research papers

SOLEIL shining on the solution-state structure of biomacromolecules by synchrotron X-ray footprinting at the Metrology beamline

A. Baud,[a,b] L. Aymé,[c] F. Gonnet,[a,b] I. Salard,[a,b] Y. Gohon,[c] P. Jolivet,[c] K. Brodolin,[d] P. Da Silva,[e] A. Giuliani,[f,g] B. Sclavi,[h] T. Chardot,[c] P. Mercère,[e] P. Roblin[g]‡ and R. Daniel[a,b]*

[a]CNRS, UMR8587, Laboratoire Analyse et Modélisation pour la Biologie et l'Environnement, 91025 Evry, France, [b]Université d'Evry-Val-d'Essonne, Laboratoire Analyse et Modélisation pour la Biologie et l'Environnement, 91025 Evry, France, [c]INRA, AgroParisTech, UMR1318, Institut Jean-Pierre Bourgin, 78000 Versailles, France, [d]CPBS, CNRS UMR 5236-UM1/UM2, BP 14491, 34093 Montpellier Cedex 5, France, [e]Metrology Beamline, Synchrotron Soleil, L'Orme des Merisiers, Saint-Aubin, BP 48, 91192 Gif-sur-Yvette, France, [f]Disco Beamline, Synchrotron Soleil, L'Orme des Merisiers, Saint-Aubin, BP 48, 91192 Gif-sur-Yvette, France, [g]INRA, UAR1008 Caractérisation et Élaboration des Produits Issus de l'Agriculture, F-44316 Nantes, France, and [h]LBPA, CNRS UMR 8113, ENS Cachan, 94235 Cachan, France. *Correspondence e-mail: regis.daniel@univ-evry.fr

Synchrotron X-ray footprinting complements the techniques commonly used to define the structure of molecules such as crystallography, small-angle X-ray scattering and nuclear magnetic resonance. It is remarkably useful in probing the structure and interactions of proteins with lipids, nucleic acids or with other proteins in solution, often better reflecting the *in vivo* state dynamics. To date, most X-ray footprinting studies have been carried out at the National Synchrotron Light Source, USA, and at the European Synchrotron Radiation Facility in Grenoble, France. This work presents X-ray footprinting of biomolecules performed for the first time at the X-ray Metrology beamline at the SOLEIL synchrotron radiation source. The installation at this beamline of a stopped-flow apparatus for sample delivery, an irradiation capillary and an automatic sample collector enabled the X-ray footprinting study of the structure of the soluble protein factor H (FH) from the human complement system as well as of the lipid-associated hydrophobic protein S3 oleosin from plant seed. Mass spectrometry analysis showed that the structural integrity of both proteins was not affected by the short exposition to the oxygen radicals produced during the irradiation. Irradiated molecules were subsequently analysed using high-resolution mass spectrometry to identify and locate oxidized amino acids. Moreover, the analyses of FH in its free state and in complex with complement C3b protein have allowed us to create a map of reactive solvent-exposed residues on the surface of FH and to observe the changes in oxidation of FH residues upon C3b binding. Studies of the solvent accessibility of the S3 oleosin show that X-ray footprinting offers also a unique approach to studying the structure of proteins embedded within membranes or lipid bodies. All the biomolecular applications reported herein demonstrate that the Metrology beamline at SOLEIL can be successfully used for synchrotron X-ray footprinting of biomolecules.

1. Introduction

The genomic era has resulted in the vast accumulation of data on DNA and protein sequences. In the current post-genomic era, attention is more focused on how these macromolecules assemble into complex edifices and intermolecular networks (Harrison, 2004). Hence, it becomes important to develop new tools to understand macromolecule assemblies and biomolecular machines, and to probe their dynamics in solution or even in the cell. These large macromolecular complexes are
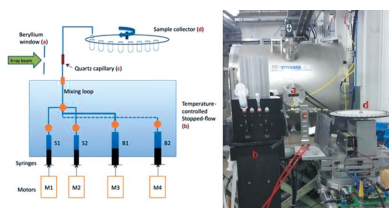
most often resistant to the usual bio-structural methods, such as NMR and X-ray crystallography, given their high molecular weight, flexible dynamic conformation, and association with membranes (Robinson *et al.*, 2007). In this context, there is a need to develop new experimental and data analysis approaches. The use of X-ray synchrotron-mediated hydroxyl radical footprinting (XFP) has increasingly gained interest in the past few years to probe the structure and dynamics of macromolecular complexes involving nucleic acids (Sclavi *et al.*, 1997; Ralston *et al.*, 2000) and proteins (Kiselar *et al.*, 2002). XFP can also be used for *in vivo* analysis as recently reported for the study of ribosomal assembly (Clatterbuck Soper *et al.*, 2013; Hulscher *et al.*, 2016). With regard to proteins, synchrotron protein footprinting together with the development of high-resolution mass spectrometry (MS) for peptides emerges as a powerful structural proteomics technique that can shed light on the conformations and dynamics of macromolecular complexes in solution, and provides information at a single amino-acid residue resolution that was previously inaccessible using traditional approaches (Maleknia & Downard, 2014).

The hydroxyl radical (•OH) is an efficient labelling reagent that reacts on short timescales with a wide range of amino acids and with the phosphodiester backbone of nucleic acids (Sclavi *et al.*, 1997). •OH radicals can be generated by the radiolysis of water molecules, and due to their small size they are an ideal reagent to finely probe the solvent accessible surfaces of macromolecules. Compared with well known chemical methods able to produce •OH (Xu & Chance, 2007; Hambly & Gross, 2005), synchrotron-mediated water radiolysis does not require extra chemical reagents such as hydrogen peroxide that could alter the biomolecule structure or interfere with the subsequent MS analysis (Kiselar *et al.*, 2002). Furthermore, synchrotron-mediated water radiolysis allows a rapid generation of •OH, which is of great interest for time-resolved structural biology (Kiselar *et al.*, 2002). Indeed, exposure of biomolecules in aqueous solution to hard X-rays leads to the production of hydroxyl radicals on the microsecond to millisecond timescales necessary for probing the structural dynamics of macromolecules. It is worth noting that, combined with rapid sample mixing, synchrotron-mediated hydroxyl radical footprinting allows for time-resolved kinetic studies of structural changes and folding events of macromolecules. Moreover, short exposure times can minimize the occurrence of secondary and parasitic reactions (Kiselar & Chance, 2010).

X-ray synchrotron protein footprinting is thus a relevant method for investigating the structures of macromolecules, as well as the sites of macromolecular interactions, provided that the labelled regions can be identified and quantified. Regarding nucleic acids, their reaction with •OH leads to backbone cleavage that can be monitored by gel electrophoresis. The oxidation of amino-acid side chains by •OH in proteins usually yields a mass increase of +16 Da (or a higher multiple), corresponding to the incorporation of oxygen (Xu & Chance, 2007). This mass increment can be easily detected by high-resolution MS. MS allows the identification of the oxidized amino-acid side chains by tandem mass spectrometry (MS/MS) analysis as well as the quantitative determination of the oxidation level by liquid-chromatography mass spectrometry (LC-MS). The differences in oxidation levels mirror changes in solvent accessibility as a result of conformational rearrangements in response to folding events, ligand binding or biomolecular interactions (Kamal *et al.*, 2007). The obtained structural and conformational data can be further used as constraints for molecular modelling (Padayatti *et al.*, 2013).

