Understanding novel protein-derived biopolymers to enable biomimetic devices

Valeria Italia

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE GRENOBLE ALPES

préparée dans le cadre d'une cotutelle entre l'*Université Grenoble Alpes* et *Swansea University*

Spécialité : **Modèles, méthodes et algorithmes en biologie, santé et environnement**

Arrêté ministériel : le 6 janvier 2005 – 25 mai 2016

Présentée par

Valeria ITALIA

Thèse dirigée par professeur **Donald Martin,** et Professor **Kenith Meissner,**
Et codirigée par dr **Marco MACCARINI** et Professor **Paolo BERTONCELLO**

préparée au sein des **Laboratoire Techniques de L'Ingénierie Médicale et de la Complexité - Informatique, Mathématiques et Applications et Faculty of Science and Engineering laboratories.**

dans **les Écoles Doctorales Ingénierie pour la Santé la Cognition et l'Environnement et Faculty of Science and Engineering et School of Science.**

Comprendre les nouveaux biopolymères dérivés de protéines pour activer les dispositifs biomimétiques

Thèse soutenue publiquement le **14 Décembre 2021**,
devant le jury composé de :

Madame Judith PETERS
Professeur, Université Grenoble Alpes, Examinatrice, Président du jury
Madame, Francesca BALDELLI BOMBELLI
Associate Professor, Politecnico di Milano, Rapportrice
Monsieur, Richard, CAMPBELL
Senior Lecturer, University of Manchester, Rapporteur
Madame, My HEDHAMMAR
Full Professor, KTH Royal Institute of Technology, Examinatrice
Madame, Sophie SHERMER
Associate Professor, Swansea University, Examinatrice

CONTENTS

Acronyms iv

1. INTRODUCTION 1

2. ULTRABITHORAX PROTEIN 4
2.1 Homeodomain Transcription Factors 4
2.2 Protein-Based Materials 6
2.2.1 Silk material 6
2.2.2 Collagen material 7
2.2.3 Ultrabithorax material 8
2.3 Ultrabithorax fusions 9

3. METHODS 11
3.1 Langmuir trough Technique 11
3.2 Brewster Angle Microscopy 15
3.3 Ellipsometry 18
3.3.1 Ellipsometry applied to ultra-thin films 19
3.4 Scanning Electron Microscopy 21
3.5 Neutron Reflectivity 26
3.5.1 Reflectivity from a single layer 27
3.5.2 Reflectivity from multiple layers 30
3.5.3 FIGARO 34

4. PROTEIN PRODUCTION RESULTS 37
4.1 Production in Swansea laboratories 37
4.1.1 Bacteria growth and protein production 37
4.1.2 Protein Purification 40
4.1.3 SDS-PAGE gel 40
4.1.4 Concentration check 43
4.2 Demonstration of industrial-scale production of Ubx 46
4.3 Production in Grenoble laboratories 47

5. MACROSCOPIC RESULTS 51
5.1 Langmuir results and discussion 52
5.1.1 Limiting Area analysis 59
5.1.2 Fiber collection 60
5.1.3 Matlab data analysis 62
5.1.4 Metal Disk coating 66
5.1.5 Antimicrobial peptide coating 69
5.2 BAM results and discussion 73

6. MOLECULAR RESULTS 77
6.1 Ellipsometry Data 77
6.1.1 Optimization of spread films 77
6.1.2 NR preparation 81
6.2 Neutron reflectometry data 85
6.2.1 Optimisation of the samples for NR experiments 86
6.2.2 Experiments 88

7. CONCLUSIONS & FUTURE WORKS 94

ACMW air contrast matched water.

AMP ampicillin.

BAM brewster angle microscopy.

BEs backscattered high-energy electrons.

BSA bovine serum albumin.

CLAMF chloramphenicol.

DNA deoxyribonucleic acid.

DTT DithioThreitol.

EGFP Enhanced Green Fluorescent Protein.

FIGARO Fluid Interfaces Grazing Angles ReflectOmeter.

Hox homeodomain.

ILL Institut Laue-Langevin.

LB Langmuir-Blodgett.

MaSp major ampullate spidroin.

m-RNA messenger ribonucleic acid.

NOEs nuclear Overhauser effects.

NR neutron reflectometry.

OD optical density.

PEs primary electrons.

SDS-PAGE Sodium Dodecyl Sulfate - PolyAcrylamide Gel Electrophoresis.

SEM scanning electron microscopy.

SEs secondary electrons.

SLD scattering length density.

TOF time of flight.

TR transcription factor.

Ubx Ultrabithorax.

LIST OF TABLES

4.1 Optical Density check during *E. coli* cell culture. OD provides information on the growth of bacteria overtime. 39
4.2 Concentration values for the purified fractions. 44
4.3 Table. 45
4.4 Optical Density check during *E. coli* cell culture. 46
4.5 Optical Density check during *E. coli* cell culture. 48

5.1 Limiting Area values $[cm^2]$ obtained by the extrapolation of the linear part of the $\pi - A$ curve to zero pressure. 59
5.2 Mean values of diameters, d, [μm] for the fibers collected with Langmuir trough at different waiting times, t, and subphases, 0.1 M NaCl and G0 buffer. Missing values are reported with the symbol '-', if it was impossible to see any. 62
5.3 Mean values of FWHM [μm] for the fibers collected with Langmuir trough at different waiting time and solutions. Missing values are reported with the symbol '-', if it was impossible to see any, with '+' if it was possible to see some which were too weak. 65

LIST OF FIGURES

2.1 Area of expression in embryo and adult life stages are shown. Reproduced from reference [62]. 5

2.2 Hierarchical structures in fibrillar collagen showing a characteristic periodic structure forming various hierarchical orders of association: fibers, fibrils and microfibrils with three helical chains. Collagen fiber diameter is from 1 to 20 μm. 7

2.3 TEM images showing the hierarchical film formation of EGFP (Enhanced Green Fluorescent Protein)-Ubx (Ultrabithorax) at the air water interface. (a)- Ubx aggregation at 15 minutes after dilution; (b)- protofibrils formation after 1h in dilution; (c)- parallel alignment of protofibrils after 1-2h; (d)- fibril formation of 50 nm diameter after 2h; (e)- lateral association of fibrils into film; (f)- Ubx film drawn into fiber. Reproduced from reference [80]. 8

2.4 Schematic gene sequence of Ubx fusion proteins used to create functionalised materials. The functional protein sequence (green, purple and red) are shown linked to flexible linker, in yellow, followed by the self-assembling protein Ubx in brown. . 9

3.1 Langmuir trough scheme. A, bath where the subphase is poured; B moving barriers to control the surface area dimension; C surface-pressure balance with Wilhelmy plate measuring the surface pressure. 12

3.2 Phospholipid of the plasma membrane with a hydrophilic head and two hydrophobic tails. 13

3.3 Scheme of Wilhelmy plate partially immersed in a water subphase where t is the plate thickness, θ is the contact angle, w is the plate width, h is the immersed depth and l is the plate length. In the inset there's the scheme of the forces acting on subphase molecules at the air-water interface. 13

3.4 Scheme of surface pressure (π) vs. area (A) and orientation of surfactants in different phases. 15

3.5 a - Schematic plane-view and end-view representation of wave nature of light . b - Light beam incident upon and reflected from a surface showing the plane of incidence and the s-direction and p-direction. 15
3.6 a - Optical setup of BAM. b - Reflection coefficients vs. incidence angle for s- and p- polarized light at air-glass interface. $n_a = 1$ and $n_g = 2$. 16
3.7 Brewster Angle Microscopy physical functioning. 17
3.8 Example of BAM image 17
3.9 Schema of ellipsometric measure. 19
3.10 Schematic diagram of a scanning electron microscope. 23
3.11 a - Schematic diagram of an electromagnetic lens. The coil windings are inside the iron shroud and the field is produced across the lens gap between polepieces. b(i) - Diagram showing how electrons are focussed pass trough an electromagnetic lens. b(ii) - Conventional optical ray diagram for comparison. 24
3.12 Interaction volume of an electron beam with a thick sample. . 25
3.13 Scattering length for different elements. The solid squares indicate values for isotopes.Reproduced from reference [98]. . 27
3.14 (a) - Reflection of an incident beam from an ideally flat interface. ki and kr are the incident and scattered wave vectors, with angles $\theta_i = \theta_r = \theta$ in the incidence plane; q is the wave vector transfer; and SLD (scattering length density) is the scattering length density of the semi-infinite substrate. (b) - Calculated neutron reflectivity profile for a beam incident on a silicon substrate where the surface is perfectly smooth. Reproduced from reference [108].$k_{z,0}$ is equal to the magnitude of the incident (and reflected) wavevector in air. 28
3.15 (a) - Reflection of an incident beam from two ideally flat interfaces, one is a thin layer of thickness d on a semi-infinite substrate (with scattering length density Nb_2 and Nb_1 for the film and substrate respectively). (b) - Calculated neutron reflectivity profile for a 500 Åthick layer of perdeuterated poly(styrene), d-PS, on a silicon substrate. The oscillations are due to the thickness of the specimen. Reproduced from reference [108]. 31
3.16 Diagram of a n-layers specimen with variable thicknesses (d_n) on a substrate. The index n + 1 corresponds to the substrate, n to the layer immediately adjacent to the substrate, down to index 0 which corresponds to the vacuum/air. 32

3.17 Three surface profiles: on the left sharp interface, at the centre a rough surface, on the right a diffuse surface. The rough and diffuse surfaces can result in identical specular reflectivities. . 33
3.18 Sketch of the reflectometer FIGARO. 34

4.1 Transformation technique scheme: using the heat shock technique, the cell membrane is made permeable to plasmid DNA which can enter into *E. coli* cell. 38
4.2 (a) - *E. coli* colonies on LB agar plate. (b) - 400 mL of autoclaved LB with 0.4 mL of $50\mu g/mL$ AMP and 0.4 mL of $30\mu g/mL$ CLAMF (c) - Harvested bacteria cells using 30 minutes of centrifugation at 3400 rpm at 4°C. 38
4.3 Column for protein purification with 10 mL nickel-nitrilotriacetic acid resine. 40
4.4 SDS-PAGE of protein samples scheme. Proteins are loaded into gel walls and electrical current is applied. Proteins run through the gel and are separated by size. 41
4.5 (a) gel containing standard, the crude and all the elution fractions (E0, E1, E2, E3, E4, E5, E6); (b) gel containing the standard, the pellet, the crude, the unbound proteins, fractions E0, E4 and E5 and the Ni-NTA column matrix. The red rectangles underline EGFP-Ubx presence or absence. 42
4.6 Absolute spectra of 1/20 dilution measured for all the fractions collected during protein purification. 43
4.7 Absorbance-concentration relationship for BSA proteins. . . . 45
4.8 Fermenter parameters during *E. coli* cell growth. Yellow line indicates the temperature, the red one indicates the pH and the pink the air flow. 47
4.9 Absorbance-concentration relationship for BSA proteins. . . . 49
4.10 Gel containing the crude, the unbound proteins, three washes W1, W2, W3, fractions E1, E2, BSA at three concentrations 1 mg/mL, 2 mg/mL, 3 mg/mL and the standard. 50

5.1 Steps of EGFP-Ubx protein-based film formation. (a) - EGFP-Ubx is diluted in a Teflon-coated tray filled with G0 buffer; (b) - after a 16 h incubation period under an aluminium pan, with a microcentrifuge tube rack EGFP-Ubx film is gently squeezed reducing its area; (c) - the aluminium pan is placed again on top of the tray leaving the film to set for an hour; (d) - a plastic inoculation loop is used to touch the surface and pull fibers (e). 51

5.2 EGFP-Ubx droplet disposition schema. Droplets are dripped at the surface gently in a symmetric way: three droplets per side, two droplets are placed near the moving barriers and one is placed near the Wilhelmy plate. 53
5.3 Isothermal compression curve measured with 10-minute waiting time at different concentrations: 1nmol, 2nmol, 3nmol and 4nmol. The subphase used was plain water (a) and 0.1 M NaCl (b). 54
5.4 Isothermal compression curve measured with 10-minute waiting time spreading 2nmol EGFP-Ubx. (a) - the subphase used has and increasing salt concentration: water, 0.1 M, 0.2 M and 0.3 M NaCl; (b) - the subphase used has different salts diluted: 0.1 M NaCl, 0.1 M KCl, 0.01 M PBS and 0.1 M NaCl + 0.05 M NaH_2PO_4. 56
5.5 Isothermal compression curve measured varying the waiting time parameter (10 minutes, 60 minutes and 120 minutes) and spreading 2nmol EGFP-Ubx with (a) - 0.1 M NaCl; (b) - 0.1 M NaCl + 0.05 M NaH_2PO_4. 58
5.6 Representation of limiting area projection. 59
5.7 Fibers collected using 2 nmol EGFP-Ubx proteins spread in 0.1 M NaCl buffer using (a) - 10 minutes waiting time and (b) - 120 minutes waiting time. It was not possible to collect fibers from 60 minutes waiting time. 61
5.8 Fiber images obtained using Langmuir trough with 2 nmol Ubx in G0 buffer and using (a) - 10 minutes waiting time, (b) - 60 minutes waiting time and (c) - 120 minutes waiting time. All the images have been taken using a Scanning Electron Microscope with 1 kV voltage. 61
5.9 Fiber flattening diagram. The view from the SEM (A) can be flattened by assuming a cylindrical shape and calculating arc lengths. The blacked-out regions are outside of the $\theta = 60°$threshold. 63
5.10 Diagram vertical autocorrelation (A) along with the graph of the computed sum values (B). 64
5.11 Ideal case of perfectly aligned fibers (a) and the corresponding autocorrelation (b) and wavelet analysis (c). 65
5.12 Wavelet analysis of the fibers reveals lateral uniformity and wrinkle sizes for waiting times of 10 (a), 60 (b), and 120 minutes (c), using 2 nmol of EGFP-Ubx spread G0 buffer. 66

5.13 Titanium discs cleaned for 15 minutes with IPA in a sonicator prior to do Plasma cleaning for 2 minutes with (a) SF_6, with 30 sscn flow rate, p = 200 mT, P = 200 W; (b) Ar, with 100 sscn flow rate, p = 70 mT, P = 50 W; (c) Ar, with 100 sscn flow rate, p = 70 mT, P = 200 W. 67
5.14 Fluorescent pictures taken with ×4 objective lens using an excitation wavelength $\lambda = 488$ nm. The disc was cleaned with SF_6 plasma, (a) and (b), and Ar plasma, (c) and (d), prior the dip coating with EGFP-Ubx. 68
5.15 Fluorescent pictures taken with ×4 objective lens using an excitation wavelength $\lambda = 488$ nm. The disc was cleaned with SF_6 plasma and checked (a) 1 week (b) 4 weeks (c) 7 weeks and (d) 11 weeks after the dip coating with EGFP-Ubx. 69
5.16 Model of the BMAP fusion protein illustrating the dityrosine bond forming residues thought to be responsible for material formation. 70
5.17 Left part: scheme of experiment. Silica surfaces are indicated with S, etched titanium discs are indicated with O and non-etched ones with N. Ubx-coated surfaces are indicated with 'x', while control surfaces are indicated with 'c'. On the top of samples bacteria solutions, C and D, and plain PBS (indicated as PBX) were pipetted. Right part: results of overnight fluorescence. 72
5.18 A-U BAM time-lapse during barrier expansion. The protein film was created using 2 nmols EGFP-Ubx proteins spread in 0.1 M NaCl solution and let rest 2 hours before the barrier closure. 73
5.19 π-A isotherm with BAM images taken during film compression-expansion in plain water. 75
5.20 π-A isotherm with BAM images taken during film compression-expansion in 0.1 M KCl. 75
5.21 Lateral association between big thick island and a small protein agglomerate. 76

6.1 Different loading approaching (a) one single droplet was let to fall from the top at the centre of Petri dish; (b) one single droplet was spread by touching gently the surface at the centre of Petri dish; (c) adsorbed layers from premixed solutions. . . 78
6.2 Time-resolved surface excess of spreading technique with adsorption (black), one droplet touching gently the surface (red), one droplet falling at the centre of the static trough (blue) . . 79

6.3 Time-resolved surface excess of adsorption technique. In the figure frame the same data are displayed on a $t^{1/2}$ scale. The dotted black lines help to visualise the linear component. . . . 80
6.4 Time-resolved surface excess of old proteins tested in 2020 (red), old proteins tested in 2021 (blue) and new proteins tested in 2021 (black). 82
6.5 Time-resolved surface excess of old proteins spread in 0.1 M NaCl using H_2O (a) and D_2O (b) as subphase. 83
6.6 Time-resolved surface excess of EGFP-Ubx proteins spread in different subphases: H_2O, 0.1 M NaCl in H_2O, D_2O and 0.1 M NaCl in D_2O. 84
6.7 Time-resolved surface excess of EGFP-Ubx proteins spread in D_2O and H_2O with two concentrations c_1 and c_2. 84
6.8 Measured reflectivity profiles at two contrasts obtained for two compression areas (300 and 110 cm^2). In the inset Corresponding SLD profiles from the NR curves. 87
6.9 Top left: Reflectivity profile the protein concentration c_1 spread in ACMW with barriers fully compressed. The curves are shifted vertically for clarity. Top right: Corresponding SLD profiles from the NR curve for the protein concentration c_1 spread in ACMW with barriers fully compressed. Bottom: Structural parameters obtained from the NR curves: d (thickness [Å]), and solvent content[%] for the protein concentration c_1 spread in ACMW with barriers fully compressed. 90
6.10 Top: Reflectivity and corresponding SLD profiles at three waiting time (10, 60 and 120 minutes) spreading c_1 of EGFP in ACMW and c_2 of EGFP in D_2O. Bottom: Structural parameters obtained from the NR curves: d (thickness [Å]), solvent content[%] and roughness (σ) [Å]. 92

ABSTRACT

Protein-based materials are currently the subject of intense research interest since they have an extended range of potential applications, such as improved bio-membrane biocompatibility for implanted medical devices and the creation of platform materials for novel biosensors. Monomers from Ultrabithorax (Ubx) transcription factor are known to spontaneously self-assemble at an air-water interface to form a monolayer, which has then been used as a basis for forming biopolymeric fibers. Here we used the Langmuir trough technique, Brewster angle microscopy (BAM), ellipsometry and neutron reflectometry (NR) to investigate the influences of different experimental conditions on EGFP-Ubx monolayer formation and the impact on biopolymeric fiber structure. We varied protein concentration, buffer properties and waiting times prior to forming biopolymeric fibers. Interestingly, we found 3 phases of material formation which brought us to a new protocol for forming fibers that reduced protein concentration by 5-fold and waiting times by 100-fold. Moreover, an in-house developed MATLAB code was used to analyze SEM images and obtain quantitative structural information about the biopolymeric fibers that were correlated directly to the surface film characteristics measured in the LB trough. These new insights into fiber formation and structure enhance the usefulness of the Ubx-based biopolymer for biomedical applications.

1. INTRODUCTION

Protein-based materials are currently the subject of intense research interest since they have an extended range of potential applications, such as improved bio-membrane biocompatibility for implanted medical devices and the creation of platform materials for novel biosensors and tissue engineering. In order to achieve these objectives, it is important to develop materials with a broad range of structural, mechanical, and functional properties optimized for each application. Self-assembly biopolymers are often used to generate the desired biomaterials with specific structure and application [61, 92, 134]. For example, spider silk has a tensile strength comparable to the structural steel one, elastin-based materials are extremely extensible and with collage it is possible to produce 3D gels [74, 127, 68]. A peculiar property of protein-based material, is the potential to incorporate other proteins with their several properties, into the material with a simple molecular engineering [61, 64, 44]. A subset of proteins naturally self-assembles into useful materials. Common examples include the already cited silks and collagen but the list of molecular proteins is in constant expansion creating novel proteins with the property to form nano-scale scaffolds. In the present dissertation we will briefly discuss the properties of silk and collagen materials. However, among all of them Ultrabithorax-based material will be used in the present work, due to its properties. Ubx protein is a Hox transcription factor of *Drosophila melanogaster* fruit fly that functions by binding DNA (deoxyribonucleic acid) to activate or repress transcription processes [76, 78]. In addition to its *in vivo* function, monomers from Ubx are known to spontaneously self-assemble at an air-water interface to form a monolayer, which has then been used as a basis for forming biopolymeric fibers and films. During this spontaneous self-assembly, nanoscopic aggregates of Ubx are formed that gradually assemble into protofibrils [80]. Lateral association among protofibrils generates a surface film that can be pulled into microscale fibers or lifted off as films [59]. The potential applications of this biopolymer have been enhanced by using an *E. coli* expression system to obtain purified fusion proteins that also self-assemble into biopolymeric membranes. Such fusion proteins utilize the Ubx portion for material formation while the biomolecule remains in an

active conformation in the membranes. Functionalised Ubx-based materials have potential impact in numerous biomedical applications including wound healing, biocompatibility coatings and environmental sensing [105, 125].
In the present work, we use a Ubx fusion with EGFP (Enhanced Green Fluorescent Protein) to investigate the mechanisms of Ubx film formation, particularly during early stages of Ubx monomer aggregation at the air-water interface. Fusion of EGFP does not affect the ability of Ubx to self-assemble. However, we discovered this fusion plays an important role in the film-formation dynamics. In fact, Ubx monomers precipitate in water and, presumably due to the positive charges (predicted net charge of +11), they explode without any counter ions. The negative charges present in EGFP prevent the explosion of Ubx monomers.
However, at the state of art, the main problem is given by the lack of understanding of the physical and chemical mechanisms underlying material formation. This means that if we do not understand how the film forms, it is not possible to fully control the process. In order to fill this gap, my work was focused on understanding the mentioned process with laboratory-based techniques.
Following previous works on other types of protein films [29], critical parameters for this aggregation include the concentration of the Ubx protein and the solution in which the protein is suspended. The composition of the solution impacts the formation of the surface film of Ubx monomers and hence the properties of resulting monolayer biopolymeric membrane. Indeed, the monomer must have a determined configuration in order to aggregate [15]. Moreover, the monomers must be at the air-water interface in order to aggregate with other monomers. Hence, conditions that promote aggregation at the surface of the water are required. Our hypothesis is that the waiting time for aggregation is an essential parameter to control to achieve an ordered structure of the biopolymeric membrane. That is, the monomers need time to change conformation, rise to the surface, then begin to aggregate and order themselves. We used the technique to investigate Ubx monomer aggregation and formation of the biopolymeric membrane by comparing the surface pressure–area (π-A) isotherms obtained under different conditions [71, 72, 12]. The shape of the surface pressure–area (π-A) isotherm depends on the nature of amphiphiles, the compression speed, the spreading conditions and the temperature. Our experiments indicated that the waiting time between the introduction of Ubx monomers and the Langmuir barrier compression was an additional critical parameter in formation of biopolymeric membranes. Moreover, we developed a new image analysis technique based on autocorrelation of SEM (scanning electron microscopy) images of the biopolymeric membranes. This autocorrelation technique is applicable

to any digital image, and in our experiments was used to quantify the linear structure of the Ubx biopolymeric membranes.
Finally, we decided to go a bit further in the film formation analysis. After some parameter optimization done with ellipsometry, I moved to neutron reflectometry. In this case, we decided to study the film formation dependence on waiting time parameter at two contrasts ACMW (air contrast matched water) and D_2O. With this method, it was able to estimate the film thickness and to corroborate the observations made previously with Langmuir technique.

2. ULTRABITHORAX PROTEIN

Protein-based biomaterials are next generation technology due to their extended range of potential applications, such as drug delivery, tissue engineering, photonics, implanted medical devices and the creation of platform materials for novel biosensors[55, 63]. This work will be focused on the study of film formation of Ubx (Ultrabithorax) protein-based material. Ubx is a Hox transcription factor in the fruit fly *Drosophila melanogaster*, that spontaneously self-assembles at an air-water interface to form a monolayer, which has then been used as a basis for biomaterial uses.

2.1 Homeodomain Transcription Factors

TR (transcription factor) are proteins that control the rate of transcription of genetic information from DNA to m-RNA (messenger ribonucleic acid) by binding to a specific DNA sequence[113, 99]. In other words, their role is to guarantee that genes are expressed in the right cell, at the right time and in the right amount. Some TFs are characterized by the presence of Hox (homeodomain), 60 amino helix-turn-helix DNA-binding domain[99, 58, 81]. A homeobox is a DNA sequence that encodes the homeodomain and Hox genes are homeobox-containing genes. Homeodomain Transcription Factors, known also as Hox proteins, operate during the embryo growth of bilaterally symmetric animals, to generate tissues, organs and appendages specifying the fate of structures in every tissue layer[58]. A given Hox protein can be involved in the specification of multiple body structures, that is it can control the expression of a variety of target genes, which may number into the hundreds [62]. For example, *Drosophila melanogaster* Hox protein Ubx specifies the posterior-most legs, the halteres (which are organs that take the place of the second pair of wings and regulate the balance during flight) and the posterior aorta, portions of the midgut, ectoderm, musculature and central and peripheral nervous systems[102, 11, 58, 62], as shown in Figure 2.1.Misexpressions and mutations of Hox proteins during embryo development are linked to birth defects, such as the growth of legs instead of antennae out of the head of *Drosophila melanogaster* if gain-of-

function mutations occur in *Antennapedia* gene. Mutations and deregulations in adulthood are associated with increased cancer risk, lower survivability, and tumorigenesis[75, 16].

Hox proteins are intrinsically disordered which means that segments or

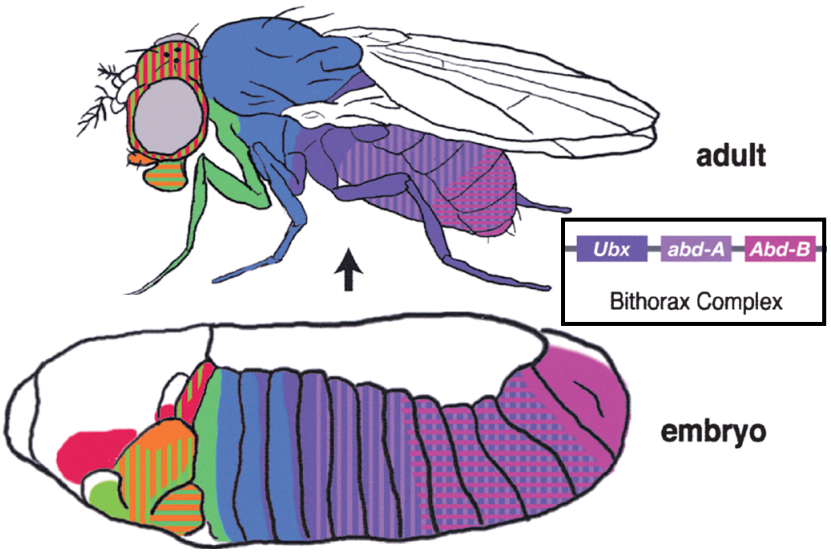


Fig. 2.1: Area of expression in embryo and adult life stages are shown. Reproduced from reference [62].

the whole protein fails to self-fold into a fixed 3D structure[104]. The disorder causes unique problems in protein characterization and study. In X-ray diffraction residues are missing from electron density maps; in NMR spectroscopy, sharp peaks, absence of NOEs (nuclear Overhauser effects) characteristic of secondary structure or negative values for ^{1}H-^{15}N heteronuclear NOEs are an indication of disorder[6, 40, 32, 86, 104]. Ubx is also the only Hox protein in which predicted intrinsically disordered regions have been experimentally validated[58, 77]

In addition to its in vivo function as a transcription factor, Ubx also self assembles into biomaterials through intermolecular dityrosine bonds[57]. Bondos's lab in Texas (Texas A&M University) has characterized the materials and explored potential uses in a variety of areas [15, 17, 16, 57, 56]. My work focuses on an in-depth study of Ubx film-formation.

2.2 Protein-Based Materials

In the last four decades, following the pioneering work of Stanley Cohen and Herbert Boyer[26] the design and production of protein-based materials has been highly studied thanks to their versatility in medicine and biomedical devices in combination with good biocompatibility and biodegradability.[130, 7, 117]. In fact, protein-based materials have different applications ranging from drug delivery, tissue engineering, biosensors and cell culture studies[55, 28]. Several proteins have the ability to self-assemble at air-water interface, forming a film that can be re-arranged into other shapes, such as fibers and 3D gels [4, 92, 19, 80, 96].
In the following 3 sections brief outlines of the properties of Ubx and 2 other commonly-used protein materials (silk and collagen) will be given.

2.2.1 Silk material

Silk proteins are produced by animals of the phylum of arthropods among which silkworms and spiders are the most prominent silk producers. Two of the most studied silk proteins are the fibroins produced by the domesticated silkworm *Bombyx mori*, used during the metamorphosis from larvae to moth, and the MaSp (major ampullate spidroin) produced by spiders and used for different purposes such as web building, prey swathing and reproduction[3, 92]. Even though spiders and silkworms are distantly related, there are some features in common in spinning silk. Both of them produce and store in glands, highly repetitive, soluble and large silk proteins ($>$ 300 kDa). When these proteins are pultrused through a narrowing conduct, which changes conformation in response to pH variations, they assemble forming silk fibers under ambient conditions[3, 92]. The pH decrease along the conduct is a well known factor that regulates fiber formation, bringing the protein conformation to change from disordered and partly helical into densely stacked β-sheet crystals[35, 36, 2]. An important role in silk assemble is played also by ion concentration gradients, dehydration of the protein dope and hydrodynamic forces along the conduct[118, 67, 114]. Silk fibroin extracted from *Bombyx mori* assemble at the air-liquid interface [123, 122]. However, it is difficult to extract sufficient quantities from spiders to meet the demands. Hence recombinant technology has been used to produce spidroins in heterologous hosts (host organisms, such as bacteria, yeasts, plants, and animals, which do not naturally have the expressing gene or one fragment) ensuring constant quality and biological safety as well as self-assembly property[19]. However, recombinant silks are difficult to produce because of genetic instability of long and highly repetitive

DNA sequences, translation inhibition by complex mRNA secondary structures, a high demand for glycine and alanine transfer RNAs, and overall metabolic burden caused by spidroin overexpression. Moreover, protein functionalization with other functional proteins or tags can affect self-assembly property[19, 111]. To solve this problem and produce native-sized silk proteins with high yields,the silk production process is engineered. For example, high molecular weight silk proteins are produced by splitting intein-mediated ligation via hosts like *E. coli*[19] or plants[129].

2.2.2 Collagen material

Collagen is one of the most abundant proteins in the extracellular matrices of many mammals connective tissues, comprising about 25–35% of the dry weight of the body[34, 8]. Collagen is typically found in tendons, skin, ligaments, cartilages, bones, corneas, gut and blood vessels[73]. There are 29 recognized and categorized types of collagen with unique characteristics and functions. Collagen proteins have complex hierarchical structures containing different levels of structure. Firstly three amino-acid chains are wound together in a triple helical configuration forming a microfibril. Microfibrils are then organised into larger elongated fibrils. Finally, collagen fibers are formed by fibril self-assembly (see Figure 2.2).

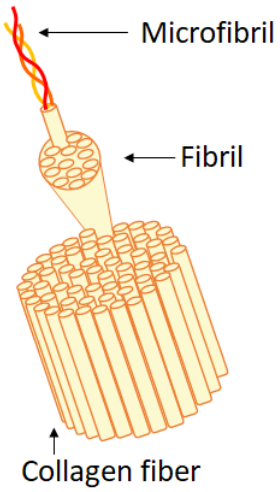


Fig. 2.2: Hierarchical structures in fibrillar collagen showing a characteristic periodic structure forming various hierarchical orders of association: fibers, fibrils and microfibrils with three helical chains. Collagen fiber diameter is from 1 to 20 μm.

Although collagen proteins are abundant in animals, once purified they lose some properties (e.g. mechanical strength) because of the loss of hierar-

chical structure. Physio-chemical properties (intrinsic physical and chemical characteristics of a substance) from α-chain are kept unaltered. Examples of these physiochemical properties are the ability to enhance the carrier systems of other materials, when they are processed as composite fibers, sponge or film [8]. For this reason, over the past decade, researcher have been working towards the possibility of mimicking the structure of collagen fibrils in native tissues, especially with the spinning technique.

2.2.3 Ultrabithorax material

As already mentioned, Ultrabithorax proteins are Hox transcription factors. It has been proved that Ubx monomers spontaneously self-assemble under mild, aqueous conditions, forming biocompatible ordered materials at an air-water interface [17] (see Figure 2.3).

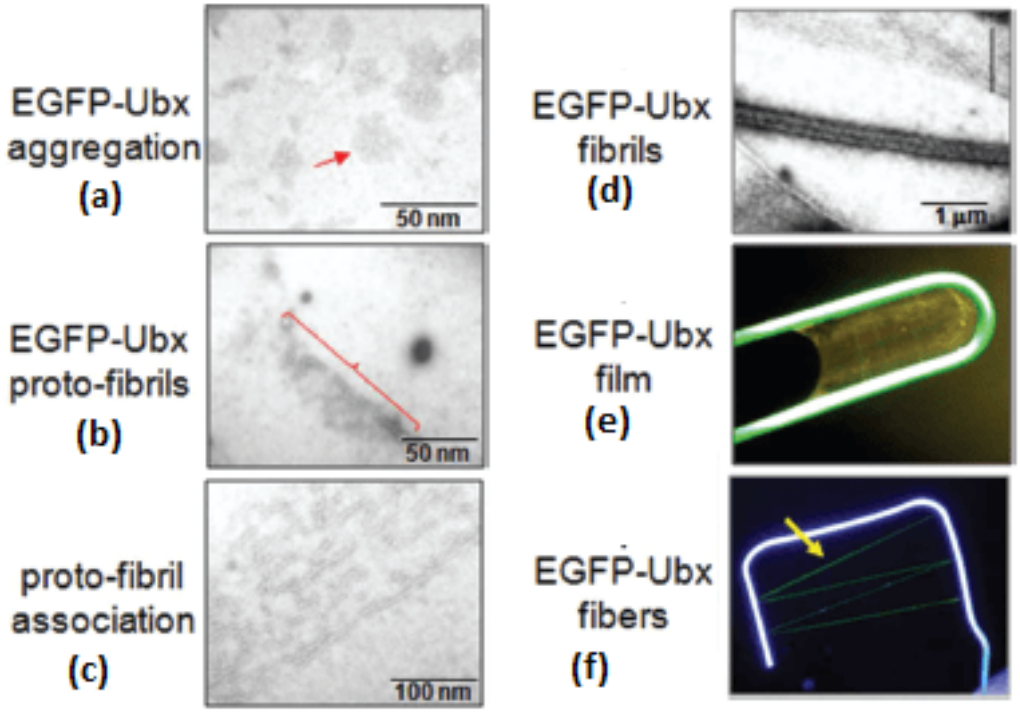


Fig. 2.3: TEM images showing the hierarchical film formation of EGFP-Ubx at the air water interface. (a)- Ubx aggregation at 15 minutes after dilution; (b)- protofibrils formation after 1h in dilution; (c)- parallel alignment of protofibrils after 1-2h; (d)- fibril formation of 50 nm diameter after 2h; (e)- lateral association of fibrils into film; (f)- Ubx film drawn into fiber. Reproduced from reference [80].

