CHARLES UNIVERSITY

Faculty of Science

Department of Biochemistry

Study program: Biochemistry

Mgr. Marek Polák

Použití pulsního značení pro studium dynamiky proteinů a proteinových komplexů

Utilization of pulse labelling techniques for studying the dynamics of proteins and protein complexes

DOCTORAL THESIS

Supervisor: RNDr. Petr Novák, Ph.D.

Consultant: RNDr. Petr Man, Ph.D.

Prague, 2024

DECLARATION

I declare that I have worked on this thesis under the guidance of my supervisor and that all sources of the previous knowledge are properly cited. No part of this work was used or will ever be used to obtain any other academic degree than Ph.D. from Charles University.

In Prague, ……

……

Signature

DECLARATION OF AUTHORSHIP

I declare that Mgr. Marek Polák contributed significantly to the experiments and to all scientific publications contained in this Ph.D. thesis. He performed most of the experiments, substantially contributed to their planning, and took a significant part in the primary data interpretation and their preparation for publication.

In Vestec, ……

……

RNDr. Petr Novák, Ph.D.

Acknowledgments

I would like to thank my academic supervisors, Dr. Petr Novák, Dr. Petr Man, and Professor Michael Volny for guiding me through this research and for their time and willingness to help me. Many thanks also go to all other members of the Laboratory of Structural Biology and Cell Signaling at Biocev for creating such a friendly atmosphere.

I sincerely thank my parents and sister for their unwavering support throughout my studies.

I would also like to acknowledge all our funding sources, including the Charles University Grant Agency (359521), the Czech Science Foundation (19-16084S) and NPO-NEURO-EX- CELLES (LX22NPO5107).

Abstract (In English)

Mass spectrometry (MS) techniques are routinely used to probe the structure and dynamics of proteins and protein complexes. Although MS techniques lack the high resolution of data provided by X-ray crystallography, NMR, or cryo-EM, they excel in providing insights into analyte dynamics, structure, and interactions with other components, such as ligands.

This doctoral thesis presents a contribution to the field of structural biology employing and extending covalent labelling approaches, namely Fast Photochemical oxidation of Proteins (FPOP) and oxidation by singlet oxygen ($^1O_2$). These approaches were followed to study the structure, dynamics, and interaction of proteins, nucleic acids, and protein-DNA complexes in solution. Initially, FPOP was used to investigate the interaction interface of FOXO4 and DAF16-DNA response element and to show the possibilities of analyzing such a complex using both 'bottom-up' and 'top-down' approaches. Furthermore, an isotope depletion strategy combined with multiCASI-ECD proved effective in delivering structural information with the highest possible resolution for mapping protein-DNA interfaces. This research showcases how information derived from structural proteomic methods can guide the construction of *in-silico* models for protein-DNA complexes with dynamic structures or interactions, which remain unclear. Moreover, this thesis reports the first adoption of FPOP to induce hydroxyl radical-induced DNA damage coupled to high-resolution MS analysis. By studying damage to double-stranded IRE and FOXO4-IRE complex, we have elucidated the principles of DNA damage analysis by high-resolution LC-MS. By studying a ternary complex consisting of TEAD1-FOXO4-DNA, we have highlighted how FPOP can easily capture even minor conformational changes on DNA upon protein binding. And by MS structural analysis of the protein LOV2, we have further demonstrated that top-down technology is an effective tool for analyzing protein oxidative damage, as shown on a model protein covalently labelled by singlet oxygen ($^1O_2$).

Key words: Structural mass spectrometry, Fast photochemical oxidation of proteins (FPOP), protein-DNA complexes, FOXO4, TEAD1, DNA damage, photosensitizers, AsLOV2.

Abstract (In Czech)

Hmotnostní spektrometrie (MS) se stala častou technikou pro sledování struktury a dynamiky proteinů a proteinových komplexů. I když techniky založené na MS detekci nemohou poskytovat informace s vysokým rozlišením, jako je např. rentgenová difrakce, nukleární magnetická rezonance (NMR), nebo kryo-elektronová mikroskopie (cryo-EM), jsou dobrým nástrojem z hlediska poskytování informaci ohledně dynamiky, struktury a interakcí analytů s dalšími molekulami, např. ligandy.

Tato doktorská práce představuje přínos do oblasti strukturní biologie při využití kovalentního značení, tzv. rychlé fotochemické oxidace proteinů (FPOP) a oxidace singletním kyslíkem ($^1O_2$) pro studium struktury, dynamiky a interakcí proteinů a proteinových komplexů s dvoušroubovici DNA.V první části práce byla metoda FPOP využita k studiu struktury komplexu FOXO4-DAF16, která poukázala na možnosti analýzy takového komplexu pomocí přístupů "bottom-up" a "top-down". Dále je zde představena strategie izotopové deplece pro strategii „top-down“, kde je demonstrována její výhoda ve spojení s multiCASI-ECD za účelem získaní nejvyšší možné strukturní informace ohledně interakce FOXO4 s IRE. Studie ukazuje, jak mohou informace získané ze strukturní proteomiky vést k vytvoření *in-silico* modelů protein-DNA komplexů, jejichž struktury o vysokém rozlišení jsou náročné na přípravu. Dále, práce představuje použití hydroxylových radikálu pro mapování poškození DNA, které je poprvé kvantifikováno MS analýzou o vysokém rozlišení. Studium poškození dvouvláknového oligonukleotidu IRE bez a v přítomnosti proteinu FOXO4 objasňuje analytické principy analýzy poškození DNA pomocí hmotnostní spektrometrie a naznačuje její další využití pro studium vyšších nadmolekulových struktur nukleových kyselin a pro určení interakčního rozhraní. Studium ternárního komplexu TEAD1-FOXO4-DNA ukazuje, jak snadno chemické sondy v podobě hydroxylových radikálů můžou zachytit i malé konformační změny na DNA po vazbě transkripčního faktoru. MS analýza LOV2 proteinu pomocí MS dále ukazuje, že je top-down efektivním nástrojem pro kvantitativní analýzu oxidativního poškození proteinu. V tomto případě je to demonstrováno na kovalentně označeném modelovém proteinu pomocí singletového kyslíku, $^1O_2$.

Klíčová slova: Strukturní hmotnostní spektrometrie, rychlá fotochemická oxidace proteinů, protein-DNA komplexy, FOXO4, TEAD1, poškození DNA, fotosensitizér, protein AsLOV2.

Contents

ABBREVIATIONS

$^1O_2$	Singlet oxygen
AnPEP	*Aspergillus Niger* Prolyl Endoprotease
CASI	Continuous accumulation of selected ions
CD	Circular dichroism
CID	Collision-induced dissociation
Cryo-EM	Cryo-electron microscopy
DBD	DNA-binding domain
DBE	DNA-binding element
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
DSC	Differential scanning calorimetry
dsDNA	Double-strand DNA
DTT	1,4-dithiothreitol
ECD	Electron-capture dissociation
EDC	N-ethyl-3-N',N'-dimethylaminopropylcarbodiimide
EMSA	Electrophoretic mobility shift assay
ESI	Electrospray ionization
ETD	Electron-transfer dissociation
ExD	Electron-based dissociation methods
FMN	Flavin mononucleotide
FPOP	Fast photochemical oxidation of proteins
FRET	Förster resonance energy transfer
FT	Fourier transform
FT-ICR	Fourier-transform ion cyclotron resonance mass spectrometry
GEE	Glycine ethyl ester
GEPSs	Genetically encoded photosensitizers
H_2O_2	Hydrogen peroxide
HDX	Hydrogen/deuterium exchange
HFIP	1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP)
IAA	Iodoacetic acid

IAM	Iodoacetamide
ID	Isotopically depleted
IN	Isotopically natural
IR	Infrared
IRE	Insulin Response element
IRMPD	Infrared multiphoton dissociation
LC	Liquid chromatography
LOV2	light, oxygen and voltage sensing domain 2
mAb	Monoclonal antibody
MALDI	Matrix-assisted laser desorption/ionization
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
nESI	Nano-ESI
NHS	*N*-Hydroxysuccinimide
NMR	Nuclear magnetic resonance
PS	Photosensitizer
RNA	Ribonucleic acid
S/N, SNR	Signal to noise ratio
SASA	Solvent accessible surface area
SAXS	Small angle X-ray scattering
SDS-PAGE	Polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate (SDS-PAGE)
SNAP	Sophistical Numerical Annotation Procedure
TCEP	tris(2-carboxyethyl)phosphine
TEA	Triethylamine
TFs	Transcription factors
TOF	Time of flight
tRNA	Transfer ribonucleic acid
UV	Ultraviolet
UVPD	Ultraviolet photodissociation
VKOR	Vitamin K epoxide reductase

1. Introduction

Cellular processes are intrinsically linked to biomolecules, such as proteins, nucleic acids, lipids, and polysaccharides, which form complex networks through their interactions. Each of these biomolecules plays a distinct role in cells, steering various pathways that activate, block, modulate or regulate a wide range of cellular responses. For instance, nucleic acids encode genetic information, undergoing replication during cell division and transcription into RNA. In turn, RNA can either serve its own function or be further translated into proteins during proteosynthesis[1]. But while proteins are constructed from only 20 amino acids, posttranslational modifications give arise to a plethora of proteoforms with different functions. Thus, deciphering their roles in health and disease requires elucidating protein structures and understanding protein dynamics and molecular interactions with other biomolecules[2,3].

1.1 Experimental methods for studying protein structure.

Over the past few decades, a toolkit of experimental techniques leveraging both chemical and physical properties have been developed to investigate the structure of biomolecules. While these methods provide high-resolution images, they often yield only a static snapshot, which is biased toward energetically favorable and stable biomolecular configurations[4,5]. Currently, the most common approaches for providing high-resolution images include X-ray crystallography, nuclear magnetic resonance (NMR), and cryogenic electron microscopy (cryo-EM).

Since the first 3D structure determination of sperm-whale myoglobin in 1958[6], X-ray crystallography has become widely used to obtain high-resolution structural models of proteins and DNA[5]. This method utilizes X-rays which pass through a substance, causing elastic bending (diffraction), albeit with a very weak diffraction signal of a single molecule. When molecules become arranged in the well-structured configuration of a protein crystal, the intensity multiplies as a result of constructive interferences in the crystal grid. The intensities of these reflections are then used to generate an experimental electron density, from which an atomic coordinate is derived, providing an experimental protein structure. However, some challenges still persist, mostly regarding signal phase determination and successful

crystal preparation[5]. Another obstacle is that X-ray crystallography provides only static snapshot of a molecule. Therefore, we cannot study protein dynamics in solution by X-ray crystallograpy[2].

In contrast to X-ray crystallography, NMR enables us to study protein dynamics in solution. Briefly, NMR utilizes a strong magnetic field to align spins of certain nuclei (^{1}H, ^{15}N, ^{13}C, and ^{31}P) in a sample atom. By irradiating the sample with a sequence of radiofrequency pulses, the energy is absorbed by nuclei occupying "lower spin state" where they excite to a "higher energy spin state". The signal released by nuclear relaxation is measured in NMR. Spatially distanced constraints are derived from these signals to reconstruct a three-dimensional structure of a biomolecule using *"isotopically enriched"* proteins. These proteins are recombinantly expressed in a bacteria culture growing in a minimal media containing glucose and ammonium sulphate as sources of ^{13}C and ^{15}N, respectively. Unlike X-ray crystallography, NMR omits the crystallization step, enabling us to measure proteins in solution and thus study their dynamics on a millisecond timescale. This high-resolution technique for 3D structure determination has, however, several limitations, including (i) higher sample consumption, (ii) the price of isotopically labelled products, and (iii) size limitations of studied samples, which is currently ~40-50 kDa, and (iv) the complexity associated with interpreting the resulting data[7,8].

Broadly speaking, *Cryo-EM* is a type of transmission electron microscopy involving a procedure known as 'vitrification'. Vitrification entails flash-freezing a protein sample in a liquid ethane, which results in the formation of amorphous ice surrounding a sample while preserving its 3D structure. During sample measurements, electron beams with wavelengths similar to atoms (~1 Å) interact with thin layers of a frozen sample. Transmitted electrons carry the information regarding the structure, which is then transformed into an image. A thousand 2D images are acquired from all possible orientations. With this information, computational software is able to reconstitute the 3D structure of a molecule. As in X-ray crystallography, one of the last steps in Cryo-EM consists of manually fitting a polypeptide chain with a known sequence into the reconstituted model to obtain a high-resolution image of the molecule.

As shown in Figure 1, several biophysical techniques, including small-angle X-ray scattering (SAXS)[9], and spectroscopic methods, such as Förster resonance energy transfer (FRET)[10] and circular dichroism (CD)[11], provide structural information with lower resolution. Among them, mass spectrometry (MS) techniques stand out. MS-based approaches can deliver a wealth of information on biomolecules, encompassing composition, stoichiometry, interactome, interaction interfaces, conformational envelope, solvent-accessible surface, topology, conformational changes, dynamics, kinetics, and posttranslational modifications[2]

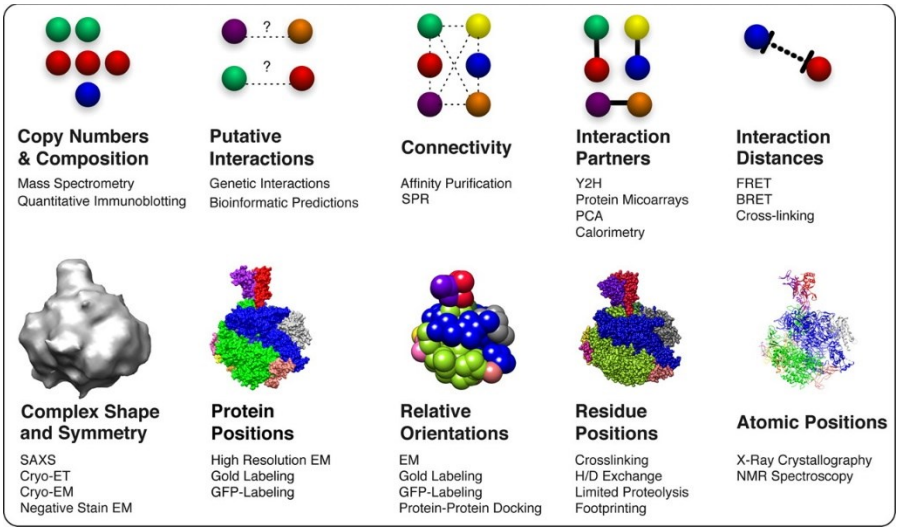


Figure 1 | The image illustrates diverse biophysical techniques used to collect data for obtaining different information regarding the studied biomolecule(s). Reprinted from ref [12].

1.2 Mass spectrometry approaches for studying biomolecules.

Currently, two complementary approaches are available for mass spectrometry analysis of biomolecules, namely the 'bottom-up' and the 'top-down' approaches (Figure 2).

Bottom-up approach. The bottom-up approach refers to characterizing proteins by analyzing peptides generated by proteolytic digestion (~0.5 kDa – 3 kDa) prior to MS analysis. Proteolytic digestion with commercially available proteases, such as Trypsin, LysC, LysN, GluC, ArgC, Chymotrypsin, AspN or AnPEP (*Aspergillus Niger* Prolyl Endoprotease),[13] produces peptides, which are analyzed either directly or first separated by liquid chromatography (LC) or using other separation techniques. Subsequently, the peptides are introduced to a mass spectrometer via a

suitable ionization source. The mass of the peptides is recorded, and if necessary, precursors are fragmented to obtain information about the sequence, often by collision-induced dissociation (CID)[14–18]. As a hybrid of these two approaches, the 'middle-down' approach is used to analyze longer peptides[18,19].

Top-down approach. *Top-down approach*. The ‘top-down’ approach enables us to analyze intact proteins without prior digestion. Protein samples can be directly introduced to a mass spectrometer using ESI or MALDI ionization techniques. The ion of interest is isolated in the gas phase, and fragmentation can be induced with various fragmentation techniques. The resulting fragment ions provide information on protein isoforms, posttranslational modifications, and sequence characterization, making it possible to identify and quantify novel proteoforms.

Typically, the analyte of interest is denatured, desalted on a reversed phase column, and sprayed with a volatile solvent (methanol, acetonitrile, or isopropanol) containing a small amount of an organic acid (~0.1-1%, formic or acetic acid) before being introduced to a mass spectrometer through an electrospray ionization source (ESI). In the gas phase, the mass spectrum of the analyte is recorded on a *m/z* axis. The selected charge state of the analyte of interest is introduced to several fragmentation techniques to gather information. Popular ion-activation methods for biomolecule fragmentation include collision induced dissociation (CID), electron-capture dissociation (ECD), electron-transfer dissociation (ETD), infrared multi-photon dissociation (IRMPD) and ultraviolet photo dissociation (UVPD)[15,20–23].

In recent years, the top-down approach for biomolecular analysis has become popular due to (i) the increased availability of both FT-ICR and Orbitrap instruments, (ii) technological advancements in ionization techniques and fragmentation methods, and (iii) the possibility to analyze larger biomolecules, such as monoclonal antibodies (mAb)[15,24]. Despite its drawbacks, such as low throughput, limited sensitivity, and a lack of residue-level sequence coverage to gather residue-level information, the top-down approach is particularly attractive when bottom-up struggles to provide any structural information due to missing peptides or overdigested regions. Moreover, in a bottom-up approach, information on the “specific proteoform of interest” is lost when digesting the whole bulk of protein(s), which can be overcome using the top-down approach. Hence, combining results from both approaches seems to be the most reasonable option for acquiring more comprehensive information on an analyte[25].

Collision-induced dissociation. Gas-phase collisional activation remains the most popular method for characterizing proteins or peptides. In CID, the kinetically activated precursor molecule collides with inert gas molecules, converting kinetic energy into internal energy, which is redistributed throughout the polypeptide backbone, thereby fragmenting the molecule[26]. This fragmentation occurs on the weakest/lowest energy bond, that is, the peptide bond, generating *b*- and *y*-fragment ions, including internal fragment ions (Figure 3). CID is performed in a device known as collision cell, where ions are accelerated through the collision cell with a defined potential.

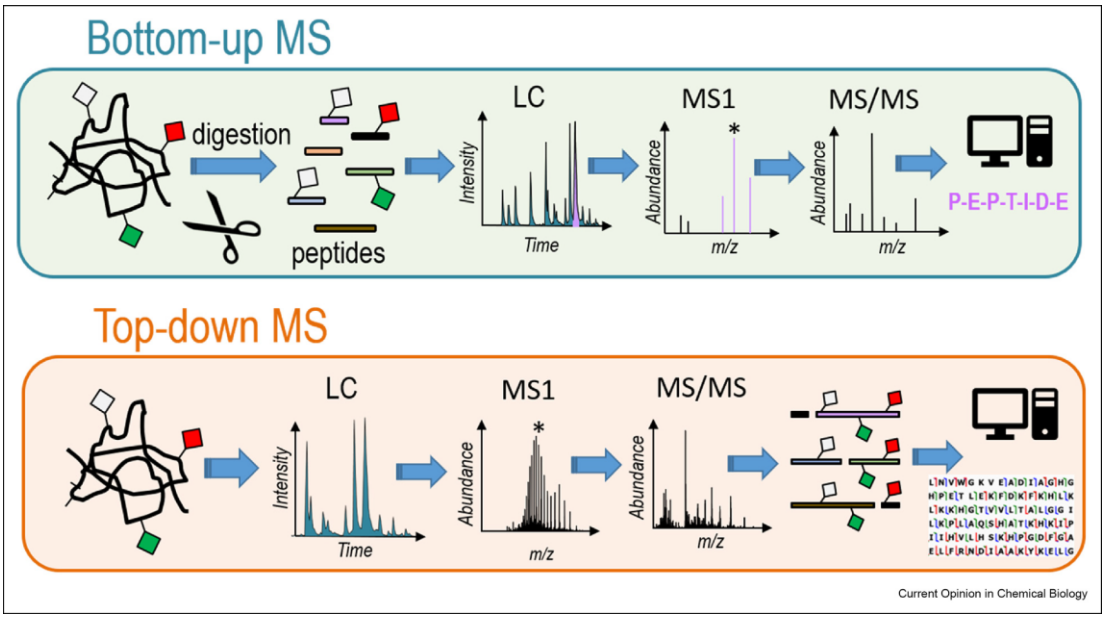


Figure 2 | The top-panel displays the usual bottom-up workflow for the analysis of biomolecules, while the bottom panel displays the typical top-down workflow. Reprinted from[27].

Notwithstanding its popularity, CID has some limitations: (i) the high-energy collision leads to fragmentation and loss of posttranslational modifications (e.g. phosphorylation, O-sulfation N-/O-glycosylation, hydroxylation, nitration, tyrosine nitration), including neutral loss of water, resulting in both qualitative and quantitative loss of information, (ii) fragmentation may prevail between specific residues, such as proline and aspartic acid, and in the N- and C-termini of proteins[20,21]. Although CID is tremendously robust for small peptides and readily implemented to each MS platform, its low efficiency has spurred the exploration and implementation of other activation/dissociation methods described below.

Electron-based dissociation methods (ExD). Several electron-based fragmentation techniques are available to researchers, but the most popular remain *electron-capture dissociation (ECD)* and *electron-transfer dissociation (ETD).* These methods were simultaneously developed over the years to overcome the limitations of CID[28,29]. ECD was first described in 1998 by the McLafferty group. This technique involves transferring ions to the ICR cell and capturing a low-energy thermal electron at a protonated backbone amide group. This process results in a charge-reduced aminoketyl radical, dissociating the N-C(α) bond, which is then converted into a more stable amide, and the C-terminal side of the radical. ECD has long been regarded as an "non-ergodic" process, meaning that the bond cleavage occurs on a timescale that prevents energy redistribution over the polypeptide backbone. However, recent studies have suggested that when aminoketyl radial is formed, the N-C(α) bond is actually thermodynamically very labile. For this reason, cleavage occurs directly at this position[30,31].

Electron-transfer dissociation (ETD). ETD was first described by John Syka and colleagues in 2004[32]. In ETD, an anionic reagent, such an anthracene, azobenzene, azulene and fluoranthene, is introduced to a collision cell when trapping positively charged proteins or cations. The interaction between the negatively charged anion and the positively charged analyte results in electron transfer to a positively charged analyte, ultimately leading to charge reduction and fragmentation. Both ECD and ETD fragmentation generate a set of complementary fragment ions, namely *c* and $z^{\bullet}$ ions (Figure 3) and cleave disulphide bonds. Thanks to their low vibrational energy redistribution, ETD and ECD preserve posttranslational modifications, so they are particularly suitable for proteomics[30,33,34].

ECD was developed not only for FT-ICR instruments but also for Orbitrap and Omnitrap platforms instruments [20,30,35,36]. ETD is mostly performed in quadrupole type ion traps and collision cells (triple-quadrupole instruments) and thus easier to implement [34].

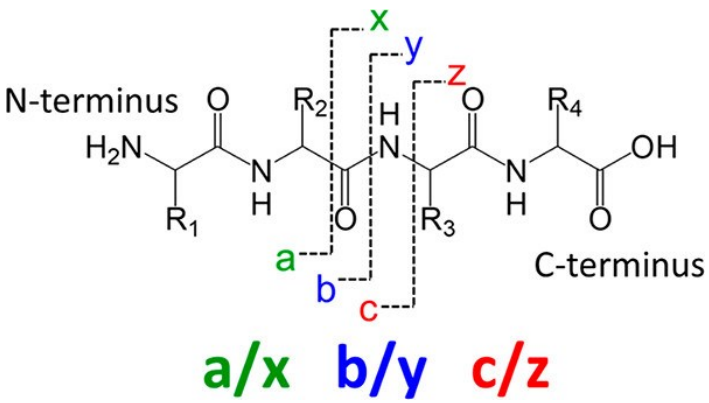


Figure 3 | Nomenclature for the peptide and protein backbone fragmentation in the gas-phase. Reprinted from[20].

Photon-based dissociation laser. Given the increasing need for alternative ion activation methods to solve complex biological problems, photon-based activation methods have been under development for the past 20 years. In photon-based dissociation, gas-phase trapped ions are exposed to a photon varying in energy (and thus wavelength), leading to the accumulation of internal energy and subsequent fragmentation by infra-red multiphoton dissociation, IRMPD (whereby ions are exposed to tens or hundreds of lower-energy photons with ~0.1 eV/photon), or ultraviolet photodissociation (UVPD), involving only a few higher-energy photons with ~2-10 eV/photon.

Various lasers were developed for gas-phase dissociation of biomolecules over the years, including CO_2 (10.6 μm), excimer (F_2 – 157 nm, ArF – 193 nm, Nd:YAG lasers – 266 nm or 355 nm) and optical parametric oscillator Nd:YAG (OPO-Nd:YAG) lasers, which offer wavelengths ranging from 205 to 2550 nm[20,37–40].

Integrating lasers with mass spectrometers requires (i) a suitable source of photons for dissociation, (ii) introducing a laser beam to a mass spectrometer, and (iii) aligning the ion cloud with the laser beam to perform dissociation[20]. The first lasers were integrated with FT-ICR mass spectrometers[38,41], but they have also been coupled to time-of-flight instruments[42], to both linear and quadrupole ion trap systems and to orbitrap instruments[37,43]. In combination with high-vacuum systems, which restrict additional collisional activation, the laser type, its energy or wavelength and the number of pulses (one-shot pulse vs. multiple shots) can be selected to modulate the fragmentation efficiency and the type and distribution of fragment ions[21,44–49].

Ion activation using photons at IR wavelength leads to the photoabsorption, vibrational excitation matching the collisional activation (heating the molecule in the gas-phase), which ultimately results in fragmentation and generates *b*- and *y*- type ions, including internal fragment ions (Figure 3)[38]. Complementary *a*–B, *w*-, and c- and y-fragment ions are generated by IRMPD in deoxyribooligonucleotides (DNA) and ribonucleotides (RNA), respectively[37,38,47,50–52].

In contrast, activation of gas-phase trapped ions by high-energy UV light causes excitation of molecules. Because conversion into the ground electronic state occurs throughout multiple dissociation channels, extensive fragmentation ensues, generating all types of ions, namely *a*-, *a*+1-, *b*-, *c*-, *x*-, *x*+1, *y*, *y*-1 and *z*-type, including fragmentation in side chains or disulfide bond cleavage[49]. Even the wavelength of UVPD is broader, with several lasers operating at different wavelengths. Currently, a 193nm dissociation is widely applied in the field[47,49,53,54].

1.3 Labelling techniques used to study higher-order structure of biomolecules.

Although genetic information is translated from mRNA to the primary sequence of proteins during proteosynthesis, their functional properties are encoded in their 3D structure. Individual protein units oligomerize and interact with other proteins and ligands in their “quaternary protein structure”, a “higher-order structure”, which is one of the main subjects of structural proteomics. Hence, structural proteomics, or structural mass spectrometry, refers to techniques which utilize mass spectrometer as an analytical tool for characterizing macromolecules.

Covalent and radical labelling plays a key role in mapping protein surfaces and macromolecular interactions. The rationale behind this experiment is labile or stable covalent labelling of solvent-accessible protein components. After labelling, the mass of modified peptides bearing labelling changes. Mass spectrometry can detect these changes, both qualitatively (localization in MS/MS mode) and quantitatively (intensity of mass increment).

1.3.1 Footprinting techniques

Advances in mass spectrometry, including the introduction of electrospray ionization (ESI)[55], coupled with chromatographic separation, have given rise to radical footprinting techniques. The first footprinting techniques, namely limited proteolysis and DNA footprinting, provided 'footprints' in terms of irreversible protein cleavage and DNA fragmentation, respectively. Current footprinting techniques induce covalent modifications in proteins to map their structural rearrangement or interaction with other biomolecules. Footprinting techniques are mainly associated with MS analysis.

1.3.2 Limited proteolysis

Introduced 60 years ago[56–58], limited proteolysis is considered to be one of the oldest footprinting techniques for probing higher-order structure of proteins and protein-protein interaction. The simple concept of limited proteolysis relies on the fact that cleavage site depends not only on the protein sequence but also on the accessibility of particular amino acid residues towards the protease (Figure 4). This technique has been employed for separating individual domains in larger proteins, with smaller protein domains often used for crystallization.

Limited proteolysis is used to analyze global protein structural changes and to specifically identify ligand-induced structural changes. In addition, in-cell proteolysis, protein-protein binding site mapping, protein aggregation analysis and protein quality control also involve limited proteolysis, and this technique can also be used to monitor virus-capsid conformational dynamics. The first limited proteolysis experiment was performed through trypsin digestion coupled with gel electrophoresis separation. Currently, other proteases, like Proteinase K, can also be used. However, since the first attempts to couple this technique with MS analysis[59], the products of reaction are today analyzed exclusively in this way[60–62].

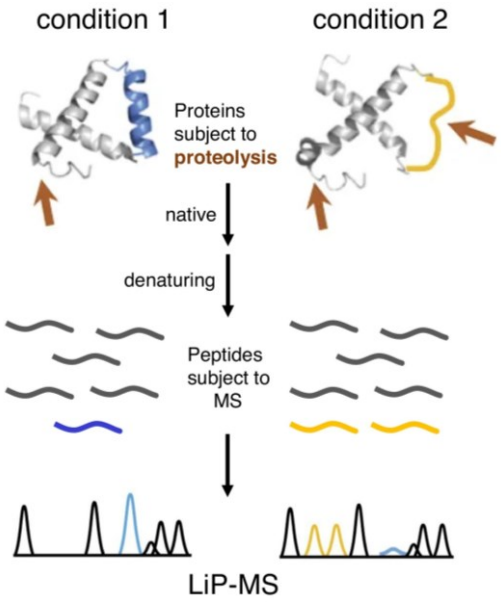


Figure 4 | A limited proteolysis typical workflow. Adapted and modified from ref. [63].

1.3.3 DNAse footprinting

DNAse I footprinting was developed to monitor protein-DNA interaction. First described in 1978, DNAse footprinting consists of (i) monitoring protein-DNA interactions and (ii) locating specific binding motif of transcription factor on the DNA. However, this technique uses deoxyribonuclease I enzyme to cleave double-stranded DNA. In pioneering studies, either the 5' or the 3' terminus was labeled, and the DNA sample was digested by DNAse I into (n-1, n=base) smaller fragments, which were separated by denaturing gel electrophoresis and visualized. When performing the reaction in the presence of the transcription factor, which physically restricted digestion around its binding motif, these bands were observed as "diminished", making it possible to perform the readout directly from gel. This technique was eventually replaced by radical Fenton chemistry[64,65].

1.3.4 Covalent labelling techniques

1.3.4.1 Hydrogen/Deuterium exchange

Hydrogen/deuterium exchange coupled to mass spectrometry is widely regarded as a valuable tool for mapping protein structure and protein interactions across academia and the biopharmaceutical industry. The fundamentals of exchanging hydrogens for deuterium were first described in the 1950s by Kaj Ulrik Linderstrøm-Lang and

colleagues. Initially measuring the stability of hydrogen bonds involved in protein secondary structures, deuterium incorporation acts as a label in a protein backbone, providing data on solvent access to amides and on the overall structure of the native protein[66,67]. While coupled with NMR in pioneering experiments conducted in the 1980s, H/D exchange is currently coupled with bottom-up MS analysis[68,69].

The typical HDX experiment is initiated by diluting a native protein into a deuterium-based buffer, at a physiological pH, to allow H/D exchange. Solvent-exposed amide hydrogens (protium, ^{1}H) are replaced by deuterium (^{2}H) over time. This substitution is strongly influenced by factors such as secondary structure and solvent access to proteins or ligands[70]. Theoretically, every hydrogen can be exchangeable for deuterium. However, hydrogens of carbon atoms are too strong to be exchanged. Hydrogen atoms of labile groups (hydroxy –OH, primary amine –NH, carboxy –COOH, sulfhydryl –SH) exchange at a fast kinetic rate, complicating their MS detection. In contrast, hydrogens of histidine (C(2)–H) and proline exchange at an excessively slow kinetic rate. Therefore, amide hydrogens of the protein backbone are suitable for monitoring because they are (i) exchangeable on a broader timescale, ranging from milliseconds to months, (ii) uniformly distributed throughout the polypeptide chain, and (iii) sensitivity to minor structural changes[4,68].

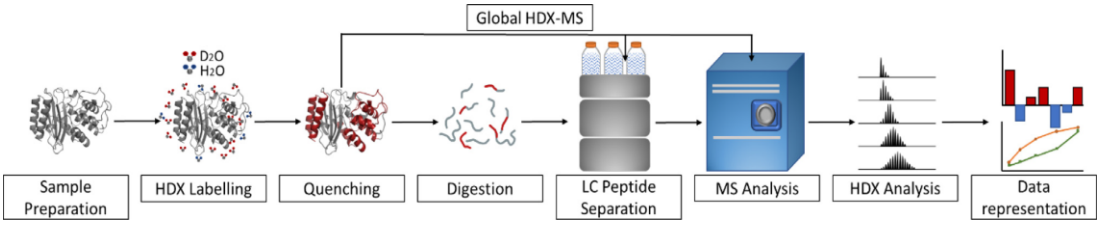


Figure 4 | The usual HDX bottom-up workflow. Reprinted from ref. [68].

The usual workflow of H/D exchange is initiated by protein incubation in a deuterium-based buffer, from which samples are collected at specific time points. Deuteration is immediately quenched by mixing with a buffer at a pH ~2.5-3, quickly flash freezing aliquots in liquid nitrogen to ensure minimal exchange kinetics[70]. The thawed, but still acidified, sample is then injected into a column containing immobilized protease. Despite remaining the gold standard protease for HDX experiments, pepsin may not be particularly useful for all proteins, e.g., membrane proteins. As a result, a portfolio of proteases is now available,[16,70–77] offering versatility in HDX experiments. Once digested, the peptides are then desalted and separated on the reverse-phase

analytical column of an LC system. The LC system is then coupled to MS to analyze peptides and their deuterated forms. With the data on the aliquots collected at different time points, we can construct deuteration plots to gain detailed insights into protein regions which undergo structural changes, such as ligand binding[70,78].

1.3.5 Chemical covalent labelling

Stable covalent labelling of biomolecules using chemical labelling agents might also provide information related to protein structure or higher-order structure of biomolecules. Chemical reagents, on one hand, target specific functional groups of amino acids. On the other hand, the reaction is not fast and must be performed at native-like conditions to avoid protein denaturation. After labeling experiment, LC-ESI-MS/MS analysis is preferred to identify modified residues and their extent of modification. This section will briefly summarize labelling compounds.

Cysteine alkylation: Thank to its nucleophilic thiol side chain, cysteine is a very reactive amino acid. In proteins, cysteine most likely creates cystine bridges with other cysteine, thus helping to preserve the tertiary protein fold. This residue is also very important in cellular responses to changes in the redox environment, such as oxidative stress[79]. The most common cysteine reaction is reduction and alkylation using various reagents prior to proteolytic digestion. Since cystine bridges help to preserve the tertiary structure of proteins, they limit protease access to the protein backbone, that is why reduction is such a crucial step prior to proteolysis. Cystines are reduced predominantly using 1,4-dithiothreitol (DTT), β-mercaptoethanol or tris(2-carboxyethyl)phosphine (TCEP). To prevent self-renaturation, cysteines are then alkylated using iodoacetamide (IAM) or iodoacetic acids (IAA), in a covalent reaction (Figure 5A)[80].

Lysine alkylation: lysine contains a primary amine group, mostly located on the protein surface, facilitating protein footprinting. One of the first covalent footprinting experiments was performed in 1991 by Fenselau, Vestling and colleagues[81], who acetylated lysins with acetic anhydride while studying interactions between melittin and calmodulin. Another pioneering study[82] was conducted by Professor Przybylski, who studied egg-white lysozyme surface using lysine acetylation coupled with MS

analysis. A convenient acetylation reagent is *N*-hydroxysuccinimide ester (NHS-ester), which was first used to modify a protein in 1975[83]. Although various reagents can acetylate lysines, the most commonly used is N-hydroxysuccinimide ester (NHS-ester)[84]. In acetylation, NHS ester is also used as a reactive crosslinking leaving group (Figure 5B)[84].

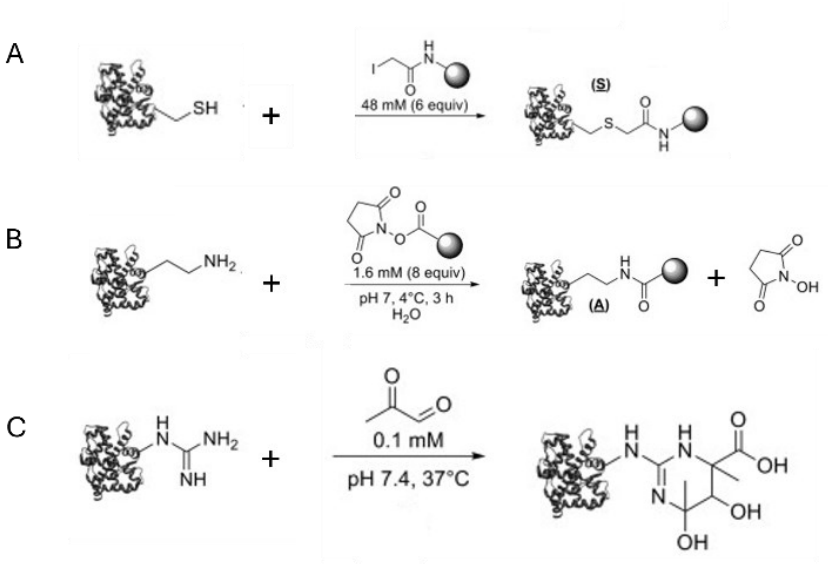


Figure 5 | A. Representative reaction of an alkylation of free cysteine. B. Representative acetylation reaction of primary lysine amine with NHS-ester. C. Reaction of arginine forming a stable pyrimidine ring. Adapted and modified from ref.[85].

Targeting *Aspartic* acid and *glutamic acids* functional groups is a useful method for protein footprinting as they bear functional carboxy groups. Carboxy groups can be esterified, as presented in the first experiments in 1945[86]. It's also possible to covalently label both amino acids using iodoacetamide (IAM)[87]. Today, for the needs of protein footprinting, the first step for labelling it's a carboxy group activation with almost exclusively used EDC reagent: N-ethyl-3-N',N'-dimethylaminopropylcarbodiimide. Activation of groups then afford improved labelling efficiency by another reagent being confidently characterized. For instance, glycine ethyl ester was used in works by Gross group[88] and recently in a Chance group[89]. Glycine ethyl ester (GEE), upon previous activation by EDC covalently labels Asp and Glu with characteristic product bearing mass shift +57.0215 Da.

Histidine has a unique character due to its imidazole ring, often found in enzyme active sites and in interaction interfaces[90]. The most extensively used histidine labelling reagent is diethylpyrocarbonate (DEPC), first described in 1960s[91]. Although DEPC labelling is irreversible, this compound is insoluble in water and

thus must be dissolved in organic solvents, such as methanol. Thus, careful mixing organic-based DEPC with protein is crucial to prevent organic-based denaturation. In 1996, Glocker et al. pioneered the use of DEPC in MS-based histidine footprinting[92]. However, DEPC labelling is not amenable to only His, also targeting Arg, Cys, Lys, Ser, Thr and Tyr. As such, DEPC is a convenient covalent probe for protein surface mapping[93,94].

Many reagents also target other amino acids. A particularly common reaction is tyrosine nitration with Tetranitromethane[95]. In addition, the reaction of Tryptophane with *N*-Chlorosuccinimide forms a keto group on an indole ring, and the reaction with the guanidyl group of arginine forms pyrimidine derivate (Figure 5C)[85]. Serine and Threonine are mostly modified by DEPC or NHS-esters at a low extent during labelling reactions of other residues[15].

1.3.6 Radical labelling techniques

Radical labelling techniques were developed during the last century to facilitate the study of the structure and interactions of biomolecules, more specifically proteins, DNA, and RNA. Among several other reactive oxygen species (ROS), hydroxyl radicals serve as a reactive probe. Their mode of action on proteins and nucleic acids has been extensively investigated, yielding several methods for its *in*-situ generation, which are used to study its impact on biomolecules. In cells, hydroxyl radicals react with proteins differently from DNA/RNA (Figure 6).

In reacting with DNA/RNA, the hydroxyl radical promotes hydrogen subtraction from ribose, causing DNA fragmentation into two parts (Figure 6A). DNA fragmentation with Fenton chemistry has been extensively used to assess the DNA-binding sequence of transcription factors, as shown in Figure 6A. Known as ‘hydroxyl radical footprinting of the DNA’, this technique will be described in detail below.

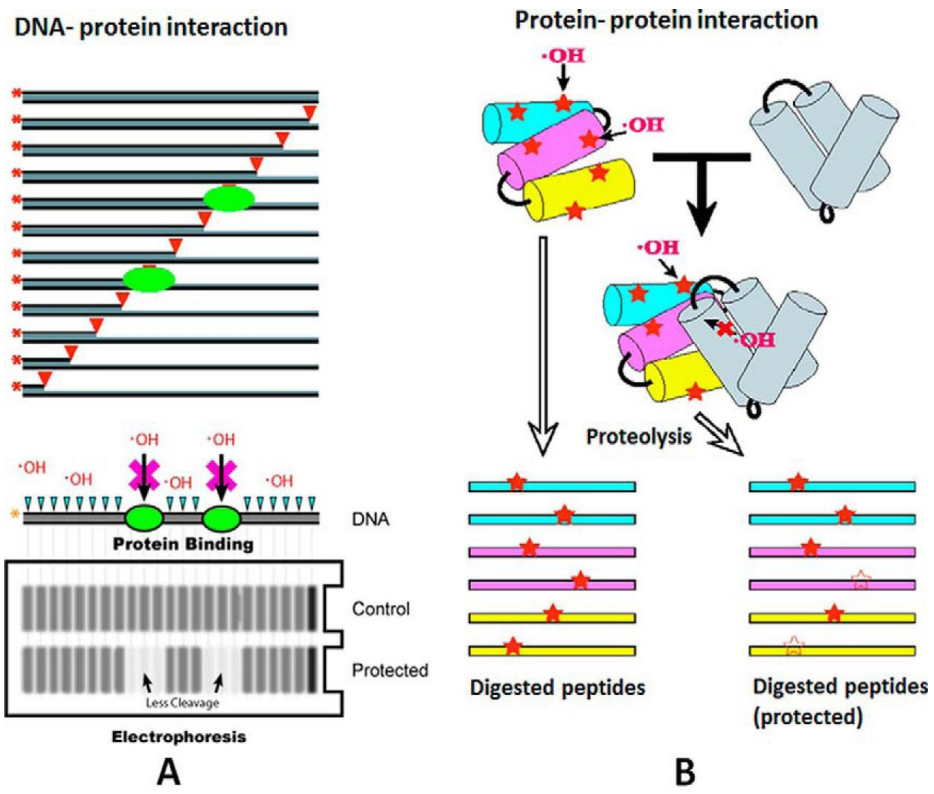


Figure 6 | A. Principles of hydroxyl radical footprinting of DNA, B. Principles of using the hydroxyl radical to study protein-protein interactions. Adapted from ref. [63].

Proteins undergo varying degrees of oxidation (Figure 6B). However, the most common modification involves the addition of +16 Da or +32 Da, represented as single and double oxygen addition, respectively. Nevertheless, hydroxyl radicals may induce various other modifications, as discussed in detail below.

In addition to the aforementioned Fenton chemistry, other techniques, such as synchrotron water radiolysis, electron pulse radiolysis, and hydrogen peroxide dissociation, have been employed to generate hydroxyl radicals. These approaches enable us to study proteins, DNA, and RNA, and their underlying chemistry is described in the following sections.

1.3.6.1 Fenton chemistry

The 'Fenton' reaction was first described by Henry Fenton in 1894[96]. However, the real presence of hydroxyl radical agents and the reaction mechanism was uncovered by Haber and Weiss in the first half of last century[97]. The reaction relies on generating hydroxyl radical by mixing iron ions with hydrogen peroxide. In addition to the sample and hydrogen peroxide, the reaction mixture also contains (i) EDTA to increase the solubility of Fe ions and (ii) ascorbic acid to induce a chain reaction by

reducing Fe(III) to Fe(II)[98]. Briefly, iron(II) present in the reaction mixture is oxidized to iron(III) by hydrogen peroxide, forming hydroxide ion and hydroxyl radical (equation 1 and 2). The reaction in equation 1 is an initiation reaction, reactions 3-6 are propagation reactions and may be chain reactions, and reaction 7 is the termination reaction[15,99]:

$$Fe(II) + H_2O_2 \rightarrow Fe(III) + OH^- + \bullet OH \quad \text{(eq. 1)}$$

$$Fe(III) + H_2O_2 \rightarrow Fe(II) + HO_2^{\bullet} + H^+ \quad \text{(eq. 2)}$$

$$H_2O_2 + \bullet OH \rightarrow HO_2^{\bullet} + H_2O \quad \text{(eq. 3)}$$

$$HO_2^{\bullet} \rightarrow O_2^{\bullet -} + H^+ \quad \text{(eq. 4)}$$

$$Fe(III) + HO_2^{\bullet} \rightarrow Fe(II) + O_2 + H^+ \quad \text{(eq. 5)}$$

$$Fe(III) + O_2^{\bullet -} \rightarrow Fe(II) + O_2 \quad \text{(eq. 6)}$$

$$Fe(II) + \bullet OH + H^+ \rightarrow Fe(III) + H_2O \quad \text{(eq. 7)}$$

Fenton chemistry-generated radicals were utilized to study protein-DNA interactions in a pioneering approach known as 'hydroxyl radical footprinting of DNA', first described by Thomas Tullius in 1986. As with DNAse I footprinting described above, hydroxyl radical footprinting of DNA can be used to monitor protein-DNA interactions (Figure 6). But unlike DNAse footprinting, which relies on enzymes for DNA cleavage, here the radicals are generated *in situ*[97,99]. The resulting hydroxyl radicals then subtract the most solvent accessible DNA (or RNA) hydrogens in the B-form duplex DNA[100,101], namely C4-H and C5-H, thereby fragmenting the nucleic acid (Figure 7). Transcription factor binding decreases solvent access to hydrogens around the binding motif, which decreases DNA damage. When gel electrophoresis is performed to separate DNA fragments, the gel readout is displayed as a "diminished" number of gel bands around the binding motif. DNA and RNA footprinting remains a popular method as hydroxyl radicals can also be generated by X-ray synchrotron water radiolysis[98,102–106]. Moreover, hydroxyl radicals generated by Fenton chemistry have also been used to oxidize protein in mapping protein surfaces[107].

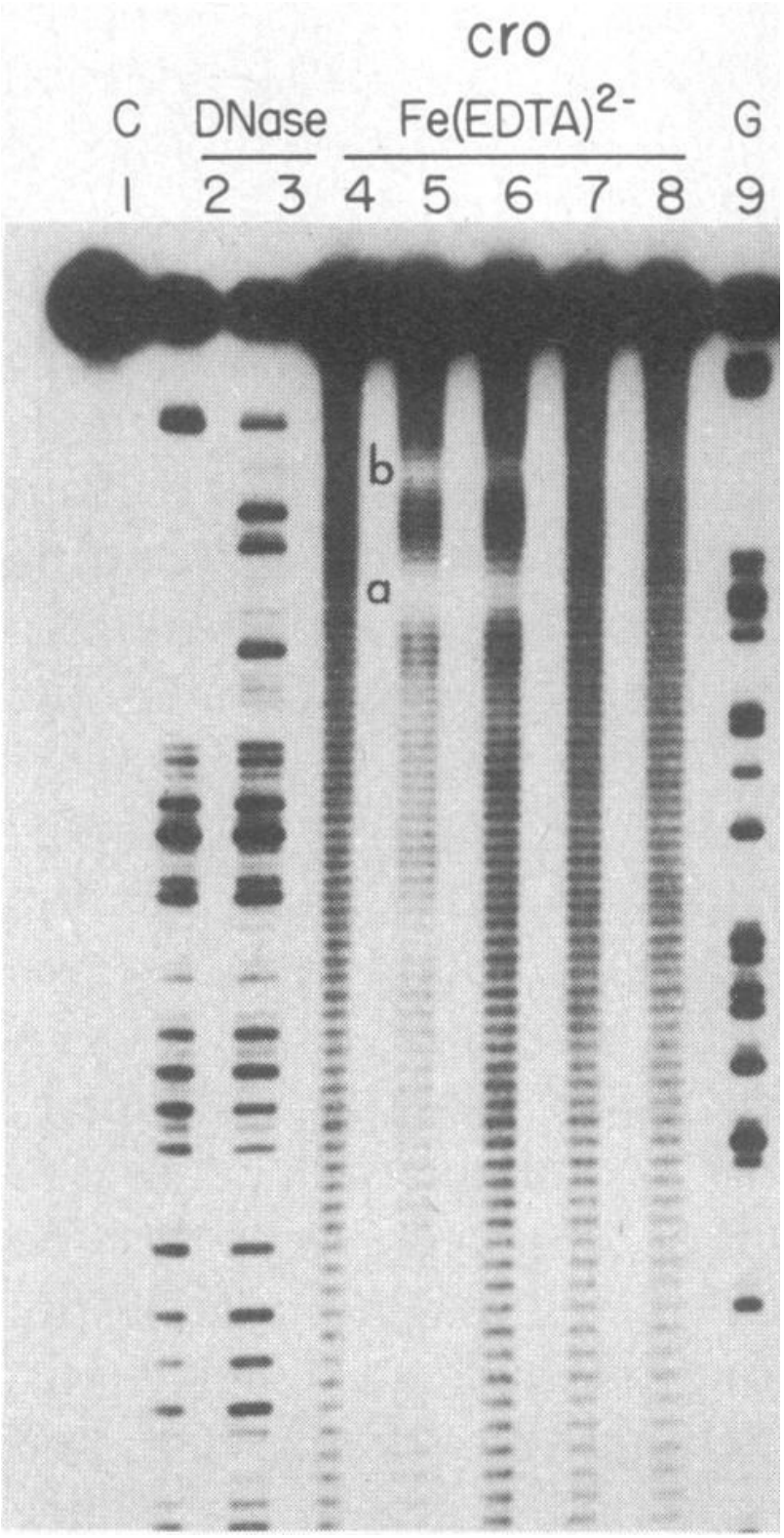


Figure 7 | Representative gel from a hydroxyl radical footprinting experiment. C – unfragmented control. 2 – DNA (120 bp) + Cro protein treated by DNAse footprinting, 3 – DNA treated by DNAse footprinting, no protein, 4 and 8 – DNA fragmented by Fenton chemistry, no protein bounded. 5-7 – DNA + Cro protein fragmented by Fenton chemistry at a different protein concentration (3.5μM, 350nm, 35nm), G – Maxam-Gilbert guanine specific sequencing reaction. Reprinted from ref. [108].

1.3.6.2 Radiolysis of water by a synchrotron

Synchrotron devices generate ~10^{14} to10^{15} photons per second with energies ranging from ~3 keV to 30 keV. Considering that analytes are exposed to pulses only for a couple of milliseconds, samples are not overdamaged. The reaction mechanism of water radiolysis is depicted in equations 11-15[109]:

$$H_2O + h\nu \longrightarrow H_2O^{\bullet+} + e^-_{dry} \quad \text{(eq. 11)}$$

$$H_2O^{\bullet+} + H_2O \longrightarrow OH^{\bullet} + H^+ \quad \text{(eq. 12)}$$

$$e^-_{dry} + H_2O \longrightarrow e^-_{aq} + H_2O^* \quad \text{(eq. 13)}$$

$$H_2O^* \longrightarrow OH^{\bullet} + H^{\bullet} \quad \text{(eq. 14)}$$

$$OH^{\bullet} + e^-_{aq} \longrightarrow OH^- \quad \text{(eq. 15)}$$

$$2\,OH^{\bullet} \longrightarrow H_2O_2 \quad \text{(eq. 16)}$$

In principle, the source of the radical is a water molecule. Water radiolysis generates a "dry electron" and a water radical-cation (equation 11) immediately reacting with another water molecule, producing a hydroxyl radical (equation 12). The radical then irreversibly modifies proteins on a surface. Then, the reaction of the dry electron results in a hydrated electron e^-_{aq} and excited water molecules (equation 13), which subsequently produce hydroxyl radicals and hydrogen atom radicals (equation 14). The resulting hydroxyl radicals oxidize solvent-accessible residues. Under anaerobic conditions, radicals from equations 15 and 16 self-quenched[99]. One advantage of X-ray is the use of water as a source of radicals in a surplus of ~55M over the analyte. Thus, this technique enables us to control the dose of radicals, simultaneously generated at a concentration ~1μM, and the time of reaction (~milliseconds) and, as such, is very powerful and useful even today[99].

The initial footprinting studies focused on a DNA or RNA damage/folding[109,110]. Over time, though, Mark Chance and others pioneered synchrotron beam use to map protein surfaces[111], interactions[112,113], conformational changes upon ligand binding[114–116], and dynamics[116,117].

1.3.6.3 Electron pulse radiolysis

Electron pulse radiolysis employs accelerated electrons with high energy at a megaelectronvolt scale, typically ~3 MeV, delivered in reproducible nanosecond pulses. The reaction mechanism is akin to that of X-ray synchrotron radiolysis. The source of radicals is the water surrounding protein analyte in the tube[99]. The first γ-ray footprinting was conducted in the 1980s to study nucleic acids damage [118,119], and later experiments proceeded with protein oxidation[120]. However, due to safety concerns associated with ionizing γ-rays, this approach is rather limited and other ways of generating radicals are preferred currently.

Despite efforts to use γ-rays, this technique was replaced by X-ray synchrotron footprinting and even in-situ Fenton chemistry. But because the lifetime of radicals cannot be controlled in Fenton chemistry and X-ray synchrotron availability remains limited, scientists have explored alternative, easier, and more readily approaches to generating hydroxyl radicals, such as laser UV photolysis of hydrogen peroxide.

1.3.6.4 Laser photolysis of hydrogen peroxide

Currently the most widely used technique to generate •OH radical is hydrogen peroxide photolysis. After exposing hydrogen peroxide to a UV light at a wavelength of ~250 nm (the highest absorption maximum of peroxide), peroxide is homolytically cleaved into two independent hydroxyl radicals, as shown in equations 17-19. This process also includes self-quenching step (equation 20)[99,121].

$$H_2O_2 + h\nu \rightarrow 2\,HO^{\bullet} \quad \text{(eq.17)}$$

$$2\,HO^{\bullet} + H_2O_2 \xrightarrow{h\nu} H_2O + HO_2^{\bullet} \quad \text{(eq. 18)}$$

$$HO_2^{\bullet} + H_2O_2 \xrightarrow{h\nu} H_2O + HO^{\bullet} + O_2 \quad \text{(eq. 19)}$$

$$2\,HO^{\bullet} \rightarrow H_2O_2 \quad \text{(eq. 20)}$$

Protein oxidation by UV flash photolysis was first reported in 2004, when Sharp et al. oxidized Lysozyme (Figure 8, top panel) and β-globulin with a UV lamp in the presence of 15% H_2O_2 for 5 minutes (Figure 8, middle panel)[122]. Prolonged irradiation for one hour (Figure 8, bottom panel) resulted in protein overoxidation,

extensive damage and structure rupturing, with extensive oxidation of almost all residues.

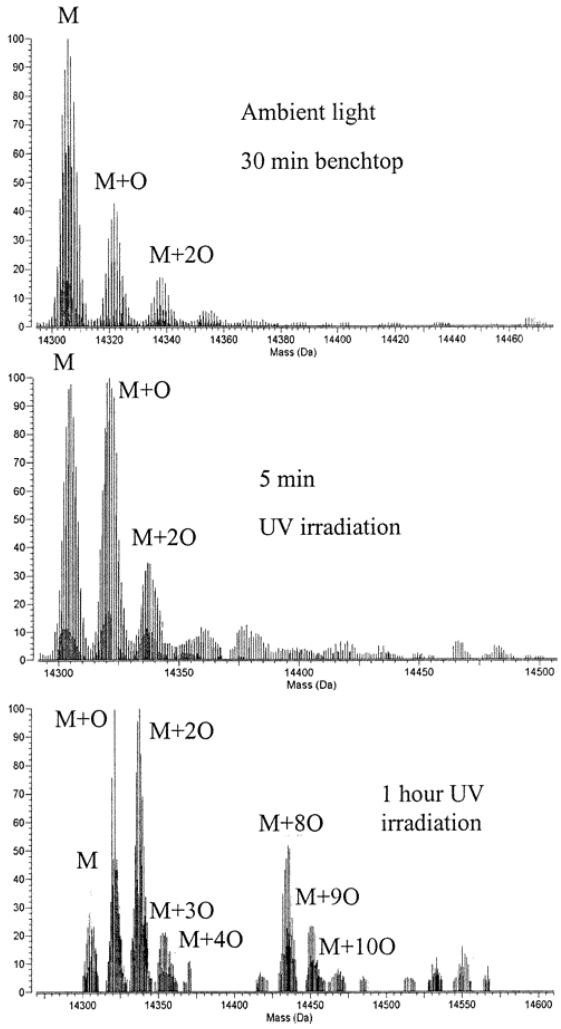


Figure 8 | MS spectra showing zoom on a charge state of lysozyme. The top spectrum represents unmodified lysozyme control (no UV light). Middle and bottom spectrum represents incubation for 5 minutes and 1 hour, respectively. Reprinted from ref [122].

To shorten the irradiation period, Aye, Low and Sze generated hydroxyl radicals using an excimer laser operated at 266 nm wavelength in their study published in 2005[123]. Excimer laser at a higher energy than a UV lamp can shorten the oxidation time and thus the duration of the experiment. The study focused on the oxidation of a small model protein, ubiquitin, in a quench flow capillary system, using a minimal concentration of H_2O_2 (~0.3 %), including one or more reproducible laser shots at an energy of 2 mJ/pulse (Figure 9). However, this system had a limited dose of energy since no scavenger was added, so the platform had to be tuned for each experiment/model protein.

Also published in 2005, a study by Hambly and Gross[124] reported similar results. The study described myoglobin oxidation in a quench-flow capillary system. At a lower concentration, 0.04%, hydrogen peroxide was dissociated under short, 17-ns laser pulses, with 50 mJ/pulse energy. But in contrast to Aye, Low and Sze, Hambly and Gross introduced a revolutionary quench-flow capillary system (Figure10A) with syringes, where protein and H_2O_2 were mixed in a T-mixer at a very short timescale. The mixture flowed through the transparent window (silica fused capillary with removed polyimide coating), where a UV laser shot initiated H_2O_2 dissociation into •OH radicals, ultimately oxidizing surface amino acids. A small amount of radical

scavenger (e.g., histidine or glutamine) was added to the analyte to introduce "one-laser shot irradiation", which affected the rate of oxidation. Moreover, behind the transparent window, the reaction was continuously quenched by immediate mixing with radical scavenger[125]. Samples were collected in an Eppendorf tube with catalase to decompose any remaining H_2O_2 that could overoxidized the analyte during proteolytic digestion.

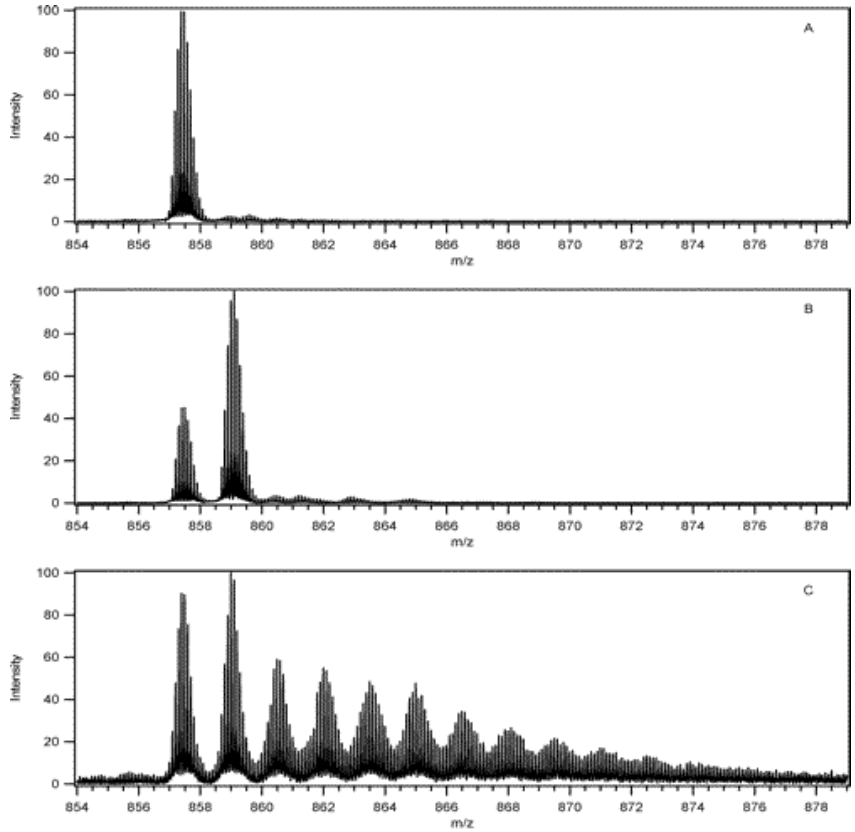


Figure 9 | ESI-MS spectra of ubiquitin 10+ charge state. A – unmodified control, B – one laser shot, C – 100 laser shots. Adapted and modified from ref[123].

The merits of this platform included (i) ensuring a short period of contact between the analyte and hydrogen peroxide, avoiding artificial overoxidation (Figure 10B), (ii) mixing the analyte with a radical scavenger in a syringe to perform "real" single-hit kinetics to oxidize analyte surface, (iii) using a very low concentration of peroxide during the reaction, albeit enough to successfully oxidize protein surface, (iv) and a T-mixer where reaction mixture was immediately quenched, with catalase in the collecting tube, and (v) enabling a detailed proteomic analysis of oxidized samples by MS. This experimental platform was a milestone in protein radical labelling and is referred to as ***Fast Photochemical Oxidation of Proteins (FPOP)***[124].

Table 1 | Rate constants of reactions of individual amino acids with hydroxyl radicals during FPOP. All possible modifications are denoted alongside mass shifts. Adapted from ref. [126].

Residue	Rate constant ($M^{-1} s^{-1}$)	Mass shifts (Da)				
Cys	3.5×10^{10}	−15.9772	+31.9898	+47.9847		
Trp	1.3×10^{10}	+3.9949	+15.9949	+31.9898	+47.9847	
Tyr	1.3×10^{10}	+15.9949	+31.9898	+47.9847		
Met	8.5×10^{9}	−32.0085	+15.9949	+31.9898		
Phe	6.9×10^{9}	+15.9949	+31.9898	+47.9847		
His	4.8×10^{9}	−23.0160	−22.0320	−10.0320	+4.9789	+15.9949
Arg	3.5×10^{9}	−43.0534	+13.9793	+15.9949		
Cystine	2.1×10^{9}	+31.9898	+47.9847			
Ile	1.8×10^{9}	+13.9793	+15.9949			
Leu	1.7×10^{9}	+13.9793	+15.9949			
Val	8.5×10^{8}	+13.9793	+15.9949			
Pro	6.5×10^{8}	+13.9793	+15.9949			
Gln	5.4×10^{8}	+13.9793	+15.9949			
Thr	5.1×10^{8}	−2.0157	+15.9949			
Lys	3.5×10^{8}	+13.9793	+15.9949			
Ser	3.2×10^{8}	−2.0157	+15.9949			
Glu	2.3×10^{8}	−30.0106	+13.9793	+15.9949		
Ala	7.7×10^{7}	+15.9949				
Asp	7.5×10^{7}	−30.0106	+15.9949			
Asn	4.9×10^{7}	+15.9949				
Gly	1.7×10^{7}	—				

Theoretically, hydroxyl radicals promote the modification of all amino acids (Table 1). Possible modifications were thoroughly studied by Xu, Takamoto and Chance[127–129]. As expected, the results showed that the most reactive residues are sulfur-containing residues, cysteine and methionine, followed by aromatic residues, namely tyrosine, tryptophane, phenylalanine and histidine. These studies also described various modification channels in which amino acids could slide after a hydroxyl radical attack. For instance, the most common modifications are arginine deguanidylation and several histidine modifications accompanied by ring opening. Nevertheless, the most common modification remains the addition of one or two oxygens (+15.9949 Da, +31.9898 Da).

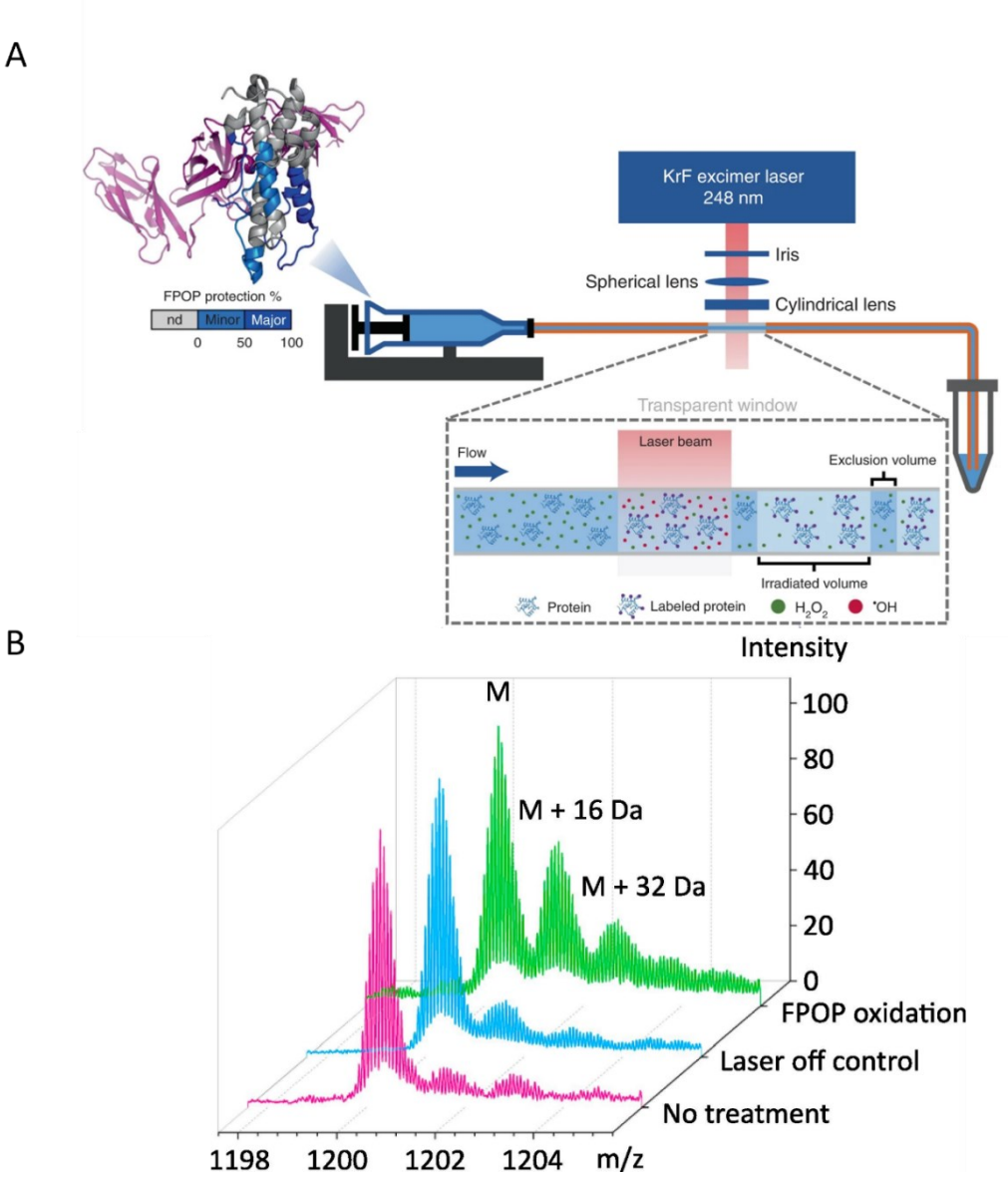


Figure 10 | A. A capillary quench flow apparatus for FPOP. B. MS spectra of a myoglobin charge state displaying an unmodified control (pink spectrum), a sample with peroxide but with a disabled laser (blue spectrum) and oxidized myoglobin after FPOP (green spectrum). Adapted and modified from ref.[130,131].

Several other methods can also generate hydroxyl radicals, including high voltage electric discharge[132], fast neutrons[133], peroxynitrous acid decomposition[134], electron beam[120] and ozonolysis[135]. These methods were mostly used in the last century to study nucleic acid, but not commonly for protein footprinting.

1.3.4.5 Other radical footprinting methods

Although the hydroxyl radical is a convenient labeling agent, its precursor, hydrogen peroxide, is highly reactive. Biomolecules can be oxidized by peroxide on a benchtop, without laser irradiation (Figure 10B). Moreover, the combination of

products (Table 1) complicates the analysis of samples. For this reason, considerable research efforts were made over the years to develop new reagents with different reactivities to residues (Figure 11). These reagents will be briefly described in this section.

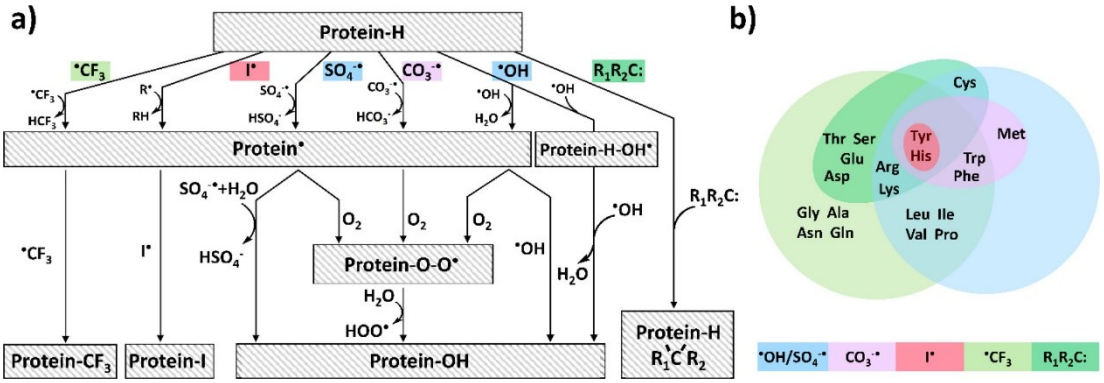


Figure 11 | A summary of radical-based footprinting techniques, B. residue specificity. Reprinted from ref[15].

