	148.0	0.6	743.0	2.7	1.6	34.0	19.2	34424.0	1.0	27.4

Experiment: rABA-1 in HEPES at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.5	745.0	3.2	1.8	49.0	24.6	18599.0	1.0	28.9
2	23.0	0.5	741.0	3.2	1.8	50.0	24.9	18754.0	1.0	23.1
3	47.0	0.5	734.0	3.2	1.8	51.0	24.7	19006.0	1.0	32.1
4	158.0	0.5	752.0	3.2	1.8	48.0	24.8	19119.0	1.0	36.4
5	180.0	0.5	768.0	3.1	1.7	45.0	24.6	18348.0	1.0	18.7

Experiment: rABA-1 in TRIS at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.5	417.0	3.3	1.8	52.0	4.8	16140.0	1.0	32.8
2	25.0	0.5	428.0	3.2	1.8	50.0	5.1	15962.0	1.0	33.6
3	52.0	0.5	390.0	3.5	2.0	62.0	4.9	15719.0	1.0	29.1
4	78.0	0.5	436.0	3.1	1.8	48.0	5.1	15557.0	1.0	41.6
5	106.0	0.5	421.0	3.3	1.8	52.0	5.1	14894.0	1.0	19.9
6	133.0	0.5	413.0	3.3	1.9	55.0	5.1	14899.0	1.0	27.5
7	161.0	0.5	437.0	3.1	1.8	48.0	5.1	14911.0	1.0	38.5

Experiment: rABA-1 in TRIS at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.5	470.0	3.3	1.9	55.0	9.5	18962.0	1.0	25.1
2	23.0	0.5	445.0	3.5	2.0	64.0	9.8	19152.0	1.0	48.9
3	46.0	0.5	500.0	3.2	1.8	48.0	9.9	17988.0	1.0	27.4
4	69.0	0.5	426.0	3.7	2.1	71.0	9.9	19399.0	1.0	30.8
5	93.0	0.5	435.0	3.6	2.0	68.0	9.9	18008.0	1.0	14.0
6	114.0	0.5	419.0	3.8	2.1	74.0	9.9	19771.0	1.0	24.1
7	139.0	0.5	507.0	3.1	0.0	47.0	9.9	17583.0	1.0	23.9
8	162.0	0.5	467.0	3.4	1.9	57.0	9.9	18073.0	1.0	48.1

Experiment: rABA-1 in TRIS at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.5	609.0	2.9	1.6	40.0	14.0	16422.0	1.0	24.4
2	26.0	0.5	587.0	3.1	1.7	45.0	14.7	17017.0	1.0	24.0
3	52.0	0.5	579.0	3.1	1.7	46.0	14.5	18129.0	1.0	31.0
4	100.0	0.5	555.0	3.3	1.8	52.0	14.8	18519.0	1.0	27.1
5	124.0	0.5	595.0	3.1	1.7	45.0	14.9	18339.0	1.0	41.3
6	146.0	0.5	560.0	3.2	1.8	51.0	14.7	18197.0	1.0	17.2
7	170.0	0.5	605.0	3.0	1.7	43.0	14.9	17761.0	1.0	28.7
8	195.0	0.5	552.0	3.3	1.8	53.0	14.8	18264.0	1.0	17.6
9	217.0	0.5	553.0	3.3	1.8	53.0	14.9	18615.0	1.0	23.1
10	242.0	0.5	447.0	4.1	2.3	89.0	14.8	21655.0	1.0	58.6

Experiment: rABA-1 in TRIS at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.5	660.0	3.1	1.7	45.0	18.8	17293.0	1.0	29.1
2	67.0	0.5	587.0	3.5	2.0	63.0	19.6	18831.0	1.0	32.5
3	90.0	0.5	606.0	3.4	1.9	58.0	19.5	18836.0	1.0	30.9
4	111.0	0.5	647.0	3.2	1.8	51.0	19.9	18557.0	1.0	26.4
5	135.0	0.5	654.0	3.2	1.8	49.0	19.8	18776.0	1.0	22.1
6	158.0	0.5	645.0	3.2	1.8	51.0	19.8	18884.0	1.0	31.4

Experiment: rABA-1 in TRIS at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	48.0	0.5	51.0	45.7	51.0	30653.0	24.2	50290.0	1.1	168.6
2	124.0	0.6	73.0	32.1	35.9	13083.0	24.3	45526.0	1.1	222.5
3	149.0	0.5	48.0	48.7	46.8	35798.0	24.2	53357.0	1.1	168.7
4	465.0	0.5	41.0	56.6	54.4	51576.0	24.2	53390.0	1.0	113.6
5	604.0	0.6	36.0	64.8	62.2	71442.0	24.0	71169.0	1.1	132.2

Experiment: rGpFAR1 in PIPES at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	47.0	28.8	16.0	10071.0	4.6	129961.0	1.0	94.8
2	55.0	0.8	42.0	32.6	18.1	13587.0	4.8	134564.0	1.0	99.4
3	110.0	0.8	49.0	27.8	15.4	9198.0	4.9	104457.0	1.0	94.9
4	165.0	0.8	60.0	22.5	12.5	5530.0	4.6	93085.0	1.1	102.9
5	180.0	0.8	60.0	22.9	12.7	5783.0	5.1	91824.0	1.0	89.3
6	196.0	0.8	61.0	22.5	12.5	5542.0	4.9	89621.0	1.0	117.8

Experiment: rGpFAR1 in PIPES at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	64.0	24.3	13.5	6625.0	8.9	131238.0	1.0	108.5
2	97.0	0.8	54.0	29.1	16.2	10266.0	9.4	140623.0	1.1	93.9
3	120.0	0.8	63.0	24.9	13.9	7070.0	9.5	118036.0	1.1	119.3
4	148.0	0.8	72.0	21.9	12.2	5166.0	9.6	103048.0	1.0	117.0
5	181.0	0.7	64.0	24.5	13.6	6763.0	9.4	111452.0	1.1	111.5
6	214.0	0.8	56.0	28.4	15.8	9704.0	9.9	126372.0	1.1	98.2

Experiment: rGpFAR1 in PIPES at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	79.0	22.7	12.6	5617.0	14.1	125080.0	1.0	92.5
2	55.0	0.8	70.0	25.6	14.2	7506.0	14.3	133717.0	1.0	114.3
3	82.0	0.8	70.0	25.8	14.3	7703.0	14.2	132885.0	1.1	108.9
4	108.0	0.8	72.0	25.0	13.9	7113.0	14.4	129394.0	1.0	117.0
5	134.0	0.8	69.0	26.1	14.5	7935.0	14.2	135463.0	1.0	105.4
6	161.0	0.8	79.0	22.9	12.7	5758.0	14.5	123336.0	1.0	89.4
7	187.0	0.8	69.0	26.3	14.6	8060.0	14.4	130620.0	1.0	102.1
8	214.0	0.7	65.0	28.0	15.5	9353.0	14.5	144952.0	1.1	90.2

Experiment: rGpFAR1 in PIPES at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.6	84.0	24.5	13.6	6802.0	19.0	151576.0	1.0	66.0
2	27.0	0.6	83.0	24.7	13.7	6932.0	19.2	148208.0	1.0	67.2
3	51.0	0.7	88.0	23.2	12.9	5952.0	19.0	129197.0	1.0	100.0
4	75.0	0.8	99.0	20.8	11.5	4544.0	19.2	113848.0	1.0	112.3
5	106.0	0.7	92.0	22.3	12.4	5421.0	19.0	124239.0	1.0	103.5
6	134.0	0.8	82.0	25.0	13.9	7090.0	19.0	133696.0	1.0	107.7

Experiment: rGpFAR1 in PIPES at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	91.0	25.7	14.3	7615.0	24.0	127199.0	1.1	98.8
2	59.0	0.8	90.0	26.1	14.5	7909.0	24.2	129564.0	1.0	102.4
3	88.0	0.8	101.0	23.2	12.9	5934.0	23.8	123569.0	1.0	93.9
4	119.0	0.8	97.0	24.3	13.5	6636.0	24.1	126821.0	1.0	101.7
5	147.0	0.8	94.0	24.7	13.7	6925.0	23.8	130727.0	1.0	106.7
6	174.0	0.8	105.0	22.3	12.4	5387.0	23.8	118220.0	1.0	99.5

Experiment: rGpFAR1 in HEPES at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	30.0	45.1	25.1	29773.0	5.1	157690.0	1.1	59.7
2	43.0	0.8	33.0	41.7	23.2	24640.0	4.9	136393.0	1.0	58.1
3	137.0	0.8	30.0	45.8	25.4	30775.0	4.9	151221.0	1.0	48.8
4	177.0	0.8	35.0	38.6	21.4	20342.0	4.9	129365.0	1.0	43.3
5	227.0	0.8	31.0	43.6	24.2	27346.0	4.8	146161.0	1.0	42.4

Experiment: rGpFAR1 in HEPES at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	37.0	42.7	23.7	25994.0	9.8	116505.0	1.0	40.9
2	57.0	0.8	32.0	48.2	26.7	34838.0	9.6	148886.0	1.1	37.8
3	121.0	0.8	39.0	40.2	22.3	22474.0	9.4	123285.0	1.0	47.3
4	175.0	0.8	31.0	50.6	28.1	39219.0	9.6	165762.0	1.1	40.6
5	228.0	0.8	39.0	40.4	22.4	22709.0	9.6	126283.0	1.0	48.3
6	276.0	0.8	42.0	37.5	20.8	19004.0	9.6	114117.0	1.0	49.3

Experiment: rGpFAR1 in HEPES at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	40.0	45.6	25.3	30458.0	14.5	153949.0	1.0	46.3
2	47.0	0.8	37.0	48.9	27.1	36113.0	14.4	178799.0	1.0	45.2
3	85.0	0.8	42.0	42.4	23.5	25581.0	14.3	150566.0	1.0	46.4
4	133.0	0.8	46.0	38.9	21.6	20797.0	14.4	145224.0	1.0	50.2
5	178.0	0.8	52.0	34.8	19.3	15904.0	14.4	145482.0	1.0	42.2

Experiment: rGpFAR1 in HEPES at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	55.0	37.3	20.7	18787.0	19.0	170832.0	1.0	55.8
2	26.0	0.7	48.0	42.7	23.7	25976.0	19.2	226458.0	1.1	33.3
3	42.0	0.8	51.0	40.4	22.4	22793.0	19.0	210991.0	1.1	40.6
4	60.0	0.7	42.0	48.1	26.7	34723.0	19.0	276419.0	1.0	31.9
5	120.0	0.8	51.0	40.1	22.2	22322.0	19.2	187071.0	1.0	55.6
6	164.0	0.8	51.0	40.5	22.5	22844.0	19.3	207829.0	1.0	44.3

Experiment: rGpFAR1 in HEPES at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	61.0	38.3	21.3	20039.0	23.7	213020.0	1.0	41.8
2	20.0	0.8	65.0	35.7	19.8	16900.0	23.9	192449.0	1.0	52.4
3	39.0	0.8	68.0	34.3	19.1	15327.0	23.9	194127.0	1.0	46.1
4	57.0	0.8	58.0	40.2	22.3	22479.0	23.9	227212.0	1.0	55.2
5	74.0	0.8	56.0	41.2	22.9	23890.0	23.8	250317.0	1.0	44.4
6	89.0	0.8	55.0	42.0	23.3	25002.0	23.8	245193.0	1.0	50.9