When Ubx protein droplets are diluted in a tray containing G0 buffer

(50 mM sodium phosphate buffer, pH 8.0, 5% glucose w/v, 500 mM sodium chloride), Ubx monomers migrate to the air-water interface, interact with each other and self-assembly. After 15 minutes of incubation, monomers start to form disordered aggregates with sub-25 nm size (2.3(a)). These clusters in their turn, interact reshaping them into linear protofibrils of $\sim 50nm$ size (2.3(b)). After 1-2 hours, protofibrils further laterally interact (2.3(c)) aggregating into fibrils with length variable from hundreds of nanometres to tens of micrometers (2.3(d)). Fibrils then form lateral association resulting in macroscopic films that are the building blocks for further applications (e.g. films, 2.3(e), and fibers, 2.3(f)).

2.3 Ultrabithorax fusions

On of the most important advantage of protein-based materials is the possibility to incorporate novel functions to them by producing a fusion protein capable of both self-assembly and the desired new activity that can be engineered for many applications. Although there are different techniques to incorporate one protein with another such as physical adsorption[9], cross-linking[87, 69], direct mixing with polymer solutions[70] and co-axial electrospinning[88, 112], the gene fusion approach is used to produce Ubx fusions. With this technique the DNA encoding one specific protein is placed in tandem with the DNA encoding Ubx, see Figure 2.4 where the functional protein sequences EGFP(green), m-Cheery(purple) and myoglobin (red) are linked to the Ubx sequence (brown) via a flexible linker (yellow). When

Fig. 2.4: Schematic gene sequence of Ubx fusion proteins used to create functionalised materials. The functional protein sequence (green, purple and red) are shown linked to flexible linker, in yellow, followed by the self-assembling protein Ubx in brown.

the resulting gene si placed in a living cell for protein production, the cell will produce a single polypeptide composed of the amino acid sequence of both proteins. The single polypeptide is called protein chimera or protein

fusion. During the film formation process, the functional protein is automatically embedded throughout the material maintaining its original activity [61]. There are several advantages using this technique. First of all, the two proteins don't have to be produced separately, reducing costs and time. Second, the resulting material is fully functionalised and uniform, that is each Ubx protein is attached to one functional protein and vice versa. Third, there is just one attachment between the two proteins (the flexible linker in Figure 2.4) leading to a uniform orientation of the resulting protein. Finally, the film formation brings a stabilization of functional proteins preserving their activities even in harsh conditions (e.g. heat, acid and other denaturating conditions)[57].

Bondos' group extended the gene fusion technique to incorporate functionalised proteins with different sizes into Ubx. They produced Ubx protein chimeras with several proteins such as Enhanced Green Fluorescent Protein (EGFP), mCherry and Sperm Whale Myoglobin. Each chimera can self-assemble at mild conditions keeping the functional activity of the parental components, even though the functionalised proteins have different sizes (from 18 kDa to 61 kDa), different charges at pH 8.0 (from -9 to + 2) and distinct secondary structure contents[61]. The proteins fused with Ubx have different properties: EGFP and mCherry are fluorescent (EGFP is green and mCherry is pink), and myoglobin absorbs visible light.

3. METHODS

3.1 Langmuir trough Technique

Langmuir trough technique is a a well-established and sophisticated method to control interfacial molecular orientation creating a films at air-water interface called Langmuir film [121]. Langmuir film study seems to date back to the 18th century BC in Babylon, but it's with the American statesman Benjamin Franklin that the history of Langmuir films began. In 1774 he made some simple quantitative calculations to find out that a few-nanometres-thick layer of oil had calming influence on the water in the Clapham ponds [42]. In the 19th century Angela Pockels set the fundamentals for the air-water-interface monolayers quantitative characterisation. She studied the contamination of water surface as a function of the area using a rudimentary surface balance[97]. At the same time, Lord Rayleigh hypothesized the maximum extension of an oil film represents a monolayer. It was about 25 years later that Irving Langmuir performed a systematic study of monolayers at air-water interface[71]. These studies led to him being awarded the Nobel Prize in 1932. Although Langmuir reported film transfer onto solid substrate, the first detailed study of sequential monolayer transfer was carried out by Blodgett[12, 13].
In order to prepare a Langmuir film, a given molecule, for this work EGFP-Ubx will be used, is dissolved in a water-insoluble solvent and then carefully spread onto the surface of a water solution. As the time passes and the solvent evaporates, a monolayer forms at the surface. When the available area, where the molecules are spread, is too large compared to the monolayer dimensions, the distance between adjacent molecules is so huge that their interaction is too weak. In this case we can treat the system as a two-dimensional gas. On the other hand, if the surface area is reduced, the molecules start to interact with each other and form a stronger film and the system can be treated as a liquid. Since Pockels studies in 1935, scientists studied and developed a mechanised computer-controlled instrumentation for the film manipulation at air-water interface. The instrumentation is called Langmuir trough, a schematic description is shown in Figure 3.1. In this scheme, A is a bath

where the subphase is poured, B are the moving barriers and C is a balance that measures the surface pressure. Both bath and barriers are usually made of Teflon, first used by Fox and Zisman [41]. This material, in fact, enables scientists to clean properly the instrumentation using aggressive chemicals to ensure a complete removal of organic contaminants. The surface-pressure measuring balance used is the Wilhelmy plate[132].

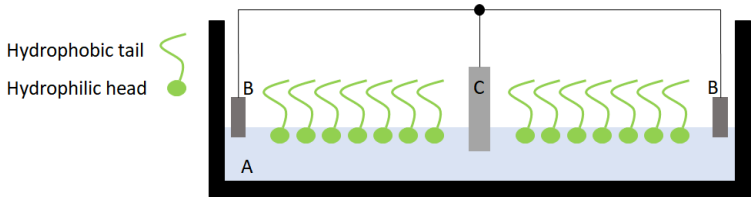


Fig. 3.1: Langmuir trough scheme. A, bath where the subphase is poured; B moving barriers to control the surface area dimension; C surface-pressure balance with Wilhelmy plate measuring the surface pressure.

The most commonly used subphase is deionised water to which it is possible to add non organic salts in different concentrations such as Sodium Chloride (NaCl) and Potassium Chloride (KCl). Once the subphase is poured in the trough, it is important to check the its purity before adding the sample. Any contamination in the aqueous solution may result in an error measure or in a damage in film formation. Once everything is clean, it is possible to spread at the surface the amphiphile molecule. A molecule is an amphiphile, when it is insoluble in water, with one end that is hydrophilic and the other one hydrophobic. An example of amphiphile is given by phospholipids present in cell membranes, Figure 3.2.
The hydrophobic tail preferentially resides in the air and the hydrophilic head is preferentially immersed in the subphase. Since the amphiphiles have one hydrophobic tail and one hydrophilic head, they like to locate at the interfaces (e.g. air-water or oil-water). For this reason they are also called surfactants (surface active).
Once one o more droplet of a dilute solution of an amphiphile molecule is spread at the subphase surface, the barriers move and compress the molecules while the pressure and the area are recorded. The surface-pressure is described as:

$$\pi = \sigma_0 - \sigma \tag{3.1}$$

Where π is the surface pressure, in mN/m, σ_0 is the surface tension of pure subphase and σ is the surface tension of the all system (subphase plus monolayer). The surface tension is a water property that measures the difference in energy of molecules between the bulk and the surface. All the molecules

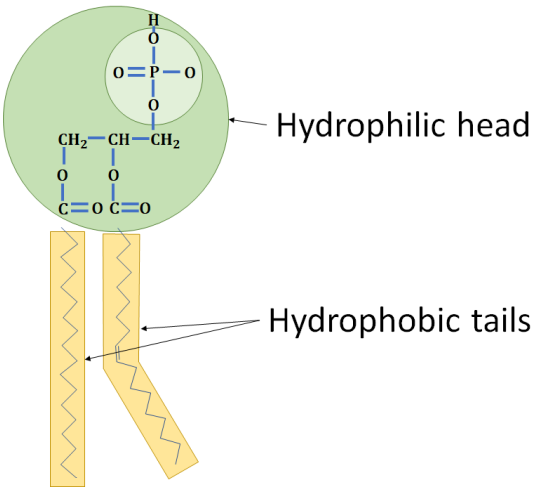


Fig. 3.2: Phospholipid of the plasma membrane with a hydrophilic head and two hydrophobic tails.

in the liquid exert an attractive van der Waals force between each other in all directions. However, surface molecules experience an imbalance of forces due to the reduced number of neighbours and hence reduced attractive interactions than a molecule in the bulk, as seen in Figure 3.3. Under equilibrium conditions, the surface tension of water is 72 mN/m.

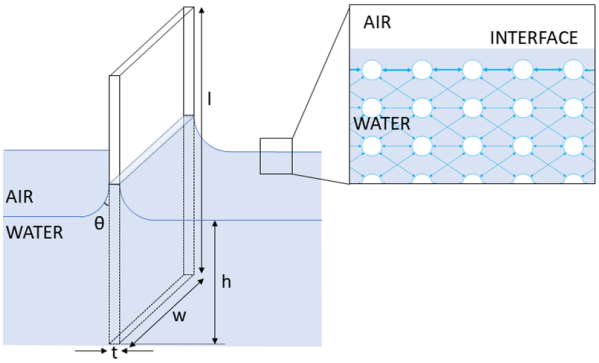


Fig. 3.3: Scheme of Wilhelmy plate partially immersed in a water subphase where t is the plate thickness, θ is the contact angle, w is the plate width, h is the immersed depth and l is the plate length. In the inset there's the scheme of the forces acting on subphase molecules at the air-water interface.

With the Wilhelmy plate- method, we can calculate the force due to surface tension on a partially immersed plate in the subphase, Figure 3.3. There are three forces acting on the plate: in the downward there are the gravity and

surface tension force , while in the upward direction there is the force caused by the displaced water. The net downward force for a l × w × t Wilhelmy plate, with a material density ρ_p, immersed to a depth h in a liquid whose density is ρ_l is:

$$F = glwt\rho_p + 2\gamma(tw)cos\theta - \rho_l gtwh \tag{3.2}$$

where γ is the liquid surface tension, θ is the contact angle and g is the gravitational constant. To determine the surface pressure it is just necessary to measure the force changes for a stationary plate between a clean surface and the same surface with an amphiphile monolayer. When the plate is fully wetted, $cos\theta = 1$ and the surface pressure is:

$$\pi = -\Delta\gamma = -\frac{\Delta F}{2(t+w)} \tag{3.3}$$

$$\pi = -\frac{\Delta F}{2w}, if w >> t \tag{3.4}$$

The force is determined by measuring the changing of plate mass, which is directly coupled to a sensitive electro-balance. The measure of surface pressure as a function of area at a constant temperature is also known as surface-pressure isotherm, or π-A and is an important indicator of the monolayer properties such as stability, reorientation of molecules, phase transitions and conformational transformations, Figure 3.4. Discontinuities in the π-A isotherms identify these changes indicating transition to different phases. The phase behaviour is a consequence of several factors such as physical and chemical properties of the amphiphile and the temperature or composition of the subphase. For this reason, in many cases phase transitions are not observed.

When the surface pressure is low, the area between molecules is large and ideally there are no interaction. In this case, the molecules have a negligible size compared to the interface area, and thus will obey the equation $\pi A = \kappa T$,where T is the absolute temperature and κ is the Boltzmann constant. This is a two dimensional ideal gas equation, for this reason, this phase is called "gas phase". As the barriers move, the molecules are compressed in a smaller area and the intermolecular distance is decreased and consequently the surface pressure increases. A phase transition is observed in the isotherm: from gas phase molecules rearrange into a liquid phase, where the monolayer is coherent even though the molecules have a large degree of freedom. When the film is further compressed, a second phase transition occurs, from the liquid to the solid state. In this condensed phase the molecules are packed and well oriented. If the monolayer compressed more and more, it collapses due to mechanical instability and a sharp pressure decrease is recorded[84].

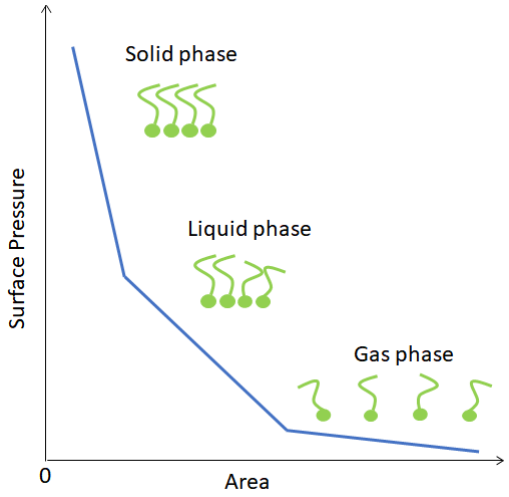


Fig. 3.4: Scheme of surface pressure (π) vs. area (A) and orientation of surfactants in different phases.

3.2 Brewster Angle Microscopy

With the use of Langmuir trough technique, it is possible to monitor the film formation and the molecular orientation, as already seen. However, it is possible to obtain other information during the film formation if in combination with a Langmuir trough, a BAM (brewster angle microscopy) is used. It was developed at the beginning of 90s [51, 53, 54] to study Langmuir films and to overcome the drawbacks of other techniques (e.g. fluorescent microscopy) whose goal was the observation of coexisting domains of different phases in monolayers. This method takes advantage of Brewster angle condition. Light beam is an electromagnetic wave where electric and magnetic fields are

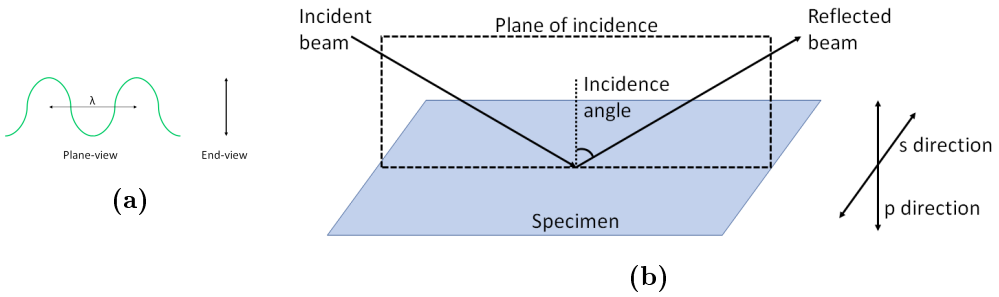


Fig. 3.5: a - Schematic plane-view and end-view representation of wave nature of light . b - Light beam incident upon and reflected from a surface showing the plane of incidence and the s-direction and p-direction.

transverse to the direction of wave propagation and also perpendicular to each other. By convention, the polarization of electromagnetic waves refers to the direction of the electric field. If the electric field oscillates perpendicularly to the plane of incidence (the plane in which the light travels before and after reflection) than the light polarization is called s-polarization. At the contrary, if the electric field oscillates along the plane of incidence, the polarization is called p-polarization, see Figure 3.5b. The reflectivity of a incident light beam at a plane between two media, with refractive index $n_1 and n_2$, depends on its polarization and incidence angle θ, Figure 3.6a.

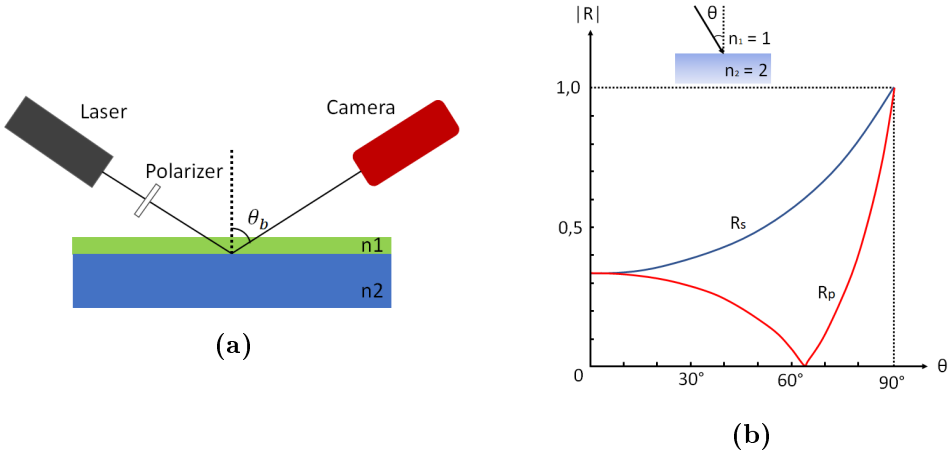


Fig. 3.6: a - Optical setup of BAM. b - Reflection coefficients vs. incidence angle for s- and p- polarized light at air-glass interface. $n_a = 1$ and $n_g = 2$.

For a Fresnel interface (an interface where the refractive index changes steeply from n_1, to n_2) and for a p-polarization light (electric field is parallel to the incidence plane), the reflectivity, R_p, is null at Brewster angle, θ_b. Figure 3.6b shows the plot of R_s and R_p versus θ at air-glass interface, or rather when $n_1 = 1$ and $n_2 = 2$. As θ varies from 0° to 90° (from normal incidence to grazing), R_s increases monotonically reaching the maximum value 1 at $\theta = 90$. On the other hand, R_p decreases initially, reaches the minimum value $R_p = 0$ at $\theta = \theta_b$ and then increases up to $R_p = 1$ at $\theta = 90$. Thus, $R_s > R_p$ for all angles θ and there's no reflection for parallel polarized waves at the Brewster's angle. The reflectivity Fresnel coefficient for non-magnetic media is:

$$R_p = \frac{n_2 \cos\theta_i - n_1 \cos\theta_t}{n_2 \cos\theta_i + n_1 \cos\theta_t} \tag{3.5}$$

where θ_i is the incidence angle and θ_t is the transmitted one. The Brewster angle can be calculated from Equation(3.5) letting $R_p =$ at $\theta_i = \theta_b$:

$$n_1 \cos\theta_t = n_2 \cos\theta_b \tag{3.6}$$

with the use of Snell's law

$$n_2 \sin\theta_t = n_1 \sin\theta_b \tag{3.7}$$

it is possible to eliminating θ_t from Equation(3.6), and the Brewster angle is:

$$\theta_b = \arctan\frac{n_2}{n_1} \tag{3.8}$$

In the case of air-water interface ($n_a = 1$ and $n_w = 1.333$), the Brewster angle is $\theta_b = 53.13$ at a wavelength $\lambda = 532nm$. Even though in reality the refractive index doesn't change steeply between two media, the reflected light has a minimum at θ_b. Spreading a thin molecular film at the air-water interface leads to a change in the refractive index and thus, to a measurable change in reflectivity[101, 124] (Figure 3.7).

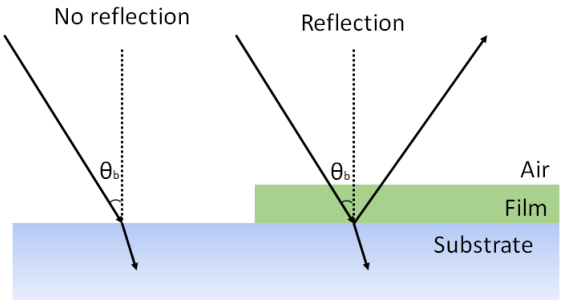


Fig. 3.7: Brewster Angle Microscopy physical functioning.

With this technique, both the shape of the condensed phase domains and their inner texture can be visualized. In fact, anisotropy plays an important role in the BAM images as inner textures and can be used to obtain structural analysis [25, 103]. The Reflectivity, in this case, is described by the following equation:

$$R_p = \left(\frac{\pi d}{\lambda}\right)^2 \frac{1}{1-n_2^2}\left(n_1^2 - n_2^2 - 1 + \frac{n_2^2}{n_1^2}\right)^2 \tag{3.9}$$

where λ is the wavelength and d is thin film thickness. Brewster angle microscope is used to visualize Langmuir monolayers or adsorbate films at the air-water interface (see Figure 3.8)for example as a function of packing density, to study the properties of Langmuir layers, or to indicate a suitable deposition pressure for Langmuir-Blodgett deposition. Usually, with BAM it is possible to observe the phase transitions of a Langmuir monolayer by the appearance of domains due to reflection

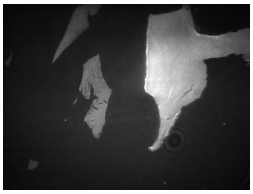


Fig. 3.8: Example of BAM image

from polar headgroups[103]. If anisotropy is absent in the Langmuir film, the domains are related to a condensed phase [103]. Special attention is paid to inner textures visible in the domains. Inner textures are related with the anisotropy and their appearance is indicative of a high molecular ordering in the Langmuir monolayer [103]. This microscope has revolutionized the visualization of monolayers at interfaces by allowing label-free real-time images of fully hydrated films. Interactions that only resulted in minor changes of area-surface pressure isotherms still resulted in significant changes of the lateral film organization.

3.3 Ellipsometry

Even though at the beginning of nineteenth century, Augustin Jean Fresnel (1788 - 1827) and Dominique François Arago (1786–1853) have mathematically studied and characterized the light polarization, the concept of partially polarized light arose only around the year 1852 when Sir George Gabriel Stokes demonstrated that it was possible to completely characterize any arbitrary polarization state from four experimental measurements [116, 126]. Thanks to Stokes' work, in 1946, Subrahmanyan Chandrasekhar formulated the equations for radioactive transfer of partially polarized light[27, 126]. In addition to Stokes theory, new developments made possible the polarization states measure as a function of wavelengths that brought to ellipsometry. Ellipsometry is a very sensitive technique which examines the optical properties of surfaces and thin films. It is based on the use of polarized light. Along with frequency and wavelength, light polarization is one of fundamental parameter used to describe transverse waves and specifies the geometrical orientation of the oscillations(perpendicular to the direction of motion of the wave)[5], Figure 3.5a.

Measurements are based on the change in polarization of light reflected at a surface, where the relative amplitude and phase of the p- and s-polarized components change by different amounts[65, 21].
Light with p- and p- polarization remains linearly polarized upon reflection. Incident and reflected bean can hence be described by their corresponding Jones vectors:

$$\vec{E}_{inc} = \begin{pmatrix} |E_p^i| e^{i\delta_p^i} \\ \\ |E_s^i| e^{i\delta_s^i} \end{pmatrix} \vec{E}_{ref} = \begin{pmatrix} |E_p^r| e^{i\delta_p^r} \\ \\ |E_s^r| e^{i\delta_s^r} \end{pmatrix} \tag{3.10}$$

To describe the changes in the state of polarization, it is possible to use two quantities:

$$\Delta = (\delta_p^r - \delta_s^r) - (\delta_p^i - \delta_s^i) \tag{3.11}$$

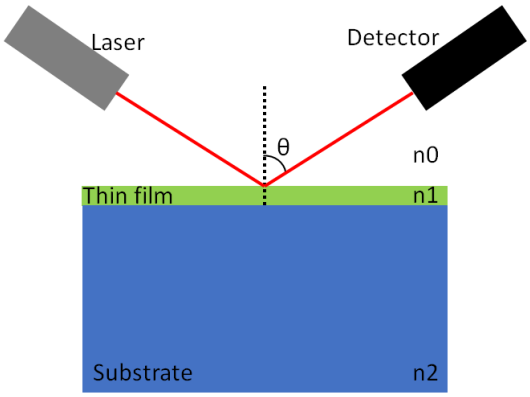


Fig. 3.9: Schema of ellipsometric measure.

$$\tan \Psi = \frac{|E_p^r|/|E_p^i|}{|E_s^r|/|E_s^i|} \tag{3.12}$$

The reflectivity properties of a sample within a given experiment are given by the corresponding reflection coefficients r_p and r_s:

$$r_p = \frac{|E_p^r|}{|E_p^i|} e^{i\delta_p^r - \delta_p^i} \qquad r_s = \frac{|E_s^r|}{|E_s^i|} e^{i\delta_s^r - \delta_s^i} \tag{3.13}$$

If we assume that there is no coupling between the orthogonal p- and s-polarizations (e.g. isotropic samples), the changes in the attenuation ψ and phase Δ can be described with the following equation[79]:

$$\tan \psi e^{j\Delta} = \frac{r_p}{r_s} = \rho \tag{3.14}$$

3.3.1 Ellipsometry applied to ultra-thin films

In the presence of an organic non-absorbing monolayer (refractive index $n \sim 1.3 - 1.6$) with a thickness below $2.5nm$ the sensitivity of ψ is negligible compared to the *Delta* one, so it is possible to consider just the phase shift. For a single homogenous film with refractive index $n_1 = \sqrt{\epsilon_1}$ in between two infinite media with $n_0 = \sqrt{\epsilon_0}$ and $n_2 = \sqrt{\epsilon_2}$ (see Figure 3.9), the reflectivity coefficients can be written as:

$$\Delta = \arctan \frac{Im\{\frac{r_p}{r_s}\}}{Re\{\frac{r_p}{r_s}\}} \tag{3.15}$$

The reflectivity coefficients r_p and r_s can be written using the Fresnel equations which describe the reflection at the surfaces $n_0 - n_1$ and $n_1 - n_2$ [21]:

$$r_p = |r_p|e^{i\delta_p^r} = \frac{r_{0,1,p} + r_{1,2,p}e^{-i2\beta}}{1 + r_{0,1,p}r_{1,2,p}e^{-i2\beta}} \quad r_s = |r_s|e^{i\delta_s^r} = \frac{r_{0,1,s} + r_{1,2,s}e^{-i2\beta}}{1 + r_{0,1,s}r_{1,2,s}e^{-i2\beta}} \tag{3.16}$$

where $\beta = 2\pi\frac{h}{\lambda}\sqrt{n_1^2 - n_0^2\sin^2\theta}$, λ is the wavelength of incident light, h is the layer thickness and θ is the incident angle. In the thin film limit ($h << \lambda$), it is possible to expand the complex reflectivity coefficients as a power series in terms of h/λ[85, 21]:

$$\Delta \sim \frac{4\sqrt{\epsilon_0}\epsilon_2\pi\cos\theta\sin^2\theta}{(\epsilon_0 - \epsilon_2)[(\epsilon_0 + \epsilon_2\cos^2\theta) - \epsilon_0]} \cdot \frac{(\epsilon_1 - \epsilon_0)(\epsilon_2 - \epsilon_1)}{\epsilon_1} \cdot \frac{h}{\lambda} \tag{3.17}$$

Δ is linearly proportionally to the ellipsometric thickness, η:

$$\Delta = \frac{g(\theta)}{\lambda}\eta \tag{3.18}$$

$g(\theta)$ is a function depending on bulk properties ($n = \sqrt{\epsilon}$) and incident angle:

$$g(\theta) = \frac{4\pi n_0 n_2^2\cos\theta\sin^2\theta}{(n_0^2 - n_2^2)[(n_0^2 + n_2^2\cos^2\theta) - n_0^2]} \tag{3.19}$$

In our case $n_2 = n_{water} = 1.33$, $n_0 = n_{air} = 1$, $\theta = 50$ and $g(\theta) = 75.70$.

$$\eta = \frac{(n_1^2 - n_0^2)(n_2^2 - n_1^2)}{n_1^2}h \tag{3.20}$$

Generally, the material density in adsorbed organic layers decreases with increasing distance from the interface, so the refractive index of the layer varies over its height and η is calculated as an integral across the interface[37, 85, 21, 33]:

$$\eta = \int \frac{(n_1^2(z) - n_0^2)(n_1^2(z) - n_2^2)}{n_1^2(z)}dz \tag{3.21}$$

Using the Gibbs convention, it is possible to introduce the parameter Γ, surface absorption per unit of surface area, as[33]:

$$\Gamma = \int \Delta c_1(z)dz \tag{3.22}$$

where $\Delta c_1(z) = c_1(z) - c_2$ is the excess concentration of the solute in the adsorbed layer, $c_1(z)$ is the absolute concentration in the layer, and c_2 is the

solute concentration in the bulk solution. De Feijter's equation can be used to derive a relation between the surface excess Γ and Δ[21, 33]:

$$\Gamma = \frac{n_1^2(n_1 - n_2)}{(n_1^2 - n_0^2)(n_1^2 - n_2^2)} \cdot \frac{\lambda}{\frac{dn}{dc} g(\theta)} \cdot \Delta \tag{3.23}$$

where $\frac{dn}{dc}$ is the refractive index increment. In this work the thin film studied with ellipsometry is protein-based so the refractive index increment can be estimated from the protein compositional prediction[82, 134, 66]. The refraction per gram R_p of the protein is calculated as the weight average of the contributions from the individual amino acids R_i:

$$R_p = \frac{\sum_i R_i M_i}{\sum_i M_i} \tag{3.24}$$

where M_i is the residue molecular mass. In the same way, it is possible to calculate the protein partial specific volume, $\bar{v}_p$ as:

$$\bar{v}_p = \frac{\sum_i \bar{v}_i M_i}{\sum_i M_i} \tag{3.25}$$

using the Lorentz-Lorenz formula [134, 66], $R = \bar{v}(n^2 - 1/(n^2 + 2))$, the refractive index of the protein n_1, is:

$$n_1 = \sqrt{\frac{2Rp + \bar{v}_p}{\bar{v}_p - R_p}} \tag{3.26}$$

Assuming volume additivity, the refractive index of protein in solution can be calculated from the Wiener equation[131, 134, 66]:

$$\frac{dn}{dc} = \frac{3}{2} \bar{v}_p n_2 \frac{n_1^2 - n_2^2}{n_1^2 + 2n_2^2} \tag{3.27}$$

where n_2 is the refractive index of the solvent, see Figure 3.15. Corrections for the wavelength were implemented[134, 66]:

$$\left(\frac{dn}{dc}\right)_\lambda = \left(\frac{dn}{dc}\right)_{578nm} \left(0.94 + \frac{20000nm^2}{\lambda^2}\right) \tag{3.28}$$

3.4 Scanning Electron Microscopy

The common optical microscope is really useful to image biological samples. However, its resolution is physically limited by Abbe limit[1]:

$$d_{min} = \frac{\lambda}{2nsin(\theta)} = \frac{\lambda}{2NA} \tag{3.29}$$

where λ is the wavelength used, n is the refractive index of the medium, θ is the incident angle and NA is the numerical aperture. So, the smallest resolvable distance between two points is half the wavelength of the imaging light divided by NA. Since visible light ($400nm < \lambda < 700nm$) is used as source in the optical microscope, only samples whose characteristic dimension is bigger than 0,2 µm can be imaged with this technique. To overcome Abbe limit an electron source can be used, in fact electrons have shorter wavelength than photons (e.g. $\lambda = 2.74$ pm at E = 200 keV)
Electron microscopy history is correlated to the development of electron optics. In 1926, Busch laid the basis of geometrical electron optics with the study of charged particles trajectories in axially symmetric electric and magnetic fields, showing that such fields could act as particle lenses[14, 93]. De Broglie, simultaneously, introduced the concept of wave-particle duality, that is every particle exhibits a wave-like behaviour . Hence, a wavelength was associated with charged particles[48, 14]. With these two discoveries, the idea of an electron microscope began to take shape. Several years later, in 1930s, Ruska made the first attempt with an electron microscope prototype, facing many technical limits to overcome[107, 106]. Ruska and Knoll tried to implement Busch's lens formula experimentally constructing the first transmission electron microscope (TEM) in 1931, with a magnification of 16 [47]. The first scanning electron microscope was built in 1935 by Knoll and his group. However, the resolution was low ($\sim 100\mu m$) due to the missing demagnification that would have produced a finer probe. Since then, many scientists worked to improve SEM (scanning electron microscopy) to the instrument we know and use nowadays.
In Figure 3.10 there's a schematic diagram of a SEM. In electron microscopy, a high-energy electron beam is focused on a specimen, and scanned along a pattern of parallel lines. A thermionic gun emits electrons, called PEs (primary electrons), when a very high electric field applied. PEs are formed into a finely focused beam passing trough magnetic lenses and sent to the sample.The beam probe interacts with specimen surface and as a result signals are generated, collected and analysed to form an image. Interactions produce mainly SEs (secondary electrons), with lower energies compared to the primary ones, BEs (backscattered high-energy electrons) and characteristic X-rays[14, 115].
To produce an electron beam, a high voltage is applied to a filament, which can be either a thermionic emitter or a field emission source. The source brightness, $\beta(A/m^2sr)$, is given by the current density per unit of solid angle:

$$\beta = \frac{j_b}{\pi(\alpha)^2} = \frac{4i_b}{(\pi)^2 d_0(\alpha)^2} \tag{3.30}$$

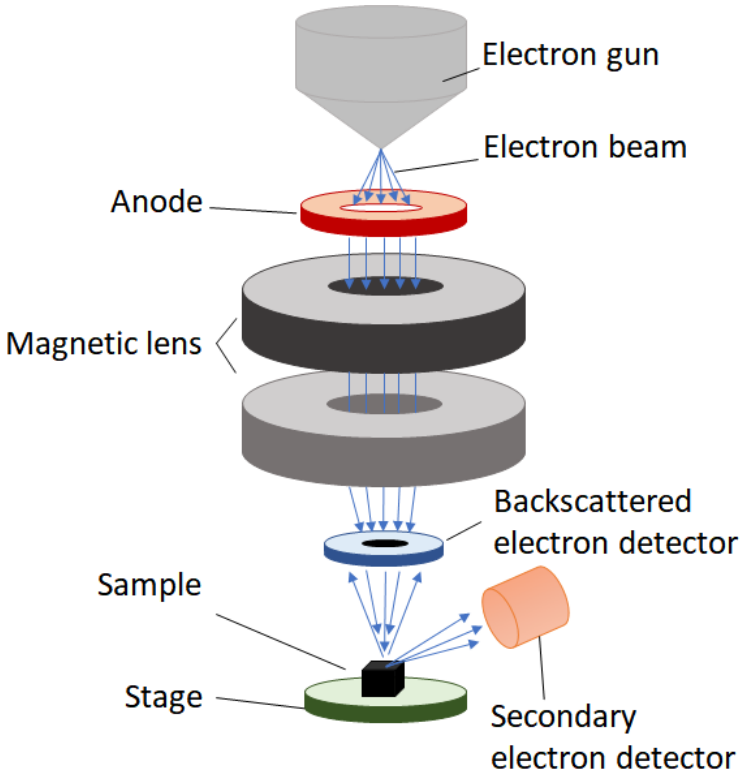


Fig. 3.10: Schematic diagram of a scanning electron microscope.

where j_b is the current density, i_b is the emission current, α is the convergence semi-angle an d_0 is the crossover beam diameter. In conclusion, the brightness increases linearly with the accelerating voltage V_0[115].
There are two types of electron gun in general use, thermionic emitters and Field emission sources.
Thermionic emitters are usually V-shape tungsten filaments or pyramidal lanthanum hexaboride, LaB_6, single crystal. They are heated, by passing a current through them, to over 25000°C to produce thermal emission of electrons from the tip. In fact, their work function Φ (the minimum thermodynamic energy needed to remove an electron from a solid,), is low: $\Phi_{tungsten} = 4.5e.V.$ and $\Phi_{LaB_6} \sim 2.5e.V.$. Since $\Phi_{LaB_6} < \Phi_{tungsten}$, the LaB_6 filaments provide larger maximum beam current which means a brighter beam. Field emission sources, know also as field emission guns or cold cathode emitters, do not require a heating but an high negative potential is applied until quantum mechanic tunnelling effect occur and electrons escape the surface work function barrier. They are brighter and more stable than thermionic emitters, with very small deviation in electronic energy [115, 50]. Moreover, they exhibit a

much narrower energy spread, and thus a larger fraction of emitter electrons are used. Since the electron extraction process is unconditioned by the final beam voltage, field emission sources are perfect for low-voltage imaging SEM. Compared to thermionic sources, field emission ones offer higher resolution ($\sim 1nm$), excellent performances over a wide range of accelerating voltages (0.5 - 30 kV) and higher resolution at low voltages. However, a better vacuum is required in this case to preserve the tip (10^{-10} instead of 10^{-6} for thermionic sources). Once the electronic beam is emitted, it passes through an anode to be accelerated and then two or three magnetic lenses demagnify the beam into a fine probe which is finally scanned across the specimen, for example we pass from a $d_0 \sim 50\mu m$ after the thermionic electron gun emission to a final spot size $d_0 \sim 10nm$. Even thou electrons can be focused by both electrostatic and electromagnetic fields, the aberration given by the latter are smaller which brings to their extensive application in SEM. Electromagnetic lenses consist of many thousands of cooper wire windings enclosed in a soft iron shell as seen in Figure 3.11a.

