Analytical Methods

Enabling simultaneous photoluminescence spectroscopy and X-ray Footprinting Mass Spectrometry to study protein conformation and interactions

Journal:	*Analytical Methods*
Manuscript ID	AY-ART-09-2024-001670.R2
Article Type:	Paper
Date Submitted by the Author:	18-Dec-2024
Complete List of Authors:	Gupta, Sayan; Lawrence Berkeley Laboratory, Russell, Brandon; Lawrence Berkeley National Laboratory Kristensen, Line G.; Lawrence Berkeley National Laboratory Costello, Shawn M.; University of California at Berkeley, Department of Molecular and Cell Biology Tyler, James; Lawrence Berkeley National Laboratory Ralston, Corie; Lawrence Berkeley National Laboratory Rad, Behzad ; E O Lawrence Berkeley National Laboratory, Materials Sciences Division Marqusee, Susan; University of California Berkeley

Enabling simultaneous photoluminescence spectroscopy and X-ray Footprinting Mass Spectrometry to study protein conformation and interactions

Sayan Gupta[1], Brandon Russell[1], Line G. Kristensen[1], Shawn M. Costello[2], Behzad Rad[3], James Tyler[3], Susan Marqusee[2], Corie Y. Ralston[3*]

[1]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

[2]Department of Molecular and Cell Biology, University of California at Berkeley, Berkeley California 94720, United States

[3]Molecular Foundry Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

Correspondence email: cyralston@lbl.gov

Abstract

X-ray Footprinting Mass Spectrometry (XFMS) is a structural biology method that uses broadband X-rays for *in situ* hydroxyl radical labeling to map protein interactions and conformation in solution. However, while XFMS alone provides important structural information on biomolecules, as we move into the era of the interactome, hybrid methods are becoming increasingly necessary to gain a comprehensive understanding of protein complexes and interactions. Toward this end, we report the development of the first synergetic application of inline and real-time fluorescent spectroscopy at the Advanced Light Source's XFMS facility to study local protein interactions and global conformational changes simultaneously. To facilitate general use, we designed a flexible and optimum system for producing high-quality spectroscopy-XFMS hybrid data, with rapid interchangeable liquid jet or capillary sample delivery for multimodal inline spectroscopy, and several choices for optofluidic environments. To validate the hybrid system, we used the

covalently interacting SpyCatcher-SpyTag split protein system. We show that our hybrid system can be used to detect the interaction of SpyTag and SpyCatcher via FRET, while elucidating key structural features throughout the complex at the residue level via XFMS. Our results highlight the usefulness of hybrid method in providing binding and structural details to precisely engineer protein interactions.

Introduction

Biomolecular spectroscopy, one of the most well-established methods, remains one of the most rapid, sensitive, and widely used approaches for protein characterization in terms of analysis of chemical bond types as well as overall three-dimensional structure[1]. In comparison, X-ray footprinting mass spectrometry (XFMS) is a specialized approach that provides structural information on biological molecules in the form of solvent accessibility maps at the residue level, as well as location and dynamics of bound waters[2]. XFMS uses a dilute protein solution in a buffer, and the method is not limited by specialized sample preparation requirements and the protein molecular weight[2, 3]. The method is often used in conjunction with other methods such as crystallography[4-7] and cryo-EM[7-9], with the advantage of solution state, or with SAXS[10-12] and HDX-MS[11, 13, 14] to give complementary structural information. XFMS was developed in the late nineties based on the well-established chemistry of water radiolysis and has evolved substantially in the last few years at synchrotron X-ray facilities[3, 15, 16]. Other non-radiolytic •OH labeling methods such as Fenton chemistry[17-20] or UV-laser-induced •OH labeling (known as fast photochemical oxidation of proteins, FPOP)[21-25] provide solvent accessibility information similar to XFMS, and are excellent methods in the absence of a synchrotron facility. All these footprinting methods are similar in that they rely on the production of hydroxyl radical, but they differ in implementation. For instance, non-radiolytic methods require the addition of H_2O_2 while for XFMS, no additional reagents are required. Recently, XFMS has achieved a new level of routine use through the implementation of an automated Alexa dosimetry[26] together with a high-dose

containerless approach using liquid jet microfluidics[3, 27]. Recently, integrated inline OH radical dosimetry has also been introduced to a UV-laser-induced •OH labeling method[28, 29]. However, important concerns, such as characterizing and validating the state of a protein or protein structural integrity at the time of the XFMS experiment, remain unaddressed, which can result in uninterpretable data and loss of user time.

Further, researchers can face hurdles when carrying out XFMS due to the limited availability and time for orthogonal spectroscopic characterization of their samples at the time of data collection at X-ray facility. The synergetic use of spectroscopic tools such as UV-visible absorption, X-ray absorption spectroscopy, fluorescence, and Raman spectroscopy has previously been established at several synchrotron beamlines, mainly for simultaneous use with crystallography, and has highlighted the importance of validation of the state of the protein in the crystal structure[30-32]. From the beamline-development perspective, the recent advances in optical engineering and data acquisition technologies using microfluidics[33-35], combined with the containerless approach using a liquid jet[3], provide an excellent platform to implement simultaneous photoluminescence spectroscopy with XFMS, and have several technical advantages. First, when using the containerless approach, there is no glass or other material enclosing the sample. This results in the increased sensitivity of probing the sample using a variety of light sources, high-resolution optics, and an efficient data acquisition system. Second, the very high-speed flow velocity of the jet enables microsecond mixing and simultaneous time-resolved spectroscopy and XFMS. Third, for photoexcitable proteins such as the orange carotenoid protein[4, 11], rhodopsin[36], and photosystem II[37], a pump and probe experiment is feasible to monitor photoactivated and transient intermediates by both XFMS and spectroscopy at the same time.

From the spectroscopy perspective, fluorescence can provide information on changes in size, shape, flexibility, and conformation, as well as knowledge about the proximity of binding partners[38-42], whereas Raman can fingerprint the changes associated with secondary structure and protein-pigment interactions[43-48]. These photoluminescence spectroscopy methods are non-

invasive, highly sensitive, and adaptable to the microfluidic configuration of XFMS and, therefore, most appropriate to complement solvent accessibility data. For these reasons, we constructed an automated inline fluorescence data collection platform coupled with high-dose XFMS using microfluidics as a first step toward the implementation of this hybrid data collection approach. The fluorescence emission from the sample was collected inline immediately prior to sample exposure, either inside the micro-capillary or from the liquid jet.

We chose to validate the new inline spectroscopy capabilities using the known split-protein system SpyCatcher-SpyTag, which is a designed protein peptide pair that can covalently bind to each other and form a stable complex by forming an isopeptide bond[49] (**Figure 1**). This powerful system in molecular biology and biotechnology is used primarily for protein labeling, immobilization, and other bioengineering applications. The SpyCatcher-SpyTag system was originally engineered by splitting the CnaB2 domain of fibronectin-binding protein from *S. pyogenes* into a 13 residue SpyTag and a 116 residue SpyCatcher[49]. This system was termed SpyCatcher001-SpyTag001, which is the only crystallographically characterized split protein system among all other engineered variants[50]. A phage display platform and selective mutation were used to develop a tighter affinity for intermolecular isopeptide bonds between these two polypeptides, resulting in the SpyCatcher002-SpyTag002 and SpyCatcher003-SpyTag003 systems, with progressively increasing speed of isopeptide bond formation as shown by stopped-flow kinetics[49]. Although hydrogen-deuterium exchange mass spectrometry (HDX) compared the key features of the backbone stability that could explain the rapid binding to SpyTag, the extent of local solvent accessibility change that drives rapid bond formation was not studied. As such, this covalently interacting split-protein system was ideal for testing our new instrument because we were able to fluorescently label SpyCatcher and SpyTag and measure inline fluorescence energy transfer (FRET[36]) to verify complete complexation while also collecting standard XFMS data. We produced the SpyCatcher003-SpyTag001 cross variant, which provided some new structural information, while also serving as an excellent model system for verification of our

hybrid XFMS structural data. The hybrid mode of data collection validated the complex formation by its FRET response using inline fluorescence measurements, and simultaneously, XFMS identified the site-specific solvent accessibility changes at the binding site, which provided residue specific information in the high-affinity complex formation. The new instrumentation for hybrid data collection described here to integrate local changes in solvent accessibility with global and local information from spectroscopy is now available more generally for protein systems under study at the XFMS facility.