Sulfate radical anion labelling (SO_4^-•). The sulfate radical anion is highly reactive, with a lower reductional potential (2430 mV) than the hydroxyl radical (1900 mV). As such, SO_4^-• can modify all types of biomolecules, including lipids, carbohydrates, proteins, and nucleic acids. However, its mechanism of attack remains unclear. This anion is created via a persulfate precursor ($Na_2S_2O_8^{2-}$) upon irradiation by a 248nm excimer laser in a quench-flow capillary system, identical to FPOP. The reactivity was determined and slightly differs from FPOP: M>Y=W>F=E=H>S>P> D=T>K=Q>L=V=I. Similarly to FPOP, the reactivity towards residues is broad, and the conditions in the FPOP platform are also tunable. Moreover, its footprinting capabilities were demonstrated on myoglobin, calmodulin and β-lactoglobin[136].

Carbonate radical anion (CO_3^-•) footprinting. Carbonate radial anion is found in living cells and primarily formed from bicarbonate-carbon dioxide buffer at physiological pH. However, superoxide dismutase, hemoprotein-Fe(III) and xanthine oxidase, among other enzymes, can decompose and neutralize this radical anion[15]. The study by Zhang et al.[137] showed the ability of FPOP platform to generate CO_3^-• radicals using a 248nm excimer laser in a carbonate/bicarbonate buffer. When present (created from hydrogen peroxide) in solution, the hydroxyl radical reacts

with CO_3^{2-}/HCO_3^- buffer, yielding $CO_3^{-}\bullet$. The hydroxyl radical targets only electron-rich residues in the following order: Met~Trp>Tyr>His~Phe. The ability to perform these modifications, including the reactivity towards residues was studied on myoglobin, ubiquitin and several small peptides[137].

Carbene radical footprinting. Carbene radical footprinting was first described by Richards and colleagues in 2000(ref.[138]). They generated methylene radical from diazirine gas (CH_2N_2) upon UV light irradiation. However, this methylene radical yielded low footprinting efficiency. In 2011, Schriemer and colleagues published study of carbene footprinting, where they used reactive photoleucine as a precursor to generate carbene radicals[139]. Footprinted samples were analyzed using a bottom-up MS approach, identifying the typical +115.03 Da shift of modified peptides in a tryptic digest.

In 2016, Gross and colleagues published a study[140] on footprinting with reactive carbene radicals performed in a FPOP platform (Figure 12). Careful manipulation of carbene radicals ensured single hit-kinetics and restricted protein overmodifications. The laser wavelength of the platform was slightly adjusted to 355 nm. After calmodulin irradiation, LC-MS/MS analysis revealed that Asp and Glu were the most reactive residues, followed by Arg, Tyr, Thr and Ser[140,141].

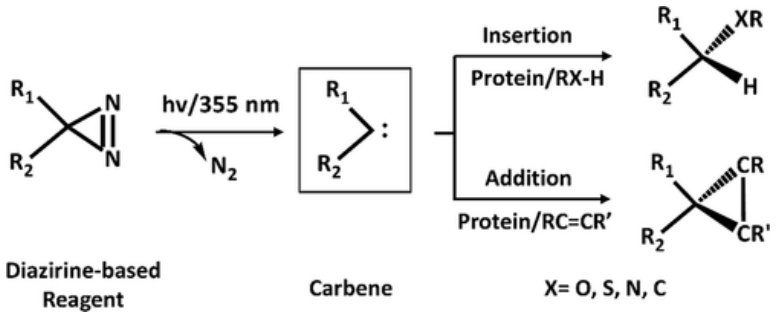


Figure 12 | Proposed mechanism for generating carbenes with a 35nm excimer laser. Reprinted from ref. [15].

Footprinting by iodine radicals (I•). Iodine radicals effectively footprint tyrosine and histidine residues. Tyrosine iodination is a common physiological reaction as thyroid peroxidase iodinates thyroid hormones[142]. The first attempts at studying electrophilic iodination of proteins were published in 2008. Ly and Julian used covalent carbon–iodine bond photodissociation in iodinated proteins at a wavelength of 266 nm to

induce residue-specific radical protein cleavage in a linear ion trap. Nevertheless, footprinting was conducted by the Gross group in 2012. These researchers implemented a FPOP platform to selectively modify proteins (Figure 13) using iodobenzoic acid as a source of I• radicals. In this study, myoglobin, carbonic anhydrase II, and human insulin were tested to show that iodination can be used in protein footprinting[143].

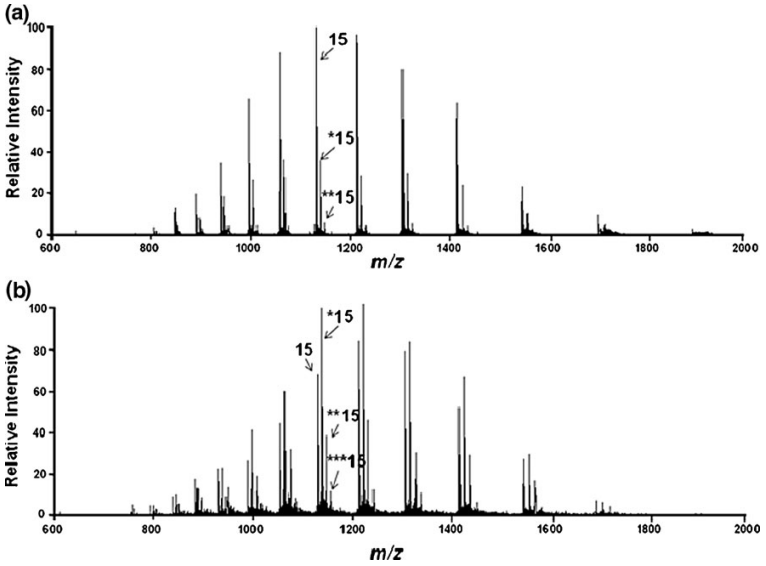


Figure 13 | ESI-MS spectra of mono-, di- and tr-iodinated holomyoglobin (a) and apomyoglobin (b). Reprinted from ref.[143].

Trifluoromethyl radical footprinting (CF_3•). Although oxidation in a FPOP platform is a powerful approach, trifluoromethylation of molecules is a commonly used technique. Fluorination is applied to produce agrochemicals and pharmaceutical drugs [144]. And while fluorine is nearly absent in nature and difficult to incorporate into some molecules because of its high electronegativity, recent advancements have shown that trifluoromethylradical footprinting is achievable.

In 2017, Zhang et al. preformed the first laser-initiated trifluoromethylation of biomolecules using a FPOP platform[145]. Combining H_2O_2 with $NaSO_2CF_3$ generated CF_3•, inducing protein footprints. By analyzing three proteins, namely myoglobin, β-lactoglobulin and vitamin K epoxide reductase (VKOR), they observed that 18 of 20 naturally occurring amino acids were modified, demonstrating the usefulness of this approach for mapping protein surfaces and protein-ligand interactions. More recently, Fojtik et al. showed that fluoroalkyl radicals, including trifluoromethyl

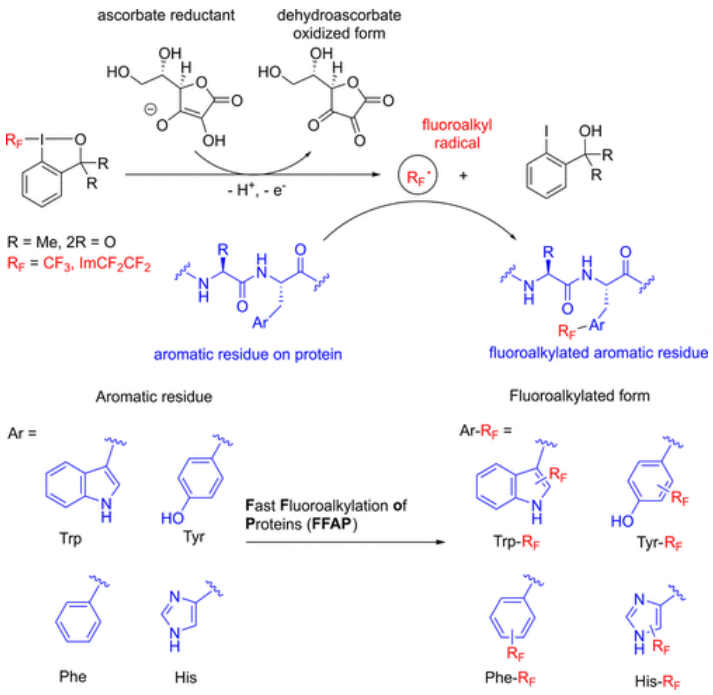


Figure 14 | Reaction scheme of fast fluoroalkylation of proteins (FFAP). Reaction of Togni reagent(s) with ascorbic acids results in fluoroalkyl radical (colored in red) generation, which subsequently modify protein residues. Reprinted from ref. [146].

radical, generated from molecules referred to as "Togni reagents", may also serve as labelling agents for biomolecules (Figure14)[146]. The reaction was performed in a quench-flow capillary setup, wherein the analyte (myoglobin, hemoglobin-haptoglobin complex and tryptophan repressor-binding protein A, WrbA) and the Togni reagent were mixed with ascorbic acid to generate fluoroalkyl radicals and footprint the sample. Subsequently, the targeted residues were identified by bottom-up MS analysis as Trp, His, Tyr and Phe and Cys, demonstrating that fluoroalkyl chemistry is useful for mapping protein surfaces and monitoring structural changes upon ligand binding.

2 Aim of the study

The primary goal of this thesis is to contribute to the ongoing development of Fast Photochemical Oxidation of Proteins (FPOP) coupled to high-resolution MS analysis for studying protein-DNA interactions.

The specific aims of this research were to:

1. Optimize FPOP to study protein-DNA complexes.
2. Optimize a top-down approach for the analysis of FPOP samples.
3. Assess the potential of FPOP for the structural characterization of nucleic acids.
4. Test singlet oxygen ($^1O_2$) as a potential labelling probe.

3 Methods

The attached publications and manuscripts include detailed descriptions of all experimental methods and procedures necessary for experimental reproduction. Thus, this chapter includes only a list of used methods.

List of experimental methods used in this research:

- Protein expression and purification in a nutrient-rich media,
- Protein expression and purification in a nutrient-minimal media,
- Electromobility shift assay (EMSA) and SDS-PAGE,
- Fast photochemical oxidation of proteins (FPOP),
- Native mass spectrometry,
- LC-MS (timsTOF Pro, FT-ICR MS) analysis oxidized peptides and nucleic acids,
- Top-down analysis of oxidized proteins (FT-ICR MS), including collision-induced dissociation (CID), electron-capture dissociation (ECD), electron-transfer dissociation (ETD), ultraviolet photo dissociation (UVPD)

4 Results and discussion

The aim of this thesis was to apply FPOP oxidative labelling technique to the study of protein-DNA complexes. Transcription factors are known to bind double-stranded DNA (dsDNA) by recognizing and interacting with a specific genomic sequence, thereby modulating gene expression through activation or repression. Therefore, a deeper understanding of interactions between transcription factors and cognate response elements may provide us with key insights into underlying mechanisms regulating their structure, dynamics and function[1].

Despite the intriguing nature of interaction between transcription factors and DNA, obtaining high-resolution structural information is often challenging due to the dynamic nature of its interaction. For instance, conventional methods like X-ray crystallography may fall short in capturing such dynamics. By contrast, structural MS may offer valuable information on protein complexes as their microsecond/millisecond labelling timescale aligns with the timescale of biomolecular interactions[125]. Moreover, this technique focuses on labelling solvent-accessible amino acids, especially covalent labelling techniques such as FPOP.

The FPOP platform has long stood out for its ability to examine protein complexes for nearly two decades[124]. During these period, numerous studies have been published on the structural characterization of proteins and their complexes [15,124,147–154].

In turn, DNA damage originally mapped protein-DNA interactions using hydroxyl radicals generated by Fenton chemistry[108]. However, to our knowledge, no studies have applied FPOP to investigate protein complexes with DNA. Thus, this thesis aimed at leveraging the potential of the FPOP platform to characterize transcription factor-dsDNA complexes, coupled to mass spectrometry (MS) data analysis, to extract comprehensive information regarding their interaction.

In **publication I,** I specifically explored the fundamental aspects of protein-DNA complex oxidation, demonstrating the ability of the FPOP platform to analyze such complexes from protein site. By combining a bottom-up approach with a top-down approach, I surveyed structural rearrangements on protein conformation induced by ligand binding. The limitations of the top-down approach prompted me to further investigate this technique, leading to the concept presented in the **publication II**.

In **publication II**, I identified the best possible way to analyze FPOP-labeled proteins using the top-down approach. Combining an idea of protein isotope depletion with testing gas-phase fragmentation techniques provided information regarding oxidation of most of residues involved directly or indirectly in protein-DNA binding. After addressing protein oxidation by hydroxyl radicals throughout this study, one question remained unanswered – what happens to the DNA? Initial studies on protein-DNA complexes demonstrated that DNA is damaged[155], but combining FPOP of DNA with MS would enable us to tackle this problem on a whole new level, benefitting from a much higher resolution than gel electrophoresis could ever provide. This concept served as the central theme of **publication III**.

In **publication III**, I elucidated the fundamentals of DNA fragmentation and data processing and the outcomes of FPOP-initiated DNA damage experiments. Currently under review, this study first demonstrated 'hydroxyl radical footprinting of DNA' coupled to high-resolution MS analysis.

Publications I and II also reports the development of in-house software designed to assign unmodified/oxidized fragment ions and to calculate the extent of oxidation of fragment ions following the implementation of a top-down approach to analyze oxidized samples by FPOP. In the last publication (**publication IV),** I assessed whether a top-down technology could be applied to analyze proteins oxidatively damaged by singlet oxygen ($^1O_2$).

4.1 Publication I

Title: Utilization of Fast Photochemical Oxidation of Proteins and Both Bottom-up and Top-down Mass Spectrometry for Structural Characterization of a Transcription Factor–dsDNA Complex.

The primary objective of **Publication I** was to demonstrate the utility of FPOP in studying protein-DNA interactions. For the initial experiment, I selected the DNA-binding domain of FOXO4 transcription factor with a short (13-bp) dsDNA sequence known as DAF16[156]. This model system was selected mainly for two reasons: (i) the FOXO4-DBD-DAF16 complex has a known crystal structure [157], and (ii) its interaction with DNA has been well-studied and documented in details[158–161]. Following FOXO4-DBD protein expression and purification, this protein was mixed

with DAF16 to form the desire complex. EMSA (Figure 15A) and native nESI in ammonium acetate buffer (Figure 15B) confirmed the formation of a complex between this transcription factor and its cognate DAF16 element.

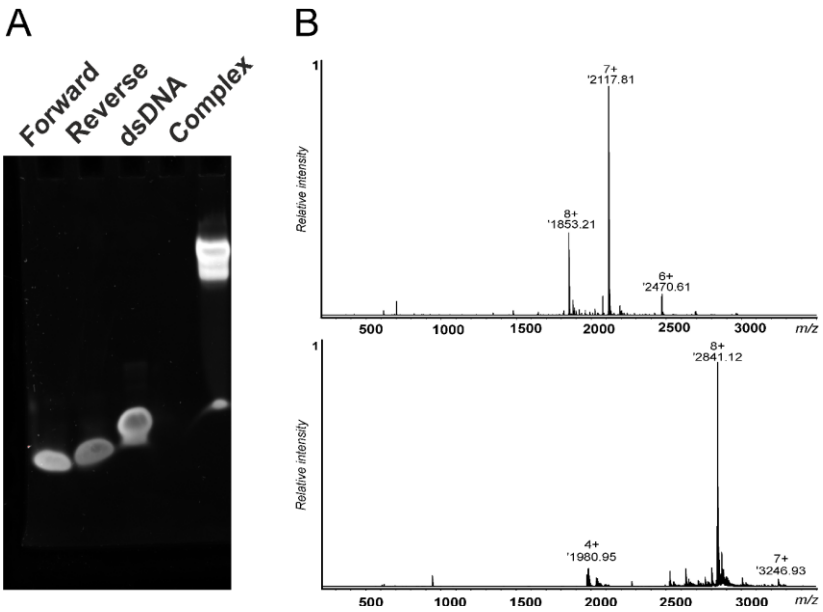


Figure 15 | A. EMSA of FOXO4-DAF16 complex, B. Native nESI of FOXO4-DBD in 150mM ammonium acetate (top panel) and its complex with DAF16 (bottom panel).

Figure 16 shows the zoomed +16 charge state of FOXO4-DBD during various phases of the experiment. Samples were desalted on a protein micro-trap prior to the spraying via nESI source into 15T-FT-ICR mass spectrometer. The top black spectrum in Figure 16 represents the unoxidized control spectrum of FOXO4-DBD. It is evident that the protein was not artificially oxidized at the beginning of the experiment, an artifact that can occur due to oxidation during protein expression and/or purification (referred to as 'background oxidation').

In the middle spectrum, FOXO4-DBD was continuously mixed with H_2O_2 in a FPOP apparatus, but the dissociation of H_2O_2 was not triggered by the UV laser. The collected samples were desalted and sprayed. Consequently, this middle spectrum reveals a level of background oxidation of protein samples by H_2O_2 and singly and doubly oxidized proteoforms. These findings demonstrate that some proteins are particularly sensitive to H_2O_2.

Subsequently, FOXO4-DBD and FOXO4-DAF16 complex were subjected to FPOP oxidation. During the FPOP experiment, UV laser triggered the dissociation of H_2O_2, thereby oxidizing amino acid side chains of a protein. Figure 16, bottom, shows oxidized FOXO4-DBD without and with DAF16 as purple and orange spectra,

respectively. When DAF16 is present in the sample, oxidized proteoforms are less abundant because the amino acids side chains involved either directly or indirectly in protein-DNA interactions are protected (compare the purple and orange spectra in Figure 16.

Bottom-up analysis of the samples started with protein digestion with Trypsin and LysC. Modified residues were identified by LC-MS/MS, and their extent of oxidation was quantified from the LC-MS trace. Overall, 19 residues were identified as modified by FPOP, and the presence of DAF16 affected solvent accessibility of residues upon binding (Figure 17). Based on dataset, I concluded that residues W97, Y102, H152 and W174 were less oxidized/modified. By comparing this dataset with an available crystal structure[157] and with mutagenesis studies[161], I confirmed that these residues directly interact with DAF16 and play a key role in DNA binding. In contrast, residues Y124, W126, M127, Y133 and M175, which do not interact with DAF16, were modified to varying extents. These results highlight the ability of FPOP to monitor changes in residues /regions whether they are involved or not in DNA-binding. Aromatic residues W97, Y102 and W173, denoted by '#', refer to oxidized positional isomers separated

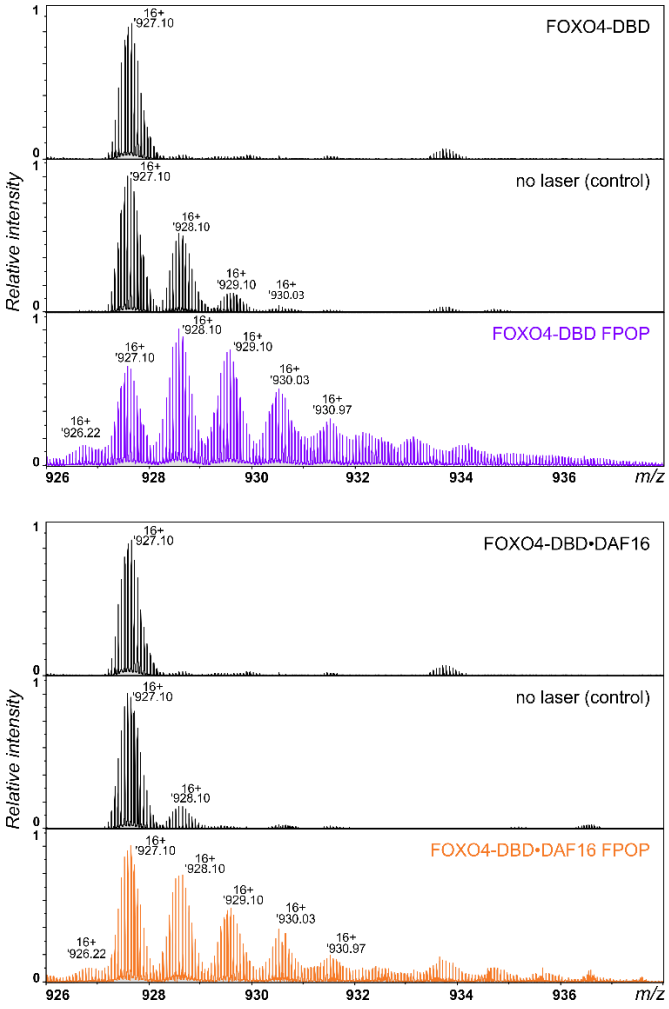


Figure 16 | Intact MS spectra of FOXO4 +16 charge state showing protein control (top spectrum), no laser control (middle spectrum) and oxidized protein with (orange spectrum) and without (purple spectrum) daf16 .

during LC, a phenomenon known as “sub amino acid resolution”[150].

When using a bottom-up approach, I also detected modifications other than +16 Da and +32 Da, such as arginine R94 and R151 deguanidylation (-43.05 Da), aspartate E179 decarboxylation (-30 Da), and His H152 and H157 conversion into Asp (-22.03 Da). Thus, a bottom-up approach provides useful information for mapping protein-DNA interaction interfaces.

.

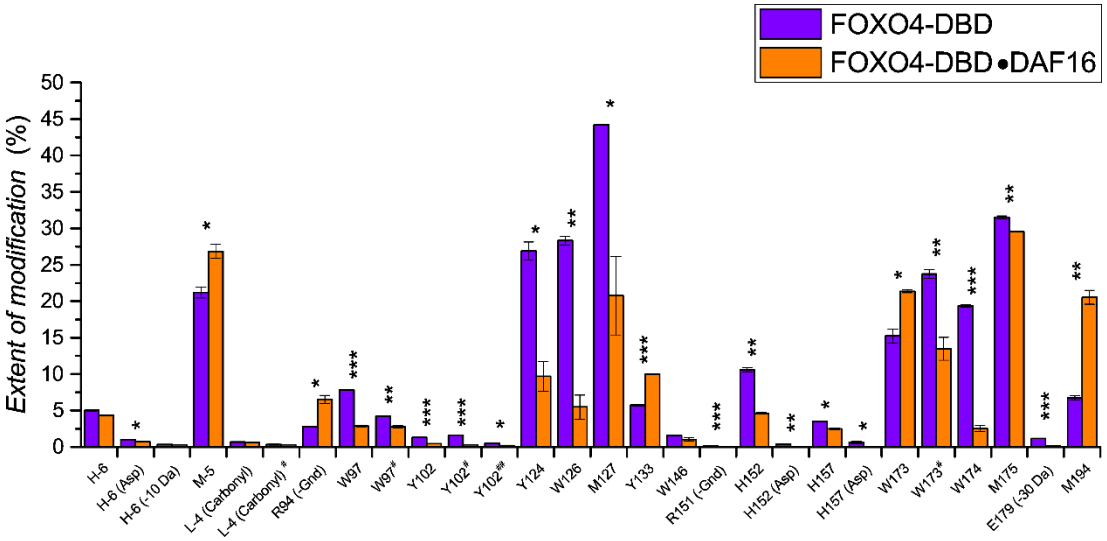


Figure 17 | Bottom-up analysis of FPOP oxidized sample. Purple histograms show the extent of modifications of oxidized residues of FOXO4-DBD without DAF16 in solution, whilst orange histograms represent the extent of modifications of FOFOX4 with DAF16.

FPOP coupled to LC-MS/MS analysis provided valuable insights into this protein-DNA interaction interface. However, a bottom-up approach enables us to analyze the entire bulk of oxidized proteoforms in a single experiment. Nevertheless, “overmodified" proteins could introduce artifacts by altering their initial structure in response to artificial modifications, as described by Sharp et al. in 2003[107]. These authors observed oxidized Met131 in their bottom-up analysis of oxidized myoglobin, albeit deriving the solvent-accessible surface area (SASA) from an X-ray structural model, at 0.00 Å^2. To overcome these invariances, I extended and optimized a top-down workflow to analyze and assign oxidized residues from a singly oxidized protein population in **publication I**. Optimizing the workflow during sample preparation exacerbated the excess of DNA present in the sample. The recommended workflow for protein sample desalting on a protein microtrap entails

protein acidification (with formic or acetic acid), which improves sample denaturation, thus enhancing trapping on a reverse-phase during desalting procedure. But when acidifying the solution, the pH drops below ~5.2, and the DNA protonates. Protonated DNA becomes hydrophobic and starts precipitating, leading to protein-DNA co-precipitation in solution. As a result, the initial sample preparation experiments led to low MS signal intensity and precluded the accumulation of oxidized protein precursors prior MS/MS fragmentation. To overcome these challenges, I optimized the workflow for protein sample preparation, which included: (i) denaturation of the FOXO4-DAF16 complex with 4M urea, a concentration deemed sufficient for protein denaturation[158], (ii) addition of Benzonase® endonuclease and $MgCl_2$ to the sample, whose activity is even enhanced in their presence and should digest residual DNA, (iii) protein desalting on a protein microtrap with LC-MS water and ammonium acetate at pH 6.8, (iv) elution of the protein sample into pure acetonitrile, and (v) spraying with an nESI source to save the protein sample.

The singly oxidized proteoforms of two charge states, namely +16 (Figure 16) and +18, with (holo form, orange spectra) and without (apo form, purple spectra) DAF16 were separately isolated in a quadrupole. Subsequently, proteoform ions were fragmented in the collision or ICR cells by CID or ECD, respectively. Unmodified fragment ions and their singly oxidized forms were identified in MS/MS spectra, and the extent of oxidation was calculated as the ratio between unmodified and oxidized fragment ion intensity. Based on my dataset, I subsequently developed and tested an in-house software (written by Dr. Daniel Kavan) for automatic fragment ion assignment. This software also assigns oxidized fragment ions with high mass accuracy and calculates the extent of oxidation for all observed fragment ions.

First, I calculated the extent of oxidation for observed fragment ions from both fragmented charge states, +16 and +18, and for both samples without and with DAF16. Then, I subtracted the extend of oxidation of the apo forms from the extent of the holo form, for instance, Δ-Extent($c28_{apo}$–$c28_{holo}$). Subsequently, I subtracted this value from the two vicinal fragment ions to assess the protection or deprotection effects between two adjacent fragment ions, for instance, Δ-Extent(*Δc28–Δc37*). This way, I was able to graphically deduced regions with changes in their solvent accessibility induced by DNA binding, as shown in **Figure 3 in Publication I**.

Despite the theoretical possibility of 266 fragment ions, only 30 CID and 60 ECD ions were intense enough for quantification. Despite failing to achieve the single-residue resolution of the bottom-up approach, the top-down approach highlighted regions affected by DAF16 binding. Figure 18 summarizes structural information in a colored X-ray structural model of FOXO4-DBD with DAF16 (PDB entry: 3L2C) based on the obtained FPOP dataset. Purple residues/regions were more modified in the apo form, whereas orange residues/regions were more modified in the holo form. Residues detected in the bottom-up approach are visualized using stick representation. Figure 18 illustrates that most residues/regions were protected upon DAF16 binding. However, some regions, containing residues Y133 and R94, were more oxidized/modified. This detailed analysis provided insights into regions either protected or more susceptible to modification in the presence of DAF16, thus fostering a more nuanced understanding of structural changes on protein induced by DNA binding.

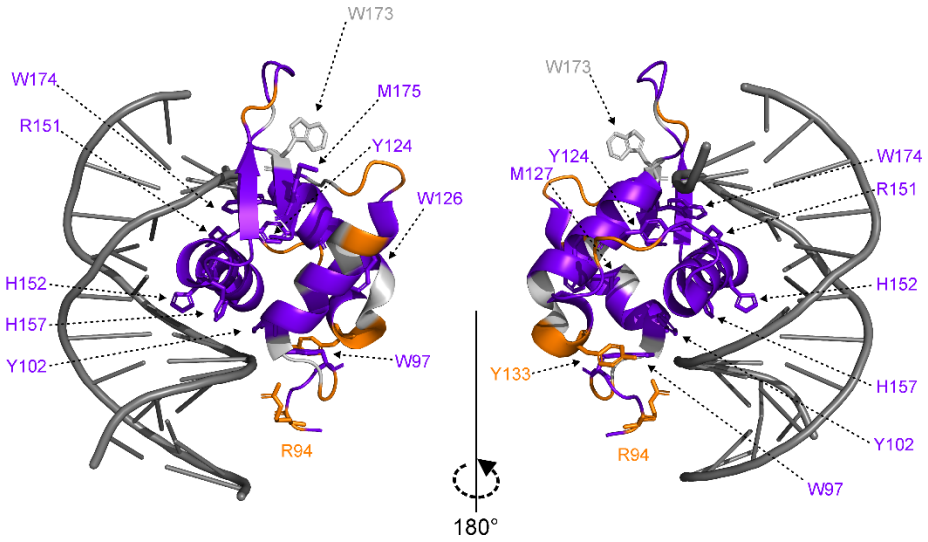


Figure 18 | X-ray structural model of FOXO4-DBD·DAF16 (PDB entry: 3L2C) [157] *with the highlighting oxidized regions detected using the top-down approach. The individual residues detected using the bottom-up approach were highlighted in these regions using a stick representation. Purple: regions/residues more modified in the apo form, orange: regions/residues more modified in the holo form.*

In **publication I**, we demonstrated that the FPOP platform can be used to oxidize protein-DNA complexes to map their interaction interface and that the top-down approach enables us to assign oxidized residues from singly oxidized proteoform and thus map protein-DNA binding interfaces. Additionally, the results from both approaches were compared with an available X-ray structural model (PDB entry:

3L2C)[157], with the results from structural and mutagenesis studies[160,161], and from MS studies, including HDX, quantitative protein crosslinking, and protein-DNA crosslinking. This comprehensive comparison revealed a high level of complementarity among all techniques, confirming that FPOP is a valuable tool for mapping protein-DNA interactions.

4.2 Publication II

Title: Isotopic Depletion Increases the Spatial Resolution of FPOP Top-Down Mass Spectrometry Analysis.

In **publication I**, we demonstrated the possibility of a top-down approach to analyze FPOP-oxidized protein-DNA samples to study its interaction interface. The top-down approach has been a benchmark technique for nearly 30 years[162]. However, with the increasing mass of analytes, the MS/MS spectra become more intricate as the number of possible fragment ions subsequently increases, including the number of isotopes in ion envelopes. Moreover, bigger the fragment, the more times will charge, which causes isotope overlap. It is known when peptides and fragments exceed ~1kDa, the monoisotopic peak ceases to be the most abundant one. Moreover, proteins and fragment ions with the mass ~10 kDa do not produce observable monoisotopic peak in MS/MS spectra[163]. These intricates MS/MS spectra, data deconvolution, leads to decrease of signal-to-noise ratio (SNR) and thus the overall quality of MS/MS spectra.

To increase the number of useful fragment ions in top-down analysis, to enhance the number of assigned oxidized residues and to acquire more structural information, I aimed at applying 'protein isotope depletion', developed over 30 years ago[164]. This well-established technique involves incubating bacteria in minimal medium containing lighter isotopes of elements. Bacteria then incorporate lighter isotopes into amino acids during metabolism and into proteins during proteosynthesis.

Consequently, in **publication II**, I explored the concept of protein depletion and its advantages for mass spectrometry detection of covalently labelled protein. To achieve this goal, I expressed and purified FOXO4-DBD in both isotopically natural (IN) and isotopically depleted (ID) media. Following the purification step, I prepared a protein-DNA complex by mixing the FOXO4-DBD construct with a 17bp dsDNA

segment termed Insulin Response Element (IRE). I subsequently oxidized proteins with and without IRE and desalted and sprayed the samples into 15T-FT-ICR mass spectrometer, as shown in Figure 19.

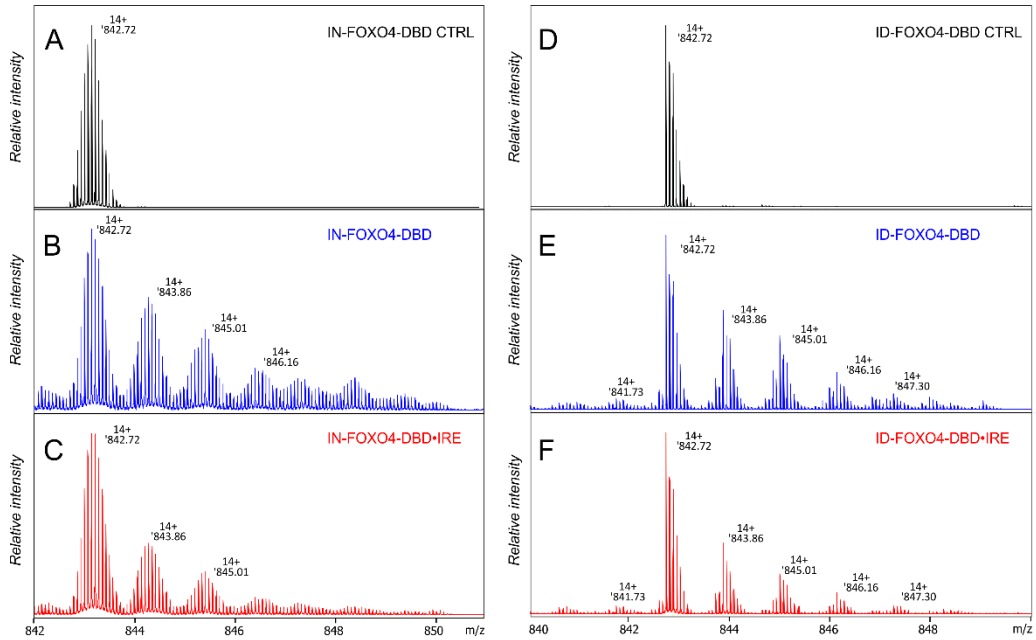


Figure 19 | ESI-MS spectrum of a zoomed-in +14 charge state showing an isotopically natural (A) and depleted (B) protein control. FPOP of an isotopically natural protein without IRE (B) and with IRE (C). FPOP of isotopically depleted protein without IRE (D) and with IRE (E).

Figure 19 shows a zoom-in +14-charge state of both an isotopically natural (A) and depleted (B) protein as an unmodified control, indicating no artificial background oxidation, which could have occurred during expression or purification. When oxidized in the FPOP platform, oxidized proteoforms appeared in both IN and ID versions, without IRE (B, E) and with IRE (C, F). Figure 19 also shows that oxidized proteoforms are less abundant when IRE is presented in the sample.

In a study by our group[147], CID, ECD and ETD was tested on a singly-oxidized ubiquitin model as a fragmentation technique to optimize sequence coverage. The results showed that CID was accompanied by neutral loss of water, thus biasing the data by lowering oxidation level. This effect was even stronger on ubiquitin fragmentation using IRMPD (Petr Novák – personal communication). ETD provides a higher yield fragments but suffers from poor sequence coverage. This study demonstrated that the best fragmentation technique is ECD, albeit requiring collecting MS/MS spectra when accumulating precursors of several charge states

(multiCASI mode in Burker's instruments). This (multiCASI) approach seemed to be the most reasonable option to address the issue at hand as each charge state generates different types of fragment ions and varying levels of oxidation[46,165,166].

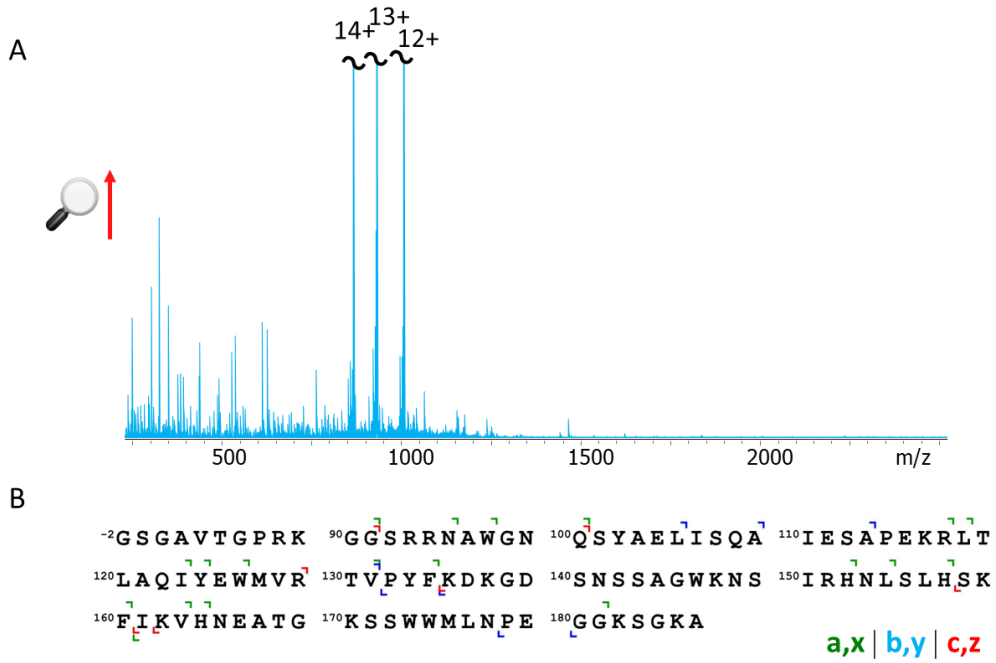


Figure 20 | A. UVPD MS/MS spectra of singly oxidized depleted FOXO4 with a zoomed-in y-axis. B. Sequence of FOXO4-DBD according to wild type FOXO4 with quantified fragment ions.

To demonstrate the advantage of protein depletion, singly oxidized proteoforms of both IN- and ID-FOXO4 were isolated in quadrupole, fragmented by fragmentation techniques, and compared with each other. To maximize sequence coverage, I accumulated three charge states prior fragmentation, as reported by Yassaghi et al.[147]. For FOXO4, singly oxidized proteoforms of three charge states, namely +14, +13 and +12, were simultaneously isolated and fragmented. I then tested three dissociation techniques for maximal sequence coverage and, thus, maximal structural information.

The first tested technique was ultraviolet photo dissociation (UVPD). UVPD was the first fragmentation technique tested in this study because UVPD is an excellent tool for sequencing proteins and oligonucleotides[45,46,49]. Furthermore, I performed UVPD experiments with the help of Dr. Alan Kádek, and our ICR cell is equipped with a BaF_2 transparent window in the back of the cell, making it possible to irradiate ions by outer photon source. Using the 193nm ArF excimer laser, aligned to the center of "shimmed" ICR cell[167], I accumulated oxidized proteoforms and

transferred them into the ICR cell. Ions were accumulated for 0.5s and irradiated by a single-shot pulse with 1.50 mJ/pulse energy. Overall, 256 scans were acquired. Figure 20A shows a raw UVPD MS/MS spectrum of an oxidized depleted protein. The y-axis is zoomed-in for a closer look at product fragment ions at a lower *m/z*. Subsequently, I processed and analyzed spectra and searched for *a*-, *b*-, *c*-, *x*-, *y*- and *z*-fragment ions, including their singly oxidized forms, in MS/MS spectra. However, only 29 fragment ions were intense enough for quantification (Figure 20B), corresponding to a sequence coverage of only ~10 %. Therefore, UVPD coverage was not broad enough to analyze FPOP-oxidized samples.

To solve the problem of limited sequence coverage, I tested *electron-transfer dissociation (ETD)*. Although Yassaghi et al.[147] demonstrated that ETD suffers from poor sequence coverage, they applied ETD to a one-case of isotopically natural model protein (ubiquitin). Conversely, Srikanth et al. demonstrated that oxidation can be tracked by ETD and that the extent of oxidation can be determined on histidine residues[168]. At the same nominal values, I isolated singly oxidized protein precursors, which were accumulated in quadrupole and fragmented by ETD. Using fluoranthene as an ETD reagent, which was ionized by an nCI source to generate radical anion at *m/z* 202, the reaction time was 15 ms after 25 ms reagent accumulation. Prior to ETD, ions were accumulated for 1.0 s, and spectra were acquired by collecting 256 MS/MS scans. Figure 21A shows an ETD MS/MS spectrum upon fragmentation of singly oxidized FOXO4 depleted precursors, displaying 3 isolated precursors, charge stripping and fragments observed at a lower intensity. I searched for *c*- and *z*-fragment ions in spectra, including their singly oxidized forms, to calculate the level of oxidation (Figure 21B). However, only 62 ions (32 *c*-ions, 30 *z*-ions) were intense enough to reliably quantify them, with sequence coverage reaching ~25% (Figure 21C). This sequence coverage precluded any attempt at acquiring structural information on most protein residues.

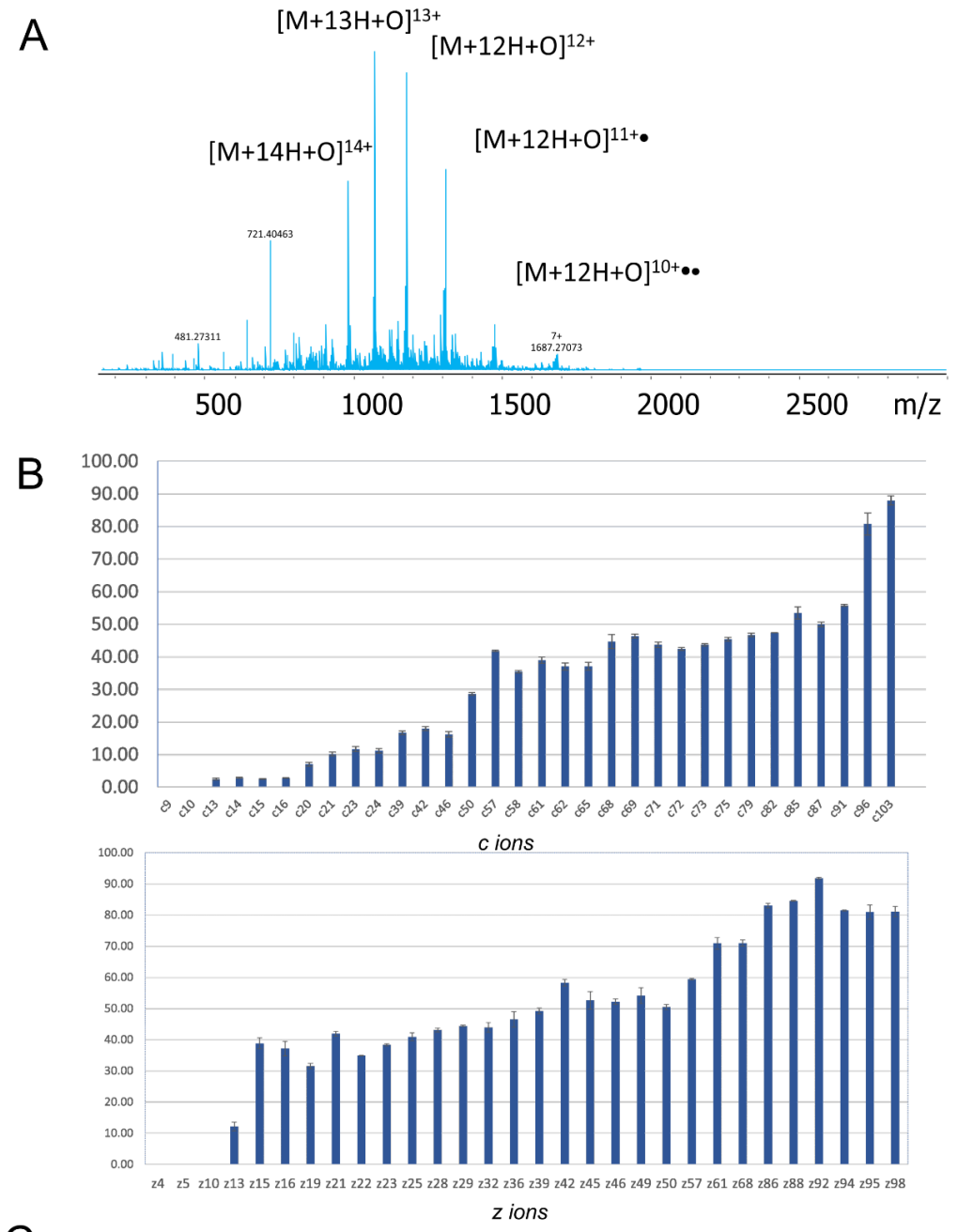


Figure 21 | A. ETD MS/MS spectra of singly oxidized isotopically deleted FOXO4-DBD. B. Quantified extent of oxidation of c- and z-fragment ions. Sequence of FOXO4-DBD according to wild type FOXO4 with denoted quantified fragment ions.

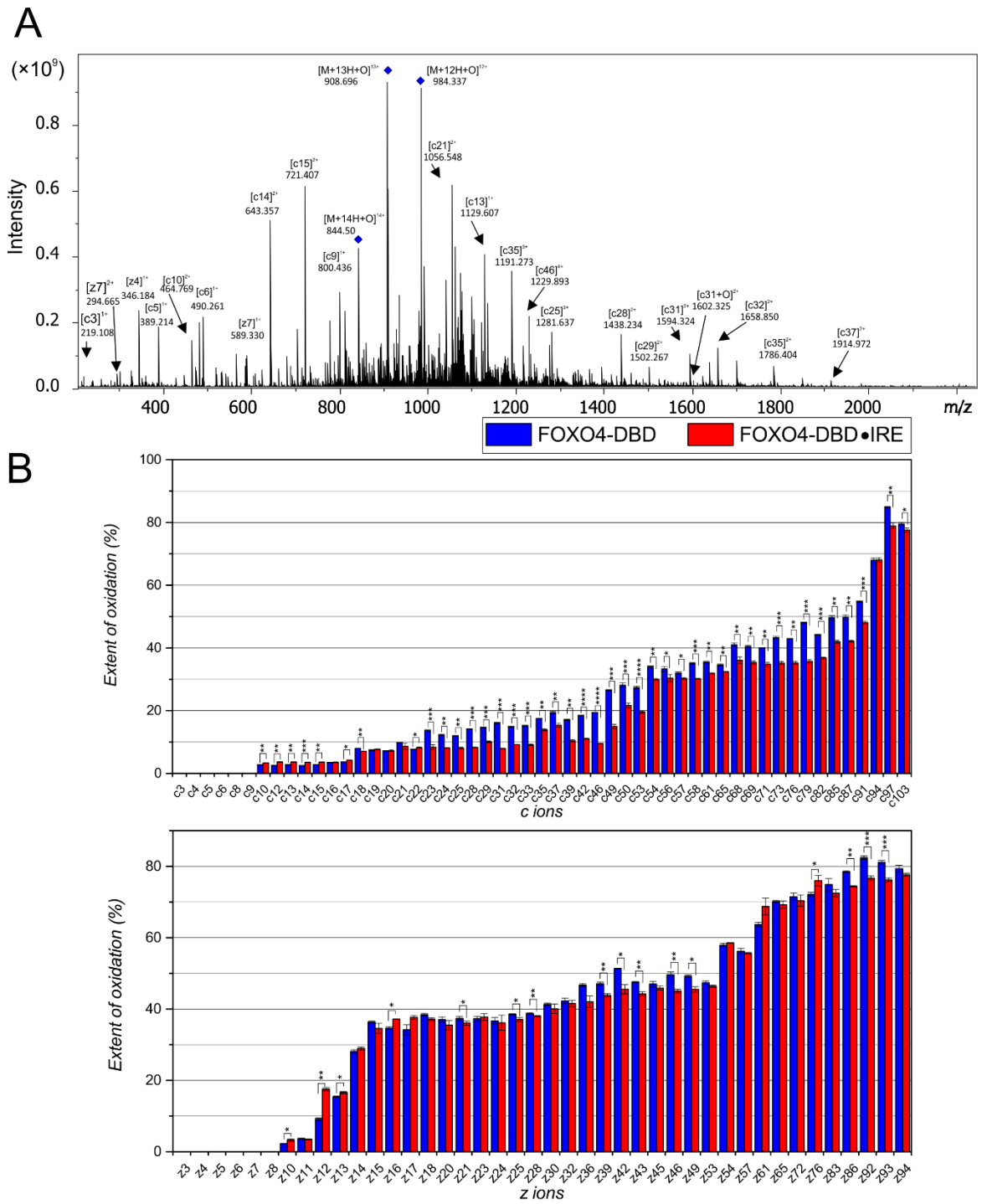


Figure 22 | A. ECD MS/MS spectra of singly oxidized isotopically deleted FOXO4-DBD. B. Quantification of the oxidation of c- and z-fragment ions. Sequence of FOXO4-DBD according to wild type FOXO4 with denoted quantified fragment ions.

Ultimately, I tested *electron-capture dissociation (ECD)*. ECD also produces complementary *c*- and *z*-fragment ions[28]. In this experiment, singly oxidized ions of IN- and ID-FOXO4 were isolated in quadrupole in a multiCASI mode, transferred to the ICR cell and irradiated using low energy electrons. Figure 22A shows ECD MS/MS a spectrum of oxidized ID FOXO4-DBD. ECD fragmentation provided a set of complementary fragment ions whose oxidation I quantified, without and with IRE (blue and red fragment ions, respectively, Figure 22B). Overall, I was able to quantify 95 fragment ions (54 *c* ions, 42 *z* ions), reaching a sequence coverage of ~45%. In contrast to IN samples, where only 57 fragment ions were quantified (30 *c* ions, 27 *z* ions, ~27 % sequence coverage), fragmentation boosted sequence coverage by almost 20 % (Figure 22C).

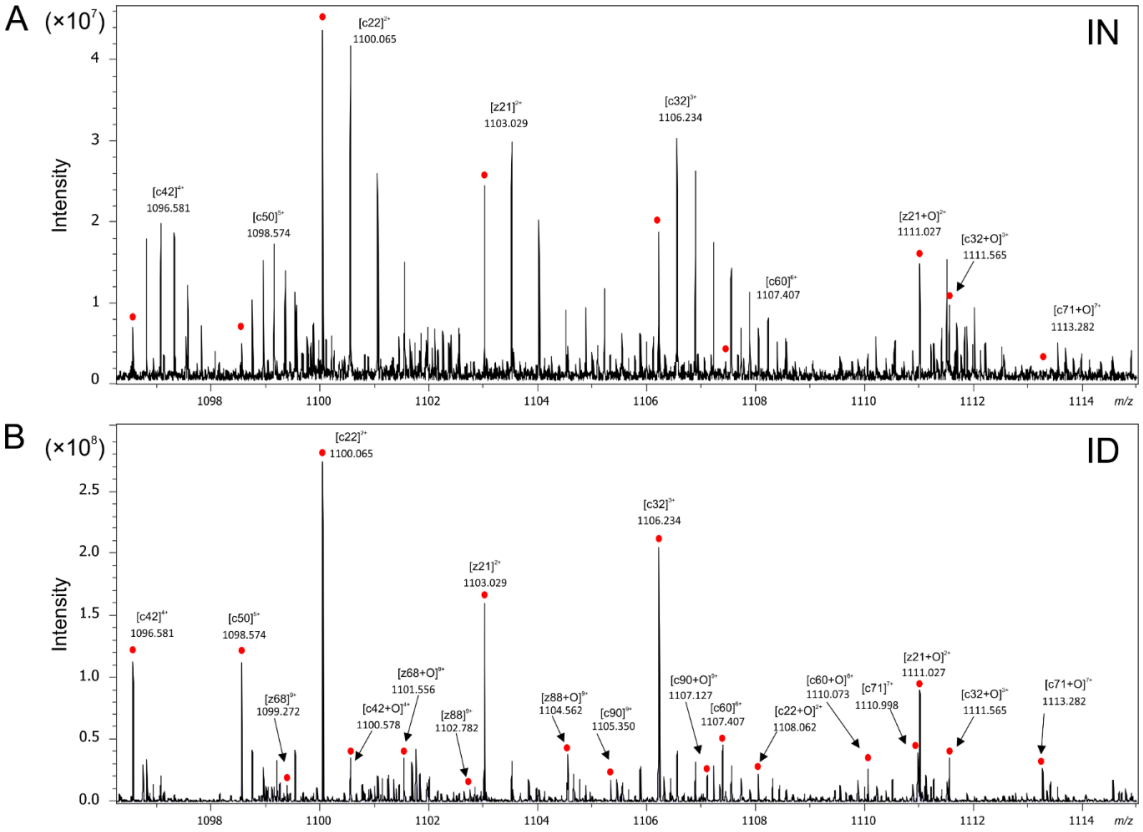


Figure 23 | A. Zoomed-in MS/MS spectra at m/z 1096–1115 upon fragmentation of isotopically natural (A) and isotopically depleted (B) samples. Fragment ions or their oxidized forms are denoted above the fragments with their monoisotopic masses. The red dot represents the position of monoisotopic peaks of particular fragment ions.

ID-MS/MS spectra displayed (i) a lower complexity than IN spectra, which greatly reduced the overlap of isotopes/fragment ions, (ii) a monoisotopic peak for all

fragment ions, and (iii) an increased signal-to-noise ratio (SNR). For instance, the zoom-in MS/MS spectra at *m/z* 1096–1115 (Figure 23) upon fragmentation of natural (A) and depleted (B) samples show that the intensity of fragment ions increased almost 10× (compare y-axis), that SNR improved, and that more ions were detected in IN samples (monoisotopic peak denoted by red dots). Combined, these results clearly demonstrate advances in protein isotope depletion.

The high-sequence coverage ECD spectra offered a more accurate oxidation assignment. Figure 24 displays an *in-silico* FOXO4-DBD•IRE structural model constructed based on the top-down dataset. When fragmenting IN-protein samples, only a few fragment ions are generated, and the overall extent of oxidation of these fragment ions is the sum of the oxidations of residues located in each region and their exposure to the solvent. In fact, fragmenting an ID-protein improved the spectra, SNR, and sequence coverage, which allowed a more thorough data analysis. Based on this analysis, I extracted higher structural information.

As shown in Figure 24, only a few residues were unambiguously assigned from the top-down dataset. However, ID protein uncovered the oxidation of 30 residues, namely K10, N16, W18, Q21, Y23,L26/I27, Q29, I31, E32, P35, Y45, R50, K58, D60, N62, S63, S64, H73, H78, I82, K83, H85, E87, S92, S93, W94, W95, M96, L97 and P99. Among them, 22 residues were de-novo deduced solely from the isotopically depleted top-down dataset. To assess this dataset, I also analyzed the IN sample using a bottom-up approach, meaning protein digestion and peptide analysis on a timsTOF Pro mass spectrometer in LC-MS(/MS) mode. These results are reported **in attached publication II**, supporting information, FigureS11. Nevertheless, 9 of the aforementioned residues, W18, Y23, Y45, Y54, H73, H78, W94#, W95 and M96 were found also in the bottom-up dataset, with the same trend, confirming the top-down dataset.

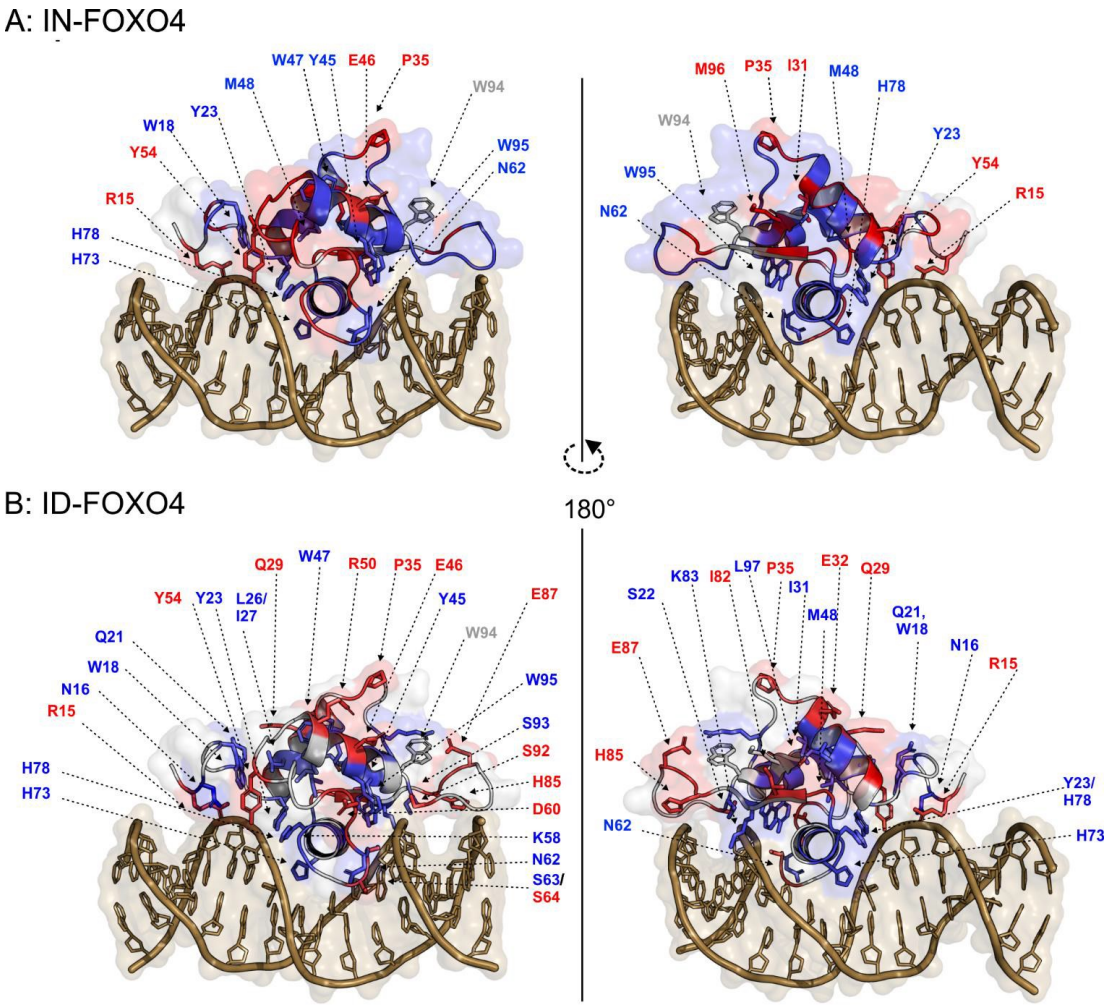


Figure 24 | An in-silico structural model of FOXO4-DBD•IRE (PDB template 3l2c)[157] *highlighting oxidized regions/residues detected by top-down analysis for natural version (A) or depleted (B) forms of FOXO4-DBD. The individual residues detected using the bottom-up approach or deduced using the top-down approach were highlighted and colored in the model. Blue – regions/residues more modified in apo form; red – regions/residues more modified in holo form.*

In summary, **publication II** identified an easy approach to gathering higher structural information by FPOP top-down analysis. Here, I showed that coupling FPOP with ECD and ion accumulation in a multiCASI mode is more than efficient than coupling FPOP with UVPD or ETD fragmentation techniques. Fragmenting the depleted sample increased the quality of the spectra and the sequence coverage, thus resulting in a more reliable oxidation assignment. These findings were highlighted in a structural model shown in Figure 24, which demonstrates marked advances in protein-isotope depletion applied to the analysis of covalently modified proteins.

4.3 Publication III

Title: Fast Photochemical Oxidation of nucleic acids coupled to high-resolution Mass spectrometry analysis *(manuscript in preparation)*.

Figure 16 and Figure 19 displays intact MS spectra of FOXO4-DBD after FPOP oxidation in solution alone and in the presence of DAF16 and IRE, respectively. When DNA is introduced to the sample, the abundance of oxidized proteoforms decreases, which could be explained by residue protection. However, experiments conducted in the 1980s to assess DNA damage by gel electrophoresis prompted me to also analyze the DNA site of the complex.

The initial experiments mapping nucleic acid damage aimed to detect modifications on DNA using high-energy electrons pulses from a γ-ray source[119]. However, one of the footprinting studies performed by Thomas Tullius aimed to readout the transcription factor binding motif from gel electrophoresis, as shown in Figure 6, in the Introduction part[169]. Hydroxyl radicals were generated *in-situ* using Fenton chemistry. But a time-consuming obstacle remained, that is, preparing a set of experiments with either at 5'- or 3'-radiolabeled termini and treating them as two independent reactions. Additionally, the lifetime of radicals cannot be controlled as in FPOP-induced experiments, and gel electrophoresis experiments require Max-Gilbert sequencing independently performed in additional lane(s), consuming even more time.

Hydroxyl radical footprinting of DNA is still the method of choice, coupling sample analysis with either denaturing gel[98,108] or capillary[170] electrophoresis. To determine the protein-DNA interface, I initiated DNA fragmentation by H_2O_2 with UV laser irradiation. Emulating the founding principles of protein-DNA footprinting, I analyzed FPOP-induced DNA damage by high-resolution MS. The initial experiment included fragmenting dsIRE using FPOP to study the characteristic products of the experiment. Thus, dsIRE was fragmented on a FPOP apparatus, the sample was collected, and fragments were separated and detected by LC after adding triethylamine (TEA) and hexafluoroisopropanol (HFIP) in mobile phases as previously described[171,172].

The •OH radicals initiate the attack on DNA by subtracting the hydrogen atom from the sugar backbone, resulting in nucleic acid fragmentation. The radicals predominantly attack sugar-backbone hydrogens,[101,106] following the reactivity H-5'>H-4'>H-3'≈H-2'≈ H-1', which is directly proportional to their solvent accessibility. The attack subtracting H-C4 ultimately leads to 3'phospate (3'P) cleavage[101], but I also detected 3'hydroxyl (3'OH), 3'phosphoglycerate (3'PG), 5'phosphate (5'P), 5'hydroxyl (5'OH), and 5'alhedyde (5'Ald), among other products, as shown in Figure 25. Then, I calculated the monoisotopic masses of all possible DNA fragments and acquired ion chromatograms of these fragments (Figure 26). By MS/MS analysis, I confirmed the presence of DNA fragments by identifying their *(a-B)*- and *w*- fragment ions[47,173] and by assessing their chromatographic parameters, observing fragment n-1 at a shorter retention time than fragment n (See Figure 26).

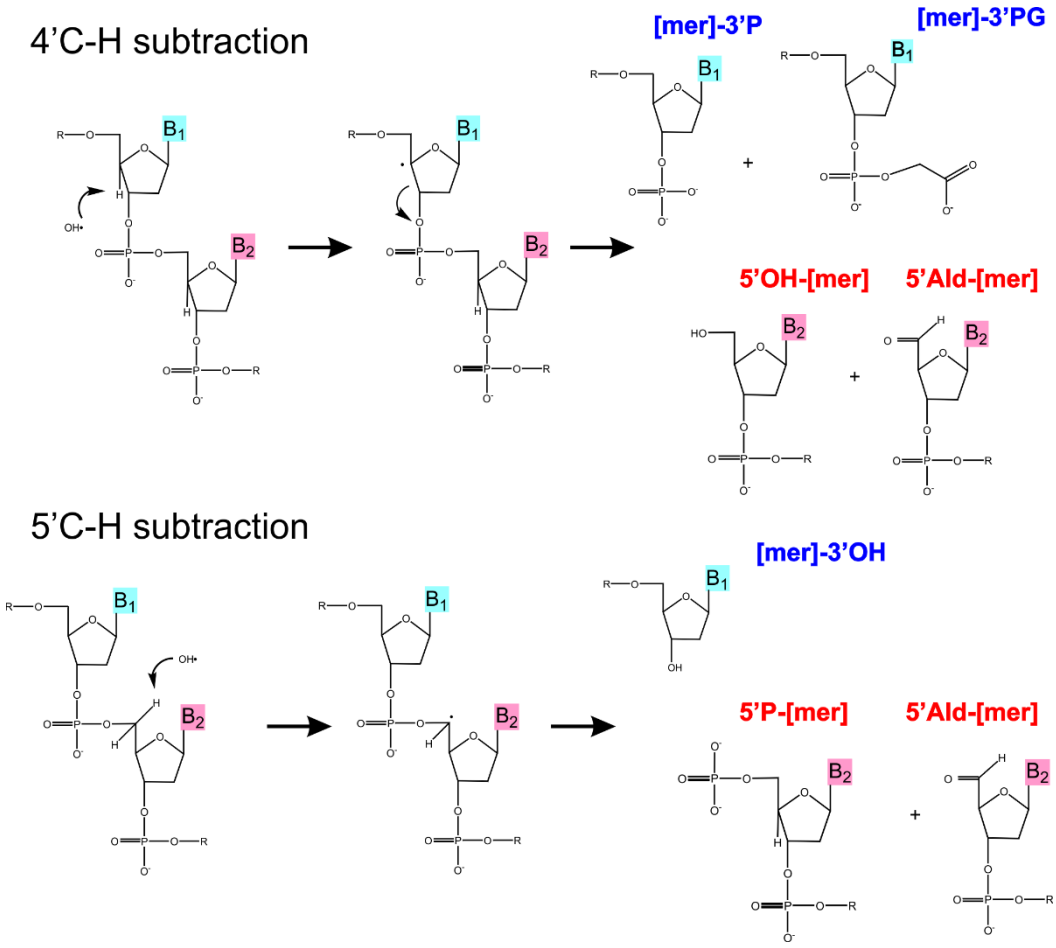


Figure 25 | Mechanism of nucleic acid fragmentation and products detected after FPOP in the LC-MS trace, in line with previously published literature [101,155,174].

3'PG products were observed in the LC-MS trace just above the level of detection. The ability to detect these products at such low amounts demonstrates that FPOP is a more valuable method for generating •OH radicals than others because the 3'PG product of fragment n is generated from fragment n+1 by 'overdamaging'[119] and that the products could not be separated and detected using classical gel electrophoresis, highlighting the merits of coupling the damaging experiment with FPOP/LC-MS. Additionally, previous studies have shown that the ultimate products are 5'-Ald DNA fragments. Nevertheless, by applying the single-hit kinetic FPOP platform, I was also able to detect 5'OH terminated fragments.

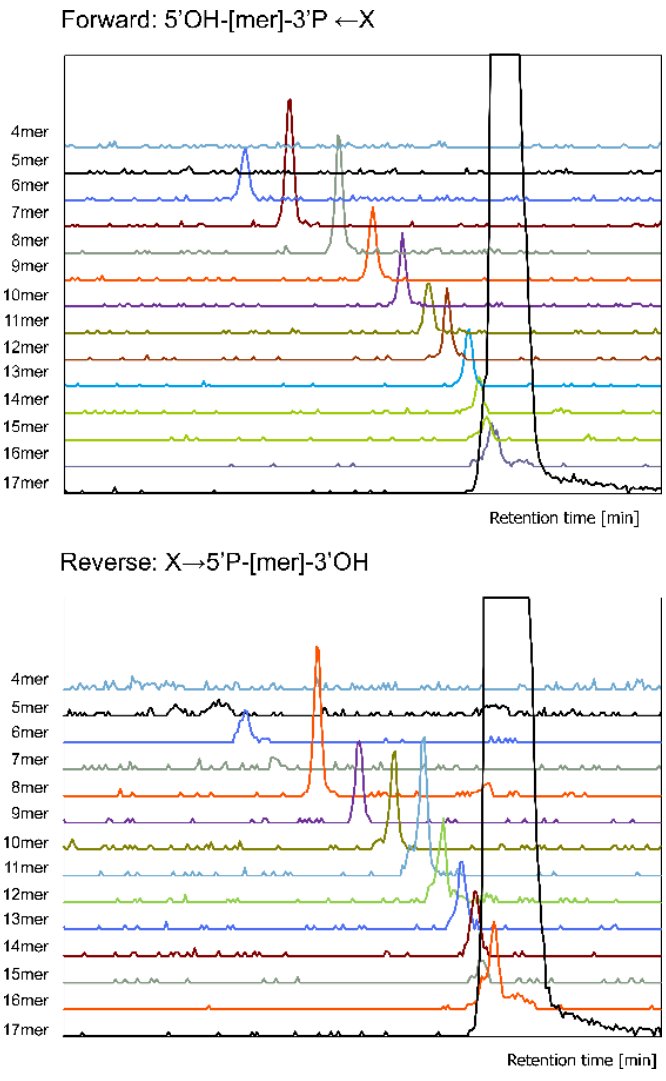


Figure 26 | Example of extracted ion chromatograms of DNA fragments originated from Forward (top panel) or Reverse strand (bottom panel). The main zoomed signal represents 17mers forward and reverse strand (colored in black).

In the next footprinting experiment, FPOP was performed on dsIRE without and with the transcription factor FOXO4-DBD at two H_2O_2 concentrations, 2 and 10 mM. I then searched for all possible DNA fragments from both forward and reverse strands by generating extracted ion chromatograms, subtracted the intensity of the fragments, calculated the extent of oxidative damage, statistically analyzed dataset and plotted the results.

Figure 27 shows quantified DNA fragments of the forward strand, which were truncated from the 3' direction, at both concentrations of peroxide, 2 and 10 mM. The reader can find the other fragments and dataset from the reverse strand in the attached manuscript.

Nevertheless, Figure 27 also displays the extent of oxidative damage to IRE fragments, and the resulting IRE fragments in solution both without (blue histograms) and with FOXO4 (pink histograms). In conclusion, fragmentation occurs at both H_2O_2 concentrations (2 and 10 mM in this study vs. previously 200mM in Fenton chemistry-coupled gel experiments)[98], (ii) dsIRE fragmentation at 2 mM H_2O_2 yielded ~6-8% of the most abundant fragments, whereas the reaction at 10 mM H_2O_2 produced ~4% fragments of the most abundant fragments, (iii) proteins significantly restricted hydrogen subtraction by •OH radicals, so FOXO4 footprinted regions around the binding motif/major and minor grooves[175]. Considering (i) and (ii), even a minor concentration of peroxide generates a set of complimentary fragments lacking base specificity but nearly the same number of fragments as a higher concentration of H_2O_2, so (iii) supports the idea of creating a 'footprint' around the binding motif, corroborating previously published studies on DNA footprinting.