Experiment: rOvFAR1 in PIPES at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	135.0	10.0	5.5	775.0	4.2	266759.0	1.0	13.7
2	32.0	0.8	133.0	10.2	5.7	807.0	4.4	267575.0	1.0	16.5
3	63.0	0.8	129.0	10.5	5.9	875.0	4.5	269275.0	1.0	24.1
4	95.0	0.8	139.0	9.6	5.3	696.0	3.7	259615.0	1.0	18.0
5	128.0	0.8	130.0	9.5	5.3	688.0	1.3	275128.0	1.0	19.6
6	160.0	0.8	138.0	8.6	4.8	545.0	0.1	258332.0	1.0	19.0
7	194.0	0.8	134.0	8.9	4.9	580.0	0.0	261957.0	1.0	20.3
8	226.0	0.8	133.0	8.9	5.0	589.0	0.0	263881.0	1.0	21.4
9	259.0	0.8	131.0	9.1	5.0	611.0	0.1	265243.0	1.0	20.8
10	291.0	0.8	134.0	8.9	4.9	582.0	0.1	259479.0	1.0	18.9

Experiment: rOvFAR1 in PIPES at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	158.0	9.9	5.5	753.0	9.3	267729.0	1.0	19.5
2	32.0	0.8	161.0	9.7	5.4	714.0	9.0	263072.0	1.0	24.5
3	64.0	0.8	167.0	9.3	5.2	651.0	9.1	261390.0	1.0	16.2
4	97.0	0.8	155.0	10.0	5.6	782.0	9.1	273327.0	1.0	24.7
5	129.0	0.8	161.0	9.7	5.4	718.0	9.1	265995.0	1.0	19.7
6	160.0	0.8	166.0	9.3	5.2	656.0	8.9	261893.0	1.0	19.6
7	194.0	0.8	164.0	9.4	5.2	664.0	8.6	262332.0	1.0	19.3
8	228.0	0.7	175.0	8.8	4.7	567.0	8.7	1024058.0	1.0	9.5
9	235.0	0.8	168.0	9.1	5.1	626.0	8.6	256594.0	1.0	19.7
10	270.0	0.8	153.0	10.0	5.6	786.0	8.8	272959.0	1.0	23.8

Experiment: rOvFAR1 in PIPES at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.5	280.0	6.4	0.0	259.0	14.2	971660.0	1.0	0.7
2	7.0	0.3	374.0	4.8	0.0	132.0	14.5	958489.0	1.0	0.3
3	14.0	0.4	345.0	5.2	0.0	160.0	14.5	969954.0	1.0	0.4
4	23.0	0.5	281.0	6.4	0.0	258.0	14.2	990266.0	1.0	0.8
5	30.0	0.4	339.0	5.3	0.0	167.0	14.5	950930.0	1.0	0.7
6	37.0	0.5	292.0	6.2	0.0	241.0	14.5	958730.0	1.0	0.9
7	45.0	0.5	293.0	6.2	0.0	240.0	14.6	258252.0	1.0	0.5
8	79.0	0.3	364.0	5.0	0.0	140.0	14.4	975863.0	1.0	0.6
9	87.0	0.4	330.0	5.5	0.0	178.0	14.5	984080.0	1.0	1.0
10	94.0	0.5	298.0	6.1	0.0	228.0	14.5	249443.0	1.0	0.3

Experiment: rOvFAR1 in PIPES at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.2	493.0	4.2	0.0	93.0	19.3	868014.0	1.0	0.5
2	7.0	0.5	338.0	6.1	0.0	236.0	19.5	883173.0	1.0	0.8
3	15.0	0.4	396.0	5.2	0.0	159.0	19.4	864565.0	1.0	0.8
4	23.0	0.4	413.0	5.0	0.0	144.0	19.4	870202.0	1.0	1.2
5	31.0	0.5	351.0	5.9	0.0	211.0	19.3	882022.0	1.0	0.9
6	38.0	0.5	339.0	6.1	0.0	234.0	19.5	891368.0	1.0	0.9
7	46.0	0.5	330.0	6.3	0.0	248.0	19.5	914155.0	1.0	0.4
8	53.0	0.5	361.0	5.7	0.0	199.0	19.4	889580.0	1.0	0.7
9	61.0	0.3	396.0	5.2	0.0	158.0	19.3	888614.0	1.0	0.3
10	68.0	0.5	331.0	6.3	0.0	248.0	19.5	885645.0	1.0	1.0

Experiment: rOvFAR1 in PIPES at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	67054.0	0.0	0.0	607.0	24.5	280028.0	1.0	26.4
2	34.0	0.7	67054.0	0.0	0.0	585.0	24.4	988544.0	1.0	24.9
3	42.0	0.7	67054.0	0.0	0.0	594.0	24.4	977796.0	1.0	18.3
4	49.0	0.7	67054.0	0.0	0.0	620.0	24.5	987976.0	1.0	28.3
5	57.0	0.7	67054.0	0.0	0.0	546.0	24.5	975221.0	1.0	16.6
6	63.0	0.7	67054.0	0.0	0.0	664.0	24.4	1019631.0	1.0	21.8
7	70.0	0.7	67054.0	0.0	0.0	589.0	24.2	963747.0	1.0	21.6
8	77.0	0.8	67054.0	0.0	0.0	622.0	24.2	278533.0	1.0	30.8
9	112.0	0.7	67054.0	0.0	0.0	573.0	24.5	976922.0	1.0	16.7
10	119.0	0.7	67054.0	0.0	0.0	560.0	24.4	948908.0	1.0	20.4

Experiment: rOvFAR1 in HEPES at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	166.0	8.2	4.6	481.0	4.7	121576.0	1.0	16.8
2	29.0	0.8	177.0	7.7	4.3	416.0	4.7	111864.0	1.0	15.8
3	59.0	0.8	170.0	8.1	4.5	463.0	4.9	109843.0	1.0	17.7
4	91.0	0.8	184.0	7.5	3.7	386.0	5.1	100745.0	1.0	11.4
5	125.0	0.8	176.0	7.7	4.3	420.0	4.6	100443.0	1.0	15.1
6	139.0	0.8	182.0	7.6	3.7	396.0	4.9	97831.0	1.0	10.2
7	154.0	0.8	189.0	7.2	3.6	350.0	4.6	95174.0	1.0	12.3
8	169.0	0.8	187.0	7.3	3.6	360.0	4.6	94041.0	1.0	11.0
9	185.0	0.8	185.0	7.4	4.0	381.0	5.0	92089.0	1.0	13.8

Experiment: rOvFAR1 in HEPES at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	219.0	7.1	3.7	335.0	9.0	128472.0	1.0	12.0
2	25.0	0.8	219.0	7.2	3.5	347.0	9.4	126601.0	1.0	11.2
3	54.0	0.8	222.0	7.0	3.5	334.0	9.4	123465.0	1.0	10.3
4	82.0	0.8	211.0	7.5	4.0	389.0	9.7	125488.0	1.0	13.2
5	110.0	0.8	204.0	7.7	4.1	409.0	9.4	125926.0	1.0	13.5
6	138.0	0.8	206.0	7.7	4.2	412.0	9.7	123014.0	1.0	13.6
7	167.0	0.8	202.0	7.8	4.4	428.0	9.7	121417.0	1.0	15.4
8	196.0	0.8	196.0	8.1	4.5	465.0	9.7	122939.0	1.0	15.7
9	223.0	0.8	195.0	8.1	4.5	466.0	9.7	121731.0	1.0	20.6

Experiment: rOvFAR1 in HEPES at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	258.0	6.9	3.4	313.0	13.7	132398.0	1.0	11.6
2	26.0	0.8	259.0	6.9	3.4	321.0	14.2	132335.0	1.0	10.0
3	53.0	0.8	259.0	7.0	3.7	323.0	14.3	131830.0	1.0	13.2
4	80.0	0.8	259.0	6.9	3.4	320.0	14.2	130301.0	1.0	10.7
5	107.0	0.8	255.0	7.1	3.8	335.0	14.3	131126.0	1.0	13.7
6	133.0	0.8	258.0	7.0	3.7	326.0	14.3	129943.0	1.0	12.8
7	161.0	0.8	255.0	7.1	3.9	338.0	14.4	130771.0	1.0	14.6
8	186.0	0.8	250.0	7.3	3.9	358.0	14.5	131046.0	1.0	13.9
9	214.0	0.8	261.0	6.9	3.7	321.0	14.4	126899.0	1.0	11.8

Experiment: rOvFAR1 in HEPES at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	298.0	6.9	3.8	318.0	19.1	135601.0	1.0	14.5
2	26.0	0.8	298.0	6.9	3.7	315.0	19.0	136316.0	1.0	13.5
3	52.0	0.8	288.0	7.1	3.9	343.0	19.0	138888.0	1.0	13.9
4	78.0	0.8	246.0	8.4	4.7	506.0	19.1	157366.0	1.0	23.2
5	105.0	0.8	283.0	7.3	4.0	360.0	19.1	139593.0	1.0	14.7
6	128.0	0.8	305.0	6.8	3.3	302.0	19.2	131842.0	1.0	9.3
7	155.0	0.8	309.0	6.7	3.6	293.0	19.2	130943.0	1.0	12.7
8	182.0	0.8	307.0	6.7	3.3	296.0	19.1	131620.0	1.0	10.7
9	209.0	0.8	296.0	7.0	3.7	324.0	19.2	133587.0	1.0	13.6
10	235.0	0.8	296.0	6.9	3.8	320.0	19.0	134639.0	1.0	15.6

Experiment: rOvFAR1 in HEPES at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	332.0	7.0	3.5	330.0	23.8	132972.0	1.0	10.7
2	26.0	0.8	347.0	6.7	2.6	296.0	23.8	130103.0	1.0	8.0
3	54.0	0.8	347.0	6.7	3.3	297.0	23.8	130117.0	1.0	10.4
4	82.0	0.8	348.0	6.7	3.3	295.0	23.8	131037.0	1.0	11.1
5	108.0	0.8	352.0	6.6	3.3	287.0	23.8	130301.0	1.0	9.9
6	136.0	0.8	355.0	6.6	3.3	285.0	24.0	130634.0	1.0	10.1
7	163.0	0.8	327.0	7.2	3.8	346.0	23.9	138659.0	1.0	13.2
8	190.0	0.8	349.0	6.7	3.3	293.0	23.8	132148.0	1.0	11.4

Experiment: rOvFAR1 in TRIS at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	171.0	8.0	4.2	448.0	4.6	96855.0	1.0	12.1
2	31.0	0.8	173.0	7.9	3.9	446.0	4.9	91528.0	1.0	10.8
3	46.0	0.8	161.0	8.5	4.7	530.0	4.9	97610.0	1.0	15.0
4	61.0	0.8	160.0	8.6	4.8	547.0	5.1	97679.0	1.0	14.4
5	76.0	0.8	164.0	8.4	4.5	509.0	4.9	92883.0	1.0	14.6
6	91.0	0.8	160.0	8.6	4.7	544.0	5.0	93965.0	1.0	13.2
7	106.0	0.8	158.0	8.7	4.6	563.0	5.1	91578.0	1.0	12.9
8	122.0	0.8	161.0	8.6	4.8	539.0	5.1	90880.0	1.0	16.7
9	136.0	0.8	162.0	8.5	4.7	525.0	5.0	89624.0	1.0	15.8