Experimental Section

Sample preparation

SpyCatcher003 S49C was purified as described previously[49] (Supporting methods), and labeled with Alexa Fluor 555 C2 maleimide (ThermoFisher Scientific). SpyCatcher003 S49C in a solution of 50 mM Tris pH 8.0, 300 mM NaCl, 5% (v/v) glycerol was reacted with 1.5-fold excess dye at 4°C for 4 hours. The reaction was then run over a BioGel P-6 column (Bio-Rad Laboratories) equilibrated in 1X PBS. The labeling efficiency was determined by UV/Vis spectrometry using the dye concentration calculated at maximum absorbance (ε_{556nm} = 158,000 M-1cm-1) and protein concentration as calculated as described in the supporting method, taking into account the dye's correction factor at 280 nm is 0.08 for Alexa Fluor 555. The degree of labeling was determined by dividing the dye concentration by the protein concentration and found to be 1.3. SpyCatcher003 was buffer exchanged to 10 mM sodium phosphate for XFMS experiments. The 13 amino acid long SpyTag001 (AHIVMVDAYKPTK) with the N-term labeled with HiLyte T Fluor 488 was purchased from Anaspec Inc. SpyTag001 lyophilized powder was dissolved in water and its concentration was determined from the molar extinction coefficient of HiLyte T Fluor 488 (73000 $cm^{-1}M^{-1}$ at 490 nm). Equimolar amounts of SpyCatcher003 and SpyTag001 in 10 mM sodium phosphate buffer were mixed at room temperature for 30 min to form the SpyCatcher003-Spytag001 complex.

Sample exposure and fluorescence data collection

Radiolysis of samples was performed at the Advanced Light Source (ALS) beamline 3.3.1, which delivers a 3 – 12 keV broadband X-ray beam from a bending magnet source with a high flux density focused beam. Because the focused beam consists of a bright spot and a lower flux density tail, we used an X-ray slit to define two regions[51]. The first region defined a high flux density beam spot size of 240 μm (V, vertical) x 80 μm (H, horizontal) for use with the liquid jet, and the second region defined a spot size of 640 μm (V) x 200 μm (H) with ~ 40 times lower flux density for use with capillary flow sample exposures. The dimensions of the slit-defined X-ray beams were chosen to maintain sample velocity within the stable jetting regime, ensuring adequate flux density and sufficient modification by hydroxyl radicals when employing a 75-100 μm jet and a 200 μm ID capillary[3]. The beam current, which is a measure of flux density, varies by a negligible amount (0.1%) during an experiment, and the focused X-ray beam used for sample exposure had a fixed dose rate[51]. The samples flow through the X-ray beam at different speeds to achieve exposure ranging from 10 – 40 μs and 250 – 1250 μs for the jet and the capillary, respectively (**Table S1**). The sample volume within the path of the beam at any given moment (1-20 nl) is negligible compared to the total sample volume exposed (50-200 μl). Therefore, for a given dose, each collected sample is effectively an average of many smaller samples. The progressive increase in the exposure time under a fixed-dose rate (Gy s-1) provides a series of samples which serve as replicates and are used to plot a dose response.

Inline fluorescence data collection and X-ray exposure of Alexa samples were carried out with both the jet and capillary sample delivery systems for comparison. Inline fluorescence data collection and X-ray exposure of protein samples, free SpyCatcher003 and the SpyCatcher003-Spytag001 complex were carried out with capillary sample delivery systems. Jet and capillary

positioning with respect to the X-ray beam and microscopes were achieved with motorized opto-mechanical components from ThorLabs similar to those described previously[27].

The sample fluorescence probe or detection area was configured vertically immediately above the point of X-ray exposure. The approximate detection area is around 0.5 mm^2 with 20 X objective magnification (~ 10mm light gathering area divided by the objective magnification), which gives a fluorescence detection volume ranging from approximately 5 to 50 nl for jet and capillary, respectively. The fluorescence detection volume is small compared to the total sample volume in the path of the beam at any given moment. Therefore, a 50-200 ms integration time of the CCD spectrometer, which has a scan frequency of 200 Hz, is required to give enough signal averaging for the flowing samples. Higher flow speeds are used for shorter exposures, but the spectroscopy data was independent of the flow speed of the sample.

LabVIEW-based software controlled the flow speed, exposure time, sample collection volumes, and spectroscopy data collection and storage. Pre-exposure fluorescence spectra collected for each X-ray exposure can be combined manually for further signal averaging. Fluorescence excitation and collection were carried out using inline cameras, lenses, and beamsplitters as described previously[27]. The Thorlabs compact CCD Spectrometer was integrated into the main XFMS LabVIEW virtual instrument (Supporting methods).

Mass spectrometry and data analysis

Exposed samples were collected in tubes containing methionine amide to immediately scavenge any secondary radical reactions[52, 53]. The SpyCatcher003 protein was digested with trypsin and AspN to obtain maximum sequence coverage. The resulting peptide fragments were analyzed by standard bottom-up LCMS data collection as previously described using Protein Metrics by Dotmatics[54]. Briefly, the fraction unmodified for each peptide was calculated as the ratio of the integrated peak area of the unmodified peptide to the sum of integrated peak areas from the modified and unmodified peptides. The dose response curves (fraction unmodified vs. X-ray exposure) were fitted to single exponential functions in Origin® (OriginLabs). The rate constant,

k (sec^{-1}), was used to measure the reactivity and solvent accessibility of sidechains towards hydroxyl radical-induced modification. The reported errors of the rate data were determined by the Origin program using 95% confidence limits of the fitting results. The R^2 of the fits were between 0.98-0.99. The ratio of rate constants provided the relative change in the solvent accessibility between the free protein and the complex[11].

Results and discussion

Design and operation of hybrid spectroscopy-XFMS data collection

The basic idea of the design includes the simultaneous capability of collecting various spectroscopy data from anywhere in the sample path (pre and post-exposure) and carrying out automated post-exposure Alexa dosimetry during the sample's X-ray exposure. The design to collect spectroscopy data from a flowing sample both at a position above the X-ray impingement point ("pre-exposure" position), and Alexa Fluor 488 fluorescence at a position below the X-ray impingement point ("post-exposure" position), was achieved by installing two laser-induced fluorescence (LIF) modules for microfluidics in the confined space around the sample jet or capillary (**Figure 2**). The integrated instrument's layout with two fluorescence imaging modules (FIMs) is shown in **Figures 3A** and **3B**. Since the jet produces stable sample flow for more than 1 cm (**Figure S1C**)[3], in the new design, accommodating two FIMs was not spatially challenging compared to our previous instrument for automated Alexa dose response analysis[27]. The pre-exposure and post-exposure FIMs contained necessary and similar optical elements to collect fluorescence emission spectra using a CCD spectrometer or PMT-based fluorescence intensity analysis[27], respectively. The configurations, functions, and software control of other components, such as the sample handling module (liquid jet or capillary), laser-assisted pre-alignment unit (LAPU), and beam alignment module (BAM), are similar but more efficient than that of the previous design as reported earlier (**Figure S1**). Signals and images from the three modules (FIMS and BAM) are viewed in a LabVIEW-based GUI, which controls the individual components, collects spectroscopy data, quantifies fluorescence emission, calculates the dose rates, and

provides overall optomechanical automation of the XFMS sample exposure experiments (**Figure S2A)**. Overall, the design provides the flexibility for probing various types of spectral characteristics, using different light sources and optical filter systems, inline with the sample flow path both above and below the X-ray exposure position. The objective lens's numerical aperture and magnification can be selected to achieve high signal collection efficiency from various photoluminescence probes, which is particularly beneficial when carrying out Raman spectroscopy. The spectral sensitivity and resolution can be increased significantly using high-end spectrometers and electron-multiplying CCD detectors. The instrument uses precision three-axis motor-controlled stages to align the spectroscopic microprobe objective lens with the liquid jet or capillary and uses wide-area camera components to observe and track the laser beam focus and location of the X-ray beam onto the samples using the LAPU and BAM (**Figure S1**).

Capturing Alexa fluorescence emission spectra from the liquid jet and capillary

In the XFMS method, Alexa 488 is used as a dosimeter molecule to determine the presence of hydroxyl radical scavenger and optimize hydroxyl radical dosage conditions for protein samples[26]. The assay, which is now fully automated[27], includes monitoring of a decrease in the Alexa fluorescence intensity in protein samples as a function of increasing exposure time and determining a rate constant of hydroxyl radical reactivity of Alexa 488, the value of which is a measure of the goodness of exposure conditions to get adequate protein modification for quantitative mass spectrometry analysis[55]. We leveraged Alexa 488 fluorescence to optimize pre-exposure spectral data collection in conjunction with actual sample exposure conditions regarding sample volume, flow speed, and exposure time range for both capillary and liquid jet sample configuration[3].