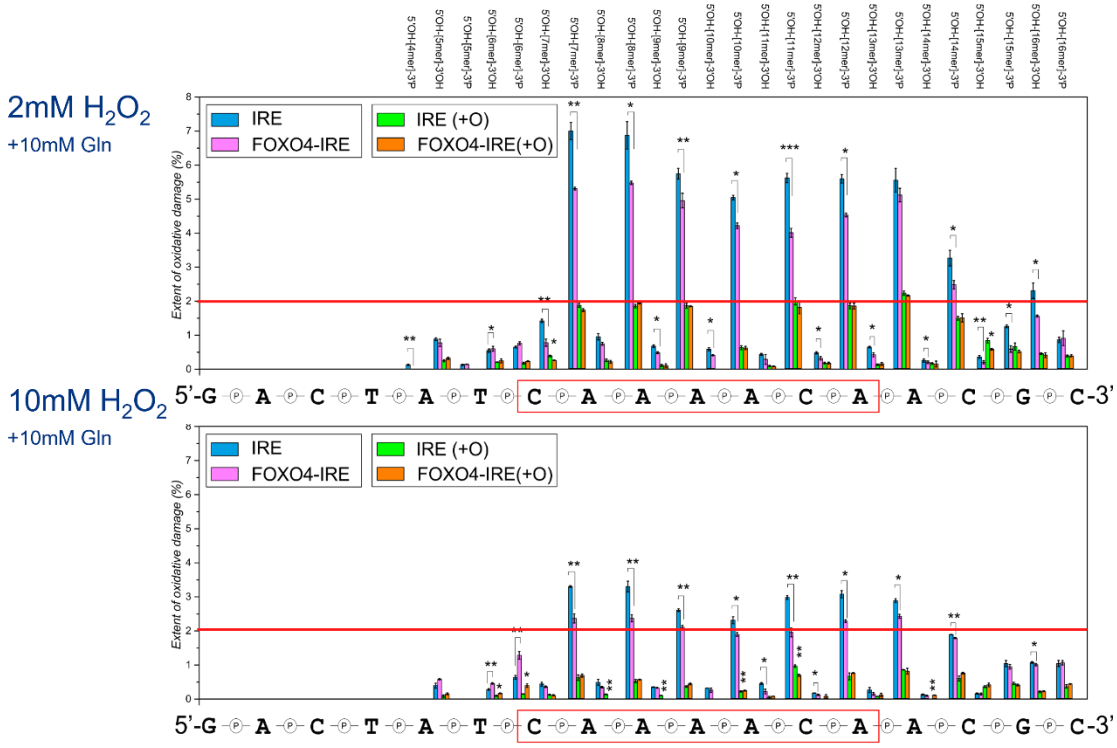


Figure 27 | Example of extracted ion chromatograms of DNA fragments originated from Forward (top panel) or Reverse (bottom panel) strands. The main zoomed-in signal represents 17mers forward and reverse strand (in black).The red line represents the informative level of the extent of damage at 2 %.

Theoretically, labelling agents should yield more fragment products at a higher concentration. However, the data showed that more intense fragments were

generated at 2 mM H_2O_2 than at 10 mM H_2O_2. Most likely, •OH radicals react more strongly with DNA, converting DNA fragments into other products, as previously described in the literature[155,176]. Accordingly, I hypothesize that DNA fragments are oxidized, dispersing the ion signal and hence adversely affecting MS detection. To address this issue, I quantified singly oxidized fragments with covalently oxidized bases detected in MS spectra. These singly oxidized fragments are highlighted in Figure 27, in green and orange histograms, respectively. At 2 mM H_2O_2, the level of oxidized fragments is ~2%, but when increasing the concentration 5×, their level decreases to ~0.5%. Thus, •OH radicals convert DNA fragments into undefined products, commonly known as DNA lesions[176].

In the next experiment, I assessed whether the FPOP platform can be used to study higher-order structures of single-stranded nucleic acids. The purpose of this experiment was to collect detailed information on interactions of single-stranded nucleic acids upon protein (protein-ssDNA and protein-RNA complexes), ligand, or single stranded nucleic acid binding, forming higher-order structures, such as G-quadruplexes, tRNA, DNA/RNA hairpins[177–180]. To this end, I induced FPOP fragmentation of forward and reverse single strands alone in solution and in their annealed duplex form. Samples were collected and analyzed by LC-MS, and the extent of oxidative damage of individual fragments was quantified, statistically analyzed, and plotted, showing that single strands are extensively fragmented in solution. However, the data on duplex DNA highlighted the 'up' and 'down' effect that is also observed in typical DNA gel-electrophoresis experiments[106]. In short, FPOP can be useful to study higher-order structures of single-strand nucleic acid in solution, but further research of single-stranded nucleic acids must be conducted to shed light on its folding in solution.

The last model system I studied in this publication consisted of two transcription factors, FOXO4-DBD and TEAD1-DBD, which interact with a 19bp oligonucleotide termed NO2. This ternary complex model system is under extensive research in our laboratory to structurally characterize their mutual interactions using MS-based techniques as one of our colleagues noticed that both binding motifs are often located near each other in the human genome. Thus, in this experiment, I induced FPOP-fragmentation of dsNO2 alone, in a complex with FOXO4 or TEAD1, and in a ternary complex with both protein domains (FOXO4-TEAD1-NO2) in solution. The data and the plots displaying the extent of damage are reported in

Publication III, in supporting information. Here, Figure 28 show heatmaps of all three complex variants created as "*Extent of damage of NO2 fragment*$_{in\ solution}$" - "*Extent of damage of NO2 fragment*$_{in\ complex}$". Non-significant changes are denoted by "×". The high fragmentation of the NO2 duplex contrasts with the major protection of the binding motifs when either FOXO4 or TEAD1 interacts with NO2. Interestingly, both 5' and 3' termini are more deprotected when proteins interact with NO2, which may be explained by deprotection of the ends when proteins interact with DNA in the major groove. We can also observe the expected protection of phosphate-terminated ends in the middle of sequence, resulting from direct hydrogen protection from the solvent by proteins.

When TEAD1 binds to NO2, more 3'OH hydroxyl ends of the forward strand located in the binding sequence are generated (higher damage, 5'ATTCC3'), as observed in the reverse strand when the FOXO4 motif is located (3'CATTTGT5') in the neighboring groove. The 3'CATTTGT5'-binding sequence displays a similar effect when bound to FOXO4. This effect weakens in a ternary complex when TEAD and FOXO4 bind to NO2 and may be explained by conformational changes in deoxyribose, which is more susceptible to the solvent while interacting with amino acid side chains of the protein. TEAD1 binding most likely melts DNA in its own major groove, with conformational effects on the second groove. In contrast, FOXO4 has a much stronger melting effect on its own groove and only a weak effect on the TEAD1 binding motif. In the ternary complex, TEAD1 still has a melting effect on its major groove, but the ability of FOXO4 to melt its own group is decreased. Therefore, (ii) both proteins induce conformational changes in DNA sugar moieties, and (ii) TEAD1 most likely decreases the ability of FOXO4 to bind to NO2. Based on these data, FPOP enables us to monitor even slight conformational changes induced by interactions with proteins and nucleic acids.

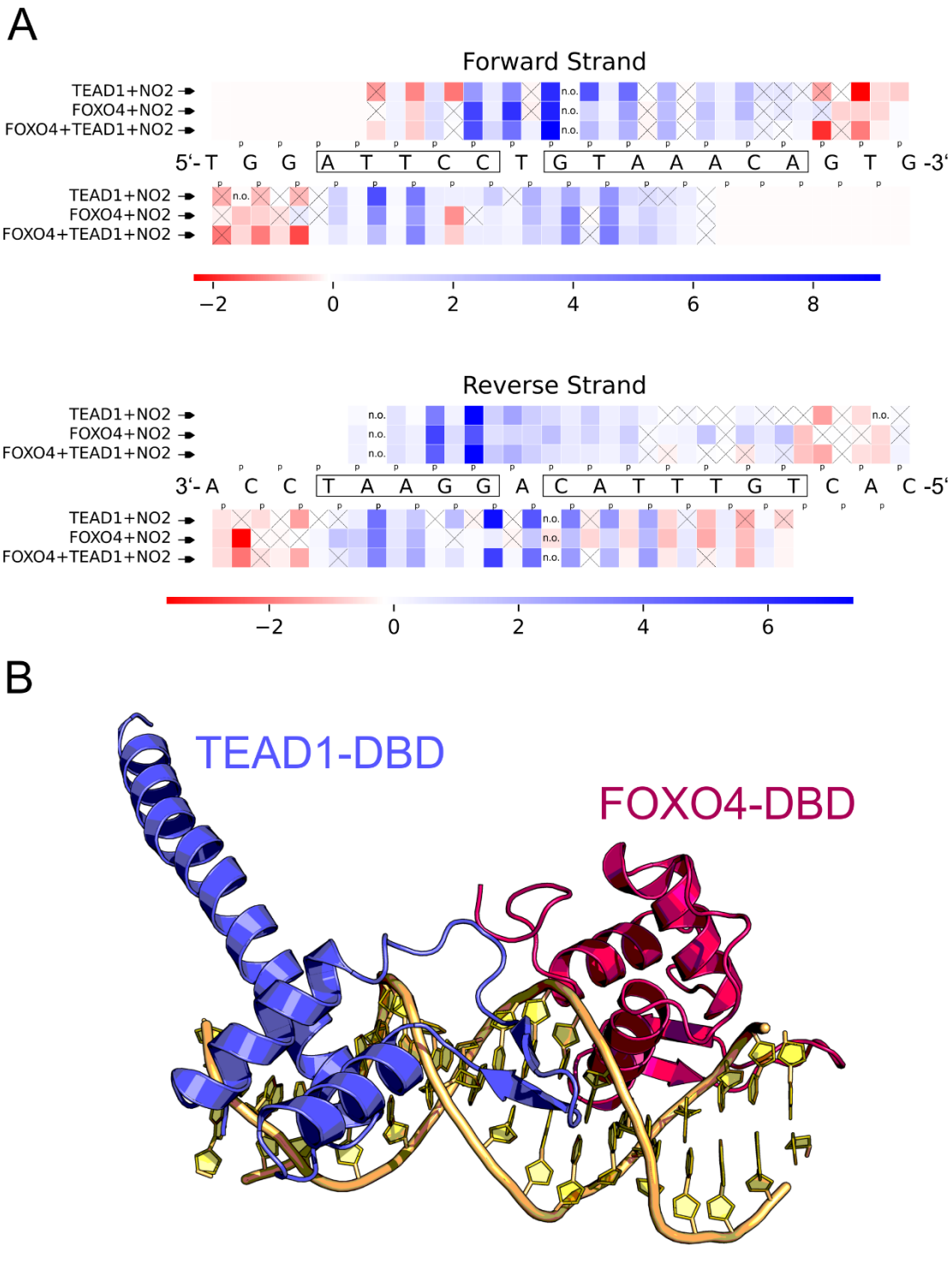


Figure 28 | A. Heatmaps of NO2 in TEAD1-NO2, FOXO4-NO2, and TEAD1-FOXO4-NO2. Non-significant changes are denoted by "×". N.O. – fragment not observed. Blue squares – protection of the position (fewer fragments), red square – deprotection of the position (more fragments).

4.4 Publication IV

Title: Design of AsLOV2 domain as a carrier of light-induced dissociable FMN photosensitizer.

Hydroxyl radicals encompass the most prevalent ROS. Among them, singlet oxygen, $^1O_2$, is a common ROS with high reactivity towards cellular components. Singlet oxygen, $O_2(a1\Delta g)$ or $^1O_2$, is the lowest excited electronic state of molecular oxygen, which is generated by energy transfer to ground (triplet) molecular oxygen. $^1O_2$ is significantly more reactive than O_2[181] and is one of the most biologically important ROS[182]. For these reasons, its reactivity, reaction mechanisms, and rate constants have been studied extensively for organic compounds[183] and for amino acids and peptides[184].

$^1O_2$ is known to react with a wide range of biomolecules, including DNA/RNA, proteins and all classes of lipids[185], causing oxidative damage in living organisms. Proteins are its main biological target due to their abundance and fast reaction rates with $^1O_2$, as consistently shown by kinetic data. $^1O_2$ primarily oxidizes Trp, His, Tyr, Met, and Cys side-chains[186] by first generating amino acid peroxides and later their decomposition products, in high yields, mostly on the aforementioned residues[187].

Photosensitizers (PSs) are organic compounds that efficiently produce $^1O_2$ when exposed to specific wavelengths. PSs are commonly employed in photodynamic, antimicrobial and antiviral therapy targeting cancer, bacterial and viral cells, respectively[188]. However, the low solubility of chemical photosensitizers limits their delivery. Nevertheless, genetically encoded photosensitizers (GEPSs) are currently undergoin extensive tests and development in cancer research as PS carriers[189].

Notwithstanding their potential, GEPSs face some challenges, such as $^1O_2$ quenching by amino acids, forming binding pockets. In addition, GEPS scaffolds also restrict oxygen diffusion to PSs, releasing $^1O_2$ out of the scaffold[190]. Previous research conducted by Petrenčáková et al.[191] demonstrated that C450 oxidation near the isoalloxazine ring is important for releasing FMN from the binding pocket. In this study performed in collaboration with Associate Professor Erik Sedlák from the Center of Interdisciplinary Biosciences in Košice, we investigated GEPS model

systems comprising FMN as a PS and LOV2 ("light, oxygen and voltage sensing") as a protein scaffold from *Avena sativa* (AsLOV2) to enhance $^1O_2$ production efficiency (Φ_Δ). The main aim of our research was to uncover the effect of mutations on protein oxidative damage upon irradiation (and to assess its extent). For this purpose, we applied MS-based techniques.

Continuing this approach, our colleagues in Košice expressed and purified three cysteine mutant protein variants, namely V416C, T418C and V416V/T418C, to accelerate FMN release and to compare variant forms with the wild-type (WT) form. Proteins were purified in the dark and then irradiated with a laser (475 nm) to induce $^1O_2$ production. Subsequently, WT and mutated variants were tested by far- and near-UV circular dichroism (CD), FMN fluorescence and DSC to assess the impact of embedded mutations in the secondary protein structure. The results showed that singly mutated variants preserve a secondary structure similar to that of the WT, whereas the V416C/T418C double mutant displayed destabilizations of the secondary structure. This finding was confirmed by CD and FMN fluorescence and by a decrease in thermal stability, which was nearly 8°C lower than that of the WT, measured by DSC. To elucidate the impact of point mutations, samples were delivered to our laboratory in Prague.

In Prague, my colleague, Dr. Dmitry Loginov, and I employed both bottom-up and top-down MS approaches to analyze proteins before and after irradiation. I analyzed the samples using the 'top-down' approach. Following this method, I offline desalted protein samples and sprayed them into a 15T-FT-ICR mass spectrometer. I then measured the mass of the analytes to confirm site-specific mutations and to ensure that no protein background oxidation had occurred. I subsequently isolated 4 charge states in multiCASI mode (multicontinuous accumulation of selected charge-states) and fragmented them by CID. Using our in-house software, developed in our laboratory (written by Dr. Daniel Kavan), I analyzed the data, searching for unmodified, singly, and doubly oxidized fragment ions in MS/MS spectra to calculate the oxidation ratio. This analysis of the top-down dataset revealed that several residues are oxidized throughout the protein structure. The most oxidized position is found in the V416C/T418C double mutant, in line with CD and DSC results demonstrating moderate destabilization of the protein scaffold, which releases FMN. The extent of oxidation ranged from y127 to y148 fragment ions, covering both mutations and F415 in WT. As shown in Figure 28, single and

double oxidations follow the increasing trend WT ~ V416C/T418C < V416C ~ T418C. The overall extent of oxidation between y127 and y148 fragment ions is driven by F415, C416 and C418 residues in all proteins. Residues at the double mutant remain almost unoxidized, which is mostly given by simple perturbation of FMN outside the binding pocket due to structural changes induced by mutations. This dataset is in strong agreement with our bottom-up results (Table 2), showing oxidation of cysteines C450, C416, C418 and F415. In the double mutant, F415 is not extensively oxidized either, confirming the perturbation effect of FMN. Overall, these results corroborate the findings of phosphorescence studies (Figure 3 in Publication IV) in that the highest phosphorescence is observed in the V416C/T418C double mutant, which is not quenched by amino acid side chains that form or fold the binding pocket.

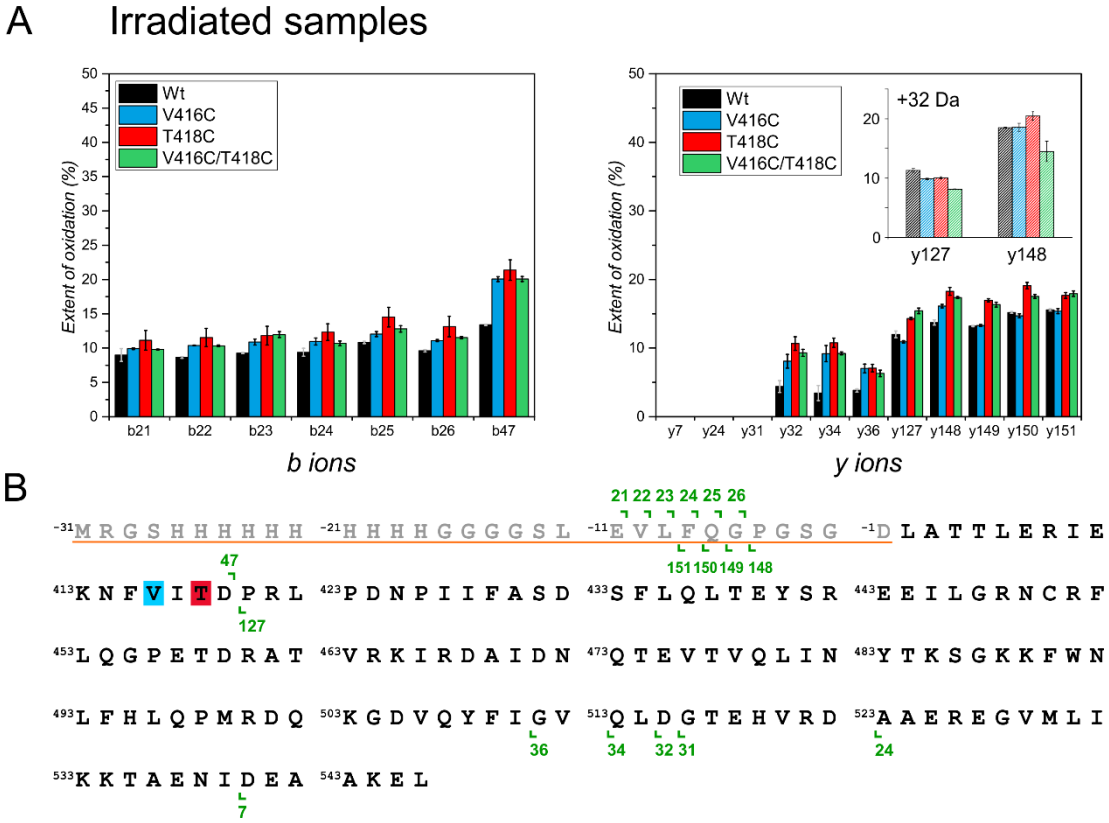


Figure 29 | (A) The results of the extent of oxidation of selected b and y fragment ions of irradiated samples are expressed as the mean ± SD of three independent measurements. The inset in the y-ions plot indicates the extent of doubly oxidized y127 and y148 fragment ions. (B) The sequence of wild type AsLOV2 protein with the denoted fragment ions is displayed at the top of the figure. The His-tag, which is not part of the AsLOV2 sequence, is colored in grey and underlined in red. The positions of the mutations V416C and T418C are shown in blue and red squares, respectively.

Table 2. List of selected oxidation products and their extent of oxidation.

Residue	Modification[a]	The extent of modification[b], %			
		WT	V416C	T418C	V416C/T418C
F415	1O	20.37	10.17	12.38	9.22
V416C	3O		21.61		5.04
T418C	3O			0.80	0.38
C450	3O	3.42	3.35	1.71	1.50

[a] - data represent values with subtracted background oxidation level ($S_{irradiated}$-$S_{control}$).
[b] - single modifications were identified using ion mobility.

In conclusion, this study demonstrates a simple and effective approach to protein surface mapping following oxidative damage caused by $^1O_2$ and its quantification. In this case study, the isolation window had to be widened to allow the transmission of oxidized clusters for CID, for mainly two reasons: (i) the size of AsLOV2 protein is ~ 20 kDa, resulting in overlapped isotope envelopes in oxidized clusters, and (ii) extensive artificial N-terminus methylation during bacterial expression. Consequently, methylated proteoforms (14.0154 amu) introduce biases to singly oxidized proteoforms (+15.9949).

Despite the need to exclude b-type fragment ions from the final data analysis and dataset evaluation, the top-down dataset matched the bottom-up results, which also showed increased oxidation of V416C and decreased oxidation of T418C position. Overall, MS data complemented biophysical techniques conducted in the laboratory in Košice, where they have shown secondary structure destabilization in double V416C/T418C mutant. Indeed, the level of oxidized cysteines in these positions was low.

5. Summary

The primary objective of this doctoral thesis is to advance the refinement of the Fast Photochemical Oxidation of Proteins (FPOP) platform for a more comprehensive study of proteins, protein-DNA complexes, and nucleic acids by high-resolution mass spectrometry analysis. The key finding and supporting evidence within my three publications as lead author and another manuscript that I co-authored are outlined as follows:

- **Successful optimization of FPOP system for studying protein-DNA complexes**: The FPOP apparatus was efficiently employed and optimized to oxidize the FOXO4-DAF16 complex. Both 'bottom-up' and 'top-down' approaches were employed to analyze the oxidized protein. Bottom-up analysis reached single-residue resolution information, whereas the top-down approach was tested to validate the bottom-up dataset, demonstrating high consistency between both approaches.
- **Protein depletion enabled enhanced spatial resolution in top-down experiments of FPOP samples**: Protein depletion was introduced as an approach to increase structural information in the top-down approach. Then, three gas-phase fragmentation techniques were systematically evaluated to maximize protein sequence coverage. Among them, multiCASI-ECD emerged as the optimal approach for boosting sequence coverage and thus structural information.
- **Structural proteomics guided model building of protein-DNA complexes**: Enhanced structural information from the top-down approach served as an effective guide for constructing *in*-silico models. Information from structural proteomic methods can effectively guide model-building operations to obtain *in-silico* protein-DNA complexes with a high-resolution structure.
- **UV-laser induced H_2O_2 decomposition can generate hydroxyl radicals for structural characterization of nucleic acids:** A FPOP apparatus with a UV laser was used to generate hydroxyl radicals from H_2O_2**,** inducing DNA damage. Preliminary resullts have shown that FPOP can be used to map protein-DNA interaction interface, higher-order structure of nucleic acids, and the dynamics of more complicated systems involving ternary complexes, including the FOXO4-TEAD1-NO2 complex.

- **Testing the potential of top-down approach to study protein oxidative damage by $^1O_2$**: The top-down approach was tested on a AsLOV2 model protein to assess the extent of oxidation caused by $^1O_2$, showcasing the potential of the top-down approach to study covalently labelled biomolecules, such as $^1O_2$.

6. List of Publications

Papers directly supporting this thesis:

1. **Polák, M.**; Yassaghi, G.; Kavan, D.; Filandr, F.; Fiala, J.; Kukačka, Z.; Halada, P.; Loginov, D. S.; Novák, P. Utilization of Fast Photochemical Oxidation of Proteins and Both Bottom-up and Top-down Mass Spectrometry for Structural Characterization of a Transcription Factor–DsDNA Complex. Anal. Chem. 2022, 94 (7), 3203–3210. doi:10.1021/acs.analchem.1c04746.

2. **Polák, M.**; Černý, J.; Novák, P. Isotopic Depletion Increases the Spatial Resolution of FPOP Top-Down Mass Spectrometry Analysis. Anal. Chem. 2024, 96 (4), 1478–1487. doi:10.1021/acs.analchem.3c03759.

3. **Polák, M.**; Brinsa, V.; Černý, J.; Kavan, D.; Volný, M.; Kádek, A.; Novák., Hydroxyl Radical Footprinting of Nucleic Acids Coupled to High-Resolution Mass Spectrometry Analysis, Nucleic Acids Res., **manuscript in preparation.**

4. Felčíková, K.; Hovan, A.; **Polák, M**.; Loginov, D.S.; Holotová, V.; Díaz, C.; Kožár, T.; Lee, O.; Varhač, R.; Novák, P., Bánó, G.; Sedlák, E. Design of AsLOV2 domain as a carrier of light-induced dissociable FMN photosensitizer. Prot. Sci., 2024, ***in press***. doi: 10.1002/pro.4921

Other publications by author:

5. **Polák, M.**; Palasser, M.; Kádek, A.; Kavan, D.; Wootton, C. A.; Delsuc, M.-A.; Breuker, K.; Novák, P.; van Agthoven, M. A. Top-Down Proteoform Analysis by 2D MS with Quadrupolar Detection. Anal. Chem. 2023, 95 (44), 16123–16130. doi:10.1021/acs.analchem.3c02225.

6. Möller, I. R.; Slivacka, M.; Hausner, J.; Nielsen, A. K.; Pospíšilová, E.; Merkle, P. S.; Lišková, R.; **Polák, M**.; Loland, C. J.; Kádek, A.; Man, P.; Rand, K. D. Improving the Sequence Coverage of Integral Membrane Proteins during Hydrogen/Deuterium Exchange Mass Spectrometry Experiments. Anal. Chem. 2019, 91 (17), 10970–10978. doi:10.1021/acs.analchem.9b00973.

7. Bibliography

(1) Pandey, P.; Hasnain, S.; Ahmad, S. Protein-DNA Interactions. In *Encyclopedia of Bioinformatics and Computational Biology*; Elsevier, 2019; pp 142–154. https://doi.org/10.1016/B978-0-12-809633-8.20217-3.

(2) Britt, H. M.; Cragnolini, T.; Thalassinos, K. Integration of Mass Spectrometry Data for Structural Biology. *Chem. Rev.* **2022**, *122* (8), 7952–7986. https://doi.org/10.1021/acs.chemrev.1c00356.

(3) Henzler-Wildman, K.; Kern, D. Dynamic Personalities of Proteins. *Nature*. Nature Publishing Group December 12, 2007, pp 964–972. https://doi.org/10.1038/nature06522.

(4) Konermann, L.; Vahidi, S.; Sowole, M. A. Mass Spectrometry Methods for Studying Structure and Dynamics of Biological Macromolecules. *Analytical Chemistry*. American Chemical Society January 7, 2014, pp 213–232. https://doi.org/10.1021/ac4039306.

(5) Dobson, C. M. Biophysical Techniques in Structural Biology. *Annu. Rev. Biochem.* **2019**, *88* (1), 25–33. https://doi.org/10.1146/annurev-biochem-013118-111947.

(6) Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C. A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* **1958**, *181* (4610), 662–666. https://doi.org/10.1038/181662a0.

(7) Gauto, D. F.; Estrozi, L. F.; Schwieters, C. D.; Effantin, G.; Macek, P.; Sounier, R.; Sivertsen, A. C.; Schmidt, E.; Kerfah, R.; Mas, G.; Colletier, J. P.; Güntert, P.; Favier, A.; Schoehn, G.; Schanda, P.; Boisbouvier, J. Integrated NMR and Cryo-EM Atomic-Resolution Structure Determination of a Half-Megadalton Enzyme Complex. *Nat. Commun.* **2019**, *10* (1). https://doi.org/10.1038/s41467-019-10490-9.

(8) Mittermaier, A. New Tools Provide New Insights in NMR Studies of Protein Dynamics. *Science (80-.).* **2006**, *312* (5771), 224–228. https://doi.org/10.1126/science.1124964.

(9) Kikhney, A. G.; Svergun, D. I. A Practical Guide to Small Angle X-Ray Scattering (SAXS) of Flexible and Intrinsically Disordered Proteins. *FEBS Lett.* **2015**, *589* (19), 2570–2577. https://doi.org/10.1016/J.FEBSLET.2015.08.027.

(10) Bajar, B.; Wang, E.; Zhang, S.; Lin, M.; Chu, J. A Guide to Fluorescent Protein FRET Pairs. *Sensors* **2016**, *16* (9), 1488. https://doi.org/10.3390/s16091488.

(11) Greenfield, N. J. Using Circular Dichroism Spectra to Estimate Protein Secondary Structure. *Nat. Protoc. 2007 16* **2007**, *1* (6), 2876–2890. https://doi.org/10.1038/nprot.2006.202.

(12) Lasker, K.; Phillips, J. L.; Russel, D.; Velázquez-Muriel, J.; Schneidman-Duhovny, D.; Tjioe, E.; Webb, B.; Schlessinger, A.; Sali, A. Integrative Structure Modeling of Macromolecular Assemblies from Proteomics Data.

Mol. Cell. Proteomics **2010**, *9* (8), 1689–1702. https://doi.org/10.1074/mcp.R110.000067.

(13) Tsiatsiani, L.; Heck, A. J. R. Proteomics beyond Trypsin. *FEBS J.* **2015**, *282* (14), 2612–2626. https://doi.org/10.1111/febs.13287.

(14) Chait, B. T. Mass Spectrometry: Bottom-up or Top-Down? *Science*. American Association for the Advancement of Science October 6, 2006, pp 65–66. https://doi.org/10.1126/science.1133987.

(15) Liu, X. R.; Zhang, M. M.; Gross, M. L. Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications. *Chem. Rev.* **2020**, *120* (10), 4355–4454. https://doi.org/10.1021/acs.chemrev.9b00815.

(16) van der Laarse, S. A. M.; van Gelder, C. A. G. H.; Bern, M.; Akeroyd, M.; Olsthoorn, M. M. A.; Heck, A. J. R. Targeting Proline in (Phospho)Proteomics. *FEBS J.* **2020**, *287* (14), 2979–2997. https://doi.org/10.1111/febs.15190.

(17) Tsiatsiani, L.; Akeroyd, M.; Olsthoorn, M.; Heck, A. J. R. Aspergillus Niger Prolyl Endoprotease for Hydrogen-Deuterium Exchange Mass Spectrometry and Protein Structural Studies. *Anal. Chem.* **2017**, *89* (15), 7966–7973. https://doi.org/10.1021/acs.analchem.7b01161.

(18) Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M.-C.; Yates, J. R. Protein Analysis by Shotgun/Bottom-up Proteomics. *Chem. Rev.* **2013**, *113* (4), 2343–2394. https://doi.org/10.1021/cr3003533.

(19) Lermyte, F.; Tsybin, Y. O.; O'connor, P. B.; Loo, J. A. Top or Middle? Up or Down? Toward a Standard Lexicon for Protein Top-Down and Allied Mass Spectrometry Approaches. *J. Am. Soc. Mass Spectrom* **2019**, *30*, 1149–1157. https://doi.org/10.1007/s13361-019-02201-x.

(20) Macias, L. A.; Santos, I. C.; Brodbelt, J. S. Ion Activation Methods for Peptides and Proteins. *Anal. Chem.* **2020**, *92* (1), 227–251. https://doi.org/10.1021/acs.analchem.9b04859.

(21) Juetten, K. J.; Brodbelt, J. S. Top-Down Analysis of Supercharged Proteins Using Collision-, Electron-, and Photon-Based Activation Methods. *J. Am. Soc. Mass Spectrom.* **2023**, *34* (7), 1467–1476. https://doi.org/10.1021/jasms.3c00138.

(22) Habeck, T.; Lermyte, F. Seeing the Complete Picture: Proteins in Top-down Mass Spectrometry. *Essays Biochem.* **2022**. https://doi.org/10.1042/EBC20220098.

(23) Chen, B.; Brown, K. A.; Lin, Z.; Ge, Y. Top-Down Proteomics: Ready for Prime Time? *Analytical Chemistry*. American Chemical Society January 2, 2018, pp 110–127. https://doi.org/10.1021/acs.analchem.7b04747.

(24) Háda, V.; Bagdi, A.; Bihari, Z.; Timári, S. B.; Fizil, Á.; Szántay, C. Recent Advancements, Challenges, and Practical Considerations in the Mass Spectrometry-Based Analytics of Protein Biotherapeutics: A Viewpoint from the Biosimilar Industry. *J. Pharm. Biomed. Anal.* **2018**, *161*, 214–238.

https://doi.org/10.1016/j.jpba.2018.08.024.

(25) Catherman, A. D.; Skinner, O. S.; Kelleher, N. L. Top Down Proteomics: Facts and Perspectives. *Biochem. Biophys. Res. Commun.* **2014**, *445* (4), 683–693. https://doi.org/10.1016/j.bbrc.2014.02.041.

(26) Johnson, A. R.; Carlson, E. E. Collision-Induced Dissociation Mass Spectrometry: A Powerful Tool for Natural Product Structure Elucidation. *Analytical Chemistry*. American Chemical Society November 3, 2015, pp 10668–10678. https://doi.org/10.1021/acs.analchem.5b01543.

(27) Brodbelt, J. S. Deciphering Combinatorial Post-Translational Modifications by Top-down Mass Spectrometry. *Curr. Opin. Chem. Biol.* **2022**, *70*, 102180. https://doi.org/10.1016/j.cbpa.2022.102180.

(28) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *J. Am. Chem. Soc.* **1998**, *120* (13), 3265–3266. https://doi.org/10.1021/ja973478k.

(29) Zubarev, R. A. Electron-Capture Dissociation Tandem Mass Spectrometry. *Curr. Opin. Biotechnol.* **2004**, *15* (1), 12–16. https://doi.org/10.1016/j.copbio.2003.12.002.

(30) Lermyte, F.; Valkenborg, D.; Loo, J. A.; Sobott, F. Radical Solutions: Principles and Application of Electron-based Dissociation in Mass Spectrometry-based Analysis of Protein Structure. *Mass Spectrom. Rev.* **2018**, *37* (6), 750–771. https://doi.org/10.1002/mas.21560.

(31) Laskin, J.; Futrell, J. H.; Chu, I. K. Is Dissociation of Peptide Radical Cations an Ergodic Process? *J. Am. Chem. Soc.* **2007**, *129* (31), 9598–9599. https://doi.org/10.1021/ja073748r.

(32) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. Peptide and Protein Sequence Analysis by Electron Transfer Dissociation Mass Spectrometry. *Proc. Natl. Acad. Sci.* **2004**, *101* (26), 9528–9533. https://doi.org/10.1073/pnas.0402700101.

(33) Qi, Y.; Volmer, D. A. Electron-Based Fragmentation Methods in Mass Spectrometry: An Overview. *Mass Spectrom. Rev.* **2017**, *36* (1), 4–15. https://doi.org/10.1002/mas.21482.

(34) Riley, N. M.; Coon, J. J. The Role of Electron Transfer Dissociation in Modern Proteomics. *Anal. Chem.* **2018**, *90* (1), 40–64. https://doi.org/10.1021/acs.analchem.7b04810.

(35) Papanastasiou, D.; Kounadis, D.; Lekkas, A.; Orfanopoulos, I.; Mpozatzidis, A.; Smyrnakis, A.; Panagiotopoulos, E.; Kosmopoulou, M.; Reinhardt-Szyba, M.; Fort, K.; Makarov, A.; Zubarev, R. A. The Omnitrap Platform: A Versatile Segmented Linear Ion Trap for Multidimensional Multiple-Stage Tandem Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2022**, *33* (10), 1990–2007. https://doi.org/10.1021/jasms.2c00214.

(36) Zhurov, K. O.; Fornelli, L.; Wodrich, M. D.; Laskay, Ü. A.; Tsybin, Y. O. Principles of Electron Capture and Transfer Dissociation Mass Spectrometry

Applied to Peptide and Protein Structure Analysis. *Chem. Soc. Rev.* **2013**, *42* (12), 5014. https://doi.org/10.1039/c3cs35477f.

(37) Brodbelt, J. S. Shedding Light on the Frontier of Photodissociation. *Journal of the American Society for Mass Spectrometry*. Springer-VerlagNew York February 2011, pp 197–206. https://doi.org/10.1007/s13361-010-0023-6.

(38) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. Infrared Multiphoton Dissociation of Large Multiply Charged Ions for Biomolecule Sequencing. *Anal. Chem.* **1994**, *66* (18), 2809–2815. https://doi.org/10.1021/ac00090a004.

(39) Fornelli, L.; Srzentić, K.; Toby, T. K.; Doubleday, P. F.; Huguet, R.; Mullen, C.; Melani, R. D.; dos Santos Seckler, H.; DeHart, C. J.; Weisbrod, C. R.; Durbin, K. R.; Greer, J. B.; Early, B. P.; Fellers, R. T.; Zabrouskov, V.; Thomas, P. M.; Compton, P. D.; Kelleher, N. L. Thorough Performance Evaluation of 213 Nm Ultraviolet Photodissociation for Top-down Proteomics. *Mol. Cell. Proteomics* **2020**, *19* (2), 405–420. https://doi.org/10.1074/mcp.TIR119.001638.

(40) R. Julian, R. The Mechanism Behind Top-Down UVPD Experiments: Making Sense of Apparent Contradictions. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (9), 1823–1826. https://doi.org/10.1007/s13361-017-1721-0.

(41) Bowers, W. D.; Delbert, S. S.; Hunter, R. L.; Mclver, R. T. Fragmentation of Oligopeptide Ions Using Ultraviolet Laser Radiation and Fourier Transform Mass Spectrometry. *J. Am. Chem. Soc.* **1984**, *106* (23), 7288–7289. https://doi.org/10.1021/ja00335a094.

(42) Liu, F. C.; Ridgeway, M. E.; Wootton, C. A.; Theisen, A.; Panczyk, E. M.; Meier, F.; Park, M. A.; Bleiholder, C. Top-Down Protein Analysis by Tandem-Trapped Ion Mobility Spectrometry/Mass Spectrometry (Tandem-TIMS/MS) Coupled with Ultraviolet Photodissociation (UVPD) and Parallel Accumulation/Serial Fragmentation (PASEF) MS/MS Analysis. *J. Am. Soc. Mass Spectrom.* **2023**, *34* (10), 2232–2246. https://doi.org/10.1021/jasms.3c00187.

(43) Madsen, J. A.; Gardner, M. W.; Smith, S. I.; Ledvina, A. R.; Coon, J. J.; Schwartz, J. C.; Stafford, G. C.; Brodbelt, J. S. Top-Down Protein Fragmentation by Infrared Multiphoton Dissociation in a Dual Pressure Linear Ion Trap. *Anal. Chem.* **2009**, *81* (21), 8677–8686. https://doi.org/10.1021/ac901554z.

(44) Macias, L. A.; Sipe, S. N.; Santos, I. C.; Bashyal, A.; Mehaffey, M. R.; Brodbelt, J. S. Influence of Primary Structure on Fragmentation of Native-Like Proteins by Ultraviolet Photodissociation. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (12), 2860–2873. https://doi.org/10.1021/jasms.1c00269.

(45) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. Complete Protein Characterization Using Top-down Mass Spectrometry and Ultraviolet Photodissociation. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651. https://doi.org/10.1021/ja4029654.

(46) Greer, S. M.; Brodbelt, J. S. Top-Down Characterization of Heavily Modified Histones Using 193 Nm Ultraviolet Photodissociation Mass Spectrometry. *J. Proteome Res.* **2018**, *17* (3), 1138–1145. https://doi.org/10.1021/acs.jproteome.7b00801.

(47) Santos, I. C.; Lanzillotti, M.; Shilov, I.; Basanta-Sanchez, M.; Roushan, A.; Lawler, R.; Tang, W.; Bern, M.; Brodbelt, J. S. Ultraviolet Photodissociation and Activated Electron Photodetachment Mass Spectrometry for Top-Down Sequencing of Modified Oligoribonucleotides. *J. Am. Soc. Mass Spectrom.* **2022**, *33* (3), 510–520. https://doi.org/10.1021/jasms.1c00340.

(48) Feider, C. L.; MacIas, L. A.; Brodbelt, J. S.; Eberlin, L. S. Double Bond Characterization of Free Fatty Acids Directly from Biological Tissues by Ultraviolet Photodissociation. *Anal. Chem.* **2020**, *92* (12), 8386–8395. https://doi.org/10.1021/acs.analchem.0c00970.

(49) Brodbelt, J. S.; Morrison, L. J.; Santos, I. Ultraviolet Photodissociation Mass Spectrometry for Analysis of Biological Molecules. *Chemical Reviews*. American Chemical Society April 8, 2020, pp 3328–3380. https://doi.org/10.1021/acs.chemrev.9b00440.

(50) Taucher, M.; Breuker, K. Top-down Mass Spectrometry for Sequencing of Larger (up to 61 Nt) RNA by CAD and EDD. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (6), 918–929. https://doi.org/10.1016/j.jasms.2010.02.025.

(51) Palasser, M.; Breuker, K. RNA Chemical Labeling with Site-Specific, Relative Quantification by Mass Spectrometry for the Structural Study of a Neomycin-Sensing Riboswitch Aptamer Domain. *Chempluschem* **2022**, e202200256. https://doi.org/10.1002/cplu.202200256.

(52) Keller, K. M.; Brodbelt, J. S. Collisionally Activated Dissociation and Infrared Multiphoton Dissociation of Oligonucleotides in a Quadrupole Ion Trap. *Anal. Biochem.* **2004**, *326* (2), 200–210. https://doi.org/10.1016/j.ab.2003.12.010.

(53) R. Julian, R. The Mechanism Behind Top-Down UVPD Experiments: Making Sense of Apparent Contradictions. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (9), 1823–1826. https://doi.org/10.1007/s13361-017-1721-0.

(54) Becher, S.; Wang, H.; Leeming, M. G.; Donald, W. A.; Heiles, S. Influence of Protein Ion Charge State on 213 Nm Top-down UVPD. *Analyst* **2021**, *146* (12), 3977–3987. https://doi.org/10.1039/D1AN00571E.

(55) Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science (80-.).* **1989**, *246* (4926), 64–71. https://doi.org/10.1126/science.2675315.

(56) Schejter, A.; Goldkorn, T.; Sokolovsky, M. Limited Proteolysis of Horse Heart Cytochrome C. *Eur. J. Biochem.* **1971**, *20* (3), 414–419. https://doi.org/10.1111/J.1432-1033.1971.TB01407.X.

(57) Ottesen, M. Induction of Biological Activity by Limited Proteolysis. *Annual review of biochemistry*. June 28, 1967, pp 55–76. https://doi.org/10.1146/annurev.bi.36.070167.000415.

(58) Cleveland, D. W.; Fischer, S. G.; Kirschner andLaemmli, M. W. U. K. Peptide Mapping by Limited Proteolysis in Sodium Dodecyl Sulfate and Analysis by Gel Electrophoresis. *J. Biol. Chem.* **1977**, *252* (3), 1102–1106. https://doi.org/10.1016/S0021-9258(19)75212-0.

(59) Cohen, S. L.; Chait, B. T.; Ferré-D'Amaré, A. R.; Burley, S. K. Probing the Solution Structure of the DNA-binding Protein Max by a Combination of Proteolysis and Mass Spectrometry. *Protein Sci.* **1995**, *4* (6), 1088–1099. https://doi.org/10.1002/pro.5560040607.

(60) Acquasaliente, L.; Pelc, L. A.; Di Cera, E. Probing Prothrombin Structure by Limited Proteolysis. *Sci. Rep.* **2019**, *9* (1), 6125. https://doi.org/10.1038/s41598-019-42524-z.

(61) Lössl, P.; Waterbeemd, M.; Heck, A. J. The Diverse and Expanding Role of Mass Spectrometry in Structural and Molecular Biology. *EMBO J.* **2016**, *35* (24), 2634–2657. https://doi.org/10.15252/embj.201694818.

(62) Schopper, S.; Kahraman, A.; Leuenberger, P.; Feng, Y.; Piazza, I.; Müller, O.; Boersema, P. J.; Picotti, P. Measuring Protein Structural Changes on a Proteome-Wide Scale Using Limited Proteolysis-Coupled Mass Spectrometry. *Nat. Protoc.* **2017**, *12* (11), 2391–2410. https://doi.org/10.1038/nprot.2017.100.

(63) de Souza, N.; Picotti, P. Mass Spectrometry Analysis of the Structural Proteome. *Curr. Opin. Struct. Biol.* **2020**, *60*, 57–65. https://doi.org/10.1016/j.sbi.2019.10.006.

(64) Brenowitz, M.; Senear, D. F.; Kingston, R. E. DNase I Footprint Analysis of Protein-DNA Binding. *Curr. Protoc. Mol. Biol.* **1989**, *7* (1), 12.4.1-12.4.16. https://doi.org/10.1002/0471142727.mb1204s07.

(65) Galas, D. J.; Schmitz, A. DNAase Footprinting a Simple Method for the Detection of Protein-DNA Binding Specificity. *Nucleic Acids Res.* **1978**, *5* (9), 3157–3170. https://doi.org/10.1093/nar/5.9.3157.

(66) Hvidt, A.; Nielsen, S. O. Hydrogen Exchange in Proteins. *Adv. Protein Chem.* **1966**, *21* (C), 287–386. https://doi.org/10.1016/S0065-3233(08)60129-1.

(67) HVIDT, A.; JOHANSEN, G.; LINDERSTRØM LANG, K.; VASLOW, F. Exchange of Deuterium and 18O between Water and Other Substances. I. Methods. *C. R. Trav. Lab. Carlsberg. Chim.* **1954**, *29* (9), 129–157.

(68) Oganesyan, I.; Lento, C.; Wilson, D. J. Contemporary Hydrogen Deuterium Exchange Mass Spectrometry. *Methods*. Academic Press July 15, 2018, pp 27–42. https://doi.org/10.1016/j.ymeth.2018.04.023.

(69) James, E. I.; Murphree, T. A.; Vorauer, C.; Engen, J. R.; Guttman, M. Advances in Hydrogen/Deuterium Exchange Mass Spectrometry and the Pursuit of Challenging Biological Systems. *Chem. Rev.* **2022**, *122* (8), 7562–7623. https://doi.org/10.1021/acs.chemrev.1c00279.

(70) Masson, G. R.; Burke, J. E.; Ahn, N. G.; Anand, G. S.; Borchers, C.; Brier, S.; Bou-Assaf, G. M.; Engen, J. R.; Englander, S. W.; Faber, J.; Garlish, R.; Griffin, P. R.; Gross, M. L.; Guttman, M.; Hamuro, Y.; Heck, A. J. R.; Houde,

D.; Iacob, R. E.; Jørgensen, T. J. D.; Kaltashov, I. A.; Klinman, J. P.; Konermann, L.; Man, P.; Mayne, L.; Pascal, B. D.; Reichmann, D.; Skehel, M.; Snijder, J.; Strutzenberg, T. S.; Underbakke, E. S.; Wagner, C.; Wales, T. E.; Walters, B. T.; Weis, D. D.; Wilson, D. J.; Wintrode, P. L.; Zhang, Z.; Zheng, J.; Schriemer, D. C.; Rand, K. D. Recommendations for Performing, Interpreting and Reporting Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) Experiments. *Nat. Methods* **2019**, *16* (7), 595–602. https://doi.org/10.1038/s41592-019-0459-y.

(71) Giansanti, P.; Tsiatsiani, L.; Low, T. Y.; Heck, A. J. R. Six Alternative Proteases for Mass Spectrometry–Based Proteomics beyond Trypsin. *Nat. Protoc.* **2016**, *11* (5), 993–1006. https://doi.org/10.1038/nprot.2016.057.

(72) Rey, M.; Yang, M.; Burns, K. M.; Yu, Y.; Lees-Miller, S. P.; Schriemer, D. C. Nepenthesin from Monkey Cups for Hydrogen/Deuterium Exchange Mass Spectrometry. *Mol. Cell. Proteomics* **2013**, *12* (2), 464–472. https://doi.org/10.1074/mcp.M112.025221.

(73) Yang, M.; Hoeppner, M.; Rey, M.; Kadek, A.; Man, P.; Schriemer, D. C. Recombinant Nepenthesin II for Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal. Chem.* **2015**, *87* (13), 6681–6687. https://doi.org/10.1021/acs.analchem.5b00831.

(74) Möller, I. R.; Slivacka, M.; Hausner, J.; Nielsen, A. K.; Pospíšilová, E.; Merkle, P. S.; Lišková, R.; Polák, M.; Loland, C. J.; Kádek, A.; Man, P.; Rand, K. D. Improving the Sequence Coverage of Integral Membrane Proteins during Hydrogen/Deuterium Exchange Mass Spectrometry Experiments. *Anal. Chem.* **2019**, *91* (17), 10970–10978. https://doi.org/10.1021/acs.analchem.9b00973.

(75) Kadek, A.; Mrazek, H.; Halada, P.; Rey, M.; Schriemer, D. C.; Man, P. Aspartic Protease Nepenthesin-1 as a Tool for Digestion in Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal. Chem.* **2014**, *86* (9), 4287–4294. https://doi.org/10.1021/ac404076j.

(76) Kadek, A.; Tretyachenko, V.; Mrazek, H.; Ivanova, L.; Halada, P.; Rey, M.; Schriemer, D. C.; Man, P. Expression and Characterization of Plant Aspartic Protease Nepenthesin-1 from Nepenthes Gracilis. *Protein Expr. Purif.* **2014**, *95*, 121–128. https://doi.org/10.1016/j.pep.2013.12.005.

(77) Brown, K. A.; Wilson, D. J. Bottom-up Hydrogen Deuterium Exchange Mass Spectrometry: Data Analysis and Interpretation. *Analyst*. Royal Society of Chemistry August 21, 2017, pp 2874–2886. https://doi.org/10.1039/c7an00662d.

(78) Masson, G. R.; Jenkins, M. L.; Burke, J. E. An Overview of Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) in Drug Discovery. *Expert Opinion on Drug Discovery*. Taylor & Francis October 3, 2017, pp 981–994. https://doi.org/10.1080/17460441.2017.1363734.

(79) Garrido Ruiz, D.; Sandoval-Perez, A.; Rangarajan, A. V.; Gunderson, E. L.; Jacobson, M. P. Cysteine Oxidation in Proteins: Structure, Biophysics, and Simulation. *Biochemistry* **2022**, *61* (20), 2165–2176. https://doi.org/10.1021/acs.biochem.2c00349.

(80) Hustoft, H. K.; Malerod, H.; Wilson, S. R.; Reubsaet, L.; Lundanes, E.; Greibrokk, T.; Hustoft, H. K.; Malerod, H.; Wilson, S. R.; Reubsaet, L.; Lundanes, E.; Greibrokk, T. A Critical Review of Trypsin Digestion for LC-MS Based Proteomics. *Integr. Proteomics* **2012**. https://doi.org/10.5772/29326.

(81) Steiner, R. F.; Albaugh, S.; Fenselau, C.; Murphy, C.; Vestling, M. A Mass Spectrometry Method for Mapping the Interface Topography of Interacting Proteins, Illustrated by the Melittin-Calmodulin System. *Anal. Biochem.* **1991**, *196* (1), 120–125. https://doi.org/10.1016/0003-2697(91)90127-F.

(82) Suckau, D.; Mak, M.; Przybylski, M. Protein Surface Topology-Probing by Selective Chemical Modification and Mass Spectrometric Peptide Mapping. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (12), 5630–5634. https://doi.org/10.1073/pnas.89.12.5630.

(83) Blumberg, S.; Vallee, B. L. Superactivation of Thermolysin by Acylation with Amino Acid N-Hydroxysuccinimide Esters. *Biochemistry* **1975**, *14* (11), 2410–2419. https://doi.org/10.1021/bi00682a022.

(84) Kalkhof, S.; Sinz, A. Chances and Pitfalls of Chemical Cross-Linking with Amine-Reactive N-Hydroxysuccinimide Esters. *Anal. Bioanal. Chem.* **2008**, *392* (1–2), 1–8. https://doi.org/10.1007/S00216-008-2231-5/FIGURES/4.

(85) Baslé, E.; Joubert, N.; Pucheault, M. Protein Chemical Modification on Endogenous Amino Acids. *Chemistry and Biology*. Cell Press March 26, 2010, pp 213–227. https://doi.org/10.1016/j.chembiol.2010.02.008.

(86) Fraenkel-Conrat, H.; Olcott, H. S. Esterification of Proteins with Alcohols of Low Molecular Weight. *J. Biol. Chem.* **1945**, *161* (1), 259–268. https://doi.org/10.1016/s0021-9258(17)41539-0.

(87) Gross, E.; Morell, J. L. Evidence for an Active Carboxyl Group in Pepsin. *J. Biol. Chem.* **1966**, *241* (15), 3638–3639. https://doi.org/10.1016/S0021-9258(18)99878-9.

(88) Zhang, H.; Wen, J.; Huang, R. Y.-C.; Blankenship, R. E.; Gross, M. L. Mass Spectrometry-Based Carboxyl Footprinting of Proteins: Method Evaluation. *Int. J. Mass Spectrom.* **2012**, *312*, 78–86. https://doi.org/10.1016/j.ijms.2011.07.015.

(89) Kaur, P.; Tomechko, S. E.; Kiselar, J.; Shi, W.; Deperalta, G.; Wecksler, A. T.; Gokulrangan, G.; Ling, V.; Chance, M. R. Characterizing Monoclonal Antibody Structure by Carboxyl Group Footprinting. *MAbs* **2015**, *7* (3), 540–552. https://doi.org/10.1080/19420862.2015.1023683.

(90) Bogan, A. A.; Thorn, K. S. Anatomy of Hot Spots in Protein Interfaces. *J. Mol. Biol.* **1998**, *280* (1), 1–9. https://doi.org/10.1006/jmbi.1998.1843.

(91) Hullán, L.; Szontagh, T.; Turtóczky, I.; Fedorcsák, I.; Hoffman, R. A.; Westerdahl, A. The Inactivation of Trypsin by Diethyl Pyrocarbonate. *Acta Chem. Scand.* **1965**, *19*, 2440–2441. https://doi.org/10.3891/acta.chem.scand.19-2440.

(92) Glocker, M. O.; Kalkum, M.; Yamamoto, R.; Schreurs, J. Selective Biochemical Modification of Functional Residues in Recombinant Human Macrophage Colony-Stimulating Factor β (RhM-CSF β): Identification by Mass Spectrometry. *Biochemistry* **1996**, *35* (46), 14625–14633. https://doi.org/10.1021/bi961199o.

(93) Kirsch, Z. J.; Blake, J. M.; Huynh, U.; Agrohia, D. K.; Tremblay, C. Y.; Graban, E. M.; Vaughan, R. C.; Vachet, R. W. Membrane Protein Binding Interactions Studied in Live Cells via Diethylpyrocarbonate Covalent Labeling Mass Spectrometry. *Anal. Chem.* **2023**, *95* (18), 7178–7185. https://doi.org/10.1021/acs.analchem.2c05616.

(94) Tremblay, C. Y.; Kirsch, Z. J.; Vachet, R. W. Complementary Structural Information for Antibody-Antigen Complexes from Hydrogen-Deuterium Exchange and Covalent Labeling Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2022**, *33* (7), 1303–1314. https://doi.org/10.1021/jasms.2c00108.

(95) Sokolovsky, M.; Riordan, J. F.; Vallee, B. L. Tetranitromethane. A Reagent for the Nitration of Tyrosyl Residues in Proteins. *Biochemistry* **1966**, *5* (11), 3582–3589. https://doi.org/10.1021/bi00875a029.

(96) Fenton, H. J. H. LXXIII.—Oxidation of Tartaric Acid in Presence of Iron. *J. Chem. Soc., Trans.* **1894**, *65*, 899–910. https://doi.org/10.1039/CT8946500899.

(97) Koppenol, W. H. The Haber-Weiss Cycle - 70 Years Later. *Redox Report*. Taylor & Francis 2001, pp 229–234. https://doi.org/10.1179/135100001101536373.

(98) Jain, S. S.; Tullius, T. D. Footprinting Protein–DNA Complexes Using the Hydroxyl Radical. *Nat. Protoc.* **2008**, *3* (6), 1092–1100. https://doi.org/10.1038/nprot.2008.72.

(99) Xu, G.; Chance, M. R. Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chemical Reviews*. American Chemical Society August 2007, pp 3514–3543. https://doi.org/10.1021/cr0682047.

(100) Tullius, T. D.; Greenbaum, J. A. Mapping Nucleic Acid Structure by Hydroxyl Radical Cleavage. *Curr. Opin. Chem. Biol.* **2005**, *9* (2), 127–134. https://doi.org/10.1016/j.cbpa.2005.02.009.

(101) Balasubramanian, B.; Pogozelski, W. K.; Tullius, T. D. DNA Strand Breaking by the Hydroxyl Radical Is Governed by the Accessible Surface Areas of the Hydrogen Atoms of the DNA Backbone. *Proc. Natl. Acad. Sci.* **1998**, *95* (17), 9738–9743. https://doi.org/10.1073/pnas.95.17.9738.

(102) Hao, Y.; Bohon, J.; Hulscher, R.; Rappé, M. C.; Gupta, S.; Adilakshmi, T.; Woodson, S. A. Time-Resolved Hydroxyl Radical Footprinting of RNA with X-Rays. *Curr. Protoc. Nucleic Acid Chem.* **2018**, *73* (1), e52. https://doi.org/10.1002/cpnc.52.

(103) Sclavi, B.; Woodson, S.; Sullivan, M.; Chance, M. R.; Brenowitz, M. Time-Resolved Synchrotron X-Ray "Footprinting", a New Approach to the Study of Nucleic Acid Structure and Function: Application to Protein-DNA Interactions and RNA Folding. *J. Mol. Biol.* **1997**, *266* (1), 144–159.

https://doi.org/10.1006/JMBI.1996.0775.

(104) Ingle, S.; Azad, R. N.; Jain, S. S.; Tullius, T. D. Chemical Probing of RNA with the Hydroxyl Radical at Single-Atom Resolution. *Nucleic Acids Res.* **2014**, *42* (20), 12758–12767. https://doi.org/10.1093/nar/gku934.

(105) Shcherbakova, I. Fast Fenton Footprinting: A Laboratory-Based Method for the Time-Resolved Analysis of DNA, RNA and Proteins. *Nucleic Acids Res.* **2006**, *34* (6), 48–57. https://doi.org/10.1093/nar/gkl055.

(106) Shaytan, A. K.; Xiao, H.; Armeev, G. A.; Gaykalova, D. A.; Komarova, G. A.; Wu, C.; Studitsky, V. M.; Landsman, D.; Panchenko, A. R. Structural Interpretation of DNA–Protein Hydroxyl-Radical Footprinting Experiments with High Resolution Using HYDROID. *Nat. Protoc.* **2018**, *13* (11), 2535–2556. https://doi.org/10.1038/s41596-018-0048-z.

(107) Sharp, J. S.; Becker, J. M.; Hettich, R. L. Protein Surface Mapping by Chemical Oxidation: Structural Analysis by Mass Spectrometry. *Anal. Biochem.* **2003**, *313* (2), 216–225. https://doi.org/10.1016/S0003-2697(02)00612-7.

(108) Tullius, T. D.; Dombroski, B. A. Hydroxyl Radical "Footprinting": High-Resolution Information about DNA-Protein Contacts and Application to λ Repressor and Cro Protein. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83* (15), 5469–5473. https://doi.org/10.1073/pnas.83.15.5469.

(109) Maleknia, S. D.; Ralston, C. Y.; Brenowitz, M. D.; Downard, K. M.; Chance, M. R. Determination of Macromolecular Folding and Structure by Synchrotron X-Ray Radiolysis Techniques. *Analytical Biochemistry*. Academic Press February 15, 2001, pp 103–115. https://doi.org/10.1006/abio.2000.4910.

(110) Sclavi, B.; Sullivan, M.; Chance, M. R.; Brenowitz, M.; Woodson, S. A. RNA Folding at Millisecond Intervals by Synchrotron Hydroxyl Radical Footprinting. *Science (80-.).* **1998**, *279* (5358), 1940–1943. https://doi.org/10.1126/science.279.5358.1940.

(111) Kiselar, J. G.; Maleknia, S. D.; Sullivan, M.; Downard, K. M.; Chance, M. R. Hydroxyl Radical Probe of Protein Surfaces Using Synchrotron X-Ray Radiolysis and Mass Spectrometry. *Int. J. Radiat. Biol.* **2002**, *78* (2), 101–114. https://doi.org/10.1080/09553000110094805.

(112) Kamal, J. K. A.; Benchaar, S. A.; Takamoto, K.; Reisler, E.; Chance, M. R. Three-Dimensional Structure of Cofilin Bound to Monomeric Actin Derived by Structural Mass Spectrometry Data. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (19), 7910–7915. https://doi.org/10.1073/pnas.0611283104.

(113) Guan, J.-Q.; Vorobiev, S.; Almo, S. C.; Chance, M. R. Mapping the G-Actin Binding Surface of Cofilin Using Synchrotron Protein Footprinting. *Biochemistry* **2002**, *41* (18), 5765–5775. https://doi.org/10.1021/bi0121104.

(114) Gupta, S.; Bavro, V. N.; D'Mello, R.; Tucker, S. J.; Vénien-Bryan, C.; Chance, M. R. Conformational Changes during the Gating of a Potassium Channel Revealed by Structural Mass Spectrometry. *Structure* **2010**, *18* (7), 839–846. https://doi.org/10.1016/j.str.2010.04.012.

(115) Kiselar, J. G.; Mahaffy, R.; Pollard, T. D.; Almo, S. C.; Chance, M. R. Visualizing Arp2/3 Complex Activation Mediated by Binding of ATP and WASp Using Structural Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (5), 1552–1557. https://doi.org/10.1073/pnas.0605380104.

(116) Kiselar, J. G.; Janmey, P. A.; Almo, S. C.; Chance, M. R. Visualizing the Ca2+-Dependent Activation of Gelsolin by Using Synchrotron Footprinting. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (7), 3942–3947. https://doi.org/10.1073/pnas.0736004100.

(117) Kiselar, J. G.; Janmey, P. A.; Almo, S. C.; Chance, M. R. Structural Analysis of Gelsolin Using Synchrotron Protein Footprinting. *Mol. Cell. Proteomics* **2003**, *2* (10), 1120–1132. https://doi.org/10.1074/mcp.M300068-MCP200.

(118) Hayes, J. J.; Kam, L.; Tullius, T. D. Footprinting Protein-DNA Complexes with γ-Rays; 1990; pp 545–549. https://doi.org/10.1016/0076-6879(90)86148-O.

(119) Henner, W. D.; Rodriguez, L. O.; Hecht, S. M.; Haseltine, W. A. Gamma Ray Induced Deoxyribonucleic Acid Strand Breaks. 3' Glycolate Termini. *J. Biol. Chem.* **1983**, *258* (2), 711–713. https://doi.org/10.1016/s0021-9258(18)33104-1.

(120) Watson, C.; Janik, I.; Zhuang, T.; Charvátová, O.; Woods, R. J.; Sharp, J. S. Pulsed Electron Beam Water Radiolysis for Submicrosecond Hydroxyl Radical Protein Footprinting. *Anal. Chem.* **2009**, *81* (7), 2496–2505. https://doi.org/10.1021/ac802252y.

(121) Li, K. S.; Shi, L.; Gross, M. L. Mass Spectrometry-Based Fast Photochemical Oxidation of Proteins (FPOP) for Higher Order Structure Characterization. *Acc. Chem. Res.* **2018**, *51* (3), 736–744. https://doi.org/10.1021/acs.accounts.7b00593.

(122) Sharp, J. S.; Becker, J. M.; Hettich, R. L. Analysis of Protein Solvent Accessible Surfaces by Photochemical Oxidation and Mass Spectrometry. *Anal. Chem.* **2004**, *76* (3), 672–683. https://doi.org/10.1021/ac0302004.

(123) Aye, T. T.; Low, T. Y.; Sze, S. K. Nanosecond Laser-Induced Photochemical Oxidation Method for Protein Surface Mapping with Mass Spectrometry. *Anal. Chem.* **2005**, *77* (18), 5814–5822. https://doi.org/10.1021/ac050353m.

(124) Hambly, D. M.; Gross, M. L. Laser Flash Photolysis of Hydrogen Peroxide to Oxidize Protein Solvent-Accessible Residues on the Microsecond Timescale. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (12), 2057–2063. https://doi.org/10.1016/j.jasms.2005.09.008.

(125) Vahidi, S.; Konermann, L. Probing the Time Scale of FPOP (Fast Photochemical Oxidation of Proteins): Radical Reactions Extend Over Tens of Milliseconds. *J. Am. Soc. Mass Spectrom.* **2016**, *27* (7), 1156–1164. https://doi.org/10.1007/s13361-016-1389-x.

(126) Niu, B.; Gross, M. L. MS-Based Hydroxyl Radical Footprinting: Methodology and Application of Fast Photochemical Oxidation of Proteins (FPOP). In *Mass Spectrometry-Based Chemical Proteomics*; Wiley, 2019; pp 363–416. https://doi.org/10.1002/9781118970195.ch15.

(127) Xu, G.; Takamoto, K.; Chance, M. R. Radiolytic Modification of Basic Amino Acid Residues in Peptides: Probes for Examining Protein−Protein Interactions. *Anal. Chem.* **2003**, *75* (24), 6995–7007. https://doi.org/10.1021/ac035104h.

(128) Xu, G.; Chance, M. R. Radiolytic Modification of Acidic Amino Acid Residues in Peptides: Probes for Examining Protein−Protein Interactions. *Anal. Chem.* **2004**, *76* (5), 1213–1221. https://doi.org/10.1021/ac035422g.

(129) Xu, G.; Chance, M. R. Radiolytic Modification and Reactivity of Amino Acid Residues Serving as Structural Probes for Protein Footprinting. *Anal. Chem.* **2005**, *77* (14), 4549–4555. https://doi.org/10.1021/ac050299+.

(130) Zhang, B.; Cheng, M.; Rempel, D.; Gross, M. L. Implementing Fast Photochemical Oxidation of Proteins (FPOP) as a Footprinting Approach to Solve Diverse Problems in Structural Biology. *Methods*. Academic Press July 15, 2018, pp 94–103. https://doi.org/10.1016/j.ymeth.2018.05.016.

(131) Liu, X. R.; Rempel, D. L.; Gross, M. L. Protein Higher-Order-Structure Determination by Fast Photochemical Oxidation of Proteins and Mass Spectrometry Analysis. *Nat. Protoc.* **2020**, *15* (12), 3942–3970. https://doi.org/10.1038/s41596-020-0396-3.

(132) Minkoff, B. B.; Blatz, J. M.; Choudhury, F. A.; Benjamin, D.; Shohet, J. L.; Sussman, M. R. Plasma-Generated OH Radical Production for Analyzing Three-Dimensional Structure in Protein Therapeutics. *Sci. Reports 2017 71* **2017**, *7* (1), 1–9. https://doi.org/10.1038/s41598-017-13371-7.

(133) Spotheim-Maurizot, M.; Charlier, M.; Sabattier, R. DNA Radiolysis by Fast Neutrons. *Int. J. Radiat. Biol.* **1990**, *57* (2), 301–313. https://doi.org/10.1080/09553009014552421.

(134) Coddington, J. W.; Hurst, J. K.; Lymar, S. V. Hydroxyl Radical Formation during Peroxynitrous Acid Decomposition. *J. Am. Chem. Soc.* **1999**, *121* (11), 2438–2443. https://doi.org/10.1021/ja982887t.

(135) Borotto, N. B.; Richards, T. K. Rapid Online Oxidation of Proteins and Peptides via Electrospray-Accelerated Ozonation. *J. Am. Soc. Mass Spectrom.* **2022**, *33* (11), 2078–2086. https://doi.org/10.1021/jasms.2c00182.

(136) Gau, B. C.; Chen, H.; Zhang, Y.; Gross, M. L. Sulfate Radical Anion as a New Reagent for Fast Photochemical Oxidation of Proteins. *Anal. Chem.* **2010**, *82* (18), 7821–7827. https://doi.org/10.1021/ac101760y.

(137) Zhang, M. M.; Rempel, D. L.; Gross, M. L. A Fast Photochemical Oxidation of Proteins (FPOP) Platform for Free-Radical Reactions: The Carbonate Radical Anion with Peptides and Proteins. *Free Radic. Biol. Med.* **2019**, *131*, 126–132. https://doi.org/10.1016/j.freeradbiomed.2018.11.031.

(138) Richards, F. M.; Lamed, R.; Wynn, R.; Patel, D.; Olack, G. Methylene as a Possible Universal Footprinting Reagent That Will Include Hydrophobic Surface Areas: Overview and Feasibility: Properties of Diazirine as a Precursor. *Protein Sci.* **2000**, *9* (12), 2506–2517. https://doi.org/10.1110/ps.9.12.2506.

(139) Jumper, C. C.; Schriemer, D. C. Mass Spectrometry of Laser-Initiated Carbene Reactions for Protein Topographic Analysis. *Anal. Chem.* **2011**, *83* (8), 2913–2920. https://doi.org/10.1021/ac102655f.

(140) Zhang, B.; Rempel, D. L.; Gross, M. L. Protein Footprinting by Carbenes on a Fast Photochemical Oxidation of Proteins (FPOP) Platform. *J. Am. Soc. Mass Spectrom.* **2016**, *27* (3), 552–555. https://doi.org/10.1007/s13361-015-1313-9.

(141) Manzi, L.; Barrow, A. S.; Hopper, J. T. S.; Kaminska, R.; Kleanthous, C.; Robinson, C. V.; Moses, J. E.; Oldham, N. J. Carbene Footprinting Reveals Binding Interfaces of a Multimeric Membrane-Spanning Protein. *Angew. Chemie Int. Ed.* **2017**, *56* (47), 14873–14877. https://doi.org/10.1002/anie.201708254.

(142) Kessler, J.; Obinger, C.; Eales, G. Factors Influencing the Study of Peroxidase-Generated Iodine Species and Implications for Thyroglobulin Synthesis. *Thyroid* **2008**, *18* (7), 769–774. https://doi.org/10.1089/thy.2007.0310.

(143) Chen, J.; Cui, W.; Giblin, D.; Gross, M. L. New Protein Footprinting: Fast Photochemical Iodination Combined with Top-Down and Bottom-Up Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (8), 1306–1318. https://doi.org/10.1007/s13361-012-0403-1.

(144) Müller, K.; Faeh, C.; Diederich, F. Fluorine in Pharmaceuticals: Looking beyond Intuition. *Science*. American Association for the Advancement of Science September 28, 2007, pp 1881–1886. https://doi.org/10.1126/science.1131943.

(145) Cheng, M.; Zhang, B.; Cui, W.; Gross, M. L. Laser-Initiated Radical Trifluoromethylation of Peptides and Proteins: Application to Mass-Spectrometry-Based Protein Footprinting. *Angew. Chemie - Int. Ed.* **2017**, *56* (45), 14007–14010. https://doi.org/10.1002/anie.201706697.

