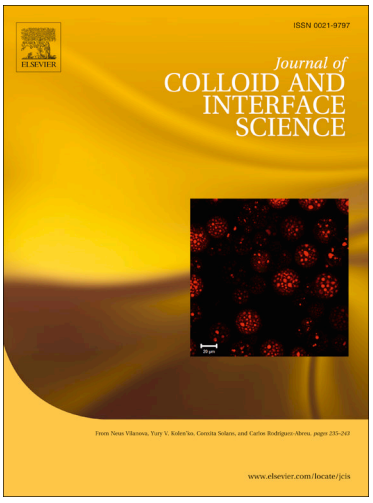



# Towards a molecular understanding of the water purification properties of *Moringa* seed proteins

M. Moulin, E. Mossou, L. Signor, S. Kieffer-Jaquinod, H.M. Kwaambwa, F. Nermark, P. Gutfreund, E.P. Mitchell, M. Haertlein, V.T. Forsyth, A.R. Rennie

PII: S0021-9797(19)30744-1
DOI: https://doi.org/10.1016/j.jcis.2019.06.071
Reference: YJCIS 25109

To appear in: *Journal of Colloid and Interface Science*

Received Date: 23 April 2019
Revised Date: 19 June 2019
Accepted Date: 20 June 2019

Please cite this article as: M. Moulin, E. Mossou, L. Signor, S. Kieffer-Jaquinod, H.M. Kwaambwa, F. Nermark, P. Gutfreund, E.P. Mitchell, M. Haertlein, V.T. Forsyth, A.R. Rennie, Towards a molecular understanding of the water purification properties of *Moringa* seed proteins, *Journal of Colloid and Interface Science* (2019), doi: https://doi.org/10.1016/j.jcis.2019.06.071

# Towards a molecular understanding of the water purification properties of *Moringa* seed proteins

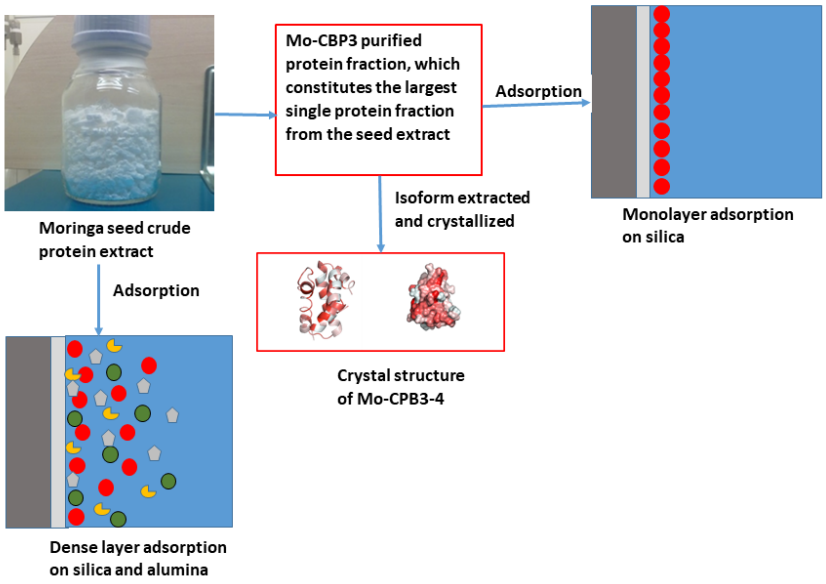
M. Moulin[1,2], E. Mossou[1,2], L. Signor[3], S. Kieffer-Jaquinod[4], H.M. Kwaambwa[5], F. Nermark[6], P. Gutfreund[1], E.P. Mitchell[2,7], M. Haertlein[1], V.T. Forsyth[1,2], A.R. Rennie[8*]

1. Institut Laue-Langevin, 71 Avenue des Martyrs, 38042 Grenoble Cedex 9, France
2. Faculty of Natural Sciences, Keele University, Staffordshire ST5 5BG, UK
3. Univ. Grenoble Alpes, CEA, CNRS, IBS, F-38000 Grenoble
4. Univ. Grenoble Alpes, CEA, Inserm, BGE U1038, F-38000, Grenoble, France
5. Faculty of Health and Applied Sciences, Namibia University of Science and Technology, Private Bag 13388, 13 Jackson Kaujeua Street, Windhoek, Namibia
6. Department of Chemistry, University of Botswana, Private Bag UB00704, Gaborone, Botswana
7. European Synchrotron Radiation Facility, 71 Avenue des Martyrs, 38043 Grenoble Cedex 9, France
8. Centre for Neutron Scattering, Uppsala University, Box 516, 75120 Uppsala, Sweden

## Abstract

*Seed extracts from Moringa oleifera are of wide interest for use in water purification where they can play an important role in flocculation; they also have potential as anti-microbial agents. Previous work has focused on the crude protein extract. Here we describe the detailed biophysical characterization of individual proteins from these seeds. The results provide new insights relating to the active compounds involved. One fraction, designated Mo-CBP3, has been characterized at a molecular level using a range of biochemical and biophysical techniques including liquid chromatography, X-ray diffraction, mass spectrometry, and neutron reflection. The interfacial behavior is of particular interest in considering water purification applications and interactions with both charged (e.g. silica) and uncharged (alumina) surfaces were studied. The reflection studies show that, in marked contrast to the crude extract, only a single layer of the purified Mo-CBP3 binds to a silica interface and that there is no binding to an alumina interface. These observations are consistent with the crystallographic structure of Mo-CBP3-4, which is one of the main isoforms of the Mo-CBP3 fraction. The results are put in context of previous studies of the properties of the crude extract. This work shows possible routes to development of separation processes that would be based on the specific properties of individual proteins.*

## Graphical abstract




Keywords: *Moringa oleifera* seeds; Mass spectrometry; Neutron reflectometry; X-ray diffraction; Water treatment

## 1. Introduction

Seeds from *Moringa* trees have long attracted attention as they can contribute to rural economies in a number of ways [1-3] as a source of oil and other products. A traditional application of the seeds has been to flocculate impurities in water using crushed seeds and this has attracted scientific studies for about 40 years. Aggregation of particulate impurities is the usual first essential step in water purification as this allows gravitational separation and filtering that removes a large fraction of the impurities and reduces the turbidity. The efficacy in terms of reducing turbidity and bacteria in treated water has been established, and the lack of toxic residues together with the ease of supply of the material makes development of this process attractive to provide clean water to small communities, particularly in developing countries. There are 14 known species of *Moringa* [2]. Most work has involved *Moringa oleifera* that is native to the Indian sub-continent and widespread in southern Africa, South America and a number of Asian countries. However, there are some reports of the use of *Moringa stenopetala* and *Moringa peregrina* that are found more commonly in other regions.

It is now clear that protein from the seeds is the active flocculating agent and a number of studies have focused on establishing an understanding at a microscopic level as to how this works. The usual means to achieve flocculation of colloids in water purification is either by screening and neutralization of charge using polyvalent salts, or by the addition of polymers that act as 'bridging' flocculants and bind to multiple particles. However, *Moringa* proteins are not thought to act in either these ways or as a depletion flocculating agent. The seed proteins are small, cationic and are difficult to denature in solution. They show a marked tendency to self-associate in aqueous solution [4] unless the concentration is very low. The model that emerges for flocculation is one that involves adsorption of protein to a wide range of different particles and this will then favor their association [5]. Essentially, the strong attractive interactions between the molecules are

transferred to the particles that are covered with protein. This allows effective coagulation and heterocoagulation of a broad range of particulate impurities.

Traditional applications of *Moringa* seeds would simply employ crushed seeds, and in some cases the seed oil will be separated for other uses. Investigations of the crude extract from seeds containing a mixture of proteins have been reported in many studies as this represents the material that would be exploited most frequently. Development of the process, optimization of the amount of material required, and studies as to how the seed protein could be used in practical conditions have been reported [6-13]. The binding of crude seed protein extracts to interfaces of silica [14] and alumina [15] from aqueous solutions has been investigated by neutron reflection and the results give a picture of a capability to adsorb to a broad range of different materials as multi-layers of protein molecules.