The system configuration is such that the spectroscopic integration time is directly dependent on the pump time of sample ejection into the X-ray path for radiolysis. The pump time is controlled by the flow rate and the total volume of sample to be ejected. Since the flow rate is

directly proportional to the exposure time, a short exposure time and a small total volume of sample require a short pump time[3]. The current data acquisition capability limits the pump time value to approximately ~ 400 ms (**Figure S2B**). However, for robust operation, we routinely use a pump time ranging between 600 ms to 3 sec for variable dose-based XFMS experiments (**Table S1**). The longer pump time provides enough signal averaging to collect high signal-to-noise spectroscopic data in parallel with XFMS sample exposure. To test the system performance, we collected Alexa emission data from glass and quartz capillaries and liquid jets before sample exposure (**Figure 4A**). We used a concentration of Alexa at 2.5 uM, which is the standard for Alexa assays for XFMS; however, we only collected 10% of the emitted light using a 90-10 beamsplitter (**Figure 2**). 90% of the light was used to capture images of the sample using the camera. Reasonable quality Alexa emission spectra were achieved using both 50 and 200 ms integration times (**Figure 4B**). We observed only a modest two-fold increase in the signal obtained from the jet sample relative to a 75 μm ID capillary sample in which the fluorescence was emitted through a ~ 143 μm glass wall (**Figure 4A**). The decrease in the emission intensities is due to the transmission properties of light through ordinary glass at the observed wavelength of Alexa emission. In contrast, there was no increase in the jet sample signal relative to a 100 μm ID quartz capillary in which the fluorescence was emitted through a ~ 130 μm quartz wall (**Figure 4A**). A slight decrease in the emission intensities at both integration times for the jet might be due to the absence of backscattered excitation light from the cylindrical glass surface as with the capillary flow but did not detract from the quality of the spectra. Therefore, we conclude that the pump times for XFMS are sufficient to collect spectroscopy data using the ordinary CCD-based spectrometer and both glass and quartz capillaries for visible wavelengths.

Validation of global conformation change in SpyCatcher-SpyTag interaction by FRET

Spectroscopy data collected on the SpyCatcher003-Alexa Fluor 555, SpyTag003-HiLyte T Fluor 488 and the fluorescently tagged SpyCatcher003-SpyTag001 complex inline immediately before microsecond X-ray exposure showed characteristic fluorescence emission spectra of the respective fluorescently tagged constructs (**Figure 1** and **5B**). We observed >30% reduction of N-terminal HiLyte T Fluor 488 labeled SpyTag001 upon binding and covalent bond formation with Alexa 555 labeled SpyCatcher003 at position S49C. This data showed the proximity of the donor HiLyte T Fluor 488 and acceptor Alexa 555 in the complex, resulting in fluorescence resonance energy transfer (FRET) and confirming a strong association of the SpyCatcher003 and SpyTag001 before sample exposure for XFMS studies (**Figure 5B**). FRET is a distance-dependent, non-radiative energy transfer process from the excited state of a donor molecule to the ground state of an acceptor molecule by a dipole-dipole coupling interaction within 10 nm proximity[42, 56]. FRET is valuable for studying protein-protein interactions, protein conformation changes in vitro, and information on cellular signaling pathways and molecular interactions in vivo[57-61]. FRET and hydroxyl radical footprinting are complementary techniques that can provide powerful insights into protein structure, dynamics, and interactions in protein complexes. The hydroxyl radical footprinting maps solvent-accessible regions of a protein by identifying solvent-exposed and protected areas based on reactivity with hydroxyl radicals, while FRET measures proximity or distances between intrinsic or labeled donor and acceptor molecules, revealing the spatial arrangement of protein domains or subunits together, they can enable detailed structural and dynamic mapping, offering insights into conformational changes upon binding or allosteric conformational shifts. A covalent modification using bulky extrinsic fluorescence molecules[62] can affect structural and functional changes, which can be further cross-validated by simultaneous FRET and solvent accessibility measurements using this hybrid approach and unlabeled protein as a control.

Our results corroborated earlier studies, which showed a 35% reduction in the intensity by FRET between a larger donor fluorescence molecule mClover3 attached to the C-terminal end of

SpyTag003 and the same SpyCatcher003 tagged with acceptor Alexa555[49]. The higher emission intensity at a lower integration time was possible by increasing the sample path length (200 μm vs. 75 / 100 μm), replacing the beamsplitter with a mirror, and collecting most of the emitted light intensity (**Figure 5B**). As this new spectroscopy system features rapid interchangeable light sources and optical components, it can cover a broad range of excitation and emission wavelengths most useful for biomolecular intrinsic and extrinsic fluorescence and can be extended to Raman spectroscopy. The approach of integration of well-established biomolecular spectroscopy methods into the XFMS data collection workflow provides options for sample monitoring for quality control as well as comprehensive global structural information together with the detailed structural insights into the amino acid side chains of a protein system and ensures that two orthogonal structural data sets are obtained under identical conditions, both in time and in the sample environment and with minimal sample usage.

Radiolytic labeling as a tool for the analysis of local protein-protein interactions

The development of flow-based high-dose microsecond radiolytic labeling[3, 27, 63] gained many advantages over conventional mechanical shutters-based millisecond exposure[26]. Apart from generating high-quality data with low sample damage induced by secondary radical reactions, the microsecond exposure is carried out using micro-capillaries or liquid jets, which enable probing the sample by spectroscopy in close proximity or near simultaneous to that of the radiolytic labeling. Spectroscopy data was collected for the SpyCatcher003 protein and the SpyCatcher003-SpyTag001 complex inline immediately prior to microsecond X-ray exposure for XFMS analysis, with a shortest estimated delay time between spectroscopy and X-ray exposure is 100 μs using the 200 ID capillary. The exposed samples were rapidly quenched in the fraction collector and processed by LCMS, as described in the methods section. Hydroxyl radicals are generated by radiolysis isotropically wherever water is present; thus, when water is in contact with side chain

residues, they become targets for labeling[2]. We obtained near > 90% sequence coverage, and experiments in dose replicates exhibited consistent radiolytic labeling on identical residues as well as the progressive extent of modification with an increase in the exposure time for both free SpyCatcher003 protein and the SpyCatcher003-SpyTag001 complex. In a controlled dose environment, the yield of side-chain modification generally aligns with pseudo-first-order reaction kinetics. Deviations from the hydroxyl radical dose-response linearity suggest a secondary reaction [2, 54, 63]. The quality of the dose response was linear in most cases, indicating that the selected X-ray dose range was adequate for footprinting, and that no secondary radiation damage caused sample perturbation (**Figure 5C**). The hydroxyl radical reactivity rate is governed by the side chain's intrinsic reactivity as well as solvent accessibility. The ratio of the site-specific rate of modification for free and complexed protein provides the most useful information by reflecting only solvent accessibility changes (**Table S2** and **Figure 6A**). Using liquid chromatography coupled mass spectrometry thus identifies the conformational hot spots and provides a ratiometric measure of solvent accessibility with residue-level resolution. An advantage of this study is that we have a high-resolution structure of the SpyCatcher001-SpyTag001 complex, which we used to model the SpyCatcher003-SpyTag001 complex and interpret the radiolytic labeling by comparing the relative solvent accessibility of modified residues in the free state to their rate of modification in the bound or complexed state. The SpyCatcher-SpyTag system was originally engineered from fibronectin-binding protein, FbaB, which contains a CnaB2 adhesin domain, and therefore, we adopted the CnaB2 secondary structural nomenclatures (**Figure S5**)[49, 64]. Those residues identified with changes in their modification rate in the bound vs. free state are consequently likely to be involved in the dynamic conformational changes that mediate SpyTag001 binding to SpyCatcher003 (**Figure 6B**). For informative comparisons, we discuss different types of ratio of hydroxyl radical reactivities or changes in solvent accessibility in the following section.

Solvent accessibility (SA) changes directly support local interactions in SpyCatcher003-SpyTag001 complex

Residues which showed a decrease in SA upon complex formation are Q11, P13, H26, K31, D33, R37, A42, T43, M44, D82, Y84, E85, residue 90 to 94, V100, T101, and residue 103 to 108 (**Figure 7A and B**). Residues D33, A42, T43, M44, Y84, and E85 showed more than 3-fold protection upon complex formation. Residue M44, on β sheet B, is projected toward the hydrophobic core and in close proximity to the isopeptide bond of the CnaB type fold. Adjacent residues A42 and T43 also showed similar protection. The isopeptide bond formation by SpyTag provides the β sheet G, thus completing the CnaB type fold and preventing the accessibility of solvent water to the core domain. In theory, we expect complete protection of such a buried methionine residue; however, any conformational fluctuation can result in access to solvent water, which results in modification of the highly reactive buried methionine residue.