(146) Fojtík, L.; Fiala, J.; Pompach, P.; Chmelík, J.; Matoušek, V.; Beier, P.; Kukačka, Z.; Novák, P. Fast Fluoroalkylation of Proteins Uncovers the Structure and Dynamics of Biological Macromolecules. *J. Am. Chem. Soc.* **2021**, *143* (49), 20670–20679. https://doi.org/10.1021/jacs.1c07771.

(147) Yassaghi, G.; Kukačka, Z.; Fiala, J.; Kavan, D.; Halada, P.; Volný, M.; Novák, P. Top-Down Detection of Oxidative Protein Footprinting by Collision-Induced Dissociation, Electron-Transfer Dissociation, and Electron-Capture Dissociation. *Anal. Chem.* **2022**, *94* (28), 9993–10002. https://doi.org/10.1021/ACS.ANALCHEM.1C05476.

(148) Vahidi, S.; Stocks, B. B.; Liaghati-Mobarhan, Y.; Konermann, L. Mapping PH-Induced Protein Structural Changes Under Equilibrium Conditions by Pulsed Oxidative Labeling and Mass Spectrometry. *Anal. Chem.* **2012**, *84* (21), 9124–9130. https://doi.org/10.1021/ac302393g.

(149) Jones, L. M.; B. Sperry, J.; A. Carroll, J.; Gross, M. L. Fast Photochemical Oxidation of Proteins for Epitope Mapping. *Anal. Chem.* **2011**, *83* (20), 7657–7661. https://doi.org/10.1021/ac2007366.

(150) Cornwell, O.; Radford, S. E.; Ashcroft, A. E.; Ault, J. R. Comparing Hydrogen Deuterium Exchange and Fast Photochemical Oxidation of Proteins: A Structural Characterisation of Wild-Type and ΔN6 B2-Microglobulin. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (12), 2413–2426. https://doi.org/10.1007/s13361-018-2067-y.

(151) Watkinson, T. G.; Calabrese, A. N.; Ault, J. R.; Radford, S. E.; Ashcroft, A. E. FPOP-LC-MS/MS Suggests Differences in Interaction Sites of Amphipols and Detergents with Outer Membrane Proteins. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (1), 50–55. https://doi.org/10.1007/s13361-016-1421-1.

(152) Yan, Y.; Chen, G.; Wei, H.; Huang, R. Y.-C.; Mo, J.; Rempel, D. L.; Tymiak, A. A.; Gross, M. L. Fast Photochemical Oxidation of Proteins (FPOP) Maps the Epitope of EGFR Binding to Adnectin. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (12), 2084–2092. https://doi.org/10.1007/s13361-014-0993-x.

(153) Li, J.; Wei, H.; Krystek, S. R.; Bond, D.; Brender, T. M.; Cohen, D.; Feiner, J.; Hamacher, N.; Harshman, J.; Huang, R. Y. C.; Julien, S. H.; Lin, Z.; Moore, K.; Mueller, L.; Noriega, C.; Sejwal, P.; Sheppard, P.; Stevens, B.; Chen, G.; Tymiak, A. A.; Gross, M. L.; Schneeweis, L. A. Mapping the Energetic Epitope of an Antibody/Interleukin-23 Interaction with Hydrogen/Deuterium Exchange, Fast Photochemical Oxidation of Proteins Mass Spectrometry, and Alanine Shave Mutagenesis. *Anal. Chem.* **2017**, *89* (4), 2250–2258. https://doi.org/10.1021/acs.analchem.6b03058.

(154) Chen, J.; Rempel, D. L.; Gross, M. L. Temperature Jump and Fast Photochemical Oxidation Probe Submillisecond Protein Folding. *J. Am. Chem. Soc.* **2010**, *132* (44), 15502–15504. https://doi.org/10.1021/ja106518d.

(155) Pogozelski, W. K.; Tullius, T. D. Oxidative Strand Scission of Nucleic Acids: Routes Initiated by Hydrogen Abstraction from the Sugar Moiety. *Chem. Rev.* **1998**, *98* (3), 1089–1107. https://doi.org/10.1021/cr960437i.

(156) Valis, K.; Prochazka, L.; Boura, E.; Chladova, J.; Obsil, T.; Rohlena, J.; Truksa, J.; Dong, L.-F.; Ralph, S. J.; Neuzil, J. Hippo/Mst1 Stimulates Transcription of the Proapoptotic Mediator NOXA in a FoxO1-Dependent Manner. *Cancer Res.* **2011**, *71* (3), 946–954. https://doi.org/10.1158/0008-5472.CAN-10-2203.

(157) Boura, E.; Rezabkova, L.; Brynda, J.; Obsilova, V.; Obsil, T. Structure of the Human FOXO4-DBD–DNA Complex at 1.9 Å Resolution Reveals New Details of FOXO Binding to the DNA. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (12), 1351–1357. https://doi.org/10.1107/S0907444910042228.

(158) Slavata; Chmelík; Kavan; Filandrová; Fiala; Rosůlek; Mrázek; Kukačka; Vališ; Man; Miller; McIntyre; Fabris; Novák. MS-Based Approaches Enable the Structural Characterization of Transcription Factor/DNA Response Element Complex. *Biomolecules* **2019**, *9* (10), 535. https://doi.org/10.3390/biom9100535.

(159) Obsilova, V.; Vecer, J.; Herman, P.; Pabianova, A.; Sulc, M.; Teisinger, J.; Boura, E.; Obsil, T. 14-3-3 Protein Interacts with Nuclear Localization Sequence of Forkhead Transcription Factor FoxO4. *Biochemistry* **2005**, *44*

(34), 11608–11617. https://doi.org/10.1021/bi050618r.

(160) Boura, E.; Silhan, J.; Herman, P.; Vecer, J.; Sulc, M.; Teisinger, J.; Obsilova, V.; Obsil, T. Both the N-Terminal Loop and Wing W2 of the Forkhead Domain of Transcription Factor Foxo4 Are Important for DNA Binding. *J. Biol. Chem.* **2007**, *282* (11), 8265–8275. https://doi.org/10.1074/jbc.M605682200.

(161) Vacha, P.; Zuskova, I.; Bumba, L.; Herman, P.; Vecer, J.; Obsilova, V.; Obsil, T. Detailed Kinetic Analysis of the Interaction between the FOXO4–DNA-Binding Domain and DNA. *Biophys. Chem.* **2013**, *184*, 68–78. https://doi.org/10.1016/J.BPC.2013.09.002.

(162) Donnelly, D. P.; Rawlins, C. M.; DeHart, C. J.; Fornelli, L.; Schachner, L. F.; Lin, Z.; Lippens, J. L.; Aluri, K. C.; Sarin, R.; Chen, B.; Lantz, C.; Jung, W.; Johnson, K. R.; Koller, A.; Wolff, J. J.; Campuzano, I. D. G.; Auclair, J. R.; Ivanov, A. R.; Whitelegge, J. P.; Paša-Tolić, L.; Chamot-Rooke, J.; Danis, P. O.; Smith, L. M.; Tsybin, Y. O.; Loo, J. A.; Ge, Y.; Kelleher, N. L.; Agar, J. N. Best Practices and Benchmarks for Intact Protein Analysis for Top-down Mass Spectrometry. *Nat. Methods* **2019**, *16* (7), 587–594. https://doi.org/10.1038/s41592-019-0457-0.

(163) Compton, P. D.; Zamdborg, L.; Thomas, P. M.; Kelleher, N. L. On the Scalability and Requirements of Whole Protein Mass Spectrometry. *Anal. Chem.* **2011**, *83* (17), 6868–6874. https://doi.org/10.1021/ac2010795.

(164) Marshall, A. G.; Senko, M. W.; Li, W.; Li, M.; Dillon, S.; Guan, S.; Logan, T. M. Protein Molecular Mass to 1 Da by 13 C, 15 N Double-Depletion and FT-ICR Mass Spectrometry. *J. Am. Chem. Soc.* **1997**, *119* (2), 433–434. https://doi.org/10.1021/ja9630046.

(165) Going, C. C.; Williams, E. R. Supercharging with m -Nitrobenzyl Alcohol and Propylene Carbonate: Forming Highly Charged Ions with Extended, Near-Linear Conformations. *Anal. Chem.* **2015**, *87* (7), 3973–3980. https://doi.org/10.1021/acs.analchem.5b00071.

(166) Juetten, K. J.; Brodbelt, J. S. Top-Down Analysis of Supercharged Proteins Using Collision-, Electron-, and Photon-Based Activation Methods. *J. Am. Soc. Mass Spectrom.* **2023**, *34* (7), 1467–1476. https://doi.org/10.1021/jasms.3c00138.

(167) Jertz, R.; Friedrich, J.; Kriete, C.; Nikolaev, E. N.; Baykut, G. Tracking the Magnetron Motion in FT-ICR Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (8), 1349–1366. https://doi.org/10.1007/s13361-015-1148-4.

(168) Srikanth, R.; Wilson, J.; Vachet, R. W. Correct Identification of Oxidized Histidine Residues Using Electron-Transfer Dissociation. *J. Mass Spectrom.* **2009**, *44* (5), 755–762. https://doi.org/10.1002/JMS.1552.

(169) Ottinger, L. M.; Tullius, T. D. High-Resolution in Vivo Footprinting of a Protein−DNA Complex Using γ-Radiation. *J. Am. Chem. Soc.* **2000**, *122* (24), 5901–5902. https://doi.org/10.1021/ja000285f.

(170) Azad, R. N.; Zafiropoulos, D.; Ober, D.; Jiang, Y.; Chiu, T. P.; Sagendorf, J. M.; Rohs, R.; Tullius, T. D. Experimental Maps of DNA Structure at

Nucleotide Resolution Distinguish Intrinsic from Protein-Induced DNA Deformations. *Nucleic Acids Res.* **2018**, *46* (5), 2637–2647. https://doi.org/10.1093/nar/gky033.

(171) Basiri, B.; van Hattum, H.; van Dongen, W. D.; Murph, M. M.; Bartlett, M. G. The Role of Fluorinated Alcohols as Mobile Phase Modifiers for LC-MS Analysis of Oligonucleotides. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (1), 190–199. https://doi.org/10.1007/s13361-016-1500-3.

(172) Liu, A.; Cheng, M.; Zhou, Y.; Deng, P. Bioanalysis of Oligonucleotide by LC–MS: Effects of Ion Pairing Regents and Recent Advances in Ion-Pairing-Free Analytical Strategies. *Int. J. Mol. Sci.* **2022**, *23* (24), 15474. https://doi.org/10.3390/ijms232415474.

(173) Gao, Y.; McLuckey, S. A. Fragmentation Reactions of Nucleic Acid Ions in the Gas Phase. In *Physical Chemistry in Action*; 2014; pp 131–182. https://doi.org/10.1007/978-3-642-54842-0_6.

(174) Kappen, L. S.; Goldberg, I. H. Deoxyribonucleic Acid Damage by Neocarzinostatin Chromophore: Strand Breaks Generated by Selective Oxidation of C-5' of Deoxyribose. *Biochemistry* **1983**, *22* (21), 4872–4878. https://doi.org/10.1021/bi00290a002.

(175) Polák, M.; Černý, J.; Novák, P. Isotopic Depletion Increases the Spatial Resolution of FPOP Top-Down Mass Spectrometry Analysis. *Anal. Chem.* **2024**, *96* (4), 1478–1487. https://doi.org/10.1021/acs.analchem.3c03759.

(176) von Sonntag, C. *Free-Radical-Induced DNA Damage and Its Repair*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2006. https://doi.org/10.1007/3-540-30592-0.

(177) Kaltashov, I. A.; Bobst, C. E.; Pawlowski, J.; Wang, G. Mass Spectrometry-Based Methods in Characterization of the Higher Order Structure of Protein Therapeutics. *J. Pharm. Biomed. Anal.* **2020**, *184*, 113169. https://doi.org/10.1016/j.jpba.2020.113169.

(178) Largy, E.; König, A.; Ghosh, A.; Ghosh, D.; Benabou, S.; Rosu, F.; Gabelica, V. Mass Spectrometry of Nucleic Acid Noncovalent Complexes. *Chemical Reviews*. American Chemical Society April 27, 2022, pp 7720–7839. https://doi.org/10.1021/acs.chemrev.1c00386.

(179) Egli, M.; Manoharan, M. Chemistry, Structure and Function of Approved Oligonucleotide Therapeutics. *Nucleic Acids Res.* **2023**, *51* (6), 2529–2573. https://doi.org/10.1093/nar/gkad067.

(180) Moumné, L.; Marie, A. C.; Crouvezier, N. Oligonucleotide Therapeutics: From Discovery and Development to Patentability. *Pharmaceutics*. Multidisciplinary Digital Publishing Institute January 22, 2022, p 260. https://doi.org/10.3390/pharmaceutics14020260.

(181) Ogilby, P. R. Singlet Oxygen: There Is Indeed Something New under the Sun. *Chem. Soc. Rev.* **2010**, *39* (8), 3181. https://doi.org/10.1039/b926014p.

(182) Auten, R. L.; Davis, J. M. Oxygen Toxicity and Reactive Oxygen Species: The Devil Is in the Details. *Pediatr. Res. 2009 662* **2009**, *66* (2), 121–127. https://doi.org/10.1203/pdr.0b013e3181a9eafb.

(183) Wilkinson, F.; Helman, W. P.; Ross, A. B. Rate Constants for the Decay and Reactions of the Lowest Electronically Excited Singlet State of Molecular Oxygen in Solution. An Expanded and Revised Compilation. *J. Phys. Chem. Ref. Data* **1995**, *24* (2), 663–677. https://doi.org/10.1063/1.555965.

(184) Michaeli, A.; Feitelson, J. Reactivity of Singlet Oxygen toward Amino Acids and Peptides. *Photochem. Photobiol.* **1994**, *59* (3), 284–289. https://doi.org/10.1111/j.1751-1097.1994.tb05035.x.

(185) Di Mascio, P.; Martinez, G. R.; Miyamoto, S.; Ronsein, G. E.; Medeiros, M. H. G.; Cadet, J. Singlet Molecular Oxygen Reactions with Nucleic Acids, Lipids, and Proteins. *Chem. Rev.* **2019**, *119* (3), 2043–2086. https://doi.org/10.1021/acs.chemrev.8b00554.

(186) Davies, M. J. Reactive Species Formed on Proteins Exposed to Singlet Oxygen. *Photochem. Photobiol. Sci.* **2004**, *3* (1), 17–25. https://doi.org/10.1039/b307576c.

(187) Gracanin, M.; Hawkins, C. L.; Pattison, D. I.; Davies, M. J. Singlet-Oxygen-Mediated Amino Acid and Protein Oxidation: Formation of Tryptophan Peroxides and Decomposition Products. *Free Radic. Biol. Med.* **2009**, *47* (1), 92–102. https://doi.org/10.1016/j.freeradbiomed.2009.04.015.

(188) O'Connor, A. E.; Gallagher, W. M.; Byrne, A. T. Porphyrin and Nonporphyrin Photosensitizers in Oncology: Preclinical and Clinical Advances in Photodynamic Therapy. *Photochem. Photobiol.* **2009**, *85* (5), 1053–1074. https://doi.org/10.1111/J.1751-1097.2009.00585.X.

(189) Gunaydin, G.; Gedik, M. E.; Ayan, S. Photodynamic Therapy for the Treatment and Diagnosis of Cancer–A Review of the Current Clinical Status. *Front. Chem.* **2021**, *9*, 686303. https://doi.org/10.3389/fchem.2021.686303.

(190) Van den Berg, P. A. W.; Feenstra, K. A.; Mark, A. E.; Berendsen, H. J. C.; Visser, A. J. W. G. Dynamic Conformations of Flavin Adenine Dinucleotide: Simulated Molecular Dynamics of the Flavin Cofactor Related to the Time-Resolved Fluorescence Characteristics. *J. Phys. Chem. B* **2002**, *106* (34), 8858–8869. https://doi.org/10.1021/jp020356s.

(191) Petrenčáková, M.; Filandr, F.; Hovan, A.; Yassaghi, G.; Man, P.; Kožár, T.; Schwer, M.-S.; Jancura, D.; Plückthun, A.; Novák, P.; Miškovský, P.; Bánó, G.; Sedlák, E. Photoinduced Damage of AsLOV2 Domain Is Accompanied by Increased Singlet Oxygen Production Due to Flavin Dissociation. *Sci. Rep.* **2020**, *10* (1), 4119. https://doi.org/10.1038/s41598-020-60861-2.

DECLARATION OF AUTHORSHIP

I. **Polák, M.**; Yassaghi, G.; Kavan, D.; Filandr, F.; Fiala, J.; Kukačka, Z.; Halada, P.; Loginov, D. S.; Novák, P. Utilization of Fast Photochemical Oxidation of Proteins and Both Bottom-up and Top-down Mass Spectrometry for Structural Characterization of a Transcription Factor–DsDNA Complex. Anal. Chem. 2022, 94 (7), 3203–3210. doi:10.1021/acs.analchem.1c04746.

I declare that Marek Polák contributed~60 % to **publication I**. He was responsible for protein expression and purification, oxidizing them, analyzing obtained datasets, and assisting in the debugging software for data interpretation. Marek also wrote the first draft of the article.

II. **Polák, M.**; Černý, J.; Novák, P. Isotopic Depletion Increases the Spatial Resolution of FPOP Top-Down Mass Spectrometry Analysis. Anal. Chem. 2024, 96 (4), 1478–1487. doi:10.1021/acs.analchem.3c03759.

I declare that Marek Polák contributed~90 % to **publication II**. He expressed and purified protein samples, oxidized them, analyzed obtained datasets, and wrote and prepared the manuscript.

III. **Polák, M.**; Brinsa, V.; Černý, J.; Kavan, D.; Volný, M.; Kádek, A.; Novák., Hydroxyl Radical Footprinting of Nucleic Acids Coupled to High-Resolution Mass Spectrometry Analysis, *Nucleic Acids Res.*, ***manuscript in preparation.***

I declare that Marek Polák contributed~70 % to **publication III**. He prepared DNA and protein samples, oxidized them, analyzed obtained datasets by LC-MS, and wrote and prepared the manuscript.

IV. Felčíková, K.; Hovan, A.; **Polák, M.**; Loginov, D.S.; Holotová, V.; Díaz, C.; Kožár, T.; Lee, O.; Varhač, R.; Novák, P., Bánó, G.; Sedlák, E. Design of AsLOV2 domain as a carrier of light-induced dissociable FMN photosensitizer. Prot. Sci., 2024, ***in press***. doi: 10.1002/pro.4921

I declare that Marek Polák contributed~15 % to **publication IV**. As a co-author, he analyzed proteins by top-down, prepared figures and contributed to the manuscript writing.

In Vestec 15.3.2024

RNDr. Petr Novák, Ph.D.

Attached Publications:

Publication I

Polák, M.; Yassaghi, G.; Kavan, D.; Filandr, F.; Fiala, J.; Kukačka, Z.; Halada, P.; Loginov, D. S.; Novák, P.

Utilization of Fast Photochemical Oxidation of Proteins and Both Bottom-up and Top-down Mass Spectrometry for Structural Characterization of a Transcription Factor–DsDNA Complex.

Analytical Chemistry, 2022

My contribution: Protein expression and purification, sample preparation, FPOP of samples, analyzing obtained data, manuscript writing

analytical chemistry

pubs.acs.org/ac

Article

Utilization of Fast Photochemical Oxidation of Proteins and Both Bottom-up and Top-down Mass Spectrometry for Structural Characterization of a Transcription Factor−dsDNA Complex

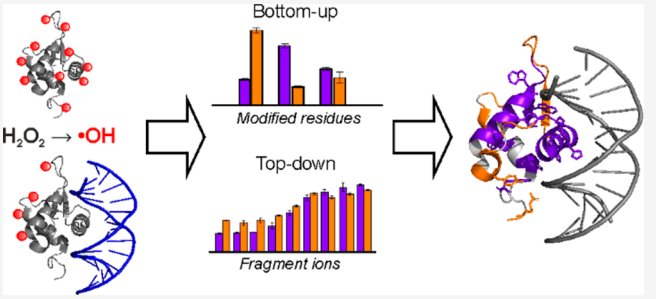
Marek Polák, Ghazaleh Yassaghi, Daniel Kavan, František Filandr, Jan Fiala, Zdeněk Kukačka, Petr Halada, Dmitry S. Loginov, and Petr Novák*

Cite This: *Anal. Chem.* 2022, 94, 3203−3210

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: A combination of covalent labeling techniques and mass spectrometry (MS) is currently a progressive approach for deriving insights related to the mapping of protein surfaces or protein−ligand interactions. In this study, we mapped an interaction interface between the DNA binding domain (DBD) of FOXO4 protein and the DNA binding element (DAF16) using fast photochemical oxidation of proteins (FPOP). Residues involved in protein−DNA interaction were identified using the bottom-up approach. To confirm the findings and avoid a misinterpretation of the obtained data, caused by possible multiple radical oxidations leading to the protein surface alteration and oxidation of deeply buried amino acid residues, a top-down approach was employed for the first time in FPOP analysis. An isolation of singly oxidized ions enabled their gas-phase separation from multiply oxidized species followed by CID and ECD fragmentation. Application of both fragmentation techniques allowed generation of complementary fragment sets, out of which the regions shielded in the presence of DNA were deduced. The findings obtained by bottom-up and top-down approaches were highly consistent. Finally, FPOP results were compared with those of the HDX study of the FOXO4-DBD·DAF16 complex. No contradictions were found between the methods. Moreover, their combination provides complementary information related to the structure and dynamics of the protein−DNA complex. Data are available via ProteomeXchange with identifier PXD027624.



Covalent labeling techniques coupled with mass spectrometry have become an alternative for analysis of protein structure and dynamics of protein−protein and protein−ligand interactions.[1] The techniques utilize reactive species to provide information concerning solvent accessibility to a protein surface.[2] The pioneering work that references chemical footprinting experiments was accomplished by Przybylski[3,4] and subsequently adopted for top-down analysis.[5] However, the spatial resolution was limited by the selective reactivity of amino acid side chains. To overcome this obstacle, hydroxyl radicals were introduced utilizing dissociation of water molecules by synchrotron irradiation.[6] Because the availability of synchrotron for footprinting experiments is rather limited, alternative methods have been successfully tested to generate hydroxyl radicals based on the photolysis of hydrogen peroxide.[7−9] Later, other radicals were utilized as well to access the solvent-accessible surface area of proteins.[10−13]

Fast photochemical oxidation of proteins (FPOP) developed in 2005 by Hambly and Gross[9] has been considered to be the most popular method.[14,15] FPOP utilizes hydroxyl radicals, commonly generated by an excimer laser from hydrogen peroxide, for modification of solvent-accessible residues on a protein surface.[16] Nowadays, FPOP is primarily performed in a quench-flow capillary system composed of syringes filled with a sample of interest, hydrogen peroxide, quencher, syringe pumps, capillary flow paths, and an excimer laser. During the sample flow, modifications of amino acid residues occur upon oxidation by generated hydroxyl radicals. Further, the reaction is quenched with different scavengers to avoid overoxidation of the sample and regulate the rate of oxidation.[17]

The range of amino acids reactive toward hydroxyl radicals is quite broad, with the most reactive residues being sulfur-containing cysteine and methionine, aromatic ring-containing tryptophan, phenylalanine, tyrosine, and to a lesser extent histidine, leucine, and isoleucine.[18,19] This gives FPOP the potential to be employed for various studies, including those involving protein conformational changes,[20] protein−protein interactions,[21] topology of membrane proteins,[22,23] or mapping

Received: November 2, 2021
Accepted: January 20, 2022
Published: February 8, 2022

https://doi.org/10.1021/acs.analchem.1c04746
Anal. Chem. 2022, 94, 3203−3210

of antibody epitopes.[24] Processes of protein folding/unfolding have also been described using FPOP.[25] Although hydroxyl radical footprinting has been originally applied to investigate protein−DNA interactions from the DNA perspective,[26] any studies utilizing FPOP for such a complex system have not been published so far.

Thus, in the present study, the DNA binding domain (DBD) of FOXO4 transcription factor with its binding element DAF16 was chosen as the model. FOXO4 with the three other members (FOXO1, -3, and -6) belongs to an "O" subclass of Forkhead box (FOX) proteins involved in the control of metabolism, cell survival, proliferation, and DNA damage repair response.[27] These proteins have been considered as potential drug targets, as they are able to induce cell cycle arrest and promote/initiate apoptosis.[28] All FOXO proteins share a conserved DNA-binding domain that follows H1-S1-H2-H3-S2-W1-S3-W2 topology[29] and recognize the consensus 5′-GTAAA(T/C)AA-3′ sequence binding element[30,31] and 5′-(C/A)(A/C)AAA(C/T)AA-3′ sequence known as DAF16 and an insulin-response element (IRE),[32] respectively.

The interaction of FOXO4-DBD along with DAF16 has been structurally well defined.[33−36] The main contact between the protein and DNA is mediated through helix H3 with the characteristic highly conserved sequence N148-X-X-R-H-X-X-S/T155 (numbering according to the FOXO4 sequence) by placing it in the major groove.

Here we describe a perspective to utilize FPOP analysis for the study of protein−DNA interactions. Analysis of the footprinted protein surface was performed by both the bottom-up classical approach and the top-down method, used for the first time for the FPOP analysis of a protein−dsDNA complex. Both results were combined and compared to a previously published X-ray structural model of FOXO4·DAF16[34] and the respective HDX study.[37]

EXPERIMENTAL SECTION

Materials and Chemicals. All solvents were purchased from Merck (Germany) in LC-MS grade. Unless stated otherwise, all chemicals were obtained from Sigma-Aldrich.

Protein Expression and Purification. The expression and purification of full length FOXO4-DBD were performed according to Slavata et al.[37] Final gel permeation chromatography was performed in helium-saturated 150 mM ammonium acetate, pH 6.8. Prior to the FPOP experiment, the protein was transferred using Zeba spin columns (7k MWCO; Thermo Fisher Scientific, Waltham, MA) into freshly degassed and helium-saturated ammonium acetate buffer (150 mM, pH 6.8).

FOXO4·DAF16 Complex Formation. Forward (5′-TTG GGT AAA CAA G-3′) and reverse (5′-CTT GTT TAC CCA A-3′) complementary strands were designed based on a previous study[31] and purchased from IDT (Coralville, IA). A stock solution of dsDAF16 was prepared by mixing both strands in an equimolar ratio in LC-MS water. The mixture was incubated at 95 °C for 1 min and cooled on the bench to form the duplex DNA. Equimolar amounts of the protein and DNA were mixed in 150 mM ammonium acetate, pH 6.8, and incubated at room temperature for 15 min to obtain FOXO4·DAF16 complex at 30 μM final concentration.

Fast Photochemical Oxidation of Proteins (FPOP). The FPOP labeling was performed in an in-house built quench flow setup, consisting of two syringe pumps (New Era Pump System Inc., Farmingdale, NY; models NE-1000 and NE-4000), three syringes (Hamilton, Reno, NV), and fused silica capillaries (Polymicro Technologies, Phoenix, AZ; Figure S1). The syringes were filled with the protein or protein·DNA complex (30 μM), H_2O_2 (30 mM), and the quencher (methionine, 75 mM), respectively, prepared in degassed 150 mM ammonium acetate, pH 6.8. Flow rates were set to 15 μL/min for the sample and H_2O_2, and 20 μL/min for the radical scavenger. The mixture of the sample and H_2O_2 was irradiated by an excimer laser (Compex 50 KrF, Coherent Inc., Salem, NH) while passing through the transparent window in a quartz capillary (i.d. 75 μm, 6.5 mm). A single shot was performed at a wavelength of 248 nm, frequency 15 Hz, energy 107 mJ, 20 ns pulse duration, and 2.24 mJ/cm^2 radiant exposure. The exclusion volume was 16%. The irradiated solution was immediately quenched by 75 mM methionine. One hundred microliters was collected per sample. Residual H_2O_2 was removed by adding 3000 U of catalase (Sigma-Aldrich, St. Louis, MO) and incubating at RT for 15 min.

Bottom-up Approach. Collected samples after the FPOP experiment were digested with LysC (Promega, Madison, WI) or a combination of Lys-C/trypsin (Promega) for 6 h at 37 °C at a protease:protein weight ratio of 1:20 with the extra addition of the respective protease (w/w 1:10) after the first 2 h of incubation. Released DAF16 was digested with 1 μL of Bal-31 endonuclease (0.5U; New England BioLabs, Ipswich, MA) for 1 h at 30 °C. The digestion was terminated by addition of TFA to the final concentration of 0.1%. All samples were desalted using reverse-phase microtrap column (Optimize Technologies, Oregon City, OR) prior to LC-MS analysis.

HPLC separation was performed using an Agilent 1200 series HPLC system (Agilent Technologies, Santa Clara, CA). The sample was loaded in a total volume of 5 μL that corresponded to 1 μg of the digested protein. The sample was injected onto a reverse-phase trap column (Luna Omega Polar C18, 0.3 × 30 mm, Phenomenex, Torrance, CA) followed by a reverse-phase analytical column (Luna Omega Polar C18, 0.3 × 150 mm, Phenomenex), both heated to 60 °C. Two mobile phases were used: A (2% acetonitrile, ACN; 0.1% formic acid, FA) and B (98% ACN, 0.1% FA). The flow rate was set to 10 μL/min with the following solution composition: 95% A, 5% B. The LC run consisted of a 35 min separation gradient of 5−40% of solvent B (98% ACN, 0.1% FA), a 3 min spike of 40−95% of solvent B, 3 min washing flow at 95% solvent B, a 1 min drop of 95−2% of solvent B, and equilibration of columns in 2% B for 10 min.

MSMS analysis was performed using a trapped ion mobility-quadrupole time-of-flight mass spectrometer (timsTOF Pro, Bruker Daltonics, Billerica, MA). The eluting peptides were interrogated by an MS acquisition method recording spectra from 100 to 1700 m/z, and ion mobility was scanned from 0.6 to 1.6 Vs/cm^2. The method consisted of a TIMS survey scan of 150 ms followed by six PASEF MS/MS scans, each 150 ms for ion accumulation and ramp time. Total cycle time was 1.08 s. Target intensity was 40 000, intensity threshold was 1000, and singly charged peptides with m/z < 800 were excluded by an inclusion/exclusion polygon filter applied within the ion mobility m/z heatmap. Precursors for data-dependent acquisition were fragmented with an ion mobility-dependent collision energy, which was linearly increased from 20 to 59 eV.

To quantify the extent of modification, exactly the same separation system was directly connected to the FT-ICR mass

https://doi.org/10.1021/acs.analchem.1c04746
Anal. Chem. 2022, 94, 3203−3210

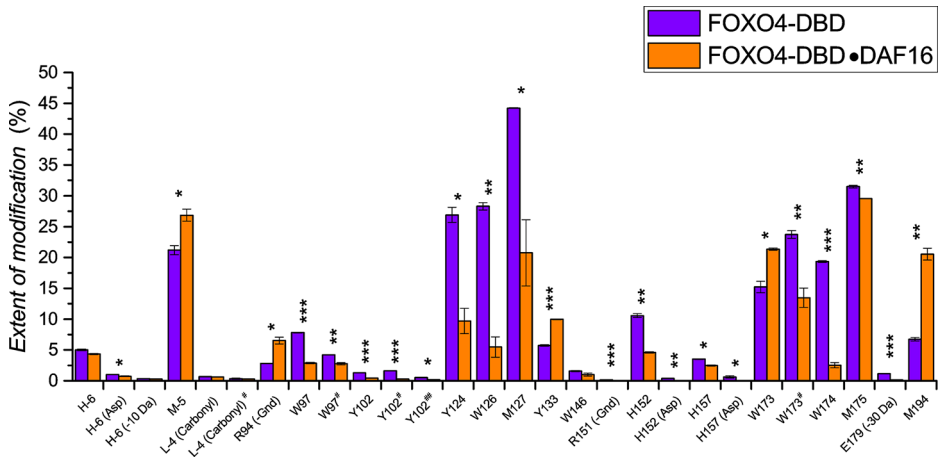


Figure 1. Quantified extents of residue modifications detected in LC-MSMS analysis. Purple: extents of peptide modification in the apo form, orange: extents of peptide modification in the holo form. Modifications other than +16 Da are indicated in brackets. Positional isomers separated in LC analysis are marked by #. The results of the *t* test are indicated by the number of asterisks above the histograms. *t* test legend: *($P \leq 0.05$), **($P \leq 0.01$), ***($P \leq 0.001$).

spectrometer. For LC-MS analysis, the acquisition m/z range was set to 200−2500 with 1 M data points transient, and ion accumulation of 0.2 s. Four scans were accumulated for each MS spectrum. Each sample was analyzed in triplicate.

LC-MSMS data were processed using Peaks X+ software (Bioinformatics Solutions Inc., Waterloo, ON, Canada) against the FOXO4-DBD sequence and FPOP modifications described in Loginov et al.[38] Negative values in the sequence of FOXO4-DBD refers to residues originating from the vector (Figure S2). Intensities of peptides were determined using DataAnalysis 5.0 (Bruker Daltonics). The extent of oxidation for the two charge states with the highest intensities was calculated as described previously.[16] The data were plotted and presented in the form of mean ± SD. Significant differences in the extent of modification were statistically analyzed by a two-tailed *t* test using GraphPad Prism 8 software. Residues with significant changes were highlighted in the crystal structure using PyMol 2.1.0 software (Schrödinger LLC, New York, NY).

Top-down Approach. To remove the oligonucleotides, urea and $MgCl_2$ were added to final concentrations of 4 M and 1 μM, respectively; 1 μL of Benzonase endonuclease (Sigma-Aldrich) was used to digest DNA at 30 °C for 30 min. Digestion was terminated by 0.1% FA. Protein samples were desalted on a reverse-phase microtrap column (Optimize Technologies) by 0.1% FA and eluted with 80% ACN/0.1% FA. Desalted protein was further diluted five times using 30% ACN/0.1% FA solution and sprayed using an nESI source in positive mode and a desolvatation temperature of 120 °C. All MS analyses were performed on a solariX XR FT-ICR mass spectrometer equipped with 15T magnet (Bruker Daltonics), calibrated using sodium trifluoroacetate to achieve sub 1 ppm mass accuracy. All data acquisition was performed in broadband mode at m/z 200−2500. Intact mass, isolation, and fragment spectra were collected using a time-of-flight 1.2 ms, 2 M data point transient starting at 200 amu as a sum of 16 and 128 scans, respectively. Singly oxidized ions of two charge states, +16 and +18, were isolated in quadrupole using CASI (continuous accumulation of selected ions) at m/z 928.90 amu and 825.80 amu, respectively, using isolation window ±1.0 amu. For control spectra, three unmodified charge states (+16, +17, +18) were isolated using multiCASI (continuous accumulation of selected ions) at m/z: +16, 927.50 amu, +17, 873.10 amu, +18, 824.50 amu, windows ±1.0 amu. Selected ions were accumulated for 5 and 15 s in the quadrupole prior to the collision-induced dissociation (CID) and electron-capture dissociation (ECD),[39] respectively. For CID, a collision voltage of 17, 15, and 13.5 V was applied for +16, +17, and +18 charge states, respectively. In ECD, the conditions were different for both charge states. For +16, the pulse length was set to 0.085 V, bias voltage to 0.7 V, and lens voltage to 13 V. For +18, the pulse length was set to 0.095 V, bias voltage to 0.8 V, and lens voltage to 13 V. ECD fragmentation of unmodified +16/+17/+18 ion species was employed in multiCASI mode by setting the pulse length to 0.090 s, bias voltage to 1.5 V, and lens voltage to 15 V. Both fragmentation experiments were carried out in triplicate for apo and holo forms at the same conditions.

Obtained data were processed using DataAnalysis 5.0 (Bruker Daltonics) and MS2Links software.[40] Detailed information about top-down data processing is included in Supporting Information. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[41] partner repository with the data set identifier PXD027624.

RESULTS AND DISCUSSION

In the present study, the DNA binding domain of the transcription factor was used as a model protein for protein−DNA complex characterization by FPOP analysis. The FOXO4-DBD·DAF16 complex formation was confirmed by an electrophoretic mobility shift assay (EMSA; Methods in Supporting Information, Figure S3A) and native nESI (Methods in Supporting Information, Figure S3B). The efficiency of FPOP on FOXO4-DBD with/without DNA was verified by intact MS (Figure S4). For further analysis, identical samples were halved: one part was used for the bottom-up analysis and the second part for the top-down analysis.

Bottom-up Analysis. The bottom-up fraction was split for two independent proteolytic digestions. Using Lys-C endoproteinase, 82% sequence coverage was achieved compared to Lys-C/trypsin when only 63% of the sequence was covered. Although the Lys-C/trypsin digestion represented the worse sequence coverage, both Lys-C and trypsin peptides were

https://doi.org/10.1021/acs.analchem.1c04746
Anal. Chem. 2022, 94, 3203−3210

utilized for calculating the extent of modification to achieve the highest spatial resolution (Figure S2). Overall, 14 peptides were identified for both digests, resulting in the detection of 19 modified residues (Table S1). The accuracy of modification assignment by Peaks X+ software was manually validated for all peptides. Only peptides bearing a single modification were considered for quantification (Figure 1). The extent of modification was finally plotted onto a crystal structure of FOXO4-DBD·DAF16 (PDB: 3L2C), showing the relevance of our findings from a structural viewpoint of the studied interaction (Figure S5).

No significant difference in the extent of modification upon protein−DNA complex formation was detected for H-6, L-4, and W146 residues. Residues W97, Y102, Y124, W126, M127, R151, H152, H157, W173, W174, M175, and E179 were shown to be protected (less oxidized) upon the complex formation. The remaining ones (M-5, R94, Y133, W173, and M194) were oxidized to a higher extent (Figure 1). Residues W97 and Y102, directly interacting with the DAF16 duplex,[34,36] were found to be oxidized at multiple sites, resulting in chromatographically resolved isomers. All species had the same tendency of a lower extent of modification in the presence of DNA (Figure 1 and Figure S5).

Residues Y124, W126, and M127, shown to be less oxidized in the protein−DNA complex (Figure 1 and Figure S5), are included in the A103−T130 region consisting of helix H1, the intervening loop, strand S1, and helix H2. In previous structural studies, this region was not described in the context of FOXO4-DBD interaction with DNA. Obtained data support a rearrangement of the entire region caused by an allosteric effect induced by DNA binding as previously reported in the HDX study.[37] A higher extent of Y133 oxidation in the presence of DNA is in agreement with data indicating a structural rearrangement of helix H4 into an unstructured region.[33,34] The next three residues, namely R151, H152, and H157, are from helix H3, which creates a larger part of the protein−DNA interface.[34] As expected, all three residues were significantly protected upon the interaction of helix H3 with the DAF16 major groove (Figure 1 and Figure S5). It should be mentioned that R151 and H152 are conserved residues that recognize the binding motif of DAF16.[29] Four different residues were found to be modified in the peptide S171−K182, including two chromatographically resolved isomers of oxidized W173 (Figure 2). The largest difference in the extent of modification was observed for the W174 residue, which was almost 10 times less oxidized in the presence of DNA, confirming its crucial role in the protein−DNA interaction, previously shown by an X-ray structural model[34] (Figure S5) and mutagenesis analysis.[36] M175 and E179 residues were also less oxidized in the presence of DAF16, although they do not directly interact with DNA. Surprisingly, two oxidized W173 isoforms had different tendencies, where the hydrophilic isomer represents the more abundant product upon DNA binding. This phenomenon can be explained by decreased flexibility of the surrounding loops. The protection of the whole region might be attributed to an allosteric stabilization in the presence of DNA. Residue M194, located in the C-terminal part of DBD, was found to be heavily oxidized in the protein−DNA complex. This might be explained by the transport function from nucleus to cytoplasm as conveyed by the C-terminus. The so-called nuclear localization sequence (NLS) of wild-type FOXO4 (182−211) protein at the C-terminus contains the S193 residue, which should be spatially

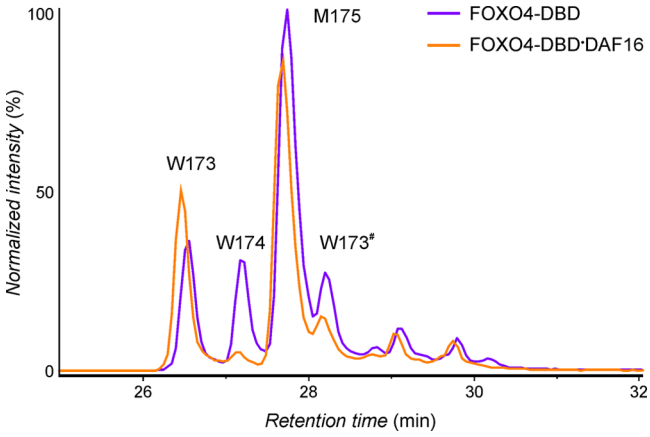


Figure 2. Overlaid extracted ion chromatograms (XICs) of singly oxidized peptide 171−182 with sequence SSWWMLNPEGGK appearing doubly charged at 704.321 amu. Different chromatographic profiles indicate changes in the protein oxidation upon DNA binding. Residue W173 oxidized at two different positions was chromatographically separated in the LC-MS analysis and is distinguished by #.

available for phosphorylation (required for the transport) in the bound form.[35,42,43]

Surprisingly, residue R94 was not protected in the presence of DNA despite its direct interaction with DNA via the guanidyl group.[34] This might be explained by the mechanism of γ-glutamyl aldehyde formation,[18] where hydroxyl radicals attack the δ-carbon of the arginine side chain, leading to the characteristic 43 Da loss. Indeed, analysis of NMR (PDB: 1E17)[33] and X-ray structural models (PDB: 3L2C)[34] revealed that, upon the interaction of the guanidyl group with the DNA backbone, the δ-carbon adopts a spatial orientation toward the solvent favorable for subsequent modification by hydroxyl radicals.

Top-down Approach. In the context of FPOP analysis, the bottom-up approach provides a very informative data set at single amino acid resolution. However, it analyzes all existing proteoforms in the sample at once, including those with multiple subsequent modifications. The extent of these modifications may vary between proteoforms, because preceding modifications may alter the initial structure of the molecule.[20] To verify the obtained data, singly oxidized forms were selected for an analysis using the top-down approach. Ions of two charge states, namely +16 and +18, were fragmented using CID and ECD techniques. Overall, 185 (20 b-ions, 27 y-ions, 138 internal ions) and 134 (77 c-ions, 51 z-ions, and 6 y-ions) fragment ions were identified by CID and ECD, respectively, resulting in 58% sequence coverage of FOXO4-DBD. Among them, 33 CID and 60 ECD ions were quantified (Figures S6−S8).

Derived extents of modifications of the apo and holo forms were subtracted from each other for the respective ions. Then the obtained values were used to visualize the differences between vicinal regions in a +16 charge state (Figure 3) and +18 charge state, respectively (Figure S9).

DNA binding significantly affects protein conformation and dynamics resulting in changes of modification extents between ions. The general trend was a stabilization of the protein molecule, i.e., a decrease in oxidation extent (Figure 4). Still, a higher level of oxidation was observed for several regions of the FOXO4-DBD molecule. Using c19, c20, and c21 ions the increased extent of oxidation was deduced for the R94 residue,

https://doi.org/10.1021/acs.analchem.1c04746
Anal. Chem. 2022, 94, 3203−3210

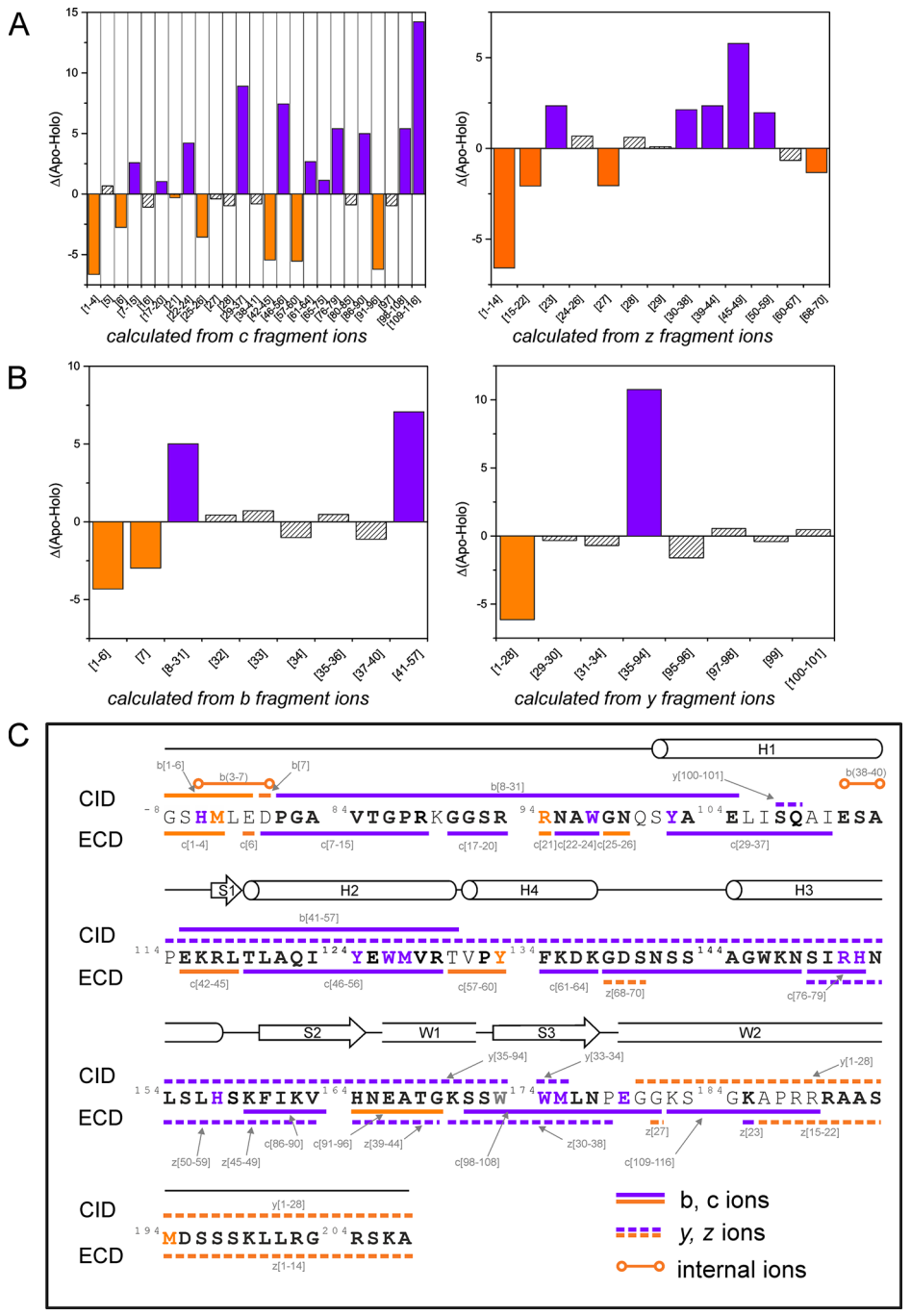


Figure 3. Plots indicating changes in oxidation rates between apo and holo forms for ECD (A) and CID (B) ions, obtained by fragmentation of the +16 charge state. Differential oxidation map of FOXO4-DBD with/without DNA (C). The bolded sequence represents spatial resolution achieved by the top-down approach (covered by complementary ECD and CID fragments). The regions between individual fragments are described above and below the sequence. Colored residues refer to the ones detected in the bottom-up approach. Secondary structure topology is denoted above the sequence. The data related to CID internal ions can be found in Figure S7.

confirming the respective bottom-up data. An increased modification extent of the N99 residue in the holo form was calculated based on the c[25−26] region and amino acid reactivity with hydroxyl radicals.[18] This residue was not detected as oxidized in the bottom-up approach. Regions c[42−45] and b(38−40) internal ion covering the end of helix H1, intervening loop, and strand S1 (E111−L118) indicated deprotection of this region upon DNA binding, probably, due to the loop rotation. Also, the region observed in the +18 charge state, c[42−43], indicates a higher oxidation level of the K116 residue. Region c[57−60] corresponding to T130−Y133 was found to be more oxidized in the presence of DNA, which is most likely attributed to the structural rearrangement of helix H4 mentioned above. A higher oxidation level of D136 residue in the presence of DNA was found in region c[63] in the +18 charge state.

Region H164−G169 was covered by regions c[90−92], c[93−97], and z[39−42], z[43−44] in the case of the +18 charge state. Region H164−N165 was more oxidized in the holo form, while the opposite trend was discovered for E166−

https://doi.org/10.1021/acs.analchem.1c04746
Anal. Chem. 2022, 94, 3203−3210

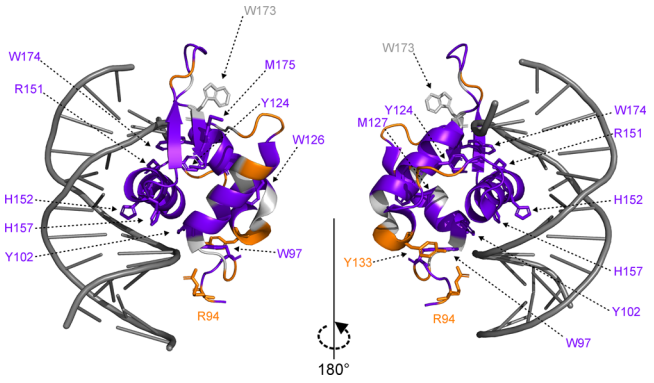


Figure 4. X-ray structural model of FOXO4-DBD·DAF16 (PDB: 3L2C)[34] with the highlighted differently oxidized regions detected by the top-down approach. The individual residues detected in the bottom-up approach were highlighted into these regions using a stick representation. Purple: regions/residues detected as more modified in the apo form, orange: regions/residues detected as more modified in the holo form.

G169. This is in agreement with the data observed in the +16 charge state, however, with lower resolution. One may conclude that an oxidation of H164 residue is deduced based on its higher reactivity.[18] Indeed, the side chain of H164 is oriented outward from the protein globule in the presence of DNA.[34] It should be mentioned that the data related to these regions were missing in the bottom-up approach. The last region found to be more oxidized upon DNA binding was the C-terminus (M194−A207) deduced from ions z[1−14], y[1−28], and y[1−25] generated from the +16 and +18 charge state, respectively. A flexibility of the C-terminus required for the implementation of biological functions[42] might explain the observed difference.

The N-terminus ((P-1)−S92) was protected upon DNA binding (Figure 3C) due to its close proximity to the DNA, as observed in a previous study.[37] Particularly, an oxidation and further protection in the presence of DNA was observed for residue S92. Also, a decreased modification level was deduced from the c[22−24] region for the most reactive residue W97, supporting our bottom-up data. Another protected region consisted of helices H1 (except its end) and H2, also confirming our findings from the bottom-up approach. Moreover, oxidation rates at single amino acid resolution deduced from the consecutive fragment ions b31−b36 and y96−y101 (Figures S7 and S8) were also in agreement with the obtained bottom-up data, indicating the lack of oxidation of E104, L105, I106, and S107 residues. The protection of the T119−R129 region (region c[46−56]) might be attributed to the Y124, W126, and M127 residues, found extensively oxidized in the bottom-up data (Figure 1 and Figure S5). Also, some contribution to this protection might be assigned to the L120 residue, which interacts with DAF16 as described previously.[34] As expected, a decrease in oxidation extents was detected for helix H3 (Figure 3C), which creates the main protein−DNA contact, supporting the respective bottom-up data for residues R151, H152, and H157. The higher modification extent of the K159−V163 region of strand S2 in the apo form (region c[86−90]) might be assigned to residue F160. The S171−G181 region belonging to the S3 strand contains highly reactive amino acids, namely W173, W174, and M175, which were shown to be protected upon DNA binding using the bottom-up approach. Top-down data (region c[98−108]) confirmed the reduced oxidation extent of this region in the holo form. The largest decrease in oxidation rates in the top-down analysis was detected for region K182−R189 of wing W2 (region c[109−116]). This might be explained by the interaction of R188 and R189 residues with DAF16.[36]

HDX and FPOP. One objective for the study of the same complex with different methods is obtaining complementary data that sheds light on the different aspects of a mechanism of interaction. In the case of FOXO4-DBD·DAF16, general structural information is available from NMR[33] and X-ray[34] studies, and HDX MS[37] data provide an overview of the interaction dynamics. In this context, FPOP data pinpoint particular residues participating or affected by protein−DNA interactions, as well as structural alterations occurring in the very early phase of these interactions (microseconds to milliseconds).

Comparison of top-down FPOP and HDX MS data revealed the complementary nature of both methods. The general trend of stabilization of a protein molecule upon DAF16 binding, observed in the HDX study,[37] was confirmed in the present work (Figure 3C and Figure S9C). A few discrepancies related to the small regions might be explained by achieving a better spatial resolution in the FPOP experiment (Figure 3C and Figure S9C). Still, there were two regions of interest, namely the H4 helix and the C-terminus. Unlike the HDX data, the C-terminus was found to be more oxidized in the presence of DNA. Its flexibility makes it easily accessible for deuterium exchange in both forms.[37] The FPOP approach provided more accurate data due to the covalent labeling of the M194 side chain (Figure 1). This result confirms the biological role of the FOXO4-DBD C-terminus described above.[35,42,43] Spatial rearrangements of the H4 helix upon DNA binding led to its higher structurization and, subsequently, to the lower deuterium exchange rate. Meanwhile, side chains of some amino acids, such as Y133 and D136, became more available to hydroxyl radicals, leading to the higher oxidation rate in FPOP analysis (Figure 4 and Figure S5). Thus, HDX MS provides an overview of the backbone structure, and FPOP adds information related to the orientation of side chains of single amino acids.

CONCLUSION

In this study, we performed FOXO4 footprinting, including the protein−DNA interface, using FPOP. The FPOP data set was resolved at single amino acid resolution using the bottom-up approach. Obtained data were verified by the top-down approach for the first time. For this purpose, singly oxidized ions were isolated using CASI and fragmented by ECD and CID. Although the top-down analysis did not reach the spatial resolution similar to that of the bottom-up data, it enabled the detection of modified residues of the selected proteoform (singly oxidized) which cannot be achieved by the bottom-up approach. Implemetation of UVPD[44,45] would increase the sequence coverage and can be beneficial for future studies. Application of bottom-up and top-down strategies allowed us to get comprehensive data reflecting the dynamics of FOXO4-DBD·DAF16 interactions. Interesting data were obtained for helix H4, the role of which remains unclear, where half of the helix was protected (F134−K135) and region V131−Y133 was more oxidized. Another region of interest was strand S3 with the highest level of protection, indicating its stabilization due to the direct interaction of side chain residues with DNA.

https://doi.org/10.1021/acs.analchem.1c04746
Anal. Chem. 2022, 94, 3203−3210

Alteration of the intervening loop and S1 strand in the presence of DNA was described for the first time. Moreover, the top-down approach enabled us to resolve modification of low reactive residues, which are not easily detectable using the bottom-up approach. Obtained FPOP data were in agreement with the HDX-MS study previously published by our group.[37] It should be mentioned that the protein and protein−DNA complex were expressed and purified in the same way in both studies, which allows direct comparison of the data and results. FPOP has the potential to be used as a valuable method for the analysis of even minor structural changes induced by DNA binding. However, it is difficult to generalize on the basis of a single example. To evaluate if FPOP really produces reliable data on protein−DNA interactions, more protein−DNA complexes should be investigated.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.1c04746.

Additional methods: Native nESI-MS, top-down data-processing, electrophoretic mobility shift assay. Additional figures: S1-FPOP workflow setup, S2-digestion map of FOXO4-DBD, S3-EMSA gel, native mass spectra of free and bound protein, S4-intact spectra of oxidized free and bound protein, S5-crystal structure of FOXO4-DAF16 complex with the highlighted residues detected in the bottom-up approach, S6-sequence of FOXO4 with depicted quantified fragment ions, S7-quantified ECD/CID fragment ions of the +16 charge state, S8-quantified ECD/CID fragment ions of the +18 charge state, S9-differential oxidation map arisen from the +18 charge state fragment analysis. Additional tables: Table S1-all modifications identified in the bottom-up approach (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Petr Novák − *Institute of Microbiology, The Czech Academy of Sciences, Prague 14220, Czech Republic;* orcid.org/0000-0001-8688-529X; Phone: +420 325 873 610; Email: pnovak@biomed.cas.cz; Fax: +420 241 062 156

Authors

Marek Polák − *Institute of Microbiology, The Czech Academy of Sciences, Prague 14220, Czech Republic; Faculty of Science, Charles University, Prague 12843, Czech Republic*

Ghazaleh Yassaghi − *Institute of Microbiology, The Czech Academy of Sciences, Prague 14220, Czech Republic;* Present Address: Department of Chemistry, University of Connecticut, 55 N. Eagleville Rd., Storrs, CT 06269, United States

Daniel Kavan − *Institute of Microbiology, The Czech Academy of Sciences, Prague 14220, Czech Republic; Faculty of Science, Charles University, Prague 12843, Czech Republic*

František Filandr − *Institute of Microbiology, The Czech Academy of Sciences, Prague 14220, Czech Republic; Faculty of Science, Charles University, Prague 12843, Czech Republic*

Jan Fiala − *Institute of Microbiology, The Czech Academy of Sciences, Prague 14220, Czech Republic; Faculty of Science, Charles University, Prague 12843, Czech Republic*

Zdeněk Kukačka − *Institute of Microbiology, The Czech Academy of Sciences, Prague 14220, Czech Republic;* orcid.org/0000-0001-7569-843X

Petr Halada − *Institute of Microbiology, The Czech Academy of Sciences, Prague 14220, Czech Republic*

Dmitry S. Loginov − *Institute of Microbiology, The Czech Academy of Sciences, Prague 14220, Czech Republic; Orekhovich Institute of Biomedical Chemistry, Moscow 119191, Russia;* orcid.org/0000-0002-4164-6145

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.1c04746

Author Contributions

The manuscript was written through the contributions of all authors. All authors have approved the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was mainly financially supported by the Czech Science Foundation (19-16084S), the European Commission H2020 (EU_FT-ICR_MS grant agreement ID: 731077 and EPIC-XS - grant agreement ID: 823839), and the Grant Agency of Charles University (359521). Additional institutional and facility support from the Academy of Sciences of the Czech Republic (RVO: 61388971), the Ministry of Education of the Czech Republic (Structural Mass Spectrometry CF - LM2018127 CIISB), and the European Regional Development Funds (CZ.1.05/1.1.00/02.0109 BIOCEV) are gratefully acknowledged. We also thank Michal Mrocek for the software development visualizing fragment ion maps.

■ REFERENCES

(1) Fabris, D.; Yu, E. T. *J. Mass Spectrom.* **2010**, *45* (8), 841−860.

(2) Liu, X. R.; Zhang, M. M.; Gross, M. L. Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications. *Chem. Rev.* **2020**, *27*, 4355−4454. DOI: 10.1021/acs.chemrev.9b00815

(3) Suckau, D.; Mak, M.; Przybylski, M. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (12), 5630−5634.

(4) Fiedler, W.; Borchers, C.; Macht, M.; Deininger, S.-O.; Przybylski, M. *Bioconjugate Chem.* **1998**, *9* (2), 236−241.

(5) Novak, P.; Kruppa, G. H.; Young, M. M.; Schoeniger, J. *J. Mass Spectrom.* **2004**, *39* (3), 322−328.

(6) Maleknia, S. D.; Brenowitz, M.; Chance, M. R. *Anal. Chem.* **1999**, *71* (18), 3965−3973.

(7) Sharp, J. S.; Becker, J. M.; Hettich, R. L. *Anal. Chem.* **2004**, *76* (3), 672−683.

(8) Aye, T. T.; Low, T. Y.; Sze, S. K. *Anal. Chem.* **2005**, *77* (18), 5814−5822.

(9) Hambly, D. M.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (12), 2057−2063.

(10) Chen, J.; Cui, W.; Giblin, D.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (8), 1306−1318.

(11) Zhang, B.; Rempel, D. L.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2016**, *27* (3), 552−555.

(12) Cheng, M.; Zhang, B.; Cui, W.; Gross, M. L. *Angew. Chemie Int. Ed.* **2017**, *56* (45), 14007−14010.

(13) Fojtík, L.; Fiala, J.; Pompach, P.; Chmelík, J.; Matoušek, V.; Beier, P.; Kukačka, Z.; Novák, P. *J. Am. Chem. Soc.* **2021**, *143* (49), 20670−20679.

(14) Hawkins, C. L.; Davies, M. J. *J. Biol. Chem.* **2019**, *294* (51), 19683−19708.

(15) Chen, G.; Tao, L.; Li, Z. *Drug Discovery Today* **2022**, *27*, 196.

(16) Li, K. S.; Shi, L.; Gross, M. L. *Acc. Chem. Res.* **2018**, *51* (3), 736−744.
(17) Zhang, B.; Cheng, M.; Rempel, D.; Gross, M. L. Implementing Fast Photochemical Oxidation of Proteins (FPOP) as a Footprinting Approach to Solve Diverse Problems in Structural Biology. *Methods* **2018**, *15*, 94−103. DOI: 10.1016/j.ymeth.2018.05.016
(18) Xu, G.; Chance, M. R. *Chem. Rev.* **2007**, *107* (8), 3514−3543.
(19) Takamoto, K.; Chance, M. R. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35* (1), 251−276.
(20) Sharp, J. S.; Becker, J. M.; Hettich, R. L. *Anal. Biochem.* **2003**, *313* (2), 216−225.
(21) Charvátová, O.; Foley, B. L.; Bern, M. W.; Sharp, J. S.; Orlando, R.; Woods, R. J. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (11), 1692−1705.
(22) Watkinson, T. G.; Calabrese, A. N.; Ault, J. R.; Radford, S. E.; Ashcroft, A. E. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (1), 50−55.
(23) Lu, Y.; Zhang, H.; Niedzwiedzki, D. M.; Jiang, J.; Blankenship, R. E.; Gross, M. L. *Anal. Chem.* **2016**, *88* (17), 8827−8834.
(24) Jones, L. M.; Sperry, J. B.; Carroll, J. A.; Gross, M. L. *Anal. Chem.* **2011**, *83* (20), 7657−7661.
(25) Stocks, B. B.; Konermann, L. *Anal. Chem.* **2009**, *81* (1), 20−27.
(26) Tullius, T. D.; Dombroski, B. A. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83* (15), 5469−5473.
(27) Calnan, D. R.; Brunet, A. *Oncogene* **2008**, *27* (16), 2276−2288.
(28) Psenakova, K.; Kohoutova, K.; Obsilova, V.; Ausserlechner, M.; Veverka, V.; Obsil, T. *Cells* **2019**, *8* (9), 966.
(29) Obsil, T.; Obsilova, V. *Oncogene* **2008**, *27* (16), 2263−2275.
(30) Furuyama, T.; Nakazawa, T.; Nakano, I.; Mori, N. *Biochem. J.* **2000**, *349* (2), 629.
(31) Valis, K.; Prochazka, L.; Boura, E.; Chladova, J.; Obsil, T.; Rohlena, J.; Truksa, J.; Dong, L. F.; Ralph, S. J.; Neuzil, J. *Cancer Res.* **2011**, *71* (3), 946−954.
(32) Biggs, W. H.; Meisenhelder, J.; Hunter, T.; Cavenee, W. K.; Arden, K. C. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (13), 7421−7426.
(33) Weigelt, J.; Climent, I.; Dahlman-Wright, K.; Wikström, M. *Biochemistry* **2001**, *40* (20), 5861−5869.
(34) Boura, E.; Rezabkova, L.; Brynda, J.; Obsilova, V.; Obsil, T. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (12), 1351−1357.
(35) Boura, E.; Silhan, J.; Herman, P.; Vecer, J.; Sulc, M.; Teisinger, J.; Obsilova, V.; Obsil, T. *J. Biol. Chem.* **2007**, *282* (11), 8265−8275.
(36) Vacha, P.; Zuskova, I.; Bumba, L.; Herman, P.; Vecer, J.; Obsilova, V.; Obsil, T. *Biophys. Chem.* **2013**, *184*, 68−78.
(37) Slavata, L.; Chmelík, J.; Kavan, D.; Filandrová, R.; Fiala, J.; Rosůlek, M.; Mrázek, H.; Kukačka, Z.; Vališ, K.; Man, P.; Miller, M.; McIntyre, W.; Fabris, D.; Novák, P. *Biomolecules* **2019**, *9* (10), 535.
(38) Loginov, D. S.; Fiala, J.; Chmelik, J.; Brechlin, P.; Kruppa, G.; Novak, P. *ACS Omega* **2021**, *6* (15), 10352−10361.
(39) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *J. Am. Chem. Soc.* **1998**, *8*, 3265−3266. DOI: 10.1021/ja973478k
(40) Kellersberger, K. A.; Yu, E.; Kruppa, G. H.; Young, M. M.; Fabris, D. *Anal. Chem.* **2004**, *76* (9), 2438−2445.
(41) Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; Pérez, E.; Uszkoreit, J.; Pfeuffer, J.; Sachsenberg, T.; Yılmaz, Ş.; Tiwary, S.; Cox, J.; Audain, E.; Walzer, M.; Jarnuczak, A. F.; Ternent, T.; Brazma, A.; Vizcaíno, J. A. *Nucleic Acids Res.* **2019**, *47* (D1), D442−D450.
(42) Obsil, T.; Ghirlando, R.; Anderson, D. E.; Hickman, A. B.; Dyda, F. *Biochemistry* **2003**, *42* (51), 15264−15272.
(43) Obsilova, V.; Vecer, J.; Herman, P.; Pabianova, A.; Sulc, M.; Teisinger, J.; Boura, E.; Obsil, T. *Biochemistry* **2005**, *44* (34), 11608−11617.
(44) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646−12651.
(45) Greer, S. M.; Brodbelt, J. S. *J. Proteome Res.* **2018**, *17* (3), 1138−1145.

https://doi.org/10.1021/acs.analchem.1c04746
Anal. Chem. 2022, 94, 3203−3210

Attached Publications:

Publication II

Polák, M.; Černý, J.; Novák, P.

Isotopic Depletion Increases the Spatial Resolution of FPOP Top-Down Mass Spectrometry Analysis.

Analytical Chemistry, 2024

My contribution: Protein expression and purification, sample preparation, FPOP of samples, analyzing obtained datasets, manuscript writing and preparing

analytical chemistry

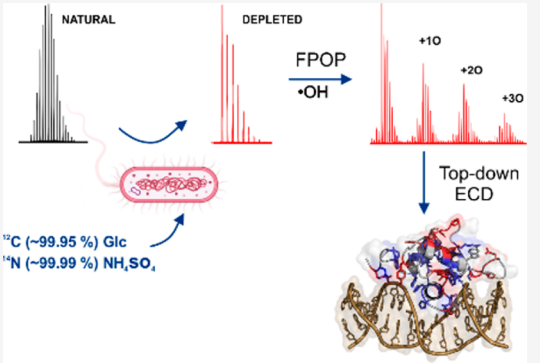
pubs.acs.org/ac
Article

Isotopic Depletion Increases the Spatial Resolution of FPOP Top-Down Mass Spectrometry Analysis

Marek Polák, Jiří Černý, and Petr Novák*

Cite This: *Anal. Chem.* 2024, 96, 1478−1487
Read Online

ABSTRACT: Protein radical labeling, like fast photochemical oxidation of proteins (FPOP), coupled to a top-down mass spectrometry (MS) analysis offers an alternative analytical method for probing protein structure or protein interaction with other biomolecules, for instance, proteins and DNA. However, with the increasing mass of studied analytes, the MS/MS spectra become complex and exhibit a low signal-to-noise ratio. Nevertheless, these difficulties may be overcome by protein isotope depletion. Thus, we aimed to use protein isotope depletion to analyze FPOP-oxidized samples by top-down MS analysis. For this purpose, we prepared isotopically natural (IN) and depleted (ID) forms of the FOXO4 DNA binding domain (FOXO4-DBD) and studied the protein−DNA interaction interface with double-stranded DNA, the insulin response element (IRE), after exposing the complex to hydroxyl radicals. As shown by comparing tandem mass spectra of natural and depleted proteins, the ID form increased the signal-to-noise ratio of useful fragment ions, thereby enhancing the sequence coverage by more than 19%. This improvement in the detection of fragment ions enabled us to detect 22 more oxidized residues in the ID samples than in the IN sample. Moreover, less common modifications were detected in the ID sample, including the formation of ketones and lysine carbonylation. Given the higher quality of ID top-down MSMS data set, these results provide more detailed information on the complex formation between transcription factors and DNA-response elements. Therefore, our study highlights the benefits of isotopic depletion for quantitative top-down proteomics. Data are available via ProteomeXchange with the identifier PXD044447.