X-ray synchrotron protein footprinting has previously been applied successfully to study the folding of nucleic acid (Sclavi *et al.*, 1997; Ralston *et al.*, 2000), the kinetics of protein–DNA interactions (Sclavi *et al.*, 2005; Rogozina *et al.*, 2009), protein–ligand binding (Kiselar *et al.*, 2003), dynamics of membrane–protein (Zhu *et al.*, 2009; Padayatti *et al.*, 2013), protein dimers (Deperalta *et al.*, 2013) and protein complexes (Kamal *et al.*, 2007) or even assemblies of cytoplasmic compounds *in vivo* (Clatterbuck Soper *et al.*, 2013). Pioneering studies were performed at beamline X28C of the Center for Synchrotron Biosciences, at National Synchrotron Light Source, Brookhaven, USA (Gupta *et al.*, 2007), and are now continued at the Advanced Light Source, Berkeley, USA (Gupta *et al.*, 2015), with an improved experimental set-up and higher labelling efficiency (Gupta *et al.*, 2014, 2016), while a new XFP beamline is awaited soon at the National Synchrotron Light Source II (NSLS-II) in partnership with the Center for Synchrotron Biosciences, Case Western Reserve University, USA (Gupta *et al.*, 2016). Comparatively very few X-ray synchrotron footprinting studies have been carried out in Europe, mainly focused on protein–DNA footprinting at the European Synchrotron Radiation Facility, Grenoble, France (Rogozina *et al.*, 2009; Sclavi *et al.*, 2005).

In the present study, we have evaluated the possibility of using the white radiation of the Metrology beamline (Ménesguen & Lépy, 2011) at the SOLEIL synchrotron radiation facility to carry out hydroxyl radical footprinting experiments on proteins. The Metrology beamline is formally dedicated to the metrology of X-ray optical components and detectors, and instrumentation development. It is located on a bending magnet and consists of two branches (a soft X-ray and a hard X-ray branch) covering an energy range between 30 eV and 40 keV, in monochromatic operational configurations (Fig. 1). Moreover, the hard X-ray branch offers the possibility to access the white beam emitted by the bending magnet, thus allowing the use of a broadband spectrum ranging from the visible wavelengths to X-rays (60 keV in Fig. 2). These X-rays features meet the requirements for footprinting experiments (Bohon *et al.*, 2014), such as high-flux-density X-rays, large beam dimensions to cover the sizes of the sample in the capillary, homogeneous intensity profiles of the beam and sufficient space in the hard X-ray branch experimental hutch to install a stopped-flow apparatus, a capillary flow cell, a water bath circulation thermostat and a sample collector. Additionally, a nearby biology laboratory provides access to all the equipment needed for preparing the experiment, such as electrophoresis chambers, spectrophotometers, freezers *etc.*

We describe here the Metrology beamline with its experimental environment and facilities to achieve successful footprinting experiments on biomolecules, and present different ways to analyze the results, using proprietary or open access software, together with examples of our recent results as a proof of concept.

2. Material and methods

2.1. Metrology beamline description

Covering the largest part of the accessible synchrotron spectrum, the Metrology beamline at the SOLEIL synchrotron is a facility specifically designed for the characterization and development of X-ray optical components and detectors. Due to its versatility, the beamline is also valuable as a general purpose beamline allowing preparing, testing and implementing a wide range of experiments, developing new techniques and their related instrumentation.

The beamline uses the synchrotron radiation emitted by a bending magnet (magnetic radius of 5.3 m), and is composed of two branches operating simultaneously: a soft X-ray branch (30–1800 eV) and a hard X-ray branch (100–40 keV) which also allows access to the white beam. Fig. 1 shows the beamline layout with its two branches. The facility is composed of three hutches: a common optical hutch, gathering the required equipment for monochromatization and primary spatial conditioning of the beams; and two independent experimental hutches for final X-ray beam conditioning on the sample and where subsequent experiments are conducted.

Biomolecule footprinting experiments were performed on the hard X-ray branch. Reduced to its simplest geometrical configuration by removing from the beam path all beamline optical components (hard X-ray monochromator, focusing mirrors and filters), this branch allows one to benefit from the high X-ray flux and broadband spectrum given by the synchrotron source.

The size of the source is 140 μm (horizontal, H) × 60 μm (vertical, V) FWHM, with a nominal beam divergence of 8 mrad (H) × 1.4 mrad (V). The angular acceptance of the hard X-ray branch is defined to 1.8 mrad (H) × 0.2 mrad (V) by a water-cooled fixed-size copper aperture located 12.4 m downstream from the source. Three pairs of motorized tungsten slits along the beam path allow then for remote adjustment of the beam divergence and size. The hard X-ray experimental hutch is equipped with three experimental stations: a two-axis high-vacuum reflectometer, an optical table and a deep X-ray photolithography (LIGA) exposure station. The optical table is generally used to carry out experiments at atmospheric pressure. It is composed of a granite table (2250 mm × 1050 mm × 200 mm) supported by three independent vertical motorized translation stages, allowing for height, pitch and roll adjustments of the table. This table is also equipped with various motorized and modular positioning systems, each offering up to six degrees of freedom. Given its high versatility, this experimental station is one of the most requested of the beamline, capable of carrying out experiments that are not possible elsewhere.

The biomolecules footprinting experiment was thus implemented on this optical table, at 31.2 m distance from the source, behind a 150 μm-thick Be window acting as an interface between the under-vacuum part of the beamline and atmospheric pressure. With the beamline slits fully opened, the beam size at this location reaches 56 mm (H) × 6.2 mm (V). The beam intensity profiles, given by the bending magnet and electron beam properties, are flat and Gaussian in the horizontal and vertical dimensions, respectively.

Fig. 2(a) shows the experimental set-up as implemented on the hard X-ray branch of the Metrology beamline. Fig. 2(b) gives the calculated beam footprint using *SpotX* software (Moreno & Idir, 2001), together with the calculated spectrum at the sample position. The integrated photon flux over the beamline overall solid angle is about 2×10^{16} photons s^{-1}; the integrated flux impinging on the solution over a 10 mm × 1 mm area is about 6.3×10^{14} photons s^{-1}.