Residue D33 and residues 84-85 are situated on the loop joining β sheets A and B and β sheets E and F, respectively, and both regions interact with the C-terminal half of SpyTag001 (**Figure 7A and B**). The proximity of D33 to R37, which also showed 1.5 fold protection, and Y119 of SpyCatcher003 suggest a potential site of interaction through a H-bonding network to stabilize this loop. The modifications of residues 84-85 were quantified from one of the modified extracted ion chromatograms of peptide 82-96, which showed Y84 modification mixed with E85. The Y84 and the E85 residues are located in proximity to Y119 and K120, respectively, of SpyTag001 and thus structurally stabilize the loop regions and therefore the complex formation by involving hydrophobic and ionic interactions as shown by the presence in preserving these residues while engineering SpyCatcher002[65]. Residue D82 , which is also part of the loop joining β sheets E and F, showed a small degree of protection upon complex formation. Overall, the decrease in the solvent accessibilities of these residues is not surprising, as it directly supports an interaction similar to that observed in the Spycatcher001-SpyTag001 crystal structure[50].

The protection at residues I90 and V100, which are in proximity to I113 and A111, respectively, indicated stabilization by a hydrophobic interaction at the N-terminal half of SpyTag001 and β sheet F of SpyCatcher003. Due to the lack of electron density in the crystal structure of the SpyCatcher001-SpyTag001 complex[50], the protection at Q11, P13, and residues 103-109 could not be directly explained in terms of binding. However, a previous report indicated an increase in the stabilization at both the N and C-terminus of SpyCatcher002 compared to SpyCatcher001, where mutation of several terminal residues showed increased binding affinity towards SpyTag[49]. This increased stability may persist in the SpyCatcher003 version, and would indicate a slightly more ordered structure in these regions, leading to the observed protections.

Peptide 22-32 from the free SpyCatcher003 sample showed modification at H26 and K31. The complex formation should generate crosslinked peptide with the SpyTag001 residues. We could not identify the native crosslinked C-terminal peptide tagged with the HiLyte T Fluor 488 and its modification product. However, we detected peptide 22-32 and its modification with 100-fold less abundance in the extracted ion chromatogram. Although the significant loss of native peptide in the digested sample suggests isopeptide bond formation, the presence of the minor peptide fraction might be due to radiation-induced isopeptide bond breakage or lack of isopeptide bond formation in some fraction of the sample population. The reduction in the abundance of native and modified peptide does not affect the quantification of the normalized fraction of unmodification as long as both native and the modification extracted ion current is well above the threshold level of the instrument's background noise performance.

The residues that showed hydroxyl radical reactivity but no change in the solvent accessibility upon complex formation are M1, L4, L7, M17, Y67, Y69, V86, A87, D97, and Q99. Most of these residues are on the non-interacting interface and are oriented toward the bulk solvent environment. Two residues, W57 and H62 located at β sheet C and loop joining sheet C to D showed a slight increase in solvent accessibilities. Surprisingly, our results show that E38

located in proximity H62 showed a strong increase in SA. These residues are located opposite to the SpyTag001 binding site, suggesting that their increase in solvent accessibility might be due to allosteric conformational changes. Most likely, E38 forms an electrostatic or H-bond interaction with H62 in SpyCatcher003 to stabilize the loop region. Upon binding of SpyTag, the β-a sheet could move to accommodate and bind to the SpyTag via H-bonding along the backbone, causing the E38 residue to become more surface accessible, but further studies would be needed to determine if this is the case. It is noteworthy that many aromatic side chains in the core hydrophobic fold remain unmodified in both the free and complexed protein. This indicated that the protein's core fold remained intact with and without the SpyTag001 peptide.

Understanding the driving force behind the formation of the SpyCatcher-SpyTag complex is challenging because of the unusual isopeptide bond formation. Mostly, mutation at the charged residues to increase electrostatic complementarity and increase the rigidity in the loop joining β sheets E and F increased the stability of the activated state structure, which might have lowered the activation energy for covalent bond formation in the complex, leading to a near diffusion controlled reaction rate. XFMS data rationalized the location of key interactions and corroborated the crystal structure and HDX studies reported earlier[49]. Several additional structural features can be gathered from solvent accessibility measurements after careful comparison of CnaB folds[64]. Cell-surface proteins from gram-positive bacteria forming these types of isopeptide bonds share low sequence similarity but broadly conserved structural motifs, in which two aromatics, Y73 and Y84, situated at the opposite ends of the hydrophobic β sheet core, are highly conserved or replaced by F (**Figure S5, Figure 7C & D**). The isopeptide bond is buried inside a hydrophobic pocket covered by Y84 in the SpyCatcher crystal structure, and is conserved among various structures. Residue Y84 is further stabilized by Y119 from β sheet G or the C-terminal half of SpyTag001. Modification of Y84 for both free and bound states indicated this region is not fully occluded from water accessibility or might have conformational flexibility that allows Y84 to

become accessible to a lesser extent to hydroxyl modification in the free state. In contrast, the hydrophobic core towards the N-terminal of SpyTag001 or β sheet, the moderately conserved Y73, as well as F75 and F92, do not show a modification in either free or bound state and are therefore completely occluded from water accessibility, indicating a rigid core formation. Previous HDX studies showed no stability change in these regions in SpyCatcher003 compared to SpyCatcher00238, indicating stability of this hydrophobic core is not influenced by mutation to generate higher affinity binding partners. Any change in stability at the other end of the β sheet core, which is opposite to Y84, might affect the activation energy of isopeptide bond formation and is a potential site of interest for further studies. The aliphatic interactions pinpointed by XFMS as described above open up the possibility of further improving binding efficiency by changing the hydrophobic interactions that affect the protein's stability. While activation energy is not a direct factor in protein-ligand efficiency and selectivity, it can indirectly impact protein-ligand interactions by affecting the kinetics of the binding mechanism and the overall thermodynamics of the interaction through structural changes or stability of the overall protein environment before and after binding interactions. Therefore, the effect of protein stability on the activation energy of ligand binding can vary widely and should be studied on a case-by-case basis. In many cases, reducing protein stability can hinder ligand binding, but the extent of this effect depends on the details of the interaction, particularly when it involves covalent bond formation. Additional time-resolved experiments, which are beyond our current studies' scope and goals, could delineate mechanistic details of isopeptide bond formation. In summary, comprehensive knowledge of global interaction by spectroscopy and local interaction by XFMS can further improve protein engineering and design.

Conclusion

The new integrated instrument combines high-speed liquid jet delivery with a high flux density micro-focused X-ray beam, and enables nondestructive fluorescence measurement nearly simultaneously with radiolysis for hydroxyl radical labeling. Here, we demonstrated the instrument using steady state FRET and XFMS on a SpyCatcher-SpyTag system to globally validate complex formation and identify key interacting residues simultaneously. Overall, the system provides a new integrated structural mapping platform that will be useful in design of new engineered protein systems such as the SpyCatcher-SpyTag system, and for monitoring global structural changes in protein conformation during or prior to irradiation. This instrument is the first prototype for a hybrid spectroscopy-XFMS data collection platform.

For future development, this instrument has the ability to capture absorption, fluorescence, and Raman spectrometry mono- or multimodally, at wavelengths from the UV to IR with single digit microliter to larger sample volumes for both destructive and nondestructive spectroscopy methods. The instrument can also be extended to accommodate pump-probe type photoactivation, time resolved hybrid multimodal spectroscopy-XFMS or temperature jump experiments with delays as short as double-digit microseconds to study transient intermediates.

Since a method for measuring dose rates and doses was also recently determined at this beamline[51], the new inline capabilities will be useful for monitoring radiation-induced events under controlled dose rates and doses, such as fluorescence monitoring of radiolysis products in various environments. The instrumentation can also be extended to spectroscopically monitoring radical chemistry during X-ray irradiation or post-X-ray irradiation on inorganic and organic materials that can be used for drug development, screening, and assay for radiotherapy treatments[66]. This type of analysis may be particularly informative for radiobiology and radiation therapy fields in which radiation damage must be precisely measured. This instrument design can also serve as a prototype for fluorescence imaging-based monitoring of intact cellular samples under controlled doses and dose rates, with a potential to follow up with standard proteomics approaches to monitor the effect of radiation damage on protein components and pathways[67, 68].

Conflict of interest

There are no conflicts to declare.

Data availability

The authors will supply the relevant data in response to reasonable requests.