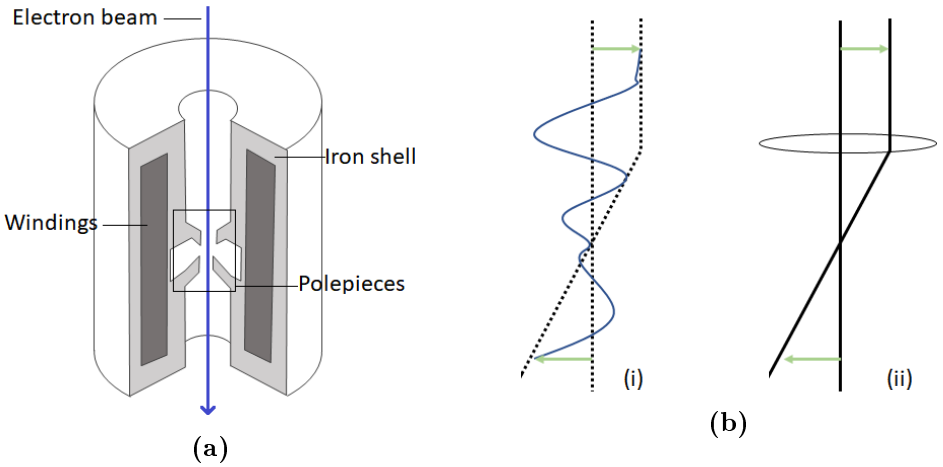


Fig. 3.11: a - Schematic diagram of an electromagnetic lens. The coil windings are inside the iron shroud and the field is produced across the lens gap between polepieces. b(i) - Diagram showing how electrons are focussed pass trough an electromagnetic lens. b(ii) - Conventional optical ray diagram for comparison.

When current passes trough the coil, an electric (E) and magnetic (B) field are generated and according to the Lorentz equation the force exerted on electrons is:

$$\vec{F} = e(\vec{E} + \vec{v} \otimes \vec{B}) \tag{3.31}$$

where $\vec{v}$ is the electron velocity in the beam. As electrons pass through the lens, they feel fringing magnetic field lines that extend across the lens gap

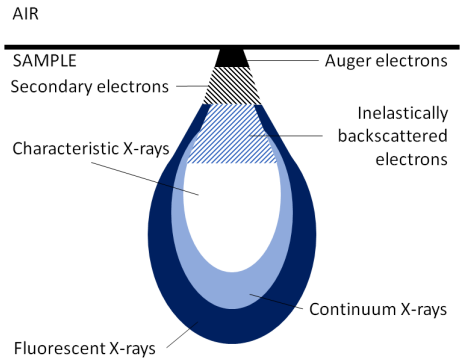


Fig. 3.12: Interaction volume of an electron beam with a thick sample.

between the polepieces and off-axis electrons begin to spiral toward the optic axis[45]. This is schematically shown in Figure 3.11b. When the current intensity is increased, the magnetic field force increases and the crossover point of the beam moves upwards. At the contrary, when the current decreases, the crossover point moves downwards.
As already mentioned, when the electron beam is scanned across the sample surface, electrons interact and penetrate into the specimen in a tear-drop shape3.12. The penetration dimensions depend on several factors including incidence angle, atomic masses of the elements in the specimen and primary electrons energy. The penetration depth increases the higher the beam energy, incident angle and the lighter the atomic mass (e.g. for a GaAs specimen with E = 20 keV at normal incidence, electrons penetrate $\sim 1\mu m$ into the sample) [50].
SEM imaging technique is based on SEs detection, which are usually electrons from or near the sample surface whose energy is typically below 50 eV[50]. Since SEs escape from a very shallow specimen area whose dimension is comparable to the beam size, they provide the highest spatial resolution images. Topographic information are obtained but, thanks to some BEs collected from the secondary electron detector, some compositional contrast is also present. BEs are those electrons that approach the atom nuclei sufficiently close to be scattered through a large angle and re-emerge from the surface. Their energy is higher compared to SEs one, but they are not as numerous. Moreover, since they come from a sample area bigger than beam size, their spatial resolution is lower. However, their have a key role in compositional information (elements of higher atomic mass give brighter signals) and in crystallographic information (because electron channelling occurs).Other secondary effects due to incident PEs energy loss can be produced such as X-rays, Auger electrons and visible light.

3.5 Neutron Reflectivity

Although all the techniques previously described are useful to analyse samples at air-water interface, the spatial resolution is not suitable to study deeply the mechanisms of film formation or to measure sample thickness or the chemical composition. Thanks to wavelengths comparable to interatomic distances (typical $\lambda = 2-30$Å), NR (neutron reflectometry) is a powerful tool for the surface behaviour investigation providing a spatial resolution below $\sim 1.0nm$ with penetration depth over $\sim 100nm$. First observed in 1946 by Fermi and Zinn[39] after the experimental discover made by James Chadwick in 1932[23], NR (neutron reflectometry) has specific properties that make this technique useful and interesting. In the present work the nomenclature of Russell[108, 109] will be used, for further details it is possible to consult its works.
Neutrons are subatomic particles with neutral charge, their electric dipole moment is either zero or too small to be measured and their mass is comparable to the proton mass. The absence of charge allows the neutrons to avoid any interaction with the electron cloud of an atom and interact, instead, with the atom core, in other words there's no coulombic interaction with matter. Since nuclei are several orders of magnitude smaller than atoms and nuclear forces are very short range (on the order of a few femtometers), neutrons penetration depth is high and they can travel large distances through most materials without being scattered or absorbed. This allows the study of volume effects and bulky components. Moreover, since neutron scattering is possible only with nuclei, NR is sensitive to the isotopic composition of matter.

Neutrons-nuclei scattering contains only s-wave components, which implies that the scattering is isotropic and can be described by the scattering length parameter, b (typically $b \sim 10^{-12}cm.$). It is a quantity that depends on the details of the interaction between the neutron and nucleus components. Both the sign and value of b change irregularly with Z (atomic number) and A (atomic weight), in contrast to x-ray scattering for which the atomic scattering length is a monotonically increasing function of Z[98]. This phenomenon is shown in Figure 3.13. A striking example is the case of hydrogen and deuterium which are two different isotopes of the same element whose scattering length values and even sign are completely different (b = -3.74 fm and 6.67 fm respectively). Moreover, b is different for the different spin states of the neutron-nucleus system,$I \pm \frac{1}{2}$, this difference is most significant for hydrogen nuclei, leading to incoherent (background) scattering from hydrogenated samples. With the scattering length it is possible to quantify the scattering

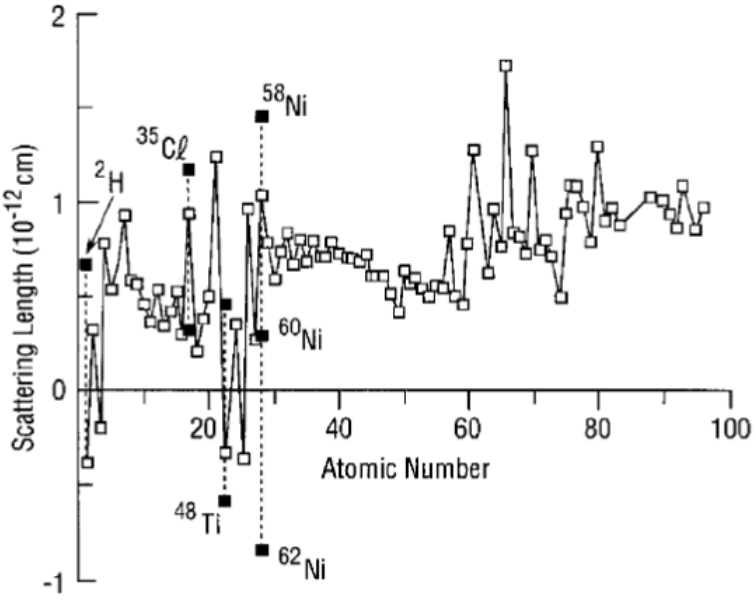


Fig. 3.13: Scattering length for different elements. The solid squares indicate values for isotopes.Reproduced from reference [98].

cross-section, σ:

$$\sigma = 4\pi b^2 \tag{3.32}$$

To measure the scattering power of a material, the SLD (scattering length density) parameter is used. SLD values depend on how tightly packed the nuclei are and how much scattering power each target nucleus has. SLD is defined as:

$$SLD = \sum_j b_j \rho_j \tag{3.33}$$

where b_j is the coherent scattering length of a nucleus and ρ_j is the atom density of the respective nucleus. SLD can be either positive or negative depending the sign of b_j.

3.5.1 Reflectivity from a single layer

In the present work, we'll discuss specular reflection, where the incident angle of the neutron beam is equal to the reflected angle, (Figure 3.14 $\theta_i = \theta_r = \theta$). Defining the $\perp$ direction as the direction normal to the surface, for specular reflection the scattering vector has only a $\perp$ component. This is the only scattering that occurs for samples or systems without any structure in the

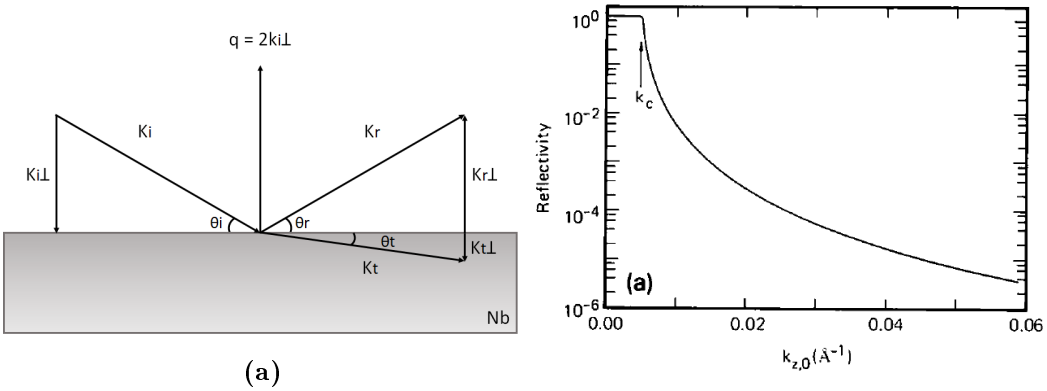


Fig. 3.14: (a) - Reflection of an incident beam from an ideally flat interface. ki and kr are the incident and scattered wave vectors, with angles $\theta_i = \theta_r = \theta$ in the incidence plane; q is the wave vector transfer; and SLD is the scattering length density of the semi-infinite substrate. (b) - Calculated neutron reflectivity profile for a beam incident on a silicon substrate where the surface is perfectly smooth. Reproduced from reference [108].$k_{z,0}$ is equal to the magnitude of the incident (and reflected) wavevector in air.

plane of surface (lateral structure).
Neutron reflectivity is usually described with neutron momentum transfer vector, q, since the scattering is elastic $|q_{incident}| = |q_{reflected}|$:

$$q_{\perp 0} = 4\frac{\pi}{\lambda}\sin\theta \tag{3.34}$$

$$q_{\perp 0} = 2k\sin\theta, k = \frac{2\pi}{\lambda} \tag{3.35}$$

where θ is the angle of incidence of the neutron beam with the surface and λ is the associated neutron wavelength, calculated from de Broglie equation:

$$\lambda = \frac{h}{m_n v} \tag{3.36}$$

where h is the Planck constant ($h = 6.626 \cdot 10^{-34} J \cdot Hz^{-1}$), m_n is the neutron mass and v is the neutron velocity. When the neutron beam coming from air impinges on a semi-infinite medium that consists of a single type of atom with a scattering length b, the neutron wave interacts with a constant potential V:

$$V = \frac{2\pi\hbar^2}{m_n} Nb \tag{3.37}$$

where Nb is the scattering length density. The neutron kinetic energy is:

$$E = \frac{\hbar^2 k^2}{2m_n} \tag{3.38}$$

If the kinetic energy associated to the perpendicular component of k is smaller than the potential V(i.e. $E_\perp < V$), the neutron energy is not enough to overcome the potential barrier and the neutron beam is totally reflected from the surface. It is possible to calculate the critical angle, θ_c, at which the total reflection occurs:

$$\frac{\hbar^2 {k_\perp}^2}{2m_n} = \frac{2\pi\hbar^2}{m_n} Nb \tag{3.39}$$

$$\frac{2\hbar^2\pi^2}{m_n\lambda^2}\sin^2\theta_c = \frac{2\pi\hbar^2}{m_n} Nb \tag{3.40}$$

$$\sin\theta_c = \sqrt{\frac{\lambda^2 Nb}{\pi}} \tag{3.41}$$

When the SLD has negative values, no total reflection can occur. If $E_\perp > V$, neutron transmission occurs. The transmitted beam energy corresponds to the energy of the incident beam minus the potential:

$$\frac{\hbar^2 {k_{\perp,t}}^2}{2m_n} = \frac{\hbar^2 {k_{\perp,i}}^2}{2m_n} - \frac{2\pi\hbar^2}{m_n} Nb \tag{3.42}$$

which gives

$${k_{\perp,t}}^2 = {k_{\perp,i}}^2 - 4\pi Nb = {k_{\perp,i}}^2 - {k_c}^2 \tag{3.43}$$

where k_c is the critical value of $k_{\perp,i}$.
If we consider the system from a classical optics point of view, incident radiation on a surface can be refracted and reflected due to the difference of refractive indices at the interface. In general, the refractive index of a specimen is slightly less than 1 and is given by[30]:

$$n^2 = \frac{k_t^2}{k_i^2} = 1 - \frac{\lambda^2 Nb}{\pi} \tag{3.44}$$

For most materials, $Nb << 1$, so for $\lambda < 60$ Åit is possible to approximate n as:

$$n \approx 1 - \frac{\lambda^2 Nb}{2\pi} \tag{3.45}$$

It is the change in the refractive index from one medium to another that cause neutron reflection. This will occur at the interface between two dissimilar materials.

3.5.2 Reflectivity from multiple layers

In a system containing multiple layers the reflection coefficient at an interface between the ith and $(i+1)th$ layer is:

$$r_{i,i+1} = \frac{k_{\perp,i} - k_{\perp,i+1}}{k_{\perp,i} + k_{\perp,i+1}} \tag{3.46}$$

In the case of air-substrate interface the reflection coefficient, $r_{0,1}$, is simply:

$$r_{0,1} = \frac{k_{\perp,0} - k_{\perp,1}}{k_{\perp,0} + k_{\perp,1}} \tag{3.47}$$

The reflectivity, R (which corresponds to the ratio of reflected to incident neutrons in layer 0.), is given by

$$R = r_{0,1} r_{0,1}* \tag{3.48}$$

where the asterisk denotes the complex conjugate. From (3.47),(3.43) and (3.48), the reflectivity for the air-substrate case is:

$$R(k_{\perp,0}) = \left[\frac{1 - [1 - (k_{c,1}/k_{\perp,0})^2]^{1/2}}{1 + [1 - (k_{c,1}/k_{\perp,0})^2]^{1/2}}\right]^2 \tag{3.49}$$

remembering (3.35) the reflectivity can be written also as:

$$R(q) = \left[\frac{q - [q^2 - q_c^2]^{1/2}}{q + [q^2 - q_c^2]^{1/2}}\right]^2 \tag{3.50}$$

When $q >> q_c$ it is possible to approximate the above expression as

$$R(q) \approx \frac{16\pi^2}{q^4} N_b^2 \tag{3.51}$$

Therefore, for an infinitely sharp interface, the reflectivity will vary with q^{-4} at high values of q. In Figure 3.14a the calculated neutron reflectivity profile for a silicon substrate with a perfectly flat surface is shown. Total external reflection occurs below $k_{\perp,0} = k_{c,1}$[108]. The scattering length density can fall off even more rapidly than q^{-4} in presence of diffuse interfaces or roughness at the air interface. These deviations are directly related to the width of the interface[108].
Consider a film of thickness d on a semi-infinite substrate, both having uniform scattering length density. In this case there are 2 step changes in the neutron refractive index, or the SLD, one at each interface. Each interface

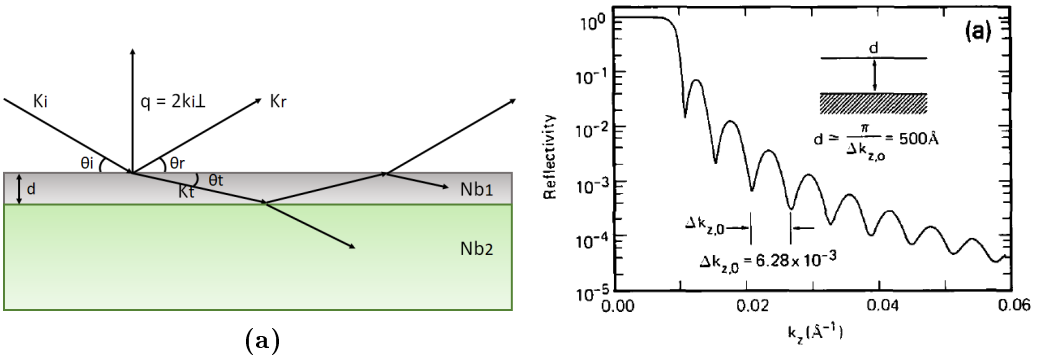


Fig. 3.15: (a) - Reflection of an incident beam from two ideally flat interfaces, one is a thin layer of thickness d on a semi-infinite substrate (with scattering length density Nb_2 and Nb_1 for the film and substrate respectively). (b) - Calculated neutron reflectivity profile for a 500 Åthick layer of perdeuterated poly(styrene), d-PS, on a silicon substrate. The oscillations are due to the thickness of the specimen. Reproduced from reference [108].

then has its own ('Fresnel') reflection coefficient, given by equation (3.46), at the air/film and film/substrate interfaces, separated by a distance d, as shown in Figure 3.15. The total reflection coefficient (reflectance) from the sample is given by summing the amplitudes of all the waves that emerge from the 0-1 interface following repeated reflection and refraction at the two interfaces. This summation results in a total reflection coefficient, r, given by:

$$r = \frac{r_{0,1} + r_{1,2}exp(2ik_{\perp,1}d)}{1 + r_{0,1}r_{1,2}exp(21k_{\perp,1}d)} \tag{3.52}$$

From (3.48), the reflectivity is

$$R = \frac{|r_{0,1}|^2 + |r_{1,2}|^2 + 2Re\{r_{0,1}r_{1,2}exp(2ik_{\perp,1}d)\}}{1 + |r_{0,1}r_{1,2}|^2 + 2Re\{r_{0,1}r_{1,2}exp(2ik_{\perp,1}d)\}} \tag{3.53}$$

in the case where the reflectances are real, (3.53) reduces to

$$R = \frac{r_{0,1}^2 + r_{1,2}^2 + 2r_{0,1}r_{1,2}\cos(2k_{\perp,1}d)}{1 + r_{0,1}^2 r_{1,2}^2 + r_{0,1}r_{1,2}\cos(2k_{\perp,1}d)} \tag{3.54}$$

This reflectivity shows an oscillation behaviour due to the cosine, which is directly related to the thickness d of the layer by:

$$d \propto \frac{\pi}{\Delta k_{\perp,0}} = \frac{2\pi}{\Delta q} \tag{3.55}$$

where Δq is the 'distance' in q space between adjacent minima in a plot of R(q). In Figure 3.14b a calculated reflectivity profile for a film, 500 Åin thickness, on a substrate where the interfaces between the film and the air and film and substrate are infinitely sharp is shown. The oscillations are due to the thickness of the specimen and their amplitude depends on the constrants contrasts at both interfaces[108]. It is possible to generalize the formalism above for the reflectivity calculation of any arbitrary number of layers. Considering the case shown in Figure 3.16: a n-layer specimen is on the top of the (n+1)th layer corresponding to the substrate. The 0 layer represents the vacuum/air surrounding the sample. The reflectance is calcu-

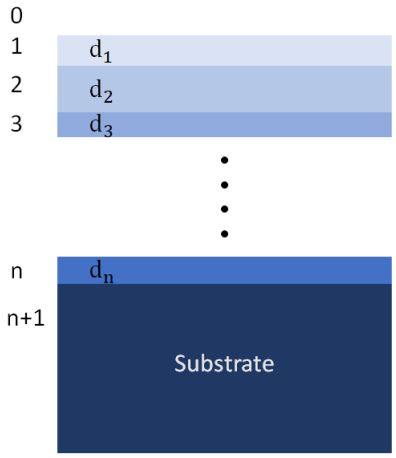


Fig. 3.16: Diagram of a n-layers specimen with variable thicknesses (d_n) on a substrate. The index n + 1 corresponds to the substrate, n to the layer immediately adjacent to the substrate, down to index 0 which corresponds to the vacuum/air.

lated starting from the substrate and the closest layer to the substrate, i.e. from the (n+1)th and nth layers respectively:

$$r_{n,n+1} = \frac{k_{\perp,n} - k_{\perp,n+1}}{k_{\perp,n} + k_{\perp,n+1}} \tag{3.56}$$

The reflectance by the (n-1)th and nth layer can then be written as:

$$r_{n-1,n} = \frac{r'_{n-1,n} + r'_{n,n+1} exp(2id_n k_n)}{1 + r'_{n-1,n} r'_{n,n+1} exp(2id_n k_n)} \tag{3.57}$$

where the prime indicates the reflectance given by equation(3.56), $r_{n-1,n}$ includes reflections at the (n-1)-n and n-(n+1) interfaces.Using equation(3.57) the reflectance at the (n-2)/(n-1) interface is calculated as:

$$r_{n-2,n-1} = \frac{r'_{n-2,n-1} + r_{n-1,n} exp(2id_{n-1} k_{n-1})}{1 + r_{n-2,n-1} r_{n-1,n} exp(2id_{n-1} k_{n-1})} \tag{3.58}$$

This recursion is then continued until the air/specimen interface is reached. Many methods have been established to solve the problem for a multilayer-stratified medium on top of an infinitely thick substrate, based on Parratt's recursive algorithm[94] and Abèle's optical matrix method[49]. In the optical matrix method described by Born and Wolf[18], a characteristic matrix M for reflection and transmission is defined for each interface:

$$M_i = \begin{bmatrix} \cos(k_i d_i) & -\sin(k_i d_i)/k_i \\ k_i \sin(k_i d_i) & \cos(k_i d_i) \end{bmatrix} \tag{3.59}$$

Multiplying these matrices for each layer in a system leads to the the total reflectivity curve for a series of layers with known properties. Using this method, different combinations of layers can lead to the same reflectivity profile, which is an effect of the well known phase problem caused by the fact that only intensities can be experimentally accessed, but not the phase.
The reflectivity can be calculated exactly for perfect layered structures, with sharp interfaces between layers. However, for most systems of interest the interfaces are not perfectly sharp[30], as a result of either material diffusion between interfaces (see Figure 3.17 a) or roughness with peaks and troughs over a large range of length scales (see Figure 3.17 b) [30]. As a consequence, both the rough and diffuse cases cause a faster decline of reflectivity with momentum transfer than the q^{-4} dependency described in Equation 3.51. There are several approaches to mathematically model roughness. The method used in the Motofit software (used in the present work to analyse the data) is to treat roughness as an error function $erf(\frac{z}{\sqrt{\sigma}})$ [90], where σ is the roughness measure, which corresponds to a multiplication of the reflectivity with a Debye Waller factor [30]:

$$R(q) \approx \frac{16\pi^2}{q^4} N_b^2 \cdot e^{-q^2\sigma 2} \tag{3.60}$$

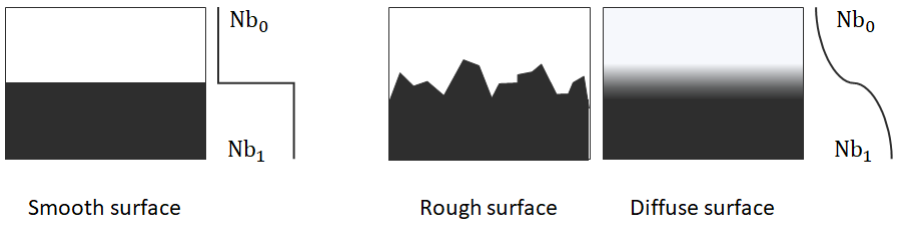


Fig. 3.17: Three surface profiles: on the left sharp interface, at the centre a rough surface, on the right a diffuse surface. The rough and diffuse surfaces can result in identical specular reflectivities.

3.5.3 FIGARO

In NR experiments, the specular reflected intensity is measured as a function of momentum transfer q:

$$\frac{R}{R_i}(q) \tag{3.61}$$

Remembering that the perpendicular component of the momentum transfer vector is given by

$$q_\perp = 2k_\perp \sin\theta = \frac{4\pi}{\lambda}\sin\theta \tag{3.62}$$

It is possible to vary $q_\perp$ by varying either the wavelength λ or the incident angle θ leading to two different methods of measurements: monochromatic reflectometry or TOF (time of flight) reflectometry. In the first case the wavelength is fixed and the incidence angle is varied during the experiments. One advantage of this technique is the possibility to chose the wavelength that maximizes the flux from the neutron source and therefore the resolution is the most efficient. Although this is a simple and very useful configuration for solid samples, where the incidence angle is varied by varying the sample position, for liquid samples θ can not be easily varied. In this cases, and many others, it is usually used the TOF reflectometry: the incident angle is kept constant during the measurements while the incident neutron beam contains all the possible wavelengths[108, 30]. In the present work, experiments have been conducted with neutron reflectometer called FIGARO (Fluid Interfaces Grazing Angles ReflectOmeter) at ILL (Institut Laue-Langevin), Grenoble France, with TOF configuration where neutrons strike the interface from above the sample, see Figure 3.18.

Once neutrons are produced by nuclear reaction (a continuous neutron beam

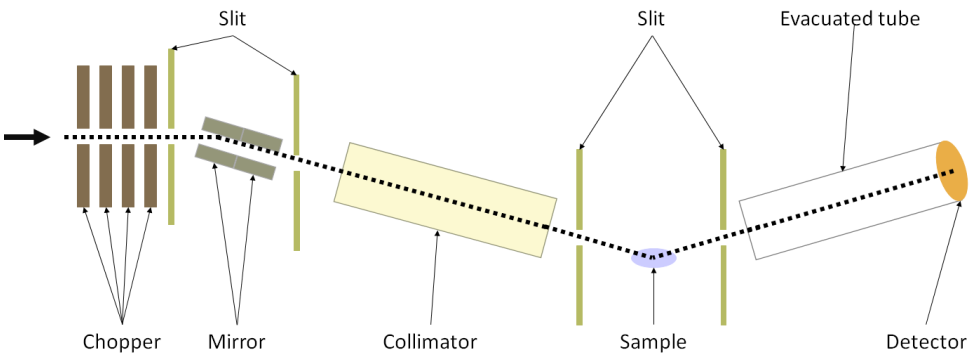


Fig. 3.18: Sketch of the reflectometer FIGARO.

is produced by nuclear fission), they travel with different velocities depending

on the associated energy and thus wavelength:

$$E = \frac{1}{2} m_n v^2 \qquad and \qquad E = \frac{2\pi^2 \hbar^2}{m_n \lambda^2} \tag{3.63}$$

then

$$v = \frac{2\pi\hbar}{m_n \lambda} \tag{3.64}$$

Since the distance between the source and the detector is known and all the neutrons start at the same time (choppers are used to cut a continuous beam into multiple pulses, thus a common starting time is determined), it is possible to measure the time at which each neutron arrives at the detector and calculate the associated wavelength. Therefore, neutrons are sorted according to wavelengths and, following that, momentum transfers upon interaction with the sample. In this way, it is possible to measure the reflectivity profile R(q) from neutrons reflected at the same incidence angle. In order to increase the q range, multiple angles are generally chosen for an experiment.
In principle, the wavelength resolution is fixed at:

$$\frac{\Delta\lambda}{\lambda} = \frac{\Delta t}{t} \tag{3.65}$$

However, the pulses duration is one of the limitations for resolution, as well as divergence beam and scattering background. In general, the background is subtracted from each measurement using the area detector. The range of q covered for a fixed angle θ is inversely proportional to the λ range, for the cold neutrons used at the ILL $\lambda \sim 1 - 30$ Å. For soft matter, the TOF technique is much more applicable than the monochromatic one, and data can be recorded much faster (hence more efficiently). The intensity at a given wavelength is less, but the efficiency is gained by measuring a broad wavelength range simultaneously, which can be helpful for samples that may change with time.
As already mentioned, neutrons are delivered to FIGARO by a nuclear reactor producing an intense continuous neutron flux (1.5×10^{15} neutrons per second per cm^2, with a thermal power of 58.3 MW). Since the energy of neutrons is too high to be used in large scale, it is important to direct the neutron beam to a moderator tank (e.g. D_2 at 25K) where neutrons have multiple scatterings losing energy during each elastic collision. In this way, neutrons reach the thermal equilibrium and their energy is of the order of $k_B T$, where T is the temperature of the water moderator, and k_B is the Boltzmann coefficient.
FIGARO is equipped with a 2D 500X150 mm^2 ^3He tubular aluminium monoblock detector. The resolution used for the experiments was 7% $\Delta\lambda/\lambda$. Each

sample was measured at two different incident angles of 0.62° and 3.8°, with a wavelength range of $\in$ [2Å, 30 Å] and thus the q-range $\in$ [0.009Å^{-1},0.2Å^{-1}].

4. PROTEIN PRODUCTION RESULTS

4.1 Production in Swansea laboratories

After a two-week period of training in Texas A&M University at Bondos' laboratories and following published protocols[80, 52, 96] it was possible to start Ubx protein production in Swansea laboratories for the first time. Producing a protein in new laboratories is not easy as the conditions change. To do so, it was necessary to adapt the protocols to our necessity.
The overall process, to obtain pure proteins, requires 5 days and is divided in two steps: bacteria growth and protein purification.

4.1.1 Bacteria growth and protein production

The first step is to incorporate the genetic material that codifies EGFP-Ubx protein, called plasmid, into the bacteria DNA. The plasmid is incorporated into the bacterial cell, which is called competent cell, through the cell membrane by the process known as transformation. In molecular biology, the genetic alteration is called transformation. For this work, EGFP-Ubx protein was cloned into a pET19b vector (plasmid) and transformed into Rosetta(DE3)pLysS *E. coli* competent bacterial cells using the heat shock technique. The bacterial cell membrane is made passively permeable to DNA by exposing it to cold conditions (temperature $\sim 0°C$) before being exposed to a heat pulse (42°C for 45 seconds), see Figure 4.1. The cells are then placed in LB agar plate with the antibiotics which select for the growth of those cells which have acquired the plasmid. For this work, two antibiotics are used all over the protein-production process: AMP (ampicillin) and CLAMF (chloramphenicol). CLAMF is used in bacteria growth to maintain pLysS plasmid into *E. coli* cells. This plasmid is essential to prevent early unwanted expression of Ubx, which is toxic to the chosen bacteria and it must be expressed only when an appropriate number of cell is grown. When the *E. coli* cells are transformed with the Ubx expression plasmid, the AMP

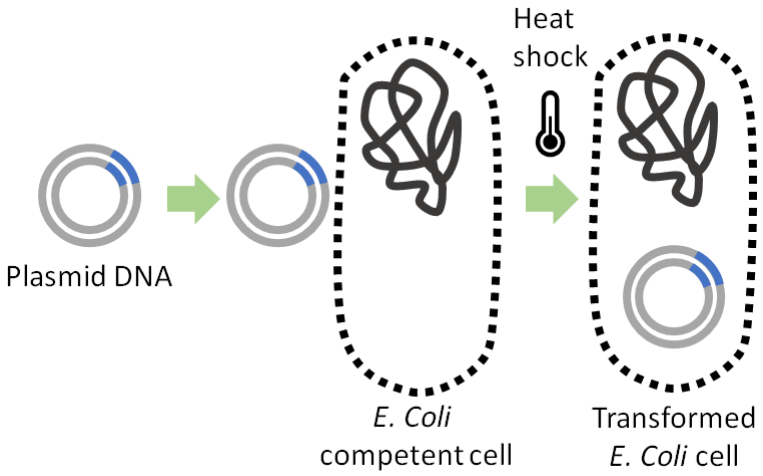


Fig. 4.1: Transformation technique scheme: using the heat shock technique, the cell membrane is made permeable to plasmid DNA which can enter into *E. coli* cell.

resistance gene is also conferred to those cells. This means that the successfully transformed *E. coli* are now AMP resistant. So this antibiotic is used to selectively eliminate bacteria that have not been transformed.

Once plated, the cells are put in incubator overnight and let grow forming

(a) **(b)** **(c)**

Fig. 4.2: (a) - *E. coli* colonies on LB agar plate. (b) - 400 mL of autoclaved LB with 0.4 mL of $50\mu g/mL$ AMP and 0.4 mL of $30\mu g/mL$ CLAMF (c) - Harvested bacteria cells using 30 minutes of centrifugation at 3400 rpm at 4°C.

colonies. At this point, it was possible to proceed with inoculation step. After the sterilization of 400 mL of LB (Langmuir-Blodgett), (10 g trypton, 5 g yeast extract, 5 g NaCl, 1 NaOH pellet in 1L distilled water) media with autoclave, 0.4 mL of $50\mu g/mL$ AMP and 0.4 mL of $30\mu g/mL$ CLAMF were

added following sterile procedure and 8 small pots have been filled with $\sim$ 2 cm media. Using one sterile pipette tip for each pot, some colonies were picked from the plate and put in each pot. Bondos' group prefers to pick one single colony for each pot, see Figure 4.2a. However, for the present work more than one colony was picked from the plate because it was assumed that all the colonies were the same. The pots were then left overnight in incubator at 37°C. The day after, 8 2L-flasks containing 1L of autoclaved LB media were inoculated with 7 mL of overnight *E. coli* cell culture from the pots (each flask was inoculated with cell culture coming from one pot) after the addition of 1 mL of AMP and CLAMP. It was important not to add more than 10% of the final volume (10 mL to 1 L) otherwise the cells would have reached mid-log too quickly and would not be vigorous enough to produce much proteins. The flasks used in Swansea are baffled, which are designed for culturing organisms that require a large surface area to volume ratio and guarantee a good ventilation. In Texas, flat-bottom flasks are used because the cell growth must be slow and controlled. For this reason, the flasks were placed in the incubator at 37 °C with a shaker speed equal to 180 rpm, instead of 200 rpm. *E. coli* cell culture was grown to an OD (optical density) between 0.6 and 0.8 at 600 nm. Measurements of OD were taken frequently over time, see Table4.1

Tab. 4.1: Optical Density check during *E. coli* cell culture. OD provides information on the growth of bacteria overtime.