2.2. Experimental set-up and X-ray footprinting experiments

The QFM-400 Micro-Volume Quench Flow (Bio-Logic SAS) stopped-flow system was used to deliver the samples in front of the beam. The stopped-flow was modified to create a protocol that allows
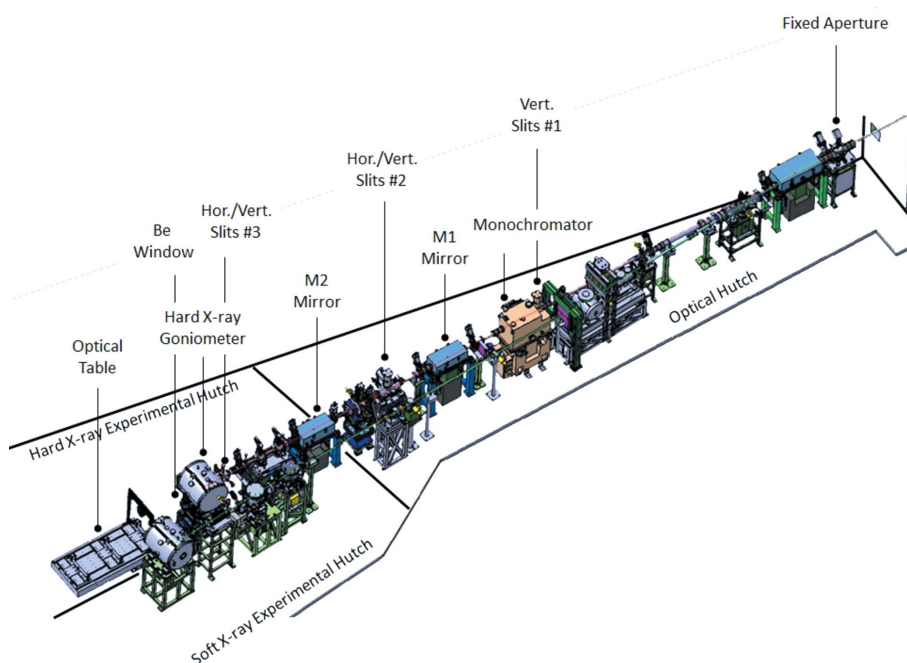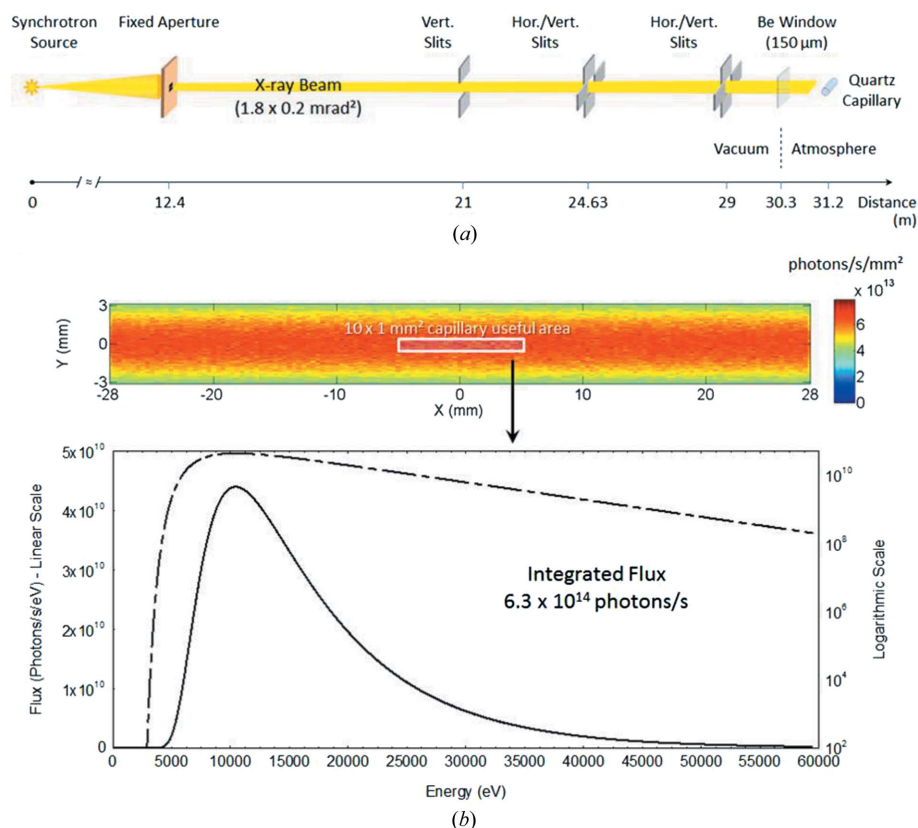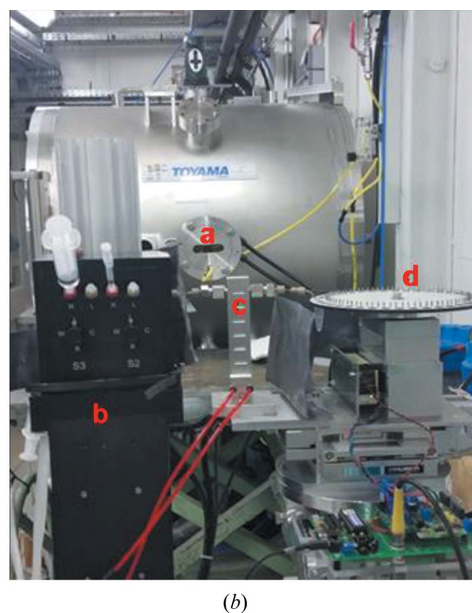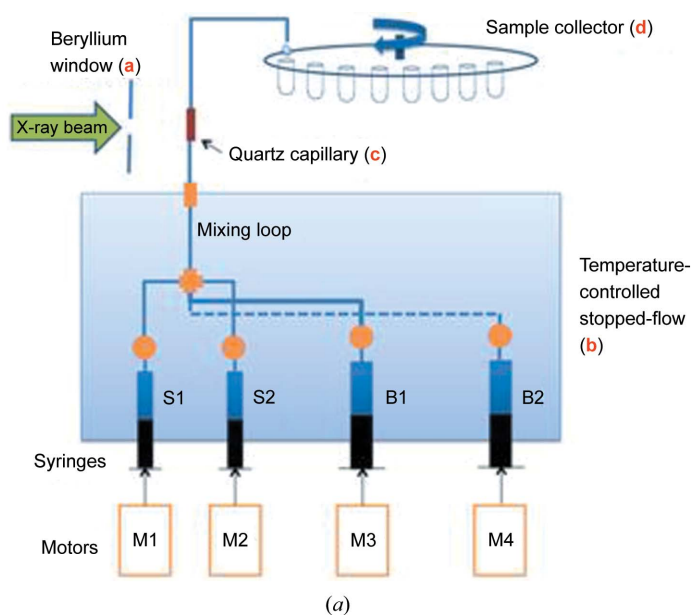

Figure 1
Layout of the Metrology beamline at SOLEIL (see details in text).

research papers

Figure 2
Experimental set-up implemented on the hard X-ray branch of the Metrology beamline. (*a*) Geometrical set-up for the biomolecules footprinting experiments on the hard X-ray branch of the Metrology beamline; (*b*) calculated pink beam footprint at the sample position for 430 mA current and considering the transmissions of a 150 µm-thick Be window, a 900 mm air path and a 10 µm-thick quartz capillary. Curves were calculated using the *SpotX* software, as described previously (Moreno & Idir, 2001). The white rectangle represents the useful area of the capillary. The corresponding spectrum impinging on the solution is given on linear (continuous line) and logarithmic (dashed line) scales.

for sequential exposure of multiple samples and an intervening washing step with water or buffer between one sample and the next. The QFM has four independent syringes, two 1 ml, one 5 ml and one 10 ml. The sample(s) are loaded in the 1 ml syringes (S1 and S2 in Fig. 3*a*), while buffer or water is loaded in the other two (B1 and B2 in Fig. 3*a*). A new exit hole was added on the side of the machine so that the sample could be rapidly discharged from the stopped-flow into the X-ray exposed quartz capillary (1 mm outer diameter, 0.8 mm inner diameter, 8 mm length). The samples are collected in a home-made sample collector whose movement is synchronized with the stopped-flow from a signal produced by the QFM controller (Fig. 3). The samples in syringes S1 and S2 can be mixed for variable amounts of time starting from 10 ms. The resulting sample mix is then pushed in front of the beam by syringe B1. The exposure time is controlled by the speed at which the sample passes across the X-ray beam in the capillary. Syringe B2 is used to rinse the tubing before another sample is mixed and exposed. The sample collector rotates to collect either irradiated sample or the outflow from the washing step. For each

Figure 3
Sample delivery set-up at the Metrology beamline. (*a*) Schematic representation of the stopped-flow apparatus and sample collector. (*b*) Photograph of the stopped-flow set-up placed in front of the beam: a, Be window insuring beamline exit at atmospheric pressure; b, stopped-flow system for sample delivery; c, temperature-controlled sample holder and quartz capillary; d, sample collector.