20

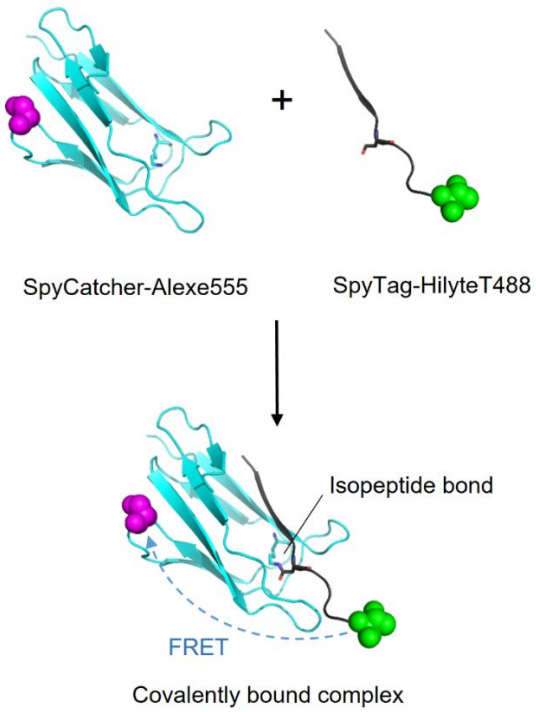

Figure 1. SpyCatcher-SpyTag split protein system. Schematics showing fluorescence tagged SpyTag001-Hilytie T488 binds to fluorescence tagged SpyCatcher003-Alexa555 that results in FRET between donor Hylite T-488 (green) and Alexa555 (purple)

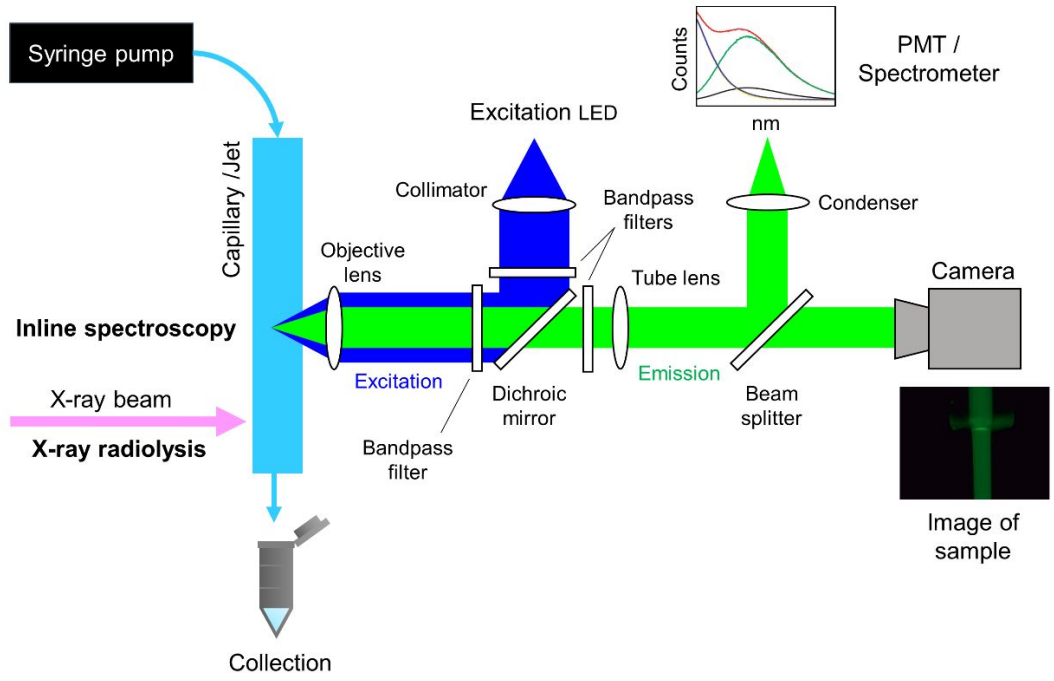

Figure 2. Schematics of a fluorescence imaging module (FIM) for XFMS. FIM consists of a light source and various optical components, including an objective lens, collimating lens, condenser lens, filter cube containing a wavelength-specific dichroic mirror, band pass excitation and emission filter, reflective mirror, and imaging camera. The position of the FIM is interchangeable for pre and post-exposure CCD or PMT-based data collection. More than one FIM can be accommodated on the sample pathway for multimodal inline spectroscopy XFMS.

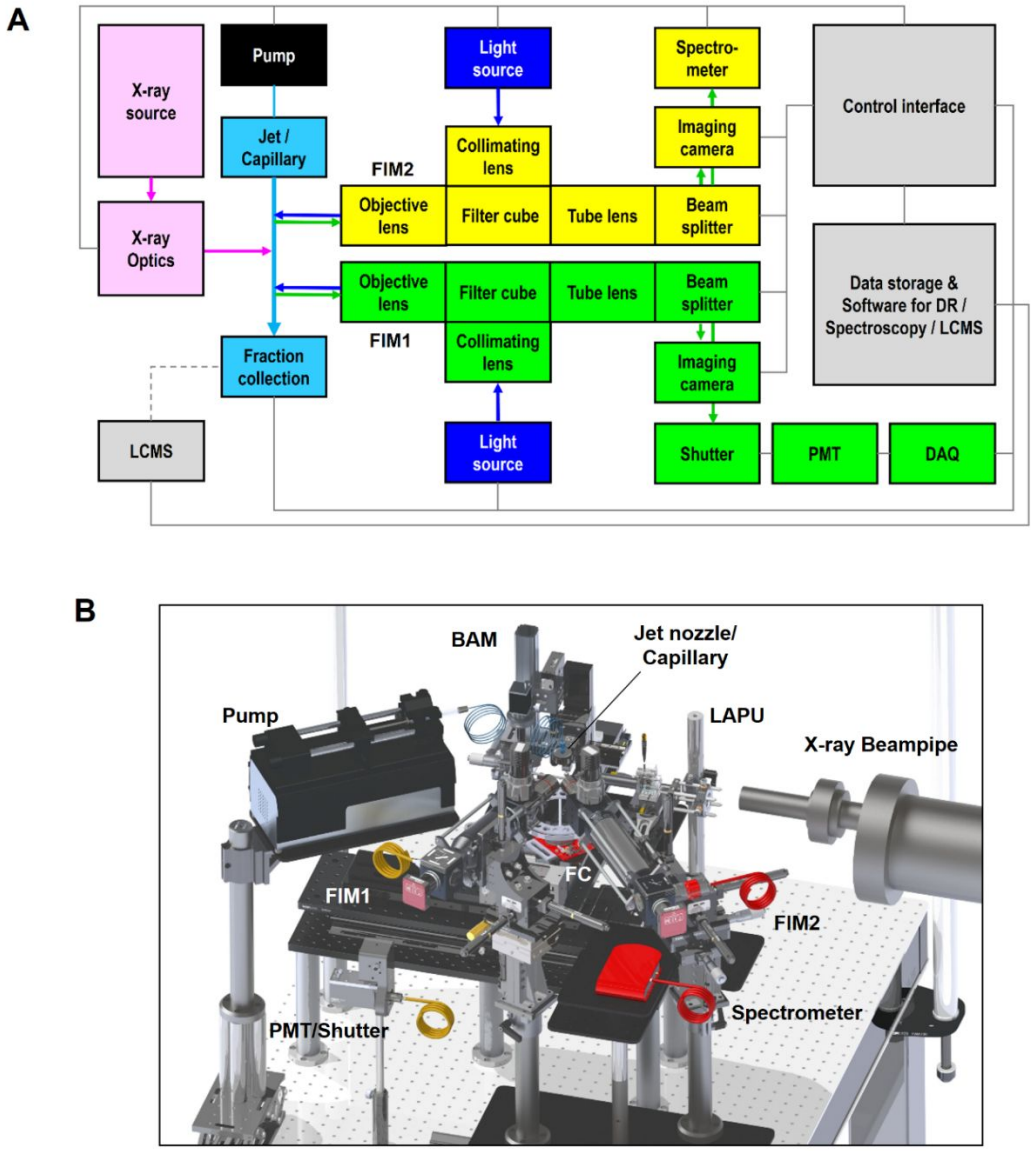

Figure 3. Integrated instrument design for hybrid spectroscopy-XFMS (A) Block diagram showing simplified system architecture of integrated components operated by the LabView based controller. Two fluorescence imaging modules (FIM) are colored grey and blue. Each FIM has similar optical components, including a light source, objective and other lens, filter cube containing wavelength-specific mirror and filter components, and imaging camera. The FIM2 and FIM1 designated for collecting emission spectrum and Alexa dose response, is focused on the sample vertically above and below the X-ray exposure point, respectively. The focused synchrotron X-ray is aligned with the sample using a laser assisted pre-alignment unit (LAPU). The data collected by FIMs for spectroscopy and Alexa dose-response are coupled to a CCD spectrometer and shutter-PMT-DAQ system, respectively. Data is stored in the G-drive for shared access through the control interface. (B) The position of critical modules and components – beam alignment module (BAM), LAPU, jet – capillary module, FIMs, , PMT shutter, and X-ray beampipe is shown by the computer-aided design of the actual instrument assembly, which is under user operation at beamline 3.3.1 at the Advanced Light Source.