Time	15.00	18.30	18.45
OD	0.00	0.59	0.79

When the mid-log was reached (period of growth of a bacterial cell population during which numbers increase exponentially) the flasks were put on ice for 30 minutes to cool, in this way proteolysis was minimized. Then, 1 mL of 1M isopropyl b-D-1-thiagalactopyranoside (IPTG) was added to each flask and the culture was placed again in the incubator at 26°C overnight. IPTG and cool temperature induce EGFP-Ubx production: transcription of the Ubx gene is controlled by the T7 promoter, a potent bacteriophage promoter and the protein cannot be transcribed unless the T7 polymerase protein is present. Expression of the T7 gene is controlled by the lac operon to add a manual on / off switch. Under normal growth conditions, endogenous lac repressor protein (LacI) binds the lac operon and prevents gene transcription. When transcription of the Ubx gene has to be activated, IPTG is used to bind the LacI protein, causing LacI to release DNA. Consequently, transcription can begin.

The day after, cells were harvested using 30 minutes of centrifugation at 3400 rpm at 4°C and pellets containing *E. coli* cells and EGFP-Ubx protein, were stored at -20°C until purification, see Figure 4.2c.

4.1.2 Protein Purification

Fig. 4.3: Column for protein purification with 10 mL nickel-nitrilotriacetic acid resine.

Each pellet was lysed in 40 mL of lysing bufer (50 mM sodium phosphate buffer, pH 8.0, 5% glucose w/v, 500 mM sodium chloride, 1 Roche Complete Mini Protease Inhibitor Cocktail tablet, 0.8 mg/mL DNase I) for 10 minutes and centrifuged at 18000 rpm at 4°C for 30 minutes. The resulting supernatant was loaded in a 10 mL nickel-nitrilotriacetic acid (Ni-NTA) agarose resin column (Fisher), that was previously equilibrated with 50 mL equilibration buffer (50 mM sodium phosphate buffer, pH 8.0, 5% glucose w/v, 500 mM sodium chloride). After the addition of supernatant (which contains both EGFP-Ubx proteins and other *E. coli* proteins), EGFP-Ubx attaches to the resin thanks to HIS-tags present at end of the protein and, in order to remove the spurious proteins coming form bacteria cells, the column was washed 4 times with equilibration buffer. Each washing solution contained an increasing amount of Imidazole, from 0 mM to 80 mM at pH 8.0. At the end of the process, 25 mL of elution buffer containing 300 mM Imidazole was added to released EGFP-Ubx proteins. 2 mL elution fractions were collected and some DTT (DithioThreitol) was added to prevent unwanted binding of Ubx and stored at -20°C.

4.1.3 SDS-PAGE gel

Once the purification step was completed, it was important to check the purity of the purified proteins with the SDS-PAGE (Sodium Dodecyl Sulfate - PolyAcrylamide Gel Electrophoresis). With this technique an electrical current is applied to negatively-charged proteins which run through a porous gel. Smaller proteins move faster than larger proteins in the gel because they

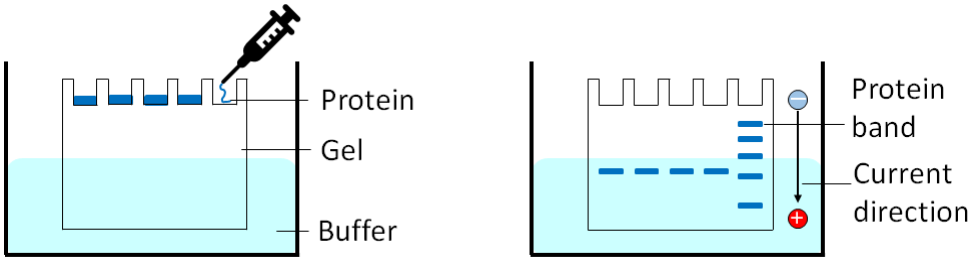
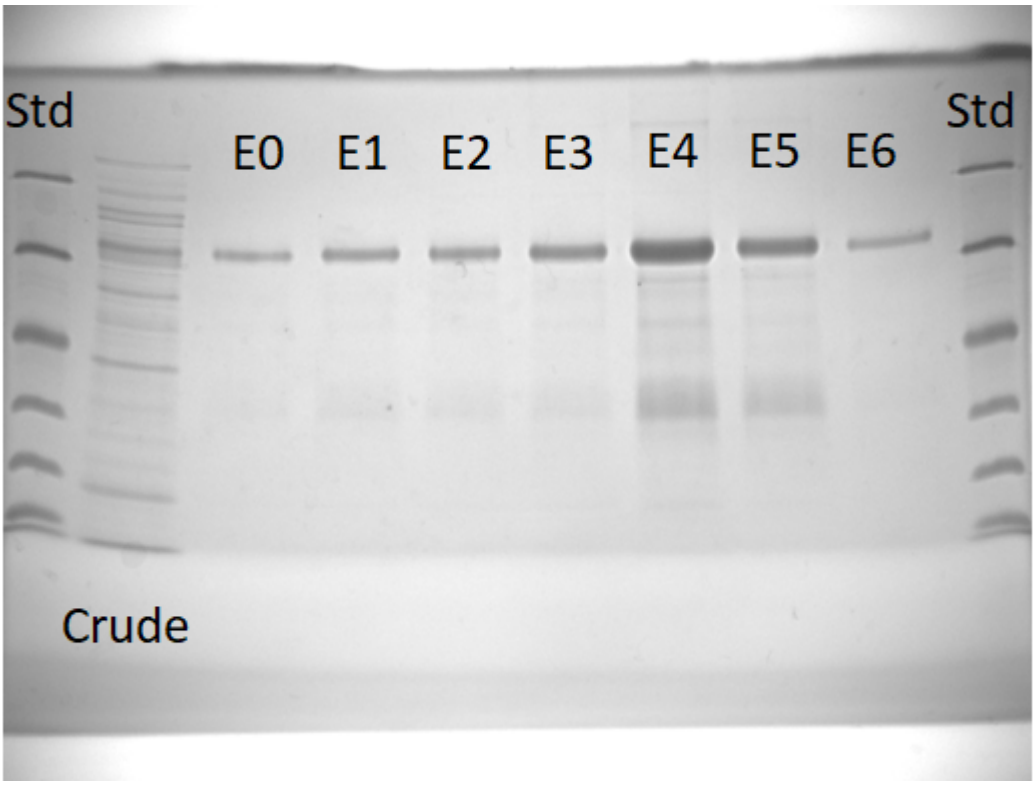
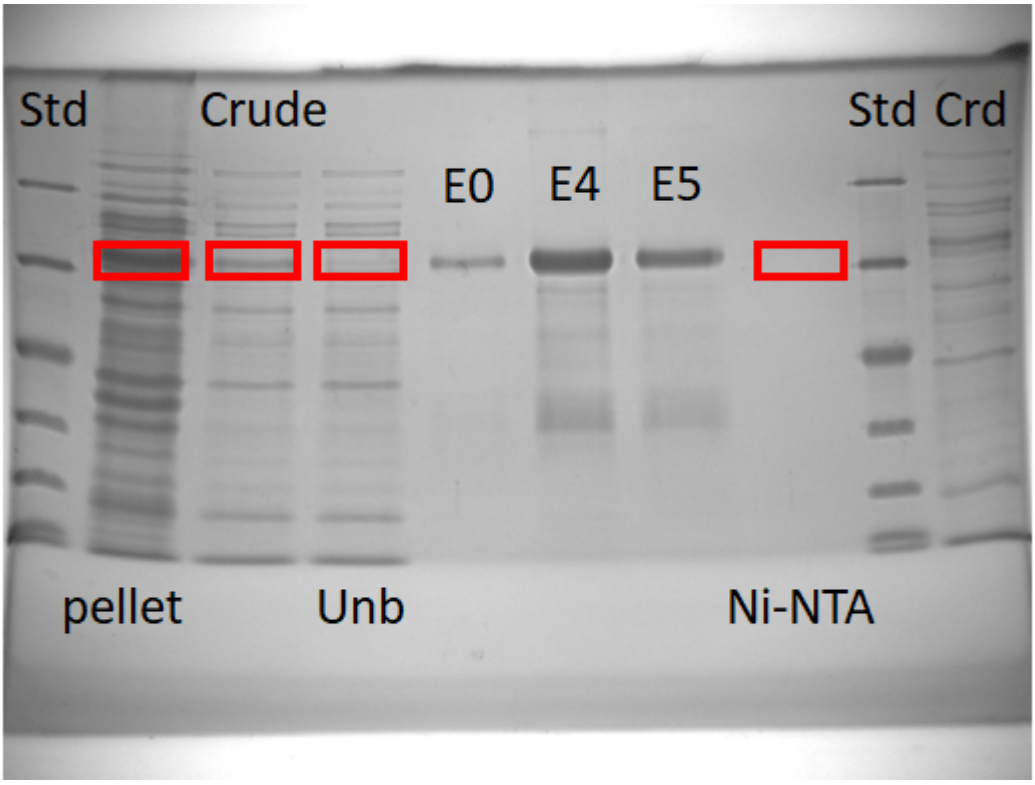


Fig. 4.4: SDS-PAGE of protein samples scheme. Proteins are loaded into gel walls and electrical current is applied. Proteins run through the gel and are separated by size.

can slip through the holes in the matrix easier. At the end of the process, proteins are separated based on size, see Figure 4.4. The determination of the molecular weight of proteins separated in gels takes place by comparison with a series of known size protein standards in the range of 10-116 kDa. Once the process is finished, it is possible to compare the position of the chosen proteins with the one of standard.
Protein were subjected to denaturing and reducing conditions so that they unfold completely and acted as long rods. The proteins were then coated with SDS, which made all the proteins negatively-charged, and loaded into the gel wells. In the present work, precast gels were used (Biorad). Loaded gels were place into the power supply box and voltage was set at 150 V for 70 minutes. Once the run was finished it was possible to image the protein bands using ChemiDoc Imaging System.
Two SDS-PAGE gels of the first protein production done in Swansea are shown in Figure 4.5. The first gel (Figure 4.5a) had the goal to check how pure the proteins were and so a standard, the crude (which was the supernatant loaded into the column containing both EGFP-Ubx proteins and other suspended proteins coming from *E. coli*) and all the elution fractions (E0, E1, E2, E3, E4, E5, E6) were loaded. From the image it was possible to observe that each elution contained proteins with molecular weight $\sim$ 70 kDa and, since EGFP-Ubx molecular weight is 69.21 kDa, thus EGFP-Ubx was present in the fractions. Moreover, in elution 1, E1, there were much less protein than in elution 4, E4. In E4 and E5 there was not only EGFP-Ubx band but also some other spurious one that indicated the internal digestion caused by the lysing buffer. The other fractions did not show internal digestions because the EGFP-Ubx concentration was lower.
With the second gel (Figure 4.5b) it was possible to check the protein purification process.



(a)



(b)

Fig. 4.5: (a) gel containing standard, the crude and all the elution fractions (E0, E1, E2, E3, E4, E5, E6); (b) gel containing the standard, the pellet, the crude, the unbound proteins, fractions E0, E4 and E5 and the Ni-NTA column matrix. The red rectangles underline EGFP-Ubx presence or absence.

The loaded proteins were: the standard, the pellet (the solid part discarded after the centrifugation before the column step), the crude, the unbound proteins collected during the column step when the supernatant was loaded, fractions E0, E4 and E5, the Ni-NTA column matrix. In the crude, EGFP-Ubx was present, as expected since it was the liquid collected before the column purification, but in the unbound the Ubx band was missing, meaning that the majority of Ubx proteins was trapped by the matrix. At the same time, Ni-NTA matrix did not show the EGFP-Ubx presence because after the column wash and the elution process all the proteins were released. However, in the pellet there was still some EGFP-Ubx, so the protein extraction was not totally efficient. Unfortunately this passage could not be improved because one possible consequence would be the protein degradation. In this case, it was preferable to lose proteins rather than have higher concentrations with degraded EGFP-Ubx.
At the end of the process, it was possible to conclude that the proteins were pure with a ratio over 95%, which was a great result especially because it was the first protein production done in Swansea laboratories.

4.1.4 Concentration check

The concentration of each elution was determined by calculating the concentration of EGFP protein, assuming that the EGFP:Ubx ratio was 1:1. Since EGFP has green fluorescence, the absorbance measurement was used, that is the measurement of attenuation of transmitted radiant power, caused by absorption, reflection, scattering and other physical process. Absorbance is linearly related to the concentration of sample, according to the

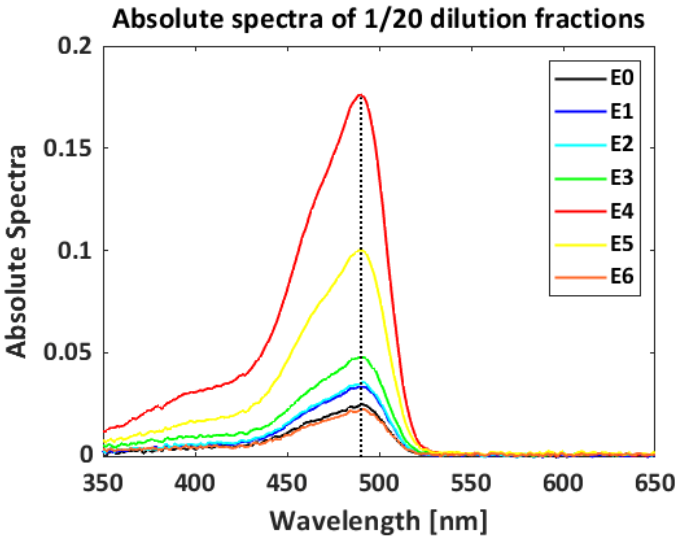


Fig. 4.6: Absolute spectra of 1/20 dilution measured for all the fractions collected during protein purification.

Lambert-Beer law, equation 4.1:

$$A = e \cdot c \cdot L \tag{4.1}$$

where A is absorbance (maximum value of absolute spectrum), e is the molar extinction coefficient of the absorbing species, c is the concentration of protein (our case EGFP) and L is the path length of light. Assuming that EGFP extinction coefficient is $e \sim 56 mM^{-1} sm^{-1}$ at 488 nm[120, 91], and the cuvette light path is 1 cm, we could measure the absolute spectra and calculated the concentration. Absolute spectra measured for all the fractions, collected after the first protein purification, are shown in Figure 4.6. The dotted black line helps to indicate the maximum value at $\lambda = 490$ nm. Once the maximum value of absolute spectra is measured, it is possible to calculate protein concentration as:

$$c = \frac{A \cdot 1/dilution}{e} \tag{4.2}$$

For example, in the case described in Figure 4.6, the maximum value for the fraction E1 is A = 0.033, so the concentration will be calculated as c = 0.033·20/56 = 0.0119 mM = 11.9 μM. In Table4.2 concentrations for the purified fractions are listed.

Tab. 4.2: Concentration values for the purified fractions.

Fraction	Concentration [μM]
E0	08.80
E1	11.90
E2	12.65
E3	17.02
E4	62.92
E5	35.76
E6	07.98

The method described above was optimal for EGFP-Ubx fusion as EGFP was a fluorescent protein.However, if plain Ubx were used or fused with non-fluorescent proteins, it would be impossible to calculate the concentration measuring absolute spectra, instead, Bradford assay was used. This was a simple procedure for determining the protein concentration by comparison with a pre-diluted protein whose concentration is known. In this work, BSA (bovine serum albumin) was use as comparison, because the molecular weight of both proteins was similar. This method was based on absorbance shift of

the dye "Coomassie Brilliant Blue G-250" when diluted with different protein concentrations.

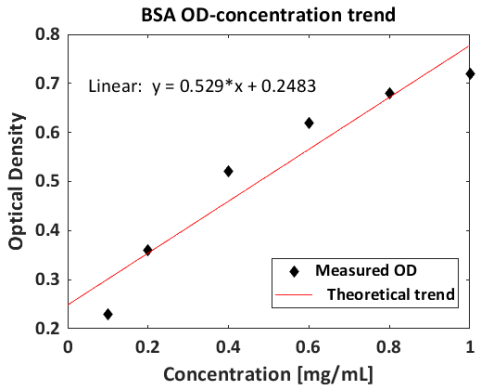


Fig. 4.7: Absorbance-concentration relationship for BSA proteins.

First of all, absorbance for known BSA concentration was measured at 590 nm and a linear absorbance-concentration trend (y = mx + q) was calculated, see Figure 4.7. Then, OD was measured for EGFP-Ubx proteins and the concentration was calculated, see Table 4.3. With this method, the protein concentration was calculated in mg/mL, in order to switch to μM it is necessary to divide the concentration by molecular weight.

Tab. 4.3: Table.

Elution	OD	c[mg/mL]	c[μM]
E0	0.46	0.44	06.7
E1	0.59	0.68	10.3
E2	0.62	0.74	11.2
E3	0.70	0.88	13.3
E4	0.83	1.10	16.6
E5	0.77	1.00	15.2
E6	0.44	0.42	06.4

Comparing the concentrations listed in Table 4.2 and in Table 4.3 it is possible to see that the two methods agree for small concentrations, but for higher values, such as for fractions E4 and E5, the two methods do not. This is due to the fact that the linear range of the assay for BSA in between 0.2 to 0.9 mg/mL. In our case, the concentration is higher so the values are not reliable. For this reason, we prefer to use the first method, when possible.

4.2 Demonstration of industrial-scale production of Ubx

In anticipation of a future increasing demand of Ubx proteins, it was decided to start the process of protein production in large volumes. In this regard, the entire process was reviewed and adjusted for the use of a fermenter. This is a container used to grow bacteria and fungi in large amounts, some fermenters have the capacity to hold thousands of litres. In the present work, a 15-L fermenter was used.
The transformation process was done as described in section 4.1.1, as this step can not be changed. Once the bacteria colonies were ready, it was possible to start the small growth. Instead of 8 small pots, 1L of autoclaved LB media was inoculated with some colonies picked from the LA agar plate, after have added AMP and CLAMP (1 mL of each). Bacteria were let grow overnight at 37°C shaking at 250 rpm. The following day, 6 L of LB media were prepared following the receipt and were put inside the fermenter where it was sterilised prior to using it. Once the media was autoclaved, it was let cool down to 37°C. Then 6 mL of each antibiotics and 7 mL of cultured LB media were added. In this case only $\sim 1\%$ of the final volume of bacteria cell culture was added because the fermenter was designed to optimize bacteria and fungi growth, facilitating the oxygenation and agitation and resulting in a faster cell growth. For the purpose of this work, however, it was important to not reach mid-log too quickly or *E. coli* would not be vigorous enough to produce proteins. Moreover, 10 mL PEG were added to avoid the foam formation. The parameters set were:

- temperature T = 37°C, yellow line in Figure 4.8;
- Air = fault(4), pink line in Figure 4.8;
- Agitation = 350 rpm;
- pH = 7.00, red line in Figure 4.8.

Once the inoculation was started, the OD was checked, see Table4.4.

Tab. 4.4: Optical Density check during *E. coli* cell culture.

Time	12.35	14.25	15.20
OD	0.00	0.32	0.68

While the last OD was measured, the culture was cooled down: in 10 minutes the temperature of the media reached 17°C, as it is shown in Figure

4.8. Then, it was possible to add IPTG to induce the EGFP-Ubx gene transcription and the temperature was set to 26°C for the overnight growth. The following day it was possible to harvest the cells. Unfortunately, the laboratory was not equipped with a large centrifuge and the harvesting process was done with a bench centrifuge that could process only 6 falcon tubes of 40 mL each, so the maximum volume used was 240 mL. To spin down 6 L of bacteria cells it took 3 hours. However the media was left on the bench for 3 hours, so the pellets weren't good and were thrown away.
Although the first attempt was unsuccessful, an other PhD student in our group was able to produce a large batch of EGFP-Ubx protein following the described procedure.

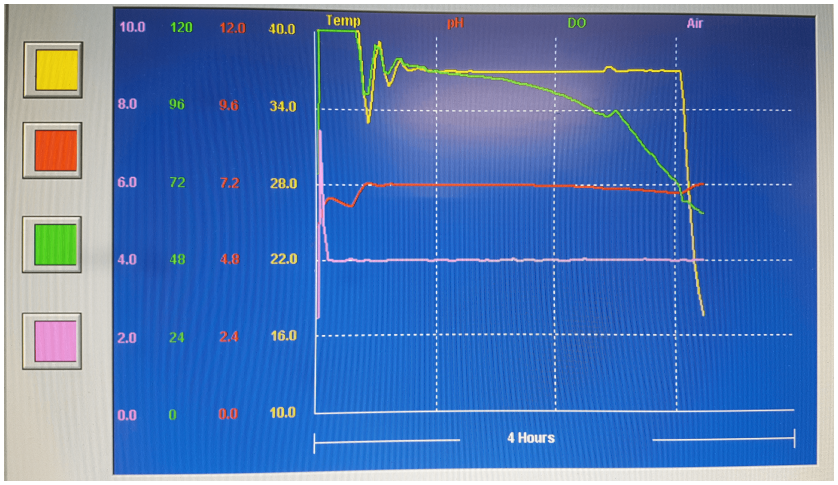


Fig. 4.8: Fermenter parameters during *E. coli* cell growth. Yellow line indicates the temperature, the red one indicates the pH and the pink the air flow.

4.3 Production in Grenoble laboratories

In August 2020, during the Covid-19 pandemic, a fire had broken out at Swansea University's campus and the laboratory, where EGFP-Ubx proteins were produced and stored, was destroyed. For this reason, the protein production was moved to Grenoble to produce at least one batch used to finish the experiments. The protocol used was the same described earlier and just few changes were made.
The transformation was done using an old plasmid that survived the fire in Swansea. The following day, 200 mL of LB media were inoculated with one single colony scratched from the plate and let grow overnight in the incubator at 37°C. Then 2 baffled flasks, each containing 1 L of media, were inoculated

and placed in the shaker at 37°C. Since the LB media used was a bit richer with nutrients to what the normal protocol suggested (15 g tryptone instead of 10 g), and in the shaker the flasks were producing too much foam, the speed was lowered to 60 rpm. From this point the OD was checked until the mid-log was reached, see Table4.5.

Tab. 4.5: Optical Density check during *E. coli* cell culture.

Time	9.50	11.50	13.50
OD	0.00	0.11	0.75

Comparing to the process done in Swansea, in this case the growth was a bit slower, but it was better than having a too fast growth. Once the flasks were let cool down and induced with IPTG, they were placed in the shaker at 26°C overnight. The day after cells were harvested and frozen at - 20°C until use.
For this purification, the lysing buffer was prepared using 4 aliquots of DNase I at 20 u/μL, 10 mM of manganese and magnesium, 2 pellets of complete, 0.5 mg of Pefabloc (instead of PMSF) and two spatula tips of lysozime. Once the pelled was lysed, it was centrifuged and the supernatant was loaded in the column. Unfortunately during the purification process the matrix was disturbed a bit and the proteins had been eluted during the last wash (G80 buffer) instead of during the elution step (G300 buffer). This brought a problem of pH, because G80 buffer had pH = 8.0 while G300 buffer had pH = 7.0, and a problem of salt concentration.
Soon after the purification, EGFP-Ubx concentration was calculated with BSA method because a spectrometer was not available, see Figure 4.9.

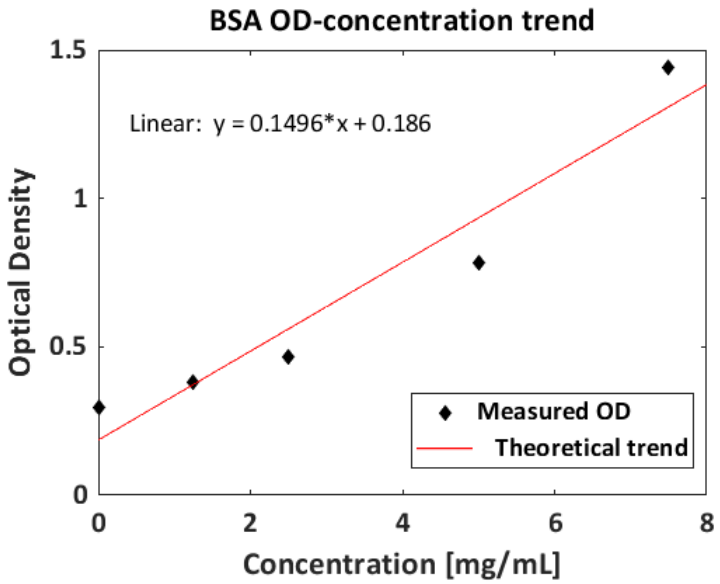


Fig. 4.9: Absorbance-concentration relationship for BSA proteins.

We decided to verify protein concentration with SDS-PAGE gel. The protocol used was the same used previously and the result is shown in Figure 4.10. In the gel it was decided to load the crude, the unbound, wash number 1, wash number 2, wash number 3, elution 1, elution 2, BSA (whose molecular weight is equal to EGFP-Ubx one) at concentration 1 mg/mL; 2 mg/mL and 3 mg/mL and the standard. As visible, there are no proteins in the three washes, as expected. The three columns with BSA indicated respectively one big band at around 70 kDa, which was the molecular weight expected. In elution 2, E2, it seemed there were no proteins while in elution 1, E1, there was a big band indicating a high concentration of proteins. However, the band indicated the presence of proteins with molecular weight $\sim$ 25 kDa. This meant that all the proteins produced were degraded and could not be used for the planned experiments and I was forced to use stored proteins.

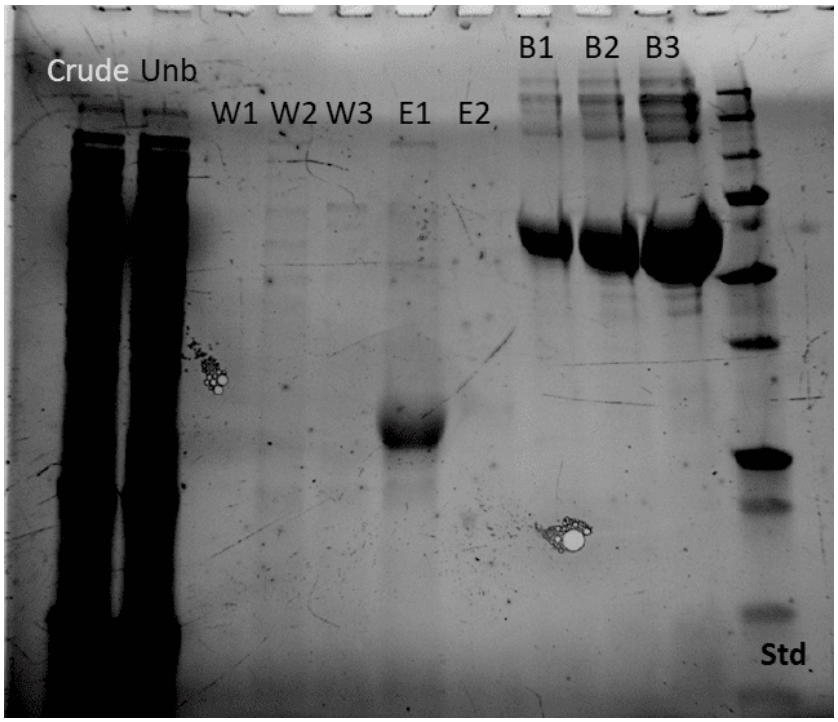


Fig. 4.10: Gel containing the crude, the unbound proteins, three washes W1, W2, W3, fractions E1, E2, BSA at three concentrations 1 mg/mL, 2 mg/mL, 3 mg/mL and the standard.

5. MACROSCOPIC RESULTS

As already mentioned, Ubx monomers spontaneously self-assemble *ex vivo* at air-water interface forming a biomaterial that can be reshaped (into micro-scale fibers or films) and used in many applications. Although both Ubx monomers and Ubx films have been highly studied and characterized (e.g. tissue engineering scaffolds [56]), film formation step has not been fully studied leading to a lack of understanding the mechanisms which bring monomers to aggregate into a film. This missing knowledge can allow us to optimize materials assembly and yield. In fact, mechanical properties can be tuned by altering the Ubx, eliminating tyrosines in the protein structure [57]. Moreover, the nature of these early stages can impact on film structure and behaviour of the resulting film. The concept Ubx study can be applied to all self-assemble proteins. In the present work, EGFP-Ubx fusion has been used.

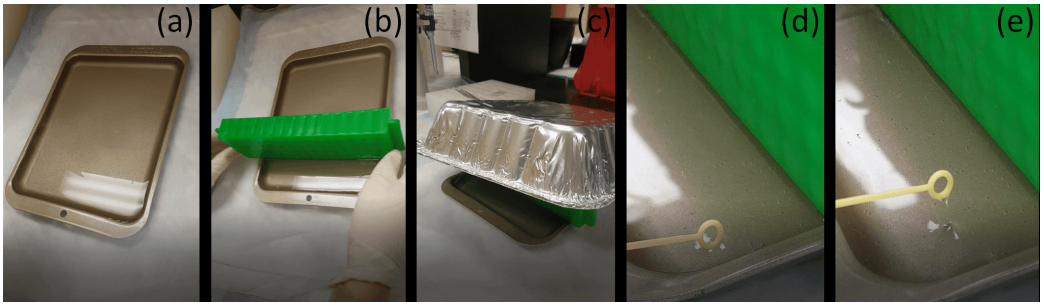


Fig. 5.1: Steps of EGFP-Ubx protein-based film formation. (a) - EGFP-Ubx is diluted in a Teflon-coated tray filled with G0 buffer; (b) - after a 16 h incubation period under an aluminium pan, with a microcentrifuge tube rack EGFP-Ubx film is gently squeezed reducing its area; (c) - the aluminium pan is placed again on top of the tray leaving the film to set for an hour; (d) - a plastic inoculation loop is used to touch the surface and pull fibers (e).

Since Bondos' group has been studied these materials for over 10 years and is the expert group in Ubx protein-based materials, my starting point was their works and protocols [15, 17, 16]. They generally use 2 mg ($\sim$ 15 nmol) of proteins spread in a metal tray filled with 600 mL of G0 buffer (0.5 M

NaCl + 0.05 M NaH_2PO_4 + 5% w/v D-Glucose, pH = 8.0) so that the liquid domes above the tray rim (Figure 5.1 (a)). Proteins form a film after a 6 hour incubation at room temperature. However, in order to have a more robust film, a 16-hours long incubation time is usually used. During the incubation time, the tray is covered by upside-down aluminium foil steam table to prevent any contamination from air dust. After this time, a film forms at the air-water and is gently squeezed with a microcentrifuge tube rack so that the available area is reduced (Figure 5.1 (b)). At this point the film is visible at the air-water interface resembling an oil slick. The aluminium pan is placed back at the top of the tray and the film is let sit for one hour (Figure 5.1 (c)) before a plastic loop is used to touch the surface (Figure 5.1 (d)) and pull fibers (Figure 5.1 (e)) or wrap films depending on the chosen application.
The goal of this work is to understand the early mechanisms of film formation and the first step toward this direction is the study of influence of external parameters upon film formation. This is the reason why the Langmuir technique has been used, since the film formation process used by Bondos' group is similar to the sample preparation used in Langmuir technique. This method enables us to measure the surface tension as a function of surface area at constant temperature, recording a $\pi - A$ isotherm.

5.1 Langmuir results and discussion

Using the Langmuir trough prevented me from following the protocol developed by Bondos' group, because the amount of proteins used is too high and the G0 buffer contains a high concentration of salts. As a result of the high amount of both EGFP-Ubx and salts, a saturation of π-A measurements at early stages is visible. It was important to avoid this condition as in this case it could not be easy to compare isotherms coming from different parameter changes. For this reason, we drastically reduced both the amount of proteins utilized and the concentration of salts and glucose in the buffer. Considering that no Langmuir studies have been conducted with Ubx protein or one of its fusion, I referred to previous works done with Bovine Serum Albumin (BSA)[119, 29, 43], whose molecular weight is comparable to that of EGFP-Ubx, 69.07 kDa.
The measurements of $\pi - A$ isotherms were carried out on a KSV NIMA Langmuir trough Large KN2003 (PTFE; total area 841 cm^2) equipped with a dipping well and a dipping mechanism. The Langmuir trough was cleaned prior to use with warm water and acetone to remove any trace of contaminants. The stock spreading solution was prepared by diluting EGFP-Ubx

purified proteins with Millipore water in order to reach a 10 μM concentration. The amount of added proteins in each experiment differs depending on the parameter-effect study.

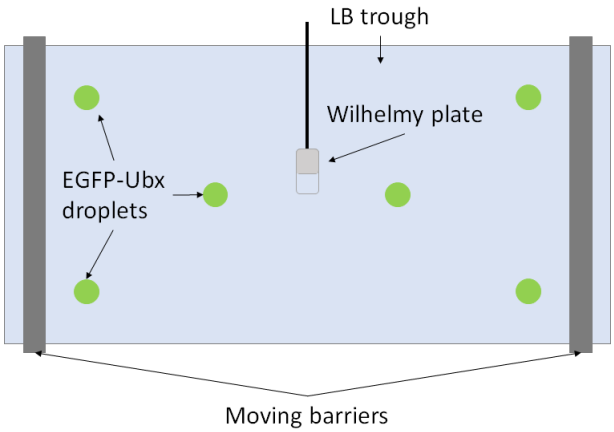
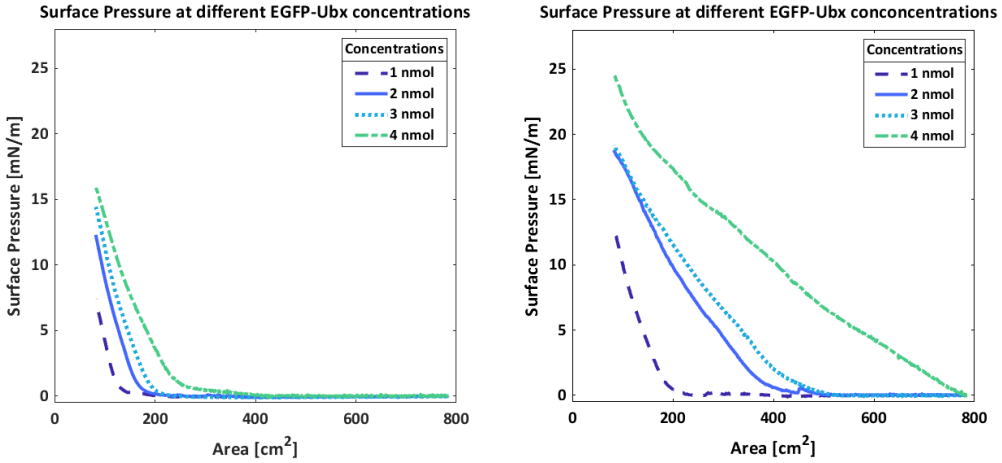


Fig. 5.2: EGFP-Ubx droplet disposition schema. Droplets are dripped at the surface gently in a symmetric way: three droplets per side, two droplets are placed near the moving barriers and one is placed near the Wilhelmy plate.