Experiment: rOvFAR1 in TRIS at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	194.0	8.2	4.6	477.0	9.7	69587.0	1.0	18.1
2	17.0	0.8	182.0	8.6	4.8	544.0	9.4	74072.0	1.0	22.0
3	37.0	0.8	198.0	8.0	4.5	459.0	9.9	71800.0	1.0	15.0
4	55.0	0.8	197.0	8.0	4.3	458.0	9.7	73023.0	1.0	12.1
5	73.0	0.8	145.0	10.9	6.1	961.0	9.8	98561.0	1.0	50.3
6	89.0	0.8	168.0	9.4	5.3	677.0	9.9	86672.0	1.0	28.1
7	105.0	0.8	205.0	7.7	3.0	412.0	9.6	73447.0	1.0	8.4
8	123.0	0.8	196.0	8.0	4.0	453.0	9.4	78308.0	1.0	11.2
9	142.0	0.8	192.0	8.2	4.6	477.0	9.4	78794.0	1.0	17.6
10	161.0	0.8	187.0	8.4	4.7	507.0	9.4	81409.0	1.0	19.6

Experiment: rOvFAR1 in TRIS at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	226.0	8.0	4.4	448.0	14.2	75886.0	1.0	16.2
2	18.0	0.8	228.0	7.9	4.4	439.0	14.3	77226.0	1.0	16.9
3	36.0	0.8	222.0	8.2	4.6	478.0	14.5	82484.0	1.0	17.1
4	54.0	0.8	197.0	9.2	5.2	643.0	14.7	93953.0	1.0	30.1
5	70.0	0.8	225.0	8.0	4.5	451.0	14.2	86527.0	1.0	18.1
6	86.0	0.8	240.0	7.5	4.2	394.0	14.5	83428.0	1.0	15.8
7	104.0	0.8	239.0	7.6	3.7	402.0	14.7	85555.0	1.0	9.5
8	120.0	0.8	222.0	8.1	4.5	468.0	14.2	92444.0	1.0	18.0
9	137.0	0.8	210.0	8.7	4.8	554.0	14.7	98190.0	1.0	23.5
10	153.0	0.8	216.0	8.4	4.7	512.0	14.6	97067.0	1.0	20.4

Experiment: rOvFAR1 in TRIS at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	265.0	7.8	4.3	422.0	19.1	91111.0	1.0	15.2
2	16.0	0.8	239.0	8.6	4.8	547.0	19.2	105968.0	1.0	24.6
3	31.0	0.8	242.0	8.5	4.7	525.0	19.1	107918.0	1.0	29.2
4	62.0	0.8	253.0	8.2	4.6	476.0	19.2	107297.0	1.0	21.2
5	94.0	0.8	278.0	7.4	4.1	381.0	19.3	103790.0	1.0	14.3
6	125.0	0.8	266.0	7.8	4.4	430.0	19.5	112365.0	1.0	20.7
7	158.0	0.8	265.0	7.8	4.3	422.0	19.1	115483.0	1.0	18.7
8	188.0	0.8	272.0	7.5	4.0	393.0	19.0	116402.0	1.0	12.5
9	218.0	0.8	258.0	7.9	4.4	445.0	19.0	124521.0	1.0	22.1

Experiment: rOvFAR1 in TRIS at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	322.0	7.2	4.0	355.0	23.7	115801.0	1.0	16.6
2	29.0	0.8	328.0	7.1	3.5	343.0	23.9	119291.0	1.0	12.1
3	60.0	0.8	315.0	7.4	4.1	378.0	23.9	125486.0	1.0	15.8
4	88.0	0.8	321.0	7.3	3.6	360.0	23.8	128777.0	1.0	12.4
5	115.0	0.8	319.0	7.4	3.9	370.0	24.0	130916.0	1.0	13.2
6	141.0	0.8	294.0	7.9	4.4	444.0	23.8	139959.0	1.0	20.7
7	166.0	0.8	323.0	7.2	3.9	352.0	23.8	135957.0	1.0	14.9

Experiment: rCeFAR2 in ACETATE at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.7	10.0	146.2	0.0	513698.0	8.7	392813.0	1.0	2.2
2	717.0	0.7	10.0	140.8	39.3	468848.0	8.6	407432.0	1.0	2.1
3	1173.0	0.7	10.0	147.1	0.0	521162.0	8.6	416243.0	1.0	1.1
4	1281.0	0.7	10.0	144.6	0.0	500321.0	8.6	323910.0	1.0	0.2
5	1495.0	0.7	11.0	139.8	0.0	461265.0	8.8	360552.0	1.0	0.5
6	1958.0	0.7	10.0	146.7	33.0	517750.0	8.8	366577.0	1.0	0.7

Experiment: rCeFAR2 in ACETATE at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.7	12.0	138.1	53.9	447471.0	13.0	355915.0	1.0	5.1
2	76.0	0.7	11.0	146.8	57.3	519310.0	13.1	458864.0	1.0	4.3
3	141.0	0.7	12.0	136.5	39.4	434744.0	13.2	315694.0	1.0	2.1
4	199.0	0.7	12.0	146.1	56.2	513099.0	13.4	378154.0	1.0	4.1
5	282.0	0.7	12.0	141.4	0.0	473791.0	13.2	382815.0	1.0	1.8
6	334.0	0.7	13.0	129.7	50.6	384479.0	13.0	424050.0	1.0	3.4
7	393.0	0.7	12.0	143.0	50.3	487029.0	13.2	349545.0	1.0	3.4
8	452.0	0.6	12.0	139.9	40.4	461416.0	13.4	531168.0	1.0	1.2
9	522.0	0.7	12.0	139.8	52.5	461224.0	13.2	381484.0	1.0	4.2
10	578.0	0.8	12.0	144.6	0.0	500504.0	13.1	304759.0	1.0	2.8

Experiment: rCeFAR2 in ACETATE at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.7	14.0	133.5	52.1	412063.0	17.4	273552.0	1.0	6.5
2	86.0	0.8	15.0	129.8	37.5	385194.0	17.6	354135.0	1.0	1.9
3	168.0	0.7	13.0	142.1	55.5	479782.0	17.4	433407.0	1.0	6.2
4	233.0	0.8	14.0	133.2	35.3	409801.0	17.6	300959.0	1.0	1.3
5	315.0	0.8	16.0	117.9	0.0	305020.0	17.6	226174.0	1.0	0.6
6	459.0	0.7	15.0	130.5	0.0	390237.0	17.6	317400.0	1.0	0.8
7	598.0	0.7	13.0	143.2	50.3	488248.0	17.8	468315.0	1.0	3.4
8	668.0	0.7	15.0	129.8	0.0	384746.0	17.6	370391.0	1.0	1.1
9	715.0	0.7	15.0	129.0	0.0	379229.0	17.7	357397.0	1.0	1.6
10	773.0	0.7	15.0	127.0	36.7	364916.0	17.6	360239.0	1.0	1.9

Experiment: rCeFAR2 in ACETATE at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	14.0	156.2	61.0	602780.0	22.6	312927.0	1.1	7.2
2	174.0	0.8	16.0	134.7	38.9	421190.0	22.5	266294.0	1.0	2.9
3	214.0	0.7	16.0	134.3	47.2	418101.0	22.4	354517.0	1.1	3.5
4	285.0	0.7	17.0	132.6	64.7	405259.0	22.6	352036.0	1.1	8.2
5	358.0	0.8	17.0	129.2	50.4	380718.0	22.4	321701.0	1.1	3.6
6	429.0	0.8	18.0	125.3	48.9	353815.0	22.4	340338.0	1.0	4.3
7	482.0	0.8	18.0	122.8	35.5	336783.0	22.4	286454.0	1.0	2.4
8	556.0	0.7	17.0	127.7	28.9	369998.0	22.4	336840.0	1.0	0.6
9	633.0	0.8	19.0	116.0	0.0	293359.0	22.6	230189.0	1.0	4.5
10	686.0	0.8	17.0	127.3	49.7	367142.0	22.4	286778.0	1.0	6.0

Experiment: rCeFAR2 in TRIS at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.7	28.0	48.5	26.9	35396.0	4.7	431152.0	1.1	33.2
2	19.0	0.7	28.0	48.5	26.9	35414.0	4.5	426273.0	1.0	26.1
3	36.0	0.7	29.0	46.7	25.9	32303.0	4.5	443075.0	1.0	24.2
4	53.0	0.7	29.0	45.9	25.5	31032.0	4.5	371623.0	1.1	34.5
5	75.0	0.8	32.0	42.3	23.5	25487.0	4.5	328932.0	1.1	16.3
6	113.0	0.8	34.0	39.9	21.2	22025.0	4.4	336366.0	1.0	11.3
7	180.0	0.7	28.0	48.0	26.6	34557.0	4.1	360705.0	1.1	28.8
8	203.0	0.7	21.0	62.9	34.9	66651.0	3.8	507523.0	1.0	45.7
9	227.0	0.8	21.0	63.2	35.1	67304.0	3.7	409049.0	1.0	67.6
10	279.0	0.7	25.0	52.5	29.2	43027.0	3.4	425943.0	1.0	26.8

Experiment: rCeFAR2 in TRIS at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	35.0	40.3	22.4	22657.0	9.6	322809.0	1.0	28.2
2	21.0	0.7	28.0	48.6	27.0	35570.0	9.4	421090.0	1.1	37.5
3	43.0	0.7	24.0	56.8	31.5	51907.0	9.5	453161.0	1.0	69.1
4	70.0	0.7	29.0	48.2	26.8	34927.0	9.7	418917.0	1.0	36.1
5	90.0	0.8	28.0	49.6	27.5	37394.0	9.5	372990.0	1.0	67.1
6	114.0	0.8	33.0	42.2	23.4	25250.0	9.4	347648.0	1.1	43.3
7	137.0	0.7	29.0	48.5	26.9	35422.0	9.7	433284.0	1.1	23.8
8	158.0	0.7	30.0	45.6	25.3	30551.0	9.6	399571.0	1.0	30.2
9	180.0	0.7	27.0	50.9	28.3	39848.0	9.5	523237.0	1.1	24.7
10	229.0	0.7	21.0	66.1	36.7	75053.0	9.4	558569.0	1.1	68.6

Experiment: rCeFAR2 in TRIS at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	37.0	0.7	34.0	52.0	28.9	41943.0	13.7	428108.0	1.1	39.6
2	53.0	0.7	34.0	52.5	29.1	42939.0	13.7	426228.0	1.1	44.5
3	71.0	0.7	45.0	39.5	22.0	21614.0	13.8	333458.0	1.0	27.4
4	130.0	0.7	43.0	41.1	22.8	23704.0	13.7	340432.0	1.0	29.9
5	149.0	0.7	38.0	46.8	26.0	32456.0	13.8	389682.0	1.0	35.8
6	171.0	0.7	41.0	42.7	23.7	26031.0	13.8	371111.0	1.0	25.1
7	191.0	0.8	41.0	43.4	24.1	27049.0	13.6	344235.0	1.0	64.5
8	212.0	0.7	35.0	50.3	27.9	38798.0	13.4	423782.0	1.0	37.8
9	236.0	0.7	35.0	50.3	27.9	38678.0	13.7	447223.0	1.1	36.9
10	256.0	0.7	35.0	50.6	28.1	39258.0	13.6	437413.0	1.0	43.7