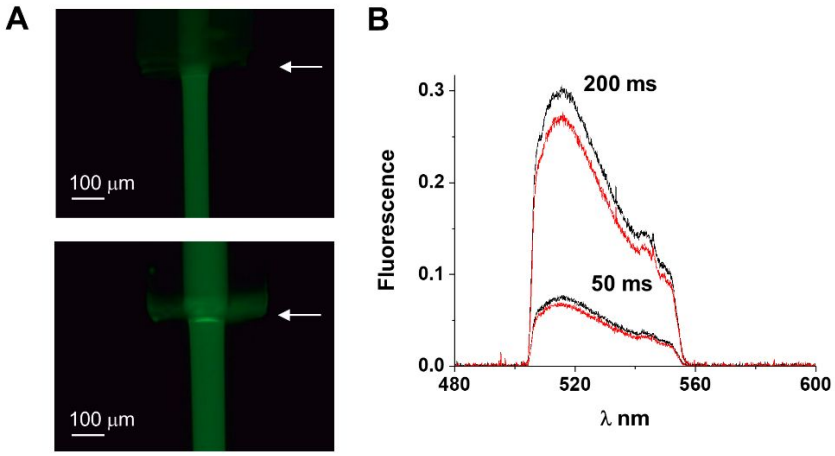

Figure 4. Comparison of Alexa fluorescence emission spectra from the liquid jet and capillary (A) FIM generated inline microscopic view of Alexa fluorescence emitted from the 360 μm OD / 75 μm ID glass capillary (top) and the 360 μm OD / 100 μm ID quartz capillary (bottom) and the ejected 75 μm and 100 μm liquid jet from the respective jet nozzles, which are indicated by arrows. (B) Alexa emission spectra from a 100 μm liquid jet (red) and through quartz of 360 μm OD / 100 μm ID quartz capillary (black) at two integration times as described in the methods section.

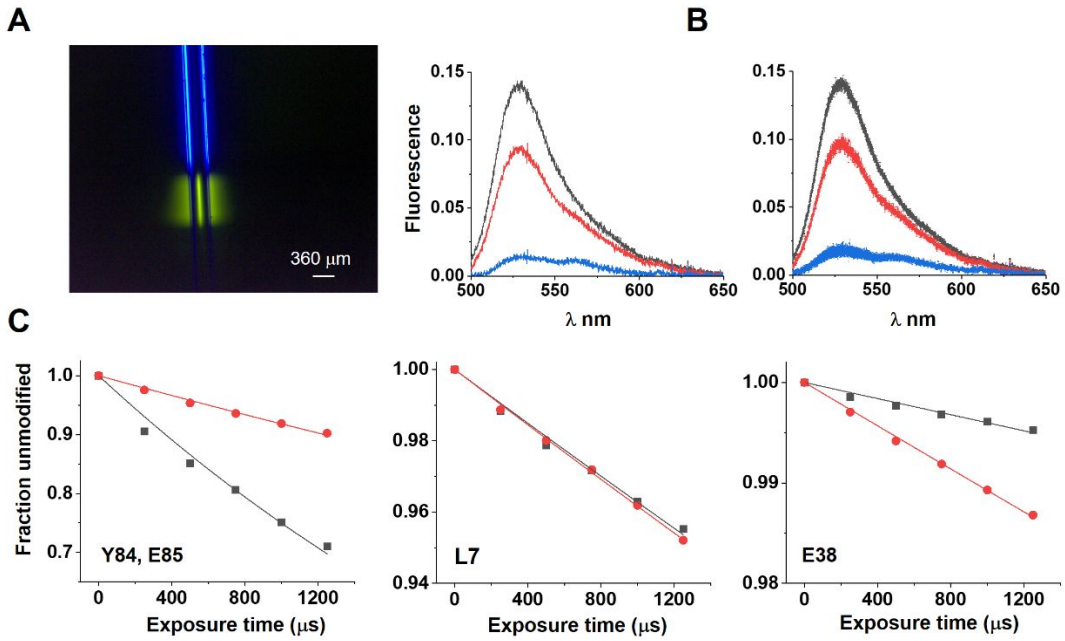

Figure 5. Simultaneous spectroscopy and XFMS data collection (A) Microscope image of the location of spectroscopy probe illuminated by the ~ 490 nm excitation wavelength LED light through the objective lens, and the location of X-ray exposure illuminated by Nd:YAG fluorescence just in front of the 200 μm ID capillary. (B) Fluorescence emission spectra showing FRET in the SpyCatcher003-SpyTag001 complex for one flow speed (right) and different flow speeds averaged (left). The spectra are obtained from SpyTag001 alone (black), SpyCatcher003-SpyTag001 complex (red) and SpyCatcher003 alone (blue) at 5uM concentration with 40 ms integration time.. The superimposed SpyCatcher001 is shown in cyan and SpyTag001 is shown in dark grey. The location of Alexa 555 and Alexa 488 labeling is shown in purple and green space fill cartoon, respectively. (C) Representative residue specific dose response of SpyCatcher003 in the presence (red) and absence (black) of SpyTag001, collected using the capillary system. Points are collected at a fixed dose rate for various exposure times, and fit to a single exponential fit (solid lines) to determine the hydroxyl radical reactivity k (s^{-1}), as described in the methods section.

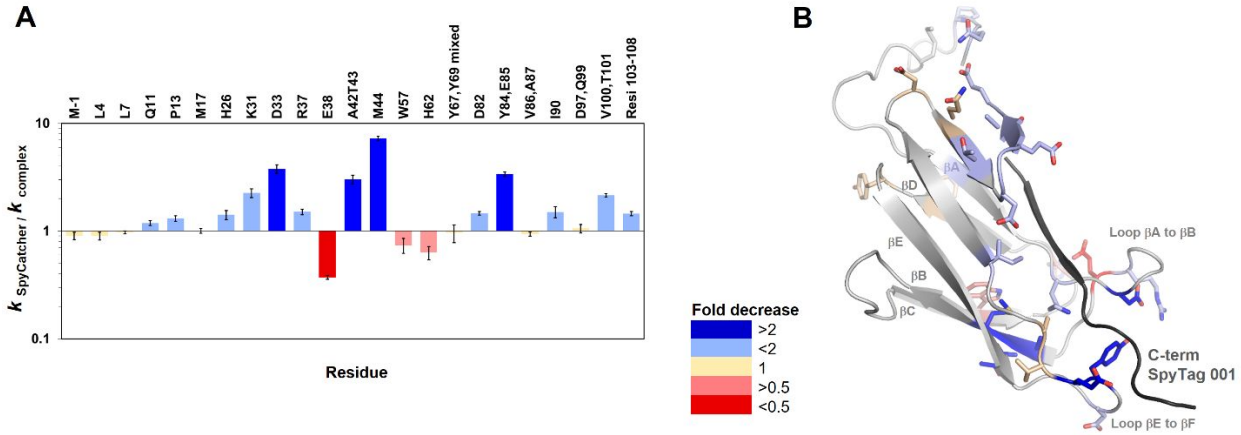

Figure 6. XFMS derived residue specific solvent accessibility change (A) Bar plot showing SA change of SpyCatcher003-SpyTag001interactions. The color profile indicated fold change in the solvent accessibility from free SpyCatcher003 to SpyCatcher003-SpyTag001 complex. The error bar shows the max and min values of the ratio shown in **Table S1.** (B) Residues highlighted with the color profile indicated fold change in the solvent accessibility from free SpyCatcher003 to SpyCatcher003-SpyTag001 complex. The structural model of SpyCatcher003 in grey color was generated by homology modeling using PDB 2X5P. The superimposed SpyTag001 is shown in dark grey.