The details of the extraction of protein from the seeds have been identified as important. The most effective flocculating agent was obtained using a 1 M saline solution, although the specific type of cation appears to be less significant than its concentration [16,17]. The presence of a mixture of proteins with a composition that might depend in part on naturally variable growth conditions and extraction methods can explain partially the various reports of different molecular masses of the *Moringa* seed proteins.

It has been reported previously that there are two major protein components isolated from crude extract of M. *oleifera*: $MO_{2.1}$ the sequence of which was determined by Gassenschmidt et al. [18], and *Mo*-CBP3-1, which is the only *Moringa* seed protein with a crystal structure available [19]. *Mo*-CBP3-1 has been shown to belong to the 2S albumin family of proteins [20], which is characterized by a compact fold due to the presence of a small chain and a large chain linked together by four disulfide bonds. This arrangement is key to the marked stability of the protein and its heat and proteolysis resistance. Several studies have demonstrated the thermostability of both proteins which remain active for coagulation [21] and/or antifungal activities [22] after heat

treatment. This structural stability has been studied by circular dichroism, showing that there is no effect of heat treatment or pH variation on the secondary structure [23].

Beside their use as flocculation agents, *Moringa* seed extracts show antimicrobial activity [24]. This activity has been ascribed to *Moringa* synthesized derivatives of benzyl isothiocyanates, known antibacterial compounds, as well as to *Moringa* seed derived peptides [25]. *Moringa* seed powder can remove over 90% of the bacterial load of raw water samples. Water-soluble peptides released from the crushed seed kernels belong to a group of cationic peptides often displaying antimicrobial activity through its interaction with negatively charged surfaces and its amphiphilic structure allowing their incorporation into cellular membranes. The antibacterial activity of such a *Moringa* seed peptide (MO2.1 also called Flo) against human pathogens including *Pseudomonas aeruginosa* and *Streptococcus pyogenes* has been studied in detail [25]. The authors suggest distinct molecular mechanisms for sedimentation and bactericidal activity involving membrane destabilization by a hydrophobic loop. *Moringa* seed peptides might be promising alternatives to chemical water disinfection, chemical food preservatives or antibacterial treatments using antibiotics especially when the pathogen has become antibiotic-resistant. The behavior of the proteins as an antibacterial agent has also been investigated by cryo-EM on *E. coli* cells, showing a fusion of inner and outer membranes [26]. One strategy to exploit antibacterial activity of *M. oleifera* cationic peptide (MOCP) involves their immobilization on sand [27].

Comparative studies of a range of plant materials, and in particular other seed proteins, have indicated that *M. oleifera* is the most effective natural flocculating agent for water purification [27-29]. An understanding as to why particular seeds and particular proteins are more effective than others as a flocculating agent, or as an antimicrobial, will be important for the future development of applications from other plant varieties or from trees grown under different conditions. The interaction of *M. oleifera* proteins with amphiphiles [4,31,32] and competitive adsorption to interfaces with displacement from selected interfaces by rinsing with a cationic surfactant [15]

offers the prospect of wider applications for mineral separations. It has also been shown that the crude extract can be used to bind particles to substrates [34]. That study also reported quartz-crystal microbalance data for adsorption that confirmed that the bound layer of protein crude extract contained a significant fraction of water.

The present study uses a combination of biophysical techniques and activity assays to relate the properties and behavior of highly positively charged proteins extracted and purified from *M. oleifera* to those of the crude seed extract.

## 2. Materials and Methods

Unless specified otherwise, all chemicals used in this work were purchased from Sigma-Aldrich and were of analytical grade.

### 2.1 Protein extraction from seeds

The initial extraction of *Moringa* protein involves use of petroleum to remove oil, extraction of the protein with water, precipitation of protein with ammonium sulfate, filtration, dissolution of the precipitate in water, dialysis to remove excess of ammonium sulfate, adsorption on a carboxymethyl cellulose column and elution with 1 M NaCl, dialysis and finally freeze-drying. This procedure has been described by Kwaambwa and Maikokera [33].

Protein purification by ion exchange and size exclusion chromatography: 20 mg of freeze-dried crude seed extract was resuspended in 2 mL of carboxymethyl cellulose (CM) Sepharose buffer (10 mM ammonium acetate, pH 6.7), and loaded on 5 mL of CM handmade Sepharose column. A gradient from 0 to 60 % of CM Sepharose buffer with 1 M NaCl for 8 column volumes (CV) was applied. The elution fractions were analyzed on a 4-20% Tris-Tricine gradient gel Bio-Rad and fractions corresponding to the main peak were pooled together.

The pooled fractions from CM Sepharose were concentrated using Amicon ultra centrifugal units (Millipore) (MWCO 3K) to a final volume of 1 mL and then loaded on a Superdex 75 HR 10/ 30 gel filtration column (GE Healthcare) which was equilibrated with distilled water running at 0.5 mL $min^{-1}$.

## 2.2 Mass spectrometry (MS) analysis

MALDI-TOF mass spectrometry measurements were carried at the Institut de Biologie Structurale in Grenoble. Nano liquid chromatography mass spectrometry (nano LC-ESI) and tandem MS data were collected at the CEA Grenoble. Complete descriptions of the instrumental settings that were used and the data acquisition parameters are reported in the Supporting Information.

## 2.3 Determination of flocculation and coagulation activities

Simple tests of flocculation were made by adding purified *Moringa* protein or crude extract to dispersions of sulfonated polystyrene latex and observing cluster formation with an optical inverted microscope (Leica DMI1). The latex, PS3, was the same as that used in previous studies of flocculation using ultra small-angle neutron scattering made by Hellsing *et al* [5]. The latex has been extensively characterized using a variety of techniques [34] and has a zeta potential of about -35 mV. The final concentration of *Moringa* proteins was 2 mg $mL^{-1}$ and the concentration of particles was 4% by weight.

To test the coagulation activity on *E. coli* and *Nannochloropsis gaditana* cells stationary phase cultures were supplemented with 0.4 mg $mL^{-1}$ of *Moringa* proteins and the coagulation observed under a microscope.

## 2.4 Crystallization and X-ray diffraction

The screening conditions for crystallization of *Mo*-CBP3-4 protein were optimized using the HTX platform of the Partnership for Structural Biology, Grenoble France. 6 x 96 conditions were tested

at 20°C, varying buffer, salt and protein concentration [35]. The resulting conditions for crystallization were established to be 0.1 M citric acid; 2.4 M sodium formate pH5.0.

For data collection at 100 K, individual crystals were briefly soaked in 2.75 M sodium formate, 0.1 M citric acid, pH 5.0 with 10% glycerol prior to cryo-cooling. X-ray diffraction data were collected at ID29 [36] at the European Synchrotron Radiation Facility (ESRF), Grenoble, to a resolution of 1.68 Å for the high-resolution dataset and to 2.06 Å for the sulfur single-wavelength anomalous diffraction (SAD) dataset. Table S1 (Supporting Information) summarizes the data collection statistics. A helical data collection strategy was used to minimize the effects of radiation damage.

The data were indexed using space group $I4_122$ with $a = b = 108.07$ Å and $c = 43.62$ Å, consistent with the previously published crystal structure of Ullah *et al*. [19].

## 2.5 Neutron reflection

The interaction of the protein with mineral surfaces was investigated by measuring neutron reflectivity with the D17 reflectometer [37] at the Institut Laue-Langevin in Grenoble. The time-of-flight mode was used allowing reflectivity data to be measured at two incident beam angles, $\theta$, of 0.7 and 3.2 degrees with wavelengths between 2.5 and 25 Å. These parameters gave a useful range of momentum transfer, perpendicular to the interface, $Q$ (= $(4\pi/\lambda) \sin \theta$) of 0.006 to a maximum of 0.25 $Å^{-1}$, which depends on the level of the signal over the background. Typical measurement times for each reflectivity curve were about 1 h. In order to make measurements with a small amount of sample at two different interfaces, a special sample cell was used that had a reservoir for solution formed by a PTFE gasket between one crystal of silicon and another of sapphire [38]. The sample holder could be rotated by 180 degrees and translated so that either the silicon oxide surface on the silicon crystal or the alumina surface that were both vertical and in contact with the liquid sample could be used for reflection measurements. Data were reduced

to reflectivity curves by normalizing to measurements of the incident spectra that were transmitted through the relevant crystal using the *cosmos* software [39,40] available at the instrument. The program also subtracts the background that is observed on the instrument detector around the specularly reflected signal.