Fresh protein solution was dripped directly and symmetrically on the whole trough interface using a 1 mL pipette, three droplets per side, two droplets are placed near the moving barriers and one is placed near the Wilhelmy plate, see Figure 5.2. The droplets were carefully put in contact to the surface, avoiding any kind of disturbance, with the barriers fully opened. The spreading process plays an important role in film formation [22]. In this phenomenon inertia, gravity and surface tension are the driven forces, balanced by viscous one [46, 24, 10]. Initially, inertia and gravity dominate the spreading process and then, as the proteins expand at air-water interface, surface tension gradients act as dominant force [24, 10]. The phenomenon of fluid flow caused by surface tension gradient is called Marangoni spreading [46, 10, 21, 22]. When this phenomenon occurs, the resulting protein film at air-water interface remains stable for many hours (the duration of measurements) [21, 22] and EGFP-Ubx monomers eventually dissolve into the subphase and can adsorb slowly to the interface [21, 22].
Once the proteins were loaded, they were allowed to equilibrate for a variable time (chosen following Majithia's work [80]), called waiting time, before the compression was started. Films were obtained at a barrier compression rate of 25 cm^2/min at $pH \sim 8.0$ and room temperature (~ 25°C). Since the protein is partially soluble in water, we chose to use cm^2 as unity of measure for the Area, instead of $nm^2/molecule$.

At the beginning I focused on finding the perfect conditions to obtain $\pi - A$ isotherms with all the significant features (e.g. gas phase and liquid phase[29], as described in section 3.1) that would allow an accurate comparison across all the measured curves. In the present work, a film is considered formed when fibers can be polled at barriers fully closed. Starting with plain water and a standard waiting time, 10 minutes, I varied the monomer concentration from 1 nmol up to 4 nmol, Figure 5.3a. The waiting time was chosen following Majithia's work [80] (see Figure 2.3) where Ubx aggregates were observed after 15 minutes waiting time.



(a) (b)

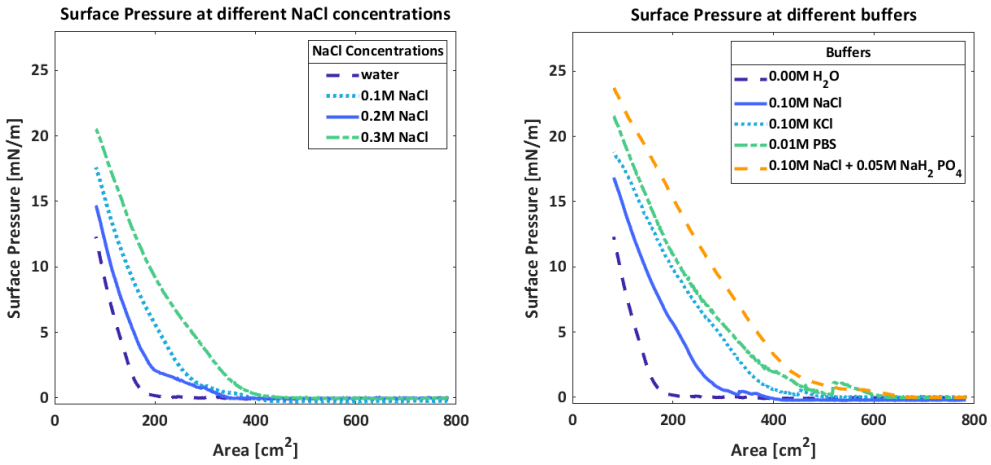
Fig. 5.3: Isothermal compression curve measured with 10-minute waiting time at different concentrations: 1nmol, 2nmol, 3nmol and 4nmol. The subphase used was plain water (a) and 0.1 M NaCl (b).

As can be seen, the isotherms collected using plain water as subphase are very similar to each other in shape and values for the four different EGFP-Ubx concentrations (even though for the 1 nmol concentration the Surface Pressure values are really small). They present the required features of gas and liquid phase, however, no film formed at the air-water interface. The absence of film formation could be explained in two ways: the concentration of proteins is too low and monomers interact forming small islands of proteins that can not laterally associate; plain water can not prevent the monomers precipitation leading to the formation of aggregates that disperse in the bulk reducing the amount of available proteins at air-water interface. In order to distinguish which of the two hypothesis is most relevant, I repeated the same experiment using, instead, a common aqueous solution for Langmuir technique, 0.1 M NaCl, and the same EGFP-Ubx monomer concentrations,

Figure 5.3b. NaCl, in fact, stabilises the protein conformation helping, in this way, film formation [15]. Thus, if no film formed at air-water interface, the problem was definitely the presence of inadequate number of proteins. The very first difference between the previous case is that $\pi - A$ values are higher in this case (we go from 15 to 25 mN/m), even though at 1 nmol in 0.1M NaCl the curve is comparable to the one obtained for 2 nmol in water. For this reason, in the case of 1 nmol of proteins spread in NaCl subphase, the hypothesis that the concentration is too low is the most reasonable. Moreover, at this concentration, it is impossible to wrap fibers or films and a monolayer was not visible to the naked eye. Thus, the formed film is not strong and compact enough. On the other hand, an increase in protein concentration results in higher surface pressure values and fibers could be drawn from the compressed films for the 2, 3, and 4 nmol concentrations. However the 4 nmol concentration creates a very strong material that can be easily pulled but it is is too close to the saturating the Langmuir trough system and the measurement cannot be stabilized (e.g. a film can be detected at the very beginning with all the barriers fully open, a further compression would lead to a more compact film). Thus, this concentration could not be used to study parameter effects since small variations would not be measurable due to saturation.. As the 2 nmol and 3 nmol concentrations behave very similarly, the 2 nmol EGFP-Ubx concentration was chosen to enable stable film formation while allowing maximum use of the dynamic range of Langmuir while varying other parameters.

The most relevant and important result from this first experiment was the ability to pull fibers after waiting only 10 minutes from the spreading of such low concentration of EGFP-Ubx. This result had not been achieved by Bondos group identifying a new phase of material formation that has never been studied. This upends everything we knew about the film formation process [80].

Once the protein concentration has been fixed, the next step was the study of influence on film formation at different salt concentrations trying to reach the same NaCl concentration (0.5 M) used by Bondos group with G0 buffer. Although I could not use 0.5 M NaCl buffer because at 0.4 M the system could not been stabilized and the measure was saturating, three dilutions have been explored using 2 nmol of EGFP-Ubx: 0.1 M NaCl, 0.2 M NaCl and 0.3 M NaCl, Figure 5.4a. In this case we can see a similar trend seen in Figure 5.3b with increasing $\pi - A$ values as the salt concentration increases. We decided anyway to proceed in my experiments with 0.1 M NaCl concentration because increasing the amount of salt does not bring major advantages and we preferred to follow previous works in protein film-formation at air-water interface [119, 110, 29].



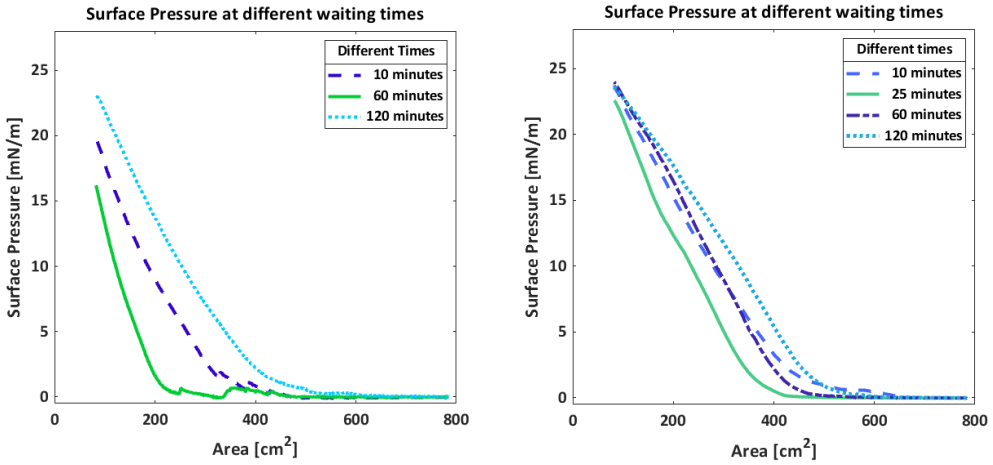
(a) (b)

Fig. 5.4: Isothermal compression curve measured with 10-minute waiting time spreading 2nmol EGFP-Ubx. (a) - the subphase used has and increasing salt concentration: water, 0.1 M, 0.2 M and 0.3 M NaCl; (b) - the subphase used has different salts diluted: 0.1 M NaCl, 0.1 M KCl, 0.01 M PBS and 0.1 M NaCl + 0.05 M NaH_2PO_4.

The following set of experiments has as a goal the study of influence of other salts present in buffers. The goal of buffer is the stabilization of intramolecular protein interaction, thus the out-competing of the intermolecular interactions whose output is aggregation. Cosolvents facilitate proper protein folding and solubility by either destabilizing aggregates or enhancing native protein stability. Among cosolvents, kosmotropes generally stabilize the native state of proteins without interacting with them. In fact, the solvent surface tension increases and kosmotropes are excluded from protein-solvent surface. Consequently, the entropic cost of disturbing the distribution of small molecules to form the solvent cavity increases. Since the native state of globular proteins usually creates the smallest solvent cavity possible, the effect of kosmotropes is to enhance the cost of hydrating intermediate or denatured states relative to the native state. These are crucial experiments to understand better the film formation mechanisms as non optimal external conditions would lead to alteration in protein structure and the consequent appearance of aggregates and precipitates. This phenomenon is one of the cause of properties loss, in our case self-assembly. To prevent this problem, it is important to study the effects of cosolvents upon protein solubility and aggregation. In order to chose the cosolvent to be studied, I referred to Bondos' article[15] in which a list of possible cosolvent for Ubx are listed. The

effect of specific cosolvent can be seen by direct comparison with other co-solvents. In this work water (as a contrast), 0.1 M NaCl, 0.1 M KCl, 0.01 M PBS (which contains sodium phosphate, sodium chloride, potassium chloride and potassium phosphate) and 0.1M NaCl + 0.05 M NaH_2PO_4 buffers were used (see Figure 5.4b). The cosolvent chosen are acting as kosmotropes in a range between 0 - 1 M for Ubx monomers and their fusions even though sodium phosphate monobasic (NaH_2PO_4) is not classified as cosolvent but helps the thermal stability of the protein. As can be seen in the graph, the shape of the curves changes as well as the $\pi - A$ values, meaning that fewer proteins aggregate and more proteins are able to reach the air-water interface to form a more compact film. It is clear that potassium chloride can stabilize better solvent-solvent interactions with respect to sodium chloride, in fact surface tension is higher for the same area values. At the same time, PBS effects are comparable to KCl one although its concentration is smaller. This can be described by the fact that PBS contains both NaCl and KCl plus their phosphate versions, thus we have a combined action of different co-solvents. The most unexpected result relates to the sodium chloride plus sodium phosphate buffer. I chose the same concentration for NaH_2PO_4 as in the G0 buffer used for both EGFP-Ubx purification and film-formation by Bondos' group. We can see that the film is more compact and a film is detected already at 600 cm^2, as it is for PBS buffer, but, contrary to what happens to latter, surface pressure does not drop to zero but monotonically increase. A last observation on Figure 5.4b can be done on the appearance of sawtooth during gas phase for all subphases, except water. We hypothesised that they were caused by EGFP-Ubx islands sliding over each other during the barrier closure, they rotated away from it. So, the increase in π-A values didn't indicate a film formation rather a collision of protein-base islands.
After having explored the effects of cosolvents, we decided to proceed the study using two aqueous solutions: 0.1 M NaCl and 0.1 M NaCl + 0.05 M NaH_2PO_4 because the first one gives film that can be pulled even with small concentration of salts and the last one because it is the combination of salts that gives the strongest and more compact film.
The last set of experiments done with Langmuir technique was focused on the exploration of waiting time dependence, Figure 5.5a and 5.5b. This parameter has never been taken in account as until now the proteins has been left to form film for hours in order to obtain an organised, compact and strong film. At the beginning I explored the film formation using 2 nmol of proteins dissolved in 0.1 M NaCl solution. Following Majithia's work[80], the film have been left to equilibrate for 10 minutes, 60 minutes and 120 minutes, Figure 5.5a. Interestingly it is possible to see a non-monotonic film-growth varying the equilibration time. At a given trough area the surface pressure

first decreases from 10 to 60 minutes, and then we see an increase in $\pi - A$ values. In order to prove the generality of this phenomenon, the same experiment have been conducted using a different buffer, 0.1 M NaCl + 0.05 M NaH_2PO_4, see Figure 5.5b.



(a) (b)

Fig. 5.5: Isothermal compression curve measured varying the waiting time parameter (10 minutes, 60 minutes and 120 minutes) and spreading 2nmol EGFP-Ubx with (a) - 0.1 M NaCl; (b) - 0.1 M NaCl + 0.05 M NaH_2PO_4.

In this case the waiting times chosen are not enough as at 10 and 60 minutes the $\pi - A$ isotherms are comparable, loosing the decreasing effect. After two hours, instead, the surface pressure values are higher as expected. So, the non-monotonic effect disappears if we consider only these three measures. For this reason, I decided to explore film-formation at 25 minutes too. From the graph it is visible that at the latter waiting time we find again the non-monotonic effect. Thus, the time frame in which the descending-ascending behaviour is altered.
Our hypothesis is that at the early stages of film formation (e.g. minutes) the film is partially disorganised as a lot of islands of fibrils are formed at air-water interface all with different orientations in 2D space, as observed by Majithia et all [80]. When the trough barriers are closed, they have not got enough time to interact and orientate and, being forced to stay in a small area, they form a film that can be anyway pulled into fibers even without a proper periodicity. As the time passes, monomers interact also with the buffer and cosolvent and amount of material adsorbed at the air-water interface decreases over time (e.g. 30 to 60 minutes, depending on cosolvents chosen). In this time frame the ability to pull fibers is reduced and no sam-

ples can be collected because there are no fibers or because the fibers are too thin to be collected. For longer waiting times (e.g. 120 minutes), the amount of monomers at the surface increases and the film structure results well organised as the fibrils islands have enough time to interact and to rotate so that the fibril result orientate with each other before the film is compressed [80]. For this reason, $\pi - A$ isotherms indicate a more compact film.

5.1.1 Limiting Area analysis

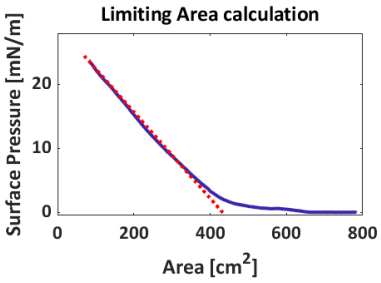


Fig. 5.6: Representation of limiting area projection.

When the trough barriers are closed, the $\pi - A$ isotherm show the transition between different phases: from the gas phase, where available area per molecule is large and the proteins do not interact with each other, the film go to the liquid phase, where the proteins are forced to become more closely packed and interact with each other until they are close to their maximum packing. Following what is usually done in electronics with cut-off frequencies, the transition point between gas and liquid phase, called Limiting Area (LA)[119, 128, 38, 133], can be calculated from Langmuir curves by projecting the linear increasing part of the curve to x-axis, see Figure 5.6. With this parameter it is not possible to predict the film formation behaviour, since it is strongly dependent on the solution used. However, it is possible to use LA values to predict the quality of the films at given parameters.

Tab. 5.1: Limiting Area values [cm^2] obtained by the extrapolation of the linear part of the $\pi - A$ curve to zero pressure.

Water	0.1 M NaCl			0.1 M NaCl + 0.05 M NaH_2PO_4		
10 min	10 min	60 min	120 min	10 min	60 min	120 min
128 ± 1	271 ± 2	209 ± 3	417 ± 4	448 ± 4	433 ± 5	489 ± 3

LA values have been calculated for the different waiting times and the two aqueous solutions described above, as listed in Table5.1. In general, it is possible to say that for both aqueous solutions LA follows the non monotonic effect seen for $\pi - A$ isotherms.

5.1.2 Fiber collection

With Langmuir technique it is possible to obtain macroscopical information about the resulting film by comparing different $\pi - A$ isotherms and to make some hypothesis regarding the film organization and structure. However, in order to have a complete comparison between the results obtained with Langmuir trough and what have been fully characterized by Bondos' group, it was important to collect fibers from each Langmuir experiment, in particular for the waiting-time comparison.
Following standard procedures, using a 10μL sterile plastic inoculation loop, EGFP-Ubx fibers have been collected at the end of film compression so that all the films had in common the same available surface area[83]. By gently touching of the plastic loop on the air-water interface, it is possible to pull fibers that can be wrapped around it. The wrapped fibers were left dehydrate slowly in a sterile plastic Petri dish sealed with parafilm. The plastic loop was also fixed at the surface of the container. In this way any violent shaking was prevented (that could break the fibers) and at the same time any air dust contamination was avoided. The Petri dish was stored at room temperature until use. First of all, we took fibers using 2 nmol of EGFP-Ubx proteins spread in 0.1 M NaCl subphase using the Langmuir trough at different waiting times. We were able to collect fibers at 10 and 120 minutes waiting time, but none at 60 minutes. Fibers were analysed using SEM (1 kV voltage and 5 µA current with a 2.5k magnification) to check the fiber structure (see Figure 5.7). Although the two diameters are comparable (see Table 5.2), the first visible difference we can see in the fiber structure is the wrinkles alignment: when the waiting time is high, 120 minutes, the fiber presents a well aligned structure with uniform parallel wrinkles (Figure 5.7 (b)). The latter have also comparable dimensions. However, when the waiting time is drastically reduced, 10 minutes, fibers present fewer wrinkles, if compared with the previous case, with different dimensions (see Figure 5.7 (a)). It is clear that the different waiting times have an impact on fiber structures and, in general, on film organization. For this reason, in order to analyse the differences in film formation when varying the waiting time, measurements using the G0 buffer were carried out using the Langmuir trough. In this way, thanks to the high density and viscosity of the buffer, fibers could be collected also in the descendent portion of Langmuir data (using the other types of buffer would not have gave us the opportunity to). Moreover, using G0 buffer enables the comparison with earlier results from the Bondos group [60]. As already mentioned, the $\pi - A$ isotherms cannot be used since the instrument was saturated, but the Langmuir trough was used for controlled compression and not to get measurements. Using G0 buffer was not only en-

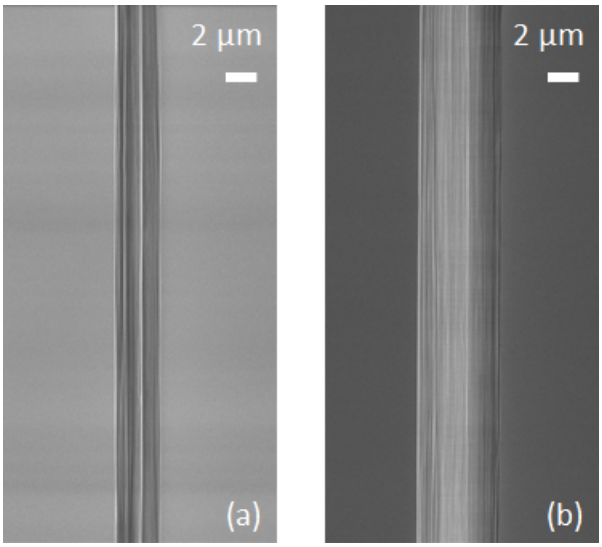


Fig. 5.7: Fibers collected using 2 nmol EGFP-Ubx proteins spread in 0.1 M NaCl buffer using (a) - 10 minutes waiting time and (b) - 120 minutes waiting time. It was not possible to collect fibers from 60 minutes waiting time.

abling us to have a good comparison with what have been previously done, but also allowed us to . Images taken with SEM are visible in Figure 5.8. The first thing that stands out is the huge difference in fiber diameter (see

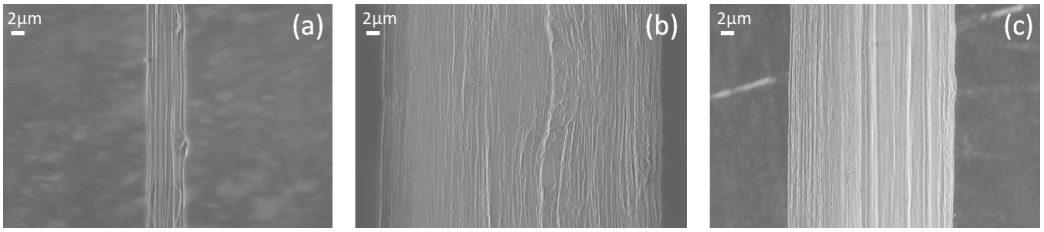


Fig. 5.8: Fiber images obtained using Langmuir trough with 2 nmol Ubx in G0 buffer and using (a) - 10 minutes waiting time, (b) - 60 minutes waiting time and (c) - 120 minutes waiting time. All the images have been taken using a Scanning Electron Microscope with 1 kV voltage.

Table 5.2), d: for 10 min $d = 4.5\mu m \pm 1\mu m$; for the 60 minutes waiting time parameter its value is $42.8\mu m \pm 2.8\mu m$ and 120 min-waiting-time diameter it is equal to $25.5\mu m \pm 2\mu m$. Although the amount of spread proteins (2nmol) and the buffer (G0) were the same, the material formed at air water interface is different meaning that the waiting-time parameter extensively affects the protein organization.
The second visible difference in all three images is the wrinkle alignment.

Tab. 5.2: Mean values of diameters, d, [μm] for the fibers collected with Langmuir trough at different waiting times, t, and subphases, 0.1 M NaCl and G0 buffer. Missing values are reported with the symbol '-', if it was impossible to see any.

	0.1 M NaCl			G0 buffer		
t [min]	10	60	120	10	60	120
d [μm]	3.9 $\pm$ 0.2	-	5.2 $\pm$ 0.4	4.5 $\pm$ 1.0	42.8 $\pm$ 2.8	25.5 $\pm$ 2.0

At 120 minutes we can see a highly ordered pattern with large and well defined wrinkles (whose dimensions are in the range of 0.7-1.7μm). On the other hand, fibers obtained after 60 minutes present a very disordered pattern where wrinkles are non-continuous and winding. In the middle of these two extremes, there's the 10 minutes waiting time case, whose fibers present some wrinkle-formations that are not perfectly aligned. Comparing what we have obtained with Bondos' analysed fibers[80], we can conclude that 2-hours waiting-time fibers are really similar to what already studied, but the remaining two fibers are different in diameter and/or in wrinkle structure.

5.1.3 Matlab data analysis

Once these studies had been concluded, we decided to go forward and analyse samples taken using 0.1 M NaCl and 0.01 M NaCl + 0.05 M NaH_2PO_4 solutions with an in-house developed MATLAB code. In this way it was possible to obtain quantitative structural information from an image. The goal of our analysis was to measure the alignment and orientation of wrinkle structure in the collected fibers. The concept on which the analysis is based is the axial autocorrelation and the full width at half maximum (FWHM) height, see Figure 5.10.
Once the fiber image is loaded in MATLAB and the wrinkle orientation in defined (fixing the y-axis), a threshold is set based on the background to create a binary mask of the fiber. The filter removes small areas captured by the threshold because considered noise. On the other hand, in order to remove the background noise, all the intensities outside the mask area on the original image are set to zero. At this point it was essential to use an image flattening algorithm to separate and differentiate the fibrils (Figure 5.9). Assuming a cylindrical shape for UBX fiber, the fiber visible part is considered to be a semicircle with the diameter equal to the width of the fiber. The centre point of the fibre was defined as the midpoint between the two widest points on the fibre. From that centre point, pixel intensity values are expanded horizontally as if the cylinder was being unrolled into a 2D

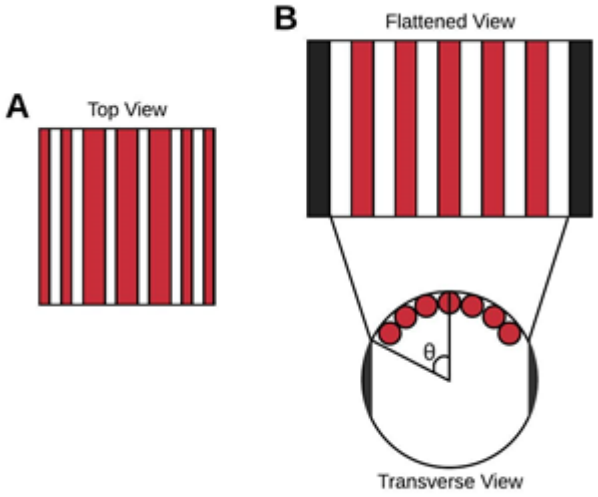


Fig. 5.9: Fiber flattening diagram. The view from the SEM (A) can be flattened by assuming a cylindrical shape and calculating arc lengths. The blacked-out regions are outside of the $\theta = 60°$threshold.

representation of its surface area. This is done by the width of a pixel to an arc length on a circle defined by the distance of the pixel from the centre line using the equation 5.1

$$s = \theta d \tag{5.1}$$

where θ is the angle between the virtual centreline of the fibre and the pixel of interest (Figure 5.10b) and d is the fiber diameter. In order to define the width per pixel, we used the image resolution, 20 nm/pixel. The width per pixel parameter is important for the calculation of the arc length between a pixel and its neighbour, based on Equation 5.1. This arc length is converted back into pixels and rounded to the nearest whole value. The image intensity value at the original pixel is then repeated as many times as necessary to match the length of the arc. This was repeated for each horizontal line of pixels. When the radius of the fiber is much larger than the radius of the wrinkle, we can cut off the portion of the image that extends beyond a 60°arc in each direction of the origin (Figure 5.10B).
At this stage, the wrinkles are indistinguishable from each other and a vertical autocorrelation and a wavelet transform help visualize the structure uniformity. The vertical autocorrelation is nothing but the translation of one image over the other (Figure 5.10(a)). After each step of the translation,

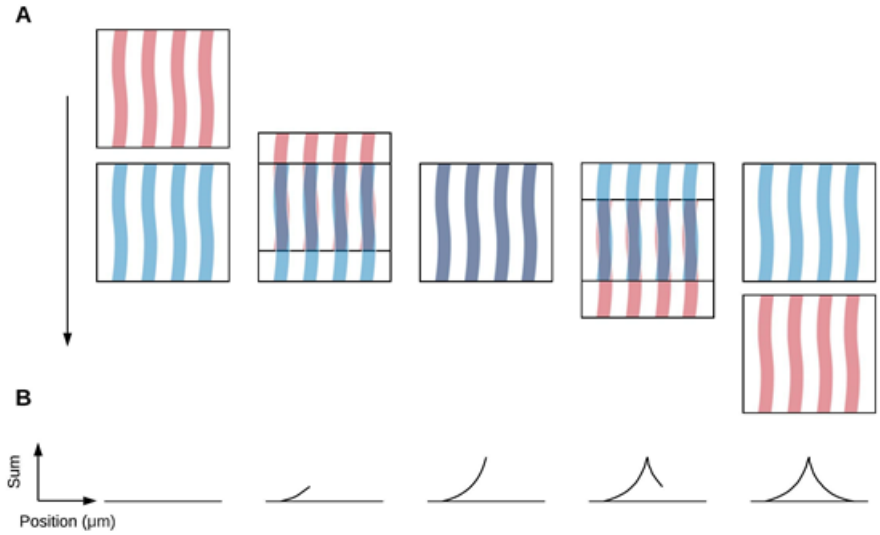
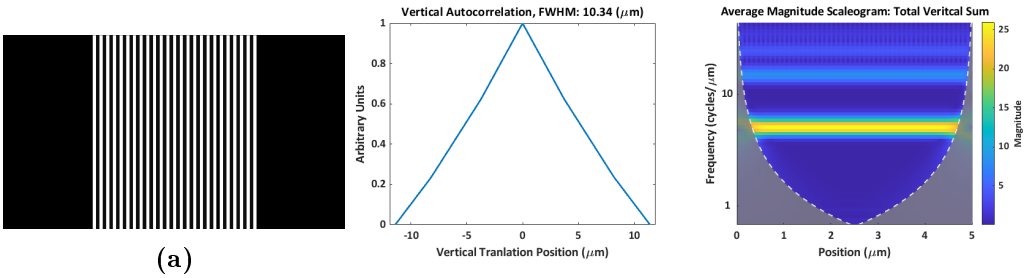


Fig. 5.10: Diagram vertical autocorrelation (A) along with the graph of the computed sum values (B).

each individual pixel is multiplied with its overlapping counterpart in the copied image. The resulting numbers are then summed across the portion of the image that is overlapped. This sum is plotted against the position of the translating image (Figure 5.10(b)). The full width at half of the maximum height (FWHM) of the graph was used to quantize the alignment. If the fibrils were well aligned (see Figure 5.11a), they would stay on top of themselves for a large portion of the autocorrelation, creating a wide FWHM, as shown in Figure 5.11b. If the wrinkles were randomly distributed, the autocorrelation would not start to overlap until the images were directly on top of one another, creating a narrow FWHM. Binary images enhance the sensitivity of the autocorrelation because only when fibrils overlap will there be any value added to the sum.

The continuous wavelet transform takes a small portion of a waveform and translates it across the data looking for where it matches. It does this for a range of frequencies from 0-50 cycles per μm. The results are displayed in a colour coded scalogram where a higher magnitude (more yellow) indicates a strong relation of that frequency around that area of the fibre, Figure 5.11c. Also, by looking at which frequencies are present in the signal, the approximate size of the fibrils can be determined by taking the inverse of the frequency.

The mean values of FWHM can be utilized to quantify and compare the



(a) (b) (c)

Fig. 5.11: Ideal case of perfectly aligned fibers (a) and the corresponding autocorrelation (b) and wavelet analysis (c).

structure of different fibers collected with Langmuir trough, shown in Table 5.3.

Tab. 5.3: Mean values of FWHM [μm] for the fibers collected with Langmuir trough at different waiting time and solutions. Missing values are reported with the symbol '-', if it was impossible to see any, with '+' if it was possible to see some which were too weak.

Buffer	0.1 M NaCl			0.1M NaCl+0.05M NaH_2PO_4			G0 buffer		
Time [min]	10	60	120	10	60	120	10	60	120
FWHM [μm]	6.4	-	7.7	+	+	5.2	4.6	2.2	9.4

Since for some waiting times it was not possible to collect fibers because of the the inability to produce them, and for some waiting times it was not possible to wrap samples, due to the break caused by the small diameter (e.g. $d \leq 2\mu m$), we differentiate the ability to collect small fibers from the inability to produce them. The symbol '-' is used if it was impossible to see any, '+' if it was possible to see some which was too weak to be collected. The vertical autocorrelation values corroborate the qualitative observations seen in Figure 5.8. There is a significant increase in the structural linearity and continuity of the fibers pulled after 120 minutes of waiting time as compared to both the 10- and 60-minute wait time groups. As already mentioned, our main hypothesis explains the different wrinkle structure with the fibrils islands orientation. When the film is formed and squeezed in minutes-long period of time, the island are not parallel and when fibers are pulled, the applied mechanical force pulls a straight fiber, but the film structure is not well organized and the wrinkles are non-uniform. On the other hand, in hours-long waiting time, the island interact with each other and rotate so that they are parallel. When the film is compressed and fibers are pulled structure is more organized. For this reason, we can conclude that the island

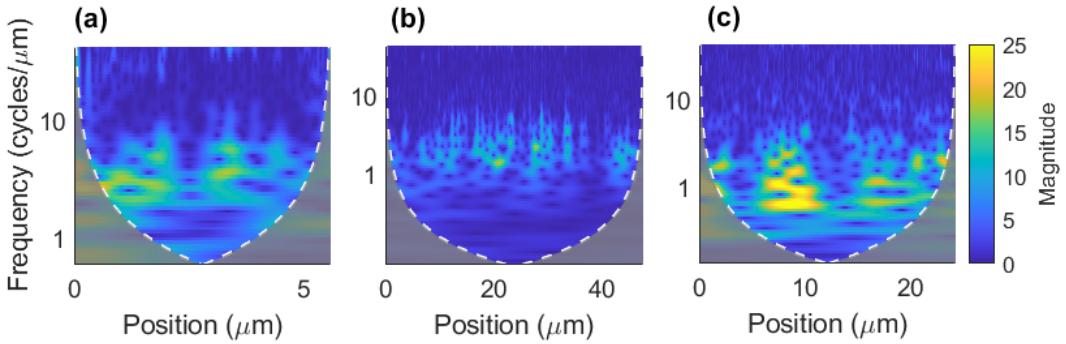


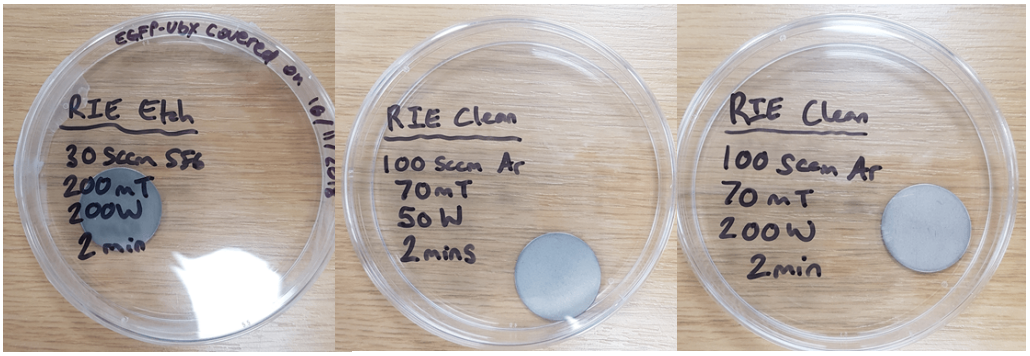
Fig. 5.12: Wavelet analysis of the fibers reveals lateral uniformity and wrinkle sizes for waiting times of 10 (a), 60 (b), and 120 minutes (c), using 2 nmol of EGFP-Ubx spread G0 buffer.

dimensions and direction determine the wrinkle uniformity.
The further step was the application of wavelet analysis in order to investigate more the wrinkle-structure uniformity. With this technique it is possible to identify characteristic frequencies within a signal and their strength, Figure 5.12. In Figure 5.12 the scalogram obtained for the fibers described in Figure 5.8 (2 nmol EGFP-Ubx spread in G0 buffer). The average pixel intensity values used was 0.1μm. The 10 minutes-waiting-time fiber presents consistently thin wrinkles ($\sim$0.25 μm in dimensions) with at least 2 areas along the fiber where they are structured, Figure 5.12(a). After a 60-minutes waiting period, there is no consistent size or structure in the image (Figure 5.12(b)), reflecting the fact that wrinkles are disordered. Finally, at 120-minutes waiting time, the structure of the wrinkles is uniform and organized and their size increases in dimension up to $\sim$1 μm.