J. Synchrotron Rad. (2017). **24**, 576–585   A. Baud *et al.* • Solution-state structure of biomacromolecules by X-ray footprinting **579**

experiment at least two samples are collected in the absence of irradiation for the controls.

2.3. Beamline settings validation

The production of hydroxyl radicals at the Metrology beamline was evaluated by monitoring the radiolytic degradation of Alexa 488 Fluorophore (Thermo-Fisher) upon X-ray exposure as described elsewhere (Gupta *et al.*, 2007). 18 μl aliquot of 1.5 μ*M* solution of Alexa 488 in MilliQ water (Millipore) was loaded into the QFM stopped-flow apparatus and injected into the quartz capillary for irradiation. Exposure times ranged from 2.5 to 100 ms, and experiments were carried out at 18°C. After exposure, free hydroxyl radicals were immediately quenched by methionine at a final concentration of 10 m*M*, and samples were stored in ice before fluorescence measurements using a Fluoromax-4 apparatus (Horiba Jobin Yvon, Longjumeau France) with 1.4 ml quartz cuvettes (Hellma, France).

2.4. Biomolecule samples and synchrotron X-ray exposure

2.4.1. Factor H and C3b proteins from the human complement system. All XFP experiments were performed in triplicate. Complement system proteins factor H (FH, UniProtKB P08603) and C3b (UniProtKB P01024) purified from human plasma, obtained from Calbiochem, were diluted in phosphate-buffered saline (PBS; 10 m*M* phosphate buffer, 2.7 m*M* potassium chloride and 137 m*M* sodium chloride, pH 7.4) (Sigma-Aldrich) to obtain a concentration of 2 μ*M*. A 14 μl aliquot of the 2 μ*M* FH solution was loaded into the instrument and mixed with 0.01 *M* PBS or C3b solution at a 1:1 molar ratio for 1 s in the mixing chamber. Given the previously determined K_d value (Perkins *et al.*, 2012), the 2 μ*M* FH and C3b preparations insured the quantitative formation of the FH–C3b complex. Complex formation was also confirmed by cross-linking studies (data not shown). The reaction mixture was delivered to the quartz capillary placed in front of the beam, and FH samples in the absence or presence of C3b were exposed to various radiation times ranging from 0 to 150 ms. All experiments were carried out at 21°C. Immediately after exposure, free hydroxyl radicals were quenched with 20 μl of 100 m*M* methionine (10 m*M* final concentration) and the samples were stored at −80°C.

2.4.2. S3 oleosin: solvent accessibility of a protein anchored in a half membrane. S3 oleosin (At4g25140, accession P29525 OLEO1_ARATH in UniProtKB data bank) is the major protein found in *Arabidopsis thaliana* lipid droplets. As described previously (Vindigni *et al.*, 2013), we have designed an expression system targeting S3 oleosin to *Saccharomyces cerevisiae* lipid droplets permitting to characterize the protein fold in native-like yeast lipid droplets environment using various biophysical methods. After purification, yeast lipid droplets harboring S3 oleosin were dialyzed against 10 m*M* sodium phosphate buffer at pH 7.5.

The emulsion of lipid droplets containing ∼5 μ*M* of S3 oleosin was irradiated from 2.5 to 100 ms. Subsequently, samples were collected and immediately mixed with methionine (10 m*M* final) and stored at −80°C.

2.5. Post-footprinting processing and bio-analysis

The irradiated protein samples were analysed using a proteomic bottom-up approach combining proteolytic digestion and nanoLC-MS/MS analysis in proteomic facilities located within 30 km radius of SOLEIL. After irradiation, samples were stored at −80°C and transported on dry ice to ensure their stability. In-depth proteomic bottom-up analysis of the samples was not possible at the SOLEIL site, since it requires long sample preparation and analysis time.

With regard to the FH, irradiated FH samples in PBS were vacuum-dried and diluted in 14 μl ultra-pure water. Trypsin digestion was carried out as previously described (Baud *et al.*, 2016) and was terminated by freezing the samples. Prior to nanoLC-MS/MS analysis, protein digests were vacuum-dried and re-suspended in 14 μl of 98% H_2O, 2% CH_3CN, 0.1% HCOOH (solvent A). NanoLC-MS/MS analyses were performed at the LAMBE MS platform (Laboratoire Analyse et Modélisation pour la Biologie et l'Environnement, Evry, France) on a Dual Gradient Ultimate 3000 chromatographic system (Dionex) using a C18 pre-column and a C18 capillary column (Acclaim PepMap C18, 15 cm length × 75 μm inner diameter, 3 μm particle size, 100 Å porosity, Dionex) coupled to a LTQ-Orbitrap™ XL mass spectrometer (Thermo-Fisher Scientific) with a nanospray source operating in positive ionization mode. LC-MS/MS parameters were set for acquiring a data-dependent scan for side-chain modifications induced by •OH radicals, as described before (Baud *et al.*, 2016).

For S3 oleosin, vacuum-dried irradiated samples were dissolved in 15 μl 4X Laemmli buffer and loaded on 12% ready-to-use Nu PAGE gels (Novex, San Diego). SDS-PAGE was carried following the Laemmli protocol (Laemmli, 1970) and using 3-[*N*-morpholino]propane sulfonic acid (MOPS) running buffer. The gel was stained with Coomassie blue (Brilliant blue G-250, Sigma-Aldrich) according to Neuhoff *et al.* (1988). We verified that protein backbone integrity was not affected during irradiation by scanning the gels using an EPSON Perfection 4990 PHOTO scanner, and the TIFF resulting files were analyzed using *MultiGauge* software from Fujifilm (St Quentin en Yvelines, France) (Fig. S1 of the supporting information). The protein bands were excised and destained using 50 m*M* NH_4HCO_3/CH_3CN (1/1 *v*/*v*). In-gel digestion by trypsin or chymotrypsin (sequencing grade, Roche Diagnostics, Meylan, France) was carried out as previously reported (Vermachova *et al.*, 2011). Peptides were extracted using 5% HCOOH in water/CH_3CN (1/1 *v*/*v*). Prior to nanoLC-MS/MS the protein digests were vacuum-dried and resuspended in 20 μl of 0.1% (*v*/*v*) trifluoroacetic acid. NanoLC-MS/MS analysis of S3 oleosin was carried out on the PAPPSO platform (Jouy-en-Josas, France) with a similar protocol as above except that the nanoLC system was coupled to a Q-Exactive mass spectrometer (Thermo-Fisher Scientific) using a nanoelectrospray interface.

2.6. Bio-informatics data analysis

Mass spectrometric raw data of the FH peptides were processed using *ProtMapMS* software (Kaur *et al.*, 2009) and manually validated using *Xcalibur* software (Thermo-Fisher Scientific) and the Mascot server.

The proportion of unmodified peptide defined as the ratio of unmodified to the total amount of peptide (modified and unmodified) was calculated for each exposure time. The rate of oxidation of a given peptide was determined from the dose-response plots of the proportion of unmodified peptide as a function of X-ray exposure time (ms). The dose-response plots were fit to

$$y = \exp(-kt), \qquad (1)$$

where k is the first-order rate constant and t is the exposure time. The dose-response curves generated for each FH peptide were used to compare the oxidation extent of FH in both states (with and without C3b) and calculate the degree of protection. The solvent accessible surface area of modified FH residues was calculated using *GETAREA* software (Fraczkiewicz & Braun, 1998) based on the available in-solution and crystal structural data of modules and whole FH.