The aqueous solvent could be exchanged automatically using a Knauer HPLC pump connected to the filling port. In order to minimize the volumes of protein solution, the sample solutions were injected manually using a syringe and a three-way valve at the bottom of the sample cell. Initial measurements were made using pure $D_2O$ and pure $H_2O$ to characterize the oxide layer on the silicon crystal and any gel layer or roughness at the alumina surface. When changing contrast, at least 20 mL of solvent was flushed through the sample holder to an exit tube at the top of the cell. It was sufficient to inject 4 mL of purified protein solution at various concentrations to fill the sample cell or to increase the concentration.

## 3. Results

### 3.1 Protein purification by ion exchange and size exclusion chromatography

Two types of liquid chromatography were used to fractionate the proteins present in the crude extract, as shown in Figure 1. The first step, using ion exchange chromatography results in the identification of three main peaks. The peak of proteins, eluting at the highest salt concentration (peak C) carrying the highest positive charges, was then further purified using size exclusion chromatography (peak D). This material was used for all subsequent characterization apart from the mass spectrometry analysis.

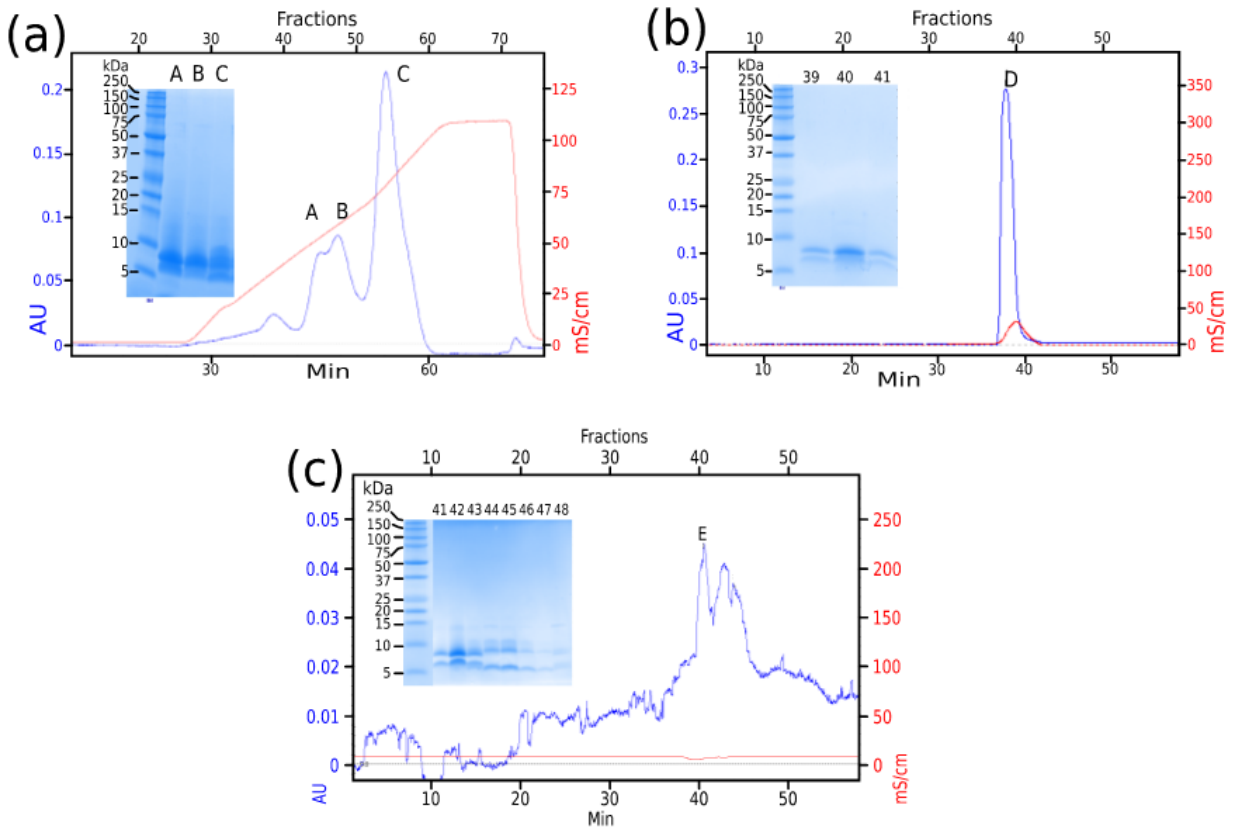



Figure 1: (a) Ion exchange chromatography results that identify three main peaks from the crude extract (labeled A, B, C). Peak C eluted at a concentration of 0.6 M NaCl. (b) Size exclusion chromatography carried out on material from peak C using a water-equilibrated S75 column. (c) Further fractionation obtained using the same column equilibrated with 50 mM NaCl. In all plots, the red trace shows the measured conductivity profile and the blue trace the absorbance at 280 nm.

These results show an unexpected change in the apparent size distribution of the protein components when run through the 50 mM NaCl equilibrated column. This difference in the two elution profiles (*i.e.* peaks D and E) could be attributed to salt-dependent interactions between *Moringa* proteins and the dextran polysaccharide component of the size exclusion column. Fractions 42 and 46 were further analyzed by mass spectrometry as described below.

## 3.2 Mass spectrometry (MS)

Different samples obtained from *M. oleifera* were analyzed by mass spectrometry, using both MALDI TOF MS and nano LC-ESI MS and MS/MS (See figure S2 in Supporting Information), to characterize and identify the main proteins present in the different extracts. The samples were

measured before and after reduction with TCEP (tris(2-carboxyethyl)phosphine). A summary of the mass spectrometric data obtained is given in Table 1. Analyzed fractions are shown in Figure 1 (c). Mass spectrometry analysis of the different fractions reveals a complex mixture of N- and C-terminal processed species of various *M. oleifera* seed protein isoforms (*Mo*-CBP3). Tandem MS measurements on the reduced chains (S: small chain and L: large chain) allowed the identification of the abundant isoforms in each fraction (see Table 1). In gel tryptic digestion of the crude extract resulted in the identification of three forms of *Mo*-CBP3 with the concentration decreasing in the following order: *Mo*-CBP3-4> *Mo*-CBP3-3>> *Mo*-CBP3-2. A pyroglutamate residue was identified at the N-terminal position of each individual chain for all *Mo*-CBP3 isoforms. Mass spectrometry analysis demonstrates the separation of different species after the size exclusion purification, with a predominance of *Mo*-CBP3-3 component present in fraction 42 and of *Mo*-CBP3-4 component present in fraction 46.