5.1.4 Metal Disk coating

During the Langmuir experiments, we decided to study the interaction between Ultrabithorax protein-based films and titanium substrates. The goal of these observations was to understand the coating durability on titanium surface and the possibility to use Ubx-based material for cell growth applications. In fact, the titanium substrates used, have great applications in cell biology and previous works done by Bondos group, showed that Ubx monomers fused with vascular endothelial growth factor (VEGF) proteins generate fibers able to help tissue regeneration [56]. However, no studies have been done with Ubx films, so I started with EGFP-Ubx material.So I performed some preliminary test using sterile titanium discs. The first step was to clean the surface in a sonicator with some IPA for 15 minutes to re-

move any dirtiness from the surface and then do a plasma cleaning. Since the goal of this step was to sterilize the plates before the coating and not to remove TiO_2 from the surface, I opted for Sulfur hexafluoride (SF_6) and Argon (Ar) plasma cleaning instead of HF, which is more aggressive and dangerous.
As I had never done plasma cleaning and I was not sure where to start from, I have tried with different parameters. The first set of parameters used was: SF_6 for 2 minutes with a gas flow rate of 30 sccm, pressure equal to 200 mT, a power equal to 200 W. In this case, the gas is too aggressive and the cleaned surface is really dark compared to the initial one, so I was not sure if I was removing a thick film from the substrate (Figure 5.13a). To solve this problem I have tried Ar and SF_6 together for 2 minutes with 100 sccm flow rate, p = 70 mT, P = 200 W. The resulting surface is a bit dark and less uniform.



(a) **(b)** **(c)**

Fig. 5.13: Titanium discs cleaned for 15 minutes with IPA in a sonicator prior to do Plasma cleaning for 2 minutes with (a) SF_6, with 30 sscn flow rate, p = 200 mT, P = 200 W; (b) Ar, with 100 sscn flow rate, p = 70 mT, P = 50 W; (c) Ar, with 100 sscn flow rate, p = 70 mT, P = 200 W.

Then, I used 100 sccm Ar, Pressure = 70 mT, Power = 50 W and time = 2 minutes. However, the disc seemed to be as bright as it was initially (Figure 5.13b), so I tried using the same parameters used previously except the power which was increased from 50 W to 200 W. However, the disc was again as bright as it was previously (Figure 5.13c). At this point I decided to proceed with Ubx film coating with the first and the last cleaned discs, so that I could compare the resulting film.
For the metal coating I decided to use Langmuir trough in order to have controlled process. I used buffer G0 with 5% w/v Glucose and I doubled the quantity of NaH_2PO_4 so that it was possible to obtain the strongest film

possible with a 2-hours waiting time parameter instead of 16-hours one. For this purpose, I increased also the protein concentration using 10 nmol instead of 2 nmol. With the aid of a dipping arm equipped with Langmuir trough, it was possible to coat the discs uniformly in both sides of the piece even if only one surface was cleaned. Once coated, the discs were put in a plastic petri dish, sealed with parafilm, and let rest to dehydrate. After a week the discs have been re-hydrated with some PBS buffer and the film was checked with a fluorescent microscope, using a $\times 4$ objective lens and a $\lambda = 488$ nm as excitation source.

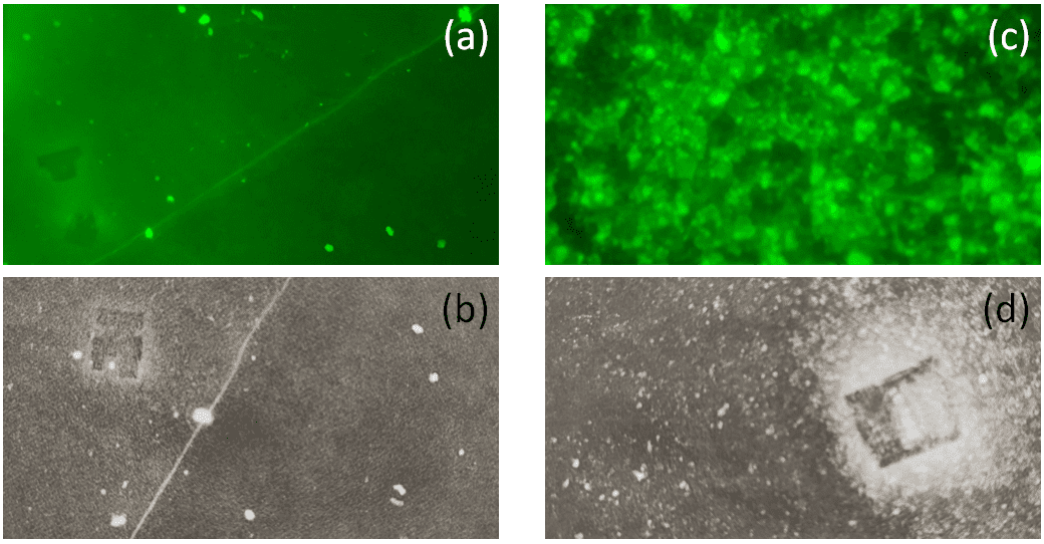


Fig. 5.14: Fluorescent pictures taken with $\times 4$ objective lens using an excitation wavelength $\lambda = 488$ nm. The disc was cleaned with SF_6 plasma, (a) and (b), and Ar plasma, (c) and (d), prior the dip coating with EGFP-Ubx.

As visible in Figure 5.14, the cleaning process with SF_6 gives a smoother and more uniform film (Figure 5.14 (a) and (b)) compared to the one cleaned with Ar plasma. This may be explained with two observations: 1- the gas flow rate used for Ar plasma (100 sccm) was so high that the ions collision on the disc created small craters on titanium surface without properly cleaning it (this could be the reason why the surface remained bright and shiny); 2- SF_6 plasma removed a thin but uniform layer of material from the surface revealing the dark part of titanium. With both explanations it is possible to conclude that the SF_6 plasma has a better cleaning effect on the titanium surface compared to Ar plasma.
Focusing on the last disc, I followed the protein-based film evolution over time. For this reason I monitored the green fluorescence at 488 nm after 1, 4, 7 and 11 weeks after film coating, see Figure 5.15. The film was re-hydrated

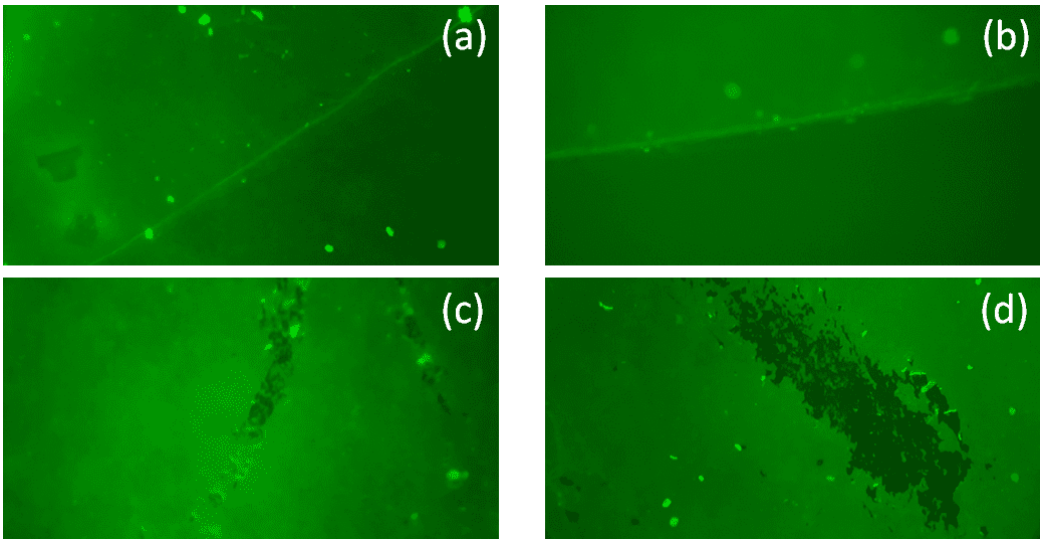


Fig. 5.15: Fluorescent pictures taken with ×4 objective lens using an excitation wavelength $\lambda = 488$ nm. The disc was cleaned with SF_6 plasma and checked (a) 1 week (b) 4 weeks (c) 7 weeks and (d) 11 weeks after the dip coating with EGFP-Ubx.

before the each observation using PBS buffer. As observable in Figure 5.15, the film remains stable for 4 weeks with almost no changes. Nevertheless, after 7 weeks a film degradation starts to appear (Figure 5.15c) becoming more extensive after 11 weeks (Figure 5.15d). For this reason, we can conclude that for future experiments it is recommended to coat the titanium disks and use them in 4 weeks.

5.1.5 Antimicrobial peptide coating

The EGFP-Ubx film coating work was particularly useful to test two new Ubx fusions and their antimicrobial action.
Our group in Swansea designed two expression constructs to create fusion proteins positioning antimicrobial peptides at the N- terminus of Ubx, using a flexible linker and SUMO tag to ensure protected expression and precise excision of the peptides. The aim was to use the Ubx film-formation ability combined with the peculiar ability to kill bacteria and fungi of specific peptides. In this way it could be possible to coat surfaces (e.g. door handle) with a thin film of these Ubx fusions and inhibit the growth of bacteria, stains, odours and product degradation. Cathelicidin LL-37 (LL37) and Bovine myeloid antimicrobial peptide-18 (BMAP) were chosen for their well-studied antimicrobial properties [31, 20]. In Figure 5.16 a model of the

BMAP-Ubx fusion protein is shown illustrating the dityrosine bond forming residues thought to be responsible for material formation. To measure

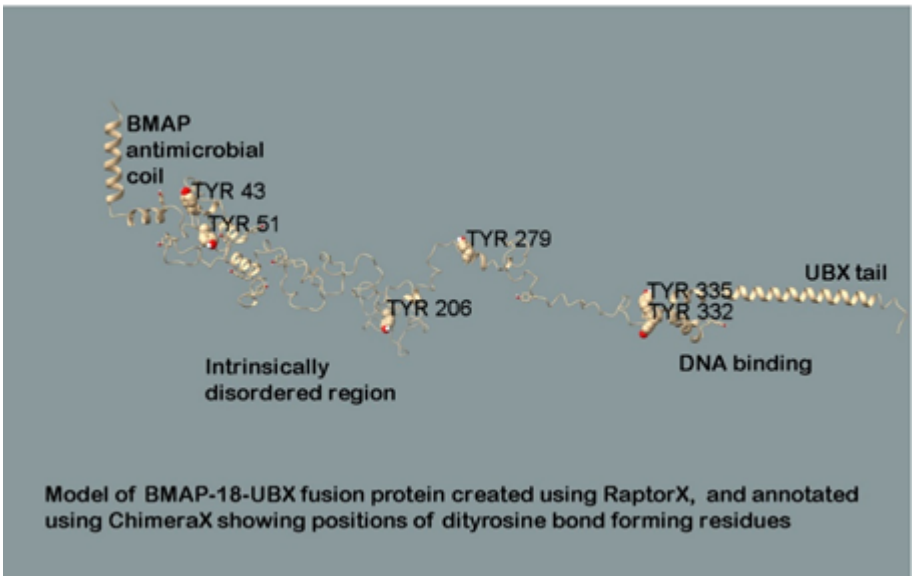


Fig. 5.16: Model of the BMAP fusion protein illustrating the dityrosine bond forming residues thought to be responsible for material formation.

the antimicrobial effect of coated surfaces with Ubx, a quantitative kinetic antimicrobial screening assay was used. This is based on ISO 22196 standards and is appropriate for use with protein covered surfaces. The assay is based on the resazurin microplate method, using kinetic analysis of growth of Bacillus subtilis, a commonly found non-pathogenic fast-growing gram-positive bacteria, as the initial test organism. Resazurin reduction, resulting from cell metabolism, produces a highly fluorescent dye which was quantified in a 96 well format using a Biotek 96 well plate reader, excitation wavelength λ_{ex} = 520 nm and emission wavelength λ_{em} = 590 nm. Three substrates were chosen for this experiment:

• titanium Grade 2 discs (diameter ∼6 mm), chosen due to its high resistance to corrosion and biocompatibility, as a stable and an inert oxide layer spontaneously forms when its surface is exposed to oxidising media;

• titanium Grade 2 discs (diameter ∼6 mm) that were previously coated with EGFP-Ubx films and etched with piranha solution (a mixture of sulfuric acid,H_2SO_4, water and hydrogen peroxide, H_2O_2) to remove Ubx film. This substrate was chosen with a view to a possible recycling of the materials;

• silicon wafer squares (side = 6 mm), chosen because it is a suitable support for growing and/or mounting cells.

The experiments were set up aseptically in a Hepa–filtered flow hood. The surfaces to be tested were initially sterilised by immersion in 60% ethanol for 10 mins followed by rinsing 4 times in sterile PBS. Then, they were coated with Ubx film, dried, placed in a sterile Petri dish and aligned on a labelled perforated sterile rack for ease of manipulation. The Petri dish was situated in a high humidity chamber. Test solutions (5μl) consisting of bacterial suspension (dilutions C and D, respectively 6×10^5 and 6×10^4 cells per mL in PBS) or PBS controls were pipetted onto surfaces which were immediately covered with autoclaved 6×8 mm pieces of thin clear polypropylene using flame sterilised forceps. The Petri dish lid was replaced and the chamber closed, then the experiment was incubated for a defined time period. When the incubation time was completed, each experimental surface and plastic film were transferred together into a labelled 2 mL microfuge tube using flame sterilised forceps. Bacteria were harvested by mixing and incubation in culture media: to each tube 500 μL LB + resazurin, made by adding 125 μL of 0.5% filter sterilized resazurin to 20 mL sterile LB, was added. Then, each tube was vortexed, and placed in a rack in heated shaker for 20 minutes, along with no-incubation controls for each bacterial suspension used. 200 μL of the samples were then transferred into a Corning™ 96-Well clear bottom black microplate for fluorescence analysis. Alongside this, a serial dilution of bacterial culture was assessed for each experiment, and two dilutions of this plated out to determine exact bacterial counts. The plate was sealed with a breathable lid and the kinetic analysis was run overnight for 16h using a predefined protocol which incubates and mixes the plate at 30 °C and takes fluorescent readings every 15 minutes to monitor growth. Data was analysed by measuring onset time at a predefined RFU of 1000 for each sample. Details of an experiment to compare titanium and silica surfaces with and without Ubx coating are shown in Figure 5.17, on the left part the scheme of experiment is depicted: silica surfaces are indicated with S, etched titanium discs are indicated with O and non-etched ones with N. Ubx-coated surfaces are indicated with 'x', while control surfaces are indicated with 'c'. On the top of samples bacteria solutions, C and D, and plain PBS (indicated as PBX) were pipetted. In Figure 5.17 on the right part, the results of overnight fluorescence are shown.

X	1	2	3	4	5	6
A	BS(B)	BS(B)	Blank	Blank		
B	BS(C)	BS(C)	Blank	Blank		
C	BS(D)	BS(D)				
D	BS(E)	BS(E)	CtC	CtC	CtC	
E	BS(F)	BS(F)	CtD	CtD	CtD	
F	SxC	NxC	OxC	ScC	NcC	OcC
G	SxD	NxD	OxD	ScD	NcD	OcD
H	SxPbx	NxPbx	OxPbx	ScPbx	NcPbx	OcPbx

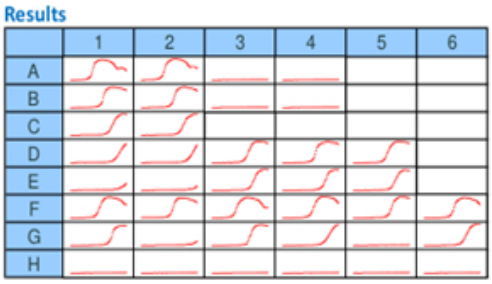


Fig. 5.17: Left part: scheme of experiment. Silica surfaces are indicated with S, etched titanium discs are indicated with O and non-etched ones with N. Ubx-coated surfaces are indicated with 'x', while control surfaces are indicated with 'c'. On the top of samples bacteria solutions, C and D, and plain PBS (indicated as PBX) were pipetted. Right part: results of overnight fluorescence.

As can be seen, Ubx coating in general do not inhibit bacterial growth, on the contrary it seems to slightly improve it. At the same time, compared to other two samples, non-etched titanium substrates are inhibiting a bit more the growth. This effect is removed when the surface it cleaned with piranha solution, probably due to chemical change at the surface. In general, bacteria grew better on etched and Ubx coated titanium surfaces.
In conclusion, it is better to use titanium grade 2 disks coated with Ubx rather

5.2 BAM results and discussion

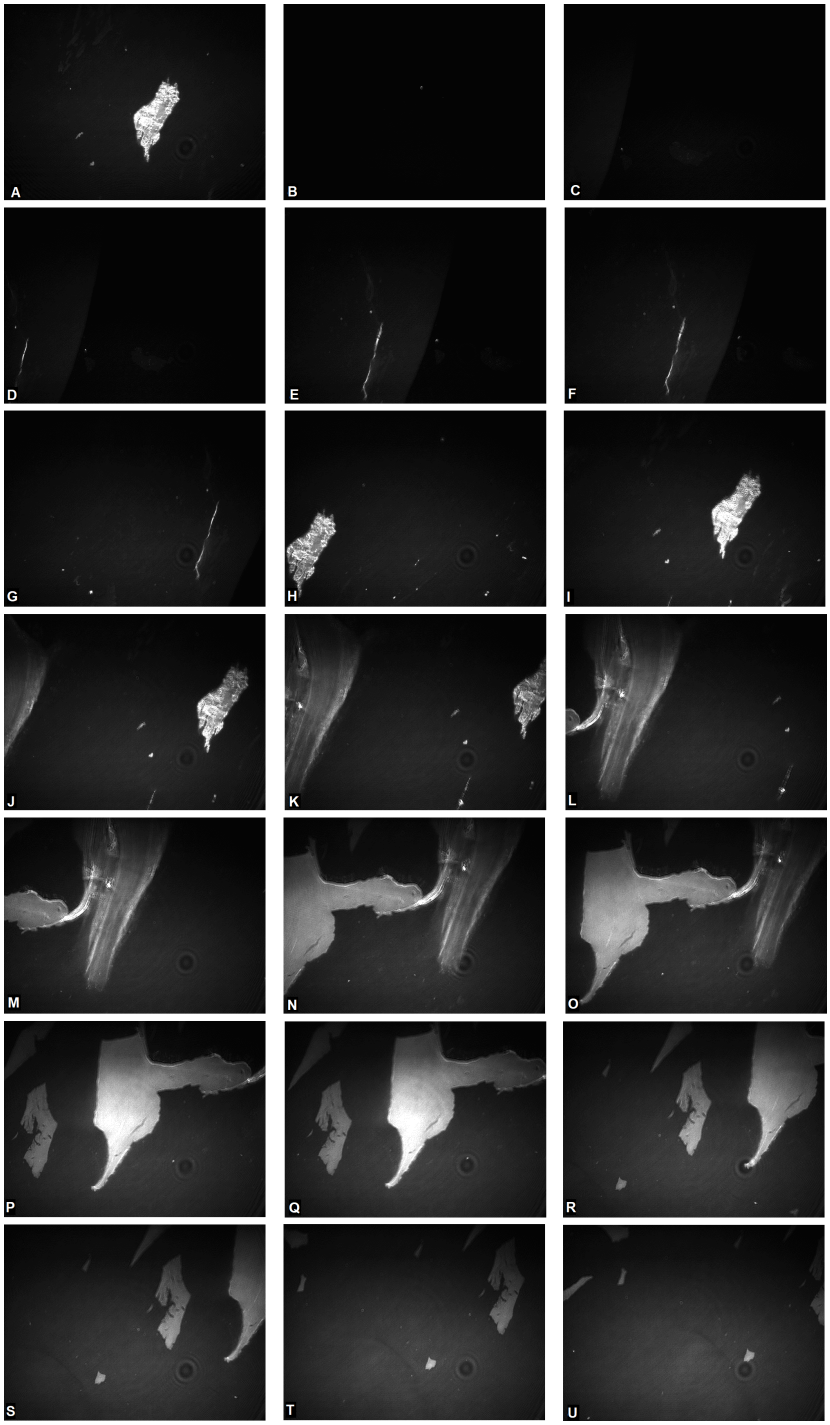


Fig. 5.18: A-U BAM time-lapse during barrier expansion. The protein film was created using 2 nmols EGFP-Ubx proteins spread in 0.1 M NaCl solution and let rest 2 hours before the barrier closure.

BAM studies started in Swansea using MicroBAM in combination with the Langmuir trough. In this way, it was possible to make observations during compression/expansion at known surface pressures and to verify the monolayer homogeneity during the film formation. MicroBAM camera has a resolution ~ 12 μm and the light source wavelength is $\lambda = 659$ nm with a power P = 50 mW. I have done the experiments using 2 nmol EGFP-Ubx proteins spread in 0.1 M NaCl solution and I let the system equilibrate for 2 hours before the barrier compression. Unfortunately, I could not image the film during barrier compression, however I was able to see what happened during the barriers expansion, which caused the film break (Figure 5.18 A-U). In the series of images, in fact, it is possible to see bright islands floating on the top of what seems to be a thin film, whose refractive index is different from the subphase one. My hypotheses is that a film starts to form as soon as proteins are spread at the air-water interface (it is not possible to switch on the BAM laser while spreading proteins because of safety reasons), and the microBAM camera is saturated at the very beginning. But, when the barriers are opened, the surface tension of the subphase is strong enough to break the film leading to portions of air-water interface where the film is not present and the camera can clearly recognise thick and thin islands floating around.

The data collected was encouraging and pushed us to investigate further once I moved to Grenoble. In this case, I was able to use the Brewster Angle Microscope Nanofilm EP3, in Partnership for Soft Condensed Matter (PSCM) laboratories. The lateral resolution was $\geq 1\mu$m, so I was hoping to be able to see film formation. In this case, I investigated air-water interface with two different sub-phases: plain water (Figure 5.19) and 0.1 M KCl solution (Figure 5.20). Since the Langmuir trough used in ILL was smaller than the one in Swansea, in order to have comparable results with the ones already commented, I scaled the quantity of EGFP-Ubx proteins proportionally to the scaled quantity of subphase used. For the two experiments I have used 1.2 nmol of proteins diluted in 350 mL of buffer, I let the proteins sit for 2 hours.

The first experiment was done with plain water, shown in Figure 5.19. As soon as I spread proteins, the ROI/sig immediately arose and the luminous intensity was linearly increasing, meaning that more and more proteins were collecting at air-water interface. However, in general during the Langmuir cycle there is not a clear structure variation even though the surface pressure indicates the presence of a film. Probably a really thin film formed at air-water interface but it was not strong enough and no fibers were visible at the end of the process.

The further experiment was done using 0.1 M KCl subphase. In this case

as soon as the proteins were spread, thick protein-based islands were visible. Moreover, during the compression-expansion cycle, it was possible to see two types of structures: a single compact film floating on non-compact material that shows a sort of lateral structure that is not complete, as can be seen in picture number 4 and 5 in Figure 5.20.

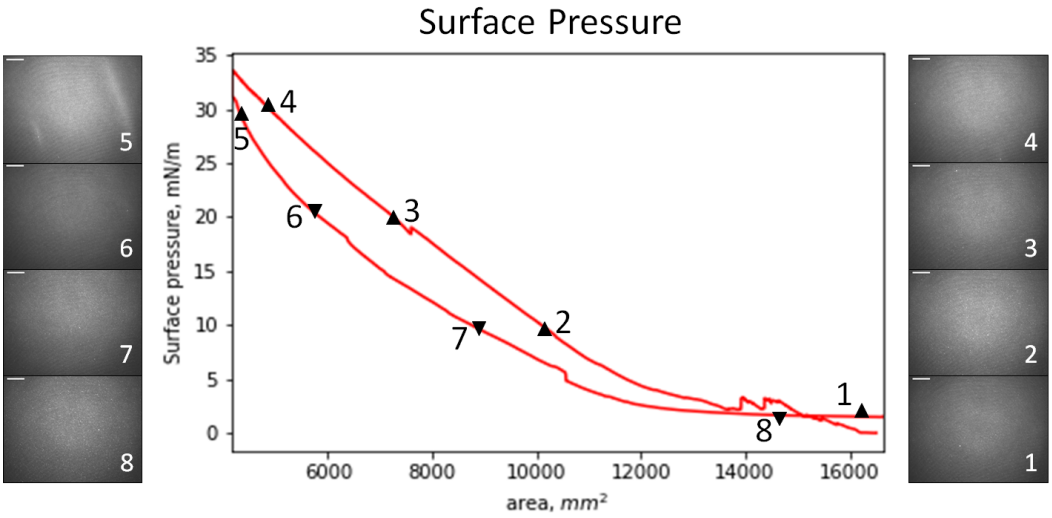


Fig. 5.19: π-A isotherm with BAM images taken during film compression-expansion in plain water.

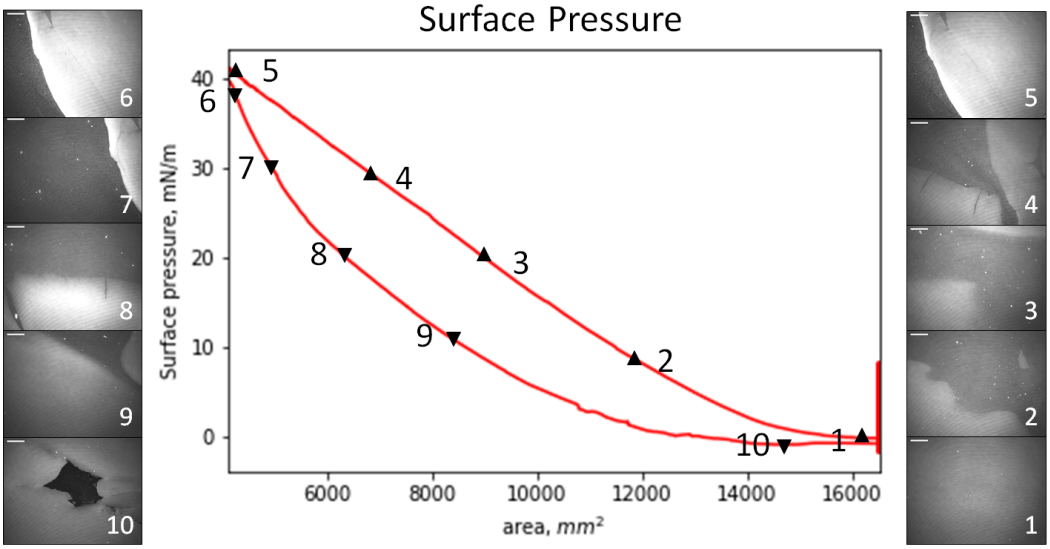


Fig. 5.20: π-A isotherm with BAM images taken during film compression-expansion in 0.1 M KCl.

During the equilibration time, it was possible to monitor film formation at air-water interface. Interestingly, lateral association between a big thick island and a small agglomerate od EGFP-Ubx fibrils, as can be seen in Figure 5.21.

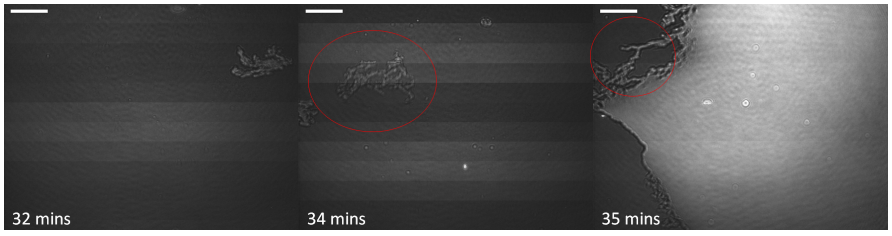


Fig. 5.21: Lateral association between big thick island and a small protein agglomerate.

6. MOLECULAR RESULTS

6.1 Ellipsometry Data

Ellipsometry results were used to help the planning and the interpretation of NR experiments, as they provide complementary informations about the surface excess during the formation of protein films.

6.1.1 Optimization of spread films

To simplify the calculation of surface excess Γ from ellipsometry data [mg/m^2], following the calculations presented in the supporting information of reference [21], it is possible to chose one model for the considered system. In this case, there are two possible models :

- The model changes in the surface coverage as changes to the layer thickness at constant density ('oil-like' behaviour). As layer density is constant, there is no solvent in the surface layer, so n_{surf} is a constant, hence the relation between the Γ and Δ_{surf} is linear [21]. In the EGFP-Ubx protein case:

$$\Gamma = 0.495\Delta_{surf} \tag{6.1}$$

- The model changes at constant thickness ('particle-like' behaviour). In this case, n_{surf} is a variable related to the density or coverage of protein in the layer, hence the relation between Γ and Δ_{surf} is quadratic. An empirical fit to simulated data gives $\Gamma(\Delta_surf)$ is quadratic [21] and for the protein studied we have:

$$\Gamma = -0.0374\Delta_{surf}^2 + 0.6126\Delta_{surf} \tag{6.2}$$

where $\Delta_{surf} = \Delta - \Delta_0$ is the change in Δ from the specimen at the interface, Δ is the measured value, and Δ_0 is the calibration value for pure subphase. For protein-based films, it is recommended to use the 'particle like' model as they tend to form monolayers with increasing area coverage rather than increasing thickness.

The ellipsometry data were recorded using a phase modulated Picometer Light ellipsometer using a HeNe laser with $\lambda = 632$ nm at $\theta = 50°$. Measurements were made using a static trough, a glass Petri dish with a diameter of 6 cm and a liquid volume of 25 cm^3. For the presented experiments we used 25 mL 0.1 M KCl solution and we added 0.591 nmol EGFP-Ubx proteins. The protein concentration chosen for the experiments, was calculated taking into account the subphase volume reduction from Langmuir trough used for the BAM (350 mL) to the Petri dish used (25 mL). The concentration was, thus, scaled from 1 to 0.6 nmol.

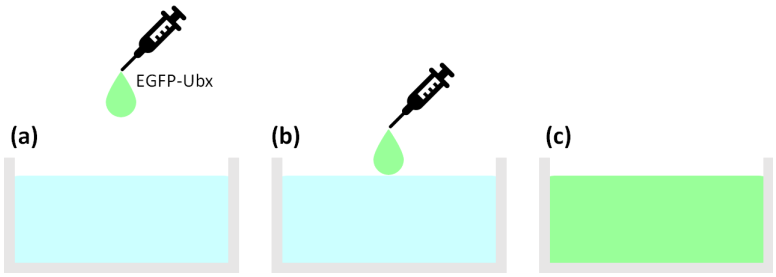


Fig. 6.1: Different loading approaching (a) one single droplet was let to fall from the top at the centre of Petri dish; (b) one single droplet was spread by touching gently the surface at the centre of Petri dish; (c) adsorbed layers from premixed solutions.

The very first study conducted with this technique, was aimed to enhance the surface excess at air-water interface when different loading approaching were used to dilute EGFP-Ubx proteins onto an aqueous subphase, see Figure 6.1: (a) one single droplet was dropped at the centre of the Petri dish from the top (height above the surface $\sim$ 1 cm), see Figure 6.1(a); (b) one single droplet was spread at the centre of the static trough by touching gently the air-water interface,see Figure 6.1(b) ; (c) subphase and proteins were premixed apart (in a volume slightly bigger than was used for the other two cases) and transferred to the trough and part of the solution was cleaned by aspiration for 3 seconds, see Figure 6.1(c). For each different loading technique, the time resolved surface excess, Γ [mg/m^2], was calculated and plotted. To investigate the consistency of the results, each experiment was repeated at least twice.
In Figure 6.2 the time-resolved surface excess is shown for the two spreading techniques (from the top and by touching the surface). Excluding an initial equilibrating period in which the surface is disturbed by the EGFP-Ubx droplet, both curves reach a steady state after several minutes. Both tend to Γ values $\sim$ 1.85 mg/m^2. So it is possible to conclude that the two techniques

are equivalent and the way to drop proteins doesn't influence the final result. Moreover, considering that the surface excess for both cases tend immediately to plateau, it is possible to state that EGFP-Ubx proteins are spread as stable films at air-water interface, despite their solubility in aqueous solutions [21]. The EGFP-Ubx protein-based film remains at the air-water interface for the duration of the measurements. In fact, protein films can be spread from concentrated droplets[21]. This can be the reason why, using BAM, it was not possible to image clearly the film formation during the waiting time or during the barrier compression process. Instead, a bright solid protein-based island was visible from the very beginning. However, remembering Langmuir trough data, it is clear that this surface-loading process depends on protein concentration. In fact, if there are not enough proteins, they dissolve into the bulk and then slowly adsorb at air-water interface again.