S3 oleosin *Xcalibur* raw data were transformed to mzXML open source format using *ProteoWizard* software version 3.0. Subsequently, data were processed using PAPPSO's open source software: *!Xtandem* pipeline (available at http://pappso.inra.fr/bioinfo/) for peptide identification and detection of oxidized residues and *MassChroQ* (Valot *et al.*, 2011). The latter was used to quantify oxidized and non-oxidized peptide by peak area integration under the extracted ion chromatogram.

3. Results and discussion

3.1. Beamline evaluation

A fluorescent assay based on the Alexa 488 fluorophore was used to determine the reproducibility of the X-ray exposure conditions over different experimental sessions and when using different sample holder configurations. The radiolysis of Alexa 488 induced a decrease of fluorescence, which can be easily measured as an indirect evaluation of the X-ray irradiation dose (Gupta *et al.*, 2007). An overall good reproducibility was achieved for experiments carried out six months apart (Fig. 4), showing the reliability of our set-up. The intensity of Alexa 488 fluorescence was decreased by half as a result of exposure to the X-ray beam for only 2.5 ms; no more significant remaining fluorescence (<5%) could be measured after 15 ms of irradiation. The rate constant for the decrease in fluorescence with exposure time was calculated using the following equation: $y = y_0 + A^{-kt}$, where k is the rate constant and t is the exposure time. The value of A was calculated to be 0.994 ± 0.0256, and $k = 292.1 \text{ s}^{-1}$, $R^2 = 0.992$.

3.2. Applications to a large spectrum of biomolecules

3.2.1. Hydroxyl radical protein footprinting of the full size FH molecule. The complement system is a biological cascade
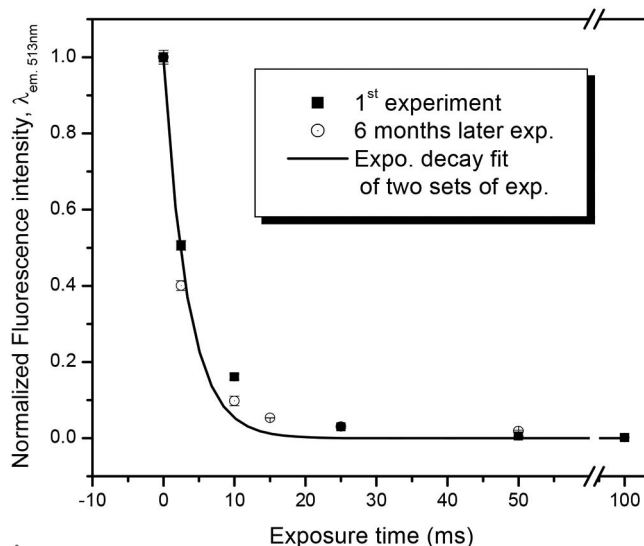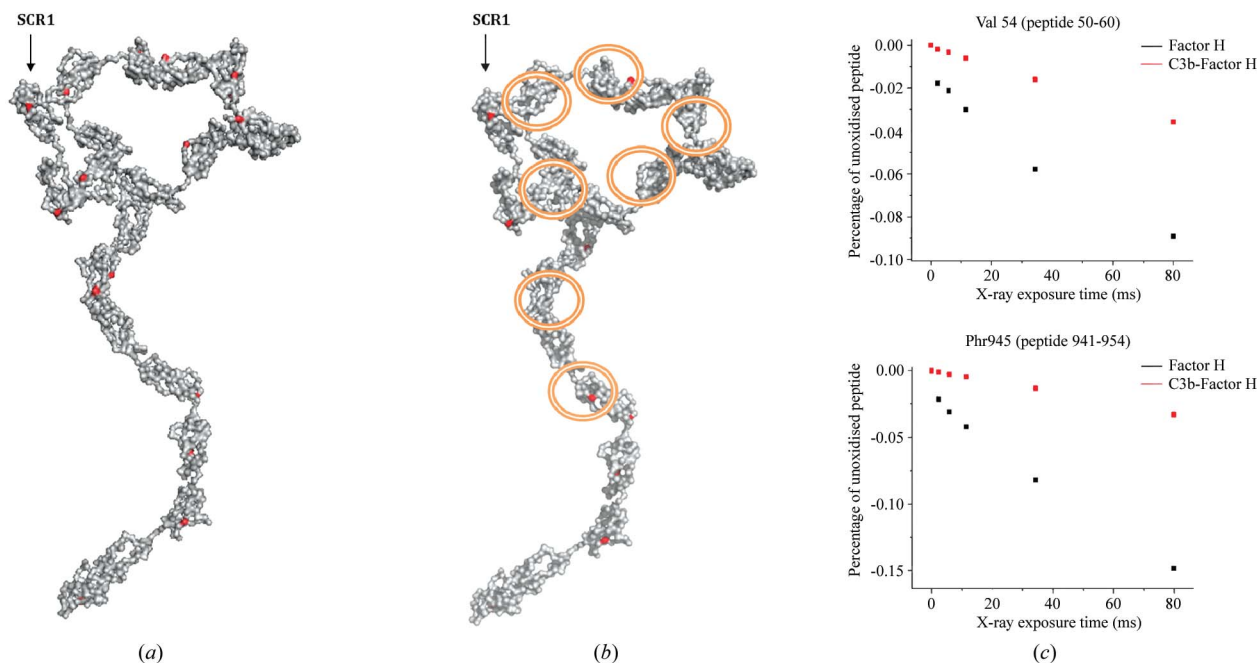

Figure 4
Reproducibility of the X-ray irradiation over time evaluated by the Alexa 488 fluorescence decay upon X-ray exposure. Fluorescence counts were normalized by the initial fluorescence content.

that holds a key role in human innate immunity since it is involved in opsonization and destruction of pathogens and altered host cells. However, it requires an efficient and permanent down regulation to avoid auto-immune damage of host cells and tissues. Among the few complement regulators, the glycoprotein FH is the major circulating regulator of the alternative pathway of the human complement system (Ferreira *et al.*, 2010). FH exhibits unique structural organization made of 20 homologous globular modules of about 60 residues in length and called short consensus repeats (SCRs) that are connected by peptide linkers of three to eight amino acids (Makou *et al.*, 2013). This organization leads to a molecule that resembles a string of beads (Fig. 5*a*), which due to its flexibility can adopt many different conformations (Perkins *et al.*, 1991; Okemefuna *et al.*, 2009). Its marked flexibility in addition to its high molecular weight (150 kDa) and multiple glycosylation sites preclude X-ray crystallography of the entire FH molecule. Crystal structures of few concatenated modules of FH have been reported, and only low- to medium-resolution data of full-length FH have been reported by X-ray and neutron scattering and analytical ultracentrifugation (DiScipio, 1992; Aslam & Perkins, 2001; Perkins *et al.*, 2002), giving only a partial insight into the structural organization of the whole molecule. As a consequence, the structural organization of the full-length FH and of its non-covalent complex with the complement protein C3b has been poorly characterized to date, which hinders a better understanding of the regulation of the complement system. In this context, XFP combined with high-resolution MS is a powerful approach providing new structural insight into the full-length FH with a single amino-acid resolution.