Experiment: rCeFAR2 in TRIS at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	16.0	0.8	40.0	50.4	28.0	38923.0	18.3	410766.0	1.0	52.9
2	33.0	0.7	41.0	48.8	27.1	35960.0	18.5	386823.0	1.0	50.5
3	53.0	0.8	39.0	51.3	28.5	40559.0	18.6	384510.0	1.0	49.2
4	72.0	0.7	48.0	41.8	23.2	24752.0	18.4	348355.0	1.0	19.9
5	156.0	0.7	35.0	57.3	31.8	52971.0	18.3	530927.0	1.1	31.4
6	184.0	0.7	36.0	55.4	30.8	49000.0	18.4	459827.0	1.0	32.1
7	201.0	0.7	39.0	51.7	28.7	41400.0	18.5	393378.0	1.1	56.4
8	218.0	0.7	43.0	46.7	25.9	32417.0	18.2	382310.0	1.1	46.3
9	258.0	0.8	41.0	48.9	27.1	36126.0	18.5	376225.0	1.0	30.6
10	277.0	0.8	44.0	45.3	25.1	30036.0	18.4	359345.0	1.0	33.9

Experiment: rCeFAR2 in TRIS at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coeff	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	61.0	36.9	20.5	18257.0	22.4	278043.0	1.0	42.1
2	62.0	0.7	65.0	34.3	19.0	15297.0	22.3	272567.0	1.0	27.9
3	75.0	0.7	57.0	39.5	21.9	21529.0	22.5	311947.0	1.1	28.3
4	88.0	0.8	58.0	38.6	21.5	20440.0	22.2	296181.0	1.0	25.2
5	115.0	0.7	35.0	63.9	35.5	69208.0	22.2	536684.0	1.0	39.1
6	163.0	0.7	32.0	68.6	38.1	82162.0	22.4	442392.0	1.1	105.3
7	181.0	0.7	37.0	60.8	33.7	61277.0	22.2	440840.0	1.1	59.2
8	203.0	0.8	52.0	43.3	24.0	26903.0	22.4	303888.0	1.0	41.4
9	224.0	0.8	50.0	44.5	24.7	28700.0	22.4	299850.0	1.0	52.8
10	246.0	0.8	59.0	37.8	21.0	19319.0	22.4	268240.0	1.1	29.5

Experiment: rCeFAR3 in ACETATE at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	306.0	4.4	1.7	109.0	4.6	84984.0	1.0	9.6
2	16.0	0.8	304.0	4.5	1.8	114.0	4.9	83126.0	1.0	6.6
3	32.0	0.8	297.0	4.6	1.8	120.0	4.9	82855.0	1.0	6.0
4	49.0	0.8	293.0	4.7	1.8	123.0	4.8	82256.0	1.0	5.2
5	67.0	0.8	295.0	4.6	1.6	120.0	4.7	82686.0	1.0	4.1
6	83.0	0.8	287.0	4.7	1.9	128.0	4.6	84230.0	1.0	8.1
7	100.0	0.8	300.0	4.5	1.8	114.0	4.6	81927.0	1.0	7.1
8	116.0	0.8	292.0	4.7	1.8	124.0	4.7	83117.0	1.0	9.1
9	133.0	0.8	297.0	4.6	1.8	118.0	4.6	82189.0	1.0	9.1
10	149.0	0.8	289.0	4.7	1.9	126.0	4.6	84053.0	1.0	4.5

Experiment: rCeFAR3 in ACETATE at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	329.0	4.7	1.9	126.0	9.1	83046.0	1.0	8.2
2	17.0	0.8	344.0	4.5	1.8	115.0	9.3	80388.0	1.0	7.2
3	34.0	0.8	345.0	4.5	1.8	115.0	9.4	80426.0	1.0	8.2
4	51.0	0.8	332.0	4.7	1.9	127.0	9.5	81958.0	1.0	7.6
5	69.0	0.8	350.0	4.5	1.8	111.0	9.4	79064.0	1.0	5.9
6	87.0	0.8	346.0	4.5	1.8	113.0	9.3	80164.0	1.0	4.4
7	105.0	0.8	347.0	4.5	1.8	113.0	9.3	80191.0	1.0	6.8
8	123.0	0.8	345.0	4.5	1.8	115.0	9.4	81402.0	1.0	8.1
9	139.0	0.8	342.0	4.6	1.8	117.0	9.4	80474.0	1.0	7.6
10	157.0	0.8	355.0	4.4	1.7	106.0	9.3	79792.0	1.0	5.7

Experiment: rCeFAR3 in ACETATE at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	400.0	4.5	1.8	112.0	14.2	81853.0	1.0	8.1
2	17.0	0.8	411.0	4.4	1.7	105.0	14.3	81074.0	1.0	4.8
3	35.0	0.8	415.0	4.3	1.7	103.0	14.3	80158.0	1.0	7.5
4	52.0	0.8	414.0	4.3	1.7	103.0	14.3	79749.0	1.0	5.0
5	70.0	0.8	406.0	4.4	1.7	108.0	14.3	81101.0	1.0	7.1
6	87.0	0.8	403.0	4.5	1.8	111.0	14.3	81627.0	1.0	7.1
7	105.0	0.8	409.0	4.4	1.7	106.0	14.2	81574.0	1.0	7.9
8	122.0	0.8	422.0	4.2	1.6	97.0	14.1	79772.0	1.0	4.6
9	139.0	0.8	400.0	4.5	1.8	112.0	14.2	81791.0	1.0	8.5
10	157.0	0.8	403.0	4.4	1.8	110.0	14.2	81271.0	1.0	7.6

Experiment: rCeFAR3 in ACETATE at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	484.0	4.2	1.7	97.0	19.1	81538.0	1.0	8.2
2	16.0	0.8	490.0	4.2	1.6	94.0	19.0	81495.0	1.0	7.8
3	33.0	0.8	486.0	4.2	1.7	96.0	19.0	80554.0	1.0	5.2
4	50.0	0.8	491.0	4.2	1.5	94.0	19.0	80687.0	1.0	3.7
5	67.0	0.8	476.0	4.3	1.7	102.0	19.1	82161.0	1.0	5.5
6	85.0	0.8	507.0	4.0	1.6	87.0	19.1	79567.0	1.0	7.3
7	102.0	0.8	494.0	4.1	1.6	92.0	19.0	79769.0	1.0	5.2
8	120.0	0.8	475.0	4.3	1.7	102.0	19.0	81426.0	1.0	9.1
9	137.0	0.8	481.0	4.2	1.7	99.0	19.0	79169.0	1.0	5.4
10	155.0	0.8	483.0	4.2	1.7	98.0	19.0	80074.0	1.0	5.4

Experiment: rCeFAR3 in ACETATE at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	572.0	4.0	1.6	88.0	23.7	76130.0	1.0	5.5
2	18.0	0.8	561.0	4.1	1.6	91.0	23.6	75749.0	1.0	9.0
3	53.0	0.8	559.0	4.1	1.6	93.0	23.8	75233.0	1.0	5.8
4	71.0	0.8	565.0	4.1	1.6	91.0	23.7	75018.0	1.0	6.4
5	90.0	0.8	568.0	4.1	1.6	89.0	23.7	73927.0	1.0	4.3
6	108.0	0.8	555.0	4.2	1.6	94.0	23.7	74737.0	1.0	6.7
7	127.0	0.8	560.0	4.1	1.6	92.0	23.7	73868.0	1.0	6.9
8	146.0	0.8	547.0	4.2	1.7	98.0	23.7	74269.0	1.0	5.1
9	164.0	0.8	554.0	4.2	1.6	95.0	23.7	73407.0	1.0	5.0
10	184.0	0.8	561.0	4.1	1.6	92.0	23.7	72358.0	1.0	4.3

Experiment: rCeFAR3 in PIPES at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	443.0	3.0	0.7	44.0	4.4	117738.0	1.0	1.1
2	29.0	0.8	442.0	3.0	0.7	44.0	4.5	115400.0	1.0	0.7
3	59.0	0.9	439.0	3.1	0.7	45.0	4.5	115377.0	1.0	1.0
4	89.0	0.9	432.0	3.1	0.8	46.0	4.3	114842.0	1.0	1.5
5	118.0	0.9	435.0	3.1	0.7	45.0	4.3	115165.0	1.0	1.0
6	149.0	0.8	437.0	3.1	0.7	44.0	4.2	114683.0	1.0	0.9
7	178.0	0.9	433.0	3.1	0.7	45.0	4.1	114945.0	1.0	0.9
8	208.0	0.8	431.0	3.0	0.9	43.0	3.2	115977.0	1.0	1.7
9	237.0	0.9	435.0	2.7	0.6	33.0	0.1	115004.0	1.0	0.7
10	267.0	0.9	436.0	2.7	0.6	33.0	0.0	115089.0	1.0	0.7

Experiment: rCeFAR3 in PIPES at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	517.0	3.0	0.9	42.0	9.2	110190.0	1.0	1.2
2	31.0	0.9	528.0	2.9	0.6	40.0	9.2	109764.0	1.0	0.4
3	63.0	0.8	516.0	3.0	0.6	42.0	9.2	111587.0	1.0	0.5
4	94.0	0.9	520.0	3.0	0.7	41.0	9.0	111245.0	1.0	1.1
5	124.0	0.9	527.0	2.9	0.5	40.0	9.1	111066.0	1.0	0.5
6	155.0	0.9	529.0	2.9	0.8	39.0	9.0	111672.0	1.0	1.4
7	185.0	0.8	528.0	2.9	0.8	40.0	9.2	111737.0	1.0	1.2
8	247.0	0.8	520.0	3.0	0.6	41.0	9.1	112604.0	1.0	0.4
9	277.0	0.9	526.0	2.9	0.6	40.0	9.1	112955.0	1.0	0.4
10	308.0	0.9	522.0	3.0	0.8	41.0	9.2	112655.0	1.0	1.0

Experiment: rCeFAR3 in PIPES at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	622.0	2.8	0.6	36.0	13.4	114584.0	1.0	0.9
2	29.0	0.9	629.0	2.8	0.6	36.0	13.7	114438.0	1.0	0.6
3	59.0	0.9	631.0	2.8	0.6	36.0	13.7	115435.0	1.0	0.6
4	89.0	0.8	628.0	2.8	0.6	36.0	13.7	115964.0	1.0	0.8
5	119.0	0.8	626.0	2.8	0.5	36.0	13.7	116295.0	1.0	0.5
6	149.0	0.8	618.0	2.8	0.7	37.0	13.6	116782.0	1.0	0.8
7	178.0	0.8	624.0	2.8	0.5	36.0	13.6	116063.0	1.0	0.5
8	208.0	0.9	626.0	2.8	0.0	36.0	13.7	116547.0	1.0	0.7
9	237.0	0.8	629.0	2.8	0.5	35.0	13.5	116851.0	1.0	0.4
10	267.0	0.9	625.0	2.8	0.5	36.0	13.5	117003.0	1.0	0.4

Experiment: rCeFAR3 in PIPES at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	730.0	2.7	0.6	33.0	18.1	116497.0	1.0	0.9
2	30.0	0.9	728.0	2.7	0.6	34.0	18.1	116997.0	1.0	0.7
3	59.0	0.9	736.0	2.7	0.5	33.0	18.1	116295.0	1.0	0.4
4	89.0	0.8	732.0	2.7	0.6	33.0	18.2	115623.0	1.0	0.8
5	118.0	0.9	732.0	2.7	0.0	33.0	18.2	115303.0	1.0	0.6
6	148.0	0.8	729.0	2.7	0.5	34.0	18.2	115029.0	1.0	0.5
7	178.0	0.9	737.0	2.7	0.6	33.0	18.1	114485.0	1.0	0.6
8	209.0	0.9	728.0	2.7	0.6	34.0	18.1	114454.0	1.0	0.6
9	239.0	0.8	734.0	2.7	0.7	33.0	18.3	114010.0	1.0	1.0
10	270.0	0.9	731.0	2.7	0.6	34.0	18.2	113116.0	1.0	0.5