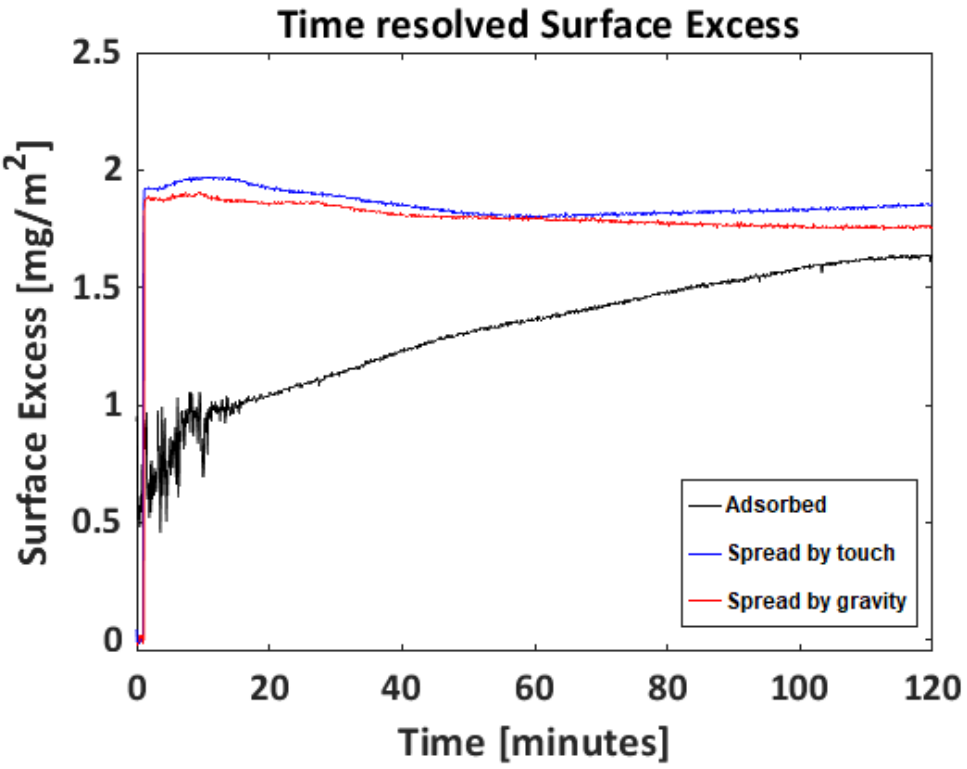


Fig. 6.2: Time-resolved surface excess of spreading technique with adsorption (black), one droplet touching gently the surface (red), one droplet falling at the centre of the static trough (blue)

The next step was the kinetic study of film formation at the air-water interface observing the time-resolved surface excess of adsorbed layers formed

when proteins and subphase were premixed, see Figure 6.1(c). In this case 30 mL of 0.1 M KCl were mixed apart with 0.71 nmol of EGFP-Ubx monomers. The solution was poured into the glass Petri dish and the surface was aspirated for 3 seconds. The surface excess was then monitored for 4 and 12 hours (Figure 6.3). From the graph it is visible that, except an initial fluc-

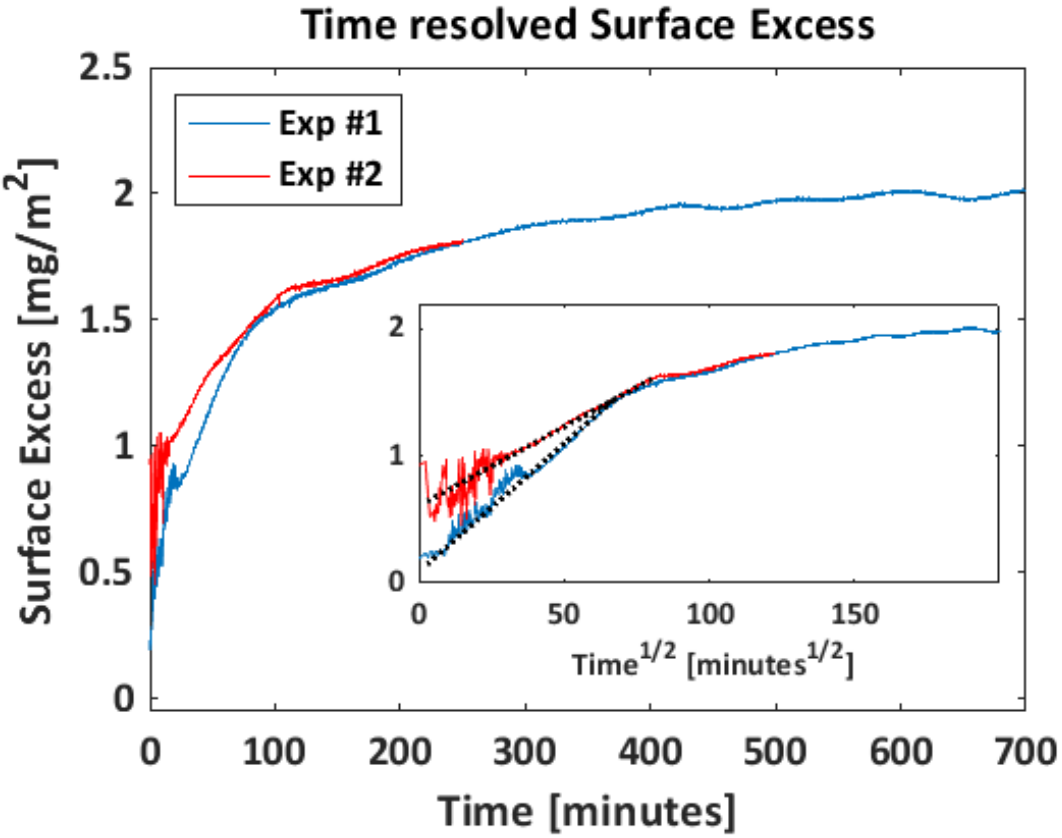


Fig. 6.3: Time-resolved surface excess of adsorption technique. In the figure frame the same data are displayed on a $t^{1/2}$ scale. The dotted black lines help to visualise the linear component.

tuation due to the surface aspiration, the two measures are consistent. This means that the kinetics of EGFP-Ubx adsorption from a premixed solution at a clean air-water interface is stable and not subjected to external parameter influence (e.g. temperature or humidity). In both cases, after an initial rapid increase, the surface excess seems to stabilize and reaches a plateau. The temporal fluctuations visible at the plateau may be related to the fusion of protein islands as they form and grow laterally at the air/water interface, passing in and out of the laser beam [21]. However, the quantity of subphase

is so small that after 3 hours it evaporates creating a laser misalignment that can contribute to signal oscillation. For this reason, for our study the experiments won't be longer than 2 hours.
In Figure 6.3 it is interesting to notice that the surface excess tends to values close to the ones seen in the spreading technique, $\Gamma\sim$ 1.79 mg/m^2. Thus, it is possible to conclude that EGFP-Ubx film formed at air-water interface after the adsorption from a premixed solution tends to be similar to the one obtained when proteins are spread at the surface. However the amount of time required to the two films to be similar is different: 2 hours against 12 hours.
Following Ramsden work[100], the surface excess was plot also on a $t^{1/2}$ scale, see box in Figure 6.3. He described the initial protein adsorption as a function of diffusion coefficient, $D^{2/3}$, bulk concentration, c, and time, $t^{1/2}$.As can be seen in the box plot, both Γ curves show a linear behaviour with positive slope, the dotted black lines help to visualise the linear tendency, before the slope decreases. The change of the slope is explained with the subsequent "jamming" of molecules after a critical surface density[100, 21].
It is possible to conclude that the film formed at air-water interface by adsorption from premixed solution resembles the film formed by spreading. However, since in the adsorption case the time-frame required is 6 times longer compared to the spreading case, in the present work only the latter case will be used for film formation.

6.1.2 NR preparation

After the forced lock down caused by Covid-19 pandemic the goal of further experiments was the characterisation of spread film at air-water interface to corroborate NR results. Moreover, a fire spread at Swansea University and destroyed the laboratories in which the EGFP-Ubx aliquots were produced and stored. For this reason, it was necessary to move the protein production in Grenoble and, more important, to test the new proteins in view of NR experiments.
The first step was to compare the surface excess of old and new proteins, see Figure 6.4. As can be seen, the new batch of proteins (Figure 6.4 black curve) shows a different behaviour compared to the old proteins (red curve in Figure 6.4). This can be explained with the fact that the new batch of proteins was completely degraded causing the absence of tyrosine sites which are responsible of di-tyrosine bonds, essential for film-formation. Since there was not enough time to start another protein production process, it was decided to test the old proteins and understand if it were possible to use them, despite the long storage at -20°C. In fact, a long storage in freezer

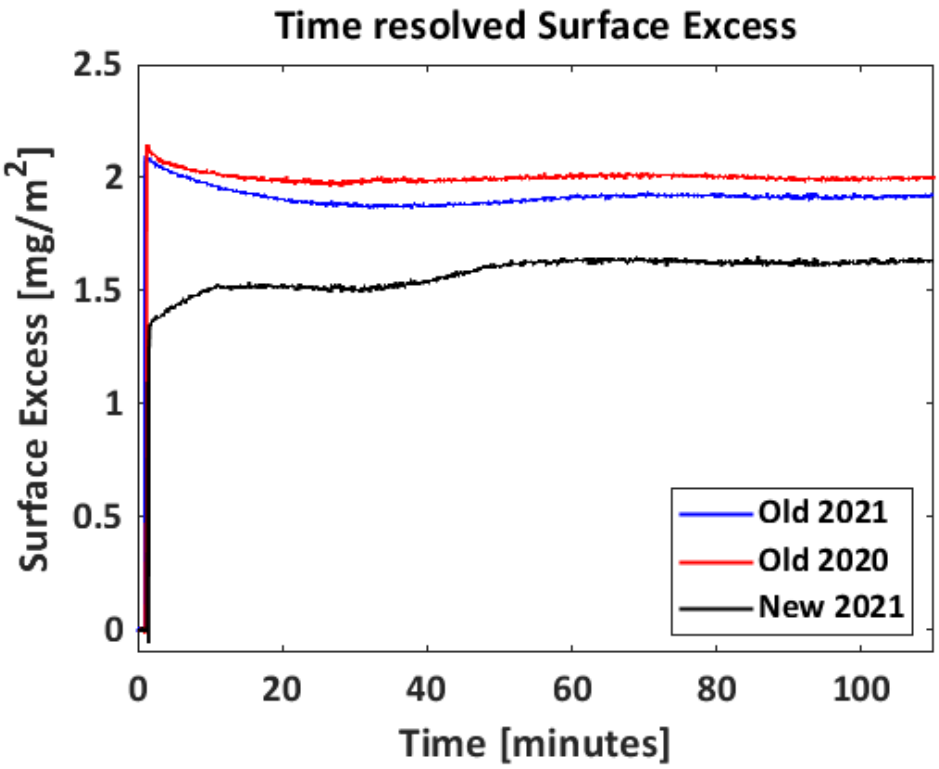


Fig. 6.4: Time-resolved surface excess of old proteins tested in 2020 (red), old proteins tested in 2021 (blue) and new proteins tested in 2021 (black).

might contribute to protein degradation. From Figure 6.4 it is possible to verify that the protein behaviour is consistent, meaning that no important degradation occurred over time and the tyrosine sites were preserved. After having verified the ability to pull fibers, it was decided to proceed with the experiments using the old proteins.

At the beginning it was important to check the consistency of old proteins spread in the subphases used for neutron reflectivity. Thus, experiments using 0.1 M NaCl solutions were repeated three times using both H_2O (Figure 6.5a) and D_2O (Figure 6.5b) as solvents.

As can be seen in the two diagrams, the time-resolved surface excess is consistent over time, in the limit of ellipsometer resolution. This means that EGFP-Ubx monomers self-assemble at air-water interface with a consistent lateral association. However, when EGFP-Ubx proteins are spread in H_2O solution, the curves differ in shape at initial times and then converge for longer time. At the contrary, when they are spread in D_2O solvent, surface excess values are quite superimposable from the very beginning. The differ-

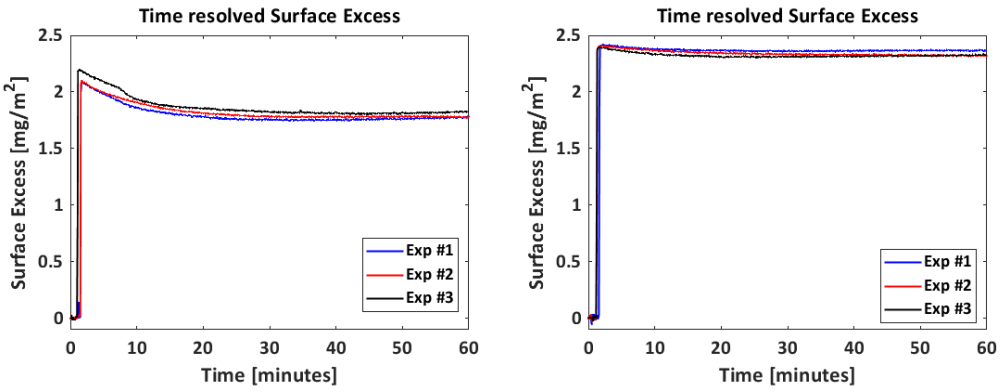


Fig. 6.5: Time-resolved surface excess of old proteins spread in 0.1 M NaCl using H_2O (a) and D_2O (b) as subphase.

ences of behaviour caused by H_2O and D_2O were studied in more depth. In Figure 6.6 a comparison between different aqueous solutions is shown. First of all, it is interesting to notice that in both cases of plain D_2O and H_2O, the addition of a kosmotrope, NaCl, helps to prevent proteins aggregation so that more EGFP-Ubx monomers are able to reach the air-water interface and participate in film formation. The observations made with Langmuir technique are thus confirmed with ellipsometry. For this reason, it was decided to use 0.1 M NaCl solutions for further experiments. Moreover, it is possible to point out that Γ values for D_2O are higher than the one for H_2O (although $\Gamma_{D_2O} \sim \Gamma_{H_2O+NaCl}$). The main cause of this behaviour can be give by isotope-specific effects from the difference in energy of H-bonding according to the isotopic composition of the subphase. Thus, it is clear that by using the same concentration of protein in the solution, it ends up with two different protein films at the air-water interface: the film in H_2O being less thick than the one in D_2O. This film difference brings to two different situations that can not be compared during neutron reflectivity data analysis. Therefore, it was decided to use a higher protein concentration in H_2O so that the resulting EGFP-Ubx film at the air/water interface has similar features (thickness, hydration) to the one forming at D_2O surface. The new protein concentration was calculated following the equation:

$$c_2 = \frac{\Gamma_{D_2O}}{\Gamma_{H_2O}} \cdot c_{Ubx} \tag{6.3}$$

where Γ_{D_2O} and Γ_{H_2O} are the corresponding surface excess values when the film is spread in D_2O and H_2O, respectively; and $c_1 = c_{Ubx}$ is the lower protein concentration, used until now. The results are shown in Figure 6.7

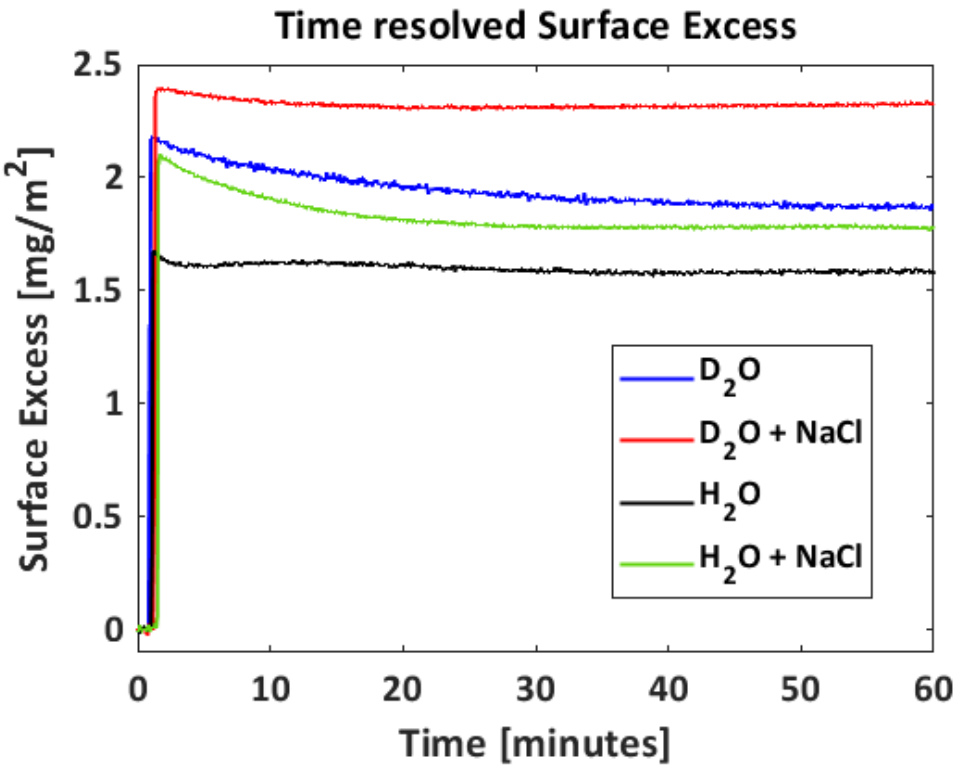


Fig. 6.6: Time-resolved surface excess of EGFP-Ubx proteins spread in different subphases: H_2O, 0.1 M NaCl in H_2O, D_2O and 0.1 M NaCl in D_2O.

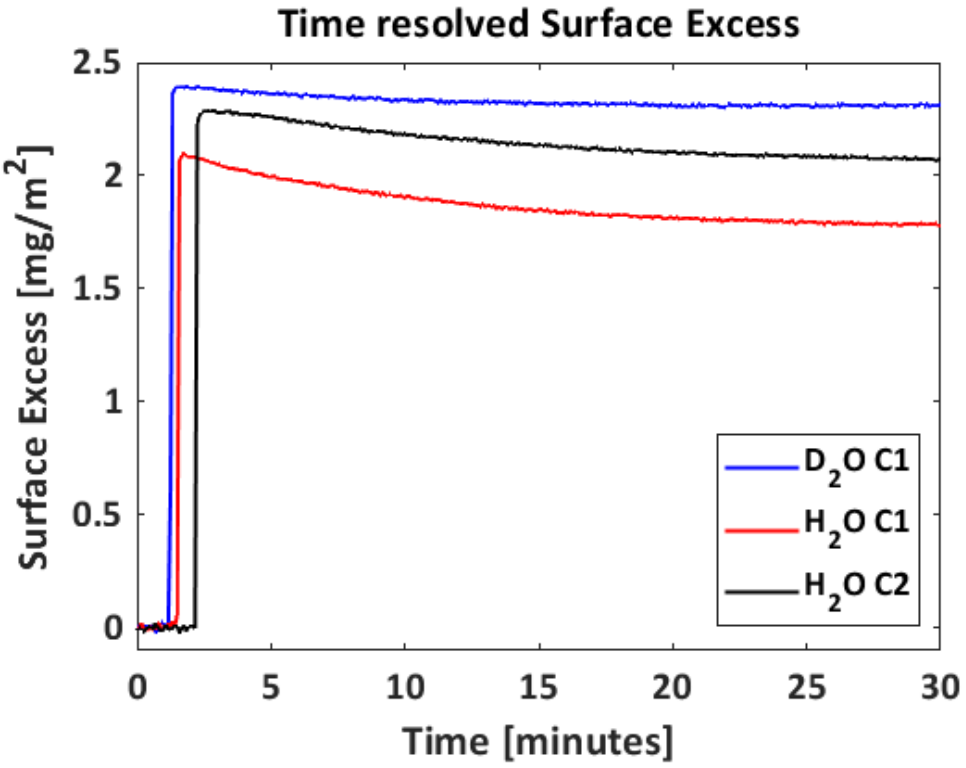


Fig. 6.7: Time-resolved surface excess of EGFP-Ubx proteins spread in D_2O and H_2O with two concentrations c_1 and c_2.

As shown in the diagram, increasing the protein concentration the Γ values for H_2O are closer to the values for D_2O using c_1. In case of c_2 there is also a change in curve shape which resembles the one for D_2O and c_2. With all these results it was possible to proceed with neutron reflectivity experiments.

6.2 Neutron reflectometry data

As already mentioned in the chapter 3.5, during NR experiment a neutron beam was directed toward the air-water interface where the proteins were spread and the intensity of reflected beam was recorded in specular set up. The reflectivity profile, $R(Q_z)$, was then calculated as ratio between the reflected neutron beam and the incident one as a function of momentum transfer perpendicular to the surface $Q_z = (4\pi \ / \ \lambda) \cdot \sin(\theta)$, where λ is the neutron beam wavelength and θ is the incidence angle. The data were reduced using COMSOS [89]. $R(Q_z)$ is a function depending on the profile along the z direction of the scattering length density (SLD),ρ. The data were analysed with the software Motofit [89]. The system was modelled as a single layer of proteins, characterized by and average ρ, a thickness d and a roughness σ, floating on a semi infinite liquid. The parameters found with this model, were then used to calculate a reflectivity profile that could be compared to the measured one. The fitting process was done using a 5% $\Delta\chi^2/\chi^2$. A common problem in neutron reflectometry is the absence of a biunivocal relationship between the reflectivity profiles and the length density profiles. More than one scattering length density profiles can be related to a reflectivity profile. To solve this issue more experiment are performed on the same system having a different isotopical composition. This method, referred as *contrast variation*, is based on the hypothesis that upon isotopic substitution of one or more components of the experimental system, the main nanostructural features does not change significantly. As we will see in the following sections, this hypotesis is not strictly fulfilled in the systems that we are studying here. Hence, in the case of Ubx films different technical expedients are required to use. The mixtures used in the following experiments are D_2O and air contrast matched water (ACMW). The latter solvent is a mixture of 8.13% v/v D_2O + 91.87% v/v H_2O, and its SLD resulted equal to zero being therefore equivalent to air. A bare ACMW-air interface does not reflects neutrons,ensuring that the only scattering comes from the EGFP-Ubx. Hence, this contrast is particularly sensitive to the properties of the protein layer. To both ACMW and D_2O different salts were added in order to obtain comparable results with what was previously done.

6.2.1 Optimisation of the samples for NR experiments

In January 2020, the first NR experiment was performed with EGFP-Ubx. Experiments were performed at FIGARO reflectometer (Institut Laue-Langevin, Grenoble) equipped with the Langmuir trough. The instrument was used in kinetic mode. Since no one has ever tried NR with the protein in question, the goal of this experiment was to explore and optimise the parameters to obtain good reflectivity signal. Even if one day (out of two) was lost due to technical problems of the instrument, it was possible to investigate an initial structural characterisation of the protein adsorption at the air-water interface at before beginning the compression (trough area = 300 cm^2) and full compression (area 110 cm^2) using two contrasts D_2O and ACMW with 0.1 M KCl. For the experiments 140 mL of subphase were used and the amount of EGFP-Ubx proteins used, 50 μl of 10μM, was calculated to match ellipsometry results shown previously. The NR curves in Figure 6.8 represented two different stages of film formation and the fits corresponded to hydrated protein layers.

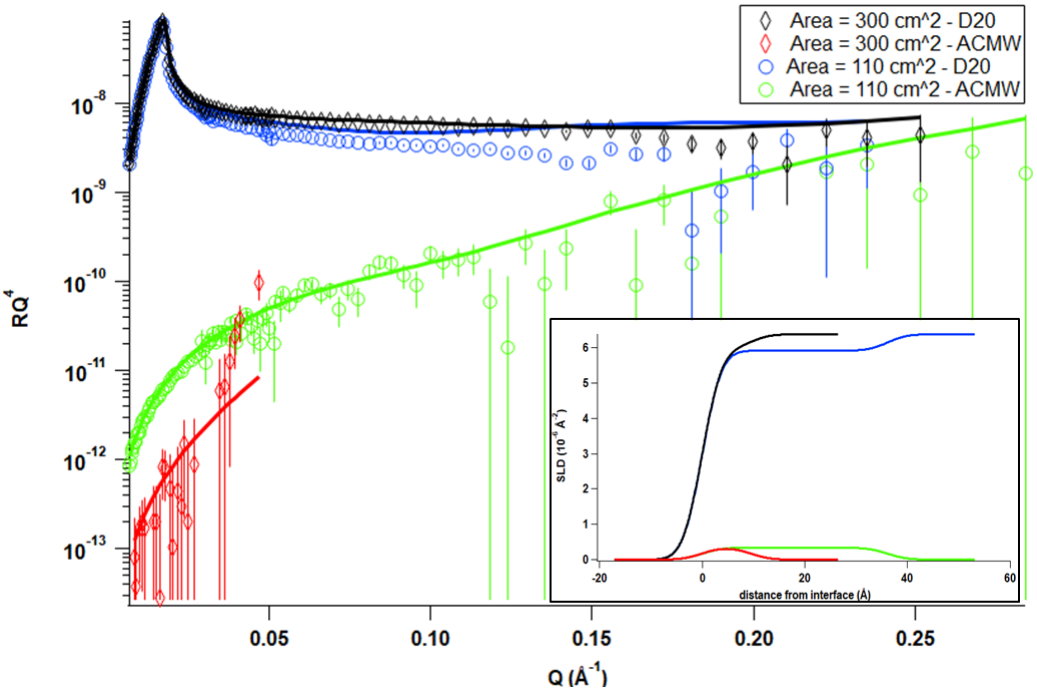


Fig. 6.8: Measured reflectivity profiles at two contrasts obtained for two compression areas (300 and 110 cm^2). In the inset Corresponding SLD profiles from the NR curves.

The first visible result was the difference between the reflectivity for ACMW at $A_1 = 300$ cm^2 and $A_2 = 110$ cm^2. When the barriers were fully open, there was a noisy signal coming from neutrons at low Q, that would become even noisier for higher values of Q (for this reason we did not measure it). On the other side, after the film compression the resulting signal was less noisy and stronger. This result was coherent with surface-pressure (π-A) values: at the beginning, the Ubx-monomers interactions were minimal, and the monomers started to aggregate forming small islands. As the barriers moved closer, the interactions became stronger and stronger because of the decrease of area per molecule. This resulted in an increase of π-A values and consequently in a more compact film structure. So at this protein concentration a film was able to form at the surface.
Differences can be seen also when the proteins were spread in D_2O. In this case, the signal is higher mainly due to the fact that the contrast between D_2O-air is higher (ACMW-air contrast is zero), so the intensity of the reflection is high. Moreover, we know that with D_2O the films are thicker, because this subphase promotes the monomers interaction, hence a film formation could be seen from the very beginning. The compression enhanced

the interactions and we could observe a variation in reflectivity which meant a thickness variation maintaining the other parameters constant.
In conclusion, we were able to detect the presence of EGFP-Ubx film at air-liquid interface. A film was measured even without compression, which perfectly agreed with the observations made with BAM. Moreover, we were able to set the parameters that let us measure neutron reflectivity.

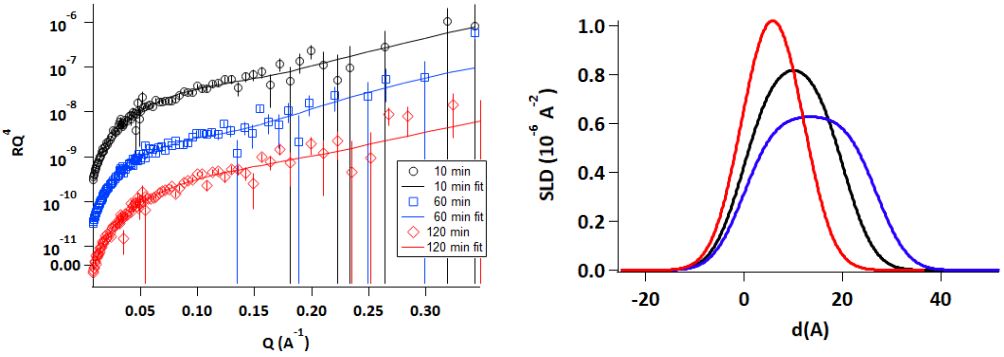
6.2.2 Experiments

After that optimum conditions for sample preparations were achieved, it was possible to to better plan a neutron reflectometry experiment to investigate the film formation. For this purpose, a new beam time was obtained in February 2021. From preliminary investigations using laboratory-based techniques, it seemed that the key differences with waiting-time parameter corresponded to either sub-micron lateral structural developments, or structural changes normal to the interface. Thus, the goal of this new set of experiment was the probe of these changes.
NR experiments were performed at FIGARO reflectometer equipped with the same Langmuir trough used previously (trough area at barriers fully open equal to 256 cm^2 and area at full compression equal to 127 cm^2). The characterisation of the protein adsorption at the air-water interface was done using two subphases, D_2O and ACMW (140 mL), to maximise the contrast between the liquid and the protein. D_2O is a contrast that generally provides a good reflectivity signal, while ACMW is a contrast that maximise the sensitivity to the structural aspects of the protein films, although it has generally a lower reflectivity signal. In order to match the Langmuir experiments, previously described, we decided to investigate the three waiting-times responsible for the non monotonic effect (10, 60 and 120 minutes) of π-A values, previously described using subphases with 0.1 M NaCl. For simplicity, in this work we won't specify the salt concentration as it will remain the same for all the results shown. After the preliminary experiments with ellipsometer, we noticed that EGFP-Ubx protein film had two different behaviours in ACMW and D_2O. That is, by using the same concentration of protein in the solution, two different films were forming at the air-water interface (in ACMW was less thick than the one in D_2O). Therefore, we chose a higher concentration in ACMW that would result in a protein film at the air/water interface that had similar features (thickness, hydration), at least as a result of the ellipsometry experiments. In order to be able to compare NR data and Langmuir results, we decided to scale the EGFP-Ubx concentration based on surface-Area (at barriers fully open) reduction. So, since the trough area scales from 800 cm^2 to 256 cm^2, the amount of proteins was scaled from 2 nmol to 0.6 nmol.

For this reason, we chose to use c_1 = 33.5 μL of 17.76 μM stoking solution of EGFP-Ubx. Moreover, a second protein concentration, and c_2 = 50 μL of 17.76 μM, was used for ACMW subphase, to match what was observed with ellipsometry. Proteins were spread at the surface by gently touching the interface and the Langmuir trough barriers were closed after the selected equilibration time. Moreover, we hypothesized that there were no structural differences between the film obtained in ACMW and in D_2O subphases due to the fact that dytirosine bonds, involved in film film-formation, are very specific. This was important in order to perform NR co-refinement at two contrasts with two films that in first approximation were similar. Neutron reflectivity was measured as a function of time during both the equilibration and the compression period. In this work, measures at full compression are shown.

The first set of experiments done was with EGFP-Ubx concentration equal to c_1 in ACMW. In Figure 6.9 on the top, the reflectivity and SLD are shown, while the bottom part shows the model parameters used to fit the data. As can be seen, the three reflectivity curves, which were shifted vertically for clarity, present similar shapes, but the differences between them are sufficiently large to obtain significant difference in results of the data fitting, as it can be seen from the SLD graph reported in Figure 6.9 right. The black line represents the SLD values associated with 10-min waiting time, the blue one is associated to 60-min and the red to 120-min. As can be seen, the high-peak SLD values seem to follow a sort of non-monotonic effect seen previously for Langmuir curves: from SLD = $0.8{\cdot}10^{-6}$ Å^{-2} at 10 minutes, the value decreases to $0.6{\cdot}10^{-6}$ Å^{-2} for 60 min and the reaches $1{\cdot}10^{-6}$ Å^{-2} for 120 minutes waiting-time. The same trend can be seen for film thickness, visible in the bottom part of Figure 6.9. The thickness of EGFP-Ubx film formed after 10 minutes was 17.1±2.8Å with a hydration of 53.0±0.2 %. When the monomers equilibrated for 60 minutes before the barrier compression, the surface film was 26.9±1.6 Å thick, with a solvent component equal to 70.3±1.8 %. Finally, for longer waiting time periods, 120 minutes, the thickness of the resulting film was 11.6±1.0 Å with a hydration of 36.5±0.4%. This behaviour can be explained with film organization structure. When the fibrils composing the protein-based material are well aligned to each other, the film resulted to be thin and the fibril structure avoid to incorporate solvent. On the other hand, a disordered material present a thick film in which there is a higher quantity of solvent. As discussed in the previous chapter, at 120 minutes the film presented a highly ordered pattern resulting in a thin thickness, while, on the other hand, after 60 minutes of equilibration, the material structure presented disordered pattern where fibrils are non-continuous and winding and the thickness was higher compared to the other cases. In

the middle of these two extremes, at 10 minutes waiting time, film present is not perfectly aligned and its thickness is more similar to the 120-minutes waiting-time than to the 60-minutes one.



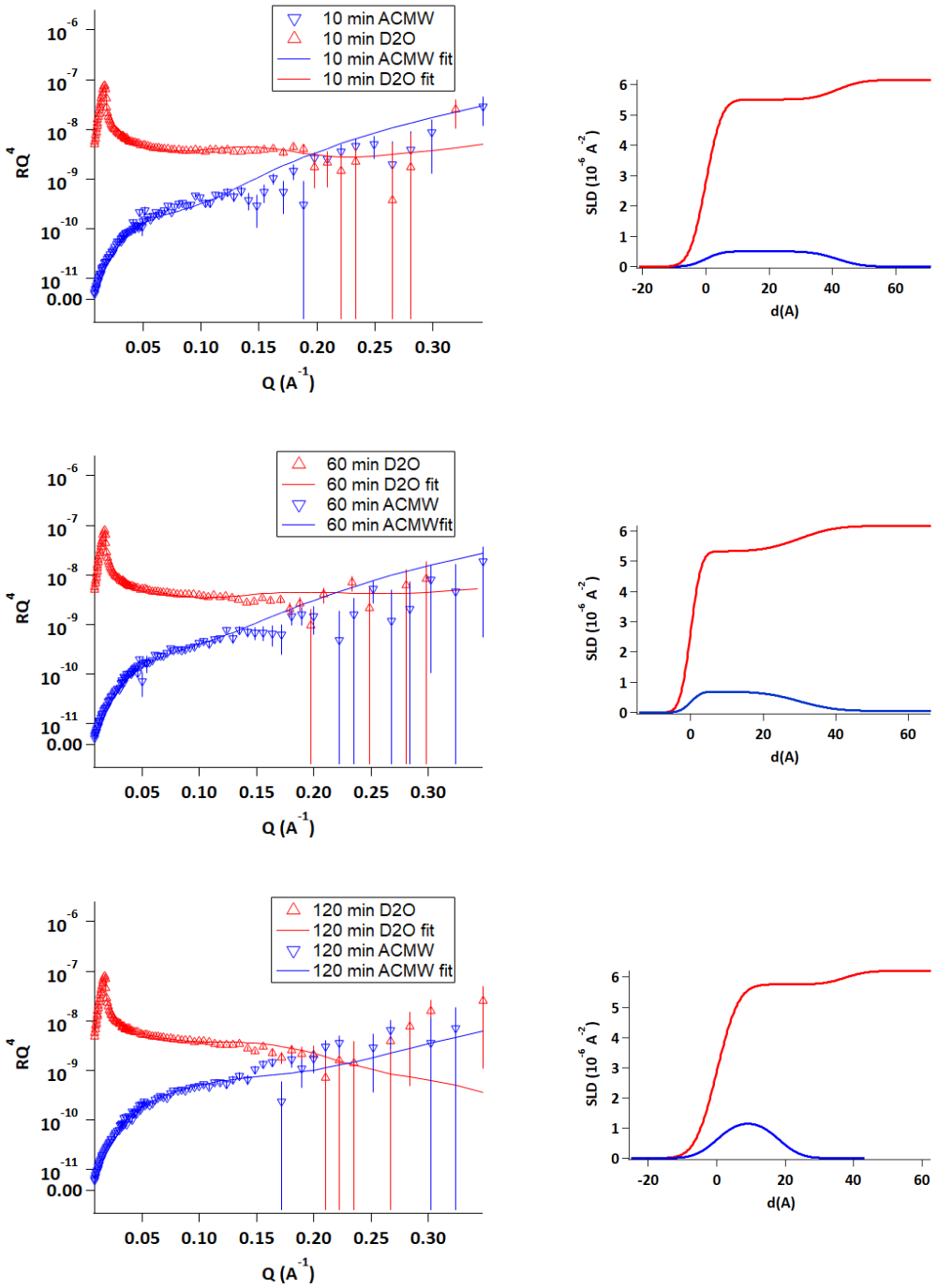
	d [Å]	Solvent [%]
10 min	17.1±2.8	53.0±0.2
60 min	26.9±1.6	70.3±1.8
120 min	11.6±1.0	36.5±0.4

Fig. 6.9: Top left: Reflectivity profile the protein concentration c_1 spread in ACMW with barriers fully compressed. The curves are shifted vertically for clarity. Top right: Corresponding SLD profiles from the NR curve for the protein concentration c_1 spread in ACMW with barriers fully compressed. Bottom: Structural parameters obtained from the NR curves: d (thickness [Å]), and solvent content[%] for the protein concentration c_1 spread in ACMW with barriers fully compressed.