**Table 1**: Identification of the main protein species observed in different samples from *Moringa oleifera* by mass spectrometry analysis.

| **Sample** [(1)] | **Experimental mass (Da)** [(2)] | **Theoretical mass (Da)** [(3)] | **Sequence assignment** [(4)] | **Modifications** | **Protein isoform** |
|---|---|---|---|---|---|
| **without reduction** [(5)] | | | | | |
| **Fraction 42** | 11959.02 | 11956.78 | S + L | Gln→Pyr<br>(N-ter of S and L chains) | ***Mo*-CBP3-3** |
| **Fraction 46** | 11786.88 | 11785.63 | S + L | Gln→Pyr<br>(N-ter of S and L chains) | ***Mo*-CBP3-4** |
| **Crude extract** | 11785.90<br>11955.96 | 11785.63<br>11956.78 | S + L | Gln→Pyr<br>(N-ter of S and L chains) | ***Mo*-CBP3-3**<br>***Mo*-CBP3-4** |
| **Crystals** | 11824 | 11785.63 | S + L | Gln→Pyr<br>(N-ter of S and L chains)<br>2 oxidations on L chain | ***Mo*-CBP3-4** |
| **after reduction** [(5)] | | | | | |
| **Fraction 42** | 4083.04<br>7881.99 | 4082.57<br>7882.27 | small chain (S)<br>large chain (L) | Gln→Pyr<br>(N-ter of S and L chains) | ***Mo*-CBP3-3** |

| | | | | | |
|---|---|---|---|---|---|
| **Fraction 46** | 3777.90<br>8016.00 | 3777.28<br>8016.41 | small chain (S)<br>large chain (L) | Gln→Pyr<br>(N-ter of S and L chains) | ***Mo*-CBP3-4** |
| **Crude extract** | Nano LC-MS/MS analysis [6] | | | | ***Mo*-CBP3-2**<br>***Mo*-CBP3-3**<br>***Mo*-CBP3-4** |
| **Crystals** | 3776<br>8045 | 3777.28<br>8016.41 | small chain (S)<br>large chain (L) | Gln→Pyr<br>(N-ter of S and L chains)<br>2 oxidation sites on L chain | ***Mo*-CBP3-4** |

[1] Fractions 42 and 46 and crude extract were analyzed by nanoLC ESI MS and MS/MS; crystals were analyzed by MALDI TOF MS.
[2] Mass values: the mass of the highest specie of isotopic pattern for nanoLC ESI MS spectra and average mass for MALDI TOF MS spectra are reported respectively.
[3] Theoretical mass values calculated considering the formation of four disulfide bridges between the eight available cysteine residues present on the small (S) and large (L) chains.
[4] Mo-CBP3-3 isoform:
Small chain (S): Pyr-QQGQQQQQCRQQFLTHQRLRACQRFIRRQTQGGG
Large chain (L): Pyr-QARRPAIQRCCQQLRNIQPRCRCPSLRQAVQLAHQQQGQVGPQQVRQMYRLASNIPAICNLRPMSCPFG
Mo-CBP3-4 isoform:
Small chain (S): Pyr-QQQRCRHQFQTQQRLRACQRVIRRWSQGGGP
Large chain (L): Pyr-QARRPPTLQRCCRQLRNVSPFCRCPSLRQAVQSAQQQQGQVGPQQVGHMYRVASRIPAICNLQPMRCPFR
[5] Samples were measured without reduction and after reduction with TCEP.
[6] Nano LC-ESI-MS/MS analysis were carried out on the crude extract sample after running it on a 1D-SDS gel; twelve bands were cut and submitted to in-gel reduction and trypsin digestion before mass spectrometry.

## 3.3 Flocculation

Flocculation activity of the purified fraction D was visualized by standard light microscopy. Figure 2A shows that clear flocculation effects (large visible aggregates) are introduced by the addition of the purified seed proteins for the algae *N. gaditana*; the same effect is noted in 2B for *E. coli* bacteria. Figure 2C shows that, as with the crude extract, the purified protein causes flocculation of anionic polystyrene latex particles. Simple inspection with optical microscopy allows large aggregates to be identified. The particles are the same as those described in previous scattering studies [5] where compact flocs with a high fractal dimension were observed. These results demonstrate the flocculating activity in relation to a wide variety of materials.

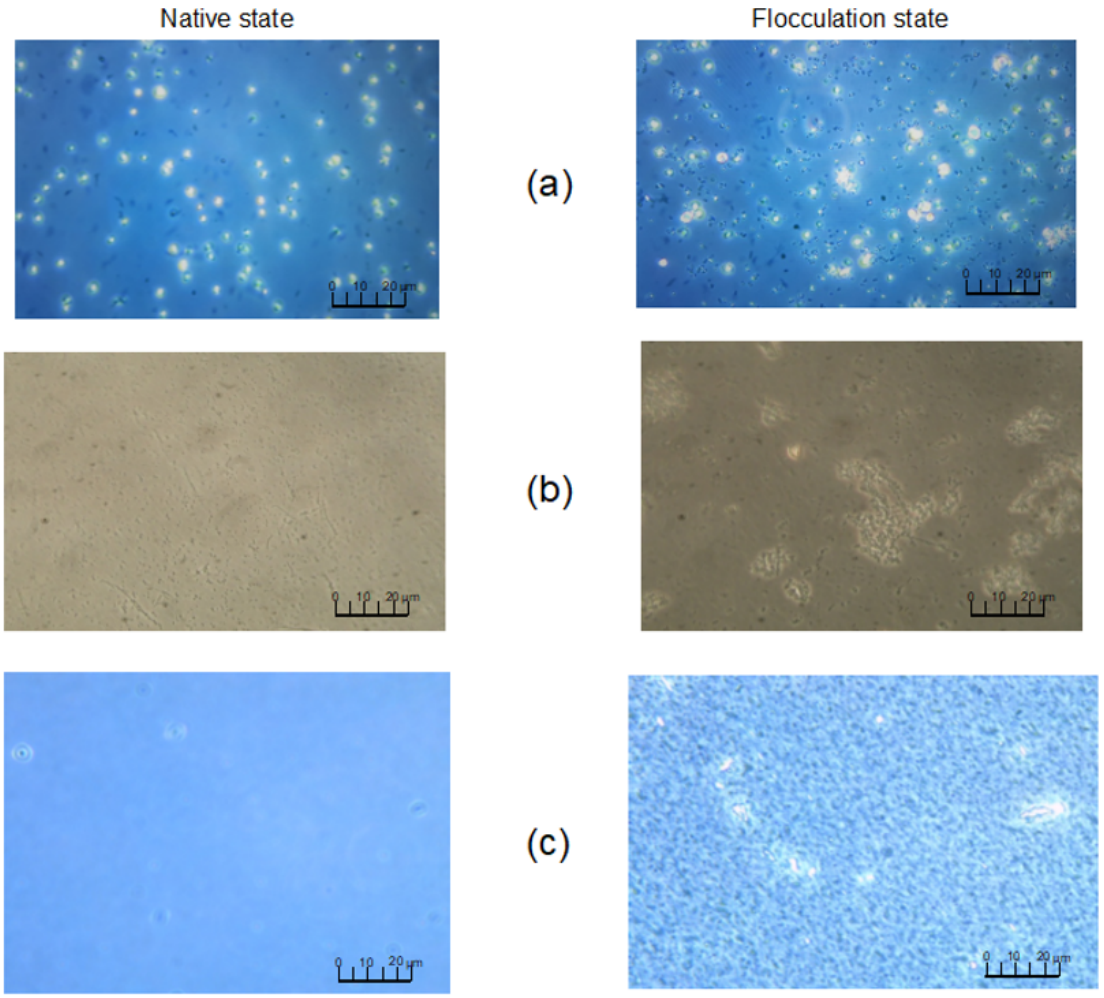



Figure 2: Microscopic images showing flocculation by the purified protein (peak D in Figure 1) for the three different materials: (a) the algae *Nannochloropsis gaditana*, (b) the bacterium *Escherichia coli* (BL21), (c) latex particles

### 3.4 Crystallographic structure of *Mo*-CBP3-4

In addition to the previous biochemical and biophysical characterization described above, crystallization trials were carried out in an attempt to get further structural insight into the water purification mechanism of the *Mo*-CBP3 proteins. Fraction D (Figure 1(b)) was successfully crystallized, and analysis of the crystals by MALDI TOF mass spectrometry demonstrated that the material corresponds to *Mo*-CBP3-4 (see Supporting Information, Figure S1).

The crystal structure determined here is essentially the same as that described by Ullah et al.[20] where the sample was mistakenly referred to as *Mo*-CPB3-1 instead of *Mo*-CPB3-4. Both sequences differ in two amino acid residues (highlighted in Figure 3a), according to Freire and co-workers [20] who have characterized the four isoforms of *Mo*-CBP3 (*Mo*-CPB3-1, *Mo*-CPB3-2, *Mo*-CPB3-3, *Mo*-CPB3-4).