In order to probe the solution structure of the entire FH molecule, we submitted the full-length FH molecule in solution alone and in interaction with C3b (330 kDa) to X-ray radiation at the Metrology beamline of the SOLEIL

Figure 5
Results of hydroxyl radical protein footprinting of FH. (a) Oxidation sites of factor H irradiated alone. (b) Oxidation sites of FH irradiated in the presence of C3b. The differences in oxidation are shown in orange circles. Factor H is represented as a grey surface. Red spheres indicate oxidized residues. PDB: 3GAV-1. (c) Representative dose-response plots showing the fraction of unmodified FH peptide as a function of X-ray exposure time (ms). [This figure has been adapted from one published in *Biochem. J.* (Baud et al., 2016).]

synchrotron. We then identified and localized the resulting oxidized amino acids (mass shift of +16 Da resulting from the incorporation of oxygen) by nanoLC coupled to high-resolution mass spectrometry. This identification of oxidized side-chains provided clues on the solvent-accessible zones on the whole FH molecule, particularly in several central modules of FH whose structure is not yet known. Moreover, assuming solvent-accessibility modification of the FH surface upon interaction with C3b, the accurate measurement of the oxidation pattern was used to construct a map of interaction sites on FH.

It is important to note that the integrity of FH was preserved after X-ray exposure as confirmed by MALDI-TOF analysis (Fig. S2 in the supporting information). The average 73% of coverage of the FH sequence was obtained throughout the entire molecule for the free and C3b-bound FH, allowing for confident analysis of vast portions of the FH. Synchrotron X-ray radiation resulted in 23 FH residues modified by hydroxyl radicals and dispatched in 17 SCRs (shown in Fig. 5a). Among them, six oxidized residues were detected in central modules (SCRs 10–14), indicating solvent accessibility of this central region of FH for which no crystallographic structure is available to date. On the other hand, only eight residues among the 23 residues detected in free FH were found oxidized in the presence of C3b (Fig. 5b) and only one residue was newly oxidized, providing information on interaction zones with C3b as well as on conformational changes of FH upon interaction. One of the most important variations was observed in the N-terminal domain which comprised eight oxidized residues in SCRs (1–3, 5) in FH irradiated alone and

only one in the presence of C3b. Only one oxidized residue in SCR 7 (compared with two in unbound FH) was found in the central part of FH in the C3b–FH complex.

In addition to the identification of the oxidized residues, the extent of oxidation according to the radiation dose has been measured leading to characteristic dose-response plots from which oxidation rates were determined. They showed a significant decrease of oxidation rate for the eight oxidized peptides found in the protein in complex with C3b (representative dose-response plots for peptides 50–60 and 941–954 are shown in Fig. 5c). This result indicates a decreased accessibility of these residues in the presence of C3b, and is in line with the lower footprinting of FH in the presence of C3b as mentioned above. Moreover, the detection of oxidized residues in all technical replicates and the linearity of dose-response plots ensure that the protections were specific to the C3b–FH complex. Together, these results suggest that C3b shields these residues and that FH potentially undergoes a conformational change upon binding with C3b (Baud et al., 2016).

3.2.2. Solvent accessibility of S3 oleosin, a plant protein anchored in yeast lipid droplets. S3 oleosin is the major lipid droplet-bound protein found in *Arabidopsis thaliana*. Oleosins have a unique tri-domain structure involving two amphiphilic regions and a hydrophobic core containing a proline knot, the latter being the longest hydrophobic protein segment known to date. Oleosins are inserted into the phospholipid monolayer surrounding the neutral lipid core of lipid droplets. They are co-purified by a drastic purification protocol including use of detergents and urea to remove non-specifically associated

proteins (Jolivet *et al.*, 2004; Vermachova *et al.*, 2011). Due to their low solubility in aqueous solution, their high-resolution structure is unknown and literature data concerning their secondary structure are contradictory (either mainly α or β folded). We have previously implemented an expression system of S3 oleosin in yeast, addressing S3 oleosin to the lipid droplets of *Saccharomyces cerevisiae* (Vindigni *et al.*, 2013). Based on this expression system and using synchrotron radiation circular dichroism spectroscopy, we have previously proposed a secondary structure model for S3 oleosin in which the central hydrophobic region adopts an original β fold among eukaryotic proteins (Gohon *et al.*, 2011; Vindigni *et al.*, 2013). This central domain contains a proline knot motif in its centre, which was proposed to be responsible for the targeting of oleosin to the lipid droplets (Abell *et al.*, 2004). Two hypotheses about the location of this motif in the lipid monolayer can be raised: it is either mainly embedded in the lipidic core of droplets, or it is in contact with the phospholipid monolayer and their polar head groups, thus probably accessible to water. XFP combined with the yeast expression system offers a unique approach to tackling this question, by purifying an organelle made up of lipid droplet harboring S3 oleosin, and performing the •OH footprinting of the lipid-associated protein S3 to determine which domains are in contact with water molecules and which are found at the interface with lipids.

The densitometry measurements on SDS-PAGE of S3 oleosin indicate that the protein was not degraded by X-ray irradiation, even after 100 ms exposure time (Fig. S1 and Table S1 in the supporting information). The average intensity of the band corresponding to the irradiated oleosin represents 104% (±19%) of that of non-irradiated samples. SDS-PAGE analysis of the wash fractions collected after each irradiation revealed that less than 10% was contaminated with the protein. In-gel digestion by trypsin and chymotrypsin and LC-MS/MS analysis led to an overall protein coverage of 85% (Fig. 6).

It is worth noting that chymotrypsin cleavage allows the coverage of the central hydrophobic domain of S3, which does not contain Arg or Lys residues. Using this approach, the S3 oleosin sequence coverage increases from 47% with trypsin to 72% with chymotrypsin leading to an overall sequence coverage of 85%. Ten amino-acid residues were identified as oxidized upon X-ray irradiation. These modifications mainly concern Met residues (Met19, Met20 and Met29) for which we considered only the over-oxidation (+32 Da mass shift) and His residues (His8, His9, His123, His160) (+16 Da mass shift).

The Met residues concerned were never found non-oxidized, therefore only over-oxidation was calculated. In addition to these modifications, Tyr118, Pro124 and Arg162 were also oxidized. Dose-response plots following oxidation of S3 oleosin were constructed for the peptides containing modified amino acids (Table 1 and Fig. 7).

Different oxidation rate constants were observed depending on the peptide considered (Table 1). Without any available structural data, it is challenging to link the oxidation rate of some peptides to their solvent accessibility and further to shielding effects. For example, the behaviour of Met residues must be pointed out. Met19 and Met20 (peptide 2; initial rates 0.0044 ms^{-1}) were over-oxidized more quickly than Met29 (peptide 3; 0.0006 ms^{-1}) when Met135 (peptide 6; 0.0001 ms^{-1}) was never found over-oxidized. The peptide YATGEHPQGSDKLDSAR (peptide 4) which possesses three oxidation sites (Tyr118, His123 and Pro124) was oxidized faster than peptides 5 and 6 containing only His123 and Pro124 residues. Thus, Tyr118 appeared more susceptible to oxidation.

Seventy-one percent of the hydrophobic segment sequence (Ala45 to Lys117) of S3 oleosin was covered after chymotrypsin digestion and analysis of peptides by LC-MS/MS (Fig. 6), but no oxidation was found in the identified peptides. In particular, two of the three Pro residues from the proline knot (Pro82 and Pro86) were found unmodified. The peptide harbouring Pro75 was not detected. On the other hand, Pro124 located close to the hydrophobic segment was detected as oxidized, confirming that our XFP experiment allows for observation of oxidized proline residue.