Experiment: rCeFAR3 in PIPES at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	850.0	2.6	0.5	29.0	21.7	101726.0	1.0	0.6
2	32.0	0.8	854.0	2.6	0.0	29.0	21.9	100027.0	1.0	1.1
3	47.0	0.8	842.0	2.6	0.6	30.0	21.8	99744.0	1.0	0.9
4	80.0	0.8	822.0	2.6	0.8	31.0	21.6	98638.0	1.0	2.5
5	95.0	0.8	855.0	2.6	0.0	29.0	21.9	98131.0	1.0	0.7
6	109.0	0.9	820.0	2.7	0.6	32.0	21.7	98693.0	1.0	1.5
7	124.0	0.9	832.0	2.6	0.6	31.0	21.7	97366.0	1.0	1.6
8	139.0	0.8	855.0	2.5	0.0	29.0	21.7	96615.0	1.0	0.9
9	154.0	0.8	852.0	2.6	0.6	29.0	21.8	96371.0	1.0	1.6
10	170.0	0.8	847.0	2.6	0.0	29.0	21.7	95997.0	1.0	1.3

Experiment: rCeFAR3 in HEPES at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.6	593.0	2.3	0.0	21.0	4.7	29927.0	1.0	3.1
2	15.0	0.5	595.0	2.3	0.0	21.0	4.8	29843.0	1.0	2.0
3	31.0	184.3	67054.0	0.0	0.0	0.0	4.8	29878.0	1.0	5.2
4	47.0	0.2	848.0	1.6	0.0	9.0	4.7	30396.0	1.0	6.1
5	63.0	0.2	15901.0	0.1	0.0	0.0	4.7	29956.0	1.0	7.0
6	78.0	269.9	67054.0	0.0	0.0	0.0	5.1	30236.0	1.0	2.7
7	94.0	0.0	4905.0	0.3	0.0	0.0	4.8	29654.0	1.0	5.3
8	109.0	0.6	595.0	2.3	0.0	21.0	4.6	30159.0	1.0	3.2
9	125.0	0.0	5245.0	0.3	0.0	0.0	4.6	29738.0	1.0	9.9
10	140.0	0.6	509.0	2.7	0.0	31.0	4.7	31688.0	1.1	3.5

Experiment: rCeFAR3 in HEPES at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	30.0	0.7	548.0	2.9	0.0	39.0	10.0	29963.0	1.0	6.9
2	46.0	0.7	488.0	3.2	0.0	49.0	9.4	32248.0	1.1	18.8
3	79.0	0.7	531.0	3.0	0.0	41.0	9.7	30106.0	1.0	6.7
4	95.0	0.7	566.0	2.8	1.1	35.0	9.7	29296.0	1.0	10.2
5	172.0	0.7	543.0	2.9	0.0	39.0	9.7	29084.0	1.0	5.7
6	188.0	0.7	544.0	2.9	0.0	39.0	9.7	29113.0	1.0	5.8
7	205.0	0.7	452.0	3.5	1.7	61.0	9.7	34884.0	1.0	10.5
8	221.0	0.7	517.0	3.0	0.0	43.0	9.5	29812.0	1.0	17.7
9	271.0	0.7	568.0	2.7	0.0	34.0	9.4	28844.0	1.0	9.5
10	287.0	0.7	527.0	3.0	1.2	42.0	9.7	29816.0	1.0	4.7

Experiment: rCeFAR3 in HEPES at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	47.0	0.7	649.0	2.8	0.0	35.0	14.5	30523.0	1.0	4.8
2	62.0	0.7	647.0	2.8	0.0	36.0	14.7	30492.0	1.0	4.0
3	92.0	0.7	665.0	2.7	1.1	33.0	14.7	30086.0	1.0	10.4
4	108.0	0.7	644.0	2.8	0.0	36.0	14.6	30326.0	1.0	9.5
5	123.0	0.7	665.0	2.7	0.0	33.0	14.7	30499.0	1.0	9.2
6	139.0	0.7	667.0	2.7	1.1	33.0	14.5	30542.0	1.0	4.4
7	167.0	0.7	659.0	2.7	1.1	34.0	14.7	30160.0	1.0	8.0
8	183.0	0.7	657.0	2.7	0.0	34.0	14.6	30700.0	1.0	5.9
9	244.0	0.7	644.0	2.8	0.0	36.0	14.6	30740.0	1.0	3.3
10	259.0	0.7	630.0	2.9	0.0	38.0	14.6	30493.0	1.0	8.0

Experiment: rCeFAR3 in HEPES at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	33.0	0.7	810.0	2.5	1.0	28.0	19.4	30567.0	1.0	8.2
2	49.0	0.7	818.0	2.5	1.0	28.0	19.5	30606.0	1.0	7.1
3	64.0	0.7	775.0	2.7	0.0	32.0	19.6	30557.0	1.0	7.5
4	80.0	0.7	757.0	2.7	1.1	34.0	19.5	30358.0	1.0	6.4
5	95.0	0.7	745.0	2.8	0.0	35.0	19.4	31310.0	1.0	5.5
6	139.0	0.7	784.0	2.6	0.0	31.0	19.7	30378.0	1.0	3.9
7	153.0	0.7	734.0	2.8	1.1	37.0	19.7	30586.0	1.0	7.3
8	183.0	0.7	784.0	2.6	1.0	31.0	19.7	29961.0	1.0	5.3
9	198.0	0.7	728.0	2.8	0.0	37.0	19.4	30792.0	1.0	5.7
10	215.0	0.7	746.0	2.8	1.1	35.0	19.5	30286.0	1.0	6.1

Experiment: rCeFAR3 in HEPES at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.7	916.0	2.5	0.0	29.0	24.3	30150.0	1.0	5.5
2	15.0	0.7	879.0	2.7	1.1	32.0	24.5	29820.0	1.0	6.6
3	31.0	0.7	819.0	2.9	0.0	38.0	24.5	30352.0	1.0	5.9
4	48.0	0.7	892.0	2.6	0.0	32.0	24.7	29256.0	1.0	4.9
5	80.0	0.7	899.0	2.6	0.0	30.0	24.4	29026.0	1.0	8.6
6	98.0	0.7	859.0	2.7	1.1	34.0	24.4	29495.0	1.0	12.8
7	114.0	0.7	891.0	2.6	0.0	31.0	24.5	28777.0	1.0	6.9
8	131.0	0.7	861.0	2.7	1.1	34.0	24.3	28592.0	1.0	8.5
9	164.0	0.7	912.0	2.6	1.0	29.0	24.5	28171.0	1.0	5.2
10	209.0	0.7	851.0	2.8	1.4	35.0	24.7	28129.0	1.0	14.6

Experiment: rCeFAR3 in TRIS at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	380.0	3.6	1.0	65.0	4.7	110435.0	1.0	1.4
2	28.0	0.8	378.0	3.6	1.0	65.0	4.6	110132.0	1.0	1.5
3	59.0	0.8	376.0	3.6	1.1	68.0	4.9	109735.0	1.0	1.6
4	90.0	0.8	372.0	3.6	1.1	68.0	4.6	110151.0	1.0	1.4
5	121.0	0.8	367.0	3.7	0.8	70.0	4.6	109894.0	1.0	1.3
6	153.0	0.8	377.0	3.6	1.0	66.0	4.7	108958.0	1.0	1.5
7	185.0	0.8	371.0	3.7	0.8	70.0	4.9	109433.0	1.0	1.0
8	216.0	0.8	379.0	3.6	1.0	65.0	4.6	108410.0	1.0	1.5
9	248.0	0.8	372.0	3.6	0.8	68.0	4.6	108801.0	1.0	1.1
10	278.0	0.8	369.0	3.7	1.1	69.0	4.6	109333.0	1.0	2.3

Experiment: rCeFAR3 in TRIS at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	453.0	3.5	0.8	60.0	9.6	101091.0	1.0	1.0
2	33.0	0.8	452.0	3.4	1.0	60.0	9.4	101555.0	1.0	2.0
3	66.0	0.8	452.0	3.5	0.8	62.0	9.9	102177.0	1.0	1.3
4	100.0	0.8	448.0	3.5	1.0	62.0	9.7	102432.0	1.0	1.6
5	133.0	0.8	449.0	3.5	1.0	63.0	9.9	102512.0	1.0	1.4
6	167.0	0.8	449.0	3.5	1.0	62.0	9.7	102186.0	1.0	1.6
7	199.0	0.8	445.0	3.5	1.0	63.0	9.7	102740.0	1.0	2.0
8	233.0	0.8	449.0	3.5	0.7	63.0	9.9	103252.0	1.0	0.7
9	265.0	0.8	452.0	3.5	0.8	61.0	9.7	102668.0	1.0	1.0
10	298.0	0.8	453.0	3.5	0.7	60.0	9.7	102452.0	1.0	0.9

Experiment: rCeFAR3 in TRIS at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	532.0	3.4	0.6	57.0	14.5	101060.0	1.0	0.7
2	33.0	0.8	532.0	3.4	1.0	58.0	14.7	101495.0	1.0	1.6
3	65.0	0.8	536.0	3.4	1.0	57.0	14.7	102210.0	1.0	2.1
4	98.0	0.8	534.0	3.4	0.9	57.0	14.7	102735.0	1.0	1.1
5	130.0	0.8	537.0	3.4	1.0	56.0	14.6	103118.0	1.0	1.7
6	163.0	0.8	538.0	3.4	0.9	56.0	14.7	103366.0	1.0	1.4
7	195.0	0.8	535.0	3.4	1.0	57.0	14.7	104256.0	1.0	1.4
8	228.0	0.8	533.0	3.4	0.9	58.0	14.7	104012.0	1.0	1.5
9	262.0	0.8	533.0	3.4	0.9	57.0	14.7	104355.0	1.0	1.2
10	294.0	0.8	534.0	3.4	1.0	57.0	14.7	104474.0	1.0	1.3

Experiment: rCeFAR3 in TRIS at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	621.0	3.3	0.9	54.0	19.4	105448.0	1.0	1.1
2	31.0	0.8	627.0	3.3	0.8	54.0	19.6	104995.0	1.0	1.1
3	64.0	0.8	631.0	3.3	0.7	52.0	19.4	104771.0	1.0	0.8
4	96.0	0.8	626.0	3.3	0.8	54.0	19.6	104736.0	1.0	1.4
5	129.0	0.8	629.0	3.3	0.8	53.0	19.6	104795.0	1.0	1.1
6	162.0	0.8	626.0	3.3	0.8	54.0	19.7	104090.0	1.0	1.1
7	195.0	0.8	632.0	3.3	0.7	53.0	19.7	103594.0	1.0	1.1
8	227.0	0.8	630.0	3.3	0.8	53.0	19.5	103608.0	1.0	0.9
9	260.0	0.8	633.0	3.3	0.7	52.0	19.5	103169.0	1.0	0.9
10	292.0	0.8	632.0	3.3	0.6	53.0	19.6	102791.0	1.0	0.5