Structural proteomics has considerably advanced the field of structural and molecular biology in recent years by enabling us to address questions related to the structure and dynamics of proteins and their complexes. Methods of structural proteomics are particularly useful for studying transcription factor−DNA interactions[1] such as those involved in transcription. Transcription is primarily regulated by transcription factors (TFs), proteins that specifically activate or inhibit this process by forming complexes with DNA. For instance, Forkhead transcription factor "O" 4 (FOXO4) is known to bind core motif 5′-(C/A)(C/C)AAA(C/T)A-3′ (Insulin Response Element) and further activate transcription that includes energy metabolism control.[2,3] Yet, despite extensive research in this area, our understanding of transcription regulation and interaction between TFs and their binding motifs remains limited. To better understand these processes, we must further clarify the structural mechanism underlying the formation and interaction of protein−DNA complexes.[4−6]

Protein−DNA complexes have been studied using cross-linking reactive probes,[7−9] hydrogen−deuterium exchange (HDX)[7,8,10−12] and radical covalent labeling.[1] It was demonstrated that surface mapping of biomolecules, detected by high-resolution MS analysis, offers useful information related to the protein structure or its interaction with other biomolecules.[1]

Various radical probes are currently available, with diverse reactivity toward different amino acids,[13−18] but the most commonly used labeling chemistry consists of modification by hydroxyl radicals.[19] Among the methods involving hydroxyl radicals, the most promising approach was introduced by Hambly and Gross in 2005 and is referred to as fast photochemical oxidation of proteins (FPOP).[20]

In FPOP, proteins are irreversibly labeled in a quench-flow capillary system by a single hit of hydroxyl radical, generated from hydrogen peroxide by an excimer laser. This reaction is immediately quenched by a suitable quenching reagent in the flow system. The rationale of this approach lies in the preferential oxidation of solvent-accessible side chains of the investigated molecule, and thus mapping of the protein

Received: August 22, 2023
Revised: November 8, 2023
Accepted: December 15, 2023
Published: January 16, 2024

https://doi.org/10.1021/acs.analchem.3c03759
Anal. Chem. 2024, 96, 1478−1487

structure.[21,22] Hydroxyl radicals promote modification of 14 of 20 naturally occurring amino acids. The most reactive are the sulfur-containing amino acids, namely, cysteine and methionine, followed by the residues of the aromatic amino acids, namely, phenylalanine, tyrosine, histidine, and tryptophan.[23−26] As such, FPOP has been applied to various studies, including mapping of protein conformational changes, protein−protein interactions,[27] the structure and topology of membrane proteins,[28−31] and large biomolecules[32] such as antibodies.[20,27−31,33,34]

Although bottom-up MS analysis has long been a method of choice for studying FPOP modified samples, top-down protocols have recently demonstrated their potential for analyzing singly oxidized proteoforms.[1,35,36] Originally, a top-down technology was introduced to determine not only biomolecule sequences but also post-translational and other (bio)chemical modifications of biomolecules.[37−39] One of the advantages of analyzing samples by top-down MS is the precise determination of the molecular mass of proteins and other proteoforms. However, when the molecular mass of the analyzed species exceeds ∼1 kDa, the monoisotopic peak ceases to be the most abundant signal in the spectrum of the analyzed protein or peptide. Mass spectra of multiply charged proteins and fragment ions over ∼10 kDa do not produce observable monoisotopic peaks anymore, so the ion signal is dispersed into several, commonly overlapping isotopic peaks. As a result, the MS/MS spectra are complex and show a low signal-to-noise ratio (SNR).[40,41]

Precisely designed to overcome these difficulties, the technique of protein isotope depletion was introduced in 1997.[42] Protein isotope depletion relies on incubating bacteria in media with depleted heavy isotopes, e.g., carbon and nitrogen. Several studies have described the benefits of isotope depletion for analyzing proteins up to ∼20 kDa,[43−45] but Gallagher et al. stood out for using protein isotope depletion to improve the resolution of MS/MS spectra and thus fragment ion assignment.[46] More recently, Popovic et al. applied protein depletion to analyze the proteome of cells by 21T-FT-ICR mass spectrometry.[47] In common, these studies leveraged the potential of protein depletion for improving the accuracy of mass determination and sequence coverage of biomolecules by MS.

Considering the above, this study aimed at exploring the concept of protein isotope depletion and its advantages for increasing the spatial resolution of top-down MS analysis of FPOP samples. To this end, we prepared an isotopically depleted version of FOXO4-DBD and oxidized this sample under two different conditions, that is, with and without its DNA binding element, IRE. First, we used a bottom-up MS approach to analyze isotopically natural (IN) protein samples to acquire single-amino acid information. Subsequently, we applied a top-down MS approach to analyze an isotopically depleted (ID) version of the protein, which resulted in enhanced sequence coverage and more precise assignment of labeled residues. ID fragmentation yielded new ions in the MS/MS spectra, which were not detected in the IN spectra. The additional fragment ions offered more detailed information about the solvent-accessible FOXO4-DBD surface, thus enabling *ab initio* design of a FOXO4-IRE structural model.

EXPERIMENTAL SECTION

Materials and Chemicals. All solvents (LC/MS grade) and chemicals were purchased from Merck (Germany) unless stated otherwise.

Protein Expression of Isotopically Natural (IN) FOXO4-DBD. The expression was performed in Terrific-Broth (TB) medium. All details about the expression of IN-FOXO4-DBD can be found in the Supporting Information.

Protein Expression of Isotopically Depleted (ID) FOXO4-DBD. The expression of FOXO4-DBD was performed in M9 minimal medium containing glucose (99.95% of ^{12}C, Merck) and ammonium sulfate (NH_4SO_4, 99.99% of ^{14}N, Merck) as a source of carbon and nitrogen, respectively. All details about the expression of ID-FOXO4-DBD can be found in the Supporting Information.

Protein Purification. Both IN- and ID-FOXO4-DBD were purified in the same fashion, as described in detail in the Supporting Information.

FOXO4·IRE Complex Formation. Forward (5′-GAC TAT CAA AAC AAC GC-3′) and reverse (5′-GCG TTG TTT TGA TAG TC-3′) complementary strands, whose sequence was retrieved from a previous study,[19] were obtained from IDT (Coralville, USA) in an HPLC quality. A stock solution of dsIRE was prepared by mixing both strands in an equimolar ratio in LC-MS water. The mixture was incubated at 90 °C for 3 min and cooled to room temperature to form dsIRE. The 30 μM and 35 μM samples of IN/ID FOXO4-DBD and dsIRE, respectively, were mixed to form a complex and ensure that all protein was bonded in complex with DNA. The complex was diluted in 150 mM ammonium acetate, pH 6.8, and incubated at room temperature for 15 min to obtain IN/ID-FOXO4·IRE complex at 30 μM final concentration.

Fast Photochemical Oxidation of Proteins (FPOP). The FPOP labeling was performed in an in-house built quench-flow reactor as described previously.[1] Prior to the FPOP experiment, glutamine (10 mM) was added to the protein samples. Briefly, ID/IN FOXO4-DBD (30 μM) with and without dsIRE (35 μM), in 150 mM ammonium acetate, pH 6.8, was continuously mixed in a T-mixer with H_2O_2 (15 mM during reaction). The mixture was irradiated by an excimer laser (COMPex 50 KrF, Coherent Inc., USA). A mixture of the sample and H_2O_2 was subjected to a single shot at a wavelength of 248 nm, frequency 15 Hz, energy 107 mJ, 20 ns pulse duration, and 2.24 mJ/cm^2 radiant exposure. The exclusion volume was 16%. The reaction was quenched by immediate mixing with 75 mM methionine. The samples were collected in an Eppendorf tube containing 3000 U of Catalase (Merck, USA).

Top-Down MS Detection. Protein−DNA samples were denatured by adding 4 M urea and 1 μM $MgCl_2$; the mixture was incubated at a bench for 15 min. DNA was digested by adding 1 μL of benzonase endonuclease (Merck) and incubated at 30 °C for an additional 15 min. The mixture was then loaded onto a reverse-phase microtrap column (Optimize technologies, USA), desalted using 0.1% FA, and eluted with 80% ACN, 0.1% FA. The desalted protein was further diluted 5 times using 30% ACN, 0.1% FA solution, and sprayed using a nESI source in positive mode at 120 °C desolvation temperature. MS and MS/MS analyses were performed on a solariX XR mass spectrometer equipped with 15 T magnet (Bruker Daltonics, Billerica, USA), which was externally calibrated using the sodium trifluoroacetate to

https://doi.org/10.1021/acs.analchem.3c03759
Anal. Chem. 2024, 96, 1478−1487

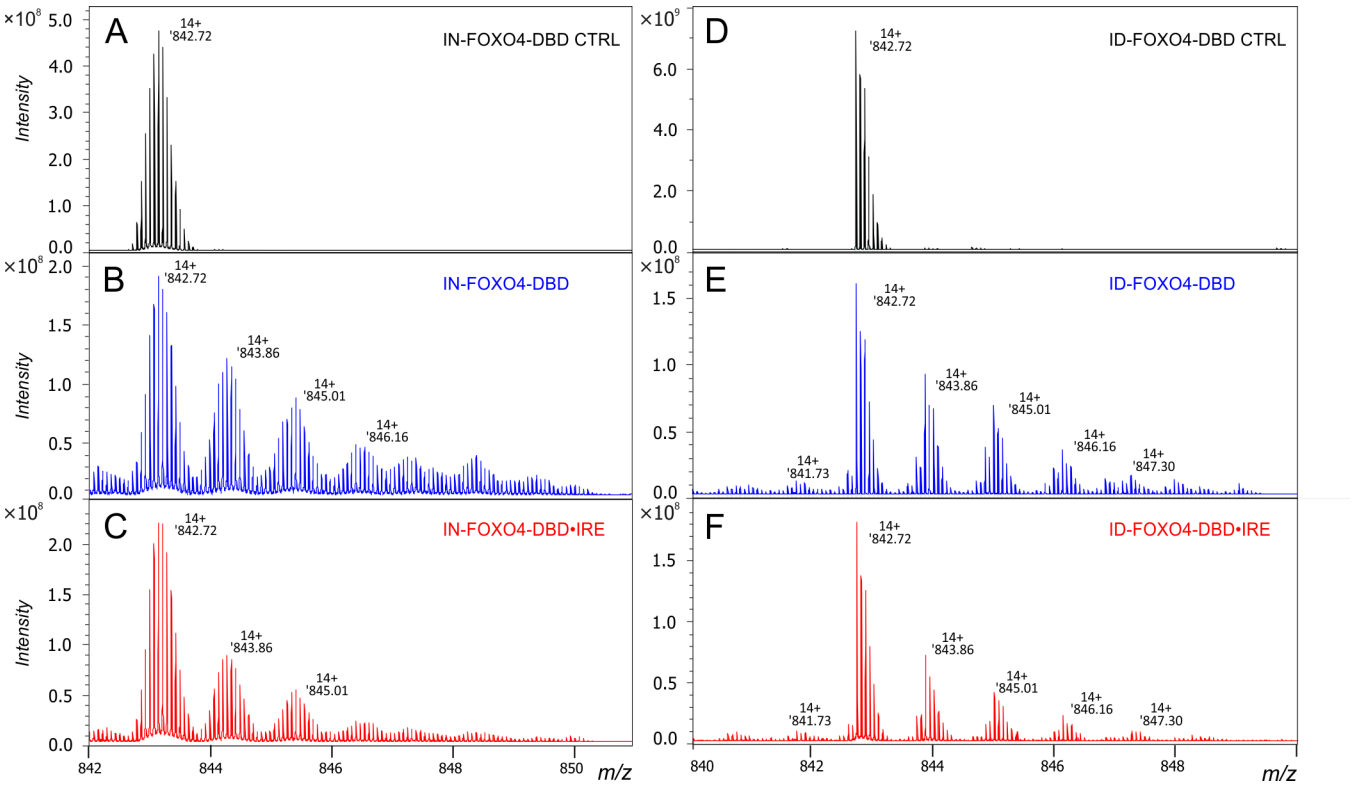


Figure 1. Zoomed mass spectrum on a +14-charge state (m/z 842−851) showing an isotopic distribution of isotopically natural (IN-, A) and isotopically depleted (ID-, D) FOXO4-DBD. Fast photochemical oxidation of IN-FOXO4 without (B) and with (C) dsIRE. Fast photochemical oxidation of ID-FOXO4 without (E) and with (F) dsIRE.

achieve 1 ppm mass accuracy. The time-of-flight was set to 1.1 ms, with the collision energy ranging between −3.0 to −5.0 V, and data were acquired with 2 M data point transient starting at 200 amu. At first, MS intact spectra were acquired in a broadband mode (m/z 200−2500) by accumulating ions for 0.1 s−0.2 s and collecting 128 scans.

For MS/MS, singly oxidized protein ions of three charge states (+14, +13, +12) were isolated using a multiCASI (multicontinuous accumulation of selected ions) in a quadrupole and then transferred to ICR cell for electron-capture dissociation (ECD). The single oxidized ions of IN-FOXO4 were isolated at 844.50, 909.30, and 984.90 amu, and the isolation window was ±1.0 amu. ID-FOXO4-DBD singly oxidized ions were isolated at nominal values 843.93, 908.77, and 984.42 amu, and the isolation window was ±0.6 amu. An ion-accumulation from 3.0 to 5.0 s was tuned to reach the intensity of the precursor ion image current of ∼10^8 prior to the ECD experiment. The ECD was done by setting the parameters to obtain optimal fragmentation as follows: ECD pulse length 0.065−0.075 s, bias 0.90−1.0 V, and lens 14.0−15.0 V. The hollow cathode current was 1.5 A. Control spectra of unmodified ions were acquired using the same condition as the oxidized ones. Data were acquired by collecting 128 scans in a technical triplicate for both apo and holo forms.

Top-Down Data Processing. Raw data were processed using Data Analysis 5.3 (Bruker Daltonics, Billerica, USA), MS2Links software,[27] and in-house built software. Details related to top-down data processing are included in the Supporting Information.

Bottom-Up LCMS Detection. Samples for bottom-up analysis were digested using Trypsin/LysC (Promega, USA) and LysC (Promega, USA). Respective protease was added at a protease:protein ratio 1:40 (m:m) and incubated overnight at 37 °C. Additional protease (m/m 1:20) was added after overnight incubation for another 6 h. IRE was digested by adding bensonase endonuclease (250 U, Merck) to the sample for 30 min at 37 °C. Digestion was terminated by addition of trifluoroacetic acid (TFA, 0.1%).

LC was performed to separate the peptides as described previously,[1] albeit with one minor modification. An LC run consisted of a 35 min linear gradient of 2−35% solvent B. LC was directly hyphenated to a trapped ion mobility-quadrupole time-of-flight mass spectrometer (timsTOF Pro, Bruker Daltonics) for MS/MS analysis. MS/MS analysis was performed as described previously.[1]

LC-MS analysis was performed on a timsTOF Pro mass spectrometer. MS analysis was performed using the same method as MSMS, but without collisional dissociation and fragment accumulation.

Data were processed using a PeaksX+ Software (Bioinformatic Solutions Inc., Waterloo, ON, Canada) against a FOXO4-DBD sequence as described previously.[1,48] Peptide intensities were extracted from LC-MS trace using Data Analysis 5.3 (Bruker Daltonics, USA) for all observed charge states, quantified, and statistically analyzed, as described previously.[49] The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[50] partner repository with the data set identifier PXD044447.

Homology Modeling. The homology model of the FOXO4/IRE complex was based on the available crystal structure 3l2c,[51] extending the 3l2c template DNA sequence "−CTATGTAAACAAC−" to the IRE "GACTATCAAAACA-ACGC" sequence. The necessary nucleotide substitutions as

https://doi.org/10.1021/acs.analchem.3c03759
Anal. Chem. 2024, 96, 1478−1487

well as residues missing at the 5′ and 3′ termini were modeled using the MMB program.[52,53] The backbone conformation of the newly built termini was set to the canonical B form DNA using dinucleotide conformation class (NtC) BB00.[54,55] The geometry of the initial model was further equilibrated during a 200 ns molecular dynamics simulation in GROMACS 2021.4[56] using the ff14SB[57] force field for FOXO4 and the tumuc1 force field[58] for DNA. The model was placed in a rectangular box with the 10 Å shortest distance from the walls. The box was filled with TIP3P model water, and Na^+ and Cl^- ions were added to reach a charge-neutral system with 100 mM salt concentration. The system was simulated with noncovalent cutoffs of 10 Å at 300 K and 1 bar with the V-rescale modified Berendsen thermostat and the Parrinello−Rahman barostat.

RESULTS AND DISCUSSION

Top-down MS analysis of FPOP enables us to determine the protein−DNA interaction interface between FOXO4 and dsDNA (namely DAF16), as shown in our recent study.[1] However, the complex spectra and lower sequence coverage prevent us from reaching single-amino acid resolution, as in the bottom-up MS approach. To increase fragment intensity and to reduce the complexity of top-down MS spectra, we prepared IN and $^{13}C/^{15}N$-doubly depleted (ID) versions of FOXO4-DBD to investigate the benefit of isotopic depletion for studying the interaction interface between FOXO4 and IRE by top-down MS analysis. We therefore expressed and purified the DNA-binding domain of FOXO4 at a length of 82−186[3] (numbering according to the FOXO4 wild-type sequence) containing two ($G^{-2}S^{-1}$) additional residues located in the N-terminus. So as to simplify the top-down MS data interpretation, we further used the 1 to 107 common numbering of *c* and *z* fragment ions. The conversion of fragment ion numbering to the wild-type FOXO4 sequence is shown in Figure S1.

Initially, TB and M9 media were used to recombinantly express IN and ID protein, respectively (Methods in the Supporting Information). Both proteins were purified by using the same protocol. The intact protein analysis revealed the homogeneity of both protein samples (Figure S2) and confirmed the isotopic depletion of FOXO4-DBD (Figure 1A,D), while electrophoretic mobility shift assay (EMSA, Methods in the Supporting Information) demonstrated the ability of both proteins to bind dsIRE (Figure S3). Subsequently, the ID and IN forms of FOXO4-DBD were oxidized with and without dsIRE using FPOP. Figure 1B,C shows multiple oxidized proteoforms of IN-FOXO4-DBD with and without dsIRE, respectively, after FPOP. Similarly, Figure 1E,F displays multiple oxidized proteoforms of ID-FOXO4-DBD with and without dsIRE, respectively. With dsIRE (Figure 1C,F), the proteoforms were less oxidized than in solution alone (Figure 1B,E). These results confirm the protection of residues directly or indirectly involved in the protein−DNA interaction.

Given the confirmation of the interaction between FOXO4-DBD and IRE by EMSA (Figure S3) and by intact MS analysis (Figure 1B,C,E,F), the investigation proceeded with top-down MS analysis. In order to demonstrate the advantage of isotopic depletion for analyzing singly oxidized proteins, singly oxidized proteoforms of IN and ID samples were isolated in the quadrupole and fragmented by ECD in the ICR cell (Figure S4).[35] By combining multiCASI simultaneous isolation of three charge states with ECD fragmentation, we were able to detect 101 nonoxidized fragment ions (57 *c* ions, 44 *z* ions) when analyzing IN samples. However, only 57 fragment ions (30 *c* ions, 27 *z* ions; see Figure S5A) were intense enough for quantification, resulting in a sequence coverage of 27% (see Figure S5B). Under the same experimental conditions, three charge states were isolated and fragmented using the ID sample. The isotopic purity of the ID protein allowed us to isolate oxidized ions in the quadrupole with a narrower isolation window of ±1 and ±0.6 Da for IN and ID samples, respectively. In total, 148 fragment ions were then annotated and manually validated, including small ions, *c*3, *c*4, and *z*3, which did not show other observable isotopic peaks (Figures S6 and S7). Unlike in the IN sample, 95 fragment ions (54 *c* ions, 42 *z* ions; see FigureS8A) were intense enough for quantification, resulting in a sequence coverage of 45% (Figure S8B). Thus, another 24 *c*-ions and 15 *z*-ions were available for quantification in ID (Figure 2) compared to in IN samples.

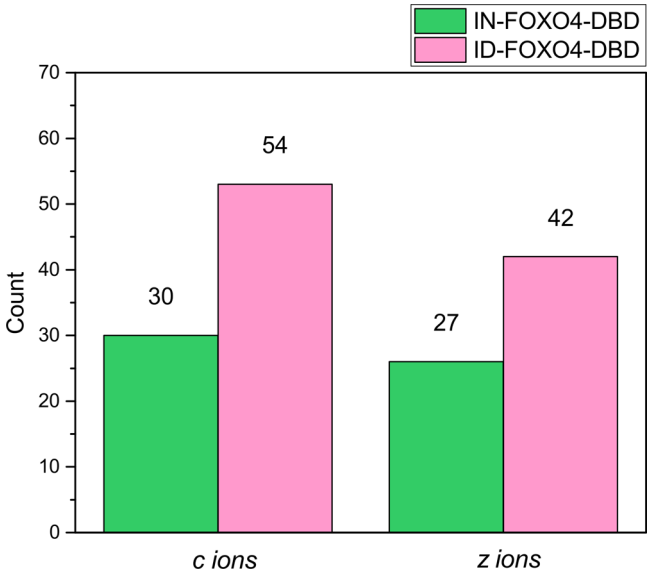


Figure 2. Histograms displaying the number of quantified fragment ions generated by ECD fragmentation of singly oxidized precursor ions of IN-FOXO4-DBD and ID-FOXO4-DBD.

ID MS/MS spectra displayed (i) lower complexity than IN spectra, which greatly reduced the overlap of isotopes/fragment ions, (ii) a monoisotopic peak for all fragment ions, and (iii) an increased signal-to-noise ratio (SNR) (Figure S9). To further illustrate these improvements, we can consider the *m*/*z* region 1056−1061, which shows the difference in the complexity of the ECD spectra between IN and ID oxidized proteins (Figure 3).

Figure 3A clearly shows the overlap of three isotopes for two fragment ions, $[c21]^{2+}$ and $[c57]^{6+}$, thus precluding reliable quantification of the $[c57]^{6+}$ fragment ion. Moreover, its oxidized form, $[c57+O]^{6+}$, has a very low signal-to-noise ratio, which is close to the limit of detection. In contrast, fragment ion intensity is approximately three times higher in ID (Figure 3B) than in IN MS/MS spectra. Hence, the $[c21]^{2+}$ fragment ion consists of only two peaks, and both $[c57]^{6+}$ $[c57+O]^{6+}$ fragments are now observed as intense ions in spectra, where the most abundant peak is the monoisotopic peak, which significantly exceeds the noise level.

During FPOP, the high reactivity of ·OH radicals promotes not only oxygen additions (+15.9949 Da) but also other

https://doi.org/10.1021/acs.analchem.3c03759
Anal. Chem. 2024, 96, 1478−1487

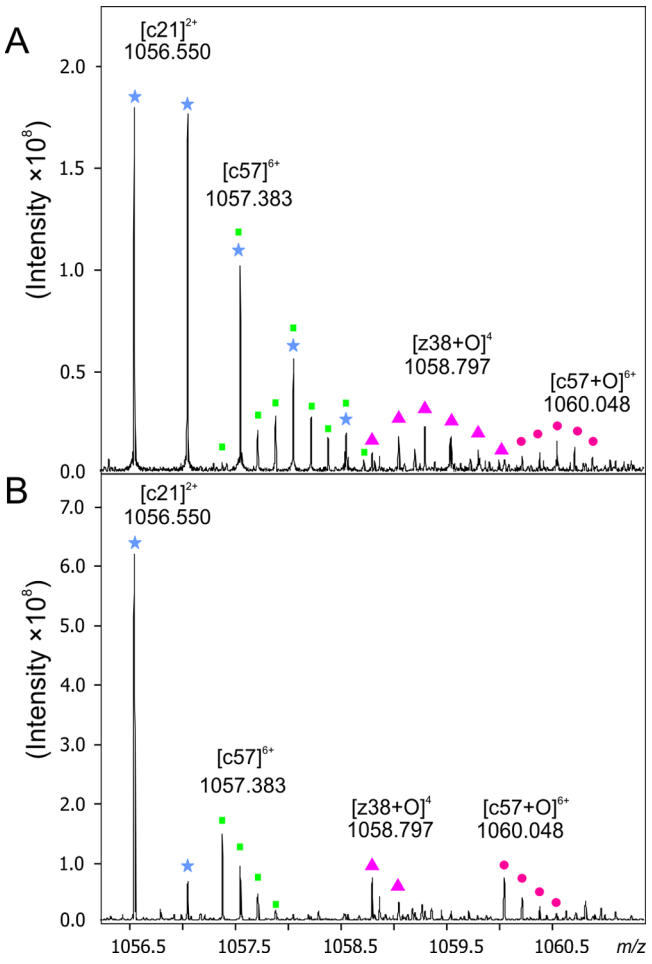


Figure 3. Zoomed ECD spectrum obtained upon fragmentation of isotopically natural (A) and isotopically depleted (B) FOXO4-DBD. The $[c21]^{2+}$ is indicated with blue asterisks; the low-abundance $[c57]^{6+}$ fragment ion is denoted by green squares; and its oxidized form, $[c57+O]^{6+}$, is indicated by pink dots. The oxidized fragment ion $[z38+O]^{4}$ is denoted by magenta triangles.

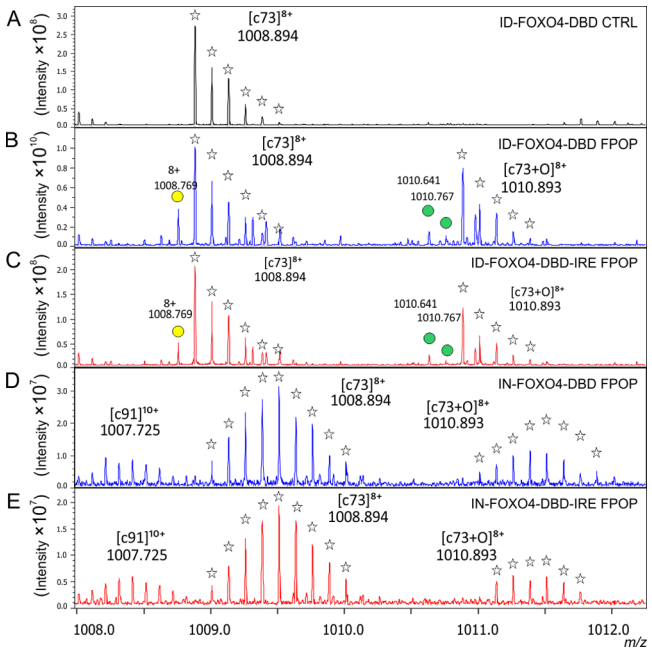


Figure 4. MS/MS spectrum zoomed in the m/z range 1008−1012.300. The control ECD spectrum of unmodified ID-FOXO4-DBD is colored in black in the top panel (A). The ECD spectrum of oxidized ID-FOXO4-DBD with (B) and without IRE (C) is colored in blue and in red. The isotopic distribution of both $[c73]^{8+}$ and $[c73+O]^{8+}$fragment ions is denoted by transparent asterisks. Yellow dots denote lysine carbonylation within the protein, represented by the loss of 1.013 Da, while the green dots represent the oxidation of protein to its keto form (+13.9793). An ECD MS/MS spectrum of IN-FOXO4-DBD without IRE (D) and with IRE (E) shows no visible lysine carbonylation or oxidation to keto form.

modifications (Table S1).[59] For instance, the conversion to the keto form[60] (addition of +13.9793 Da) or lysine carbonylation[36,61] (loss of −1.0313 Da) can also be detected during the FPOP analysis. In a typical bottom-up data analysis,[62] where small peptides are created by enzymatic digestion and subsequently separated on a reversed-phase column, these modifications are easily detected by DDA analysis. However, these modifications are difficult to detect by top-down MS analysis for several reasons: (i) they are not major products of the reaction and thus are not observed at a higher intensity, and (ii) they may overlap with other reaction products in the MS/MS spectrum.[1,35]

By simplifying the mass spectra and improving the signal-to-noise ratio (Figure S9), isotopic depletion helps to overcome the aforementioned limitations. As a result, these modifications can be detected in top-down MS/MS spectra of ID samples, as exemplified by the $[c73]^{8+}$ fragment ion (Figure 4). The control spectrum does not show any oxidized $[c73+O]^{8+}$ fragment ion (Figure 4A, black). But an MS/MS spectrum of oxidized FOXO4-DBD without (Figure 4B,D; blue) and with the IRE (Figure 4C,E; red) provided an oxidized $[c73+O]^{8+}$ fragment ion. Additionally, oxidation to a keto form (indicated by green dots) and protein carbonylation (indicated by a yellow dot) were also detected during MS analysis, albeit to a lesser extent with IRE. This result confirms both the wide reactivity of ·OH radicals toward different residues and the protection of some residues by IRE (Figure 4B,C). Panels D and E of Figure 4, conversely, both show that neither lysine carbonylation nor keto-oxidation is identified in the zoomed-in view of the MS/MS spectrum of IN-FOXO4-DBD.

To obtain structural information based on the assignment of oxidized residues and the increasing intensity of fragment ions, the extent of oxidation was calculated for both the apo and holo forms. This allowed us to visualize the differences between vicinal fragment regions (Figure 5). The difference map for isotopically natural FOXO4 represents an example of lower sequence coverage (Figure 5A). In this case, only several residues might be assigned as oxidized based solely on a sequence coverage and on amino acid reactivity toward hydroxyl radicals. Thus, the overall extent of oxidation is a combination of the sum of extents of oxidations of residues located in each region and their exposure toward solvent. In contrast, fragmenting isotopically depleted protein provides increased signal-to-noise ratio, which resulted in elevated sequence coverage, and thus, more residues might be assigned as oxidized ones (Figure 5B). Analysis of the ID protein revealed changes in the oxidation patterns of several residues of FOXO4-DBD upon binding to IRE (Figure 5C). These were further visualized in an *in silico* model of FOXO4-IRE (Figure 6, Figure S10). The first detected oxidized residues, K10, N16 and W18, were protected upon the complex formation and

https://doi.org/10.1021/acs.analchem.3c03759
Anal. Chem. 2024, 96, 1478−1487

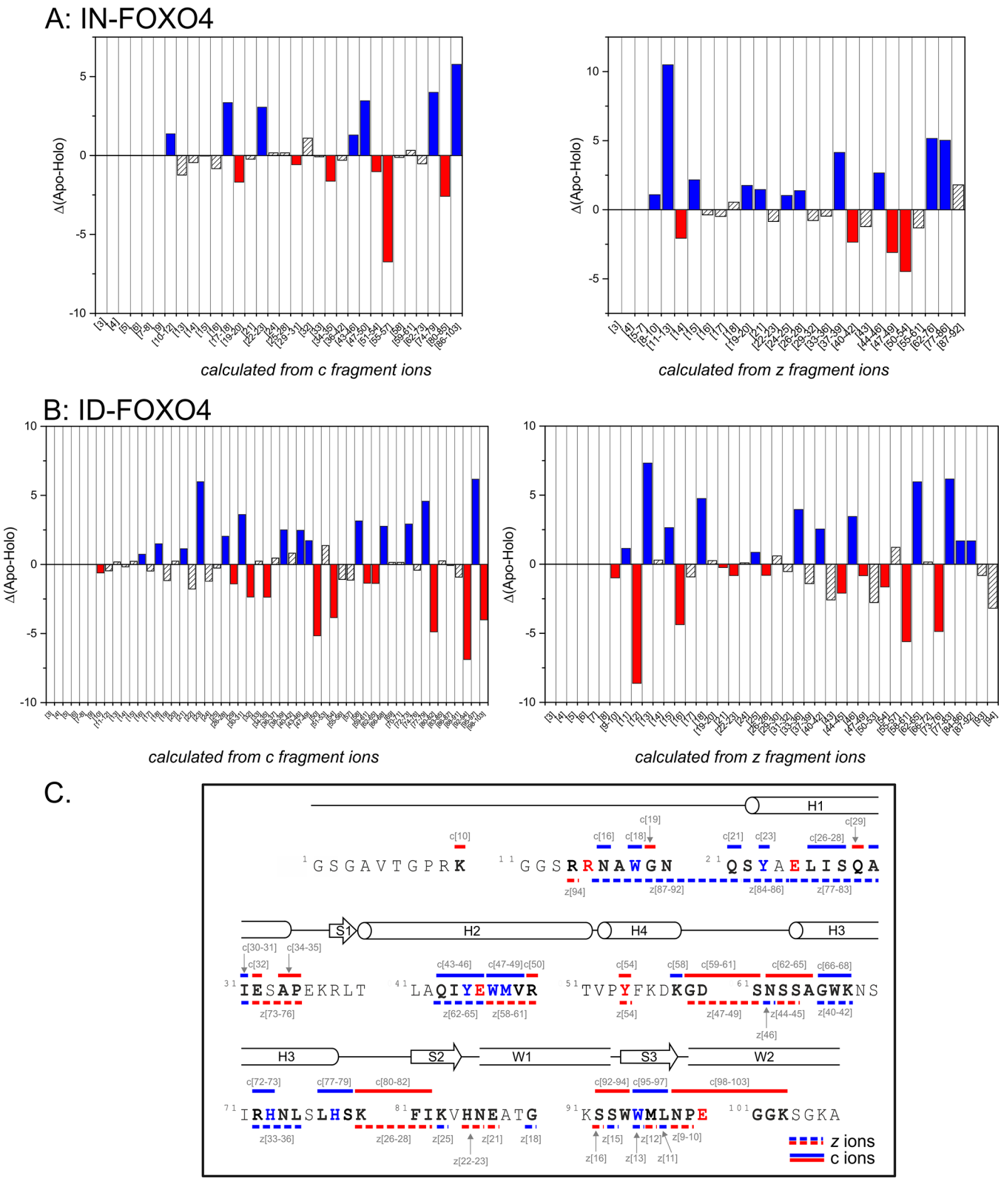


Figure 5. Plots indicating changes in oxidation rates between apo and holo forms of isotopically natural FOXO4 (A) and isotopically depleted FOXO4 (B); assessed by ECD fragmentation in multiCASI mode (Figure S5, Figure S8). Blue histograms represent changes in which region/residue was protected by IRE, and red histograms represent changes which resulted in deprotection of region/residue by IRE. (C) Changes obtained in ID-FOXO4-DBD were visualized into the differential oxidation map of FOXO4-DBD. The bold sequence represents spatial resolution achieved by fragmentation of isotopically depleted FOXO4-DBD. Colored residues were also detected by bottom-up analysis, as shown in Figure S11B and Table S1.

https://doi.org/10.1021/acs.analchem.3c03759
Anal. Chem. 2024, 96, 1478−1487

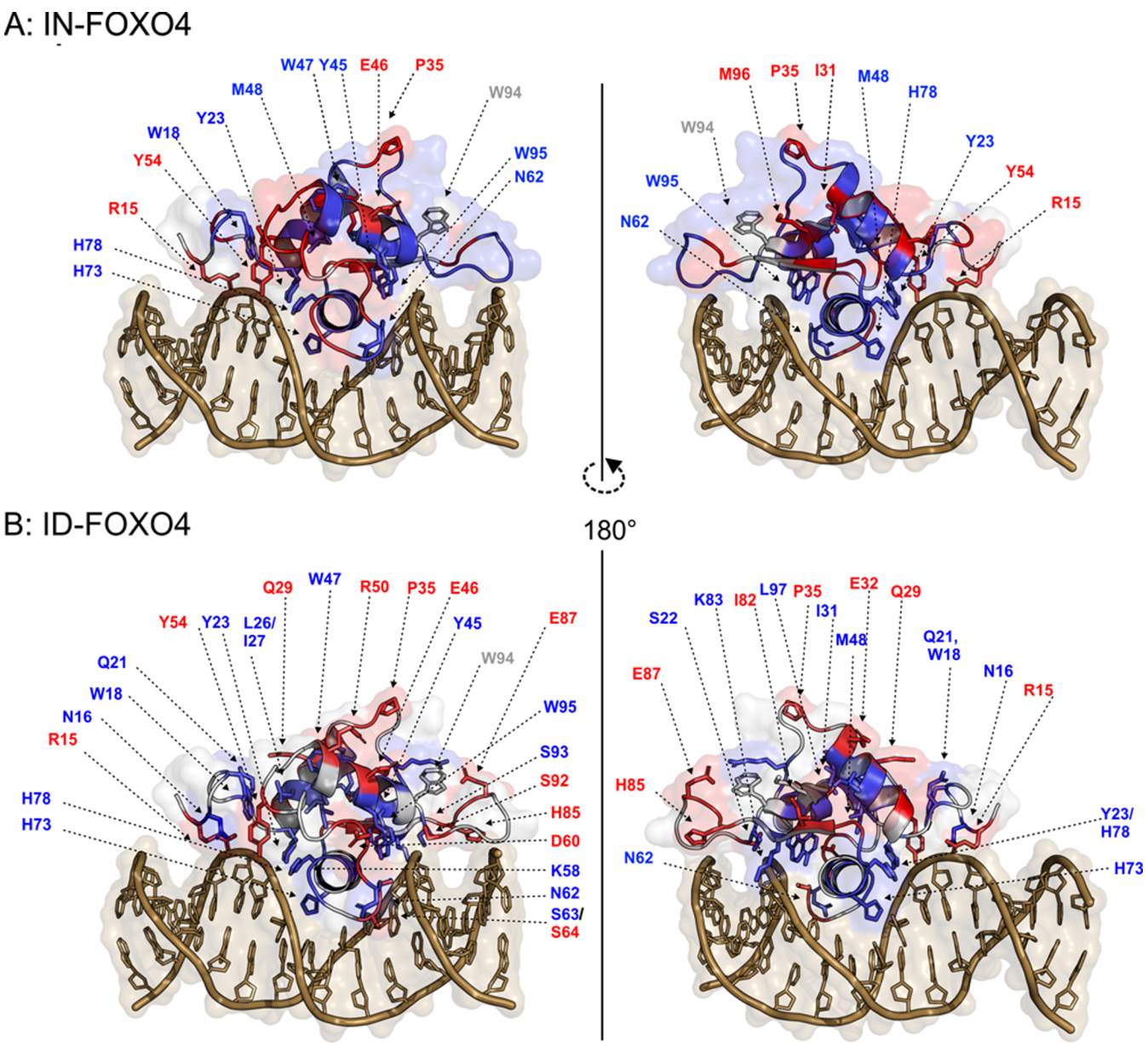


Figure 6. An *in silico* structural model of FOXO4-DBD·IRE (PDB template 3L2C)[22] with the highlighted differently oxidized regions/residues detected by both top-down analyses for natural version (A) or depleted version (B) of FOXO4-DBD. The individual residues detected in either bottom-up approach or deduced from top-down were highlighted in the model and colored. Blue: regions/residues detected as more modified in apo form; red: regions/residues detected as more modified in holo form.

deduced from c[10], c[16], and c[18] fragment ions. Residues N16 and W18 were found to be directly interacting with the IRE according to the structural model (Figure 6) and also according to the previous study.[51] This is also consistent with the z[87−92] region. Helix H1 (S22-A34) contains several residues that were protected by IRE, in particular residues Q21, Y23, L26/I27, and I31, which were covered by c[21], c[23], c[26−28], c[30−31], z[77−83], and z[84−86]. This is in agreement with the previously published HDX[8] and structural studies.[51,63] However, we have also observed a higher oxidation rate for Q29, E32, and P35 deduced from c[29], c[32], c[34−35], and z[73−76], located throughout the H1 helix and intervening loop far away from the protein−DNA interface (Figure 6, Figure S10). Residues in helices H2 and H4, namely, Y45, W47/M48, and Y54,[51] showed different oxidation patterns, with some being less oxidized and some being more oxidized upon protein binding to IRE. The residues were covered by regions c[43−46], c[47−49], z[62−65], c[54], and z[54]. Moreover, R50, oriented toward the solvent,[64] was found to be more oxidized in the presence of IRE (Figure 5C, Figure 6).[1,51] The protection of K58 and deprotection of D60 residue was deduced based on regions c[58], c[59−61], and z[47−49]. Oxidation of N62 is deduced from z[46], alongside the region c[62−65]/z[44−45] which pinpoints the oxidation to both S63 and S64 residues. One may hypothesize that both N62 and S63 are protected upon the complex formation,[3,51] while S64 residue is oriented more toward the solvent and thus might be more oxidized upon complex formation (Figure 6).[51] A helix H3 (G66-H78), the main contributor of interaction with the major groove of IRE,[51] was found to be heavily protected. Residues of helix H3, namely, W67, H73, and H78, are covered by regions c[66−

https://doi.org/10.1021/acs.analchem.3c03759
Anal. Chem. 2024, 96, 1478−1487

68], c[72−73], c[77−79], z[33−36], and z[40−42] and display the direct protection of helix by the major groove of IRE. Next, there has been a multitude of residues oxidized on strand S2 and wing W1 (F81−S93) bearing different oxidation patterns. Residues K83 and S93, covered by z[25] and z[15], were both protected. This observations are supported by a structural model and complementary mutagenesis studies[3] demonstrating direct interaction of K83 and S93 residues with the DNA. Even though F81 is located next to I82 and also is more reactive, we pinpoint hotspot to the I82, because residue F81 is located away from the solvent and interacts with IRE in our model. Thus, residues I82, H85, E87, and S92, covered by regions z[26−28], z[22−23], z[21], and z[16], were detected as more oxidized.

Finally, data from isotopically depleted samples allowed us to resolve the solvent accessibility of residues located at strand S3 at a single residue resolution. Residues W94, W95, and M96 are covered by regions z[12], z[13], z[14], c[92−94], and c[95−97]. Overall effect of IRE binding is the stabilization of strand S3 and deprotection and protection of W94 and W95 residues, respectively, as W94 interacts with IRE.[3,51,63] The analysis of smaller *z* ions led to the assignment of L97 and P99 as oxidized ones based on [z8] and [z10] fragment ions, respectively. Altogether, the isotopic depletion and selective gas-phase enrichment and fragmentation of oxidized ions lead to the detection of 30 residues, namely K10, N16, W18, Q21, Y23, L26/I27, Q29, I31, E32, P35, Y45, R50, K58, D60, N62, S63, S64, H73, H78, I82, K83, H85, E87, S92, S93, W94, W95, M96, L97, and P99. Out of them, 22 residues were deduced solely from the isotopically depleted top-down data set. The other 7 were also detected using the bottom-up analysis described below.

Bottom-Up Analysis. To assess the top-down MS data, the same isotopically natural samples were also subjected to a bottom-up analysis. The LysC and Trypsin/LysC digestion of IN samples yielded 22 peptides, providing a sequence coverage of over 96% (see Figure S11A). Across the sequence of FOXO4-DBD, 18 residues were modified (Figure S11B and Table S1). Nine residues, namely, W18, Y23, Y45, W47, M48, H73, H78, W94[#], and W95, showed decreased oxidation in the complex with IRE, indicating protection. By contrast, residues R15, E46, Y54, W94, W94[##], and E100 showed increased oxidation/modification levels upon complex formation. Five residues were not affected by IRE binding. The single-residue resolution of the bottom-up approach supported data from our *ab initio* model, in line with previously published mutagenesis studies,[3] which reported that most residues involved in the interaction with the major groove of DNA were less oxidized/modified.

The agreement between the bottom-up and top-down results demonstrates the usefulness of the top-down and complementarity of both techniques. For instance, residues W18, Y23, Y45, Y54, H73, H78, W94, W95, and M96 differed in the extent of their modification upon IRE binding in the bottom-up data set, as observed by top-down MS analysis. This confirms that top-down MS analysis is a reliable technique for detecting oxidation levels at different residues.

Notwithstanding the above, some differences were detected between the results of the two techniques: (i) Bottom-up workflow benefits from LC for resolving even isobaric/isomeric modifications prior to mass spectrometric detection, a feature that has been referred to as "sub amino acid resolution".[34] As a case in point, oxidized W94 can have several forms depending on its position on the indole ring. This finding was also detected by the top-down workflow, albeit only as a cumulative modification, represented as a sum of individual extents and without indicating the exact position of the modification (Figure S12). (ii) Limited sequence coverage, especially for larger proteins, is a known weakness of top-down techniques in general. Thus, residue oxidation determined by bottom-up can be overlooked by top-down approaches due to the lack of usable fragment ions intense enough for detection, as in the case of W47/M48 residues, both of which were detected in bottom-up but not individually in the top-down spectra. (iii) Bottom-up analysis of all proteoforms present in the sample also detects modifications other than the +16 Da[22] due to oxidation resulting from enzymatic digestion, as shown by R15/R72 deguanidylation, E25, E46, and E100 decarboxylation, and H73 and H78 conversion into aspartate in our bottom-up data set (Figure S11, Table S1). In the top-down experiment, we performed targeted gas-phase accumulation of species with mass increased by +16 Da (indicative of oxidative modification) prior to fragmentation. For this reason, all other modifications were not identified in the top-down experiment. Nevertheless, this limitation could be easily circumvented by including these other theoretical modifications in the accumulation, if necessary. Despite requiring the identification of the species to be included, selective gas-phase accumulation is a great advantage of top-down analysis for significantly enriching the included forms.

In our experiments, this enrichment enabled the detection and assignment of 22 new oxidized residues that were not detected by bottom-up MS analysis, as discussed above. These oxidations were impossible to detect by a bottom-up workflow given the low limit of detection, which is especially limiting for forms with low levels of oxidation and concentrations in the mixture of trypsinated peptides. As such, the benefits of top-down MS analysis in improving the limit of detection can actually outweigh its other limitations.

One obstacle to the broader use of isotopic depletion in MS analysis is the availability of processing software. The current software portfolio is restricted to the natural occurring isotopic distribution, and the data deconvolution mainly relies on averagine function.[45] This function does not allow the deconvolution of an altered isotopic pattern. Nevertheless, this problem may also be solved in the near future by adopting ion deconvolution using measured collision cross sections of trapped ions.[65] For larger proteins, MSMS spectra are necessarily complex, regardless of using protein isotope depletion. However, for small and midsized proteins,[41] the advantage of isotopic depletion is significant.

CONCLUSION

Protein isotope depletion improves the detection and quantification of FPOP oxidation by a top-down MS analysis. This approach has three main advantages: (i) more precisely isolating singly oxidized ions in quadrupole filter prior to fragmentation; (ii) enhancing the intensities of unmodified and oxidized fragment ions; and (iii) improving the resolution of fragment ions in MS/MS spectra by reducing the number of isotopic peaks and thus reducing the overlap of existing peaks, including those of individual isotopes.

Combining isotopic depletion with top-down analysis of FPOP samples boosts sequence coverage by 19% and identifies 22 more oxidized residues in comparison to bottom-up MS analysis. Nevertheless, bottom-up and top-down analyses show

https://doi.org/10.1021/acs.analchem.3c03759
Anal. Chem. 2024, 96, 1478−1487

highly consistent results, demonstrating their complementarity. The more detailed information on the interaction interface obtained in this study enables the *ab initio* design of FOXO4·IRE complex formation. Even beyond interactions between transcription factors and DNA, the approach reported in this study holds great promise for future research of noncovalent interactions by top-down MS, particularly for more complex biomolecular assemblies, whose interaction dynamics often remains unclear in solution.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.3c03759.

Additional methods: Expression and purification of isotopically natural and depleted FOXO4-DBD, top-down data processing, electrophoretic mobility shift assay; Additional Figures: (Figure S1) FOXO4 sequence with both top-down and wild-type numbering, (Figure S2) ESI-MS spectra of desalted protein samples, (Figure S3) EMSA gel, (Figure S4) broadband ECD spectrum of IN-/ID-FOXO4-DBD, (Figure S5) quantified ions of IN-FOXO4-DBD, (Figure S6) zoom of $[c4]^{1+}$ and $[c5]^{1+}$ ions, (Figure S7) zoom of $[z3]^{1+}$ and $[z4]^{1+}$ ions, (Figure S8) quantified ions of ID-FOXO4-DBD, (Figure S9) zoom-in of ECD MSMS spectra of IN-/ID-FOXO4-DBD, (Figure S10) ab initio model of FOXO4-IRE with wild-type numbering, (Figure S11) bottom-up analysis of IN-FOXO4-DBD, (Figure S12) quantified extent of oxidation of W94 residue; Additional table: (Table S1) all modifications identified in bottom-up approach (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Petr Novák − *Institute of Microbiology of the Czech Academy of Sciences, 14220 Prague, Czech Republic; Department of Biochemistry, Faculty of Science, Charles University, 12843 Prague, Czech Republic;* orcid.org/0000-0001-8688-529X; Phone: +420 325 873 610; Email: pnovak@biomed.cas.cz; Fax: +420 241 062 156

Authors

Marek Polák − *Institute of Microbiology of the Czech Academy of Sciences, 14220 Prague, Czech Republic; Department of Biochemistry, Faculty of Science, Charles University, 12843 Prague, Czech Republic*

Jiří Černý − *Laboratory of Structural Bioinformatics of Proteins, Institute of Biotechnology of the Czech Academy of Sciences, 14220 Prague, Czech Republic;* orcid.org/0000-0002-1969-9304

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.3c03759

Author Contributions

Conceptualization - P.N. and M.P.; M.P. performed all experiments and analyzed data. J.C. performed homology modeling. M.P. and P.N. wrote and edited the manuscript. All authors have approved the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Michael Volný and Carlos V. Melo for helpful discussions and editing of the manuscript. This work was mainly financially supported by the NPO-NEURO-EXCELLES (LX22NPO5107), Czech Science Foundation (22-27695S), the European Commission H2020 (EU_FT-ICR_MS grant agreement ID: 731077 and EPIC-XS - grant agreement ID: 823839). Additional institutional and facility support from the Academy of Sciences of the Czech Republic (RVO: 61388971), Grant Agency of Charles University (359521), the Ministry of Education of the Czech Republic (Structural mass spectrometry CF - LM2018127 CIISB), ELIXIR-CZ (LM2023055), and the European Regional Development Funds (CZ.1.05/1.1.00/02.0109 BIOCEV) are gratefully acknowledged. We also thank Julie Winterová for helpful statistical data analysis.

■ REFERENCES

(1) Polák, M.; Yassaghi, G.; Kavan, D.; Filandr, F.; Fiala, J.; Kukačka, Z.; Halada, P.; Loginov, D. S.; Novák, P. *Anal. Chem.* **2022**, *94* (7), 3203−3210.

(2) Obsil, T.; Obsilova, V. *Biochim. Biophys. Acta - Mol. Cell Res.* **2011**, *1813* (11), 1946−1953.

(3) Vacha, P.; Zuskova, I.; Bumba, L.; Herman, P.; Vecer, J.; Obsilova, V.; Obsil, T. *Biophys. Chem.* **2013**, *184*, 68−78.

(4) Pandey, P.; Hasnain, S.; Ahmad, S. Protein-DNA Interactions. In *Encyclopedia of Bioinformatics and Computational Biology*; Elsevier, 2019; pp 142−154. .

(5) Lambert, S. A.; Jolma, A.; Campitelli, L. F.; Das, P. K.; Yin, Y.; Albu, M.; Chen, X.; Taipale, J.; Hughes, T. R.; Weirauch, M. T. *Cell* **2018**, *172* (4), 650−665.

(6) Hagenbuchner, J.; Obsilova, V.; Kaserer, T.; Kaiser, N.; Rass, B.; Psenakova, K.; Docekal, V.; Alblova, M.; Kohoutova, K.; Schuster, D.; Aneichyk, T.; Vesely, J.; Obexer, P.; Obsil, T.; Ausserlechner, M. J. *Elife* **2019**, *8*, e48876.

(7) Filandrová, R.; Vališ, K.; Černý, J.; Chmelík, J.; Slavata, L.; Fiala, J.; Rosůlek, M.; Kavan, D.; Man, P.; Chum, T.; Cebecauer, M.; Fabris, D.; Novák, P. *Structure* **2021**, *29* (4), 345−356.e8.

(8) Slavata; Chmelík; Kavan; Filandrová; Fiala; Rosůlek; Mrázek; Kukačka; Vališ; Man; Miller; McIntyre; Fabris; Novák. *Biomolecules* **2019**, *9* (10), 535.

(9) Scalabrin, M.; Dixit, S. M.; Makshood, M. M.; Krzemien, C. E.; Fabris, D. *Methods* **2018**, *144*, 64−78.

(10) Sperry, J. B.; Wilcox, J. M.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (6), 887−890.

(11) Ma, L.; Fitzgerald, M. C. *Chem. Biol.* **2003**, *10* (12), 1205−1213.

(12) Sperry, J. B.; Shi, X.; Rempel, D. L.; Nishimura, Y.; Akashi, S.; Gross, M. L. *Biochemistry* **2008**, *47* (6), 1797−1807.

(13) Gau, B. C.; Chen, H.; Zhang, Y.; Gross, M. L. *Anal. Chem.* **2010**, *82* (18), 7821−7827.

(14) Chen, J.; Cui, W.; Giblin, D.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (8), 1306−1318.

(15) Manzi, L.; Barrow, A. S.; Hopper, J. T. S.; Kaminska, R.; Kleanthous, C.; Robinson, C. V.; Moses, J. E.; Oldham, N. J. *Angew. Chemie Int. Ed.* **2017**, *56* (47), 14873−14877.

(16) Zhang, M. M.; Rempel, D. L.; Gross, M. L. *Free Radic. Biol. Med.* **2019**, *131*, 126−132.

(17) Cheng, M.; Zhang, B.; Cui, W.; Gross, M. L. *Angew. Chemie - Int. Ed.* **2017**, *56* (45), 14007−14010.

(18) Fojtík, L.; Fiala, J.; Pompach, P.; Chmelík, J.; Matoušek, V.; Beier, P.; Kukačka, Z.; Novák, P. *J. Am. Chem. Soc.* **2021**, *143* (49), 20670−20679.

(19) Sharp, J. S.; Becker, J. M.; Hettich, R. L. *Anal. Biochem.* **2003**, *313* (2), 216−225.

(20) Hambly, D. M.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (12), 2057−2063.

(21) Wang, L.; Chance, M. R. *Mol. Cell. Proteomics* **2017**, *16* (5), 706−716.
(22) Liu, X. R.; Zhang, M. M.; Gross, M. L. *Chem. Rev.* **2020**, *120* (10), 4355−4454.
(23) Xu, G.; Chance, M. R. *Chem. Rev.* **2007**, *107* (8), 3514−3543.
(24) Xu, G.; Chance, M. R. *Anal. Chem.* **2004**, *76* (5), 1213−1221.
(25) Xu, G.; Takamoto, K.; Chance, M. R. *Anal. Chem.* **2003**, *75* (24), 6995−7007.
(26) Xu, G.; Chance, M. R. *Anal. Chem.* **2005**, *77* (8), 2437−2449.
(27) Charvátová, O.; Foley, B. L.; Bern, M. W.; Sharp, J. S.; Orlando, R.; Woods, R. J. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (11), 1692−1705.
(28) Pan, Y.; Stocks, B. B.; Brown, L.; Konermann, L. *Anal. Chem.* **2009**, *81* (1), 28−35.
(29) Watkinson, T. G.; Calabrese, A. N.; Ault, J. R.; Radford, S. E.; Ashcroft, A. E. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (1), 50−55.
(30) Lu, Y.; Zhang, H.; Niedzwiedzki, D. M.; Jiang, J.; Blankenship, R. E.; Gross, M. L. *Anal. Chem.* **2016**, *88* (17), 8827−8834.
(31) Gupta, S.; Bavro, V. N.; D'Mello, R.; Tucker, S. J.; Vénien-Bryan, C.; Chance, M. R. *Structure* **2010**, *18* (7), 839−846.
(32) Loginov, D. S.; Fiala, J.; Brechlin, P.; Kruppa, G.; Novak, P. *Biochim. Biophys. Acta - Proteins Proteomics* **2022**, *1870* (2), No. 140735.
(33) Cornwell, O.; Bond, N. J.; Radford, S. E.; Ashcroft, A. E. *Anal. Chem.* **2019**, *91* (23), 15163−15170.
(34) Cornwell, O.; Radford, S. E.; Ashcroft, A. E.; Ault, J. R. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (12), 2413−2426.
(35) Yassaghi, G.; Kukačka, Z.; Fiala, J.; Kavan, D.; Halada, P.; Volný, M.; Novák, P. *Anal. Chem.* **2022**, *94* (28), 9993−10002.
(36) Tomášková, N.; Novák, P.; Kožár, T.; Petrenčáková, M.; Jancura, D.; Yassaghi, G.; Man, P.; Sedlák, E. *Int. J. Biol. Macromol.* **2021**, *174*, 413−423.
(37) Donnelly, D. P.; Rawlins, C. M.; DeHart, C. J.; Fornelli, L.; Schachner, L. F.; Lin, Z.; Lippens, J. L.; Aluri, K. C.; Sarin, R.; Chen, B.; Lantz, C.; Jung, W.; Johnson, K. R.; Koller, A.; Wolff, J. J.; Campuzano, I. D. G.; Auclair, J. R.; Ivanov, A. R.; Whitelegge, J. P.; Paša-Tolić, L.; Chamot-Rooke, J.; Danis, P. O.; Smith, L. M.; Tsybin, Y. O.; Loo, J. A.; Ge, Y.; Kelleher, N. L.; Agar, J. N. *Nat. Methods* **2019**, *16* (7), 587−594.
(38) Petrenčáková, M.; Filandr, F.; Hovan, A.; Yassaghi, G.; Man, P.; Kožár, T.; Schwer, M. S.; Jancura, D.; Plückthun, A.; Novák, P.; Miškovský, P.; Bánó, G.; Sedlák, E. *Sci. Rep.* **2020**, *10* (1), 1−15.
(39) Kellersberger, K. A.; Yu, E.; Kruppa, G. H.; Young, M. M.; Fabris, D. *Anal. Chem.* **2004**, *76* (9), 2438−2445.
(40) Valkenborg, D.; Mertens, I.; Lemière, F.; Witters, E.; Burzykowski, T. The Isotopic Distribution Conundrum. *Mass Spectrometry Reviews*; John Wiley & Sons, Ltd, January 1, 2012; pp 96−109. .
(41) Compton, P. D.; Zamdborg, L.; Thomas, P. M.; Kelleher, N. L. *Anal. Chem.* **2011**, *83* (17), 6868−6874.
(42) Marshall, A. G.; Senko, M. W.; Li, W.; Li, M.; Dillon, S.; Guan, S.; Logan, T. M. *J. Am. Chem. Soc.* **1997**, *119* (2), 433−434.
(43) Bou-Assaf, G. M.; Chamoun, J. E.; Emmett, M. R.; Fajer, P. G.; Marshall, A. G. *Anal. Chem.* **2010**, *82* (8), 3293−3299.
(44) Charlebois, J. P.; Patrie, S. M.; Kelleher, N. L. *Anal. Chem.* **2003**, *75* (13), 3263−3266.
(45) Zubarev, R. A.; Demirev, P. A. *J. Am. Soc. Mass Spectrom.* **1998**, *9* (2), 149−156.
(46) Gallagher, K. J.; Palasser, M.; Hughes, S.; Mackay, C. L.; Kilgour, D. P. A.; Clarke, D. J. *J. Am. Soc. Mass Spectrom.* **2020**, *31* (3), 700−710.
(47) Popovic, Z.; Anderson, L. C.; Zhang, X.; Butcher, D. S.; Blakney, G. T.; Zubarev, R. A.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2023**, *34*, 137−144.
(48) Loginov, D. S.; Fiala, J.; Chmelik, J.; Brechlin, P.; Kruppa, G.; Novak, P. *ACS Omega* **2021**, *6* (15), 10352−10361.
(49) Li, K. S.; Shi, L.; Gross, M. L. *Acc. Chem. Res.* **2018**, *51* (3), 736−744.
(50) Perez-Riverol, Y.; Bai, J.; Bandla, C.; García-Seisdedos, D.; Hewapathirana, S.; Kamatchinathan, S.; Kundu, D. J.; Prakash, A.; Frericks-Zipper, A.; Eisenacher, M.; Walzer, M.; Wang, S.; Brazma, A.; Vizcaíno, J. A. *Nucleic Acids Res.* **2022**, *50* (D1), D543−D552.
(51) Boura, E.; Rezabkova, L.; Brynda, J.; Obsilova, V.; Obsil, T. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (12), 1351−1357.
(52) Flores, S. C.; Altman, R. B. *RNA* **2010**, *16* (9), 1769−1778.
(53) Flores, S. C.; Bernauer, J.; Shin, S.; Zhou, R.; Huang, X. *Brief. Bioinform.* **2012**, *13* (4), 395−405.
(54) Černý, J.; Božíková, P.; Svoboda, J.; Schneider, B. *Nucleic Acids Res.* **2020**, *48* (11), 6367−6381.
(55) Černý, J.; Božíková, P.; Malý, M.; Tykač, M.; Biedermannová, L.; Schneider, B. *Acta Crystallogr. Sect. D Struct. Biol.* **2020**, *76* (9), 805−813.
(56) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. *SoftwareX* **2015**, *1*−*2*, 19−25.
(57) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696−3713.
(58) Liebl, K.; Zacharias, M. *J. Chem. Theory Comput.* **2021**, *17* (11), 7096−7105.
(59) Xu, G.; Chance, M. R. *Anal. Chem.* **2005**, *77* (14), 4549−4555.
(60) Niu, B.; Gross, M. L. MS-Based Hydroxyl Radical Footprinting: Methodology and Application of Fast Photochemical Oxidation of Proteins (FPOP). In *Mass Spectrometry-Based Chemical Proteomics*; Wiley, 2019; pp 363−416. .
(61) Yin, V.; Mian, S. H.; Konermann, L. *Chem. Sci.* **2019**, *10* (8), 2349−2359.
(62) Liu, X. R.; Zhang, M. M.; Gross, M. L. *Chemical Reviews* **2020**, 4355−4454.
(63) Boura, E.; Silhan, J.; Herman, P.; Vecer, J.; Sulc, M.; Teisinger, J.; Obsilova, V.; Obsil, T. *J. Biol. Chem.* **2007**, *282* (11), 8265−8275.
(64) Obsilova, V.; Vecer, J.; Herman, P.; Pabianova, A.; Sulc, M.; Teisinger, J.; Boura, E.; Obsil, T. *Biochemistry* **2005**, *44* (34), 11608−11617.
(65) James, V. K.; Sanders, J. D.; Aizikov, K.; Fort, K. L.; Grinfeld, D.; Makarov, A.; Brodbelt, J. S. *Anal. Chem.* **2022**, *94* (45), 15613−15620.

https://doi.org/10.1021/acs.analchem.3c03759
Anal. Chem. 2024, 96, 1478−1487

Attached Publications:

Publication III

Polák, M.; Brinsa, V.; Černý, J.; Kavan, D.; Volný, M.; Kádek, A.; Novák, P.

Hydroxyl Radical Footprinting of Nucleic Acids Coupled to High-Resolution Mass Spectrometry Analysis,

Nucleic Acids Research, manuscript in preparation.

My contribution: DNA and protein sample preparation, FPOP of samples, LC-MS analysis of DNA damaged samples, data analysis, manuscript writing and preparing.

Hydroxyl Radical Footprinting of Nucleic Acids Coupled to High-Resolution Mass Spectrometry Analysis

Marek Polák[$,‡], Vítězslav Brinsa[‡], Jiří Černý[†], Daniel Kavan[$,‡], Michael Volný[$], Alan Kádek[$], and Petr Novák[*,$,‡]

$Institute of Microbiology of the Czech Academy of Sciences, 14220, Prague, Czech Republic; ‡Department of Biochemistry, Faculty of Science, Charles University, 12843, Prague, Czech Republic, †Laboratory of Structural Bioinformatics of Proteins, Institute of Biotechnology of the Czech Academy of Sciences, 14220, Prague, Czech Republic

KEYWORDS: Fast photochemical oxidation of proteins, FOXO4, TEAD1, Insulin response element, IRE, LCMS, hydroxyl radical footprinting, mass spectrometry, multiCASI.

ABSTRACT: DNA transcription is primarily regulated by transcription factors (TFs), proteins that activate or inhibit this process. Protein-DNA complexes and their interactions have long been studied with various methods, such as hydroxyl radical footprinting of DNA and MS-driven approaches. Yet, despite these efforts, the lack of detailed information on their interactions continues to hinder our ability to understand their mechanism of action in depth. Here, we aim at leveraging Fast Photochemical Oxidation of Proteins (FPOP) quench-flow system to analyze hydroxyl radical-induced DNA damage. As a model system, we used double-stranded DNA (dsDNA), namely Insulin Response Element (IRE), both with and without the DNA-binding domain of the transcription factor Forkhead box O4 (FOXO4). After FPOP-induced fragmentation of IRE with and without FOXO4, the samples were analyzed by LC-MS(MS). Performing LC separation prior to MS detection enabled us to separate complimentary fragments bearing phosphates, hydroxyls, and other products, e.g., phosphoglycerate products, and singly oxidized fragments while simultaneously providing us with an unbiased high-resolution dataset. Quantification of the fragments revealed a consistent outcome, indicating that FOXO4 binding to IRE protects the binding motif. When analyzing damage to a dynamic system of two transcription factors, FOXO4 and TEA domain family member 1 (TEAD1), interacting with a single oligonucleotide, we found that TF binding may induce duplex melting, demonstrating that FPOP can capture minor conformational changes in DNA. In addition, FPOP can also be used to study the higher-order structure of DNA, as shown by single-stranded IRE in solution and in duplex. Combined, our findings highlight the potential of FPOP coupled to MS detection for studying DNA damage, providing high-resolution data from a single experiment. In future studies, this approach may be applied to analyze more complex systems, such nucleosomes, or DNA dynamics during transcription.

Protein-nucleic acid complexes play key roles in cellular processes, including replication, transcription, translation, and DNA repair. In these complexes, proteins known as transcription factors (i) bind to DNA in a specific manner and (ii) regulate transcription.[1,2] Accordingly, their specific DNA-binding sequences must be analyzed in detail to fully understand these complex interactions. However, due to their size and dynamics nature of interaction, these highly complex assemblies are particularly difficult to study with conventional biophysical techniques, such as X-ray crystallography or NMR. However, the method of choice in the first attempts at studying protein-nucleic acid interactions was footprinting.

In footprinting, DNA and protein-DNA complexes are cut by a reagent. The resulting fragments are visualized by denaturing gel electrophoresis, which reveals a "footprinted" region with faded or diminished bands in the cleavage pattern.[3] Using DNAse I to digest a DNA sequence bonded to lac repressor protein, Galas & Smichtz[4] reported the first DNA footprinting method in 1978. A few years later, the most significant breakthrough occurred when Thomas Tullius introduced Fenton chemistry to initiate oxidative DNA strand scission[5,6].

Known as a hydroxyl radical footprinting of DNA, this technique relies on the reaction between the $[Fe(II)EDTA]^{2-}$ complex and hydrogen peroxide, generating OH radicals. These radicals subsequently subtract hydrogens from a deoxyribose moiety, which ultimately leads to DNA scission[7,8]. What makes this technique powerful is its ability to provide high-resolution data on protein-DNA complexes, down to a single nucleotide[9]. In experiments, DNA samples are radiolabeled with ^{32}P, mostly at the 5' termini, and the fragments are then separated by denaturing electrophoresis or capillary electrophoresis.

Recent research has also explored alternative ways of generating OH radicals from water, such as γ-ray beam or synchrotron X-ray, to study DNA, RNA and their complexes with proteins in solution[10–16]. But while OH chemistry was initially used to map damage to nucleic acids, most studies have focused on modifying and studying protein surfaces. Thanks to advances in mass spectrometry (MS) techniques, including electrospray ionization and LC, OH radicals generated by Fenton chemistry and synchrotron beams have yielded extensive research on protein surfaces and interactions[17–21] and on amino acids[22–25]. Yet, the lifetime of OH radicals generated in Fenton chemistry cannot be controlled and synchrotron availability is rather limited. For these reasons, Hambly and Gross developed an alternative method termed Fast Photochemical Oxidation of Proteins (FPOP) [26].

In FPOP, laser dissociation of hydrogen peroxide provides OH radicals. An FPOP platform consists of syringe pumps, capillary flow paths and an excimer laser. During the continuous mutual mixing of sample and peroxide, solvent-accessible protein residues are oxidized or modified upon hydrogen peroxide dissociation. The quench-flow capillary system avoids overoxidation by controlling the lifetime of radicals and modifies every amino acid[27–29]. Since its inception, both top-down[30–33] and bottom-up[30,34–42] approaches have been applied to study protein surface and interaction interfaces and protein complexes, respectively.

In a previous study, we demonstrated that FPOP can be used to structurally characterize protein sites in protein-dsDNA complexes[30]. However, to our knowledge, no FPOP approach has been developed to analyze damage to nucleic acids in these complexes by high-resolution MS[43]. Nevertheless, MS has been extensively applied to study base modifications in nucleic acids induced by the hydroxyl radical[43] or other reactive oxygen species (ROS)[8,44,45]. In this study, we therefore aim at developing a FPOP quench-flow system coupled to high-resolution MS to analyze hydroxyl radical-induced DNA damage.

Experimental section

Materials and chemicals

All solvents (LC/MS grade) and chemicals were purchased from Merck (Germany) unless stated otherwise. Isotopically labeled water ($H_2{}^{18}O$, 97 %) was obtained from Cambridge Isotopes Laboratories, USA.

Protein expression and purification

FOXO4-DBD was expressed and purified as described previously[46], and the protein sample from the same batch was used in this experiment for a direct comparison.

Protein•DNA complex formation

Insulin Response Element (IRE) and NO2 strands bearing a binding motif and their complementary reverse strands were obtained from Integrated DNA Technologies (Coralville, USA) in HPLC quality. Duplexes were prepared by mixing both strands in an equimolar ratio, heating them to 90 °C for 3 minutes and cooling down to room temperature to form duplex DNA. Subsequently, 150 and 50μM of protein (FOXO4-DBD, TEAD1-DBD) and dsDNA, respectively, were mixed in 150 mM ammonium acetate, pH 6.8, to form a complex and to ensure that all DNA was bonded in a complex. The sequences of the forward strands bearing binding motifs were IRE 5'-GAC TAT CAA AAC AAC GC-3' and NO2 5'-TGG ATT CCT GTA AAC AGT G-3'.

Fast photochemical oxidation of DNA

Glutamine at a final concentration of 10mM was added to each sample prior to FPOP experiment. FPOP oxidation was performed as described previously[46] at a hydrogen peroxide concentration of either 2 or 10mM. Briefly, the mixtures of H_2O_2 and either DNA or protein/DNA complex were irradiated by a single shot of an KrF excimer laser (COMPex 50 type, Coherent Inc., USA) at 248 nm wavelength, 15 Hz frequency, 100 mJ energy, 20 ns pulse duration and 2.24 mJ/cm² radian exposure. The exclusion volume was 16%, and the reaction was quenched by immediate mixing with 75mM methionine. The samples were collected in an Eppendorf tube containing 3,000 U of Catalase (Merck, USA). The protein component was digested by incubating the samples with Proteinase K (1:10 proteinase K/protein ratio) at room temperature for 30 minutes.

LC-MS analysis ESI-FT-ICR analysis

LC separation was performed on an Agilent 1290 Infinity II HPLC system (Agilent Technologies, Santa Clara, CA) equipped with a trap column (ZORBAX 300SB-C18, 5 μm, 0.30 × 5 mm, (Agilent Technologies, Santa Clara, CA) and an analytical column (ZORBAX C18-Extend, 0.3 ×150 mm, 3.5 μm, Agilent Technologies, Santa Clara, CA) in a system heated to 50°C. For this purpose, 50 pmol of DNA products were desalted on a trap column using 10mM ammonium bicarbonate, pH 7, at a flow rate of 10 μl/min for 5 minutes. The DNA products were separated by a 35-minute linear gradient of 2-55% mobile phase B, followed by uptake of solution B to 99% for 3 minutes, 1 min dropping step 99-2% of B followed by column regeneration at 2% of solution B for another 10 minutes. The NO2 products were separated using the same 2-55% gradient of B, which took only 20 min. The binary pump was operated at a flow rate of 10 μL×min^{-1}, and the mobile phases had the following composition: Solution A – 8mM Triethylammonium Acetate, 50mM hexafluoroisopropanol (HFIP) in water, pH 7.0. Solution B – 8mM Triethylammonium Acetate, 50mM hexafluoroisopropanol (HFIP) in 75% methanol, pH 7.0. Ions were then introduced into a SolariX XR mass spectrometer equipped with 15T magnet. An FT was operated in negative ion mode and externally calibrated using negatively charged trifluoroacetate clusters for sub 1 ppm mass accuracy. Data were acquired in broadband mode m/z 200-3500 with a 2M data point transient starting at 200 m/z, whereas ions were accumulated for 0.2 s for each MS scan. The desolvatation temperature was 200 °C, and the time of flight was 1.2 ms.