Taken together, these results show that all the amino acids of the protein sequence have varying susceptibility for modification by hydroxyl radicals, as reported previously (Xu *et al.*, 2003; Xu & Chance, 2004). Thus, we first confirm that S3 oleosin is folded in yeast lipid droplets. Secondly, the central region is likely to be inserted within the hydrophobic core of the lipid droplet. Tyr118, located just outside the C-terminal end of the hydrophobic domain, is highly oxidized, whereas Tyr116 in the hydrophobic region remains unmodified upon X-ray radiation. N- and C-terminal regions of oleosin contain most of the oxidized residues, confirming their cytosolic orientation. Differences in the oxidation rate of cytosolic peptides may reflect
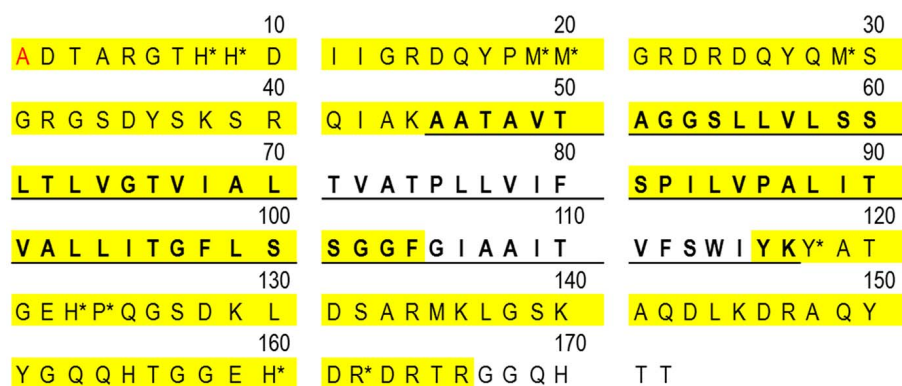

Figure 6
Amino-acid sequence of S3 oleosin expressed in *S. cerevisiae* oil bodies. The hydrophobic segment is shown in bold face and underlined. Amino acids identified by mass spectrometry are highlighted in yellow. X-ray modifications identified by MS/MS analysis are indicated by an asterisk (*). N-ter Ala is acetylated.

Table 1
Oxidized peptides from X-ray irradiated S3 oleosin.

The oxidation rate (expressed in ms^{-1}) was determined from the dose-response plots of the proportion of unmodified peptide as a function of X-ray exposure time (ms) fit to an exponential [equation (1)]. When the same modified amino acids were identified in several peptides, all of the data were averaged. Oxidized peptides were obtained by in-gel digestion by trypsin and chymotrypsin and identified by LC-MS/MS.

| Peptide number	Peptide sequence	Identified modifications	•OH modified amino acids	Initial rate (ms^{-1})	R^2
1	ADTARGTHHDIIGRDQYPM	H8: +15.99491; H9: +15.99491; M19: +15.99491; A1: +42.01057	H8–H9	−0.0005	0.7206
2	DQYPMMGR	M19: +31.9898; M20: +31.9898	M19–M20	−0.0044	0.9663
3	QMSGRGSDY	M29: +31.9898; Q28: −17.02655	M29	−0.0006	0.8549
4	YATGEHPQGSDKLDSAR	H123: +15.99491; P124: +15.99491; Y118: +15.99491	Y118–H123–P124	−0.0028	0.6575
5	ATGEHPQGSDKL	H123: +15.99491; P124: +15.99491	H123–P124	−0.0001	0.7096
6	ATGEHPQGSDKLDSARM	H123: +15.99491; M135: +15.99491; P124: +15.99491	H123–P124		
7	AQYYGQQHTGGEHDRDR	H160: +15.99491; R162: +15.99491	H160–R162	−0.0013	0.945
8	DRAQYYGQQHTGGEHDRDR	H160: +15.99491, R162: +15.99491	H160–R162		
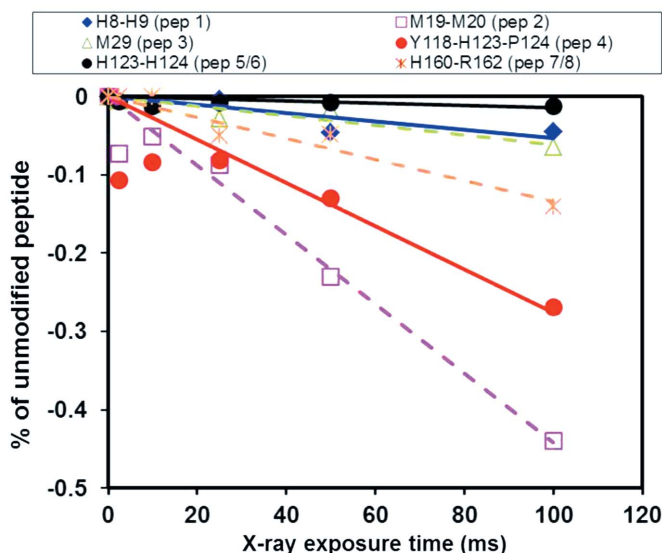
Figure 7
Representative dose-response plots showing the fraction of unmodified S3 oleosin peptides as a function of X-ray exposure time (ms). Peptides were obtained by the trypsin or chymotrypsin digestion of X-ray irradiated S3 oleosin and identified with modifications (see Table 1).

either contact with the surface of the phospholipid monolayer or protein–protein interactions. Further XFP experiments with lipid droplets from various plant seeds and a comparison of the obtained oxidation pattern of oleosins will provide new insights into oleosin insertion and location of regions involved in the protein–protein interactions.

4. Conclusions

To date, the majority of synchrotron X-ray footprinting experiments were conducted at the National Synchrotron Light Source, Brookhaven, USA, and at ESRF, Grenoble, France. Moreover, XFP has been used to study medium-size protein complexes such as rhodopsin-G-protein (Orban *et al.*, 2012), and, to our knowledge, only a few studies on complexes of over 300 kDa have been reported (Deperalta *et al.*, 2013).

Here we describe the successful employment of the Metrology beamline at the SOLEIL synchrotron facility to study the structure of the complement FH–C3b complex of molecular weight over 300 kDa and S3 oleosin protein bound to lipid droplets. For this purpose, the Metrology beamline was equipped with an experimental set-up devoted to the analysis of biomolecules by XFP.

The identification of oxidized side-chains has given insight into the solvent accessibility of FH residues, particularly important with regard to several SCR modules (*e.g.* SCR 5 and SCR 14) whose structure is not yet known. The oxidation pattern was used to construct a map of interaction sites on FH and highlighted a possible new role of the central part of FH in folding of FH upon formation of the C3b–FH complex. We also found that several SCR modules may exhibit slightly different structural arrangements in the full-size FH molecule, compared with previously concatenated modules in RX crystals. This underlines that the *in vivo*-like solution structure determination does not suffer from artificial constraints such as those that may be caused by crystal packing.

Furthermore, here we show that XFP also offers a unique approach to studying the structure of proteins embedded within membranes or lipid bodies. Here we show the results obtained for the oxidation of the S3 oleosin in its native-like environment, *i.e.* in an organelle. The obtained results confirm that the central hydrophobic domain and the proline knot are not exposed to water, and provide valuable insight into their location within the lipid layer.

Besides the protein–protein and protein–lipid complexes studied here, the application of this experimental set-up to the study of the kinetics of DNA–protein interactions should be straightforward, since the same kind of set-up was used at ESRF for this purpose. In the longer term we envisage applying this technique to the irradiation of live cells in order to probe the different macromolecular conformations present.