Experiment: rCeFAR3 in TRIS at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	728.0	3.2	0.9	51.0	24.3	98478.0	1.0	1.2
2	34.0	0.8	741.0	3.2	0.0	49.0	24.6	97623.0	1.0	2.3
3	48.0	0.8	727.0	3.2	0.0	51.0	24.5	95873.0	1.0	0.6
4	64.0	0.8	731.0	3.2	0.0	50.0	24.4	95249.0	1.0	1.3
5	79.0	0.8	732.0	3.2	0.9	51.0	24.5	95265.0	1.0	2.4
6	95.0	0.8	712.0	3.3	0.9	54.0	24.5	95404.0	1.0	1.8
7	110.0	0.8	724.0	3.2	0.0	52.0	24.4	94080.0	1.0	1.8
8	125.0	0.8	737.0	3.2	0.0	50.0	24.5	93039.0	1.0	1.8
9	141.0	0.8	726.0	3.2	0.0	52.0	24.5	93057.0	1.0	1.3
10	156.0	0.8	734.0	3.2	0.0	50.0	24.4	92494.0	1.0	1.8

Experiment: Lysozyme in ACETATE at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	584.0	2.3	0.0	22.0	4.6	196132.0	1.0	0.2
2	19.0	0.9	575.0	2.3	0.4	23.0	4.6	197349.0	1.0	0.4
3	37.0	0.9	561.0	2.4	0.6	25.0	5.0	197285.0	1.0	0.6
4	58.0	0.9	558.0	2.4	0.0	26.0	5.1	197594.0	1.0	0.4
5	77.0	0.9	559.0	2.4	0.0	25.0	4.8	198630.0	1.0	0.5
6	97.0	0.9	555.0	2.4	0.5	26.0	4.8	197994.0	1.0	0.6
7	117.0	0.9	549.0	2.5	0.6	27.0	4.9	198900.0	1.0	1.0
8	135.0	0.9	547.0	2.5	0.5	27.0	4.7	200009.0	1.0	0.6
9	155.0	0.9	547.0	2.5	0.6	26.0	4.6	199603.0	1.0	1.1
10	174.0	0.9	539.0	2.5	0.6	28.0	4.7	200227.0	1.0	1.1

Experiment: Lysozyme in ACETATE at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	673.0	2.3	0.0	22.0	9.3	168137.0	1.0	0.4
2	19.0	0.9	684.0	2.3	0.0	21.0	9.4	169210.0	1.0	0.6
3	41.0	0.9	685.0	2.3	0.3	22.0	9.6	171759.0	1.0	0.3
4	63.0	0.9	680.0	2.3	0.0	22.0	9.6	172482.0	1.0	0.3
5	84.0	0.9	681.0	2.3	0.0	22.0	9.5	174029.0	1.0	0.2
6	106.0	0.9	674.0	2.3	0.4	22.0	9.4	175619.0	1.0	0.4
7	128.0	0.9	687.0	2.3	0.4	22.0	9.7	176716.0	1.0	0.2
8	149.0	0.9	672.0	2.3	0.5	22.0	9.4	178008.0	1.0	1.1
9	170.0	0.9	683.0	2.3	0.4	22.0	9.4	178659.0	1.0	0.4
10	191.0	0.9	687.0	2.3	0.0	21.0	9.4	178474.0	1.0	0.3

Experiment: Lysozyme in ACETATE at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	958.0	1.8	0.3	13.0	14.1	148840.0	1.0	0.2
2	24.0	0.9	947.0	1.9	0.4	13.0	14.2	150209.0	1.0	0.4
3	48.0	0.9	963.0	1.8	0.0	13.0	14.2	150712.0	1.0	0.5
4	72.0	0.9	962.0	1.8	0.0	13.0	14.2	150684.0	1.0	0.3
5	95.0	0.9	952.0	1.9	0.3	13.0	14.3	151567.0	1.0	0.4
6	120.0	0.9	963.0	1.8	0.0	13.0	14.2	152341.0	1.0	0.2
7	144.0	0.9	960.0	1.8	0.0	13.0	14.2	152603.0	1.0	0.3
8	167.0	0.9	953.0	1.9	0.0	13.0	14.3	152644.0	1.0	0.8

Experiment: Lysozyme in ACETATE at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	1132.0	1.8	0.0	12.0	18.9	130442.0	1.0	0.2
2	26.0	0.9	1118.0	1.8	0.3	12.0	19.1	131046.0	1.0	0.4
3	53.0	0.9	1131.0	1.8	0.3	12.0	18.9	131665.0	1.0	0.3
4	79.0	0.9	1123.0	1.8	0.4	12.0	19.1	132500.0	1.0	0.7
5	106.0	0.9	1120.0	1.8	0.3	12.0	19.0	132806.0	1.0	0.3
6	132.0	0.9	1127.0	1.8	0.3	12.0	19.1	133151.0	1.0	0.5
7	159.0	0.9	1116.0	1.8	0.3	12.0	18.9	133837.0	1.0	0.3
8	185.0	0.9	1129.0	1.8	0.4	12.0	18.9	133962.0	1.0	0.4
9	211.0	0.9	1118.0	1.8	0.3	12.0	19.1	134877.0	1.0	0.4
10	237.0	0.9	1124.0	1.8	0.3	12.0	19.1	134872.0	1.0	0.4

Experiment: Lysozyme in ACETATE at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefl	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	1298.0	1.7	0.0	11.0	23.4	136213.0	1.0	0.6
2	26.0	0.9	1279.0	1.8	0.4	12.0	23.8	136975.0	1.0	0.5
3	51.0	0.9	1294.0	1.8	0.4	12.0	23.8	135653.0	1.0	0.6
4	77.0	0.9	1299.0	1.8	0.0	12.0	23.8	134984.0	1.0	0.3
5	102.0	0.9	1309.0	1.7	0.0	11.0	23.8	134839.0	1.0	0.4
6	129.0	0.9	1299.0	1.8	0.0	12.0	23.7	133939.0	1.0	0.5
7	155.0	0.9	1309.0	1.7	0.4	11.0	23.8	133647.0	1.0	0.8
8	182.0	0.9	1295.0	1.8	0.0	12.0	23.7	132748.0	1.0	0.4
9	208.0	0.9	1283.0	1.8	0.4	12.0	23.8	132678.0	1.0	0.7
10	235.0	0.9	1293.0	1.8	0.4	12.0	23.8	131593.0	1.0	0.6

Experiment: Lysozyme in PIPES at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	662.0	2.0	0.0	16.0	4.3	191883.0	1.0	0.6
2	19.0	0.9	648.0	2.1	0.0	17.0	4.4	191260.0	1.0	0.3
3	39.0	0.9	637.0	2.1	0.0	18.0	4.4	191179.0	1.0	0.3
4	59.0	0.9	641.0	2.1	0.4	18.0	4.4	191208.0	1.0	0.5
5	79.0	0.9	635.0	2.1	0.4	18.0	4.4	191113.0	1.0	0.5
6	98.0	0.9	635.0	2.1	0.0	18.0	4.5	190503.0	1.0	0.5
7	118.0	0.9	634.0	2.1	0.5	18.0	4.2	189782.0	1.0	0.8
8	139.0	0.9	631.0	2.1	0.5	18.0	4.5	189464.0	1.0	0.7
9	159.0	0.9	624.0	2.1	0.0	19.0	4.4	189096.0	1.0	0.4
10	178.0	0.9	632.0	2.1	0.3	18.0	4.4	189736.0	1.0	0.3

Experiment: Lysozyme in PIPES at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	773.0	2.0	0.0	15.0	8.8	176165.0	1.0	0.3
2	18.0	0.9	771.0	2.0	0.4	15.0	8.9	176369.0	1.0	0.5
3	40.0	0.9	776.0	2.0	0.3	15.0	9.1	176145.0	1.0	0.3
4	60.0	0.9	774.0	2.0	0.4	15.0	8.9	176466.0	1.0	0.5
5	81.0	0.9	770.0	2.0	0.4	15.0	8.8	177118.0	1.0	0.7
6	101.0	0.9	775.0	2.0	0.3	15.0	8.9	177372.0	1.0	0.3
7	122.0	0.9	770.0	2.0	0.4	15.0	8.8	177753.0	1.0	0.4
8	142.0	0.9	766.0	2.0	0.4	15.0	8.7	177768.0	1.0	0.5
9	163.0	0.9	774.0	1.9	0.4	15.0	8.6	178475.0	1.0	0.6
10	184.0	0.9	771.0	2.0	0.0	15.0	8.9	178808.0	1.0	0.3

Experiment: Lysozyme in PIPES at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	923.0	1.8	0.3	13.0	12.9	177061.0	1.0	0.2
2	20.0	0.9	917.0	1.9	0.4	14.0	13.4	177313.0	1.0	0.4
3	41.0	0.9	924.0	1.9	0.0	13.0	13.4	178016.0	1.0	0.4
4	62.0	0.9	922.0	1.9	0.0	14.0	13.5	178822.0	1.0	0.2
5	83.0	0.9	923.0	1.9	0.0	14.0	13.4	179853.0	1.0	0.6
6	104.0	0.9	916.0	1.9	0.3	14.0	13.5	182465.0	1.0	0.4
7	124.0	0.9	921.0	1.9	0.0	14.0	13.4	179441.0	1.0	0.7
8	145.0	0.9	924.0	1.9	0.3	13.0	13.4	179915.0	1.0	0.4
9	166.0	0.9	927.0	1.9	0.4	13.0	13.4	181253.0	1.0	0.6
10	186.0	0.9	928.0	1.9	0.4	13.0	13.5	180152.0	1.0	0.6

Experiment: Lysozyme in PIPES at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	1053.0	1.9	0.4	13.0	18.1	186027.0	1.0	0.9
2	20.0	0.9	1093.0	1.8	0.4	12.0	18.2	178580.0	1.0	0.6
3	41.0	0.9	1095.0	1.8	0.3	12.0	18.3	178721.0	1.0	0.4
4	61.0	0.9	1088.0	1.8	0.3	12.0	18.2	179065.0	1.0	0.5
5	83.0	0.9	1100.0	1.8	0.3	12.0	18.3	178833.0	1.0	0.4
6	103.0	0.9	1090.0	1.8	0.3	13.0	18.3	179112.0	1.0	0.3
7	124.0	0.9	1100.0	1.8	0.3	12.0	18.3	178286.0	1.0	0.5
8	145.0	0.9	1092.0	1.8	0.3	12.0	18.1	179067.0	1.0	0.6
9	165.0	0.9	1094.0	1.8	0.3	12.0	18.3	178866.0	1.0	0.4

Experiment: Lysozyme in PIPES at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	1258.0	1.8	0.0	12.0	22.7	176885.0	1.0	0.6
2	20.0	0.9	1270.0	1.7	0.3	11.0	22.7	175990.0	1.0	0.6
3	41.0	0.9	1279.0	1.8	0.0	12.0	23.1	174630.0	1.0	0.4
4	62.0	0.9	1277.0	1.7	0.3	11.0	22.7	173985.0	1.0	0.3
5	84.0	0.9	1275.0	1.8	0.3	12.0	23.1	172715.0	1.0	0.3
6	105.0	0.9	1276.0	1.7	0.0	11.0	22.9	172054.0	1.0	0.3
7	127.0	0.9	1280.0	1.7	0.3	11.0	22.9	170403.0	1.0	0.5
8	148.0	0.9	1290.0	1.7	0.0	11.0	23.1	168799.0	1.0	0.4
9	170.0	0.9	1285.0	1.7	0.3	11.0	23.0	167388.0	1.0	0.4