Data-dependent (DDA) MS/MS acquisition

Data-dependent acquisition was performed using the same LC separation method, at 0.3 s ion accumulation in

MS, averaging 2 spectra per MS scan, and at 1.0 s ion accumulation in MS/MS. In addition, 3 precursor ions were selected and fragmented at a fixed collision energy of 15.5 eV (overall collision energy). Active exclusion was on, and ions were excluded for fragmentation after 1 MS/MS fragmentation for one minute. MS/MS spectra were recorded in broadband mode ranging 207.24 - 2500 amu. Data were acquired using a 512k data point transient.

Data-independent (DIA) MS/MS acquisition

Data-independent acquisition was performed using the separation method described above. In MS, ion accumulation was set to 0.2 s, averaging 2 spectra per MS scan; then, ions were accumulated for 0.8 s per MS/MS scan in quadrupole and fragmented using a fixed collision energy of 16.0 eV (overall collision voltage). MS/MS scans were recorded in broadband mode, ranging 207.24 - 2500 amu, with locked central mass and an isolation window of 1900 m/z. Data were acquired using a 2M data point transient.

Data interpretation

Raw data were processed using Data Analysis 5.3 (Bruker Daltonics). Data were deconvoluted using a Sophistical Numerical Annotation Procedure (SNAP) algorithm (Bruker Daltonics) with the embedded building block, 'averabaseine', as described by Zubarev[47]: C9.75, H12.25, N3.75, O6, P1. Subsequently, a mascot generic file (.mgf) was exported for further processing in LinX software[48]. A Mongo Oligo Calculator v2.06 (available online) and ChemDraw 11 (Perkin Elmer, USA) were then used to calculate and confirm the monoisotopic masses of all possible DNA fragments, respectively. The Mongo Oligo calculator parameters were embedded sequence, monoisotopic mass, negative ion mode, DNA, 5'-OH/5'-phosphate, 3'OH/3'-phosphate termini, molecular mass, electrospray series, and CID fragments. Extracted ion chromatograms of singly-, doubly- and triply charged DNA products were then manually searched in Data Analysis software. Fragments were detected in DIA/DDA MSMS spectra by verifying the isotopic composition of each fragment in a web-based software available online at https://valkenborg-lab.shinyapps.io/pointless4dna/.

The intensities of fragment ions were manually extracted for each fragment to calculate the extent of oxidative damage according to the following equation:

$$Extent\ of\ oxidative\ damage\ of\ fragment\ n\ (\%) = \quad \text{(eq. 1)}$$

$$\frac{\sum I_{fragment\ ion\ n}}{(\sum I_{fragment\ ion\ n} + \sum I_{forward\ or\ reverse\ strand}} \times 100$$

Where $\sum I_{fragment\ ion\ n}$ represents the sum of intensities of monoisotopic peaks of all observed charge states of a particular DNA fragment, and $\sum I_{forward/reverse\ strand}$ represents the sum of intensities of monoisotopic peaks of either Forward or Reverse strands, which yielded a specific n fragment.

Data were expressed as mean ± SD (n=3), statistically analyzed using unpaired student t-test in GraphPad Prism 8.0 software and plotted.

Results and discussion

Since Tullius' pioneering study on hydroxyl radical footprinting of a protein-DNA complex, published in 1986[6], multiple research groups have studied both damage and modifications of nucleic acids, as reported, for instance, in ref. [8,9,16,44,49,50]. Over the years, DNA damage analysis by Fenton chemistry was overcome by FPOP[26], a more convenient platform using hydrogen peroxide as a source of radicals, albeit almost exclusively applied to proteins and their complexes. Based on our recent research efforts[30,46] aimed at gathering structural data on proteins in protein-DNA complexes, we hypothesized that FPOP could also be used to gain structural insights into nucleic acid sites in such complexes. Thus, this study primarily aimed at applying a FPOP platform together with high-resolution MS analysis to analyze nucleic acid damage of a 17-bp-long dsDNA sequence, the Insulin Response Element (IRE).

DNA fragmentation mechanism

The initial experiment included a simple FPOP-induced fragmentation experiment of dsIRE using •OH radicals to study the basic characteristics and the outcome of the experiment. As described above, •OH radicals initiate the oxidative attack by subtracting hydrogen atoms from the backbone and by subsequently fragmenting nucleic acids in two parts. After the experiment, the sample was collected and analyzed by LC-MS operated in negative ion mode. Radicals more frequently attack sugar-backbone hydrogens, as reported previously,[9,50] showing the following reactivity H-5'>H-4'>H-3'≈H-2'≈ H-1', which is directly proportional to their solvent accessibility. In the early stages of hydroxyl radical footprinting experiments, DNA was ^{32}P labelled either at a 5'- or 3'-termini, so the products that were visualized by electrophoresis were either 5'(^{32}P)-[n] fragments bearing 3'-phosphate or 3'-phosphooglycerate (mass increment +58.005 Da)[8,51,52] or 5'-phosphate and 5'-aldehyde, upon 3'(^{32}P) end labelling[53]. In our LC-MS analysis, we observed not only the same products but also their complementary versions, namely [mer]-3'**OH**, [mer]-3'**P**, 5'**P**-[mer], 5'**OH**-[mer], 5'**Aldehyde**-[mer], as shown in Figure 1[8,9,53]. We also detected [mer]-3'**Phosphoglycerate** (PG) products, derived from multiple fragmentation of the deoxyribose moiety of its n+1 fragment[51]. These fragments were observed above the level of detection and thus to a very low extent (Figure S1), which can be explained by single-hit kinetics of the FPOP platform. Furthermore, 5'-aldehyde[mer] was previously reported as the last product, generated after a longer reaction time of Fenton chemistry (~2 min) from its alcohol form[9]. Nevertheless, the characteristic single-hit kinetics of FPOP enabled us to identify both alcohol and aldehyde forms.

Following the last fragmentation into independent fragments, oxygen was further incorporated at a cleavage site, as shown in Figure 1. We also hypothesized that this additional oxygen originated either from the second •OH radical attack or from the water molecule. To test our hypothesis, we performed the same experiment but in a nor-

mal (e.g. $H_2{}^{16}O$) and ^{18}O isotope-enriched ($H_2{}^{18}O$) water environment. Figure S2 shows phosphate/hydroxyl terminated fragments, namely [9mer]-3'**P**/**OH** and **OH**/**P**3'-[8mer], generated during FPOP. Comparing isotopic envelopes of phosphate- or hydroxyl-terminated ends from the reaction in either $H_2{}^{16}O$ or $H_2{}^{18}O$ environment, no visible 2Da changes were detected in $H_2{}^{18}O$. Based on these data, the additional oxygen originates from H_2O_2, but its precise position (phosphate or hydroxyl end) remains unclear. Nevertheless, the monoisotopic masses of all possible fragment ions bearing either phosphate or hydroxyl ends were calculated and subsequently searched in LC-MS data, as shown in Figure 1.

Electromobility shift assay (EMSA)

After detecting the characteristic products of the FPOP reaction by MS, we performed the next experiment to obtain the characteristic 'footprint protection' of the binding motif, as observed in the initial gel electrophoresis experiments. To this end, we tested the ability of FOXO4-DBD to bind to dsIRE by EMSA. Figure S3 shows the image of an electrophoretic gel stained with both GelRed® and Coomassie Blue, revealing a mobility shift when FOXO4 is bound to IRE, in contrast to IRE alone. As a result, a 3:1 FOXO4:dsIRE excess is required for complete IRE binding. Subsequently, we conducted further experiments using this 3:1 FOXO4:dsIRE ratio.

Fast Photochemical Oxidation of DNA

Since the DNA strands were commercially obtained by solid-phase DNA synthesis and purified by HPLC, the commercial stocks might have contained smaller fragments from incomplete solid-phase synthesis. Thus, the initial experiment of protein-DNA footprinting included 3 independent LC-MS runs of both strands from which we extracted ion chromatograms of all possible fragments.

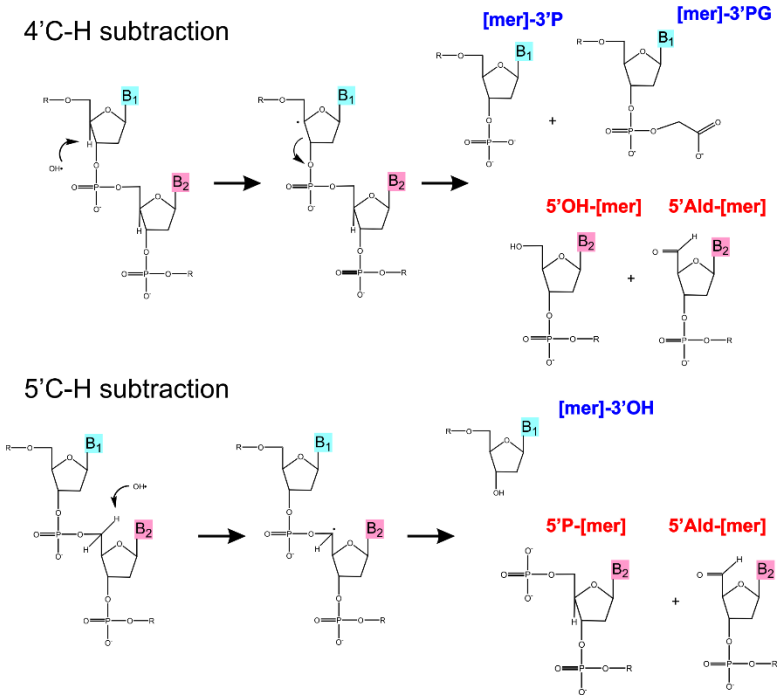


Figure 1. The fragmentation mechanism and products of nucleic acid after FPOP detected in LC-MS trace.

Extracted ion chromatograms of all possible fragment ions, including the full Forward/Reverse 17-mer strands (Figure S4 and Figure S5), showed no significant impurities (<0.5 %) in the form of smaller fragments, which could have biased the dataset. Out of all fragments, only 5'OH-[For-14mer]-3'OH was simultaneously detected in the LC-MS trace; the amount of fragments was calculated (1.99±0.36%) and subtracted from further laser-initiated experiments. We then performed FPOP experiments of IRE with and without FOXO4 at two H_2O_2 concentrations, namely 10mM, commonly used to modify proteins in FPOP experiments, and 2mM H_2O_2. When irradiated, smaller DNA fragments appeared in the LC-MS trace, with shorter retention times, owing to interactions with the reverse phase during LC (Figure 2, Figure S6, Figure S7).

Fragments were then identified using the following methods: (i) by searching DIA/DDA MS/MS spectra for CID fragment ions (*a(-B)-*, *w-*)[54,55], (ii) by comparing the isotopic envelope of DNA fragments with that of an in silico model[56] using web-based software available online, and (iii) by correlating the retention times of n fragments with that of its n+1 fragment from the same strand to facilitate fragment identification (Figure 2). Moreover, we also observed that the phosphate-terminated fragment eluted later than its hydroxyl forms because this fragment more strongly interacts with triethylamine in the mobile phase. Following these steps, out of 128 possible fragments, 47 and 51 complementary fragments originated from forward and reverse strands, respectively, were simultaneously identified after fragmenting dsIRE using 10mM H_2O_2. With 2mM H_2O_2, 48 and 47 fragments from forward and reverse strand were detected, respectively. Moreover, we did not detect

any internal fragments although they may be generated in longer DNA samples in radical reactions[54,57].

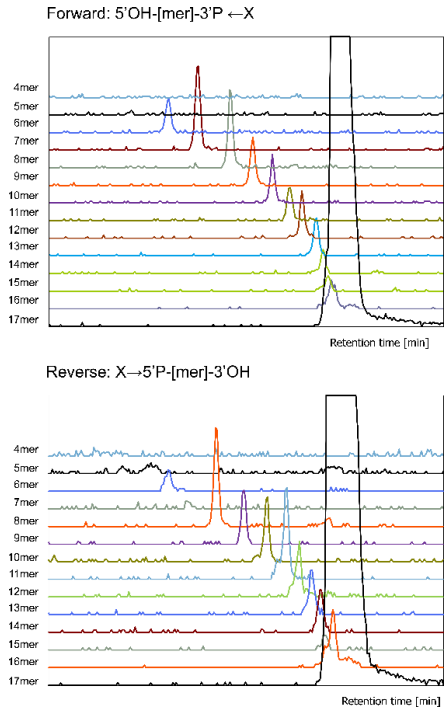


Figure 2. Examples of extracted ion chromatograms of DNA fragments originated from Forward (top panel) or Reverse (bottom panel) strands. The main, zoomed-in signal represents 17mers forward and reverse strands (colored in black).

However, we detected signals represented by n-1 fragments (n=base) with characteristic shorter retention time, as displayed in Figure 2. The arrows in Figure 2 represent the direction of cleavage. Among the products with a unique sequence and thus monoisotopic mass, only one extracted ion chromatogram corresponded to two chromatographically resolved peaks. Further investigation revealed two isobaric compounds with the same mass but scrambled sequence, 5'OH-[For-GACTAT]-3'OH and 3'-OH-[Rev-CTGATA]-5'OH, which were resolved by MS/MS analysis (Figure S8). Thus, both products were quantified individually.

The extent of oxidative damage was then calculated for each fragment ion observed at two H_2O_2 concentrations, with and without FOXO4 (represented by pink and blue histograms, respectively), statistically analyzed and plotted for both strands independently (Figure 3 and Figure 4). These figures show that (i) H_2O_2 initiated fragmentation at both concentrations (2mM and 10mM in this study vs. previously described 200mM in Fenton chemistry[3]), (ii) dsIRE fragmentation with 2mM H_2O_2 yielded ~6-8% of the most abundant fragments, whereas the reaction with 10mM H_2O_2 produced less than ~4% of the most abundant fragments, and (iii) proteins significantly restricted hydrogen subtraction by •OH radicals, thus accounting for FOXO4 footprinted regions around the binding motif/major groove and minor groove[46]. Based on (i) and (ii), even a minor peroxide concentration generates complimentary fragments lacking base specificity, yet nearly the same number of fragments as the higher concentration of H_2O_2. As explained in (iii), the 'footprint' around the binding motif corroborates the findings of previous studies on DNA footprinting.

Notwithstanding the above, these data also showed that fragments were at 2mM H_2O_2 generated at a higher extent than at 10mM H_2O_2. However, labelling agents should yield more fragment products at higher concentrations. Perhaps •OH radicals react more harshly with DNA, converting DNA fragments into other products, as previously described in the literature[8,44]. We hypothesize that fragment release from the duplex exposes bases to oxidation in solution, thus dispersing the MS ion signal. To address this issue, we quantified singly oxidized fragments with covalently oxidized bases observed in MS spectra. Thus, Figure 3 and Figure 4 also display oxidized fragments with and without FOXO4 represented by orange and green histograms, respectively. The results show almost no variation in the level of the oxidized fragments as a function of the H_2O_2 concentration, suggesting that oxidation is a secondary reaction, succeeding fragment release from the duplex. For instance, in the forward strand, the levels of oxidized fragments were ~2 and 0.5% at 2 and 10 mM H_2O_2, respectively. Accordingly, the signal must be dispersed elsewhere and is most likely converted into other unspecified DNA products/lesions[8,44].

We also observed a significantly higher amount of phosphate-terminated fragments. This result may be attributed to two factors. On the one hand, •OH radicals react with every hydrogen of the sugar moiety, yielding 3'- or 5'-phosphate-terminated fragments as the main products[9]. On the other hand, 3'-OH or 5'-OH groups are closer to the sugar moiety, which may trigger secondary reactions on this sugar moiety, thereby dispersing the signal into various products.

As shown in Figure 1, the fragmentation mechanism also indicates that 5'-terminated fragments give rise to both hydroxyl and aldehyde forms[9]. But while aldehyde and alcohol products of larger fragments (5'**OH**-[16mer] to 5'**OH**-[10mer]) were not chromatographically resolved (Figure S9A,B), those of shorter fragments, such as 5'**OH**-[8mer] (Figure S9A,C), displayed chromatographically resolved peaks, corresponding to separate aldehyde and alcohol products. To quantify the amount of both products, we (i) modelled the isotopic envelope of each non-resolved fragment ion[56], (ii) subtracted the theoretical intensity of the third aldehyde isotope from the overall intensity of the third isotope in the spectra to calculate the intensity of the pure alcohol product, and (iii) quantified both aldehyde and alcohol products with and without FOXO4. Figure S9D reveals that both alcohol and aldehyde forms sometimes exhibit different patterns and amounts, presumably due to solvent accessibility of the hydrogens that yield those products. In future experiments with more complex DNA systems, where aldehyde/ alcohol products will most likely not be chromatographically resolved, we recommend quantifying only aldehyde products to minimize the bias introduced by its third isotope in alcohol products.

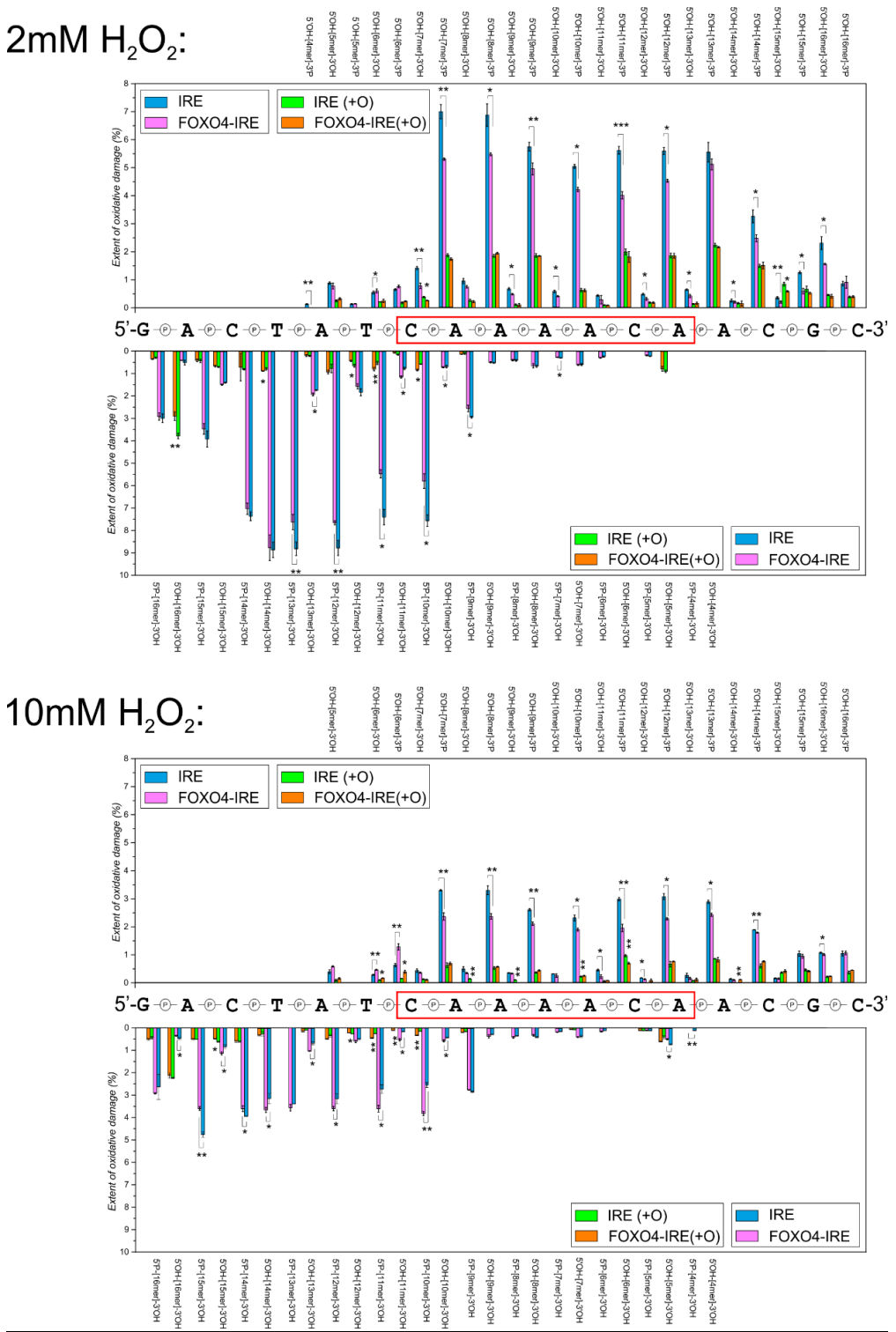


Figure 3. Quantified extent of oxidative damage of the forward strand at two concentrations, 2mM and 10 mM of H_2O_2. Pink and blue histograms represent the extent of oxidative damage with and without FOXO4, respectively, and green and orange histograms represent singly oxidized fragment forms. T-test results are indicated by asterisks above the histograms. T-test legend: * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$), **** ($p \leq 0.0001$).

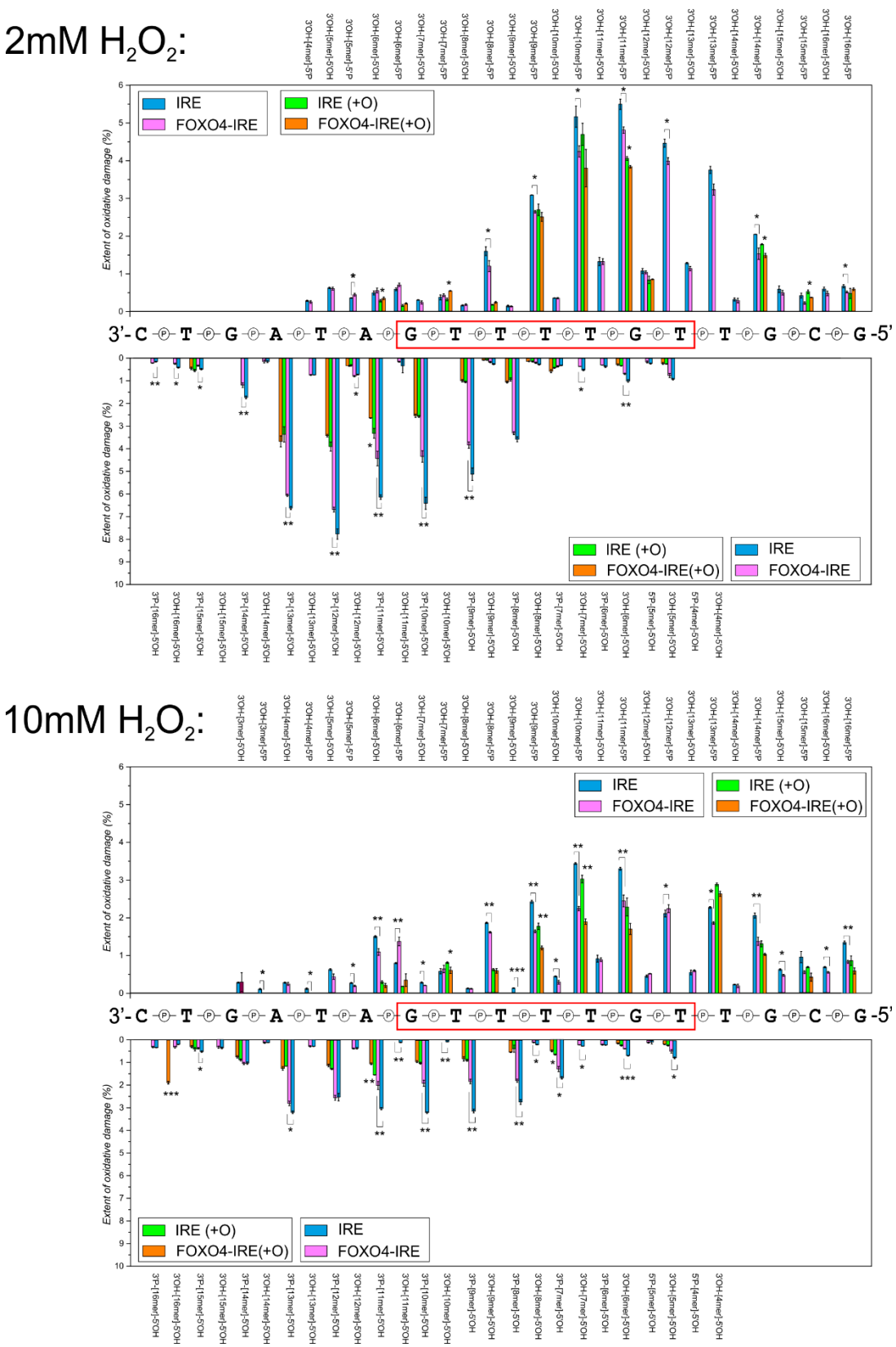


Figure 4. Quantified extent of oxidative damage of forward strand at two concentrations, 2mM and 10 mM of H_2O_2. Pink and blue histograms represent the extent of oxidative damage with and without FOXO4, respectively, and green and orange histograms represent singly oxidized fragment forms. T-test results are indicated by asterisks above histograms. T-test legend: * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$), **** ($p \leq 0.0001$).

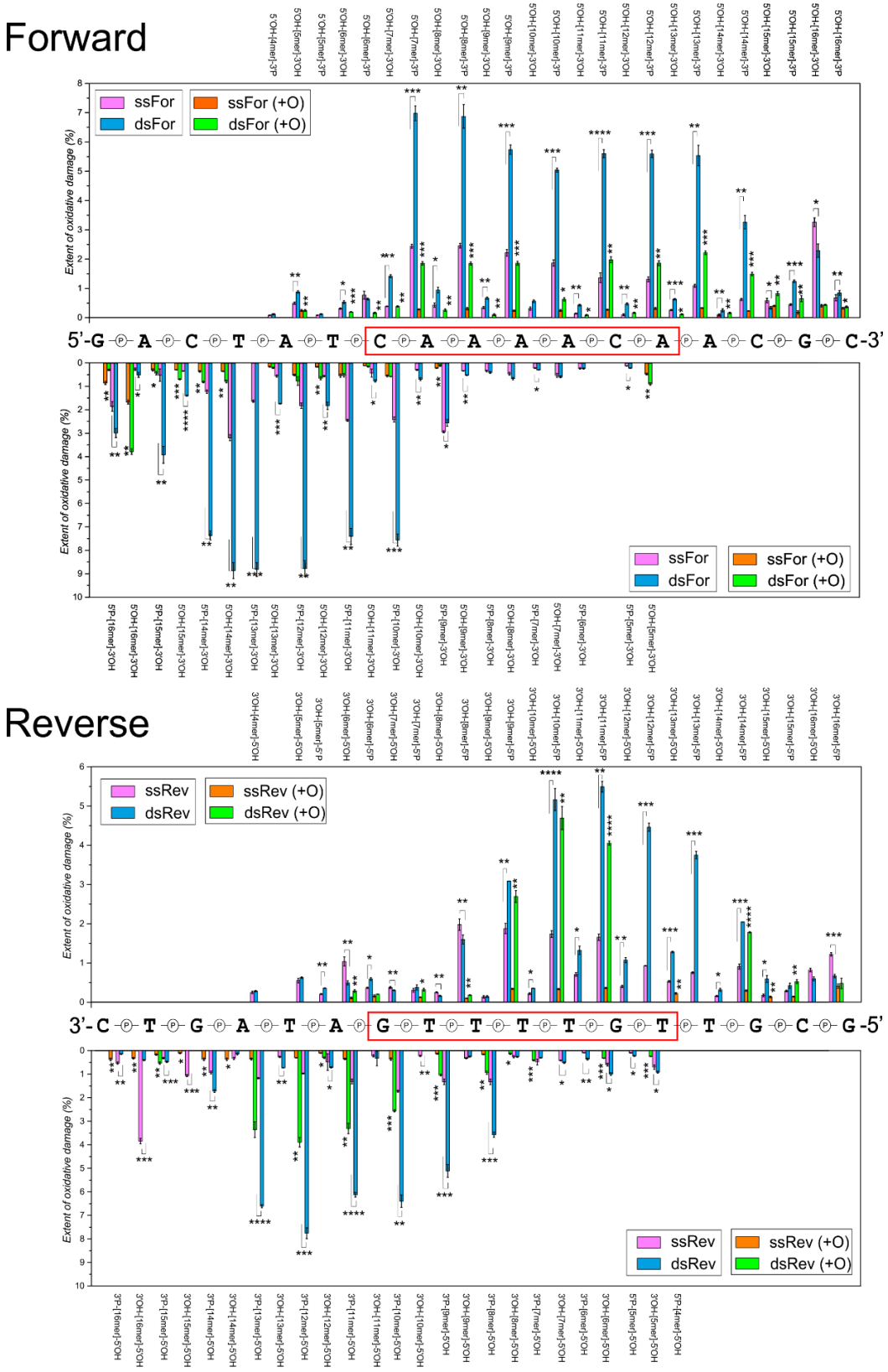


Figure 5. Quantified extent of oxidative damage of forward strand at two concentrations, 2mM and 10 mM of H_2O_2. Blue and pink histograms represent the extent of oxidative damage in the absence and the presence of FOXO4, respectively, the green and orange histograms represent observed singly-oxidized fragment forms. The results of the t-test are indicated by a number of asterisks above the histograms. T-test legend: * ($P \leq 0.05$), ** ($P \leq 0.01$), *** ($P \leq 0.001$), **** ($P \leq 0.0001$).

We also assessed whether this chemistry can be used to study structural changes in ssDNA in solution or interactions with cognate binding partners. Studying fragmentation patterns can provide detailed data on ssDNA interactions with proteins (protein-ssDNA complexes, and protein-RNA complexes) or other single-stranded nucleic acids, forming higher-order structures such as G-quadruplexes, tRNA, and DNA/RNA hairpins[43,58,59]. To this end, we induced ss and duplex IRE fragmentation in a quench-flow capillary system at 2mM H_2O_2. Figure 5 displays quantified DNA fragments, including oxidized ss and ds forms of both forward and reverse strands. Pink histograms represent the extent of damage when only ss were presented in solution, averaging ~2.0% damage. The pink histogram plots also indicate an inconsistency in ss fragmentation. The corresponding level of oxidized fragments is very low (orange histograms), especially when bases are exposed to the solvent, possibly because H'3, H'2 and H'1 are also exposed in ss, and thus the radical subtract/reacts with hydrogens at every position rather than oxidizes bases[44]. Once the strands form duplex DNA, we observed an 'up and down' effect, which should directly correspond to the secondary structure and solvent accessibility of hydrogens. Although we used only 17bp dsDNA, this effect may be also occurred in longer DNA samples, as shown in previous gel experiments[50]. In addition, we found that the level of oxidized fragments is also higher in dsDNA simply because the bases are stabilized in a secondary structure. Once H'3, H'2 and H'1 are stacked, with limited solvent exposure, the hydroxyl radical attacks H4' and H5', the most solvent accessible hydrogens, yielding the characteristic products shown in Figure 1. Therefore, our FPOP platform can also be used to study ssDNA and their higher-order structures or interactions with other biomolecules.

The initial experiment included a case study of FOXO4 and 17-bp-long IRE with a binding motif in the middle of the sequence. But to further demonstrate the applicability of this approach on a different model system, we tested DNA with two different binding motifs. The oligonucleotide NO2 is 19 bp long and combines TEAD1[60] and FOXO4 binding motifs, namely 5'-ATTCC-3' and 5'-CAAAACA-3', respectively. This FPOP experiment should lead to strong protection of the binding motifs. Beforehand, the equilibration ratio between both proteins and the NO2 was determined by EMSA gel electrophoresis.

In Figure S10, the EMSA gel shows proteins alone and mixed with NO2 in different proportions. The 2:2:1 proportion was the most conducive to ternary complex (FOXO4-TEAD1-NO2) formation. By contrast, the 3:1 ratio promoted the formation of a binary protein-DNA complex, regardless of protein. At this ratio, complexes were premixed with 10mM Gln, preincubated, and oxidized in an FPOP system (10mM H_2O_2). The product mixture was injected into a column for LC-MS analysis before manual quantification and statistical analysis of the data. Figures 11, 12, and 13 show the extent of NO2 oxidative damage with and without (grey histograms) TEAD1-NO2, FOXO4-NO2 and TEAD1-FOXO4-NO2 complexes, respectively. Dashed histograms in Figures S11, S12, and S13 represent oligonucleotide fragments identified as isobaric, meaning that they were observed at the same mass, but not chromatographically resolved. For instance, fragments [For-15mer]-3'-**OH**/**P** (missing -GTG-from 3'-termini) were observed at the same mass as X→P/OH-5'[For-15mer] (missing 5'OH-TGG- from 5'-end), so the extent of their damage was mutually biased, which explains inconsistencies in some data. For this reason, some fragments bear higher standard deviations or reverse trends.

Then, Figure S14 displays summarized fragments from all experiments plotted above and below the sequence for both forward and reverse NO2 strands. Because Figure S14 is difficult to read, we included heatmaps of all three complex variants in Figure 6A, displaying differences in variants as "*Extent of damage of NO2 fragment*$_{in\ solution}$" and "*Extent of damage of NO2 fragment*$_{in\ complex}$". Non-significant changes are denoted as "×", and "n.o." refers to not-observed DNA fragments generated using a custom python-based script. The results show protection/deprotection effects, similar for H/DX data interpretationexperiments[61].

The heatmaps in Figure 6 highlight the strong protection of positions around the binding motifs and a much smaller protection/deprotection and the end of both strands, as expected. For instance, when TEAD1 interacts with NO2, the highest differences occur around the binding motif and adjacent bases[62], expressing major groove protection, with only mild protection around the FOXO4 binding sequence. However, the opposite (deprotecting) effect was observed in forward strands of the TEAD1 fragments [For-5mer]-**3'OH**, [For-6mer]-**3'OH**, and [For-7mer]-**3'OH**, displaying higher damage to the binding motif when TEAD1 is in the sample (compare grey and pink histograms in Figure S11). We suspect that the binding helix interaction with the major groove, inducing either 2'-deoxyribose conformational changes or duplex melting. As a result, hydrogens become more susceptible to hydroxyl radicals. The same effect of higher cleavage was noted in the reverse strand, in the direction of X→5'OH cleaving, yielding more 5'**OH**-[Rev-9,8,7,6,5,4mer] fragments when only NO2 was in solution. Surprisingly, the unfolding effect of the -CATTTGT- binding motif was even stronger when FOXO4 interacted with DNA. Nevertheless, FOXO4 did not induce extensive melting of the TEAD1 binding motif, as shown in [For-5,6,7mer]-3'**OH** DNA fragments (Forward 5'ATTCC3' motif). Thus, this suggests that TEAD1 interacts more with NO2 and decreases binding ability of FOXO4 with its binding motif on NO2 oligonucleotide.

.

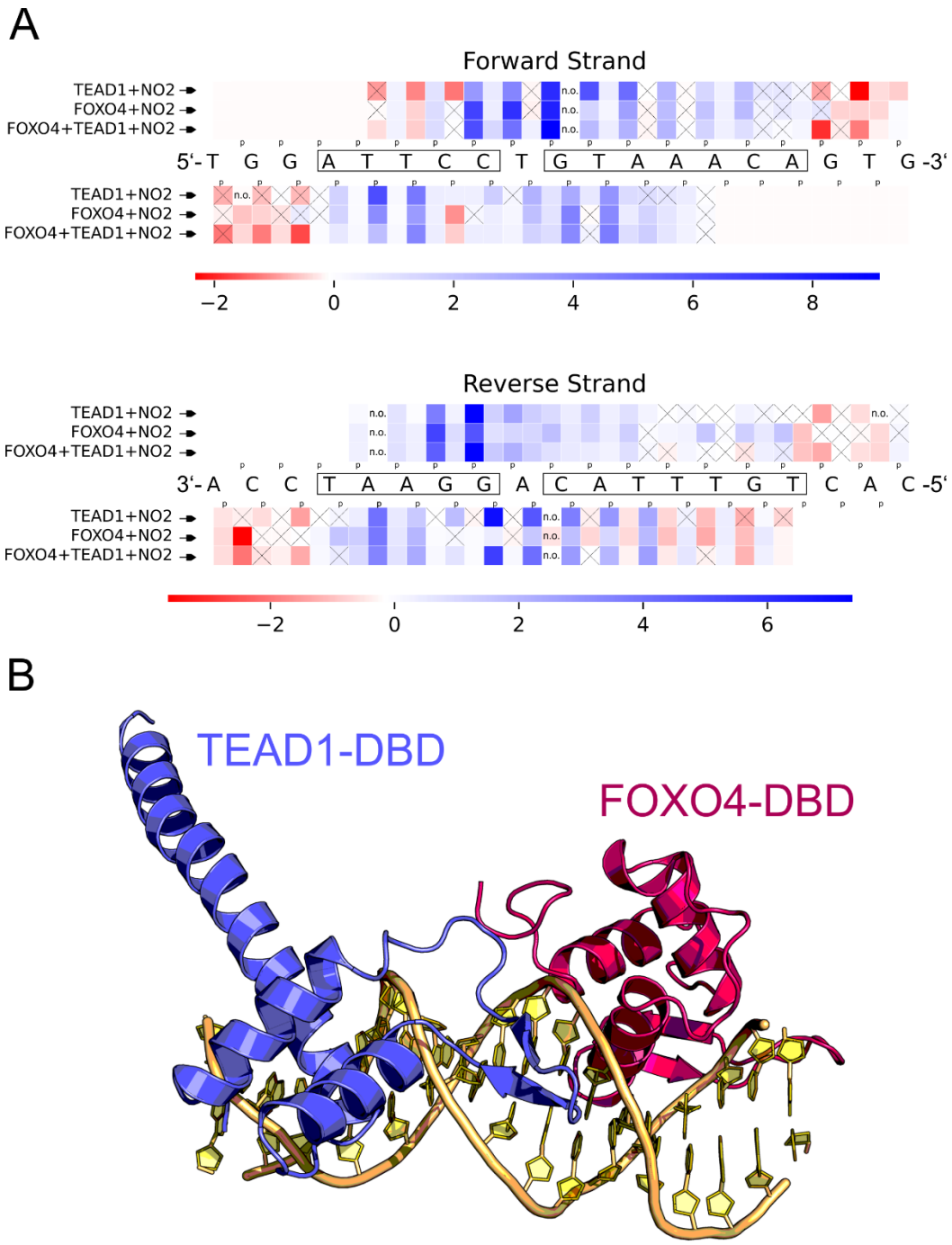


Figure 6. A. Heatmaps generated as a function of the protection/deprotection effect from Figures S11, S12 and S13 showing both strands and both generated fragments. Non-significant changes are denoted as "×", and "n.o." refers to not-observed DNA fragments. Binding motifs of both TEAD1 and FOXO4 are denoted in the black box. B. In silico model generated by computationally connecting two crystal structures, TEAD1 with DNA (PDB entry: 5GZB) and FOXO-DAF16 (PDB entry: 3L2C). TEAD1 is colored by blue, FOXO4 is colored by red.

Binding of both TFs provided information regarding the interactions of both proteins with NO_2. Mutually, binding resulted in a weaker melting effect, predominantly in FOXO4 proteins, suggesting that TEAD1 decreases FOXO4 interactions with NO_2. However, their mutual interactions and close proximity (Figure S6B) induced changes in the solvent accessibility of hydrogens at almost all positions, mostly generating a small deprotecting effect, but not as extensive and comparable only to NO_2 in solution. These results are in line with the competitive effect of both TFs caused by the spatial proximity of both binding motifs in a single oligonucleotide. We also expect a more dynamic nature of interaction in genome during transcription, where full-length TFs interacts with an unfolded chromatin. Nevertheless, FPOP coupled to LC facilitates the monitoring of even minor structural changes in DNA upon ligand binding.

Conclusion

By FPOP coupled to high-resolution MS, we can map protein surfaces and interactions with their cognate partners. For the first time, we have shown in this study that FPOP can be used to induce dsDNA fragmentation (dsIRE) in solution and with its cognate partner, FOXO4-DBD coupled with LC-MS analysis. LC analysis of the resulting fragments combined with FT-ICR MS/MS detection provides data on mass with high accuracy and resolution, enabling us to separate and analyze (i) phosphate- and hydroxyl-terminated fragments, (ii) phosphate and phosphoglycerate products, (iii) alcohol and aldehyde products, and (iv) all their singly oxidized forms, which are not distinguishable by gel electrophoresis [6]. Thus, complimentary fragments from forward and reverse strands are detected in a single experiment without prior ^{32}P labelling, yielding high-resolution information about the solvent accessibility of ribose hydrogens. Moreover, quantifying and comparing the extent of oxidative damage to fragments with and without FOXO4 shows the expected 'footprinting' effect around the binding motif, so this technique is suitable for mapping DNA damage and interactions.

FPOP can also be used to study ssDNA in solution and in B-form. By FPOP-induced damage, we can map higher-order structures of nucleic acids, including RNA. Further research may confirm whether this chemistry is also applicable to modified oligonucleotides, including therapeutic agents[55,58].

When tested on a more complex system, consisting of the TFs FOXO4-DBD and TEAD1-DBD and the oligonucleotide NO_2 with both binding motifs, our FPOP platform unveils not only the 'footprinted' effect of the two TFs around their binding motifs but also mutual corruption of the complex due to positional proximity of the binding motifs. Hydroxyl-terminated fragments exhibit higher damage in the binding motif when interacting with TFs, suggesting conformational changes in sugar moieties upon interaction with amino acids. These findings demonstrate that our approach detects even minor structural changes in DNA whether in a complex with individual TFs or in a ternary complex. In future studies, we will analyze more complex samples and design software for automated analysis of high-resolution datasets towards assessing damage to single ribonucleic acids in solution.

ASSOCIATED CONTENT

Supporting Information

Supporting information contains additional Figures: S1 Extracted ion chromatogram of 9-mer phosphate and phosphoglycerate, S2: 9mer and 8mer from the reaction performed in $H_2{}^{18}O$, S3: EMSA of FOXO4-IRE complex, S4: XICs of forward strand fragments before FPOP, S5: XICs of reverse strand fragments before FPOP, S6: XICs of forward strand fragments after FPOP, S7: XICs of reverse strand fragments after FPOP, S8: MS/MS spectra of 5'OH-[For-GACTAT]-3'OH, S9 XICs and MS spectra of For-15mer a For-8mer, S10 EMSA of FOXO4-TEAD1-NO_2, S11: extent of damage of the TEAD1-NO_2 complex, S12: extent of damage of the FOXO4-NO_2 complex, S13: extent of damage of the FOXO4-TEAD1-NO_2 complex, S14: All-in-one plot for the extent of damage of FOXO4-TEAD1-NO_2. Supporting Information is available free of charge on the ACS Publications website.

AUTHOR INFORMATION

Corresponding Author

* **Petr Novák** – Institute of Microbiology, The Czech Academy of Sciences, 14220 Prague, Czech Republic; orcid.org/0000-0001-8688-529X; E-mail: pnovak@biomed.cas.cz. Fax: +420 241 062 156. Tel.: +420 325 873 610 (P.N.).

Authors

Marek Polák – Institute of Microbiology of the Czech Academy of Sciences, Prague, Czech Republic, Faculty of Science, Charles University, 12800 Prague, Czech Republic; orcid.org/0000-0002-8929-5479

Vítězslav Brinsa – Institute of Microbiology of the Czech Academy of Sciences, Prague, Czech Republic, Faculty of Science, Charles University, 12800 Prague, Czech Republic; orcid.org/0000-0002-8929-5479

Jiří Černý – Institute of Biotechnology of the Czech Academy of Sciences, Prague, Czech Republic, https://orcid.org/0000-0002-1969-9304

Michael Volný – Institute of Microbiology, The Czech Academy of Sciences, 14220 Prague, Czech Republic

Alan Kádek – Institute of Microbiology, The Czech Academy of Sciences, 14220 Prague, Czech Republic

Author Contributions

The manuscript was written through the contributions of all authors. All authors have approved the final version of the manuscript.

ACKNOWLEDGMENT

We thank Carlos V. Melo for editing the manuscript. This work was mainly funded by NPO-NEURO-EXCELLES (LX22NPO5107), the Czech Science Foundation (22-27695S), the European Commission H2020 (EU_FT-ICR_MS grant agreement ID: 731077 and EPIC-XS - grant agreement ID: 823839). Additional institutional and facility support from the Academy of Sciences of the Czech Republic (RVO: 61388971), Grant Agency of Charles University (359521), the Ministry of Education of the Czech Republic (Structural mass spectrometry CF - LM2018127 CIISB), ELIXIR-CZ (LM2023055) and the

European Regional Development Funds (CZ.1.05/1.1.00/02.0109 BIOCEV) are gratefully acknowledged

REFERENCES

(1) Lambert, S. A.; Jolma, A.; Campitelli, L. F.; Das, P. K.; Yin, Y.; Albu, M.; Chen, X.; Taipale, J.; Hughes, T. R.; Weirauch, M. T. The Human Transcription Factors. *Cell* **2018**, *172* (4), 650–665. https://doi.org/10.1016/j.cell.2018.01.029.

(2) Pandey, P.; Hasnain, S.; Ahmad, S. Protein-DNA Interactions. In *Encyclopedia of Bioinformatics and Computational Biology*; Elsevier, 2019; pp 142–154. https://doi.org/10.1016/B978-0-12-809633-8.20217-3.

(3) Jain, S. S.; Tullius, T. D. Footprinting Protein–DNA Complexes Using the Hydroxyl Radical. *Nat. Protoc.* **2008**, *3* (6), 1092–1100. https://doi.org/10.1038/nprot.2008.72.

(4) Galas, D. J.; Schmitz, A. DNAase Footprinting a Simple Method for the Detection of Protein-DNA Binding Specificity. *Nucleic Acids Res.* **1978**, *5* (9), 3157–3170. https://doi.org/10.1093/nar/5.9.3157.

(5) Koppenol, W. H. The Haber-Weiss Cycle - 70 Years Later. *Redox Report*. Taylor & Francis 2001, pp 229–234. https://doi.org/10.1179/135100001101536373.

(6) Tullius, T. D.; Dombroski, B. A. Hydroxyl Radical "Footprinting": High-Resolution Information about DNA-Protein Contacts and Application to λ Repressor and Cro Protein. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83* (15), 5469–5473. https://doi.org/10.1073/pnas.83.15.5469.

(7) Tullius, T. D. Probing DNA Structure with Hydroxyl Radicals. *Curr. Protoc. Nucleic Acid Chem.* **2001**, *7* (1), 6.7.1-6.7.8. https://doi.org/10.1002/0471142700.nc0607s07.

(8) Pogozelski, W. K.; Tullius, T. D. Oxidative Strand Scission of Nucleic Acids: Routes Initiated by Hydrogen Abstraction from the Sugar Moiety. *Chem. Rev.* **1998**, *98* (3), 1089–1107. https://doi.org/10.1021/cr960437i.

(9) Balasubramanian, B.; Pogozelski, W. K.; Tullius, T. D. DNA Strand Breaking by the Hydroxyl Radical Is Governed by the Accessible Surface Areas of the Hydrogen Atoms of the DNA Backbone. *Proc. Natl. Acad. Sci.* **1998**, *95* (17), 9738–9743. https://doi.org/10.1073/pnas.95.17.9738.

(10) Sclavi, B.; Sullivan, M.; Chance, M. R.; Brenowitz, M.; Woodson, S. A. RNA Folding at Millisecond Intervals by Synchrotron Hydroxyl Radical Footprinting. *Science (80-.).* **1998**, *279* (5358), 1940–1943. https://doi.org/10.1126/science.279.5358.1940.

(11) Sclavi, B.; Woodson, S.; Sullivan, M.; Chance, M. R.; Brenowitz, M. Time-Resolved Synchrotron X-Ray "Footprinting", a New Approach to the Study of Nucleic Acid Structure and Function: Application to Protein-DNA Interactions and RNA Folding. *J. Mol. Biol.* **1997**, *266* (1), 144–159. https://doi.org/10.1006/JMBI.1996.0775.

(12) Adilakshmi, T.; Lease, R. A.; Woodson, S. A. Hydroxyl Radical Footprinting in Vivo: Mapping Macromolecular Structures with Synchrotron Radiation. *Nucleic Acids Res.* **2006**, *34* (8), e64. https://doi.org/10.1093/nar/gkl291.

(13) Ottinger, L. M.; Tullius, T. D. High-Resolution in Vivo Footprinting of a Protein - DNA Complex Using γ-Radiation. *Journal of the American Chemical Society*. June 21, 2000, pp 5901–5902. https://doi.org/10.1021/ja000285f.

(14) Costa, M.; Monachello, D. Probing RNA Folding by Hydroxyl Radical Footprinting. In *Methods in Molecular Biology*; Humana Press Inc., 2014; Vol. 1086, pp 119–142. https://doi.org/10.1007/978-1-62703-667-2_7.

(15) Tullius, T. D.; Greenbaum, J. A. Mapping Nucleic Acid Structure by Hydroxyl Radical Cleavage. *Curr. Opin. Chem. Biol.* **2005**, *9* (2), 127–134. https://doi.org/10.1016/j.cbpa.2005.02.009.

(16) Ingle, S.; Azad, R. N.; Jain, S. S.; Tullius, T. D. Chemical Probing of RNA with the Hydroxyl Radical at Single-Atom Resolution. *Nucleic Acids Res.* **2014**, *42* (20), 12758–12767. https://doi.org/10.1093/NAR/GKU934.

(17) Sharp, J. S.; Becker, J. M.; Hettich, R. L. Protein Surface Mapping by Chemical Oxidation: Structural Analysis by Mass Spectrometry. *Anal. Biochem.* **2003**, *313* (2), 216–225. https://doi.org/10.1016/S0003-2697(02)00612-7.

(18) Sharp, J. S.; Becker, J. M.; Hettich, R. L. Analysis of Protein Solvent Accessible Surfaces by Photochemical Oxidation and Mass Spectrometry. *Anal. Chem.* **2004**, *76* (3), 672–683. https://doi.org/10.1021/ac0302004.

(19) Chance, M. R. Unfolding of Apomyoglobin Examined by Synchrotron Footprinting. *Biochem. Biophys. Res. Commun.* **2001**, *287* (3), 614–621. https://doi.org/10.1006/BBRC.2001.5628.

(20) Guan, J.-Q.; Vorobiev, S.; Almo, S. C.; Chance, M. R. Mapping the G-Actin Binding Surface of Cofilin Using Synchrotron Protein Footprinting. *Biochemistry* **2002**, *41* (18), 5765–5775. https://doi.org/10.1021/bi0121104.

(21) Kiselar, J. G.; Maleknia, S. D.; Sullivan, M.; Downard, K. M.; Chance, M. R. Hydroxyl Radical Probe of Protein Surfaces Using Synchrotron X-Ray Radiolysis and Mass Spectrometry. *Int. J. Radiat. Biol.* **2002**, *78* (2), 101–114. https://doi.org/10.1080/09553000110094805.

(22) Xu, G.; Chance, M. R. Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chemical Reviews*. American Chemical Society August 2007, pp 3514–3543. https://doi.org/10.1021/cr0682047.

(23) Xu, G.; Chance, M. R. Radiolytic Modification of Sulfur-Containing Amino Acid Residues in Model Peptides: Fundamental Studies for Protein Footprinting. *Anal. Chem.* **2005**, *77* (8), 2437–2449. https://doi.org/10.1021/ac0484629.

(24) Xu, G.; Chance, M. R. Radiolytic Modification of Acidic Amino Acid Residues in Peptides: Probes for Examining Protein−Protein Interactions. *Anal. Chem.* **2004**, *76* (5), 1213–1221. https://doi.org/10.1021/ac035422g.

(25) Xu, G.; Chance, M. R. Radiolytic Modification and Reactivity of Amino Acid Residues Serving as Structural Probes for Protein Footprinting. *Anal. Chem.* **2005**, *77* (14), 4549–4555. https://doi.org/10.1021/ac050299+.

(26) Hambly, D. M.; Gross, M. L. Laser Flash Photolysis of Hydrogen Peroxide to Oxidize Protein Solvent-Accessible Residues on the Microsecond Timescale. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (12), 2057–2063. https://doi.org/10.1016/j.jasms.2005.09.008.

(27) Niu, B.; Gross, M. L. MS-Based Hydroxyl Radical Footprinting: Methodology and Application of Fast Photochemical Oxidation of Proteins (FPOP). In *Mass Spectrometry-Based Chemical Proteomics*; Wiley, 2019; pp 363–416. https://doi.org/10.1002/9781118970195.ch15.

(28) Zhang, B.; Cheng, M.; Rempel, D.; Gross, M. L. Implementing Fast Photochemical Oxidation of Proteins (FPOP) as a Footprinting Approach to Solve Diverse Problems in Structural Biology. *Methods*. Academic Press Inc. July 15, 2018, pp 94–103. https://doi.org/10.1016/j.ymeth.2018.05.016.

(29) Li, K. S.; Shi, L.; Gross, M. L. Mass Spectrometry-Based Fast Photochemical Oxidation of Proteins (FPOP) for Higher Order Structure Characterization. *Acc. Chem. Res.* **2018**, *51* (3), 736–744. https://doi.org/10.1021/acs.accounts.7b00593.

(30) Polák, M.; Yassaghi, G.; Kavan, D.; Filandr, F.; Fiala, J.; Kukačka, Z.; Halada, P.; Loginov, D. S.; Novák, P. Utilization of Fast Photochemical Oxidation of Proteins and Both Bottom-up and Top-down Mass Spectrometry for Structural Characterization of a Transcription Factor–DsDNA Complex. *Anal. Chem.* **2022**, *94* (7), 3203–3210. https://doi.org/10.1021/acs.analchem.1c04746.

(31) Tomášková, N.; Novák, P.; Kožár, T.; Petrenčáková, M.; Jancura, D.; Yassaghi, G.; Man, P.; Sedlák, E. Early Modification of Cytochrome c by Hydrogen Peroxide Triggers Its Fast Degradation. *Int. J. Biol. Macromol.* **2021**, *174*, 413–423. https://doi.org/10.1016/j.ijbiomac.2021.01.189.

(32) Petrenčáková, M.; Filandr, F.; Hovan, A.; Yassaghi, G.; Man, P.; Kožár, T.; Schwer, M. S.; Jancura, D.; Plückthun, A.; Novák, P.; Miškovský, P.; Bánó, G.; Sedlák, E. Photoinduced Damage of AsLOV2 Domain Is Accompanied by Increased Singlet Oxygen Production Due to Flavin Dissociation. *Sci. Rep.* **2020**, *10* (1), 1–15. https://doi.org/10.1038/s41598-020-60861-2.

(33) Yassaghi, G.; Kukačka, Z.; Fiala, J.; Kavan, D.; Halada, P.; Volný, M.; Novák, P. Top-Down Detection of Oxidative Protein Footprinting by Collision-Induced Dissociation, Electron-Transfer Dissociation, and Electron-Capture Dissociation. *Anal. Chem.* **2022**, *94* (28), 9993–10002. https://doi.org/10.1021/ACS.ANALCHEM.1C05476.

(34) Loginov, D. S.; Fiala, J.; Brechlin, P.; Kruppa, G.; Novak, P. Hydroxyl Radical Footprinting Analysis of a Human Haptoglobin-Hemoglobin Complex. *Biochim. Biophys. Acta - Proteins Proteomics* **2022**, *1870* (2), 140735. https://doi.org/10.1016/j.bbapap.2021.140735.

(35) Cornwell, O.; Radford, S. E.; Ashcroft, A. E.; Ault, J. R. Comparing Hydrogen Deuterium Exchange and Fast Photochemical Oxidation of Proteins: A Structural Characterisation of Wild-Type and ΔN6 B2-Microglobulin. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (12), 2413–2426. https://doi.org/10.1007/s13361-018-2067-y.

(36) Cornwell, O.; Bond, N. J.; Radford, S. E.; Ashcroft, A. E. Long-Range Conformational Changes in Monoclonal Antibodies Revealed Using FPOP-LC-MS/MS. *Anal. Chem.* **2019**, *91* (23), 15163–15170. https://doi.org/10.1021/acs.analchem.9b03958.

(37) Charvátová, O.; Foley, B. L.; Bern, M. W.; Sharp, J. S.; Orlando, R.; Woods, R. J. Quantifying Protein Interface Footprinting by Hydroxyl Radical Oxidation and Molecular Dynamics Simulation: Application to Galectin-1. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (11), 1692–1705. https://doi.org/10.1016/j.jasms.2008.07.013.

(38) Pan, Y.; Stocks, B. B.; Brown, L.; Konermann, L.

Structural Characterization of an Integral Membrane Protein in Its Natural Lipid Environment by Oxidative Methionine Labeling and Mass Spectrometry. *Anal. Chem.* **2009**, *81* (1), 28–35. https://doi.org/10.1021/ac8020449.

(39) Watkinson, T. G.; Calabrese, A. N.; Ault, J. R.; Radford, S. E.; Ashcroft, A. E. FPOP-LC-MS/MS Suggests Differences in Interaction Sites of Amphipols and Detergents with Outer Membrane Proteins. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (1), 50–55. https://doi.org/10.1007/s13361-016-1421-1.

(40) Lu, Y.; Zhang, H.; Niedzwiedzki, D. M.; Jiang, J.; Blankenship, R. E.; Gross, M. L. Fast Photochemical Oxidation of Proteins Maps the Topology of Intrinsic Membrane Proteins: Light-Harvesting Complex 2 in a Nanodisc. *Anal. Chem.* **2016**, *88* (17), 8827–8834. https://doi.org/10.1021/acs.analchem.6b01945.

(41) Gupta, S.; Bavro, V. N.; D'Mello, R.; Tucker, S. J.; Vénien-Bryan, C.; Chance, M. R. Conformational Changes During the Gating of a Potassium Channel Revealed by Structural Mass Spectrometry. *Structure* **2010**, *18* (7), 839–846. https://doi.org/10.1016/j.str.2010.04.012.

(42) Liu, X. R.; Zhang, M. M.; Gross, M. L. Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications. *Chemical Reviews*. American Chemical Society May 27, 2020, pp 4355–4454. https://doi.org/10.1021/acs.chemrev.9b00815.

(43) Largy, E.; König, A.; Ghosh, A.; Ghosh, D.; Benabou, S.; Rosu, F.; Gabelica, V. Mass Spectrometry of Nucleic Acid Noncovalent Complexes. *Chemical Reviews*. American Chemical Society April 27, 2022, pp 7720–7839. https://doi.org/10.1021/acs.chemrev.1c00386.

(44) von Sonntag, C. *Free-Radical-Induced DNA Damage and Its Repair*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2006. https://doi.org/10.1007/3-540-30592-0.

(45) Di Mascio, P.; Martinez, G. R.; Miyamoto, S.; Ronsein, G. E.; Medeiros, M. H. G.; Cadet, J. Singlet Molecular Oxygen Reactions with Nucleic Acids, Lipids, and Proteins. *Chemical Reviews*. American Chemical Society February 13, 2019, pp 2043–2086. https://doi.org/10.1021/acs.chemrev.8b00554.

(46) Polák, M.; Černý, J.; Novák, P. Isotopic Depletion Increases the Spatial Resolution of FPOP Top-Down Mass Spectrometry Analysis. *Anal. Chem.* **2024**, *96* (4), 1478–1487. https://doi.org/10.1021/acs.analchem.3c03759.

(47) Zubarev, R. A.; Demirev, P. A. Isotope Depletion of Large Biomolecules: Implications for Molecular Mass Measurements. *J. Am. Soc. Mass Spectrom.* **1998**, *9* (2), 149–156. https://doi.org/10.1016/S1044-0305(97)00232-8.

(48) Kukačka, Z.; Rosůlek, M.; Jelínek, J.; Slavata, L.; Kavan, D.; Novák, P. LinX: A Software Tool for Uncommon Cross-Linking Chemistry. *J. Proteome Res.* **2021**, *20* (4), 2021–2027. https://doi.org/10.1021/acs.jproteome.0c00858.

(49) Azad, R. N.; Zafiropoulos, D.; Ober, D.; Jiang, Y.; Chiu, T. P.; Sagendorf, J. M.; Rohs, R.; Tullius, T. D. Experimental Maps of DNA Structure at Nucleotide Resolution Distinguish Intrinsic from Protein-Induced DNA Deformations. *Nucleic Acids Res.* **2018**, *46* (5), 2637–2647. https://doi.org/10.1093/nar/gky033.

(50) Shaytan, A. K.; Xiao, H.; Armeev, G. A.; Gaykalova, D. A.; Komarova, G. A.; Wu, C.; Studitsky, V. M.; Landsman, D.; Panchenko, A. R. Structural Interpretation of DNA–Protein Hydroxyl-Radical Footprinting Experiments with High Resolution Using HYDROID. *Nat. Protoc.* **2018**, *13* (11), 2535–2556. https://doi.org/10.1038/s41596-018-0048-z.

(51) Henner, W. D.; Rodriguez, L. O.; Hecht, S. M.; Haseltine, W. A. Gamma Ray Induced Deoxyribonucleic Acid Strand Breaks. 3' Glycolate Termini. *J. Biol. Chem.* **1983**, *258* (2), 711–713. https://doi.org/10.1016/s0021-9258(18)33104-1.

(52) Henner, W. D.; Grunberg, S. M.; Haseltine, W. A. Sites and Structure of Gamma Radiation-Induced DNA Strand Breaks. *J. Biol. Chem.* **1982**, *257* (19), 11750–11754. https://doi.org/10.1016/S0021-9258(18)33827-4.

(53) Kappen, L. S.; Goldberg, I. H. Deoxyribonucleic Acid Damage by Neocarzinostatin Chromophore: Strand Breaks Generated by Selective Oxidation of C-5′ of Deoxyribose. *Biochemistry* **1983**, *22* (21), 4872–4878. https://doi.org/10.1021/BI00290A002/ASSET/BI00290A002.FP.PNG_V03.

(54) Gao, Y.; McLuckey, S. A. Fragmentation Reactions of Nucleic Acid Ions in the Gas Phase. In *Physical Chemistry in Action*; 2014; pp 131–182. https://doi.org/10.1007/978-3-642-54842-0_6.

(55) Santos, I. C.; Lanzillotti, M.; Shilov, I.; Basanta-Sanchez, M.; Roushan, A.; Lawler, R.; Tang, W.; Bern, M.; Brodbelt, J. S. Ultraviolet Photodissociation and Activated Electron Photodetachment Mass Spectrometry for Top-Down Sequencing of Modified Oligoribonucleotides. *J. Am. Soc. Mass Spectrom.* **2022**, *33* (3), 510–520.

https://doi.org/10.1021/JASMS.1C00340/ASSET/IMAGES/LARGE/JS1C00340_0009.JPEG.

(56) Agten, A.; Prostko, P.; Geubbelmans, M.; Liu, Y.; De Vijlder, T.; Valkenborg, D. A Compositional Model to Predict the Aggregated Isotope Distribution for Average Dna and Rna Oligonucleotides. *Metabolites* **2021**, *11* (6), 400. https://doi.org/10.3390/METABO11060400/S1.

(57) Kenderdine, T.; McIntyre, W.; Yassaghi, G.; Rollo, D.; Bunkowski, A.; Goerlach, L.; Suckau, D.; Tremintin, G.; Greig, M.; Bell, C.; Fabris, D. Integrating Internal Fragments in the Interpretation of Top-Down Sequencing Data of Larger Oligonucleotides. *J. Am. Soc. Mass Spectrom.* **2023**. https://doi.org/10.1021/jasms.3c00207.

(58) Moumné, L.; Marie, A.-C.; Crouvezier, N. Oligonucleotide Therapeutics: From Discovery and Development to Patentability. *Pharmaceutics* **2022**, *14* (2), 260. https://doi.org/10.3390/pharmaceutics14020260.

(59) Zhu, Y.; Zhu, L.; Wang, X.; Jin, H. RNA-Based Therapeutics: An Overview and Prospectus. *Cell Death Dis.* **2022**, *13* (7). https://doi.org/10.1038/S41419-022-05075-2.

(60) Filandrová, R.; Vališ, K.; Černý, J.; Chmelík, J.; Slavata, L.; Fiala, J.; Rosůlek, M.; Kavan, D.; Man, P.; Chum, T.; Cebecauer, M.; Fabris, D.; Novák, P. Motif Orientation Matters: Structural Characterization of TEAD1 Recognition of Genomic DNA. *Structure* **2021**, *29* (4), 345-356.e8. https://doi.org/10.1016/j.str.2020.11.018.

(61) Masson, G. R.; Burke, J. E.; Ahn, N. G.; Anand, G. S.; Borchers, C.; Brier, S.; Bou-Assaf, G. M.; Engen, J. R.; Englander, S. W.; Faber, J.; Garlish, R.; Griffin, P. R.; Gross, M. L.; Guttman, M.; Hamuro, Y.; Heck, A. J. R.; Houde, D.; Iacob, R. E.; Jørgensen, T. J. D.; Kaltashov, I. A.; Klinman, J. P.; Konermann, L.; Man, P.; Mayne, L.; Pascal, B. D.; Reichmann, D.; Skehel, M.; Snijder, J.; Strutzenberg, T. S.; Underbakke, E. S.; Wagner, C.; Wales, T. E.; Walters, B. T.; Weis, D. D.; Wilson, D. J.; Wintrode, P. L.; Zhang, Z.; Zheng, J.; Schriemer, D. C.; Rand, K. D. Recommendations for Performing, Interpreting and Reporting Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) Experiments. *Nat. Methods* **2019**, *16* (7), 595–602. https://doi.org/10.1038/s41592-019-0459-y.

(62) Shi, Z.; He, F.; Chen, M.; Hua, L.; Wang, W.; Jiao, S.; Zhou, Z. DNA-Binding Mechanism of the Hippo Pathway Transcription Factor TEAD4. *Oncogene* **2017**, *36* (30), 4362–4369. https://doi.org/10.1038/onc.2017.24.

Supporting information

Hydroxyl Radical Footprinting of Nucleic Acids Coupled to High-Resolution Mass Spectrometry Analysis

Marek Polák$,‡, Vítězslav Brinsa‡, Jiří Černý†, Daniel Kavan$,‡, Michael Volný$, Alan Kádek$, and Petr Novák*,$,‡

$Institute of Microbiology of the Czech Academy of Sciences, 14220, Prague, Czech Republic; ‡Department of Biochemistry, Faculty of Science, Charles University, 12843, Prague, Czech Republic,
†Laboratory of Structural Bioinformatics of Proteins, Institute of Biotechnology of the Czech Academy of Sciences, 14220, Prague, Czech Republic

*Correspondence to: pnovak@biomed.cas.cz

Table of contents

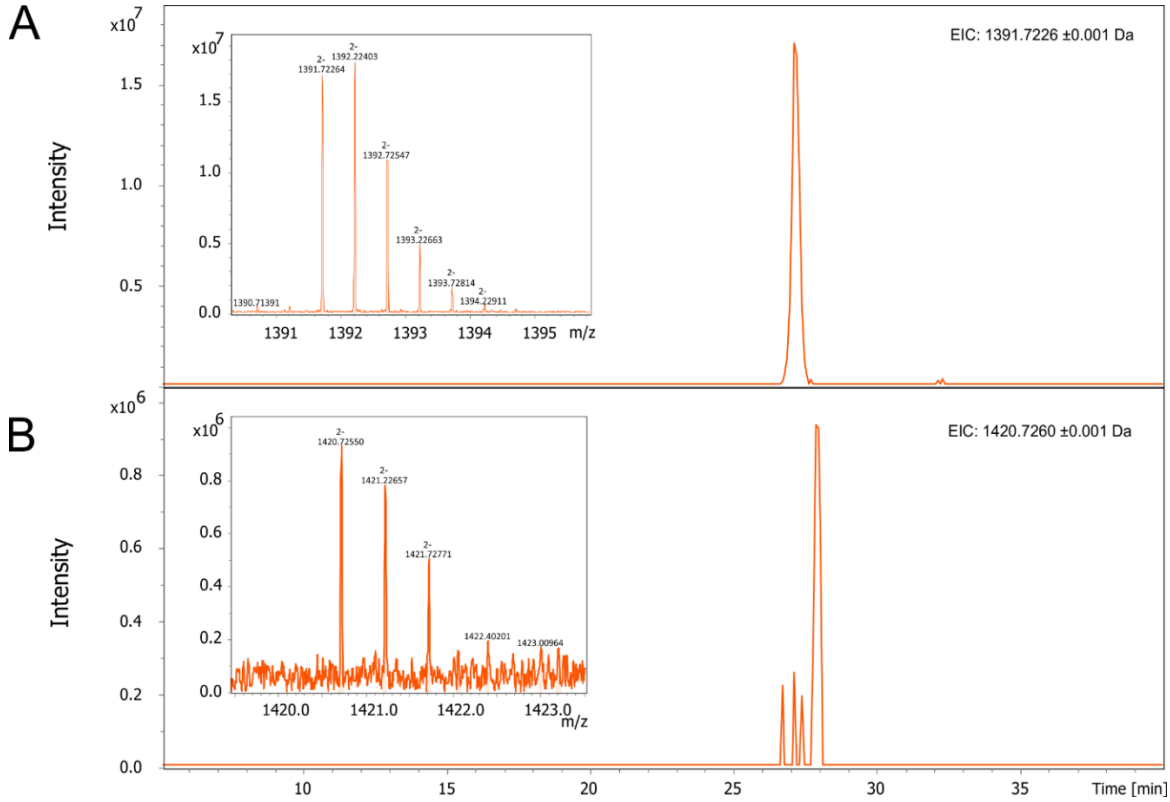


Figure S1. Extracted ion chromatogram of 5'OH-[9-mer]-3'P (A) and 5'OH-[9-mer]-3'PG (Phosphoglycerate, B) searched in a LC-MS trace with 1391.7226 ± 0.001 Da and 1420.7260 ± 0.001 Da mass, respectively. The inset represents the MS1 spectrum zoomed-in at the m/z axis where the doubly charged products were observed.