Finally, the reported experimental set-up implemented on the Metrology beamline at SOLEIL has proven to be successful for the footprinting of biomolecules and will be further developed to improve automation of sample manipulation and enhance the user experience.

References

Abell, B. M., Hahn, M., Holbrook, L. A. & Moloney, M. M. (2004). *Plant J.* **37**, 461–470.

Aslam, M. & Perkins, S. J. (2001). *J. Mol. Biol.* **309**, 1117–1138.

Baud, A., Gonnet, F., Salard, I., Le Mignon, M., Giuliani, A., Mercère, P., Sclavi, B. & Daniel, R. (2016). *Biochem. J.* **473**, 1805–1819.

Bohon, J., D'Mello, R., Ralston, C., Gupta, S. & Chance, M. R. (2014). *J. Synchrotron Rad.* **21**, 24–31.

Clatterbuck Soper, S. F., Dator, R. P., Limbach, P. & Woodson, S. A. (2013). *Mol. Cell*, **52**, 506–516.

Deperalta, G., Alvarez, M., Bechtel, C., Dong, K., McDonald, R. & Ling, V. (2013). *mAbs*, **5**, 86–101.

DiScipio, R. G. (1992). *J. Immunol.* **149**, 2592–2599.

Ferreira, V. P., Pangburn, M. K. & Cortés, C. (2010). *Mol. Immunol.* **47**, 2187–2197.

Fraczkiewicz, R. & Braun, W. (1998). *J. Comput. Chem.* **19**, 319–333.

Gohon, Y., Vindigni, J. D., Pallier, A., Wien, F., Celia, H., Giuliani, A., Tribet, C., Chardot, T. & Briozzo, P. (2011). *Biochim. Biophys. Acta*, **1808**, 706–716.

Gupta, S., Celestre, R., Petzold, C. J., Chance, M. R. & Ralston, C. (2014). *J. Synchrotron Rad.* **21**, 690–699.

Gupta, S., Feng, J., Chan, L. J. G., Petzold, C. J. & Ralston, C. Y. (2016). *J. Synchrotron Rad.* **23**, 1056–1069.

Gupta, S., Guttman, M., Leverenz, R. L., Zhumadilova, K., Pawlowski, E. G., Petzold, C. J., Lee, K. K., Ralston, C. Y. & Kerfeld, C. A. (2015). *Proc. Natl Acad. Sci. USA*, **112**, E5567–E5574.

Gupta, S., Sullivan, M., Toomey, J., Kiselar, J. & Chance, M. R. (2007). *J. Synchrotron Rad.* **14**, 233–243.

Hambly, D. M. & Gross, M. L. (2005). *J. Am. Soc. Mass Spectrom.* **16**, 2057–2063.

Harrison, S. C. (2004). *Nat. Struct. Mol. Biol.* **11**, 12–15.

Hulscher, R. M., Bohon, J., Rappé, M. C., Gupta, S., D'Mello, R., Sullivan, M., Ralston, C. Y., Chance, M. R. & Woodson, S. A. (2016). *Methods*, **103**, 49–56.

Jolivet, P., Roux, E., d'Andrea, S., Davanture, M., Negroni, L., Zivy, M. & Chardot, T. (2004). *Plant Physiol. Biochem.* **42**, 501–509.

Kamal, J. K. A., Benchaar, S., Takamoto, K., Reisler, E. & Chance, M. R. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 7910–7915.

Kaur, P., Kiselar, J. G. & Chance, M. R. (2009). *Anal. Chem.* **81**, 8141–8149.

Kiselar, J. G. & Chance, M. R. (2010). *J. Mass Spectrom.* **45**, 1373–1382.

Kiselar, J. G., Janmey, P., Almo, S. C. & Chance, M. R. (2003). *Mol. Cell. Proteomics*, **2**, 1120–1132.

Kiselar, J. G., Maleknia, S. D., Sullivan, M., Downard, K. M. & Chance, M. R. (2002). *Int. J. Radiat. Biol.* **78**, 101–114.

Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.

Makou, E., Herbert, A. P. & Barlow, P. N. (2013). *Biochemistry*, **52**, 3949–3962.

Maleknia, S. D. & Downard, K. M. (2014). *Chem. Soc. Rev.* **43**, 3244–3258.

Ménesguen, Y. & Lépy, M.-C. (2011). *X-ray Spectrum.* **40**, 411–416.

Moreno, T. & Idir, M. (2001). *J. Phys. IV*, **11**, Pr2-527–Pr2-531.

Neuhoff, V., Arold, N., Taube, D. & Ehrhardt, W. (1988). *Electrophoresis*, **9**, 255–262.

Okemefuna, A. I., Nan, R., Gor, J. & Perkins, S. J. (2009). *J. Mol. Biol.* **391**, 98–118.

Orban, T., Jastrzebska, B., Gupta, S., Wang, B., Miyagi, M., Chance, M. R. & Palczewski, K. (2012). *Structure*, **20**, 826–840.

Padayatti, P. S., Wang, L., Gupta, S., Orban, T., Sun, W., Salom, D., Jordan, S. R., Palczewski, K. & Chance, M. R. (2013). *Mol. Cell. Proteomics*, **12**, 1259–1271.

Perkins, S. J., Gilbert, H. E., Aslam, M., Hannan, J., Holers, V. M. & Goodship, T. H. (2002). *Biochem. Soc. Trans.* **30**, 996–1001.

Perkins, S. J., Nan, R., Li, K., Khan, S. & Miller, A. (2012). *Immunobiology*, **217**, 281–297.

Perkins, S. J., Nealis, A. S. & Sim, R. B. (1991). *Biochemistry*, **30**, 2847–2857.

Ralston, C. Y., Sclavi, B., Sullivan, M., Deras, M. L., Woodson, S. A., Chance, M. R. & Brenowitz, M. (2000). *Methods Enzymol.* **317**, 353–368.

Robinson, C. V., Sali, A. & Baumeister, W. (2007). *Nature (London)*, **450**, 973–982.

Rogozina, A., Zaychikov, E., Buckle, M., Heumann, H. & Sclavi, B. (2009). *Nucleic Acids Res.* **37**, 5390–5404.

Sclavi, B., Woodson, S., Sullivan, M., Chance, M. R. & Brenowitz, M. (1997). *J. Mol. Biol.* **266**, 144–159.

Sclavi, B., Zaychikov, E., Rogozina, A., Walther, F., Buckle, M. & Heumann, H. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 4706–4711.

Valot, B., Langella, O., Nano, E. & Zivy, M. (2011). *Proteomics*, **11**, 3572–3577.

Vermachova, M., Purkrtova, Z., Santrucek, J., Jolivet, P., Chardot, T. & Kodicek, M. (2011). *Proteomics*, **11**, 3430–3434.

Vindigni, J. D., Wien, F., Giuliani, A., Erpapazoglou, Z., Tache, R., Jagic, F., Chardot, T., Gohon, Y. & Froissard, M. (2013). *Biochim. Biophys. Acta*, **1828**, 1881–1888.

Xu, G. & Chance, M. R. (2004). *Anal. Chem.* **76**, 1213–1221.

Xu, G. & Chance, M. R. (2007). *Chem. Rev.* **107**, 3514–3543.

Xu, G., Takamoto, K. & Chance, M. R. (2003). *Anal. Chem.* **75**, 6995–7007.

Zhu, Y., Guo, T., Park, J. E., Li, X., Meng, W., Datta, A., Bern, M., Lim, S. K. & Sze, S. K. (2009). *Mol. Cell. Proteomics*, **8**, 1999–2010.