Experiment: Lysozyme in HEPES at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.1	670.0	1.6	0.0	14.0	4.8	201558.0	1.0	0.3
2	18.0	0.0	671.0	1.0	0.0	14.0	4.6	201934.0	1.0	0.4
3	37.0	0.0	691.0	0.4	0.0	13.0	4.9	202068.0	1.0	0.3
4	56.0	0.0	648.0	0.9	0.0	14.0	4.8	200176.0	1.0	0.6
5	74.0	0.7	651.0	3.0	0.0	14.0	4.6	198592.0	1.0	0.3
6	94.0	0.0	651.0	0.6	0.0	13.0	4.9	199474.0	1.0	0.6
7	114.0	0.0	681.0	0.3	0.0	14.0	4.9	201363.0	1.0	0.3
8	132.0	0.1	699.0	1.4	0.0	14.0	5.1	199131.0	1.0	0.3
9	152.0	0.8	659.0	3.2	0.0	14.0	4.8	202273.0	1.0	0.6
10	172.0	0.2	693.0	2.0	0.0	14.0	4.8	203439.0	1.0	0.3

Experiment: Lysozyme in HEPES at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.2	753.0	2.1	0.0	15.0	10.1	201151.0	1.0	0.2
2	21.0	0.1	773.0	1.4	0.0	15.0	9.7	200219.0	1.0	0.5
3	43.0	0.0	798.0	0.6	0.0	15.0	9.8	198387.0	1.0	0.2
4	64.0	0.2	789.0	2.3	0.0	15.0	10.1	201009.0	1.0	0.3
5	82.0	0.1	775.0	1.6	0.0	16.0	9.6	195824.0	1.0	0.4
6	100.1	0.2	748.0	2.1	0.0	15.0	10.1	196978.0	1.0	0.3
7	115.0	0.1	764.0	1.6	0.0	16.0	9.8	199383.0	1.0	0.1
8	126.0	0.0	777.0	0.2	0.0	16.0	9.7	194310.0	1.0	0.2
9	48.0	0.1	748.0	2.1	0.0	15.0	9.8	196166.0	1.0	0.5
10	162.0	0.1	746.0	2.0	0.0	15.0	9.9	195018.0	1.0	0.2

Experiment: Lysozyme in HEPES at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.8	978.0	2.6	0.0	13.0	13.7	180870.0	1.0	0.2
2	20.0	0.1	920.0	0.1	0.0	13.0	13.9	180471.0	1.0	0.3
3	41.0	0.0	938.0	0.2	0.0	13.0	13.0	180241.0	1.0	0.3
4	62.0	0.0	965.0	0.3	0.0	14.0	11.8	179618.0	1.0	0.4
5	82.0	0.0	971.0	0.7	0.0	14.0	11.2	180360.0	1.0	0.2
6	103.0	0.0	990.0	0.1	0.0	13.0	10.8	179534.0	1.0	0.2
7	124.0	76.1	954.0	0.0	0.0	14.0	10.2	179379.0	1.0	0.4
8	144.0	0.7	968.0	2.5	0.0	14.0	11.0	197467.0	1.0	0.3
9	165.0	0.6	961.0	2.3	0.0	13.0	11.0	194740.0	1.0	0.3
10	184.0	0.0	952.0	0.7	0.0	14.0	11.2	182333.0	1.0	0.1

Experiment: Lysozyme in HEPES at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.2	1100.0	1.7	0.0	12.0	19.2	175063.0	1.0	0.3
2	20.0	0.0	1037.0	0.8	0.0	13.0	19.3	172204.0	1.0	0.1
3	41.0	0.0	1011.0	0.4	0.0	12.0	19.3	170978.0	1.0	0.2
4	63.0	0.0	1020.0	0.6	0.0	13.0	19.3	172042.0	1.0	0.4
5	84.0	125.2	1054.0	0.0	0.0	13.0	19.4	170024.0	1.0	0.3
6	106.0	0.0	1080.0	0.9	0.0	13.0	19.3	171740.0	1.0	0.3
7	128.0	0.1	1020.0	0.1	0.0	12.0	19.5	170706.0	1.0	0.6
8	149.0	0.0	1064.0	0.6	0.0	12.0	19.3	171324.0	1.0	0.4
9	171.0	0.0	1069.0	0.2	0.0	13.0	19.3	171126.0	1.0	0.3
10	192.0	109.4	1054.0	0.0	0.0	13.0	19.5	171311.0	1.0	0.2

Experiment: Lysozyme in HEPES at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	41.0	0.0	1245.0	0.2	0.0	12.0	23.0	180241.0	1.0	0.3
2	62.0	0.0	1254.0	0.3	0.0	12.0	23.8	179618.0	1.0	0.4
3	82.0	0.0	1271.0	0.7	0.0	13.0	24.0	180360.0	1.0	0.2
4	103.0	0.0	1020.0	0.1	0.0	13.0	24.3	179534.0	1.0	0.2

Experiment: Lysozyme in TRIS at 5°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	684.0	2.3	0.0	21.0	4.8	169210.0	1.0	0.6
2	20.0	0.9	685.0	2.3	0.3	22.0	4.6	171759.0	1.0	0.3
3	41.0	0.9	680.0	2.3	0.0	22.0	4.8	172482.0	1.0	0.3
4	62.0	0.9	672.0	2.3	0.5	22.0	4.8	178008.0	1.0	1.1
5	82.0	0.9	683.0	2.3	0.4	22.0	4.6	178659.0	1.0	0.4
6	103.0	0.9	555.0	2.4	0.5	26.0	4.7	197994.0	1.0	0.6
7	124.0	0.9	567.0	1.8	0.0	23.0	4.9	150712.0	1.0	0.5
8	144.0	0.9	562.0	1.8	0.0	22.0	4.9	150684.0	1.0	0.3
9	165.0	0.9	652.0	1.9	0.3	23.0	4.8	151567.0	1.0	0.4
10	184.0	0.9	639.0	2.5	0.6	28.0	4.8	200227.0	1.0	1.1

Experiment: Lysozyme in TRIS at 10°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	555.0	2.4	0.5	26.0	9.3	172482.0	1.0	0.4
2	19.0	0.9	663.0	1.8	0.0	23.0	9.2	178008.0	1.0	0.6
3	41.0	0.9	662.0	1.8	0.0	22.0	9.6	150209.0	1.0	0.3
4	63.0	0.9	680.0	2.3	0.0	22.0	9.6	150712.0	1.0	0.4
5	84.0	0.9	648.0	2.1	0.0	22.0	9.5	174029.0	1.0	0.5
6	106.0	0.9	676.0	2.0	0.0	21.0	9.4	175619.0	1.0	0.4
7	128.0	0.9	684.0	2.3	0.0	21.0	9.7	176716.0	1.0	0.2
8	149.0	0.9	685.0	2.3	0.3	22.0	9.6	152603.0	1.0	0.6
9	170.0	0.9	680.0	2.3	0.0	22.0	9.4	178659.0	1.0	0.3
10	191.0	0.9	681.0	2.3	0.0	22.0	9.4	178474.0	1.0	0.6

Experiment: Lysozyme in TRIS at 15°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	958.0	1.8	0.3	13.0	14.1	148840.0	1.0	0.2
2	24.0	0.9	947.0	1.9	0.4	13.0	14.2	150209.0	1.0	0.4
3	48.0	0.9	963.0	1.8	0.0	13.0	14.2	150712.0	1.0	0.5
4	72.0	0.9	962.0	1.8	0.0	13.0	14.2	150684.0	1.0	0.3
5	95.0	0.9	952.0	1.9	0.3	13.0	14.3	151567.0	1.0	0.4
6	120.0	0.9	963.0	1.8	0.0	13.0	14.2	152341.0	1.0	0.2
7	144.0	0.9	960.0	1.8	0.0	13.0	14.2	152603.0	1.0	0.3
8	167.0	0.9	953.0	1.9	0.0	13.0	14.3	152644.0	1.0	0.8

Experiment: Lysozyme in TRIS at 20°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	1132.0	1.8	0.0	12.0	18.9	160442.0	1.0	0.2
2	26.0	0.9	1100.0	1.8	0.3	12.0	18.3	178833.0	1.0	0.4
3	53.0	0.9	1090.0	1.8	0.3	13.0	18.3	179112.0	1.0	0.3
4	79.0	0.9	1100.0	1.8	0.3	12.0	18.3	178286.0	1.0	0.5
5	106.0	0.9	1120.0	1.8	0.3	12.0	19.0	152806.0	1.0	0.3
6	132.0	0.9	1127.0	1.8	0.3	12.0	19.1	163151.0	1.0	0.5
7	159.0	0.8	1053.0	1.9	0.4	13.0	18.1	186027.0	1.0	0.9
8	185.0	0.9	1093.0	1.8	0.4	12.0	18.2	178580.0	1.0	0.6
9	211.0	0.9	1118.0	1.8	0.3	12.0	19.1	154877.0	1.0	0.4
10	237.0	0.9	1124.0	1.8	0.3	12.0	19.1	154872.0	1.0	0.4

Experiment: Lysozyme in TRIS at 25°C

Meas. #	Time (s)	Ampl ( )	Diffn Coefi	Radius (nm)	Polyd (nm)	Estd MW	Temp (°C)	Count (cnts/s)	Baseline	SOS error
1	0.0	0.9	1298.0	1.7	0.0	11.0	23.4	136213.0	1.0	0.6
2	26.0	0.9	1297.0	1.8	0.3	12.0	23.8	131279.0	1.0	0.3
3	38.0	0.9	1296.0	1.8	0.3	12.0	23.7	130884.0	1.0	0.5
4	53.0	0.9	1299.0	1.8	0.3	12.0	23.8	130466.0	1.0	0.5
5	72.0	0.9	1309.0	1.7	0.0	11.0	23.8	134839.0	1.0	0.4
6	96.0	0.9	1299.0	1.8	0.0	12.0	23.7	133939.0	1.0	0.5
7	106.0	0.9	1297.0	1.8	0.3	12.0	23.8	131279.0	1.0	0.3
8	118.0	0.9	1296.0	1.8	0.3	12.0	23.7	130884.0	1.0	0.5
9	125.0	0.9	1299.0	1.8	0.3	12.0	23.8	130466.0	1.0	0.5
10	169.0	0.9	1293.0	1.8	0.4	12.0	23.8	131593.0	1.0	0.6

Appendix 2

2.1 ABA-1 Protein parameters

rABA-1 sequence:

```
    1       11       21       31       41       51
    |       |        |        |        |        |
  1 HHFTLESSLD THLKWLSQEQ KDELLKMKKD GKAKKELEAK
ILHYYDELEG DAKKEATEHL   60
 61 KGGCREILKH VVGEEKAAEL KNLKDSGASK EELKAKVEEA
LHAVTDEEKK QYIADFGPAC  120
121 KKIYGVHTSR RRR
```

Number of amino acids: 133

Molecular weight: 15267.3

Theoretical pI: 7.92

Amino acid composition:

Ala (A)	12	9.0%
Arg (R)	5	3.8%
Asn (N)	1	0.8%
Asp (D)	8	6.0%
Cys (C)	2	1.5%
Gln (Q)	3	2.3%
Glu (E)	19	14.3%
Gly (G)	8	6.0%
His (H)	8	6.0%
Ile (I)	4	3.0%
Leu (L)	15	11.3%
Lys (K)	23	17.3%
Met (M)	1	0.8%
Phe (F)	2	1.5%
Pro (P)	1	0.8%
Ser (S)	6	4.5%
Thr (T)	5	3.8%
Trp (W)	1	0.8%
Tyr (Y)	4	3.0%
Val (V)	5	3.8%
Asx (B)	0	0.0%
Glx (Z)	0	0.0%
Xaa (X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 27
Total number of positively charged residues (Arg + Lys): 28

Atomic composition:

Carbon	C	672
Hydrogen	H	1090
Nitrogen	N	192
Oxygen	O	207
Sulfur	S	3

Formula: $C_{672}H_{1090}N_{192}O_{207}S_3$
Total number of atoms: 2164

Extinction coefficients:

Conditions: 6.0 M guanidium hydrochloride, 0.02 M phosphate buffer, pH 6.5

Extinction coefficients are in units of M^{-1} cm^{-1} .