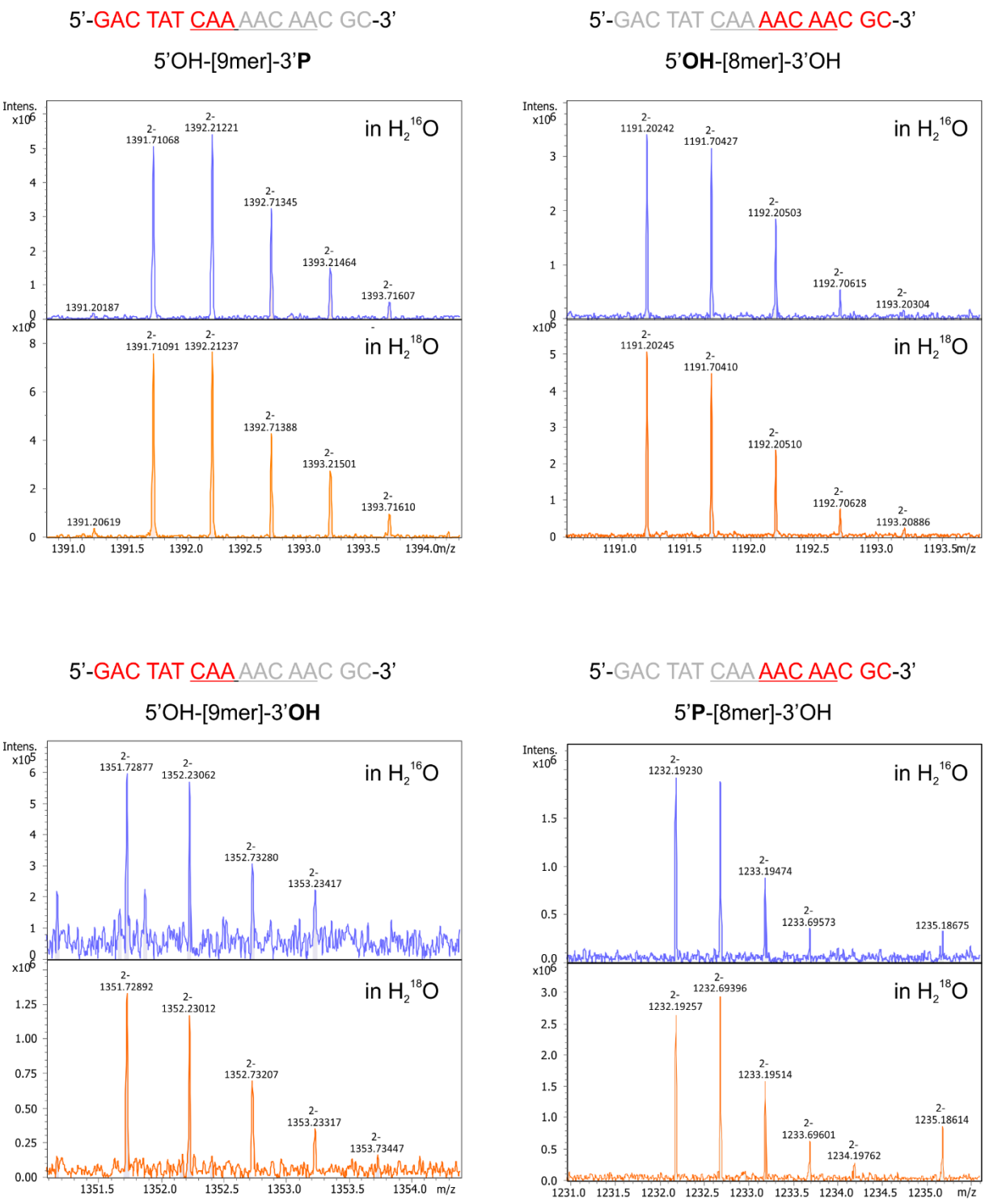


Figure S2. LC-MS spectra of forward strand fragments generated during FPOP. The whole sequence is denoted above each mass spectrum, and the sequence highlighted in red indicates the position of the fragment. Below, the sequence is denoted 9mer or 8 mer with the terminated end highlighted in bold, either phosphate or hydroxyl. Blue spectra represent reactions performed in $H_2{}^{16}O$, and orange spectra represent reactions performed in $H_2{}^{18}O$.

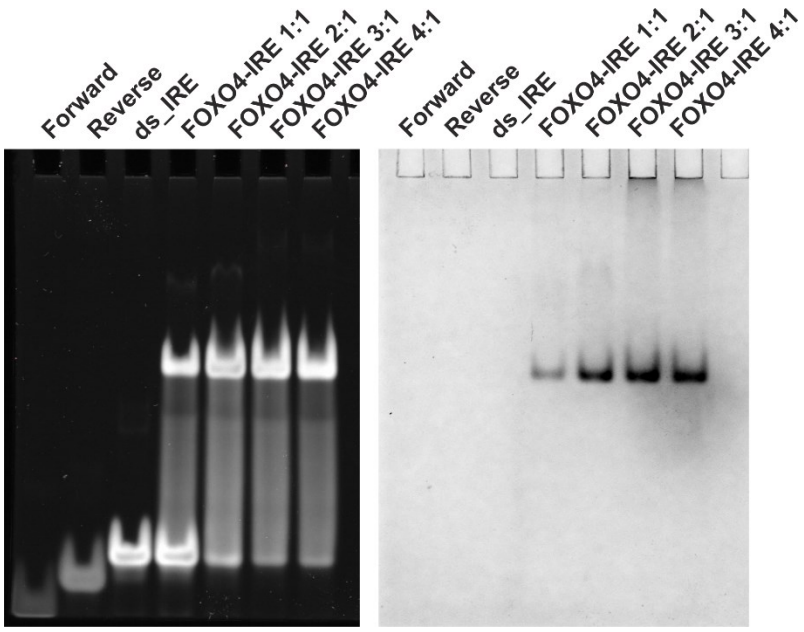


Figure S3. Electrophoretic mobility shift assay of the FOXO•dsIRE complex stained with GelRed® (left panel) for nucleic acid and with Coomassie Brilliant Blue for protein (right panel) detection. FOXO4 caused a mobility shift in the duplex in comparison with IRE alone.

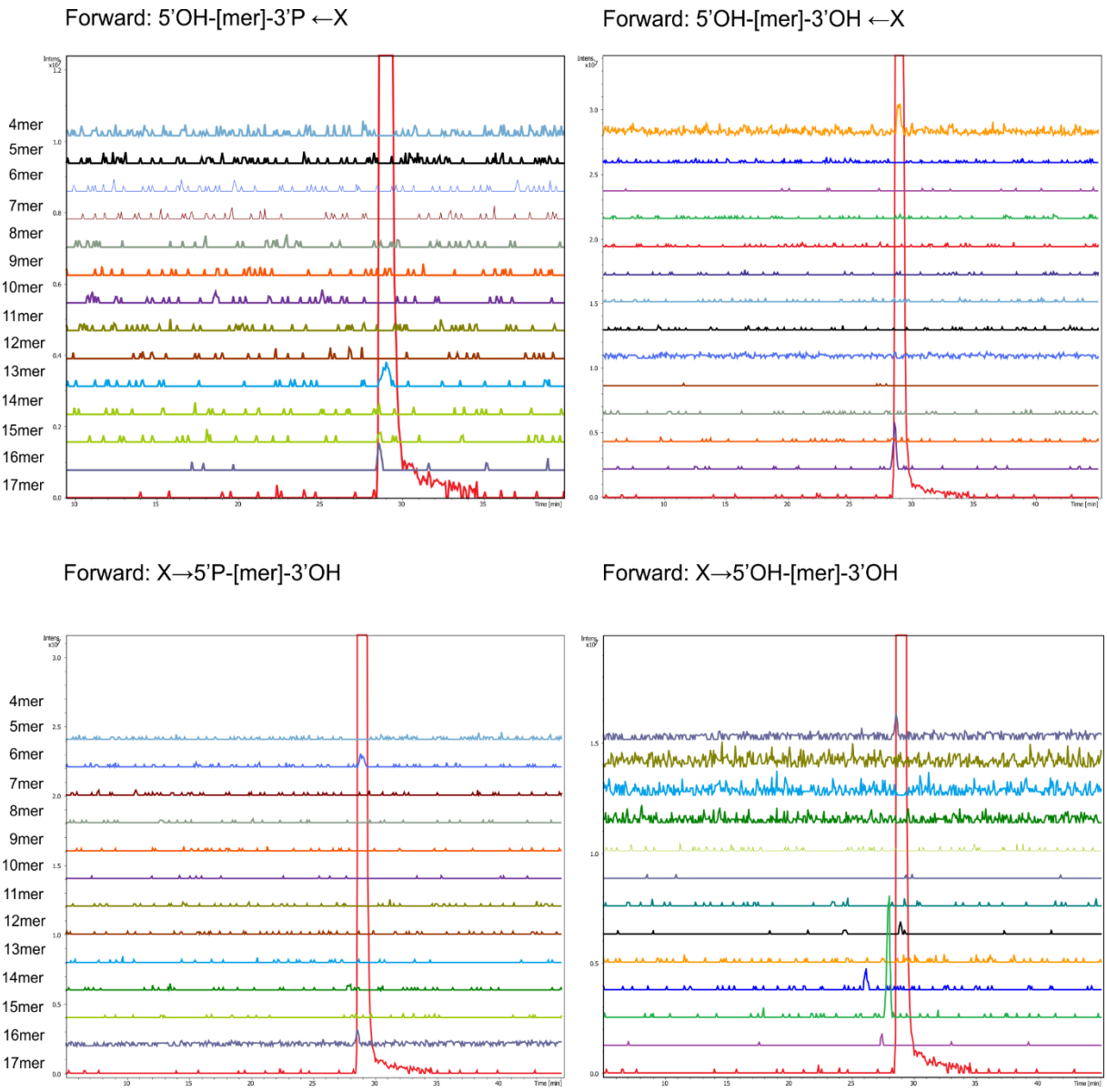


Figure S4. Extracted ion chromatograms of all potential fragments searched in a control stock solution of the forward strand. The strongest signal represents the forward strand. The X symbol represents the direction of cleavage.

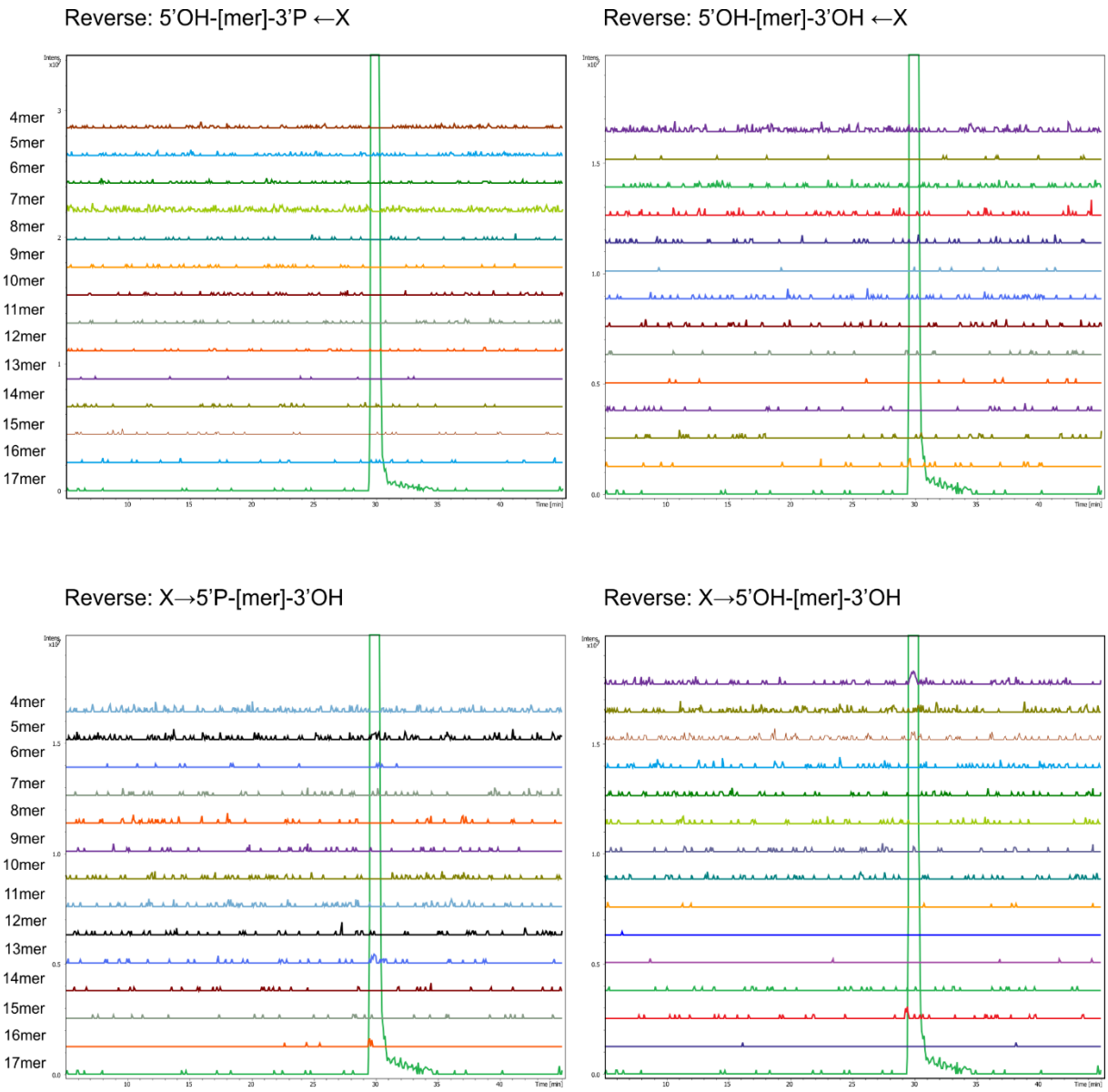


Figure S5. Extracted ion chromatograms of all potential fragments searched in a control stock solution of the reverse strand. The strongest signal represents the forward strand. The 'X' symbol represents the direction of cleavage.

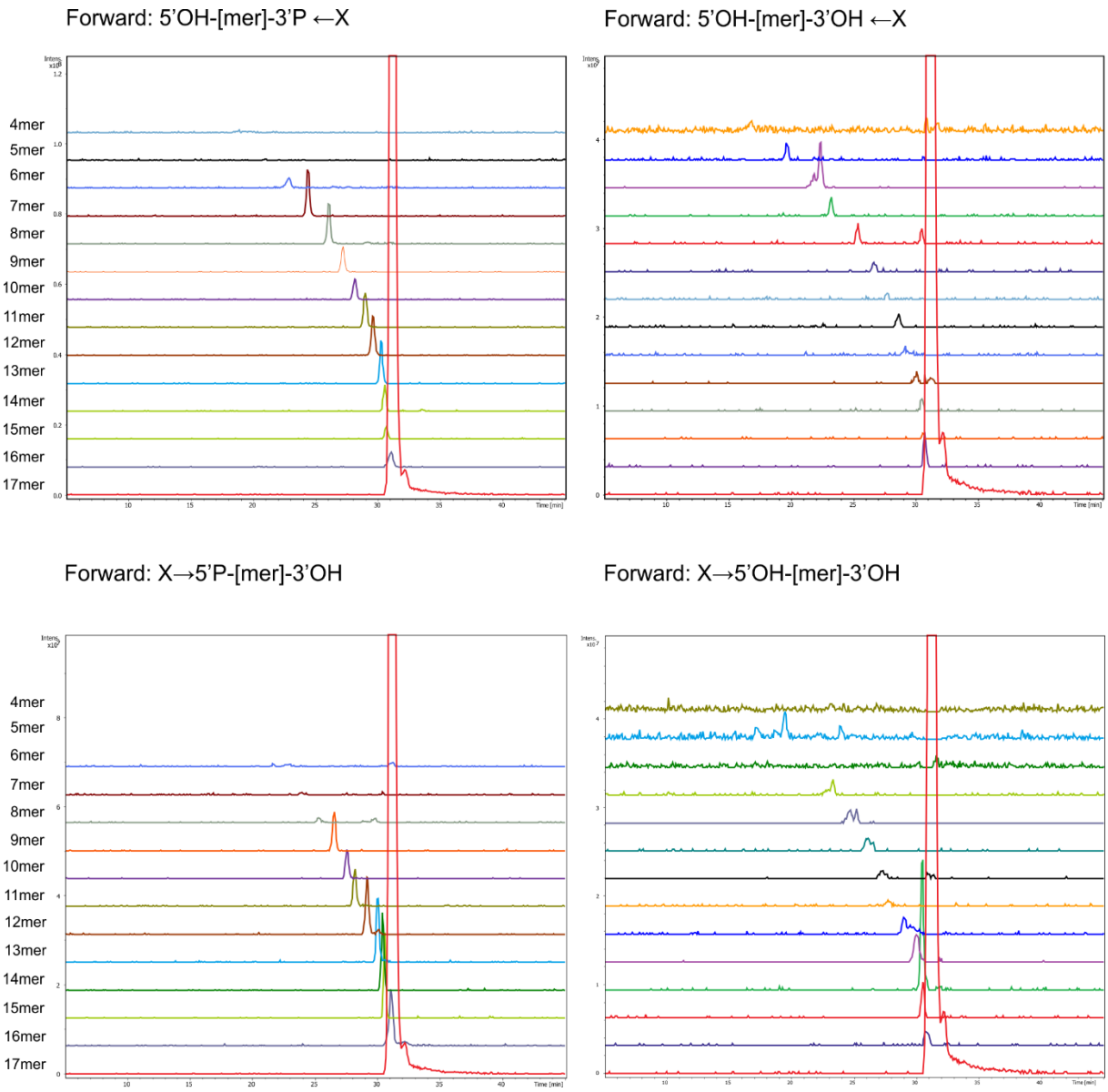


Figure S6. Extracted ion chromatograms of all potential fragments searched after FPOP with n-1 chromatographically separated fragments. The strongest signal represents the forward strand. The X symbol represents the direction of cleavage.

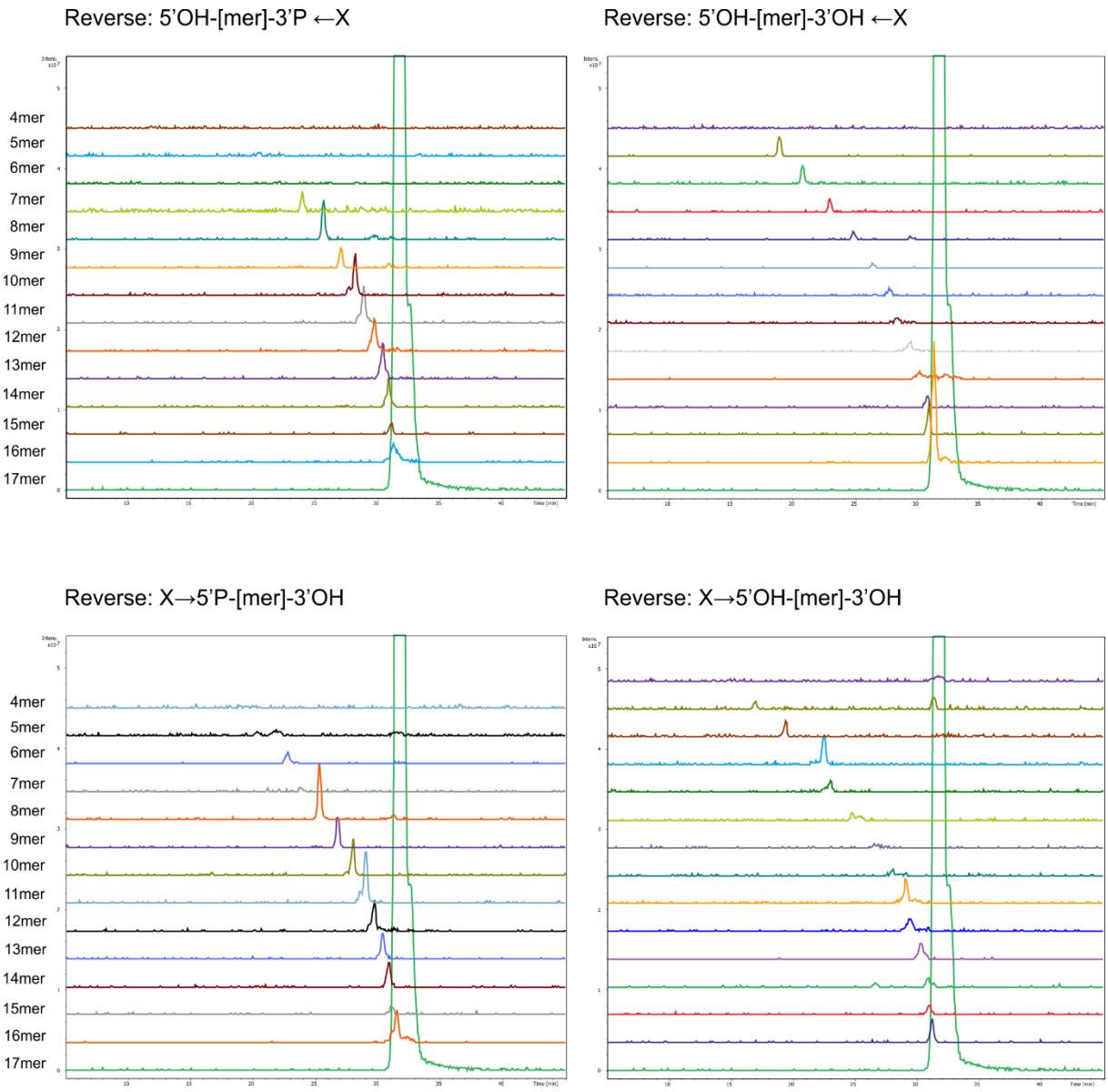


Figure S7. Extracted ion chromatograms of all potential fragments searched after FPOP with n-1 fragments are chromatographically separated. The strongest signal represents the reverse strand. The X symbol represents the direction of cleavage.

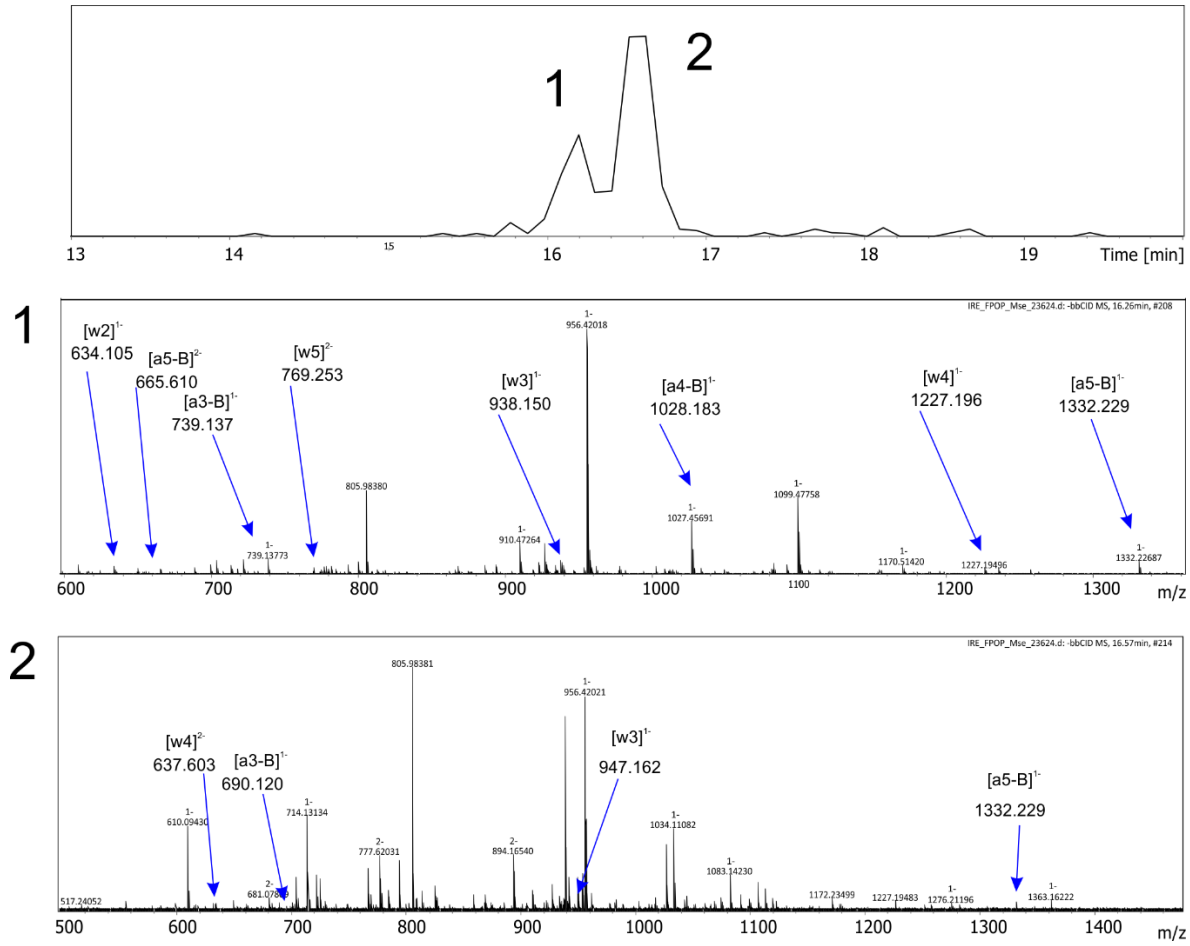


Figure S8. Extracted ion chromatogram of 894.161 Da searched mass, accurate to ±0.1 Da showing two chromatographically resolved peaks (top panel). Manual analysis of DIA MS/MS spectra showing that product 1 (5'OH-[For-GACTAT]-3'OH) originates from the forward strand and that product 2 (3'-OH[Rev-CTGATA]-5'OH) originates from the reverse strand.

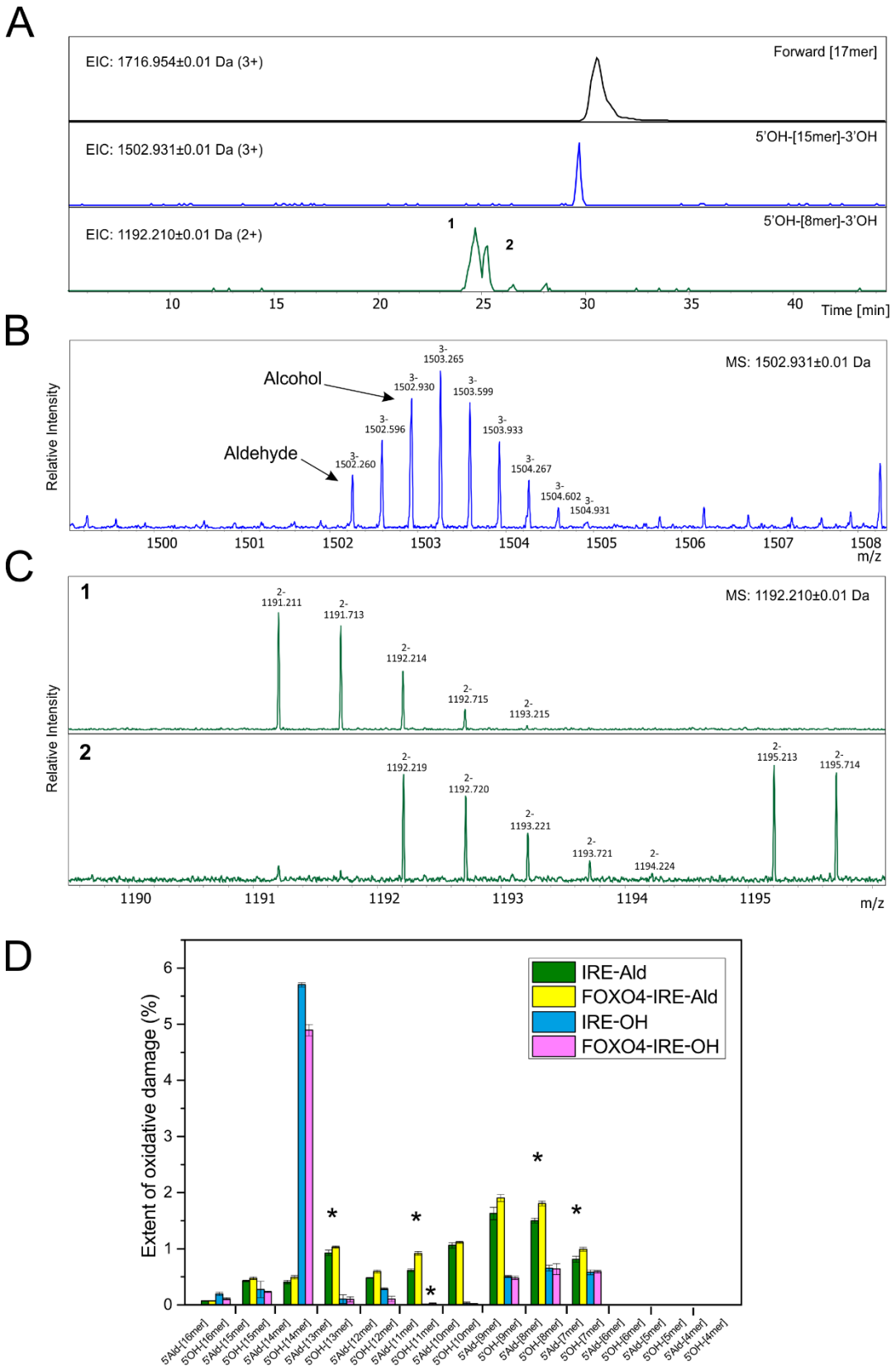


Figure S9. A. LC-MS trace of the forward strand (17 mer, top panel) and two fragments created upon its fragmentation, 5'OH-[15mer]-3'OH 5'OH-[8mer]-3'OH. B) MS spectrum zoomed-in at 5'OH-[15mer]-3'OH where alcohol and aldehyde forms were not separated on a reverse/phase during LC-MS, resulting in an overlapped product isotopic peak. C) MS spectrum zoomed-in at 5'OH-[8mer]-3'OH where aldehyde (1.) and alcohol products (2.) were chromatographically separated, also leading to different isotopic patterns of aldehyde and alcohol products. D) The extent of oxidative damage quantified for both aldehyde (green and yellow for IRE and FOXO4-IRE, respectively) and alcohol (blue and pink for IRE and FOXO4-IRE,

respectively) forms. T-test results are indicated by asterisks above the histograms. T-test legend: * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$), **** ($p \leq 0.0001$).

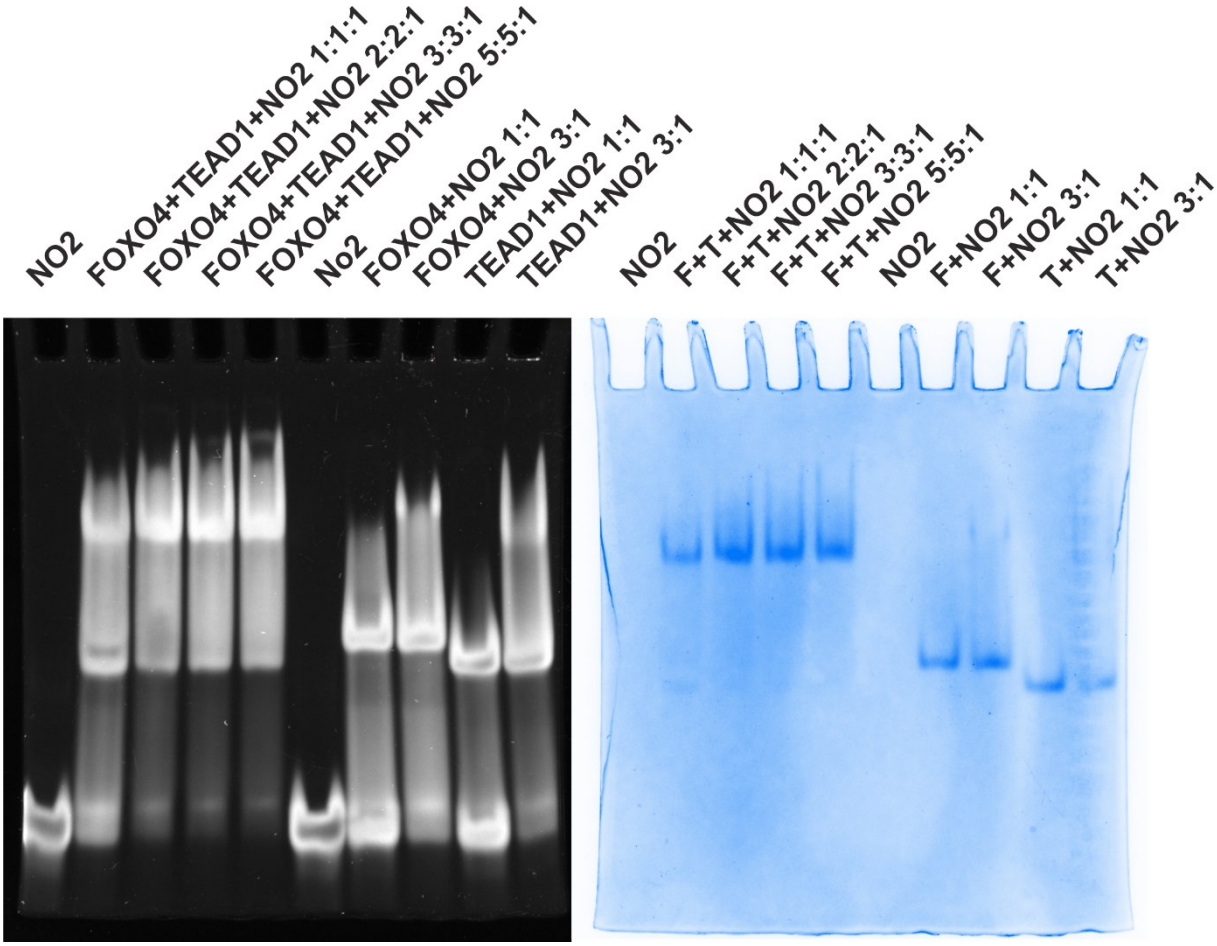


Figure S10. Electrophoretic mobility shift assay of the FOXO4-DBD•TEAD1-DBD•NO2 , TEAD1-DBD•NO2 and FOXO4-DBD•NO2 complex stained with GelRed® (left panel) for nucleic acid and with Coomassie Brilliant Blue for protein (right panel) detection. Whether alone or together, the proteins caused a mobility shift in the oligonucleotide in comparison with the NO2 duplex alone.

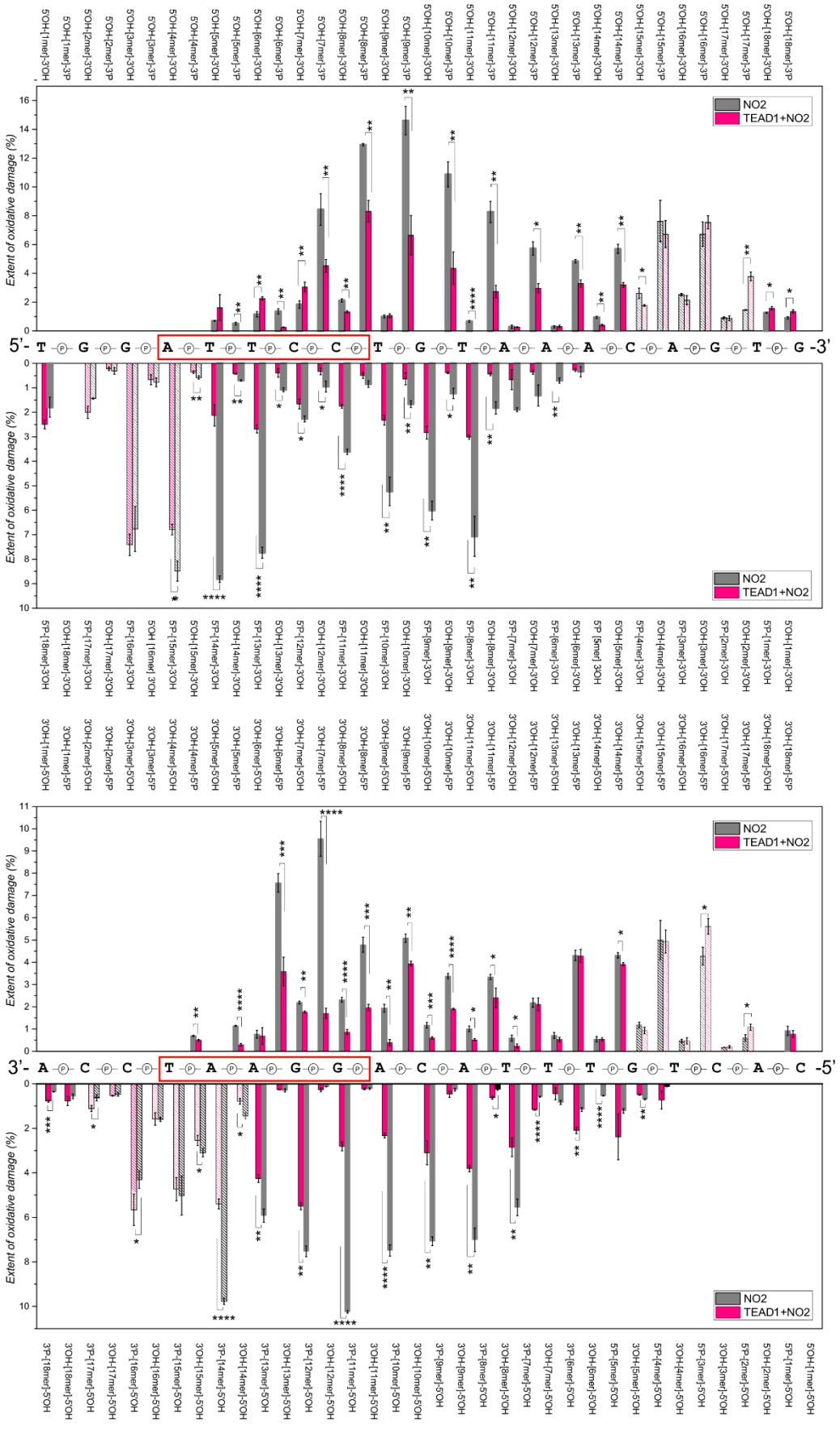


Figure S11. Quantified extent of oxidative damage of NO2 with (pink) and without (grey) TEAD1-DBD. Dashed histogram represents isobaric fragments with the same mass and retention time. T-test results are indicated by asterisks above the histograms. T-test legend: * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$), **** ($p \leq 0.0001$).

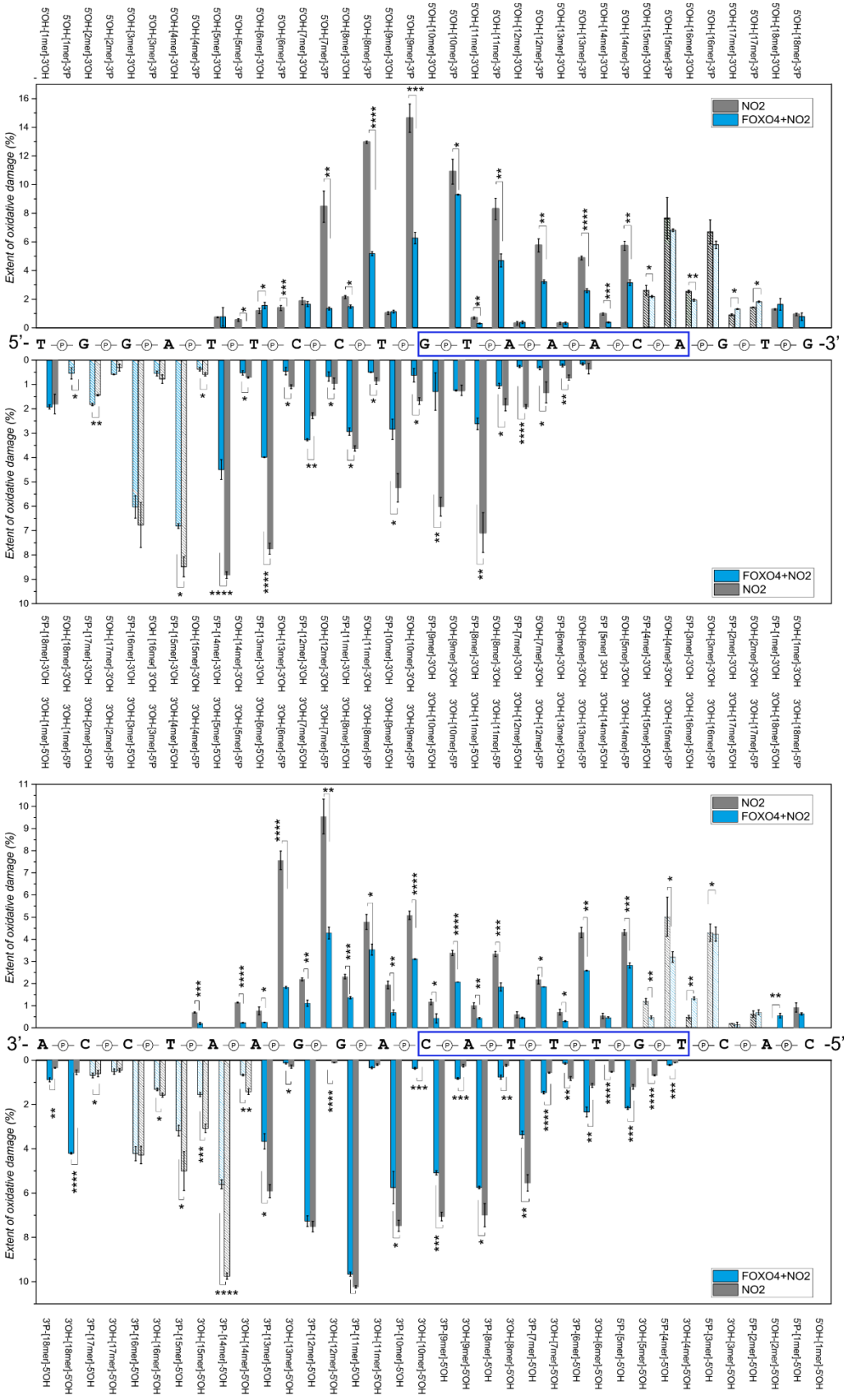


Figure S12. Quantified extent of oxidative damage of NO_2 with (pink) and without (grey) FOXO4-DBD. Dashed histogram represents isobaric fragments with the same mass and retention time. T-test results are indicated by asterisks above the histograms. T-test legend: * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$), **** ($p \leq 0.0001$).

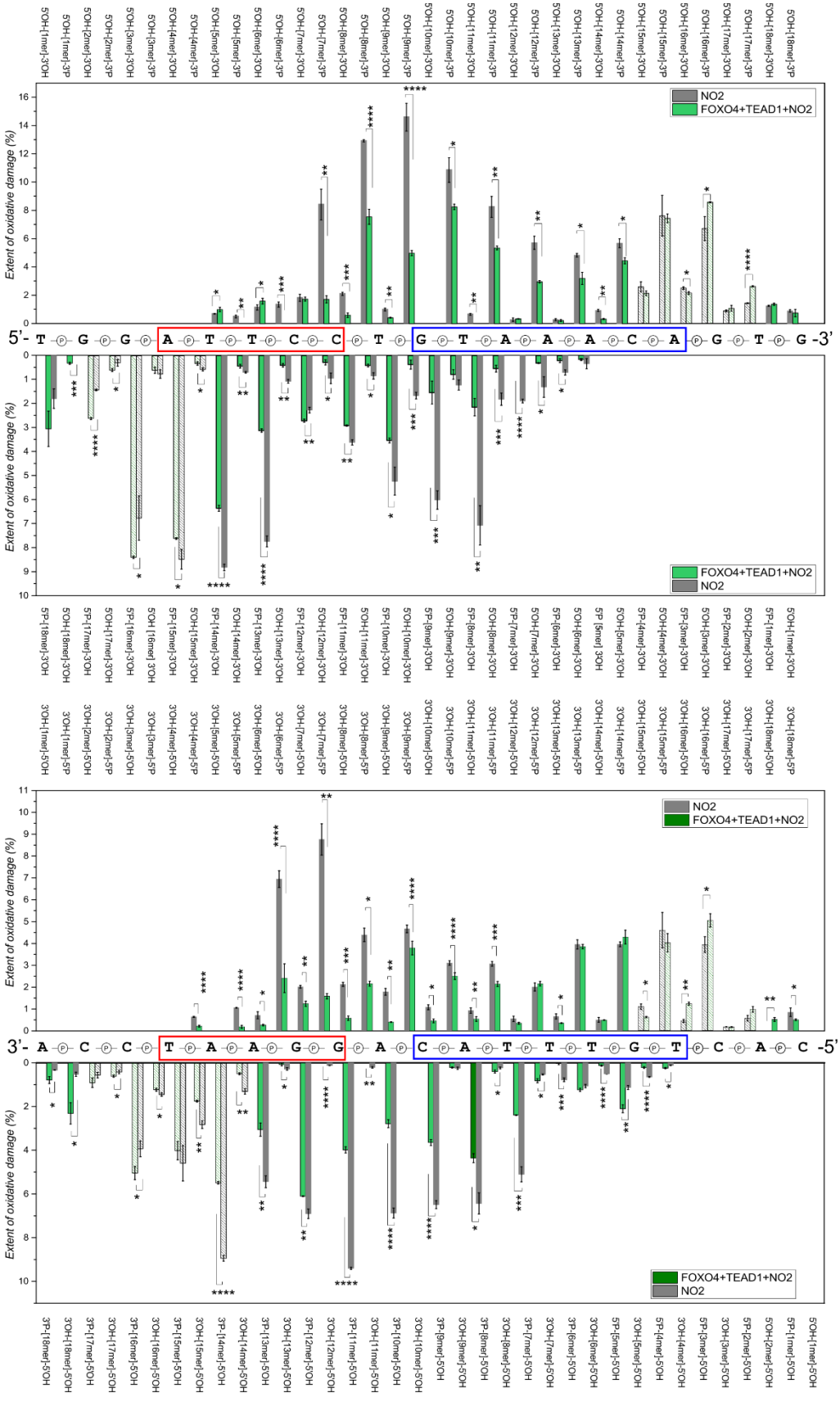


Figure S13. Quantified extent of oxidative damage of NO2 with (pink) and without (grey) FOXO4-DBD and TEAD1-DBD. Dashed histogram represents isobaric fragments with the same mass and retention time. T-test results are indicated by asterisks above the histograms. T-test legend: * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$), **** ($p \leq 0.0001$).

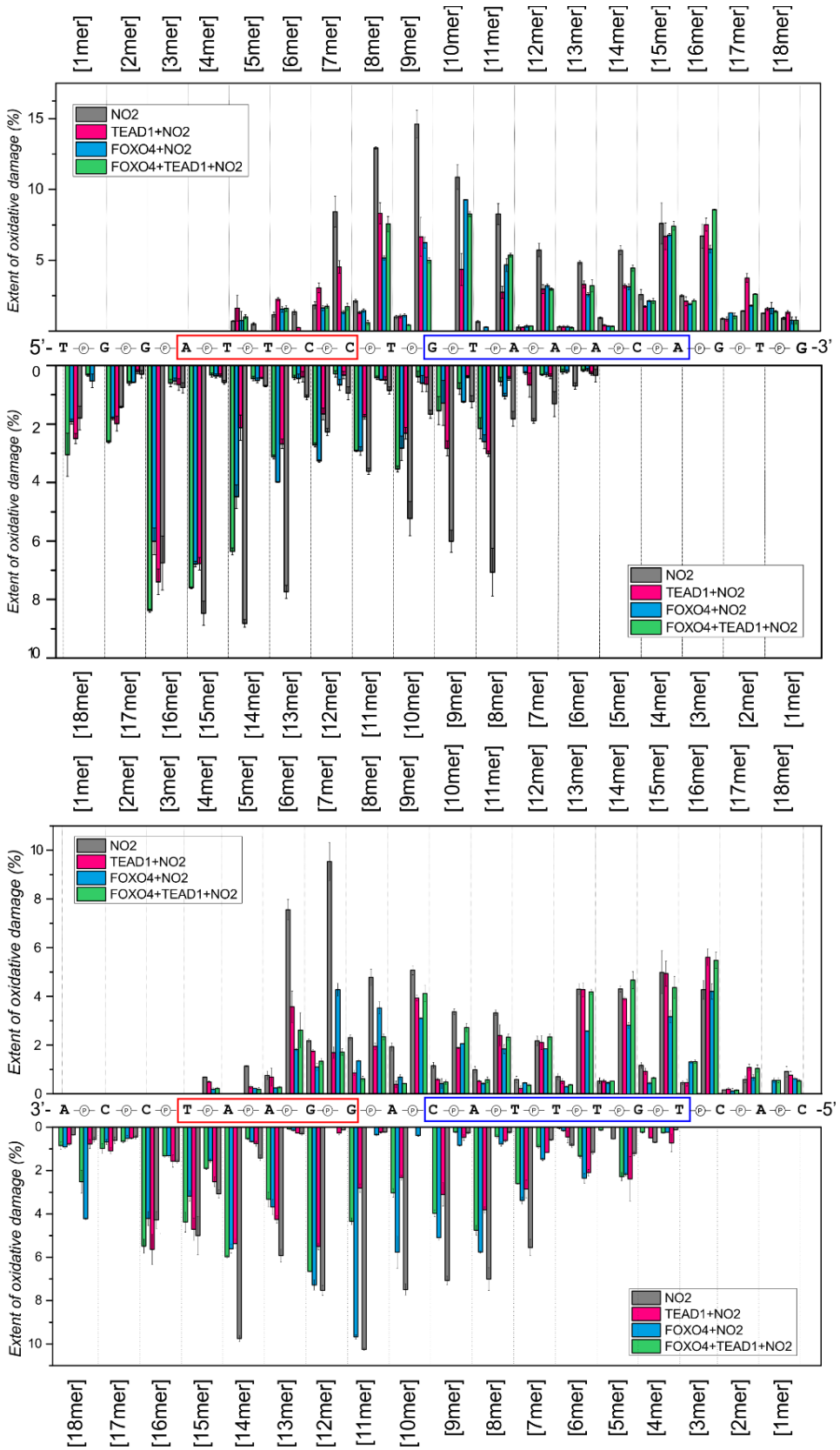


Figure S14. Quantified extent of oxidative damage of NO2 forward and reverse strands. Grey histograms – extent of oxidative damage of NO2 duplex in solution, pink histograms – extent of oxidative damage of TEAD1-NO2 in solution, blue histograms – extent of oxidative damage of FOXO-NO2 in solution, pink histograms – extent of oxidative damage of TEAD1-FOXO4-NO2 complex in solution.

Attached Publications:

Publication IV

Felčíková, K.; Hovan, A.; **Polák, M**.; Loginov, D.S.; Holotová, V.; Díaz, C.; Kožár, T.; Lee, O.; Varhač, R.; Novák, P., Bánó, G.; Sedlák, E.

Design of AsLOV2 domain as a carrier of light-induced dissociable FMN photosensitizer.

Protein Science, 2024, *in press*.

My contribution: top-down analysis of protein samples, preparing figures for publications, manuscript writing contribution.

Design of AsLOV2 domain as a carrier of light-induced dissociable FMN photosensitizer

Kristína Felčíková[1⁋], Andrej Hovan[1⁋], Marek Polák[2,3], Dmitry S. Loginov[2], Veronika Holotová[4], Carlos Díaz[4], Tibor Kožár[4], One-Sun Lee[4], Rastislav Varhač[5], Petr Novák[2], Gregor Bánó[1], and Erik Sedlák*[4,5]

[1]Department of Biophysics, Faculty of Science, P.J. Šafárik University, Jesenná 5, 041 54 Košice, Slovakia

[2]Institute of Microbiology, Průmyslová 595, 252 50 Vestec, Czech Republic

[3]Department of Biochemistry, Faculty of Science, Charles University, Prague 12843, Czech Republic

[4]Center for Interdisciplinary Biosciences, Technology and Innovation Park, P.J. Šafárik University, Jesenná 5, 041 54 Košice, Slovakia

[5]Department of Biochemistry, Faculty of Science, P.J. Šafárik University, Moyzesova 11, 041 54 Košice, Slovakia

⁋ The authors contributed equally to this work

* To whom correspondence should be addressed: Center for Interdisciplinary Biosciences, Technology and Innovation Park, P.J. Šafárik University, Jesenná 5, 041 54 Košice, Slovakia; Erik Sedlák, email: erik.sedlak@upjs.sk

Abstract

Flavin mononucleotide (FMN) is a highly efficient photosensitizer (PS) yielding singlet oxygen ($^1O_2$). However, its $^1O_2$ production efficiency significantly decreases upon isoalloxazine ring encapsulation into the protein matrix in genetically encoded photosensitizers (GEPS). Reducing isoalloxazine ring interactions with surrounding amino acids by protein engineering may increase $^1O_2$ production efficiency GEPS, but at the same time weakened native FMN-protein interactions may cause undesirable FMN dissociation. Here, in contrast, we intentionally induce the FMN release by light-triggered sulfur oxidation of strategically placed cysteines (oxidation-prone amino acids) in the isoalloxazine-binding site due to significantly increased volume of the cysteinyl side residue(s). As a proof of concept, in three variants of the LOV2 domain of *Avena sativa* (AsLOV2), namely V416C, T418C, and V416C/T418C, the effective $^1O_2$ production strongly correlated with the efficiency of irradiation-induced FMN dissociation (WT<V416C<T418C<V416C/T418C). This alternative approach enables us: (i) to overcome the low $^1O_2$ production efficiency of flavin-based GEPSs without affecting native isoalloxazine ring-protein interactions and (ii) to utilize AsLOV2, due to its inherent binding propensity to FMN, as a PS vehicle, which is released at a target by light irradiation.

Keywords: genetically encoded photosensitizers, flavin cofactor, singlet oxygen, LOV2 domain, miniSOG

Significance statement

Photosensitizer encapsulation into a protein matrix is accompanied by a significant decrease in its singlet oxygen, $^1O_2$, production efficiency (Φ_Δ). Typical examples are flavoproteins in which Φ_Δ of FMN decreases upon binding to protein by more than 10-fold as a result of the quenching of its triplet state due to isoalloxazine ring interactions with the surrounding amino acids. Consequently, the conventional approach to increase Φ_Δ is to reduce interactions of amino acids with the isoalloxazine ring. By contrast, our approach increases Φ_Δ by introducing $^1O_2$ oxidation prone amino acid close to the isoalloxazine ring with minimal effects on protein-cofactor interactions in the native state of the flavoproteins, but which actively trigger FMN dissociation upon irradiation with blue light *via* amino acid(s) oxidation. This enables proper targeting of FMN as photosensitizer by a specific flavoprotein and unleash its ability to produce $^1O_2$ by light irradiation.

Introduction

Photosensitizers (PSs) are small organic compounds that induce cytotoxicity when irradiated with a specific light source (Dougherty et al., 1998; Lee et al., 2020). PSs possess high potential in clinical application in photodynamic therapy (PDT) (O'Connor et al., 2009; Zhang et al., 2018) as well as an alternative therapy against multi-drug resistant bacteria ~~as well~~ (Hilgers et al., 2019). However, PSs' applications are significantly affected by their low solubility and stability in aqueous solvents and their inherently low specificity to target cells (Mansoori et al., 2019). These limitations have led to an intensive effort to develop an efficient genetically encoded PSs (GEPSs).

GEPSs are proteins that bind to specific chemical PSs with several biological and medical applications, including immunophotosensitizing, chromophore-assisted light inactivation (CALI) of proteins and photoablation of cells, photodynamic diagnosis (PDD), antimicrobial photodynamic inactivation (aPDI), and correlative light electron microscopy (CLEM) (Rodriguez-Pulido et al., 2016; Souslova et al., 2017; Hilgers et al., 2019;Gunaydin et al., 2021). GEPSs have several advantages over chemical PSs: (i) protein binding overcomes the low solubility of chemical PSs, (ii) the protein matrix provides an encased phototoxic PS with a constant local environment, regardless of subcellular location, fostering a reliable production of reactive oxygen species (ROS), and (iii) conjugation with proper tags (leader peptides, targeted proteins, such as antibodies, and DARPins) by genetic engineering enables the targeted accumulation of GEPSs at specific subcellular structures or compartments (Proshkina et al.,

2015; Serebrovskaya et al., 2009; Westberg et al., 2019). GEPSs are divided into two main classes. The first class includes proteins with a green fluorescent protein-like structure (Bulina et al., 2006; Gorbachev et al., 2020; Micheletto et al., 2021; Sarkisyan et al., 2015; Takemoto et al., 2013) such as KillerOrange, KillerRed, as well as improved versions of KillerRed with enhanced phototoxicity known as SuperNova and SuperNova 2. The second class encompasses flavin-binding proteins derived from the light-oxygen-voltage (LOV) photoreceptor domain of plants, algae, and bacteria (Petrencakova et al., 2020; Ruiz-Gonzalez et al., 2013; Souslova et al., 2017; Westberg et al., 2017a; Westberg et al., 2015). The main difference between the two classes of GEPS, apart from the chemical nature of PS, is that the second group of PS is dissociable.

The cofactor flavin mononucleotide (FMN), which represents naturally occurring PS as a part of the second group of GEPS, is characterized by a $^1O_2$ production efficiency (Φ_Δ) of ~ 0.51-0.65 (Baier et al., 2006; Westberg et al., 2017b). However, upon binding to mini-singlet oxygen generator (miniSOG) protein (Shu et al., 2011), its Φ_Δ decreases by more than 10 fold, $\Phi_\Delta \leq 0.05$, demonstrating a major drawback of GEPSs (Endres et al., 2018; Petrencakova et al., 2020; Ragas et al., 2011; Westberg et al., 2017a; 2017b; Westberg et al., 2015). Such significant decrease in Φ_Δ is a result of numerous static and dynamic interactions of amino acids surrounding the tightly encapsulated FMN in the protein, including hydrogen bonding, van der Waals, and π–π interactions (van den Berg et al., 2002). Nonetheless, the value of Φ_Δ depends on the lifetime of the FMN triplet state (^{3}FMN), which is, in turn, affected by how the protein matrix (i) quenches ^{3}FMN and (ii) restricts oxygen diffusion towards the PS. Thus, protein engineering methods have been used to reduce the electron affinity of ^{3}FMN by removing intermolecular H-bonds and precluding photoinduced electron transfer reactions between the isoalloxazine ring and the protein scaffold, thereby significantly increasing the Φ_Δ to values similar to those of free FMN (Westberg et al., 2017a; Westberg et al., 2015). But this approach is quite complex, requiring carefully removing protein-cofactor interactions that decrease the Φ_Δ value without significantly affecting the protein affinity of the cofactor. Moreover, this approach inevitably weakens the protein-cofactor bond and hence may lead to an undesirable cofactor dissociation.

The comparison between the Φ_Δ values of FMN and FAD, which are 0.51-0.65 and ~0.07, respectively, shows the extreme sensitivity of the isoalloxazine moiety on a close environment (Baier et al., 2006; Westberg et al., 2017b). Although numerous interactions affect the isoalloxazine moiety when encapsulated in the protein matrix of the flavoprotein miniSOG

(mini-singlet oxygen generator), the Φ_Δ of miniSOG was increased in two steps, first to 0.19-0.23 in the variant SOPP (singlet oxygen photosensitizing protein) (Westberg et al., 2015) and then to a similar value to that of free FMN (~0.6) in the variant SOPP3 (Westberg et al., 2017a). Recently, another strategy how to increase miniSOG Φ_Δ value has been identified, namely the photo-induced transformation of FMN into lumichrome, which increases the accessibility of oxygen to the isoalloxazine ring and makes protein quenching less favorable (Torra et al., 2019). These improved variants of flavin-based GEPSs were result of replacing amino acids (by mutation or as a result of oxidation) responsible for: (i) steric barriers to oxygen diffusion toward the isoalloxazine ring, (ii) FMN triplet state quenching by electron transfer, and (iii) $^1O_2$ quenching by chemical reactions. However, the increased values of Φ_Δ were offset by weakened cofactor-protein interactions, potentially releasing the cofactor, which is due to a reactive nature of free FMN an unwanted outcome. In fact, the potential deleterious effects of this highly efficient PS are typically minimized by deeply burying the isoalloxazine moiety of flavin cofactors in the protein matrix of flavoproteins or storage proteins (Meissner et al., 2007; Stockwell & Thornton, 2006).

Considering the above and the findings of our previous study in which we observed that the mutation of Cys450 localized close to isoalloxazine ring for alanine affected both FMN dissociation and Φ_Δ(Petrencakova et al., 2020), we aimed at developing an alternative approach to increasing the Φ_Δ of GEPS system. The term "GEPS system" aims to point to the fact that our approach leads to the creation of two/three-component system containing a GEPS (flavoprotein) and/or its apoform and free FMN, which arose from a single entity represented by a GEPS (flavoprotein). In the presented approach, we introduced strategic mutations to minimize the effects on protein-cofactor interactions in the native state of the flavoproteins, but which actively trigger FMN dissociation upon irradiation with blue light as a result of oxidation of the designed cysteine(s) in the binding site of isoalloxazine moiety. In such case, the potential of $^1O_2$ production by isoalloxazine ring is fully unlocked near the GEPS by avoiding the two major reasons for the decreased Φ_Δ value: (i) quenching of the cofactor ^{3}FMN state and (ii) limited oxygen diffusion due to encapsulation in a protein scaffold.

Materials and methods

Cloning, Expression, and Purification. The AsLOV2 domain and the corresponding mutants were cloned, expressed, and purified as previously reported (Petrencakova et al., 2020). Briefly, all proteins were expressed in *E. coli* strain BL21(DE3). The bacterial cells were

grown at 37 ºC in ampicillin containing (100 µg/ml) TB medium to OD_{600} ~0.6-0.8. Protein expression was induced by adding isopropyl β-D-galactopyranoside (100 µM final concentration) following a temperature downshift from 37 ºC to 25 ºC and expressed in the dark, overnight, at 25 ºC. The proteins were purified by metal ion affinity chromatography (Ni-NTA Superflow, Qiagen) followed by purification on a Superdex 75 Increase, 10/300 GL, size exclusion column and concentrated in 20 mM TrisHCl buffer, pH 7.8. All steps were performed in the dark. The 280 nm/477 nm absorbance ratio of the final protein was ~2.6, suggesting the absence of the AsLOV2 apoform (Zayner et al., 2012).

Time-resolved singlet oxygen phosphorescence detection. The experimental apparatus used for singlet oxygen detection was described in detail elsewhere (Hovan et al., 2023). Briefly, short 5-7-ns-long laser pulses with 475-nm wavelength originating from an optical parametric oscillator (OPO) pumped by an Nd: YAG Q-switched laser excited 25 µM AsLOV2 samples (2 ml) at 10 Hz frequency. The average laser power was set to 0.8 mW with estimated beam diameter of 50 µm, and the samples were stirred with an overhead glass stirrer to minimize photobleaching. The time-resolved singlet oxygen phosphorescence signal (at 1270 nm) was measured on a photomultiplier tube (Hamamatsu H10330A-75), in photon counting mode, connected to a multichannel scaler PCI card (Becker&Hickl, MSA-300). Spectral background was subtracted from the signal as explained in our previous study (Hovan et al., 2023). Two additional band-pass filters were used to detect the background in the spectral regions neighboring the 1270 nm $^1O_2$ peak. On average, 2,500 laser pulses were consecutively detected with each filter throughout the experiments. A single phosphorescence kinetic curve was obtained in 12.5 minutes (3 x 2500 s pulses). It is important to note that the sample irradiation by the pulsed blue laser light during the $^1O_2$ measurements also triggers the FMN dissociation process in our experiments.

Measurements of FMN triplet state lifetime. To monitor the FMN triplet state, we used an additional 633-nm cw laser during the time-resolved singlet oxygen phosphorescence detection, as detailed in our previous study (Petrencakova et al., 2020). Time-resolved absorption at 633 nm was then recorded on an avalanche photodiode (Thorlabs, APD110A2) connected to a digitizing oscilloscope (Tektronix, DPO 7254). During the experiment, we consecutively collected the average signal from 2,500 laser pulses throughout the irradiation process. Over 75 minutes, we constructed a total of eighteen decay curves for each sample due to utilization of three different filters to be able correctly subtract background and to determine the amplitude of the measured signals. These curves allowed us to monitor the irradiation

dependence of the triplet state amplitudes and lifetimes and obtain parameters used for an evaluation of the singlet oxygen phosphorescence data.

Spectral analysis. *Ultraviolet and visible absorption spectra* were recorded on UV-2401PC (Shimadzu) and V-750 (Jasco) UV-Vis spectrophotometers. Protein concentrations were calculated using an extinction coefficient of ε_{447}=13800 $M^{-1} \cdot cm^{-1}$ for oxidized FMN. *Fluorescence emission spectra* were recorded on RF-5301PC (Shimadzu) and FP 8550 (Jasco) spectrofluorophotometers. The emission spectra of FMN were measured using excitation wavelengths at 445 nm, and the fluorescence spectra were measured using a protein concentration of 10 μM. *Circular dichroism spectra* measurements were performed on a Jasco 810 spectropolarimetry (Jasco) at 10 μM protein concentration. The measurements in the far-UV and near-UV spectral regions were performed in a quartz cuvette with 1 mm and 0.5 cm pathlengths, respectively.

Determination of a light-induced released FMN. Relative amounts of released FMN from AsLOV2 wt and its variants were determined by FMN fluorescence after filtration using 10 kDa cut-off filter tubes. Each sample, i.e. before and after irradiation AsLOV2 wt and its variants, 900 μl of 10 μM protein, was loaded into Amicon Ultra Centrifugal filter tube and centrifuged for 5 min at 7500g. After the spin, the collected flow-through of each sample was measured for FMN fluorescence.

Thermal stability of ASLOV2 variants. The thermally induced transitions of AsLOV2 variants were measured on an RF-5301PC spectrofluorophotometer by monitoring the fluorescence emission at 495 nm (λ_{ex}=450 nm) while increasing the temperature, controlled with a Peltier block, from 25°C to 90°C and recording data at a 1.5°C/min scan rate. All measurements were performed in 25 mM phosphate buffer, pH 7.9.

Thermal denaturation was analyzed by fitting normalized experimental data according to the following equation:

$$y_{obs} = \frac{\exp\left[\frac{\Delta H_{vH}}{R}.\left(\frac{1}{T}-\frac{1}{T_m}\right)\right]}{1+exp\left[\frac{\Delta H_{vH}}{R}.\left(\frac{1}{T}-\frac{1}{T_m}\right)\right]} \quad (1)$$

where y_{obs} is the experimentally measured parameter, ΔH_{vH} is the van´t Hoff enthalpy change, R is the gas constant, and T_m is the transition temperature.

The thermal stability of the study proteins was also analyzed by differential scanning calorimetry (DSC). DSC measurements were performed on a VP-DSC differential scanning microcalorimeter (Microcal, U.S.A.), and the stabilities of the variants were assessed by

determining the temperatures at the maxima of the heat capacity change curves of the corresponding proteins.

Molecular Modeling. The Schrödinger suite of programs (Schrödinger, 2022) was used to process the AsLOV2-related protein structures downloaded from the PDB database (2V1A.pdb; 2V1B.pdb) (Halavaty & Moffat, 2007). The protein structures were treated and prepared for calculations in Maestro (*Maestro*, 2022). The updated structures were then single (V416C, V416M, T418C, T418M) and double mutated (V416C/T418C, V416C/T418M, V416M/T418C, V416M/T418M). The geometries of the mutated structures were optimized and then solvated with water molecules added with a 1 nm buffer at neutral pH around the proteins. The geometries of the resulting solvated protein structures were then minimized again and equilibrated. After equilibration, the final structures were submitted for 500 ns NPT (pressure at 1.01325 bar) molecular dynamics (MD) simulations with the Desmond program (Shaw, 2020) at 300 K. Molecular geometries resulting from simulations were saved at 100 ps intervals and used for further analysis. Molecular complexes were visualized in the Maestro program.

Free Energy of Binding. We calculated the binding free energy of FMN to the AsLOV2 domain, i.e., potential of mean force (PMF), by pulling an FMN with the constant velocity (ν_z = 1 Å/ns) through steered molecular dynamics (SMD) simulation from z=7 (inside of the AsLOV2 domain) to z=25 (outside of the AsLOV2 domain) (see Figure S1 for further details on the SMD). At each window, the system qA equilibrated for 1 ns while constraining the position of FMN and the α-atoms of the protein at 310 K using the CHARMM36m force field (Huang et al., 2017). The initial structure of the complex of FMN and AsLOV2 domain was adapted from the crystallographic structure (PDB ID: 2v0u) (Halavaty & Moffat, 2007). We also adapted three mutants (V416C, T418C, and V416C/T418C) of the AsLOV2 domain to compare the effect of the position of cysteine residue (Figure S2) on FMN binding. For this purpose, we divided the path from z = 7 to z = 15 into 8 smaller windows (the length of each window is 1 Å), performing eight independent SMD simulations in each window. From z = 15 to z = 25, we divided the path into 5 smaller windows (2 Å window length), also performing eight independent SMD simulations in each window. PMF was constructed from SMD simulations based on the Jarzynski equality equation (Jarzynski, 1997; Park & Schulten, 2004)

$$\Delta A = -\beta^{-1} \ln\langle \exp[-\beta W] \rangle \quad (2),$$

where ΔA is a free energy difference, β is the product of the Boltzmann factor and temperature, and W is the non-equilibrium work derived from SMD simulations. The non-equilibrium work of the pulling force was assessed using the following equation:

$$W = -kv \int_0^t dt' [z(t') - z_0 - vt'] \quad (3),$$

where k and v are the force constant (418.4 kJ/mol/Å^2) and velocity of pulling (1 Å/ns), and $z(t')$ and z_0 are the reaction coordinate at t' in the simulation and the initial position of the center of mass of FMN, respectively. Using a second-order cumulant expansion of equation 2, we derived the following equation:

$$\Delta A = \langle W \rangle - \frac{\beta}{2} [\langle W^2 \rangle - \langle W \rangle^2] \quad (4)$$

The system with 66 Å × 66 Å × 78 Å dimensions was filled with 10612 modified TIP3P water (Figure S3) (Price & Brooks, 2004), and the pressure of 1 atm was maintained using the Langevin piston method with a piston period of 100 fs, a damping time constant of 50 fs, and a piston temperature of 310 K (Feller et al., 1995; Martyna et al., 1994). Full electrostatics were applied using the particle-mesh Ewald method with a 1 Å grid width (Darden et al., 1993). A group-based cutoff was used to calculate nonbonded interactions with a switching function, updating every tenth time-step. Using the SHAKE algorithm, the hydrogen bond was kept rigid, with a 2 fs time step (Andersen, 1983). All simulations were performed using nanoscale molecule dynamics (NAMD) (Phillips et al., 2005), and the graphics shown in this report were prepared using visual molecular dynamics (VMD) (Humphrey et al., 1996).