The structural refinement and data statistics for the X-ray diffraction analysis are summarized in Table S1. Structure solution was determined by sulfur SAD. The values of $R_{merge}$ (0.035) completeness (99.6%), $R_{free}$ and $R_{work}$ (0.23/0.20) are good indicators of the high quality of the data. Most of the backbone of the protein was readily visible from the experimentally determined electron density map allowing unambiguous identification of 90 residues without resorting to sequence information. The molecular structure consists of 5 α-helices forming two chains, each having an intra-chain disulfide bond. In addition, there are two inter-chain disulfide bonds. The structure is compact and highly stable as with other 2S albumin proteins. Figure 3b shows the crystal structure, highlighting the hydrophobic residues (in red). While the core of the protein appears to be mostly hydrophobic, the solvent-exposed region, which contains a large number of arginine residues (10 out of 90 residues) is highly hydrophilic, as seen in Figure 4. The surface charge distribution of the protein (Figure 3c) confirms its high net positive charge, as well as the existence of both cationic and anionic residues. This is consistent with the high calculated isoelectric point of the protein (11.8). The crude extract was also observed to have a high isoelectric point, above 10, as noted by Kwaambwa and Rennie [4].

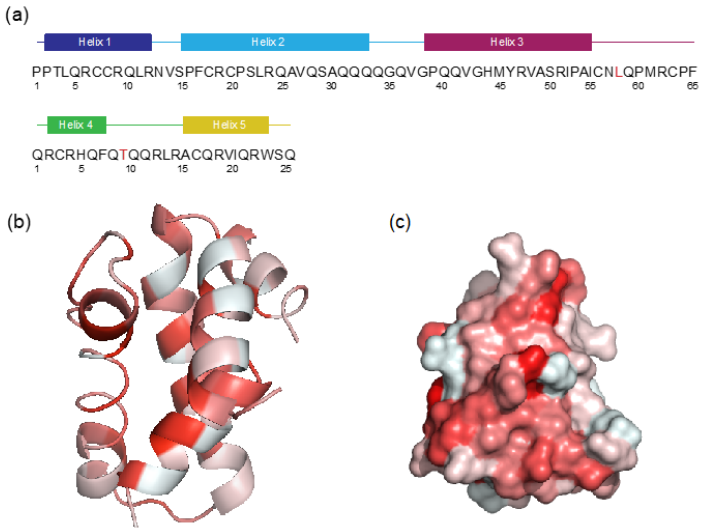



Figure 3: Primary sequence and the crystal structure of *Mo*-CPB3-4. (a) shows the primary sequence of *Mo*-CPB3-4 with the differences with *Mo*-CPB3-1 highlighted in red, the threonine on position 9 becomes a serine and the leucine in position 58 becomes a proline, in *Mo*-CPB3-1. (b) and (c) showing the helix and surface representation with polar versus non-polar areas of the protein. Red represents the most hydrophobic and white the most hydrophilic regions according to the Eisenberg hydrophobicity scale [41].

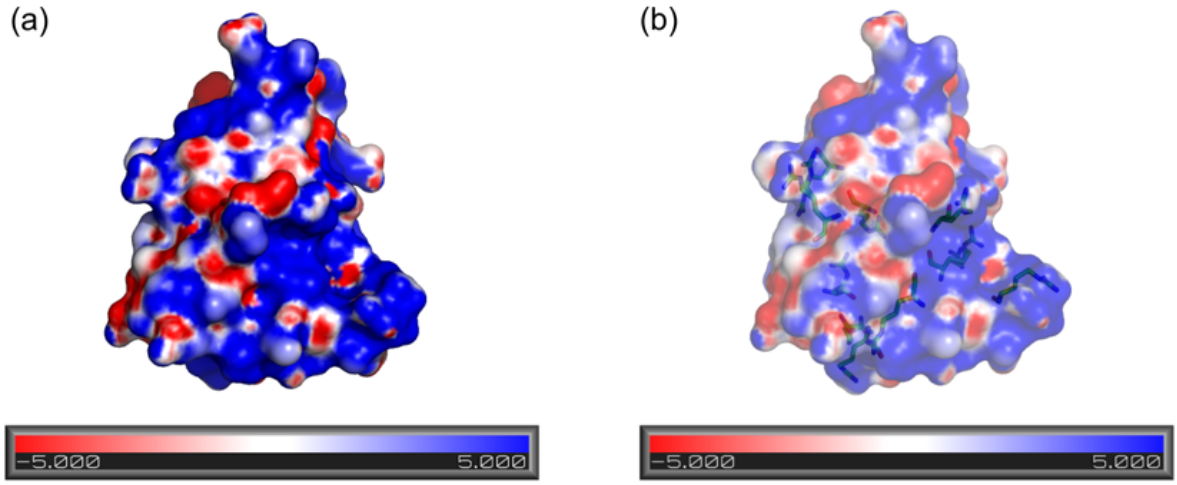



Figure 4: (a) Surface charge distribution of *Mo*-CBP3-4 where red, white and blue represent negative, neutral and positive charge respectively with the electrostatic potential isocontours going from $-5k_BT/e$ to $+5k_BT/e$, where $k_B$ is the Boltzmann constant, T is temperature and e is the electronic charge. (b) highlights the arginine residues. Both PyMOL [42] and APBS software [43] were used.

## 3.5 Surface behavior

### *Adsorption to silica*

Measurements of specular reflection of neutrons allow composition profiles at interfaces to be determined. Previous experiments [14] have identified that the crude extract binds to silica from

aqueous solution to form a dense layer with a thickness corresponding to multiple protein molecules near the surface, and a diffuse layer with decreasing density that extends towards the bulk solution. The results of the neutron reflection experiments show interesting behavior that distinguishes the purified fraction from that found previously. The data for the purified fraction are shown in Figure 5, and clearly demonstrate adsorption to the silica layer on the silicon substrate for a solution having a concentration as low as 0.025 mg mL$^{-1}$. These results show a marked difference from those observed for the clean interface, with a strong reduction in the reflectivity for the solution with protein in the Q range between 0.05 and 0.15 A$^{-1}$.

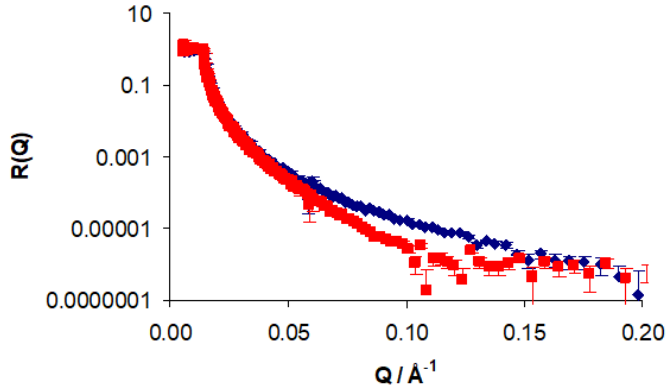



Figure 5. Neutron reflectivity data measured for the clean silicon/silica/$D_2O$ interface (♦) and the surface in contact with 0.03 mg mL$^{-1}$ solution of protein in $D_2O$ (■).

In contrast to previous results on crude extract, there was no change in the reflectivity as the concentration was raised. This indicated that there was no increase in the adsorbed amount at higher concentrations. This is illustrated in Figure 6, which shows that there is no change in the reflectivity as the concentration is increased to 0.5 mg mL$^{-1}$. In comparison, the crude extract showed an increasing adsorption until a plateau amount was reached at a concentration of 0.5 mg mL$^{-1}$.