The first table lists values computed assuming ALL Cys residues appear as half cystines, whereas the second table assumes that NONE do.

	276 nm	278 nm	279 nm	280 nm	282 nm
Ext. coefficient	11345	11327	11160	10930	10520
Abs 0.1% (=1 g/l)	0.743	0.742	0.731	0.716	0.689

	276 nm	278 nm	279 nm	280 nm	282 nm
Ext. coefficient	11200	11200	11040	10810	10400
Abs 0.1% (=1 g/l)	0.734	0.734	0.723	0.708	0.681

Estimated half-life:

The N-terminal of the sequence considered is H (His).

The estimated half-life is: 3.5 hours (mammalian reticulocytes, in vitro).
10 min (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 49.83
This classifies the protein as unstable.

Aliphatic index: 75.64
Grand average of hydropathicity (GRAVY): -1.018

2.2 OvFAR1 Protein parameters

rOvFAR1 sequence:

```
    1       11       21       31       41       51
    |       |        |        |        |        |
   1 MYHQLILMAL IGVIMANVVP FSMSNIPEEY KEFIPEEVKN
FYKNLTQEDR QILRELASKH   60
  61 ATFTNEDAAL EALKNKSDKL YQKAVELRNF VKAKIDSLKP
DAKAFVDEII AKVRSLRPED  120
 121 GQKLDMEKLK QAARDIIAKY EALNEETKEE LKATFPNTTK
IITNEKFKRI ANSFLQKN
```

Number of amino acids: 178

Molecular weight: 20580.8

Theoretical pI: 8.58

Amino acid composition:

Ala (A)	18	10.1%
Arg (R)	7	3.9%
Asn (N)	12	6.7%
Asp (D)	9	5.1%
Cys (C)	0	0.0%
Gln (Q)	7	3.9%
Glu (E)	19	10.7%
Gly (G)	2	1.1%
His (H)	2	1.1%
Ile (I)	14	7.9%
Leu (L)	17	9.6%
Lys (K)	23	12.9%
Met (M)	5	2.8%
Phe (F)	9	5.1%
Pro (P)	6	3.4%
Ser (S)	7	3.9%
Thr (T)	8	4.5%
Trp (W)	0	0.0%
Tyr (Y)	5	2.8%
Val (V)	8	4.5%
Asx (B)	0	0.0%
Glx (Z)	0	0.0%
Xaa (X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 28
Total number of positively charged residues (Arg + Lys): 30
Atomic composition:

Carbon	C	924
Hydrogen	H	1495
Nitrogen	N	245
Oxygen	O	274
Sulfur	S	5

Formula: $C_{924}H_{1495}N_{245}O_{274}S_5$
Total number of atoms: 2943

Extinction coefficients:

Conditions: 6.0 M guanidium hydrochloride, 0.02 M phosphate buffer, pH 6.5

Extinction coefficients are in units of M^{-1} cm^{-1} .

	276 nm	278 nm	279 nm	280 nm	282 nm
Ext. coefficient	7250	7000	6725	6400	6000
Abs 0.1% (=1 g/l)	0.352	0.340	0.327	0.311	0.292

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 45.01
This classifies the protein as unstable.

Aliphatic index: 91.07

Grand average of hydropathicity (GRAVY): -0.516

2.3 CeFAR2 protein parameters

rCeFAR2 sequence:

```
    1       11       21       31       41       51
    |       |        |        |        |        |
   1 MIRAFLVVAL ASVAVFSAPI PEVPQNFDDI PAEYKGLIPA
EVAEHLKAIT AEEKAALKEL   60
  61 AQNHKEYKTE EEFKAALKEK SPSLYEKAGK LEALLTAKFE
KLDATAQALV KKIIAKGREL  120
 121 HQQYLAGDKP TLDSLKELAK GYIAEYKALS DDAKATITAE
FPILTGFFQN EKIQAIVGQY  180
 181 VN
```

Number of amino acids: 182

Molecular weight: 20035.1

Theoretical pI: 5.73

Amino acid composition:

Ala (A)	30	16.5%
Arg (R)	2	1.1%
Asn (N)	4	2.2%
Asp (D)	7	3.8%
Cys (C)	0	0.0%
Gln (Q)	8	4.4%
Glu (E)	20	11.0%
Gly (G)	7	3.8%
His (H)	3	1.6%
Ile (I)	12	6.6%
Leu (L)	20	11.0%
Lys (K)	22	12.1%
Met (M)	1	0.5%
Phe (F)	8	4.4%
Pro (P)	8	4.4%
Ser (S)	6	3.3%
Thr (T)	8	4.4%
Trp (W)	0	0.0%
Tyr (Y)	7	3.8%
Val (V)	9	4.9%
Asx (B)	0	0.0%
Glx (Z)	0	0.0%
Xaa (X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 27
Total number of positively charged residues (Arg + Lys): 24

Atomic composition:

Carbon	C	917
Hydrogen	H	1464
Nitrogen	N	228
Oxygen	O	270
Sulfur	S	1

Formula: $C_{917}H_{1464}N_{228}O_{270}S_1$
Total number of atoms: 2880

Extinction coefficients:

Conditions: 6.0 M guanidium hydrochloride, 0.02 M phosphate buffer, pH 6.5

Extinction coefficients are in units of M^{-1} cm^{-1} .

	276 nm	278 nm	279 nm	280 nm	282 nm
Ext. coefficient	10150	9800	9415	8960	8400
Abs 0.1% (=1 g/l)	0.507	0.489	0.470	0.447	0.419

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 20.02
This classifies the protein as stable.

Aliphatic index: 99.40

Grand average of hydropathicity (GRAVY): -0.164

2.4 CeFAR3 protein parameters

rCeFAR3 sequence:

```
    1       11       21       31       41       51
    |       |        |        |        |        |
  1 MSRLFAFNVF CLVLLRFSAA APADDSSPFS QILKQHKDLL
PSEVVQAYQD LSPEEKAALK   60
 61 DVFKNYKSYK NEGELIAALK EKSSSLGEKA EKLQAKLQKK
VDALSPKPKD FVNELIAGGR  120
121 GLYARSVNGE KISVSEIKLL IETQVAAYKA LPAEAQDELK
KNFGGVAKFL EDDKTQTLIA  180
181 KLLEKNNNQ
```

Number of amino acids: 189

Molecular weight: 20905.0

Theoretical pI: 8.76

Amino acid composition:

Ala (A)	22	11.6%
Arg (R)	4	2.1%
Asn (N)	9	4.8%
Asp (D)	10	5.3%
Cys (C)	1	0.5%
Gln (Q)	10	5.3%
Glu (E)	16	8.5%
Gly (G)	8	4.2%
His (H)	1	0.5%
Ile (I)	7	3.7%
Leu (L)	25	13.2%
Lys (K)	25	13.2%
Met (M)	1	0.5%
Phe (F)	9	4.8%
Pro (P)	7	3.7%
Ser (S)	15	7.9%
Thr (T)	3	1.6%
Trp (W)	0	0.0%
Tyr (Y)	5	2.6%
Val (V)	11	5.8%
Asx (B)	0	0.0%
Glx (Z)	0	0.0%
Xaa (X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 26
Total number of positively charged residues (Arg + Lys): 29

Atomic composition:

Carbon	C	941
Hydrogen	H	1523
Nitrogen	N	247
Oxygen	O	284
Sulfur	S	2

Formula: $C_{941}H_{1523}N_{247}O_{284}S_2$
Total number of atoms: 2997

Extinction coefficients:

Conditions: 6.0 M guanidium hydrochloride, 0.02 M phosphate buffer, pH 6.5

Extinction coefficients are in units of M^{-1} cm^{-1} .

The first table lists values computed assuming ALL Cys residues appear as half cystines, whereas the second table assumes that NONE do.

	276 nm	278 nm	279 nm	280 nm	282 nm
Ext. coefficient	7250	7000	6725	6400	6000
Abs 0.1% (=1 g/l)	0.347	0.335	0.322	0.306	0.287

	276 nm	278 nm	279 nm	280 nm	282 nm
Ext. coefficient	7250	7000	6725	6400	6000
Abs 0.1% (=1 g/l)	0.347	0.335	0.322	0.306	0.287

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 44.41
This classifies the protein as unstable.

Aliphatic index: 94.55
Grand average of hydropathicity (GRAVY): -0.367

2.5 GpFAR1 protein parameters

rGpFAR1 (with leader sequence) sequence:

```
    1       11       21       31       41       51
    |       |        |        |        |        |
   1 MQRILLCLTG ASFIVLLFGA SLPPIDISSI PEQYRELIPK
EVIDFYNTLT AEDKQALKEV   60
  61 AERHEEFQTE EQAMEALKAK SEKLHSKAVE LRNLVKEKID
KLVPGAKTFV TETIEKLKAM  120
 121 RPKSGEKPNL EELRKGANDT IEKFKALSVE AKESLKANFP
KITGVIQSEK FQALAKSLLK  180
 181 TEGAAPAA
```

Number of amino acids: 188

Molecular weight: 20922.3

Theoretical pI: 7.75

Amino acid composition:

Ala (A)	21	11.2%
Arg (R)	6	3.2%
Asn (N)	5	2.7%
Asp (D)	5	2.7%
Cys (C)	1	0.5%
Gln (Q)	7	3.7%
Glu (E)	24	12.8%
Gly (G)	7	3.7%
His (H)	2	1.1%
Ile (I)	12	6.4%
Leu (L)	22	11.7%
Lys (K)	24	12.8%
Met (M)	3	1.6%
Phe (F)	8	4.3%
Pro (P)	9	4.8%
Ser (S)	11	5.9%
Thr (T)	10	5.3%
Trp (W)	0	0.0%
Tyr (Y)	2	1.1%
Val (V)	9	4.8%
Asx (B)	0	0.0%
Glx (Z)	0	0.0%
Xaa (X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 29
Total number of positively charged residues (Arg + Lys): 30

Atomic composition:

Carbon	C	939
Hydrogen	H	1546
Nitrogen	N	246
Oxygen	O	282
Sulfur	S	4

Formula: $C_{939}H_{1546}N_{246}O_{282}S_4$
Total number of atoms: 3017

Extinction coefficients:

Conditions: 6.0 M guanidium hydrochloride, 0.02 M phosphate buffer, pH 6.5

Extinction coefficients are in units of M^{-1} cm^{-1} .

The first table lists values computed assuming ALL Cys residues appear as half cystines, whereas the second table assumes that NONE do.

	276 nm	278 nm	279 nm	280 nm	282 nm
Ext. coefficient	2900	2800	2690	2560	2400
Abs 0.1% (=1 g/l)	0.139	0.134	0.129	0.122	0.115

	276 nm	278 nm	279 nm	280 nm	282 nm
Ext. coefficient	2900	2800	2690	2560	2400
Abs 0.1% (=1 g/l)	0.139	0.134	0.129	0.122	0.115

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 33.38
This classifies the protein as stable.

Aliphatic index: 95.59
Grand average of hydropathicity (GRAVY): -0.331