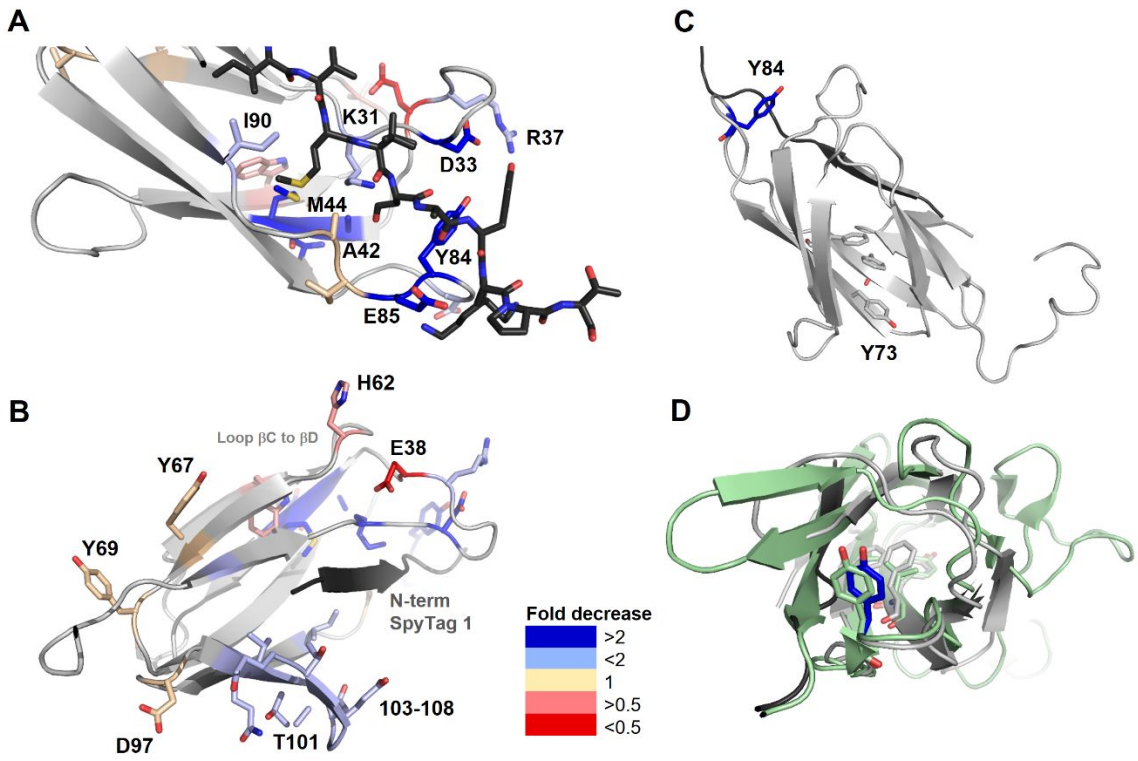

Figure 7. **Visualization of local interactions in the structural model of SpyCatcher003**. (A and B) Residues highlighted with the color profile indicated fold change in the solvent accessibility from free SpyCatcher003 to SpyCatcher003-SpyTag001 complex. The structural model of SpyCatcher003 in grey color was generated by homology modeling using PDB 2X5P. The superimposed SpyTag001 is shown in dark grey. (C) Position of conserved aromatics at the hydrophobic core. (D) Structural superposition of S. *agalactia* minor pilin Gbs52 (PDB 3PHS) and SpyCatcher001 (PDB 4MLI), and the location of conserved tyrosine, including the Y84 of SpyCatcher (red) in the hydrophobic pocket.

Acknowledgments

Work was conducted at the Advanced Light Source beamline (ALS) 3.3.1 and at the Molecular Foundry, which are supported by the Office of Science of the U.S. Department of Energy (DOE) under contract DE-AC02-05CH11231. Footprinting development and operations at ALS beamline 3.3.1 is supported by National Institutes of Health 1R01GM126218 and P30GM124169.

Supporting Information Available

Further information on protein purification, protein sequence alignment, instrument configuration and operation, LCMS data collection and analysis, and rate constant of hydroxyl radical modification.

References

1. A. Bolje and S. Gobec, *Pharmaceutics*, 2021, **13**.
2. S. Gupta, J. Feng, L. J. Chan, C. J. Petzold and C. Y. Ralston, *J Synchrotron Radiat*, 2016, **23**, 1056-1069.
3. S. Gupta, Y. Chen, C. J. Petzold, D. P. DePonte and C. Y. Ralston, *Anal Chem*, 2020, **92**, 1565-1573.
4. R. L. Leverenz, M. Sutter, A. Wilson, S. Gupta, A. Thurotte, C. Bourcier de Carbon, C. J. Petzold, C. Ralston, F. Perreau, D. Kirilovsky and C. A. Kerfeld, *Science*, 2015, **348**, 1463-1466.
5. S. Gupta, C. Merriman, C. J. Petzold, C. Y. Ralston and D. Fu, *The Journal of biological chemistry*, 2019, **294**, 13327-13335.
6. J. Bohon, L. D. Jennings, C. M. Phillips, S. Licht and M. R. Chance, *Structure*, 2008, **16**, 1157-1165.
7. S. Gupta, J. Chai, J. Cheng, R. D'Mello, M. R. Chance and D. Fu, *Nature*, 2014, **512**, 101-104.
8. C. L. Durie, J. Lin, N. W. Scull, K. L. Mack, M. E. Jackrel, E. A. Sweeny, L. M. Castellano, J. Shorter and A. L. Lucius, *Biophysical journal*, 2019, **116**, 1856-1872.
9. M. Schoof, B. Faust, R. A. Saunders, S. Sangwan, V. Rezelj, N. Hoppe, M. Boone, C. B. Billesbolle, M. Zimanyi, I. Deshpande, J. Liang, A. A. Anand, N. Dobzinski, B. S. Zha, B. Barsi-Rhyne, V. Belyy, A. W. Barile-Hill, S. Gupta, C. R. Simoneau, K. Leon, K. M. White, S. Nock, Y. Liu, N. J. Krogan, C. Y. Ralston, D. L. Swaney, A. Garcia-Sastre, M. Ott, M. Vignuzzi, P. Walter and A. Manglik, *bioRxiv : the preprint server for biology*, 2020, DOI: 10.1101/2020.08.08.238469.
10. B. N. Chaudhuri, S. Gupta, V. S. Urban, M. R. Chance, R. D'Mello, L. Smith, K. Lyons and J. Gee, *Biochemistry*, 2011, **50**, 1799-1807.
11. S. Gupta, M. Guttman, R. L. Leverenz, K. Zhumadilova, E. G. Pawlowski, C. J. Petzold, K. K. Lee, C. Y. Ralston and C. A. Kerfeld, *Proc Natl Acad Sci U S A*, 2015, **112**, E5567-5574.
12. M. A. Dominguez-Martin, M. Hammel, S. Gupta, S. Lechno-Yossef, M. Sutter, D. J. Rosenberg, Y. Chen, C. J. Petzold, C. Y. Ralston, T. Polivka and C. A. Kerfeld, *Scientific reports*, 2020, **10**, 15564.

13. T. Orban, B. Jastrzebska, S. Gupta, B. Wang, M. Miyagi, M. R. Chance and K. Palczewski, *Structure*, 2012, **20**, 826-840.
14. Y. Du, N. M. Duc, S. G. F. Rasmussen, D. Hilger, X. Kubiak, L. Wang, J. Bohon, H. R. Kim, M. Wegrecki, A. Asuru, K. M. Jeong, J. Lee, M. R. Chance, D. T. Lodowski, B. K. Kobilka and K. Y. Chung, *Cell*, 2019, **177**, 1232-1242 e1211.
15. R. Jain, D. Abel, M. Rakitin, M. Sullivan, D. T. Lodowski, M. R. Chance and E. R. Farquhar, *J Synchrotron Radiat*, 2021, **28**, 1321-1332.
16. R. Jain, N. S. Dhillon, E. R. Farquhar, B. Wang, X. Li, J. Kiselar and M. R. Chance, *Anal Chem*, 2022, **94**, 9819-9825.
17. I. Shcherbakova, S. Mitra, R. H. Beer and M. Brenowitz, *Nucleic acids research*, 2006, **34**, e48.
18. Y. Zhu, T. Guo, J. E. Park, X. Li, W. Meng, A. Datta, M. Bern, S. K. Lim and S. K. Sze, *Molecular & cellular proteomics : MCP*, 2009, **8**, 1999-2010.
19. A. M. Edwards, M. Ruiz, E. Silva and E. Lissi, *Free Radical Res*, 2002, **36**, 277-284.
20. J. S. Sharp, J. M. Becker and R. L. Hettich, *Analytical biochemistry*, 2003, **313**, 216-225.
21. T. T. Aye, T. Y. Low and S. K. Sze, *Anal Chem*, 2005, **77**, 5814-5822.
22. D. M. Hambly and M. L. Gross, *J Am Soc Mass Spectr*, 2005, **16**, 2057-2063.
23. B. B. Stocks and L. Konermann, *Anal Chem*, 2009, **81**, 20-27.
24. S. Vahidi, B. B. Stocks, Y. Liaghati-Mobarhan and L. Konermann, *Anal Chem*, 2013, **85**, 8618-8625.
25. S. Vahidi and L. Konermann, *J Am Soc Mass Spectrom*, 2016, **27**, 1156-1164.
26. S. Gupta, M. Sullivan, J. Toomey, J. Kiselar and M. R. Chance, *J Synchrotron Radiat*, 2007, **14**, 233-243.
27. M. Rosi, B. Russell, L. G. Kristensen, E. R. Farquhar, R. Jain, D. Abel, M. Sullivan, S. M. Costello, M. A. Dominguez-Martin, Y. Chen, S. Marqusee, C. J. Petzold, C. A. Kerfeld, D. P. DePonte, F. Farahmand, S. Gupta and C. Y. Ralston, *Commun Biol*, 2022, **5**, 866.
28. J. S. Sharp, S. K. Misra, J. J. Persoff, R. W. Egan and S. R. Weinberger, *Anal Chem*, 2018, **90**, 12625-12630.
29. B. Xie and J. S. Sharp, *Anal Chem*, 2015, **87**, 10719-10723.
30. A. M. Orville, R. Buono, M. Cowan, A. Heroux, G. Shea-McCarthy, D. K. Schneider, J. M. Skinner, M. J. Skinner, D. Stoner-Ma and R. M. Sweet, *J Synchrotron Radiat*, 2011, **18**, 358-366.
31. N. Shimizu, T. Shimizu, S. Baba, K. Hasegawa, M. Yamamoto and T. Kumasaka, *J Synchrotron Radiat*, 2013, **20**, 948-952.
32. J. Kern, R. Alonso-Mori, R. Tran, J. Hattne, R. J. Gildea, N. Echols, C. Glockner, J. Hellmich, H. Laksmono, R. G. Sierra, B. Lassalle-Kaiser, S. Koroidov, A. Lampe, G. Han, S. Gul, D. Difiore, D. Milathianaki, A. R. Fry, A. Miahnahri, D. W. Schafer, M. Messerschmidt, M. M. Seibert, J. E. Koglin, D. Sokaras, T. C. Weng, J. Sellberg, M. J. Latimer, R. W. Grosse-Kunstleve, P. H. Zwart, W. E. White, P. Glatzel, P. D. Adams, M. J. Bogan, G. J. Williams, S. Boutet, J. Messinger, A. Zouni, N. K. Sauter, V. K. Yachandra, U. Bergmann and J. Yano, *Science*, 2013, **340**, 491-495.
33. A. F. Chrimes, K. Khoshmanesh, P. R. Stoddart, A. Mitchell and K. Kalantar-Zadeh, *Chemical Society reviews*, 2013, **42**, 5880-5906.
34. B. Kuswandi, Nuriman, J. Huskens and W. Verboom, *Analytica chimica acta*, 2007, **601**, 141-155.
35. P. Minzioni, R. Osellame, C. Sada, S. Zhao, F. G. Omenetto, K. B. Gylfason, T. Haraldsson, Y. B. Zhang, A. Ozcan, A. Wax, F. Mugele, H. Schmidt, G. Testa, R. Bernini, J. Guck, C. Liberale, K. Berg-Sorensen, J. Chen, M. Pollnau, S. Xiong, A. Q. Liu, C. C. Shiue, S. K. Fan, D. Erickson and D. Sinton, *J Optics-Uk*, 2017, **19**.
36. T. E. Angel, S. Gupta, B. Jastrzebska, K. Palczewski and M. R. Chance, *Proc Natl Acad Sci U S A*, 2009, **106**, 14367-14372.