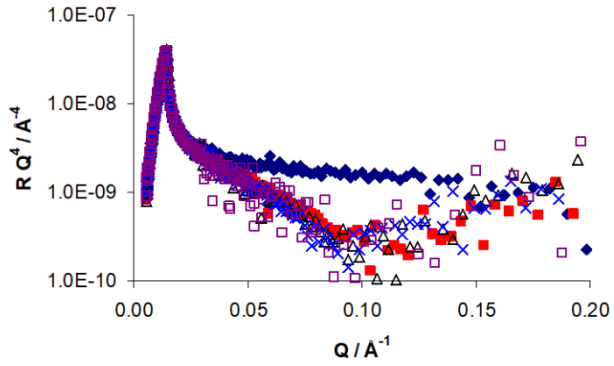



Figure 6. Reflectivity shown as RQ$^4$ against Q for different concentrations of protein (■ 0.0025 mg mL$^{-1}$, Δ 0.05 mg mL$^{-1}$, × 0.1 mg mL$^{-1}$, □ 0.5 mg mL$^{-1}$) at the oxide layer on a silicon substrate. For comparison, the data for the clean silicon/silica substrate are also shown (♦).

The clear change in reflectivity on adsorption of protein that is apparent in the data shown in Figures 5 and 6 can be interpreted quantitatively using optical matrix calculations of the reflectivity as implemented in the *cprof* program [44]. The plot of $RQ^4$ *vs.* Q allows some features in the data to be identified given that the effect of the large contrast difference between silicon and $D_2O$ is diminished. The minimum in the reflectivity at about Q = 0.12 $Å^{-1}$ indicates a well-defined layer. Rinsing the surface with $D_2O$ after adsorption did not change the reflectivity (see Supporting Information, Figure S3), showing that the adsorbed protein is not displaced with pure water. This observation also allowed measurements to be made with a second solvent contrast that can provide more details about the interfacial layer. The fit made to the combined data for the bound layer after rinsing first with $D_2O$ and then $H_2O$ is shown in Figure 7. The fitted curves (solid lines) correspond to a layer of 15.3 ± 1 Å that consists of 53 ± 3 % protein and 47 ± 3 % water. Allowance was made for exchange of protons in the protein as well as the different contrast of water in the fit to the data. The data are adequately modelled with a uniform protein layer having just 2 to 3 Å roughness at each of the interfaces.

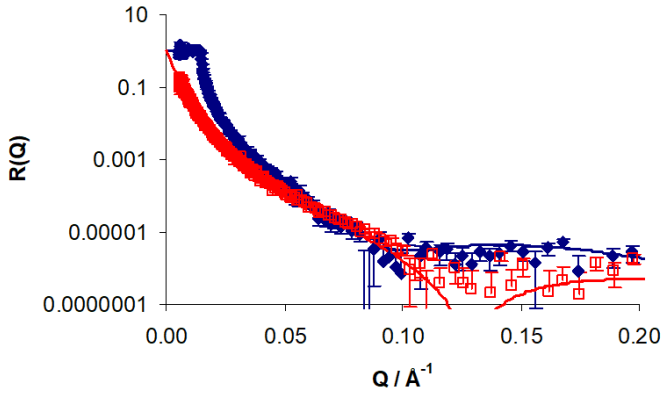



Figure 7. Reflection data for the adsorbed protein layer on the silica surface after rinsing measured in $D_2O$ (♦) and $H_2O$ (□) with the curves for the model described in the text (1.3 mg $m^{-2}$, 47% water, thickness 15 Å).

The surface coverage of protein can be calculated from the fit parameters as the product of the volume fraction and the thickness of the layer. It corresponds to 1.3 ± 0.2 mg $m^{-2}$. This is hydrated and corresponds to a single molecular layer of protein and contrasts with the multilayer adsorption that was found for the crude extract.

*Adsorption to alumina*

The design of the sample holder [34] allowed measurements to be made with the same protein solution at the alumina/solution interface by simply rotating the cell so as to reflect from a second solid/liquid interface. The data shown in Figure S4 (Supporting information) indicate that there is no significant change in the reflectivity, and this demonstrates that adsorption did not occur at this surface. This contrasts with the results found previously for the crude extract of *Moringa* seed protein [15] that had shown adsorption to alumina, albeit with a lower plateau coverage than observed at the silica interface.

## 4. Discussion

These results, which exploit measurements made using a wide variety of techniques, demonstrate clear differences in the behavior of the purified fraction of the *Moringa* seed protein from that of the crude extract. The observed flocculation activity for latex particles is in agreement with previously published work [5]. Sanchez-Martin *et al* [45] have shown that the highest flocculation activity for kaolinite corresponds to the most charged fractions of the extract. These observations imply that the valuable range of interactions that causes the *Moringa* seed proteins to flocculate a wide variety of materials may depend on a mixture of proteins being present. The surface of the latex particles has a small overall negative charge that arises from dissociation of sulfate groups: this explains why the particles are stabilized in aqueous dispersion and how flocculation occurs following combination with the positively charged protein. Similar interactions may also play a role in the flocculation/coagulation properties of the purified fraction for bacteria and algae, as seen in Figure 2.

The neutron reflection results, which show a clear interaction of the protein with the silica interface, and an essentially immeasurable interaction with the alumina interface, may be of general significance for an understanding of binding mechanisms relevant to applications involving selective separation of different particles in aqueous dispersions. These observations can be

understood in terms of the positively charged nature of the purified fraction and are consistent with the crystal structure of *Mo*-CBP3-4, the major component of the analyzed fraction. It is interesting to consider the crystal structure alongside the composition and arrangement of the bound protein at solid interfaces. The hydration of the crystal structure is about 54% - very similar to the 47% identified in the protein layer at the silica interface. We note that the interfacial behavior of the purified fraction occurs as a protein monolayer, whereas that for the crude extract appears to incorporate additional diffuse layers beyond the immediate surface, and this is apparently a consequence of the presence of a mixture of different molecules [14,15]. The molecular arrangement of the protein in the crystal, at neutral pH, suggests a fairly uniform distribution of positively charged groups; given the stability of the protein and its resistance to denaturation and degradation, the distribution of polar groups is likely to be similar in free solution. This may explain why the purified fraction does not associate as multilayers at surfaces, in contrast to the observed behavior of the crude extract. It may also explain the lack of adsorption on the uncharged alumina surface at neutral pH, in contrast to the clear binding to negatively charged silica.

Apart from providing understanding of the behavior a single protein, *Mo*-CBP3-4, at a molecular level, the present work suggests that the combination of different materials in the seed extract may be important in the overall use as a flocculating agent. The previous suggestions of a single sequence of peptides that give rise to flocculation [24,25,46] may be only part of an overall picture of interaction of multiple components together or of different peptides causing association of different impurities. Although this observation could at first be seen as a complexity in exploitation of *Moringa* seeds, it presents the possibility that different components of seed material could be used for selective flocculation and consequent separation of specific components. This could add further value to the use of seed proteins in the treatment of both drinking water and waste-water.

## 5. Conclusions

The key findings of this work provide a step in understanding the molecular basis of the water purification applications of the *Moringa* seed protein system. The results demonstrate clearly the importance of characterizing individual components of the seed proteins and their behavior. Whereas previous studies have focused on the characterization of less well-defined seed extracts, which may depend on species, harvesting conditions, and extraction methods, this study has investigated the *Mo*-CBP3 protein fraction, which constitutes the largest single protein fraction from the seed extract. It contains predominantly 2 isoforms (*Mo*-CBP3-3 and *Mo*-CBP3-4). In characterizing this specific fraction, it is important to note that *M. oleifera* seeds contain several other constituents in addition to the proteins described here and that the physical and biochemical properties of the extract such as flocculation, surface interaction, antimicrobial activity as a whole may therefore depend on multiple biomolecule interactions. The purified fraction adsorbs on silica as a single layer rather than as multilayers observed for the crude-extract [14]. This suggests that the presence of multiple components may be important to cause multilayer absorption. Compared to previous studies, the new findings have yielded substantially improved insights beyond those concerning the crude protein extract where material was found to adsorb to both silica and alumina [14,15]. For example, the explicit identification of molecular characteristics that can control adsorption to different interfaces is important for possible future exploitation of these materials in separation processes. Deeper understanding of which seed components bind to materials will also facilitate development of reusable filters, coated with *Moringa* proteins, which have been proposed for facile water treatment [48-51].