It is possible to conclude that with this experiment we were able to corroborate the observations done with experiments made starting from the same protein concentrations in the liquid phase. We could ask if the lateral morphologies determined using BAM and described in the previous chapter 5.2 might affect the interpretation of the NR data. It is known that the presence of lateral morphologies on a surface affect the neutron reflection in two different ways wheatear these morphologies extend on a length scale larger or smaller of the coherent length of the neutrons. In the case where two different morphologies are present on the surface of water whose lateral extension is larger than the coherent length of the neutrons (e.g. macroscopical regions with and without protein films), the resulting reflectivity profile would be given by a linear combination of the reflectivities of the two regions weighted by the relative fraction of the areas occupied by the two morphologies. However, that was not the case in our NR experiments as the reflectivity profiles could be well fitted with one single reflectivity profiles. We can then conclude

that our protein films at the stage in which the measurement were conducted, presented a homogeneous composition at least on the scale of the coherent length of the neutrons.

The second set of experiment was done spreading a protein concentration c_2 in ACMW and c_1 in D_2O and a co-refinement of the two measures have been done. The use of two different concentrations of EGFP-Ubx at the two contrasts was dictated by the observation that the thickness of the protein film changes significantly if we use D_2O or ACMW. We then adapted the concentration of the protein in ACMW to match the elliposmetric thickness of the protein film in D2O. In Figure 6.10 reflectivity and SLD data of EGFP-Ubx monolayers spread in ACMW and D_2O for different waiting times are shown. In the table, the structural parameters obtained from the co-refinement of NR curves are listed. In this case, it is possible to see that the reflectivity curves present really small changes. Moreover, the non monotonic effect seen previously disappears in this case, for both ACMW and D_2O subphase, returning in the regime studied by Bondos' group. The film thickness, calculated with co-refinement, decreases while the waiting time increases too, going from 41.4±1.0Å to 30.0±0.8Å and finally 22.3±1.4ÅThe same trend is followed by SLD values and roughness. The non-monotonic effect disappears also for the two experiments in ACMW and D_2O separately, demonstrating that this phenomenon is not caused by co-refinement. This clearly indicates that as time passed, the monomers first and the fibrils then, had enough time to organise and form a smooth structured material, with a decreasing content solvent. This result shows that the concentration of the EGFP-Ubx, used to deposit the film, has a key role in the final properties of the film.
The explanation of this phenomenon can be attributed to different causes to the two subphases used, ACMW and D_2O. In the first case, in fact, probably the protein concentration used (c_2) was too high for the amount of liquid (140 mL) in which it was spread. The non-monotonic behaviour was not fully explored at different protein concentrations. In the case of D_2O subphase, instead, the loss of non-monotonic behaviour can be explained with the stronger hydrophobic bonds compared to the one present in H_2O. Since Ubx proteins present hydrophobic aminoacids at the outside, their structure is affected by the change of subphase hydrophobic bonds and the stability of the monomers increased. Moreover, D_2O has a higher density compared to H_2O which is a key parameter in Ubx film-formation. In fact, G0 buffer used for film formation contains 5% w/v D-Glucose that increases the density of the liquid and helps to stabilise the protein structure. So, in this experiment, D_2O effects are similar to D-Glucose ones.



	d [Å]	Solvent [%]	σ [Å]
10 min	41.4±1.0	76.5±1.6	4.0±0.5
60 min	30.0±0.8	69.3±0.9	2.3±0.3
120 min	22.3±1.4	54.1±3.1	1.3±1.2

Fig. 6.10: Top: Reflectivity and corresponding SLD profiles at three waiting time (10, 60 and 120 minutes) spreading c_1 of EGFP in ACMW and c_2 of EGFP in D_2O. Bottom: Structural parameters obtained from the NR curves: d (thickness [Å]), solvent content[%] and roughness (σ) [Å].

In conclusion, this second set of experiments shows a key information concerning the use of the contrast variation on EGFP-Ubx films at the air water interface that might be relevant for other protein systems at the air water interface. Since the properties of the resulting film strongly depend on the isotopic composition of the liquid subphase, care must be taken in the choice of the the reflectivity curves that are used to perform co-refinement. Further studies are needed on this topic. For example, it would be useful to repeat the reflectometry experiments at lower concentration in D_2O to check if similar trends with waiting time can be obtained as in the concentration c_1.

7. CONCLUSIONS & FUTURE WORKS

Ultrabithorax (Ubx) protein is a Hox transcription factor that regulates the transcription rate of genetic information inside fruit fly cells. Interestingly, it was found that this protein presents a unique property in addition to its in vivo function as a transcription factor: Ubx self-assembles into biomaterials through intermolecular dityrosine bonds at air-water interfaces [57]. Ubx protein-based materials are non-immunogenic, biodegradable, biocompatible, and have stabilizing effects on incorporated proteins [95, 96, 58, 61]. These unique properties and the ability to incorporate functional proteins make the materials ideal for use in a variety of fields. In this dissertation, I explored EGFP-Ubx protein-based material in order to understand which parameters influence the film formation and, consequently, the resulting film structure. These results have a great impact in film customization and optimization for different applications.

The first and principal result was the ability to produce EGFP-Ubx proteins in Swansea laboratories with high purity, even though in small batches. Without this outcome it could not be possible to achieve the results described.

Once the protein production was established, it was possible to proceed with the proposed analysis. The first experiments involved the monitoring of film formation in a LB trough where three critical parameters have been studied: the concentration of monomers used, the subphase in which the protein is spread and the waiting time needed for monomers to equilibrate before the area reduction through the barrier compression. At the end of these studies we discovered a new material formation regime depending on waiting time parameter. Moreover, a new fiber-formation protocol was developed using 5-times lower protein concentration in 100-times less time compared to all the works published until now. The ability to pull fibers after just 10 minutes using such a low protein concentration was surprising and totally unexpected. This result had not been achieved by Bondos group identifying a new phase of material formation that has never been studied.

We developed an in-house Matlab tool, based on autocorrelation and wavelet

analysis, to analyse the structure of pulled fibers at different conditions. With this code it was possible to make hypotheses on film structure at air-water interface by quantifying the order of wrinkles present in the fibers. This tool can be adapted to any digital image.
With EGFP-Ubx I made some tests on surface coating analysing the degradation time frame, discovering that the film remains stable for 4 weeks with almost no changes. The coating film started to degrade after 7 weeks becoming more and more extensive after 11 weeks. With plain Ubx film coating, instead, we tested the antimicrobial property of coated surfaces finding out that the best substrate to use is titanium grade 2 and cannot be cleaned to a future reuse with piranha solution because the etching would change chemically the surface causing an environment favourable to bacterial growth.
The following step was to analyse film formation more in deep. After some parameter optimization done with ellipsometry, I moved to neutron reflectometry. I was able to estimate the film thickness and to probe the time effect of EGFP-Ubx at two contrasts, corroborating the observations made with low protein concentration.

At the end of the present work it is important to point some possible future works. First of all, it would be important to study deeper the isotopic effect on film formation and repeat the reflectometry experiments at lower concentration in D_2O as mentioned in the previous chapter 6.2.2. Nonetheless, the shown results raised several questions concerning the specific influences on the structure and properties of the materials assembled with the new protocol. A possible study in collaboration with Bondos group could be useful to determine whether similar interactions form in fibers created with the two different methods.

BIBLIOGRAPHY

[1] E. Abbe. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung: I. Die Construction von Mikroskopen auf Grund der Theorie. *Archiv für mikroskopische Anatomie*, 9(1):413–418, dec 1873.

[2] M. Andersson, G. Chen, M. Otikovs, M. Landreh, K. Nordling, N. Kronqvist, P. Westermark, H. Jörnvall, S. Knight, Y. Ridderstråle, L. Holm, Q. Meng, K. Jaudzems, M. Chesler, J. Johansson, and A. Rising. Carbonic Anhydrase Generates CO2 and H+ That Drive Spider Silk Formation Via Opposite Effects on the Terminal Domains. *PLoS Biology*, 12(8):e1001921, aug 2014.

[3] M. Andersson, J. Johansson, and A. Rising. Silk spinning in silkworms and spiders, aug 2016.

[4] C. Åstrand, V. Chotteau, A. Falk, and M. Hedhammar. Assembly of FN-silk with laminin-521 to integrate hPSCs into a three-dimensional culture for neural differentiation. *Biomaterials Science*, 8(9):2514–2525, 2020.

[5] M. Avadhanulu and P. Kshirsagar. *A textbook of engineering physics*. 1992.

[6] F. J. AVILES, G. E. CHAPMAN, G. G. KNEALE, C. CRANE-ROBINSON, and E. M. BRADBURY. The Conformation of Histone H5: Isolation and Characterisation of the Globular Segment. *European Journal of Biochemistry*, 88(2):363–371, aug 1978.

[7] N. A. Ayoub, J. E. Garb, R. M. Tinghitella, M. A. Collin, and C. Y. Hayashi. Blueprint for a High-Performance Biomaterial: Full-Length Spider Dragline Silk Genes. *PLoS ONE*, 2007.

[8] Z. Bazrafshan and G. K. Stylios. Spinnability of collagen as a biomimetic material: A review. *International Journal of Biological Macromolecules*, 129:693–705, 2019.

[9] V. Beachley and X. Wen. Fabrication of nanofiber reinforced protein structures for tissue engineering. *Materials Science and Engineering C*, 29(8):2448–2453, oct 2009.

[10] S. Berg. Marangoni-driven spreading along liquid-liquid interfaces. *Physics of Fluids*, 21(3), 2009.

[11] M. Bienz. Homeotic genes and positional signalling in the Drosophila viscera, jan 1994.

[12] K. B. Blodgett. Films Built by Depositing Successive Monomolecular Layers on a Solid Surface. *Journal of the American Chemical Society*, 1935.

[13] K. B. Blodgett. Properties of built-up films of barium stearate. *Journal of Physical Chemistry*, 41(7):975–984, 1937.

[14] A. Bogner, P. H. Jouneau, G. Thollet, D. Basset, and C. Gauthier. A history of scanning electron microscopy developments: Towards "wet-STEM" imaging. *Micron*, 38(4):390–401, jun 2007.

[15] S. E. Bondos and A. Bicknell. Detection and prevention of protein aggregation before, during, and after purification. *Analytical Biochemistry*, 2003.

[16] S. E. Bondos, G. Geraldo Mendes, and A. Jons. Context-dependent HOX transcription factor function in health and disease. In *Progress in Molecular Biology and Translational Science*, volume 174, pages 225–262. Academic Press, jan 2020.

[17] S. E. Bondos, L. Swint-Kruse, and K. S. Matthews. Flexibility and disorder in gene regulation: LacI/GalR and Hox proteins. *Journal of Biological Chemistry*, 290(41):24669–24677, 2015.

[18] M. Born, E. Wolf, A. B. Bhatia, P. C. Clemmow, D. Gabor, A. R. Stokes, A. M. Taylor, P. A. Wayman, and W. L. Wilcock. *Principles of Optics*. Cambridge University Press, oct 1999.

[19] C. H. Bowen, B. Dai, C. J. Sargent, W. Bai, P. Ladiwala, H. Feng, W. Huang, D. L. Kaplan, J. M. Galazka, and F. Zhang. Recombinant Spidroins Fully Replicate Primary Mechanical Properties of Natural Spider Silk. *Biomacromolecules*, 2018.

[20] R. Bucki, K. Leszczyńska, A. Namiot, and W. Sokołowski. Cathelicidin LL-37: A multitask antimicrobial peptide. *Archivum Immunologiae et Therapiae Experimentalis*, 58(1):15–25, 2010.

[21] R. A. Campbell, J. C. Ang, F. Sebastiani, A. Tummino, and J. W. White. Spread Films of Human Serum Albumin at the Air-Water Interface: Optimization, Morphology, and Durability. *Langmuir*, 31(50):13535–13542, 2015.

[22] R. A. Campbell, A. Tummino, B. A. Noskov, and I. Varga. Polyelectrolyte/surfactant films spread from neutral aggregates. *Soft Matter*, 12(24):5304–5312, 2016.

[23] J. Chadwick. Possible existence of a neutron, 1932.

[24] R. Chebbi. Dynamics of unidirectional gravity-inertial spreading of oil on water. *Chemical Engineering Science*, 60(23):6806–6813, dec 2005.

[25] L. A. Clifton, R. A. Campbell, F. Sebastiani, J. Campos-Terán, J. F. Gonzalez-Martinez, S. Björklund, J. Sotres, and M. Cárdenas. Design and use of model membranes to study biomolecular interactions using complementary surface-sensitive techniques. *Advances in Colloid and Interface Science*, 277:102118, 2020.

[26] S. N. Cohen, A. C. Chang, H. W. Boyer, and R. B. Helling. Construction of biologically functional bacterial plasmids in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, 70(11):3240–3244, nov 1973.

[27] E. Collett 1934-. *Polarized light : fundamentals and applications / Edward Collett.* Marcel Dekker,, New York :, 1993.

[28] F. Costa, R. Silva, and A. R. Boccaccini. *Fibrous protein-based biomaterials (silk, keratin, elastin, and resilin proteins) for tissue regeneration and repair.* Elsevier Ltd., 2018.

[29] N. F. Crawford and R. M. Leblanc. Serum albumin in 2D: A Langmuir monolayer approach, 2014.

[30] R. Cubitt and G. Fragneto. Chapter 2.8.3 NEUTRON REFLECTION: PRINCIPLES AND EXAMPLES OF APPLICATIONS SCATTERING IN MICROSCOPIC PHYSICS AND CHEMICAL PHYSICS Neutron Scattering Contents. In *Neutron Reflection: Principles and Exampes of Applications.*

[31] A. Da Costa, A. M. Pereira, P. Sampaio, J. C. Rodríguez-Cabello, A. C. Gomes, M. Casal, and R. Machado. Protein-Based Films Functionalized with a Truncated Antimicrobial Peptide Sequence Display Broad Antimicrobial Activity. *ACS Biomaterials Science and Engineering*, 7(2):451–461, 2021.

[32] G. W. Daughdrill, M. S. Chadsey, J. E. Karlinsey, K. T. Hughes, and F. W. Dahlquist. The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, σ28. *Nature Structural Biology*, 4(4):285–291, apr 1997.

[33] J. A. De Feijter, J. Benjamins, and F. A. Veer. Ellipsometry as a tool to study the adsorption behavior of synthetic and biopolymers at the air–water interface. *Biopolymers*, 17(7):1759–1772, 1978.

[34] S. N. Deshmukh, A. M. Dive, R. Moharil, and P. Munde. Enigmatic insight into collagen, may 2016.

[35] C. Dicko, F. Vollrath, and J. M. Kenney. Spider silk protein refolding is controlled by changing pH. *Biomacromolecules*, 5(3):704–710, may 2004.

[36] L. J. Domigan, M. Andersson, K. A. Alberti, M. Chesler, Q. Xu, J. Johansson, A. Rising, and D. L. Kaplan. Carbonic anhydrase generates a pH gradient in Bombyx mori silk glands. *Insect Biochemistry and Molecular Biology*, 65:100–106, oct 2015.

[37] P. Drude. Ueber die Reflexion und Brechung ebener Lichtwellen beim Durchgang durch eine mit Oberflächenschichten behaftete planparallele Platte. *Annalen der Physik*, 279(5):126–157, jan 1891.

[38] B. I. Estate. Molecules for Langmuir—Blodgett film formation. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 330(1610):141–152, 1990.

[39] E. Fermi and W. H. Zinn. Reflection of Neutrons on Mirrors. *Physical Review*, 70(1–2):103, 1946.

[40] C. M. Fletcher, A. M. McGuire, A. C. Gingras, H. Li, H. Matsuo, N. Sonenberg, and G. Wagner. 4E binding proteins inhibit the translation factor eIF4E without folded structure. *Biochemistry*, 37(1):9–15, jan 1998.

[41] H. W. Fox, E. F. Hare, and W. A. Zisman. Wetting properties of organic liquids on high energy surfaces. *Journal of Physical Chemistry*, 59(10):1067–1106, 1955.

[42] B. Franklin. XLIV. Of the stilling of waves by means of oil. Extracted from sundry letters between Benjamin Franklin, LL. D. F. R. S. William Brownrigg, M. D. F. R. S. and the Reverend Mr. Farish. *Philosophical Transactions of the Royal Society of London*, 64:445–460, dec 1774.

[43] M. J. Gálvez-ruiz. Different approaches to study protein films at air/water interface. 247(July):533–542, 2017.

[44] A. Girotti, D. Orbanic, A. Ibáñez-Fonseca, C. Gonzalez-Obeso, and J. C. Rodríguez-Cabello. Recombinant Technology in the Development of Materials and Systems for Soft-Tissue Repair. *Advanced Healthcare Materials*, 4(16):2423–2455, nov 2015.

[45] J. I. Goldstein, D. E. Newbury, P. Echlin, D. C. Joy, C. E. Lyman, E. Lifshin, L. Sawyer, and J. R. Michael. The SEM and Its Modes of Operation. In *Scanning Electron Microscopy and X-ray Microanalysis*, pages 21–60. Springer US, 2003.

[46] E. Gorter and F. Grendel. On the spreading of proteins. *Transactions of the Faraday Society*, 22(0):477–483, jan 1926.

[47] F. Haguenau, P. W. Hawkes, J. L. Hutchison, B. Satiat-Jeunemaître, G. T. Simon, and D. B. Williams. Key Events in the History of Electron Microscopy. 2003.

[48] P. Hawkes. Recent advances in electron optics and electron microscopy, 2004.

[49] O. S. Heavens and S. F. Singer. Optical Properties of Thin Solid Films. *Physics Today*, 9(3):24–26, 1956.

[50] M. Henini. Scanning electron microscopy: An introduction. *III-Vs Review*, 13(4):40–44, jul 2000.

[51] S. Hénon and J. Meunier. Microscope at the Brewster angle: Direct observation of first-order phase transitions in monolayers. *Review of Scientific Instruments*, 62(4):936–939, 1991.

[52] C. Hessinger, G. M. Technau, and A. Rogulja-Ortmann. The Drosophila Hox gene Ultrabithorax acts in both muscles and motoneurons to orchestrate formation of specific neuromuscular connections . *Development*, 144(1):139–150, 2017.

[53] D. Hönig and D. Möbius. Direct visualization of monolayers at the air-water interface by Brewster angle microscopy. *Journal of Physical Chemistry*, 95(12):4590–4592, 1991.

[54] D. Hönig and D. Möbius. Reflectometry at the Brewster angle and Brewster angle microscopy at the air-water interface. *Thin Solid Films*, 210-211(PART 1):64–68, apr 1992.

[55] D. Howell, S.-P. Tsai, K. Churion, J. Patterson, K. Bayless, and S. E. Bondos. Mechanically-Tunable, Protein-Based Materials Can be Functionalized with Other Proteins and with DNA. *Biophysical Journal*, 110(3):338a, 2016.

[56] D. W. Howell, C. L. Duran, S. P. Tsai, S. E. Bondos, and K. J. Bayless. Functionalization of Ultrabithorax Materials with Vascular Endothelial Growth Factor Enhances Angiogenic Activity. *Biomacromolecules*, 17(11):3558–3569, 2016.

[57] D. W. Howell, S. P. Tsai, K. Churion, J. Patterson, C. Abbey, J. T. Atkinson, D. Porterpan, Y. H. You, K. E. Meissner, K. J. Bayless, and S. E. Bondos. Identification of multiple dityrosine bonds in materials composed of the Drosophila protein Ultrabithorax. *Advanced Functional Materials*, 25(37):5988–5998, 2015.

[58] H. C. Hsiao, K. L. Gonzalez, D. J. Catanese, K. E. Jordy, K. S. Matthews, and S. E. Bondos. The intrinsically disordered regions of the Drosophila melanogaster hox protein ultrabithorax select interacting proteins based on partner topology. *PLoS ONE*, 9(10), 2014.

[59] J. Huang, C. Wong, A. George, and D. L. Kaplan. The effect of genetically engineered spider silk-dentin matrix protein 1 chimeric protein on hydroxyapatite nucleation. *Biomaterials*, 2007.

[60] Z. Huang, Y. Lu, R. Majithia, J. Shah, K. Meissner, K. S. Matthews, S. E. Bondos, and J. Lou. Size dictates mechanical properties for protein fibers self-assembled by the drosophila hox transcription factor ultrabithorax. *Biomacromolecules*, 2010.

[61] Z. Huang, T. Salim, A. Brawley, J. Patterson, K. S. Matthews, and S. E. Bondos. Functionalization and patterning of protein-based materials using active ultrabithorax chimeras. *Advanced Functional Materials*, 21(14):2633–2640, 2011.

[62] C. L. Hughes and T. C. Kaufman. Hox genes and the evolution of the arthropod body plan, nov 2002.

[63] J. ichiro Jo, J. Q. Gao, and Y. Tabata. Biomaterial-based delivery systems of nucleic acid for regenerative research and regenerative therapy, dec 2019.

[64] R. Jansson, N. Thatikonda, D. Lindberg, A. Rising, J. Johansson, P. Å. Nygren, and M. Hedhammar. Recombinant spider silk genetically functionalized with affinity domains. *Biomacromolecules*, 15(5):1696–1706, may 2014.

[65] H. Jerrard. Ellipsometry and polarized light. *Optics & Laser Technology*, 10(2):100, 1978.

[66] D. Khago, J. C. Bierma, K. W. Roskamp, N. Kozlyuk, and R. W. Martin. Protein refractive index increment is determined by conformation as well as composition. *Journal of Physics Condensed Matter*, 30(43), 2018.

[67] D. P. Knight and F. Vollrath. Changes in element composition along the spinning duct in a Nephila spider. *Naturwissenschaften*, 88(4):179–182, 2001.

[68] S. R. Koebley, F. Vollrath, and H. C. Schniepp. Toughness-enhancing metastructure in the recluse spider's looped ribbon silk. *Materials Horizons*, 4(3):377–382, may 2017.

[69] H. S. Koh, T. Yong, C. K. Chan, and S. Ramakrishna. Enhancement of neurite outgrowth using nano-structured scaffolds coupled with laminin. *Biomaterials*, 29(26):3574–3582, sep 2008.

[70] S. Kyle, A. Aggeli, E. Ingham, and M. J. McPherson. Production of self-assembling biomaterials for tissue engineering, jul 2009.

[71] I. Langmuir. The constitution and fundamental properties of solids and liquids. II. Liquids. *Journal of the American Chemical Society*, 1917.

[72] I. Langmuir and V. J. Schaefer. Activities of Urease and Pepsin Monolayers. *Journal of the American Chemical Society*, 1938.

[73] C. H. Lee, A. Singla, and Y. Lee. Biomedical applications of collagen. *International Journal of Pharmaceutics*, 221(1-2):1–22, jun 2001.

[74] B. Li and V. Daggett. Molecular basis for the extensibility of elastin, 2002.

[75] T. Li, K. Lilja, R. J. Morris, and G. B. Brandani. Langmuir–Blodgett technique for anisotropic colloids: Young investigator perspective, 2019.

[76] Y. Liu, K. S. Matthews, and S. E. Bondos. Multiple intrinsically disordered sequences alter DNA binding by the homeodomain of the Drosophila Hox protein ultrabithorax. *Journal of Biological Chemistry*, 2008.

[77] Y. Liu, K. S. Matthews, and S. E. Bondos. Multiple intrinsically disordered sequences alter DNA binding by the homeodomain of the Drosophila Hox protein ultrabithorax. *Journal of Biological Chemistry*, 283(30):20874–20887, 2008.

[78] Y. Liu, K. S. Matthews, and S. E. Bondos. Internal Regulatory Interactions Determine DNA Binding Specificity by a Hox Transcription Factor. *Journal of Molecular Biology*, 2009.

[79] R. Magnusson. *Mueller matrix ellipsometry studies of nanostructured materials*. 2014.

[80] R. Majithia, J. Patterson, S. E. Bondos, and K. E. Meissner. On the design of composite protein-quantum dot biomaterials via self-assembly. *Biomacromolecules*, 2011.

[81] W. McGinnis, R. L. Garber, J. Wirz, A. Kuroiwa, and W. J. Gehring. A homologous protein-coding sequence in drosophila homeotic genes and its conservation in other metazoans. *Cell*, 37(2):403–408, jun 1984.

[82] T. L. McMEEKIN, M. L. GROVES, and N. J. HIPP. Refractive Indices of Amino Acids, Proteins, and Related Substances. (25):54–66, 1964.

[83] G. G. Mendes, R. M. Booth, D. L. Pattison, A. J. Alvarez, and S. E. Bondos. Generating Novel Materials Using the Intrinsically Disordered Protein Ubx. In *Methods in Enzymology*. 2018.

[84] S. T. Milner, J. F. Joanny, and P. Pincus. Buckling of langmuir monolayers. *EPL*, 9(5):495–500, 1989.

[85] H. Motschmann and R. Teppner. Ellipsometry in interface science. In *Studies in Interface Science*, volume 11, pages 1–42. 2001.

[86] S. W. Muchmore, M. Sattler, H. Liang, R. P. Meadows, J. E. Harlan, H. S. Yoon, D. Nettesheim, B. S. Chang, C. B. Thompson, S. L. Wong, S. C. Ng, and S. W. Fesik. X-ray and NMR structure of human Bcl-xL, an inhibitor of programmed cell death. *Nature*, 381(6580):335–341, may 1996.

[87] A. R. Murphy and D. L. Kaplan. Biomedical applications of chemically-modified silk fibroin. *Journal of Materials Chemistry*, 19(36):6443–6450, sep 2009.

[88] R. A. Neal, S. G. McClugage, M. C. Link, L. S. Sefcik, R. C. Ogle, and E. A. Botchwey. Laminin nanofiber meshes that mimic morphological properties and bioactivity of basement membranes. *Tissue Engineering - Part C: Methods*, 15(1):11–21, oct 2009.

[89] A. Nelson. Co-refinement of multiple-contrast neutron/X-ray reflectivity data using MOTOFIT. *Journal of Applied Crystallography*, 39(2):273–276, mar 2006.

[90] L. Névot and P. Croce. Caractérisation des surfaces par réflexion rasante de rayons X. Application à l'étude du polissage de quelques verres silicates. *Revue de Physique Appliquée*, 15(3):761–779, 1980.

[91] S. B. Nielsen, A. Lapierre, J. U. Andersen, U. V. Pedersen, S. Tomita, and L. H. Andersen. Absorption spectrum of the green fluorescent protein chromophore anion in vacuo. *Physical Review Letters*, 87(22):228102/1–228102/4, 2001.

[92] L. Nilebäck, S. Arola, M. Kvick, A. Paananen, M. B. Linder, and M. Hedhammar. Interfacial Behavior of Recombinant Spider Silk Protein Parts Reveals Cues on the Silk Assembly Mechanism. *Langmuir*, 34(39):11795–11805, 2018.

[93] C. W. Oatley. The early history of the scanning electron microscope. *Journal of Applied Physics*, 53(2):1, aug 1982.

[94] L. G. Parratt. Surface studies of solids by total reflection of x-rays. *Physical Review*, 95(2):359–369, jul 1954.

[95] J. L. Patterson, C. A. Abbey, K. J. Bayless, and S. E. Bondos. Materials composed of the Drosophila melanogaster protein ultrabithorax are

cytocompatible. *Journal of Biomedical Materials Research - Part A*, 102(1):97–104, 2014.

[96] J. L. Patterson, A. M. Arenas-Gamboa, T. Y. Wang, H. C. Hsiao, D. W. Howell, J. P. Pellois, A. Rice-Ficht, and S. E. Bondos. Materials composed of the Drosophila Hox protein ultrabithorax are biocompatible and nonimmunogenic. *Journal of Biomedical Materials Research - Part A*, 103(4):1546–1553, 2015.

[97] A. Pockels. Relations between the surface-tension and relative contamination of water surfaces, 1893.

[98] D. L. Price and K. Skold. Introduction to Neutron Scattering. *Methods in Experimental Physics*, 23(PA):1–97, jan 1986.

[99] J. T. Prince, K. P. McGrath, C. M. DiGirolamo, and D. L. Kaplan. Construction, Cloning, and Expression of Synthetic Genes Encoding Spider Dragline Silk. *Biochemistry*, 1995.

[100] J. J. Ramsden. Concentration scaling of protein deposition kinetics. *Physical Review Letters*, 71(2):295–298, jul 1993.

[101] J. M. Rodríguez Patino, C. Carrera Sánchez, and M. R. Rodríguez Niño. Morphological and structural characteristics of monoglyceride monolayers at the air-water interface observed by Brewster angle microscopy. *Langmuir*, 15(7):2484–2492, 1999.

[102] A. Rogulja-Ortmann, S. Renner, and G. M. Technau. Antagonistic roles for Ultrabithorax and Antennapedia in regulating segment-specific apoptosis of differentiated motoneurons in the Drosophila embryonic central nervous system. *Development*, 135(20):3435–3445, oct 2008.

[103] C. Roldán-Carmona, J. J. Giner-Casares, M. Pérez-Morales, M. T. Martín-Romero, and L. Camacho. Revisiting the Brewster Angle Microscopy: The relevance of the polar headgroup, 2012.

[104] P. Romero, Z. Obradovic, X. Li, E. C. Garner, C. J. Brown, and A. K. Dunker. Sequence complexity of disordered protein. *Proteins: Structure, Function and Genetics*, 42(1):38–48, 2001.

[105] J. Rosenbloom, W. R. Abrams, and R. Mecham. Extracellular matrix 4: the elastic fiber. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 1993.

[106] E. Ruska. The Development of the Electron Microscope and of Electron Microscopy (Nobel Lecture). *Angewandte Chemie International Edition in English*, 26(7):595–605, 1987.

[107] E. Ruska and H. O. Müller. über Fortschritte bei der Abbildung elektronenbestrahlter OberflÄchen. *Zeitschrift für Physik*, 116(5-6):366–369, may 1940.

[108] T. P. Russell. X-ray and neutron reflectivity for the investigation of polymers. *Materials Science Reports*, 5(4):171–271, 1990.

[109] T. P. Russell. On the reflectivity of polymers: Neutrons and X-rays. *Physica B: Condensed Matter*, 221(1-4):267–283, apr 1996.

[110] J. Sánchez-González, J. Ruiz-García, and M. J. Gálvez-Ruiz. Langmuir-Blodgett films of biopolymers: A method to obtain protein multilayers. *Journal of Colloid and Interface Science*, 2003.

[111] M. Saric and T. Scheibel. Engineering of silk proteins for materials applications, dec 2019.

[112] A. Schneider, X. Y. Wang, D. L. Kaplan, J. A. Garlick, and C. Egles. Biofunctionalized electrospun silk mats as a topical bioactive dressing for accelerated wound healing. *Acta Biomaterialia*, 5(7):2570–2578, sep 2009.

[113] M. P. Scott, J. W. Tamkun, and G. W. Hartzell. The structure and function of the homeodomain, jul 1989.

[114] K. Spiess, A. Lammel, and T. Scheibel. Recombinant spider silk proteins for applications in biomaterials, sep 2010.

[115] D. J. Stokes. Principles and Practice of Variable Pressure / Environmental Scanning, 2008.

[116] G. G. Stokes. On the Composition and Resolution of Streams of Polarized Light from different Sources. In *Mathematical and Physical Papers*, pages 233–258. Cambridge University Press, aug 2010.

[117] F. Teulé, A. R. Cooper, W. A. Furin, D. Bittencourt, E. L. Rech, A. Brooks, and R. V. Lewis. A protocol for the production of recombinant spider silk-like proteins for artificial fiber spinning. *Nature Protocols*, 2009.

[118] E. K. Tillinghast, S. F. Chase, and M. A. Townley. Water extraction by the major ampullate duct during silk formation in the spider, Argiope aurantia Lucas. *Journal of Insect Physiology*, 30(7):591–596, jan 1984.

[119] J. M. Trillo, E. I. Jado, S. G. Fernández, and P. S. Pedrero. Monolayers of human serum albumin - II. Effect of pH and ionic strength of substrate. *Kolloid-Zeitschrift & Zeitschrift für Polymere*, 250(4):325–329, 1972.

[120] R. Y. Tsien. The green fluorescent protein. *Annual review of biochemistry*, 67:509–544, 1998.

[121] A. Ulman. *An Introduction to Ultrathin Organic Films: From Langmuir-Blodgett to Self-Assembly*, volume San Diego. 2013.

[122] R. Valluzzi and S. P. Gido. The crystal structure of bombyx mori silk fibroin at the air-water interface. *Biopolymers - Nucleic Acid Sciences Section*, 42(6):705–717, 1997.

[123] R. Valluzzi, S. P. Gido, W. Muller, and D. L. Kaplan. Orientation of silk III at the air-water interface. In *International Journal of Biological Macromolecules*, volume 24, pages 237–242. Elsevier, mar 1999.

[124] D. Vollhardt. Brewster angle microscopy: A preferential method for mesoscopic characterization of monolayers at the air/water interface, 2014.

[125] J. E. Wagenseil and R. P. Mecham. Vascular Extracellular Matrix and Arterial Mechanics. *Physiological Reviews*, 2009.

[126] F. Wahaia. *Ellipsometry - Principles and Techniques for Materials Characterization*. InTech, nov 2017.

[127] T. Wang, J. H. Lai, and F. Yang. Effects of Hydrogel Stiffness and Extracellular Compositions on Modulating Cartilage Regeneration by Mixed Populations of Stem Cells and Chondrocytes in Vivo. *Tissue Engineering - Part A*, 22(23-24):1348–1356, dec 2016.

[128] M. Watanabe, Y. Kosaka, K. Oguchi, K. Sanui, and N. Ogata. Regulation of Supermolecular Structure of Amphiphilic Polymers by Means of the Langmuir-Blodgett Technique. *Macromolecules*, 21(10):2997–3003, 1988.

[129] N. Weichert, V. Hauptmann, C. Helmold, and U. Conrad. Seed-specific expression of spider silk protein multimers causes long-term stability. *Frontiers in Plant Science*, 7(JAN2016):6, jan 2016.

[130] H. Wen, X. Lan, Y. Zhang, T. Zhao, Y. Wang, Z. Kajiura, and M. Nakagaki. Transgenic silkworms (Bombyx mori) produce recombinant spider dragline silk in cocoons. *Molecular Biology Reports*, 2010.

[131] O. Wiener. Zur theorie der refraktionskonstanten. pages 256–277, 1910.

[132] L. Wilhelmy. Ueber die Abhängigkeit der Capillaritäts-Constanten des Alkohols von Substanz und Gestalt des benetzten festen Körpers. *Annalen der Physik*, 195(6):177–217, 1863.

[133] Y. Wu, H. Fan, C. Yang, and L. Zhang. Pyrene-based amphiphile regulated C60 aggregation in monolayers and Langmuir–Blodgett films. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 2020.

[134] H. Zhao, P. H. Brown, and P. Schuck. On the distribution of protein refractive index increments. *Biophysical Journal*, 100(9):2309–2317, 2011.