37. R. Hussein, M. Ibrahim, A. Bhowmick, P. S. Simon, R. Chatterjee, L. Lassalle, M. Doyle, I. Bogacz, I. S. Kim, M. H. Cheah, S. Gul, C. de Lichtenberg, P. Chernev, C. C. Pham, I. D. Young, S. Carbajo, F. D. Fuller, R. Alonso-Mori, A. Batyuk, K. D. Sutherlin, A. S. Brewster, R. Bolotovsky, D. Mendez, J. M. Holton, N. W. Moriarty, P. D. Adams, U. Bergmann, N. K. Sauter, H. Dobbek, J. Messinger, A. Zouni, J. Kern, V. K. Yachandra and J. Yano, *Nature communications*, 2021, **12**, 6531.
38. J. L. S. Lopes, A. P. U. Araujo and D. M. Jameson, *Methods Appl Fluores*, 2015, **3**.
39. J. A. Ross, M. A. Gilmore, D. Williams, K. R. Aoki, L. E. Steward and D. M. Jameson, *Analytical biochemistry*, 2011, **413**, 43-49.
40. D. M. Jameson and J. A. Ross, *Chem Rev*, 2010, **110**, 2685-2708.
41. R. Sridharan, J. Zuber, S. M. Connelly, E. Mathew and M. E. Dumont, *Biochimica et biophysica acta*, 2014, **1838**, 15-33.
42. J. R. Lakowicz, *Springer* Third Edition
43. U. Samuni and J. M. Friedman, *Methods in molecular biology*, 2005, **305**, 287-300.
44. A. A. Bunaciu, H. Y. Aboul-Enein and V. D. Hoang, *Appl Spectrosc Rev*, 2015, **50**, 377-386.
45. S. A. Oladepo, K. Xiong, Z. M. Hong, S. A. Asher, J. Handen and I. K. Lednev, *Chem Rev*, 2012, **112**, 2604-2628.
46. E. Kish, M. M. Pinto, D. Kirilovsky, R. Spezia and B. Robert, *Biochimica et biophysica acta*, 2015, **1847**, 1044-1054.
47. K. Kajimoto, T. Kikukawa, H. Nakashima, H. Yamaryo, Y. Saito, T. Fujisawa, M. Demura and M. Unno, *The journal of physical chemistry. B*, 2017, **121**, 4431-4437.
48. G. G. Kochendoerfer, S. Kaminaka and R. A. Mathies, *Biochemistry*, 1997, **36**, 13153-13159.
49. A. H. Keeble, P. Turkki, S. Stokes, I. N. A. Khairil Anuar, R. Rahikainen, V. P. Hytonen and M. Howarth, *Proc Natl Acad Sci U S A*, 2019, **116**, 26523-26533.
50. L. Li, J. O. Fierer, T. A. Rapoport and M. Howarth, *Journal of molecular biology*, 2014, **426**, 309-317.
51. S. Gupta, J. L. Inman, J. Chant, L. Obst-Huebl, K. Nakamura, S. M. Costello, S. Marqusee, J. H. Mao, L. Kunz, R. Paisley, M. C. Vozenin, A. M. Snijders and C. Y. Ralston, *Radiation research*, 2023, **200**, 523-530.
52. G. Xu, J. Kiselar, Q. He and M. R. Chance, *Anal Chem*, 2005, **77**, 3029-3037.
53. J. Saladino, M. Liu, D. Live and J. S. Sharp, *J Am Soc Mass Spectrom*, 2009, **20**, 1123-1126.
54. L. G. Kristensen, J. M. Holton, B. Rad, Y. Chen, C. J. Petzold, S. Gupta and C. Y. Ralston, *J Synchrotron Radiat*, 2021, **28**, 1333-1342.
55. S. Gupta, V. N. Bavro, R. D'Mello, S. J. Tucker, C. Venien-Bryan and M. R. Chance, *Structure*, 2010, **18**, 839-846.
56. L. Ma, F. Yang and J. Zheng, *Journal of molecular structure*, 2014, **1077**, 87-100.
57. W. R. Algar, N. Hildebrandt, S. S. Vogel and I. L. Medintz, *Nature methods*, 2019, **16**, 815-829.
58. S. Naudi-Fabra, M. Tengo, M. R. Jensen, M. Blackledge and S. Milles, *Journal of the American Chemical Society*, 2021, **143**, 20109-20121.
59. A. Margineanu, J. J. Chan, D. J. Kelly, S. C. Warren, D. Flatters, S. Kumar, M. Katan, C. W. Dunsby and P. M. French, *Scientific reports*, 2016, **6**, 28186.
60. A. B. Ghisaidoobe and S. J. Chung, *International journal of molecular sciences*, 2014, **15**, 22518-22538.
61. P. R. Gleason, P. I. Kelly, D. W. Grisingher and J. H. Mills, *Organic & biomolecular chemistry*, 2020, **18**, 4079-4084.
62. A. Hawe, M. Sutter and W. Jiskoot, *Pharmaceutical research*, 2008, **25**, 1487-1499.
63. S. Gupta, R. Celestre, C. J. Petzold, M. R. Chance and C. Ralston, *J Synchrotron Radiat*, 2014, **21**, 690-699.
64. H. J. Kang and E. N. Baker, *Trends in biochemical sciences*, 2011, **36**, 229-237.

65. A. H. Keeble, A. Banerjee, M. P. Ferla, S. C. Reddington, I. Anuar and M. Howarth, *Angewandte Chemie*, 2017, **56**, 16521-16525.
66. J. Geng, Y. Zhang, Q. Gao, K. Neumann, H. Dong, H. Porter, M. Potter, H. Ren, D. Argyle and M. Bradley, *Nature chemistry*, 2021, **13**, 805-810.
67. A. R. Buchberger, K. DeLaney, J. Johnson and L. Li, *Anal Chem*, 2018, **90**, 240-265.
68. J. F. Cahill, V. Kertesz and G. J. Van Berkel, *Anal Chem*, 2015, **87**, 11113-11121.

Data availability

The authors will supply the relevant data in response to reasonable requests.