The composition of the bound layer close to the surface observed in this and in previous studies with both neutron reflectometry and quartz crystal microbalance measurements [34] that consists of almost 50% water corresponds approximately to the hydration within the crystal structure identified in the present work. This suggests that the maximum achievable protein density at a surface may already be achieved with present protocols. In future studies, it will be interesting to

explore whether different varieties of *Moringa* or different growth conditions for *M. oleifera* give rise to variations in the molecular composition in the seeds. Previous work has shown differences in the flocculation of particles in the presence of extracts from *M. stenopetala* and *M. oleifera* [5]. Further work may also benefit from molecular dynamics (MD) calculations that exploit the crystal structure information in identifying interfacial interactions at a molecular level. Recent MD studies of bovine serum albumin on silica surfaces have been described by Kubiak-Ossowska *et al* [52].

Additionally, recombinant expression of the different components may prove very useful for further detailed characterization; for example, the use of neutron scattering in conjunction with *in vivo* deuteration approaches [53,54] may provide valuable information on interactions within the crude extract as well as with other relevant systems such as phospholipid bilayers.

Acknowledgements

We acknowledge the platforms of the Grenoble Instruct-ERIC Center (ISBG : UMS 3518 CNRS-CEA-UGA-EMBL) with support from FRISBI (ANR-10-INBS-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). This study was also supported by the 'Investissement d'Avenir Infrastructures Nationales en Biologie et Santé' programme (ProFI project). We are also grateful for the allocation of beam time at the ILL for neutron measurements (DIR-141) and at the ESRF for macromolecular crystallography. HMK, FN and ARR thank the Swedish Research Council for support under Research Links project 348-2011-7241. V.T.F. acknowledges support from the EPSRC under grant numbers GR/R99393/01 and EP/C015452/1 which funded the creation of the Deuteration Laboratory (D-Lab) in the Life Sciences group of the ILL.

## References

[1] Morton, J. F. 1991. The Horseradish Tree, Moringa-Pterygosperma (Moringaceae) - a Boon to Arid Lands. *Economic Botany*, *45*(3), 318–333.

[2] Shindano, J., Kasase, C. 2009. Moringa (Moringa oleifera): A source of food and nutrition, medicine and industrial products. In *ACS Symposium Series* Vol. 1021, pp. 421–467.

[3] Anwar, F., Latif, S., Ashraf, M., Gilani, A. H. 2007. Moringa oleifera: A food plant with multiple medicinal uses. *Phytotherapy Research* 21, 17-25.

[4] Kwaambwa, H. M., Rennie, A. R. 2012. Interactions of surfactants with a water treatment protein from Moringa oleifera seeds in solution studied by zeta-potential and light scattering measurements. *Biopolymers*, *97*(4), 209–218.

[5] Hellsing, M. S., Kwaambwa, H. M., Nermark, F. M., Nkoane, B. B. M., Jackson, A. J., Wasbrough, M. J., Rennie, A. R. 2014. Structure of flocs of latex particles formed by addition of protein from Moringa seeds. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, *460*, 460–467.

[6] McConnachie, G. L., Folkard, G. K., Mtawali, M. A., Sutherland, J. P. 1999. Field trials of appropriate hydraulic flocculation processes. *Water Research*, *33*(6), 1425–1434.

[7] Beltrán-Heredia, J., Sánchez-Martín, J. 2009. Improvement of water treatment pilot plant with Moringa oleifera extract as flocculant agent. *Environmental Technology*, *30*(6), 525–534.

[8] Pritchard, M., Craven, T., Mkandawire, T., Edmondson, A. S., O'Neill, J. G. 2010. A comparison between Moringa oleifera and chemical coagulants in the purification of drinking water - An alternative sustainable solution for developing countries. *Physics and Chemistry of the Earth*, *35*(13–14), 798–805.

[9] Franco, M., Silva, G. K. e, Paterniani, J. E. S. 2012. Water treatment by multistage filtration system with natural coagulant from Moringa oleifera seeds. *Engenharia Agrícola*, *32*(5), 989–997.

[10] Bhatia, S., Othman, Z., Ahmad A. L. 2007. Pretreatment of Palm Oil Mill Effluent (POME) Using Moringa Oleifera Seeds as Natural Coagulant. *Journal of Hazardous Materials* 145(1–2): 120–26.

[11] Prasad, R. K. 2009. Color removal from distillery spent wash through coagulation using Moringa oleifera seeds: Use of optimum response surface methodology. *Journal of Hazardous Materials*, *165*(1–3), 804–811

[12] Santos, A. F. S., Argolo, A. C. C., Paiva, P. M. G., Coelho, L. C. B. B. 2012. Antioxidant activity of Moringa oleifera tissue extracts. *Phytotherapy Research: PTR*, *26*(9), 1366–1370.

[13] Baptista, A. T. A., Coldebella, P. F., Cardines, P. H. F., Gomes, R. G., Vieira, M. F., Bergamasco, R., Vieira, A. M. S. 2015. Coagulation-flocculation process with ultrafiltered saline extract of moringa oleifera for the treatment of surface water. *Chemical Engineering Journal*, *276*, 166–173.

[14] Kwaambwa, H. M., Hellsing, M., Rennie, A. R. 2010. Adsorption of a water treatment protein from moringa oleifera seeds to a silicon oxide surface studied by neutron reflection. *Langmuir*, *26*(6), 3902–3910.

[15] Kwaambwa, H. M., Hellsing, M. S., Rennie, A. R., Barker, R. 2015. Interaction of moringa oleifera seed protein with a mineral surface and the influence of surfactants. *Journal of Colloid and Interface Science*, *448*, 339–346.

[16] Madrona, G. S., Bergamasco, R., Seolin, V. J., Fagundes Klen, M. R. 2011. The Potential of Different Saline Solution on the Extraction of the Moringa oleifera Seed's Active Component for Water Treatment. *International Journal of Chemical Reactor Engineering*, *9*(1).

[17] Madrona, G. S., Serpelloni, G. B., Salcedo Vieira, A. M., Nishi, L., Cardoso, K. C., Bergamasco, R. 2010. Study of the effect of Saline solution on the extraction of the Moringa oleifera seed's active component for water treatment. *Water, Air, and Soil Pollution*, *211*(1–4), 409–415.

[18] Gassenschmidt, U., Jany, K. D., Bernhard, T., Niebergall, H. 1995. Isolation and characterization of a flocculating protein from Moringa oleifera Lam. *BBA - General Subjects*, *1243*(3), 477–481.

[19] Ullah, A., Mariutti, R. B., Masood, R., Caruso, I. P., Gravatim Costa, G. H., Millena de Freita, C., Arni, R. K. 2015. Crystal structure of mature 2S albumin from Moringa oleifera seeds. *Biochemical and Biophysical Research Communications*, *468*(1–2), 365–371.

[20] Freire, J. E. C., Vasconcelos, I. M., Moreno, F. B. M. B., Batista, A. B., Lobo, M. D. P., Pereira, M. L., Grangeiro, T. B. 2015. Mo-CBP3, an antifungal chitin-binding protein from Moringa oleifera seeds, is a member of the 2S albumin family. *PLoS ONE*, *10*(3), 1–24. https://doi.org/10.1371/journal.pone.0119871

[21] Ghebremichael, K. A., Gunaratna, K. R., Henriksson, H., Brumer, H., Dalhammar, G. 2005. A simple purification and activity assay of the coagulant protein from Moringa oleifera seed. *Water Research*, *39*(11), 2338–2344.

[22] Gifoni, J. M., Oliveira, J. T. A., Oliveira, H. P. H. D., Batista, A. B., Pereira, M. L., Gomes, A. S., Vasconcelos, I. M. 2012. A novel chitin-binding protein from Moringa oleifera seed with potential for plant disease control. *Biopolymers*, *98*(4), 406–415.

[23] Batista A. B., Oliveira, J. T. A., Gifoni, J. M., Pereira, M. L., Almeida, M. G. G., Gomes, V. M., Vasconcelos, I. M. 2014. New insights into the structure and mode of action of Mo-CBP3, an antifungal chitin-binding protein of Moringa oleifera seeds. *PLoS ONE*, *9*(10), 1–9.

[24] Madsen, M., Schlundt, J., Omer, E. F. E. 1988. Effect of water coagulation by seeds of Moringa oleifera on bacterial concentrations. *Journal of Ethnopharmacology*, *22*(3), 329.

[25] Suarez, M., Haenni, M., Canarelli, S., Fisch, F., Chodanowski, P., Servis, C., Mermod, N. 2005. Structure-function characterization and optimization of a plant-derived antibacterial peptide. *Antimicrobial Agents and Chemotherapy*, *49*(9), 3847–3857.

[26] Shebek, K., Schantz, A. B., Sines, I., Lauser, K., Velegol, S., Kumar, M. 2015. The flocculating cationic polypetide from moringa oleifera seeds damages bacterial cell membranes by causing membrane fusion. *Langmuir*, *31*(15), 4496–4502.

[27] Jerri, H. A., Adolfsen, K. J., McCullough, L. R., Velegol, D., Velegol, S. B. 2012. Antimicrobial sand via adsorption of cationic Moringa oleifera protein. *Langmuir*, *28*(4), 2262–2268.

[28] Bodlund, I, 2013. *Coagulant Protein from plant materials: Potential Water Treatment Agent* Licentiate Thesis. KTH, Sweden.

[29] Bodlund, I., Pavankumar, A. R., Chelliah, R., Kasi, S., Sankaran, K., Rajarao, G. K. 2014. Coagulant proteins identified in Mustard: A potential water treatment agent. *International Journal of Environmental Science and Technology*, *11*(4), 873–880.

[30] Conaghan, S. M.Sc. Thesis, Dublin City University (2013) '*Development and analysis of environmentally neutral, biodegradable, novel flocculants for drinking water treatment*'

[31] Maikokera, R., Kwaambwa, H. M. 2007. Interfacial properties and fluorescence of a coagulating protein extracted from Moringa oleifera seeds and its interaction with sodium dodecyl sulphate. *Colloids and Surfaces B: Biointerfaces*, *55*(2), 173–178.

[32] Kwaambwa, H.M, Nermark, F. M. 2013. Interactions in Aqueous Solution of a Zwitterionic Surfactant with a Water Treatment Protein from Moringa oleifera Seeds Studied by Surface Tension and Ultrasonic Velocity Measurements. *Green and Sustainable Chemistry*, *3*(4), 135–140.

[33] Kwaambwa, H. M., Maikokera, R. 2008. Infrared and circular dichroism spectroscopic characterisation of secondary structure components of a water treatment coagulant protein extracted from Moringa oleifera seeds. *Colloids and Surfaces B: Biointerfaces*, *64*(1), 118–125.

[34] Nouhi, S., Pascual, M., Hellsing, M. S., Kwaambwa, H. M., Skoda, M. W. A., Höök, F., Rennie, A. R. 2018. Sticking particles to solid surfaces using *Moringa oleifera* proteins as a glue *Colloids and Surfaces B: Biointerfaces* 168, 68-75.

[35] Rennie, A. R., Hellsing, M. S., Wood, K., Gilbert, E. P., Porcar, L., Schweins, R., Malfois, M. 2013. Learning about SANS instruments and data reduction from round robin measurements on samples of polystyrene latex. *Journal of Applied Crystallography*, *46*(5), 1289–1297.

[36] Dimasi, N., Flot, D., Dupeux, F., Márquez, J. A. 2007. Expression, crystallization and X-ray data collection from microcrystals of the extracellular domain of the human inhibitory receptor expressed on myeloid cells IREM-1. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, *63*(3), 204–208.

[37] De Sanctis, D., Beteva, A., Caserotto, H., Dobias, F., Gabadinho, J., Giraud, T., Mueller-Dieckmann, C. 2012. ID29: A high-intensity highly automated ESRF beamline for macromolecular crystallography experiments exploiting anomalous scattering. *Journal of Synchrotron Radiation*, *19*(3), 455–461.

[38] Cubitt, R., Fragneto, G. 2002. D17: the new reflectometer at the ILL *Appl. Phys.* A 74 (Suppl.) S329–S331.

[39] Rennie, A. R., Hellsing, M. S., Lindholm, E., Olsson, A. 2015. Note: Sample cells to investigate solid/liquid interfaces with neutrons. *Review of Scientific Instruments*, *86*(1).

[40] COSMOS, (2017). URL: https://www.ill.eu/instruments-support/instruments-groups/instruments/d17/more/documentation/d17-lamp-book/cosmos/

[41] Gutfreund, P., Saerbeck, T., Gonzalez, M. A., Pellegrini, E., Laver, M., Dewhurst, C., Cubitt, R. 2018. Towards generalized data reduction on a chopper-based time-of-flight neutron reflectometer. *Journal of Applied Crystallography*, *51*(3).

[42] Eisenberg, D., Weiss, R. M., Terwilliger, T. C. 1984. The hydrophobic moment detects periodicity in protein hydrophobicity. *Proceedings of the National Academy of Sciences*, *81*(1), 140–144.

[43] The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC

[44] Baker, N. A., Sept, D., Joseph, S., Holst, M. J., McCammon, J. A. 2001. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences*, *98*(18), 10037–10041.

[45] Rennie A.R, (2015) http://www.reflectometry.net/fitprogs/cprof.htm

[46] Sánchez-Martín, J., Beltrán-Heredia, J., Peres, J. A. 2012. Improvement of the flocculation process in water treatment by using Moringa oleifera seeds extract. *Brazilian Journal of Chemical Engineering*, *29*(3), 495–501.

[47] Suarez, M., Entenza, J. M., Doerries, C., Meyer, E., Bourquin, L., Sutherland, J., Mermod, N. 2003. Expression of a Plant-Derived Peptide Harboring Water-Cleaning and Antimicrobial Activities. *Biotechnology and Bioengineering*, *81*(1), 13-20.

[48] Jerri, H. A., Adolfsen, K. J., McCullough, L. R., Velegol, D., Velegol, S. B. 2012. Antimicrobial Sand via Adsorption of Cationic *Moringa oleifer*a Protein *Langmuir*, 28, 2262-2268.

[49] Williams, F. E., Lee, A. K., Orandi, S., Sims, S. K., Lewis, D. M. 2017. *Moringa oleifera* functionalised sand - reuse with non-ionic surfactant dodecyl glucoside. *J Water Health* 15, 863-872.

[50] Mndelemani, C., Song, Y., Wei, F., Chen, Y. 2018. Chemically Enhanced Primary Treatment-Trickling Filter with *Moringa oleifera* Seeds for Improved Energy Recovery from Wastewater Treatment Plants in Malawi. 2018. *International Journal of Environmental Protection and Policy* 6, 26-31.

[51] Xiong, B., Piechowicz, B., Wang, Z., Marinaro, R., Clement, E., Carlin, T., Uliana, A., Kumar, M., Velegol, S. B. 2018.*Moringa oleifera* f-sand Filters for Sustainable Water Purification. *Environ. Sci. Technol. Lett.* 5, (2018), 38-42.

[52] Kubiak-Ossowska, K., Tokarczyk, K., Jachimska, B., Mulheran, P. A. 2017. Bovine Serum Albumin Adsorption at a Silica Surface Explored by Simulation and Experiment. *Journal of Physical Chemistry B*, *121*(16), 3975–3986.

[53] Haertlein, M., Moulin, M., Devos, J. M., Laux, V., Dunne, O., Forsyth, V. T. 2016. Chapter Five - Biomolecular Deuteration for Neutron Structural Biology and Dynamics. *Methods in Enzymology* 566, 113–157.

[54] Dunne, O., Weidenhaupt, M., Callow, P., Martel, A., Moulin, M., Perkins, S. J., Forsyth, V. T. 2017. Matchout deuterium labelling of proteins for small-angle neutron scattering studies using prokaryotic and eukaryotic expression systems and high cell-density cultures. *European Biophysics Journal: EBJ*, *46*(5), 425-432.