Top-down mass spectrometry. To prevent light-induced oxidation, the denatured protein samples were desalted in the dark. Protein samples were denatured by mixing them with 0.1% formic acid in a 1:1 (v:v) ratio. Denatured protein samples (10 μg) were loaded onto a reverse-phase microtrap column (Optimize Technologies, USA), desalted using 0.1% FA (3 × 250 μl). Subsequently, the samples were eluted with 80% MeCN, 0.1% FA. The desalted samples were then diluted 50× using 30% MeCN, and 0.1% FA before being sprayed into a SolariX-XR mass spectrometer equipped with a 15T magnet (Bruker Daltonics, Billerica, USA) and an ESI source and operated in positive ionization mode. MS spectra were acquired in a broadband mode ranging from 207.2 to 3000 m/z. The collision voltage was set to -5.0 V at 150 °C desolvation temperature, 0.200 s ion accumulation time, and 1 ms time of flight, acquiring 128 scans using a 2M data point transient starting at 207.20 m/z. For MSMS, oxidized clusters were isolated in a quadrupole and transferred into the collision cell for collision-induced dissociation (CID). Oxidized ion clusters were isolated for 4 charge states

using multi-continuous accumulation of selected ions (multi-CASI) mode at the following nominal masses: 947.00 amu (21+), 996.00 amu (20+), 1046.00 amu (19+), 1108.00 amu (18+). CID was induced by setting a quadrupole selection window of 15.00 amu for each charge state and applying -13.00 / -15.00 / -17.00/ -19.00 V collision voltage for +21/+20/+19/+18 charge states, respectively, setting ion accumulation at 0.5 s. MSMS spectra were acquired by collecting 256 scans with a 2M data point transient starting at 207.20 m/z.

Bottom-up mass spectrometry. For the analysis was taken an aliquot from each sample corresponding to 0.5 µg of the protein. The samples were reduced and alkylated by adding TCEP [tris(2-carboxyethyl)phosphine] and chloroacetamide at a concentration of 10 mM and 30 mM, respectively. Then, the protein samples were incubated at 90 °C for 2 min and cooled down on the bench. Further, samples were diluted using 50 mM 4-ethylmorpholine buffer pH = 8.3 containing 5% acetonitrile (ACN) and digested by trypsin (ratio 2:1) at 37 °C for 2 h. Digestion was terminated by adding trifluoroacetic acid (TFA) to a final concentration of 0.5%. Samples were desalted using StageTip approach and dried. Prior LC-MSMS analysis samples were dissolved in containing 2% ACN and 0.1% TFA. Shotgun proteomic analysis was performed using the Vanquish liquid chromatography system (Thermo Scientific) coupled to the timsToF SCP mass spectrometer equipped with Captive spray (Bruker Daltonics) and operated in a positive data-dependent mode. Peptides were trapped using C18 trap column (PepMap Neo C18 5µm, 300 µm × 5 mm, Thermo Scientific) and separated on a C18 analytical column (DNV PepMap Neo 75 µm × 150 mm, 2 µm, Thermo Scientific) using a linear gradient of 5% to 35% ACN over 50 minutes at a flow rate of 350 nL/min. Both the trap and analytical columns were heated to 50 °C. The timsTOF SCP settings were based on the standard proteomics PASEF method. Only light-induced samples were measured in DDA mode to prepare spectral library. Quantification of modification extent was done using data acquired by DIA mode on the timsTOF SCP. DIA analysis was done using the same LC setup. All samples were run in triplicates.

Data processing. Top-down raw data were deconvoluted using Data Analysis 5.3 (Bruker Daltonics) and processed as already described (Polak et al., 2022; Yassaghi et al., 2022). Fragment ions from raw spectra were assigned using an in-house software tool, and a search was conducted against an *in-silico* library of either wild-type or mutant protein sequences with a mass accuracy of 2 ppm. The results were manually validated.

DDA data were processed by FragPipe search workflow (Kong et al., 2017; Yu et al., 2020). The search was performed against a database containing sequences of wild type and

respective mutants of AsLOV2 and cRAP contaminant sequences. Precursor ion tolerance was set at 12 ppm, and the mass tolerance for MS/MS fragment ions was set at 0.05 Da. Carbamidomethylation of cysteine and various oxidative modifications of Met, Phe, Trp, Tyr, His, Pro, and Cys (+O, +2O, $+O_2$-H_2O, etc.) were set as a variable, the false-discovery ratio was set to less than 1%. Obtained results were exported to the Skyline software (MacLean et al., 2010), where fragmentation spectra were manually validated. Quantification was done using DataAnalysis 5.3 software and in-house python scripts. Data were corrected by subtracting background oxidation. The extent of oxidative modifications and statistical analysis were performed according to the work of Loginov et al. (Loginov et al. 2022).

Results and Discussion

Design of efficient GEPS system from AsLOV2 wt. The AsLOV2 domain belongs to the superfamily of Per-Arnt-Sim (PAS) domains, which are found in various sensor proteins, in organisms ranging from archaea to eukaryotes (Taylor & Zhulin, 1999). The crystal structure of AsLOV2 displays the typical PAS fold, flanked by a conserved N-terminal turn-helix-turn motif and a C-terminal flanking region containing an amphipathic Jα helix. These regions dock on the LOV2 core domain and bury several hydrophobic residues of the central β-sheet of the core domain (Figure 1A, S3) (Halavaty & Moffat, 2007). The isoalloxazine moiety is buried deep in the protein, connected to solvent molecules through several tunnels, as shown in our recent study(Petrencakova et al., 2020). Based on our previous results, we hypothesized that suitably replacing small amino acids near the isoalloxazine ring with $^1O_2$ oxidation-prone amino acids, such as cysteine or methionine, could lead to irradiation-triggered flavin cofactor

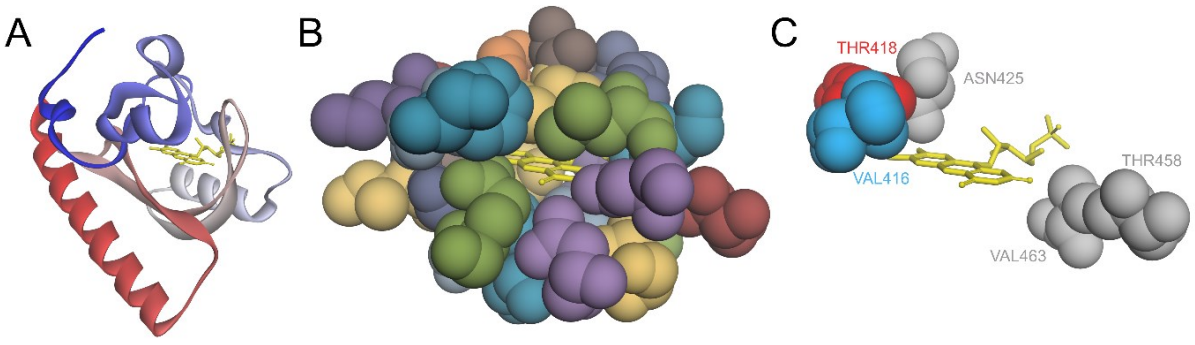


Figure 1. (A) Crystal structure of AsLOV2 (PDB: 2v0u). The N-terminal turn-helix-turn motif is shown in dark blue, and the Jα helix is shown in red. (B) Detailed view of the binding site of the isoalloxazine moiety, indicating a tight arrangement around the isoalloxazine ring. (C) Isoalloxazine ring binding site, highlighting the five amino acids that may be replaced based on the selection criteria (see main text). The amino acids Val416 and Thr418 were modified in this study.

dissociation. To identify suitable amino acids for replacement, we applied the following exclusion criteria: (i) conservative amino acids – to avoid adversely affecting flavin binding, (ii) large amino acids – replacing large with small amino acids is sterically meaningless as the increase of their volumes by oxidation may not be sufficient to create "crowding effect" in the isoalloxazine moiety binding site, and (iii) positively charged amino acids – they may stabilize negatively charged flavin. As a result, only five amino acids (Val416, Thr418, Asn425, Thr458, and Val463) were left close to the isoalloxazine ring (Figures 1C, S4). After visual inspection, we decided to replace Val416 and Thr418 by cysteinyl and methionyl residues. Of the eight possible variants, only three containing cysteines, i.e., V416C, T418C, and V416C/T418C, were expressed in *E. coli* in ample concentration for biophysical characterization. These observations matched our MD simulations (Figure S4), thus indicating that replacing any of the two positions by methionyl residues perturbs helices and most likely decreases the stability of the whole protein structure.

Conformational and stability analysis. To assess the possible impact of the mutations on the stability and conformational properties of the proteins, we performed (i) *in silico* analysis of FMN binding to AsLOV2 variants, (ii) CD measurements in the far- and near-UV spectral regions, and (iii) thermal denaturation experiments, by following the variation of FMN fluorescence as a function of temperature, and DSC measurements of all AsLOV2 variants. FMN binding to the proteins was analyzed *in silico* by calculating the potential of mean force by pulling an FMN at constant velocity in steered molecular dynamics simulations (Figure S1). No significant differences in dissociation constants of the FMN cofactor were found when comparing the free energy of pulling between wild type and mutant AsLOV2 variants (Figure 2A).

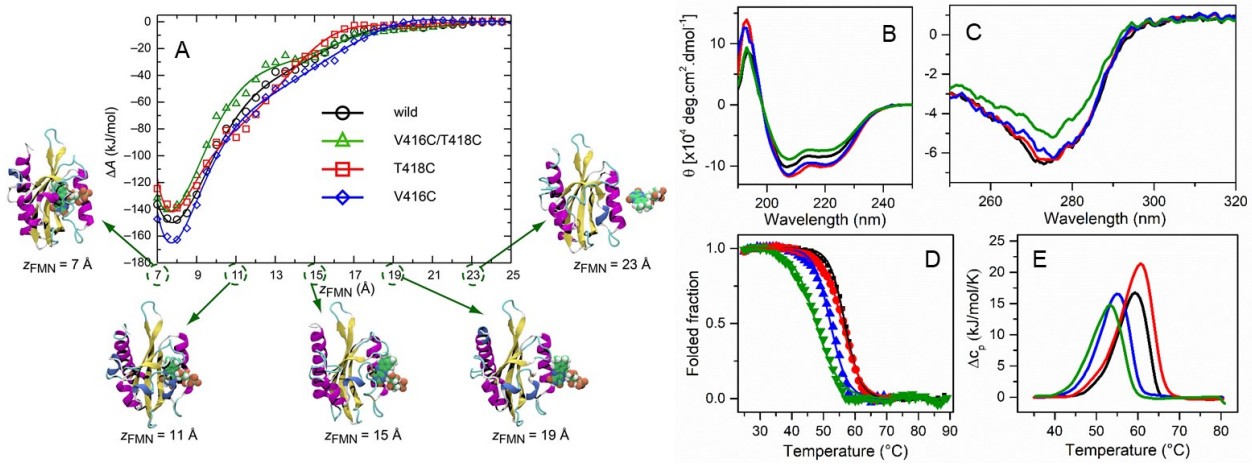


Figure 2. (A) Free energy profiles of FMN binding to wild type and mutant AsLOV2 domains are shown as a function of the position of the center of FMN. The origin of the z-coordinate (z_{FMN} = 0) corresponds to the center of the AsLOV2 domain (Figure S1). Snapshots at z_{FMN} = 7, 11, 15, 19 and 23 Å are shown. The position of FMN in the crystal structure (PDB ID = 2v0u) corresponds to z_{FMN} ≈ 7.8 Å. (B) and (C) far- and near-UV circular dichroism spectra of the AsLOV2 variants expressed as the mean residue ellipticity and molar ellipticity, respectively. (D) Thermal denaturation of AsLOV2 monitored by FMN fluorescence. The symbols and lines correspond to experimental data and fits according to Equation (1), respectively. (E) DSC thermograms of AsLOV2 variants. AsLOV2 variants are colored as follows: WT (black), V416C (blue), T418C (red), and V416C/T418C (green) (left box).

The mutations in positions 416 and 418 induced only small changes in the secondary structure of AsLOV2, as shown by CD spectra in the far-UV region. In two variants, V416C and T418C, the ellipticity in the far-UV region at 220 nm, reflecting a contribution of the α-helical conformation, slightly increased by ~20% and 15%, respectively. By contrast, in the double mutant, V416C/T418C, the ellipticity decreased by ~15% at 220 nm (Figure 2B).

In turn, the near-UV spectra of AsLOV2 wt, V416C, and T418C significantly overlap, indicating unperturbed tertiary structure around the aromatic amino acid residues (Figure 2C) (Kelly et al., 2005). The double mutant showed a slightly decreased signal in the near-UV region, which may either reflect small changes in the tertiary structure of the protein as indicated by decreased thermal stability of the double mutant and FMN fluorescence analysis.

The thermal denaturation of the proteins was analyzed by monitoring the variation of flavin fluorescence as a function of temperature to assess protein stability upon flavin binding (Figure 2D) and by DSC analysis to assess the global stability of the tertiary conformation (Figure 2E). Both approaches indicate the thermal stabilities of the single variants decreased similarly and only slightly (< 4°C), on both local and global levels, indicating unperturbed cooperativity interactions in the protein variants. The double variant transition temperature decreased by nearly 8°C, as determined by both methods, suggesting moderate destabilization

of the protein structure. Nevertheless, the cooperativity of the transition was not perturbed. Ultimately, we measured the following transition/melting temperatures (CD/DSC) of the AsLOV2: WT (57.2°C/59.3°C), V416C (53.4°C/54.9°C), T418C (56.4°C/60.6°C), and V416C/T418C (47.7°C/53.4°C).

FMN triplet state lifetimes. The analysis of the transient absorption data showed double exponential curves describing the decay of the absorption signal, except for the T418C variant of AsLOV2, in line with our previous findings for the wild type of AsLOV2 (Petrencakova et al., 2020). Based on these results, two populations of FMN were identified, one located inside the protein matrix and the other in the solvent. For the T418C variant, transient absorption was adequately represented by a single exponential curve, suggesting either one population of FMN or, more realistically, two populations with the same or very similar lifetimes. The experimental data were fitted by the following equation:

$$Absorbance = A_1 e^{-\frac{t}{\tau_{FMNin}}} + A_2 e^{-\frac{t}{\tau_{FMNout}}} \quad (5)$$

For T418C, we only used one term of Equation 5. The experimental data and the corresponding fits are shown in Figure S5. The resulting fits reproduce the measured data well. Based on our previous study (Petrencakova et al., 2020), the amplitudes of the FMN inside the proteins should decrease and the amplitudes of the free FMN in the solvent should increase with the irradiation time. These correlations were observed in all AsLOV2 variants, except for T418C because only one amplitude was assessed for this variant. However, the amplitude of T418C remained constant throughout the experiment (Figure S6). As shown by the lifetimes of individual triplet states listed in Table S1 (Figure S7) and the singlet oxygen lifetime (see below), the free FMN lifetime values of all variants were close to 2.7 μs, which is the lifetime of free FMN in air-saturated water at 37 °C (Westberg et al., 2017a,b). These results support the hypothesis that FMN dissociates from the protein matrix upon irradiation.

The lifetime of FMN inside the WT protein was similar to that reported in our previous study (1.57 μs) (Petrencakova et al., 2020). Conversely, the corresponding value of the V416C variant was half of the WT value. This finding indicates that substitution with cysteine provides an efficient pathway to quench the FMN triplet state or to produce an adduct similar to C450.

For the T418C variant, we assessed either a single triplet state lifetime or two with almost identical values. One way or the other, the cysteine at position 418 of the amino acid sequence likely affects the FMN position in the protein matrix, preventing C450 from creating a covalent

adduct with FMN. Therefore, substitution with cysteine increases the efficiency of singlet oxygen production (see below) by promoting the dissociation of FMN to the surrounding solvent.

For V416C/T418C, we found practically the same lifetime as for WT. This suggests that the effects of the V416C and T418C mutations somehow cancel each other. Based on the FMN triplet state lifetime values, either the electron is transferred from FMN to one of the cysteines or the competition between interactions is not detectable.

Singlet oxygen phosphorescence measurements. Singlet oxygen phosphorescence kinetics at different stages of irradiation are shown in Figure 3. The time between individual curves was 12.5 minutes. The data were fitted using the following equation (Jimenez-Banzo et al., 2008; Ragas et al., 2013):

$$P(t) = A_{in}\frac{\tau_{\Delta in}\tau_{FMNin}}{\tau_{\Delta in}-\tau_{FMNin}}\left(e^{-\frac{t}{\tau_{FMNin}}}-e^{-\frac{t}{\tau_{\Delta in}}}\right)+A_{out}\frac{\tau_{\Delta out}\tau_{FMNout}}{\tau_{\Delta out}-\tau_{FMNout}}\left(e^{-\frac{t}{\tau_{FMNout}}}-e^{-\frac{t}{\tau_{\Delta out}}}\right) \quad (6)$$

Even though it is possible to fit the phosphorescence kinetics just with the second term of the Equation 6, like it was shown in our previous study (Petrencakova et al., 2020), both FMN populations contribute to singlet oxygen phosphorescence. To minimize the number of free parameters in Equation 6, the lifetimes of individual FMN populations were set to the lifetimes obtained from transient absorption measurements and the value of $\tau_{\Delta out}$ to was set to be 3.5 μs, which corresponds to the singlet oxygen lifetime in pure water (Westberg et al., 2017a,b). The corresponding fits in Figure 3 reproduce the measured kinetics. The resulting singlet oxygen lifetimes measured at different stages of the experiment are listed in Table S1. In each protein variant, the effective lifetime $\tau_{\Delta in}$ increased towards the singlet oxygen lifetime in pure water (3.5 μs) (Westberg et al., 2017a,b). It is important to note, that $\tau_{\Delta in}$ represents an effective lifetime value, which reflects the changes due to gradual oxidation of singlet oxygen quenchers (i.e. amino acids) and also the diffusion of singlet oxygen out of the protein matrix.

Singlet oxygen phosphorescence increased over time, as shown in the kinetics curves (Figure 3). This result supports our assumption, based on our recent work (Petrencakova et al., 2020), that FMN dissociates to the solvent upon irradiation, producing singlet oxygen. The wild type and V416C variants displayed similar trends. After the first round of irradiation, a significant portion of FMN was already dissociated from V416C/T418C and found in the solvent.

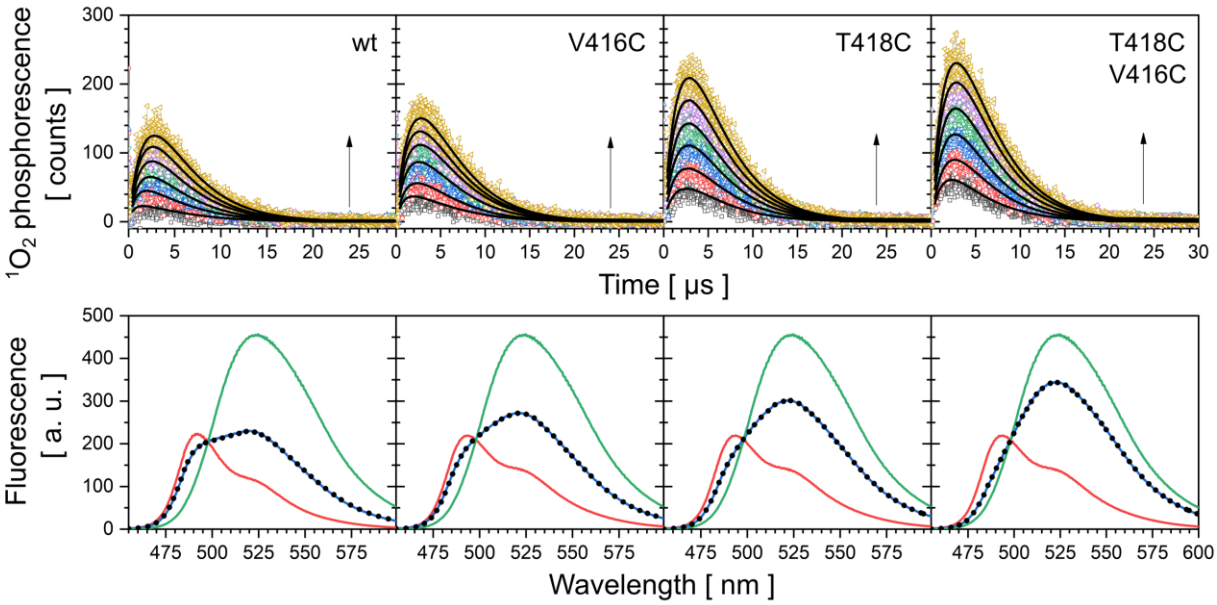


Figure 3. Upper row – time-resolved singlet oxygen phosphorescence of the AsLOV2 variants, with the same scaled y-axis for all plots. The color scheme represent the following accumulated incident energy: black open squares 0.2 J, red open circles 0.8 J, blue triangles 1.4 J, green reverse triangles 2 J, purple diamonds 2.6 J, and yellow ochre left triangles 3.2 J. The black lines correspond to fits by Equation 6. Lower row – Fluorescence spectra of the AsLOV2 variants before irradiation (red curves), free FMN (green curves) and AsLOV2 variants after irradiation (black dotes). The blue line is the linear combination of the red and green curves.

The amount of FMN released into the solvent upon irradiation was assessed by FMN fluorescence. Because FMN fluorescence spectra after irradiation can be linearly combined with the corresponding spectra of the protein before irradiation (red color) and with the fluorescence spectrum of free FMN (green color) (Petrencakova et al., 2020), we obtained closely correlative fits of the fluorescence spectra of the protein variants after irradiation, as shown in Figure 3 (lower row). The close fits (blue lines) and the presence of isosbestic points in the fluorescence spectra of the studied AsLOV2 systems validate our approach.

Based on this analysis, we were able to order the AsLOV2 variants by their efficiency of FMN release to the solvent upon 75 min irradiation (expressed as a percentage in brackets) as follows: WT(33%) < T416C(41%) < T418C(51%) < V416C/T418C(65%). The percentage of FMN dissociation from AsLOV2 wt in this study was lower than in our previous study (Petrencakova et al., 2020). This difference may be related to minor beam diameter adjustment in comparison with the previous experiments (Petrencakova et al., 2020).

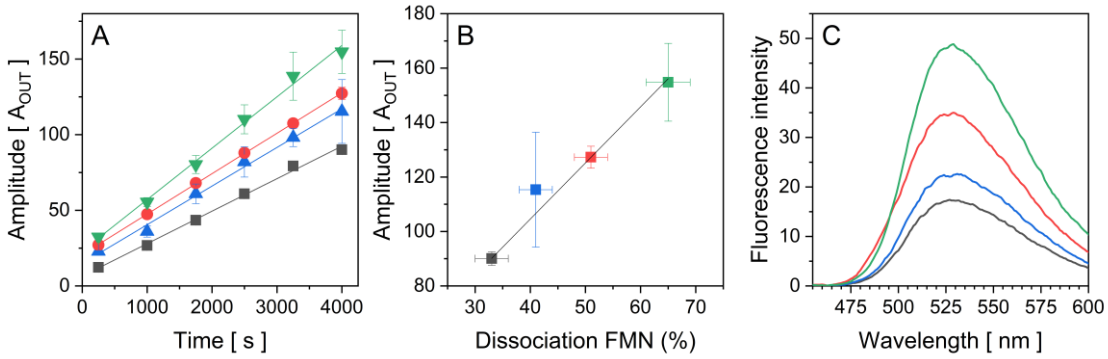


Figure 4. (A) Variation of amplitude [A_{out}] derived from Equation 6 as a function of time, reflecting the $^1O_2$ phosphorescence of each AsLOV2 variant shown in Figure 3 (upper row): wt (black squares), V416C (blue triangles), T418C (red circles), and V416C/T418C (green reverse triangles). (B) Correlation between the amplitude [A_{out}] derived from equation 6 after the last round of irradiation and FMN dissociation, as calculated from Figure 3. The correlation is described by linear equation y= 2.06x+ 22.12 with R^2= 0.998. (C) Fluorescence intensity of the irradiation-released FMN of the corresponding AsLOV2 after filtration experiment. No fluorescence was present in the filtrates of AsLOV2 wt and its variants before irradiation.

The results shown in Figure 3 were further analyzed by plotting the curves of the amplitudes obtained from Equation 6, corresponding to the amount of $^1O_2$ production as a function of time (Figure 4A and Figure S8). Assuming that most $^1O_2$ is produced only by free FMN (Petrencakova et al., 2020), the slopes provide the relative rates of FMN dissociation to the solvent. Accordingly, the slopes can be ordered as follows: wt < V416C < T418C < V416C/T418C. Variant V416C/T418C releases FMN faster than any other variant (Figures 3 and 4). In Figure S8 the irradiation dependence of [A_{in}] is also visible for each protein variant. These changes (minor increase) might be attributed to changing quantum yield of the proteins as the proteins become gradually oxidized (Pimenta et al., 2013; Torra et al., 2019). As it is clearly indicated in absorbance spectra of AsLOV2 wt and its variants measured after irradiation (Figure S9), in our experimental setup, there is no evidence of phototransformation of FMN to lumichrome as it has been previously observed by Torra et al (Torra et al., 2019). For completeness, we note that the formation of lumichrome in AsLOV2 systems studied was

observed but only upon prolonged light irradiation suggesting the phototransformation of the free FMN.

The $^1O_2$ phosphorescence of the V416C/T418C variant indicates the most efficient FMN release. However, corresponding slope in Figure 4A is the lowest, suggesting the presence of free FMN after the first round of irradiation, is most likely due to ambient light exposure induced FMN dissociation before the experiment. This result underscores the high sensitivity of the variant V416C/T418C to irradiation.

The phosphorescence values derived from Figure 3 (upper row) were also plotted as the highest amplitude of the $^1O_2$ phosphorescence as a function of the percentage of FMN dissociation from the corresponding proteins (Figure 4B). The resulting plot showed a strong correlation, suggesting a causal relationship. Combined, these findings support our model based on $^1O_2$ production by free FMN.

To assess the relative extent of FMN dissociation induced by the light irradiation, we performed filtration experiments as in our previous report (Petrencakova et al., 2020). In these experiments, the released FMN passes through the filter, while the protein is retained. Obtained results show that the relative release of FMN correspond to the presented results from phosphorescence and fluorescence experiments (Figures 3 and 4A, B).

Oxidative modification of the AsLOV2 variants. FMN release is accompanied by $^1O_2$ production and modification/oxidation of solvent-accessible protein residues, including residues in the dissociation channel, as it has been previously shown by mass spectrometry (Petrencakova et al., 2020). Halved samples were first analyzed using the top-down approach. Before and after irradiation (Figure S10), desalted protein samples were sprayed via an ESI source in broadband m/z mode to confirm the mass and thus the site-specific mutation, and to show the homogeneity of the samples, meaning that the site-specific mutations had no further effect on protein charge for ESI spraying. The zoom on 21+ charge state (Figure S11) shows predominant N-terminal methylation (+14.0156 amu) even before irradiation (grey spectra). After irradiation, small but, nevertheless, distinct changes in oxidation pattern appeared, thereby confirming $^1O_2$ oxidation of protein residues. Oxidized clusters were isolated and fragmented in multi-CASI mode by CID generating complementary b- and y-fragment ions. The extent of their oxidation was calculated for singly and doubly oxidized fragment ions (Figure 5, Figure S12, Table S2), and 7 and 11 b- and y-ions were simultaneously found in MSMS replicated spectra, respectively. We omitted the b-ions from the final interpretation for

the reason that N-terminal methionine is extensively methylated (+14.0156 amu) and thus biases with the extent the oxidation of b-type fragment ions (+15.9949 amu).

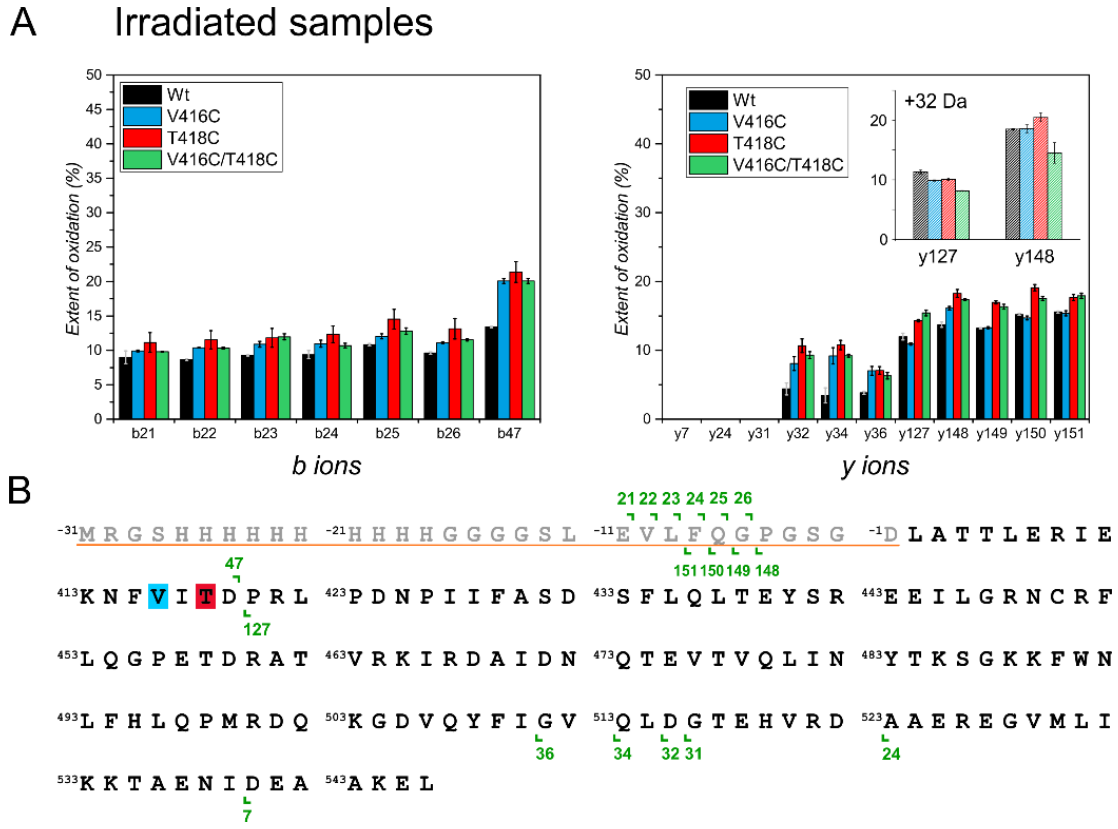


Figure 5. (A) The results of the extent of oxidation of selected b and y fragment ions of irradiated samples are expressed as the mean ± SD of three independent measurements. The inset in the y-ions plot indicates the extent of doubly oxidized y127 and y148 fragment ions. (B) The sequence of wild type AsLOV2 protein with the denoted fragment ions is displayed at the top of the figure. The His-tag, which is not part of the AsLOV2 sequence, is colored in grey and underlined in red. The positions of the mutations V416C and T418C are shown in blue and red squares, respectively.

In the y-ions, the first oxidized fragment is y32, which covers and pinpoints the oxidation to His519 residue. Fragment ion y31 also covers His519 but was not observed as oxidized. We manually validated spectra and found that unmodified y31 fragment ion was just at the level of noise that is why its oxidized form was below the level of detection. The next major increase of oxidation was observed between y36 and y127, covering P420 - I510 region. The oxidation of residues is significant in this region (Figure 5, Table S2), with the highest extent observed in V416/T418 mutant. This is in accordance with the DSC studies where this mutant displays moderate destabilization of the protein structure, which likely leads to an exposing some residues in this region to solvent and theire subsequent oxidation. Another increase in an

oxidation extent can be observed between y127 and y148 fragment ions, covering both mutations and F415 located inside the FMN binding pocket. As can be seen from Figure 5 and Table S2, single and double oxidations follow the increasing trend WT ~ V416C/T418C < V416C ~ T418C. The overall extent of oxidation between y127-y148 fragment ions is driven by F415, C416 and C418 residues for all proteins.

The second half of the samples were analyzed using the bottom-up approach. Samples were in-solution digested by trypsin for a short time to prevent oxidation during long incubation times.

Table 1. List of selected oxidation products and their extent of oxidation.

Residue	Modification[a]	The extent of modification[b], %			
		WT	V416C	T418C	V416C/T418C
F415	1O	20.37	10.17	12.38	9.22
V416C	3O		21.61		5.04
T418C	3O			0.80	0.38
C450	3O	3.42	3.35	1.71	1.50

[a] - data represent values with subtracted background oxidation level ($S_{irradiated}$-$S_{control}$).
[b] - single modifications were identified using ion mobility.

The peptides were then identified by a searching engine, including their singly, doubly, and triply oxidized forms. Out of identified peptides and residues, several residues throughout the sequence were observed as oxidized but at a low and nonsignificant extent, mostly <1 %. Nevertheless, peptides containing embedded cysteine mutations were identified and their extent of oxidation was calculated for all three mutated variants. We also identified two peptides with the sequence ^{443}EEILGRNCR and ^{449}NCRFLQGPETDR containing C450. However, generating two peptides during Trypsin digestion resulted in a decreased peptide signal in LCMS trace, and the calculated extent of oxidation, which differs in both peptides, was observed just above the level of detection. Thus, the calculated extent of oxidation for C450 in Table 1 is a weighted arithmetic average of the extents from both peptides. The extent of modification was then calculated for wt and all three variants (Table 1). The bottom-up dataset is in general in high agreement with our top-down dataset. As can be seen from Table 1, the F415 and C450, located near to and inside the binding pocket, respectively, were both observed as oxidized. When cysteine mutation is introduced inside the binding pocket, the extent of F415 oxidation decreased as cysteines, more sensitive to $^1O_2$, were observed predominantly as triply oxidized. However, the effect of mutation results in different outcomes for particular mutation. While the V416C oxidation leads to FMN stabilization and the C416

scavenges $^1O_2$, the T418C and V416C/T418C mutations more likely destabilize the FMN binding, which is supported by protecting the F415 from oxidation and not observing compensation of oxidation on C450 and mutated residues (Table 1).

Based on our MS data, we reached the following conclusion: (i) the modifications of AsLOV2 variants are light-dependent as the overall extent of the oxidative modifications is much higher in irradiated samples of all AsLOV2 variants than in non-irradiated samples (Figure S10, Figure S11), (ii) the same amino acids are oxidized in all AsLOV2 variants (Figures 5 and S12), albeit to a different extent, and (iii) the residues are oxidized at a different extent due to the conformational changes and mild protein destabilization, observable mostly on both T418C and V416C/T418C protein variant. Conclusions (i) and (ii) strongly indicate that the AsLOV2 variants share similar conformational properties, and conclusion (iii) supports more efficient FMN dissociation from the AsLOV2 variants.

Because FMN is less efficiently dissociated from AsLOV2 C450A than from AsLOV2 wt, the oxidation damage is larger in the C450 variant than in the WT protein (Petrencakova et al., 2020). The extent of oxidation (both singly and doubly) between y127 and y148 ions, both of which contain extra cysteine(s) in position(s) 416 and 418, increased similarly in WT, V416C and T418C, but not in the V416C/T418C double mutant (Figure 5, Table S2). This suggests that FMN is released more quickly from the double mutant than from the single variants and produce $^1O_2$ in the solvent, in consent with the fluorescence spectra shown in Figure 3.

Conclusions

Our approach to the design of efficient flavin-based GEPS system relies on controlled cofactor dissociation as a direct result of irradiation at a suitable wavelength. In our particular case, it depends on a natural propensity of flavoprotein AsLOV2 to bind highly efficient PS represented by FMN, which upon blue light irradiation oxidizes close $^1O_2$ prone oxidation cysteinyl residue(s) and this chemical modification triggers its dissociation. Assuming tight encapsulation of the isoalloxazine ring in its binding site, the design method is based on an identification of amino acid positions occupied by nonconservative amino acids with small side residues near the isoalloxazine moiety. Replacement of the small amino acid by amino acid, which becomes upon $^1O_2$ oxidation larger than the original one leads to a compensatory steric effect, accompanied by cofactor release. Applying this approach, we were able to increase the Φ_Δ values in all prepared AsLOV2 variant systems. In particular, the Φ_Δ value of AsLOV2

V416C/T418C system after irradiation is ~16 times higher than that of AsLOV2 wt system before irradiation without significantly compromising protein stability and FMN binding affinity. The presented proof-of-principle study demonstrates that this approach is a viable alternative to the traditional design of the efficient GEPSs. Our GEPS system fulfills two critical properties of GEPS: (i) ability to target PS by suitable tags and (ii) increased Φ_Δ value.

The potential limitations of our GEPS system is related to the limited spatial specificity due to a diffusion of the irradiation-triggered released of PS. Spatial specificity is very critical for the correlative light electron microscopy (CLEM) and photodynamic diagnosis (PDD). As such, our AsLOV2 system is unlikely suitable for use in CLEM and PDD applications. However, we believe that AsLOV2 system might be useful in chromophore-assisted light inactivation (CALI) of proteins and photoablation of cells as discussed for example in Suslova et al. (Suslova et al., 2017). In fact, our AsLOV2 system is related to currently developing drug carrier systems for PDT based on liposomes and polymer nanoparticles, which release PS loads at or in the target cells due to a changed a physico-chemical property of solution environment such as pH, ionic strength or temperature (Debele et al., 2015; Akasov et al., 2022). In the case of AsLOV2 system, the spatial specificity of FMN action is determined by the applied light as only the incident light triggers the release of PS and a production of $^1O_2$. The efficiency as well as the mechanism of action of the GEPS system proposed in this work needs to be determined in an analogous study as the one very recently performed by Mogensen et al. (Mogensen et al., 2022).

Associated content

Supporting Information. Figure S1. A schematic view of the complex of the AsLOV2 domain and FMN. Figure S2 and S3. The top and side views on AsLOV2 variants and WT, respectively. Figure S4. MD simulation structures of AsLOV2 variants. Figure S5. Transient absorption measurements. Figure S6. Time evolution of FMN triplet states amplitudes. Figure S7. Time evolution of ^{3}FMN amplitudes. Figure S8. Time evolution of FMN amplitudes related to singlet oxygen production. Figure S9. Absorption spectra of ASLOV2 wt and its variants. Figure S10. Broadband ESI-MS spectra of ASLOV2 wt and its variants. Figure S11. A zoom into the ESI-MS spectra 21+ charge state. Figure S12. The sequence of wild type AsLOV2 protein with the denoted fragment ions. Table S1. The ^{3}FMN and singlet oxygen lifetimes at different stages of the irradiation experiment. Table S2. The extent of oxidation for b- and y-

fragment ions generated by multiCASI-CID. This material is available free of charge via the Internet at http://pubs.acs.org.

Author Contributions

The manuscript was written with the contribution of all authors. / All authors have given approval to the final version of the manuscript. / ‡These authors contributed equally.

Acknowledgment

This research was funded by the Slovak Research and Development Agency (project APVV-20-0340), the Program EXCELES of National Plan of recovery and resilience Ministry of Education, Youth and Sports of the Czech republic (NPO-NEURO-EXCELLES grant LX22NPO5107), the Grant Agency of Charles University (grant 359521), and the Czech Academy of Sciences (RVO61388971). Access to instruments for mass spectrometry analysis was granted through the EU_FT-ICR_MS network (funded by the EU Horizon 2020, grant agreement ID: 731077) and the Ministry of Education, Youth and Sports of the Czech Republic (Structural Mass Spectrometry CF - LM2018127 CIISB). We thank Carlos V. Melo for editing the manuscript.

References

Akasov, R., Khaydukov, E. V., Yamada, M., Zvyagin, A. V., Leelahavanichkul, A., Leanse, L. G., Dai, T., & Prow, T. (2022) Nanoparticle enhanced blue light therapy [Review]. *Adv. Drug. Deliv. Rev.*, 184, 114198. https://doi.org/10.1016/j.addr.2022.114198

Andersen, H. C. (1983). RATTLE - A Velocity Version of the SHAKE Algorithms for Molecular-Dynamics Calculations [Article]. *J. Comput. Phys.*, *52*(1), 24-34. https://doi.org/10.1016/0021-9991(83)90014-1

Baier, J., Maisch, T., Maier, M., Engel, E., Landthaler, M., & Baumler, W. (2006). Singlet oxygen generation by UVA light exposure of endogenous photosensitizers [Article]. *Biophys. J.*, *91*(4), 1452-1459. https://doi.org/10.1529/biophysj.106.082388

Bulina, M. E., Chudakov, D. M., Britanova, O. V., Yanushevich, Y. G., Staroverov, D. B., Chepurnykh, T. V., . . . Lukyanov, K. A. (2006). A genetically encoded photosensitizer [Article]. *Nat. Biotech.*, *24*(1), 95-99. https://doi.org/10.1038/nbt1175

Darden, T., York, D., & Pedersen, L. (1993). Particle Mesh Ewald - An N.log(N) Method for Ewald Sums in Large Systems [Note]. *J. Chem. Phys.*, *98*(12), 10089-10092. https://doi.org/10.1063/1.464397

Debele, T. A., Peng, S., & Tsai, H. C. (2015) Drug Carrier for Photodynamic Cancer Therapy [Article]. *Int. J. Mol. Sci.*, 16(9) 22094-22136. https://doi.org/10.3390/ijms160922094

Dougherty, T. J., Gomer, C. J., Henderson, B. W., Jori, G., Kessel, D., Korbelik, M., . . . Peng, Q. (1998). Photodynamic therapy [Review]. *Jnci-Journal of the National Cancer Institute*, *90*(12), 889-905. https://doi.org/10.1093/jnci/90.12.889

Endres, S., Wingen, M., Torra, J., Ruiz-Gonzalez, R., Polen, T., Bosio, G., . . . Drepper, T. (2018). An optogenetic toolbox of LOV-based photosensitizers for light-driven killing of bacteria. *Sci. Rep.*, *8*, Article 15021. https://doi.org/10.1038/s41598-018-33291-4

Feller, S. E., Zhang, Y. H., Pastor, R. W., & Brooks, B. R. (1995). Constant-Pressure Molecular-Dynamics Simulation - The Langevin Piston Method [Article]. *J. Chem. Phys.*, *103*(11), 4613-4621. https://doi.org/10.1063/1.470648

Gorbachev, D. A., Staroverov, D. B., Lukyanov, K. A., & Sarkisyan, K. S. (2020). Genetically Encoded Red Photosensitizers with Enhanced Phototoxicity [Article]. *International Journal of Molecular Sciences*, *21*(22), 12, Article 8800. https://doi.org/10.3390/ijms21228800

Gunaydin, G., Gedik, M. E., & Ayan, S. (2021). Photodynamic Therapy for the Treatment and Diagnosis of Cancer-A Review of the Current Clinical Status [Review]. *Front. Chem.*, *9*, 26, Article 686303. https://doi.org/10.3389/fchem.2021.686303

Halavaty, A. S., & Moffat, K. (2007). N- and C-terminal flanking regions modulate light-induced signal transduction in the LOV2 domain of the blue light sensor phototropin 1 from *Avena sativa* [Article]. *Biochemistry*, *46*(49), 14001-14009. https://doi.org/10.1021/bi701543e

Hilgers, F., Bitzenhofer, N. L., Ackermann, Y., Burmeister, A., Grunberger, A., Jaeger, K. E., & Drepper, T. (2019). Genetically Encoded Photosensitizers as Light-Triggered Antimicrobial Agents [Article]. *Int. J. Mol. Sci.*, *20*(18), 20, Article 4608. https://doi.org/10.3390/ijms20184608

Hovan, A., Pevna, V., Huntosova, V., Miskovsky, P., & Bano, G. (2023). Singlet oxygen lifetime changes in dying glioblastoma cells [Article]. *Photochemistry and Photobiology*, *00*, 1-13. https://doi.org/10.1111/php.13828

Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B. L., . . . MacKerell, A. D. (2017). CHARMM36m: an improved force field for folded and intrinsically disordered proteins [Article]. *Nature Methods*, *14*(1), 71-73. https://doi.org/10.1038/nmeth.4067

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics [Article]. *J. Mol. Graph.*, *14*(1), 33-38. https://doi.org/10.1016/0263-7855(96)00018-5

Jarzynski, C. (1997). Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, *78*(14), 2690-2693. https://doi.org/10.1103/PhysRevLett.78.2690

Jimenez-Banzo, A., Ragas, X., Kapusta, P., & Nonell, S. (2008). Time-resolved methods in biophysics. 7. Photon counting vs. analog time-resolved singlet oxygen phosphorescence detection [Article]. *Photochem. Photobiol. Sci.*, *7*(9), 1003-1010. https://doi.org/10.1039/b804333g

Kelly, S. M., Jess, T. J., & Price, N. C. (2005). How to study proteins by circular dichroism [Review]. *Biochimica Et Biophysica Acta-Proteins and Proteomics*, *1751*(2), 119-139. https://doi.org/10.1016/j.bbapap.2005.06.005

Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D., & Nesvizhskii, A. I. (2017) MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics [Article]. *Nat. Methods.*, 14(5), 513-520. https://doi.org/10.1038/nmeth.4256

Lee, C. N., Hsu, R., Chen, H., & Wong, T. W. (2020). Daylight Photodynamic Therapy: An Update [Review]. *Molecules*, *25*(21), 16, Article 5195. https://doi.org/10.3390/molecules25215195

Loginov, D. S., Fiala, J., Brechlin, P., Kruppa, G., & Novak, P. (2022). Hydroxyl radical footprinting analysis of a human haptoglobin-hemoglobin complex [Article]. *Biochim. Biophys. Acta-Proteins and Proteomics*, *1870*(2), 6, Article 140735. https://doi.org/10.1016/j.bbapap.2021.140735

MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C., & MacCoss, M. J. (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*, 26(7), 966-968. https://doi.org/10.1093/bioinformatics/btq054

Maestro. In. (2022). 2022-2. Schrödinger, LLC.

Mansoori, B., Mohammadi, A., Doustvandi, M. A., Mohammadnejad, F., Kamari, F., Gjerstorff, M. F., . . . Hamblin, M. R. (2019). Photodynamic therapy for cancer: Role of natural products [Review]. *Photodiag. Photodyn. Therapy*, *26*, 395-404. https://doi.org/10.1016/j.pdpdt.2019.04.033

Martyna, G. J., Tobias, D. J., & Klein, M. L. (1994). Constant-Pressure Molecular-Dynamics Algorithms [Article]. *J. Chem. Phys.*, *101*(5), 4177-4189. https://doi.org/10.1063/1.467468

Meissner, B., Schleicher, E., Weber, S., & Essen, L. O. (2007). The dodecin from Thermus thermophilus, a bifunctional cofactor storage protein [Article]. *J. Biol. Chem.*, *282*(45), 33142-33154. https://doi.org/10.1074/jbc.M704951200

Micheletto, M. C., Guidelli, E. J., & Costa, A. J. (2021). Interaction of Genetically Encoded Photosensitizers with Scintillating Nanoparticles for X-ray Activated Photodynamic Therapy [Article]. *ACS Appl. Mater. Interf.*, *13*(2), 2289-2302. https://doi.org/10.1021/acsami.0c19041

Mogensen, D. J., Westberg M., Breitenbach T., Etzerodt M., & Ogilby P. R. (2021). Stable Transfection of the Singlet Oxygen Photosensitizing Protein SOPP3: Examining Aspects of Intracellular Behavior [Article]. *Photochem Photobiol.*, 97(6), 1417-1430. https://doi.org/10.1111/php.13440

O'Connor, A. E., Gallagher, W. M., & Byrne, A. T. (2009). Porphyrin and Nonporphyrin Photosensitizers in Oncology: Preclinical and Clinical Advances in Photodynamic Therapy [Review]. *Photochem. Photobiol.*, *85*(5), 1053-1074. https://doi.org/10.1111/j.1751-1097.2009.00585.x

Park, S., & Schulten, K. (2004). Calculating potentials of mean force from steered molecular dynamics simulations [Article]. *J. Chem. Phys.*, *120*(13), 5946-5961. https://doi.org/10.1063/1.1651473

Petrencakova, M., Filandr, F., Hovan, A., Yassaghi, G., Man, P., Kozar, T., . . . Sedlak, E. (2020). Photoinduced damage of AsLOV2 domain is accompanied by increased singlet oxygen production due to flavin dissociation [Article]. *Sci. Rep.*, *10*(1), 15, Article 4119. https://doi.org/10.1038/s41598-020-60861-2

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., . . . Schulten, K. (2005). Scalable molecular dynamics with NAMD [Review]. *J. Comput. Chem.*, *26*(16), 1781-1802. https://doi.org/10.1002/jcc.20289

Pimenta, F. M., Jensen, R. L., Breitenbach, T., Etzerodt, M., & Ogilby, P. R. (2013). Oxygen-Dependent Photochemistry and Photophysics of "MiniSOG," a Protein-Encased Flavin. *Photochem. Photobiol.*, *89*(5), 1116-1126. https://doi.org/10.1111/php.12111

Polak, M., Yassaghi, G., Kavan, D., Filandr, F., Fiala, J., Kukacka, Z., . . . Novak, P. (2022). Utilization of Fast Photochemical Oxidation of Proteins and Both Bottom-up and Top-down Mass Spectrometry for Structural Characterization of a Transcription Factor-dsDNA Complex [Article]. *Anal. Chem.*, *94*(7), 3203-3210. https://doi.org/10.1021/acs.analchem.1c04746

Price, D. J., & Brooks, C. L. (2004). A modified TIP3P water potential for simulation with Ewald summation [Article]. *J. Chem. Phys.*, *121*(20), 10096-10103. https://doi.org/10.1063/1.1808117

Proshkina, G. M., Shilova, O. N., Ryabova, A. V., Stremovskiy, O. A., & Deyev, S. M. (2015). A new anticancer toxin based on HER2/neu-specific DARPin. and photoactive flavoprotein miniSOG [Article]. *Biochimie*, *118*, 116-122. https://doi.org/10.1016/j.biochi.2015.08.013

Ragas, X., Cooper, L. P., White, J. H., Nonell, S., & Flors, C. (2011). Quantification of Photosensitized Singlet Oxygen Production by a Fluorescent Protein [Article]. *Chemphyschem*, *12*(1), 161-165. https://doi.org/10.1002/cphc.201000919

Ragas, X., He, X., Agut, M., Roxo-Rosa, M., Gonsalves, A. R., Serra, A. C., & Nonell, S. (2013). Singlet Oxygen in Antimicrobial Photodynamic Therapy: Photosensitizer-Dependent Production and Decay in *E. coli* [Article]. *Molecules*, *18*(3), 2712-2725. https://doi.org/10.3390/molecules18032712

Rodríguez-Pulido, A., Cortajarena, A.L., Torra, J., Ruiz-González, R., Nonell, S., & Flors, C. (2016) Assessing the potential of photosensitizing flavoproteins as tags for correlative microscopy [Article]. *Chem. Commun.*, 52, 8405-8408. http://dx.doi.org/10.1039/C6CC03119F

Ruiz-Gonzalez, R., Cortajarena, A. L., Mejias, S. H., Agut, M., Nonell, S., & Flors, C. (2013). Singlet Oxygen Generation by the Genetically Encoded Tag miniSOG [Article]. *J. Am. Chem. Soc.*, *135*(26), 9564-9567. https://doi.org/10.1021/ja4020524

Sarkisyan, K. S., Zlobovskaya, O. A., Gorbachev, D. A., Bozhanova, N. G., Sharonov, G. V., Staroverov, D. B., . . . Lukyanov, K. A. (2015). KillerOrange, a Genetically Encoded Photosensitizer Activated by Blue and Green Light [Article]. *Plos One*, *10*(12), 11, Article e0145287. https://doi.org/10.1371/journal.pone.0145287

Schrödinger, L. (2022). *Schrödinger*. In (Version Release 2022-2)

Serebrovskaya, E. O., Edelweiss, E. F., Stremovskiy, O. A., Lukyanov, K. A., Chudakov, D. M., & Deyev, S. M. (2009). Targeting cancer cells by using an antireceptor antibody-photosensitizer fusion protein [Article]. *Proc. Nat. Acad. Sci. U.S.A.*, *106*(23), 9221-9225. https://doi.org/10.1073/pnas.0904140106

Shaw, D. E. (2020). *Desmond Molecular Dynamics System*. In D. E. Shaw.

Shu, X., Lev-Ram, V., Deerinck, T. J., Qi, Y., Ramko, E. B., Davidson, M. W., Jin, Y., Ellisman, M. H., & Tsien, R. Y. (2011). A genetically encoded tag for correlated light and electron microscopy of intact cells, tissues, and organisms [Article]. *PLoS Biol.*, 9(4), e1001041. https://doi.org/10.1371/journal.pbio.1001041

Souslova, E. A., Mironova, K. E., & Deyev, S. M. (2017). Applications of genetically encoded photosensitizer miniSOG: from correlative light electron microscopy to immunophotosensitizing [Review]. *J. Biophotonics*, *10*(3), 338-352. https://doi.org/10.1002/jbio.201600120

Stockwell, G. R., & Thornton, J. M. (2006). Conformational diversity of ligands bound to proteins [Article]. *Journal of Molecular Biology*, *356*(4), 928-944. https://doi.org/10.1016/j.jmb.2005.12.012

Takemoto, K., Matsuda, T., Sakai, N., Fu, D., Noda, M., Uchiyama, S., . . . Nagai, T. (2013). SuperNova, a monomeric photosensitizing fluorescent protein for chromophore-assisted light inactivation [Article]. *Scientific Reports*, *3*, 7, Article 2629. https://doi.org/10.1038/srep02629

Taylor, B. L., & Zhulin, I. B. (1999). PAS domains: Internal sensors of oxygen, redox potential, and light [Review]. *Microbiol. Mol. Biol. Rev.*, *63*(2), 479-506. https://doi.org/10.1128/mmbr.63.2.479-506.1999

Torra, J., Lafaye, C., Signor, L., Aumonier, S., Flors, C., Shu, X. K., . . . Royant, A. (2019). Tailing miniSOG: structural bases of the complex photophysics of a flavin-binding singlet oxygen photosensitizing protein. *Sci. Rep.*, *9*, Article 2428. https://doi.org/10.1038/s41598-019-38955-3

van den Berg, P. A. W., Feenstra, K. A., Mark, A. E., Berendsen, H. J. C., & Visser, A. (2002). Dynamic conformations of flavin adenine dinucleotide: Simulated molecular dynamics of the flavin cofactor related to the time-resolved fluorescence characteristics [Article]. *J. Phys. Chem. B*, *106*(34), 8858-8869. https://doi.org/10.1021/jp020356s

Westberg, M., Bregnhoj, M., Etzerodt, M., & Ogilby, P. R. (2017a). No Photon Wasted: An Efficient and Selective Singlet Oxygen Photosensitizing Protein [Article]. *J. Phys. Chem. B*, *121*(40), 9366-9371. https://doi.org/10.1021/acs.jpcb.7b07831

Westberg, M., Bregnhoj, M., Etzerodt, M., & Ogilby, P. R. (2017b). Temperature Sensitive Singlet Oxygen Photosensitization by LOV-Derived Fluorescent Flavoproteins. *J. Phys. Chem. B*, *121*(12), 2561-2574. https://doi.org/10.1021/acs.jpcb.7b00561

Westberg, M., Etzerodt, M., & Ogilby, P. R. (2019). Rational design of genetically encoded singlet oxygen photosensitizing proteins [Article]. *Cur. Opin. Struct. Biol.*, *57*, 56-62. https://doi.org/10.1016/j.sbi.2019.01.025

Westberg, M., Holmegaard, L., Pimenta, F. M., Etzerodt, M., & Ogilby, P. R. (2015). Rational Design of an Efficient, Genetically Encodable, Protein-Encased Singlet Oxygen Photosensitizer [Article]. *J. Am. Chem. Soc.*, *137*(4), 1632-1642. https://doi.org/10.1021/ja511940j

Yassaghi, G., Kukacka, Z., Fiala, J., Kavan, D., Halada, P., Volny, M., & Novak, P. (2022). Top-Down Detection of Oxidative Protein Footprinting by Collision-Induced Dissociation,Electron-Transfer Dissociation, and Electron-Capture Dissociation [Article]. *Anal. Chem.*, *94*(28), 9993-10002. https://doi.org/10.1021/acs.analchem.1c05476

Yu, F., Haynes, S. E., Teo, G. C., Avtonomov, D. M., Polasky, D. A., & Nesvizhskii, A. I. (2020) Fast Quantitative Analysis of timsTOF PASEF Data with MSFragger and IonQuant [Article]. *Mol. Cell. Proteomics*, 19(9), 1575-1585. https://doi.org/10.1074/mcp.TIR120.002048

Zayner, J. P., Antoniou, C., & Sosnick, T. R. (2012). The amino-terminal helix modeulates lignt-activated conformational changes in AsLOV2 [Article]. *J. Mol. Biol.*, *419*(1-2), 61-74. https://doi.org/10.1016/j.jmb.2012.02.037

Zhang, J., Jiang, C. S., Longo, J. P. F., Azevedo, R. B., Zhang, H., & Muehlmann, L. A. (2018). An updated overview on the development of new photosensitizers for anticancer photodynamic therapy [Review]. *Acta Pharm. Sinica B*, *8*(2), 137-146. https://doi.org/10.1016/j.apsb.2017.09.003

Supporting information

to

Design of AsLOV2 domain as a carrier of light-induced dissociable FMN photosensitizer

Kristína Felčíková, Andrej Hovan, Marek Polák, Dmitry S. Loginov, Veronika Holotová, Carlos Díaz, Tibor Kožár, One-Sun Lee, Rastislav Varhač, Petr Novák, Gregor Bánó, and Erik Sedlák

SUPPORTING FIGURES

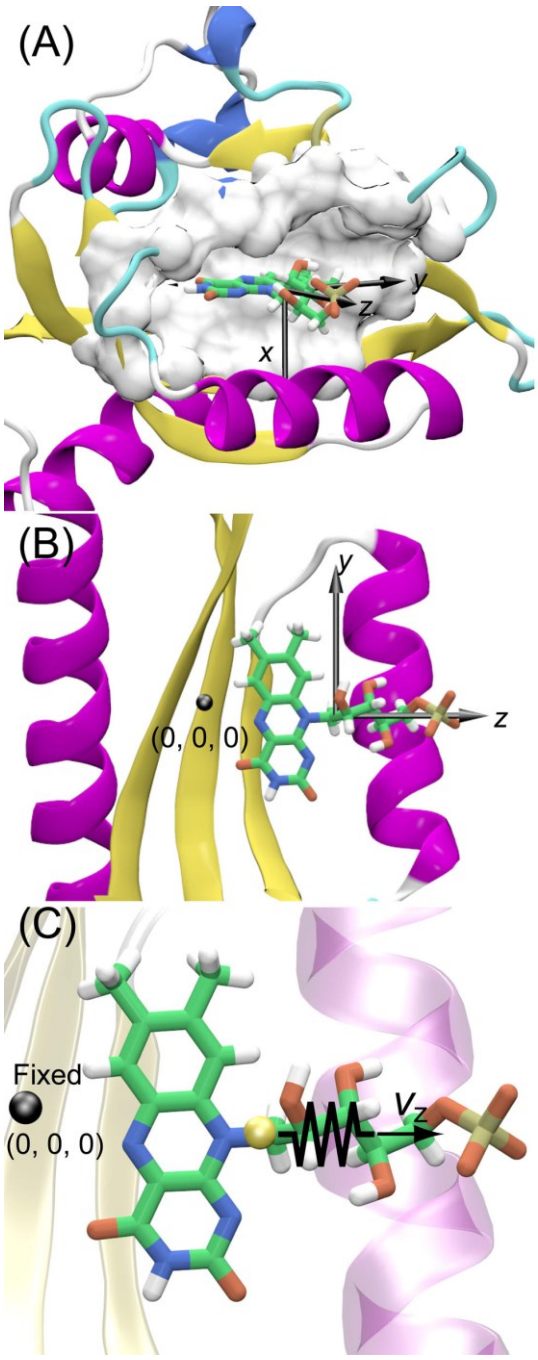


Figure S1. (A) A schematic view of the complex of the AsLOV2 wt domain and FMN. The principal axes of FMN are superimposed on the licorice model of FMN. (B) The center of AsLOV2 wt (black sphere) is set as (0,0,0) and the longest principal axis of FMN is taken as *z* coordinate. (C) A force is applied to the center of FMN along the z axis to steer the translation of FMN while the position of the center of AsLOV2 wt is fixed.

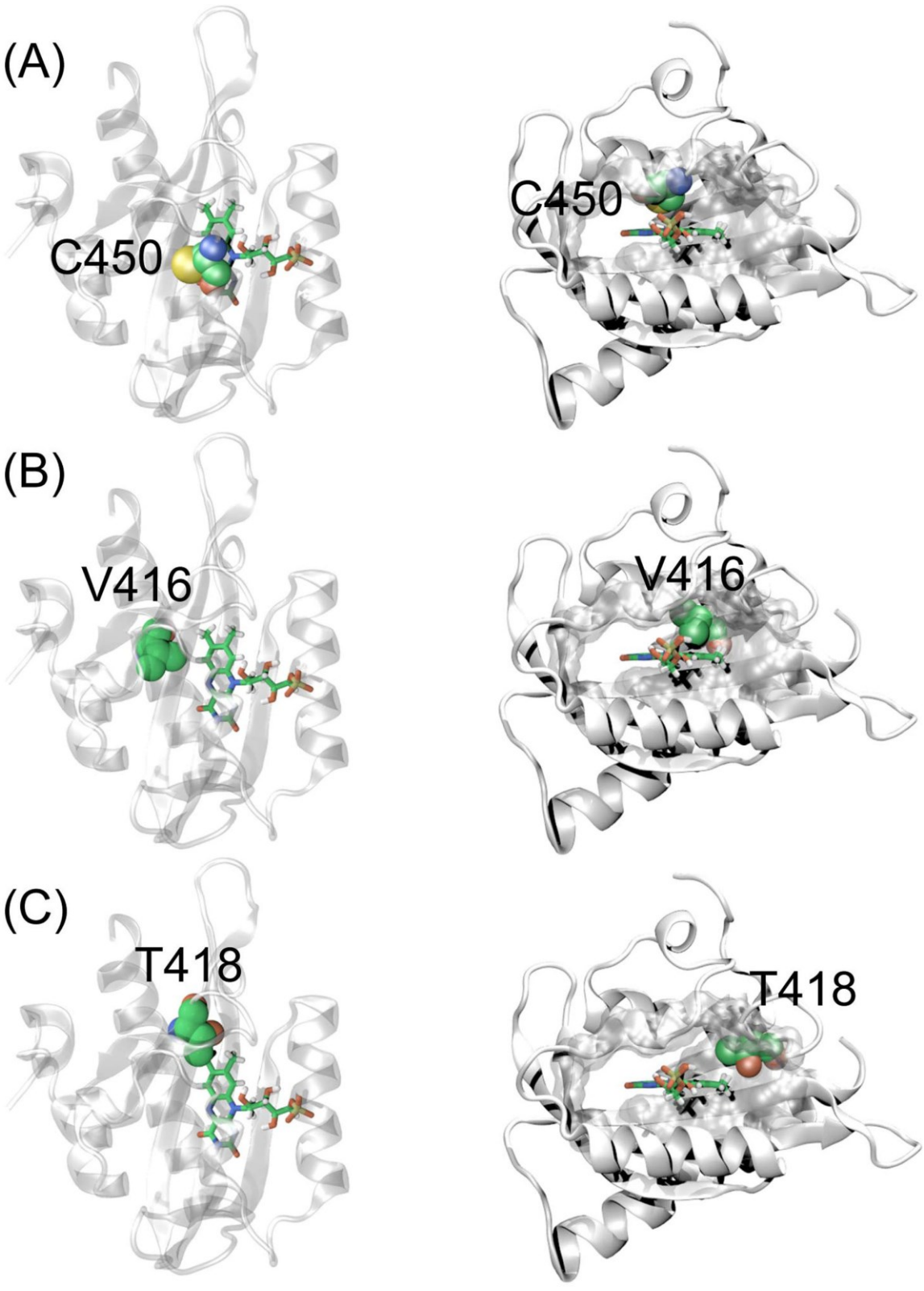


Figure S2. The top and side views for representing the positions of (A) C450 (B) V416 and (C) T418. Each residue is represented by the space filling model and FMN is represented by the licorice model. The structures of the AsLOV2 wt domain (cartoon model) and its cavity (surface model) are shown in grey.

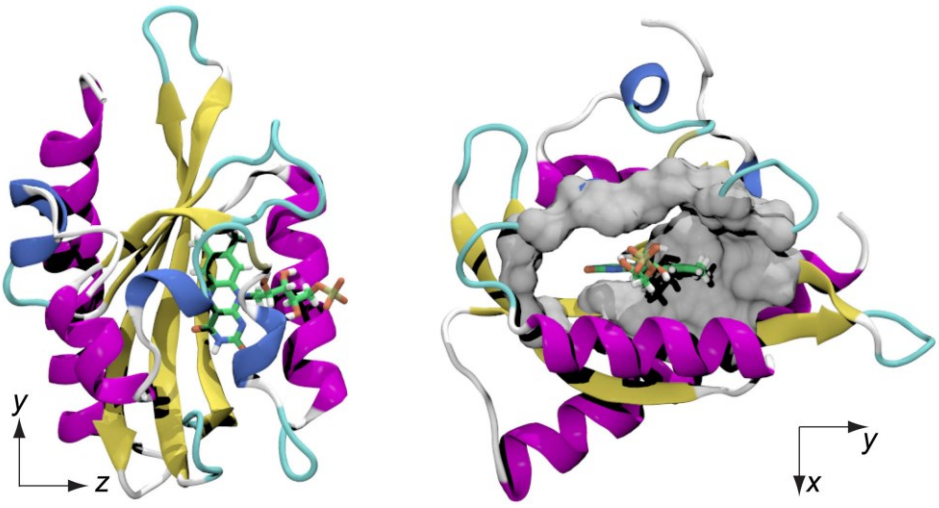


Figure S3. Top and side views of the complex of AsLOV2 wt domain and FMN (PDB ID: 2v0u). The Cartesian coordinates are based on the principal axes of FMN.

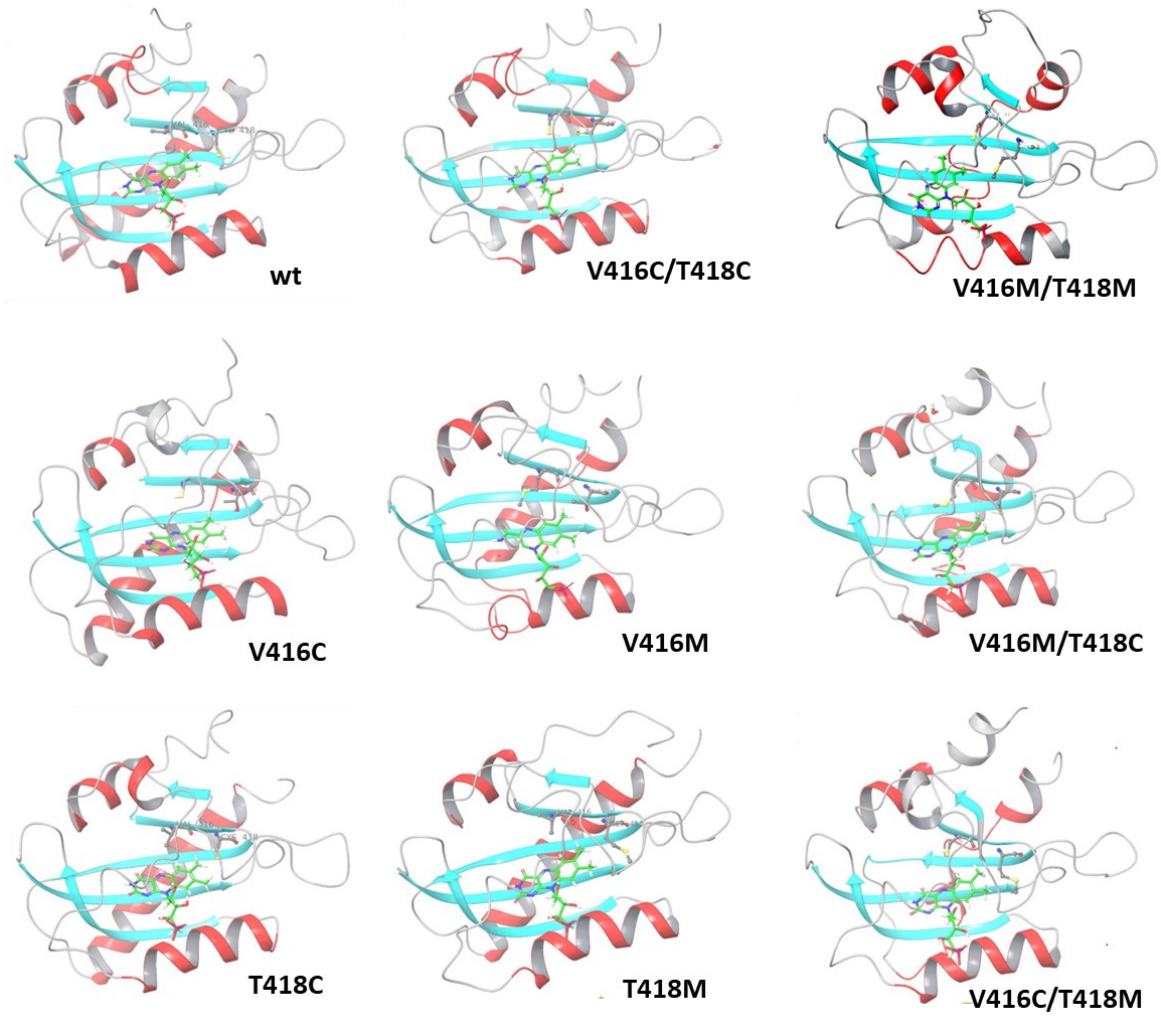


Figure S4. The structures of the corresponding AsLOV2 variants after 500 ns MD simulations. It is noticable that the variants containing methionyl residue(s), besides T418M, exhibit significantly perturbed conformations, particularly regarding their α helical secondary structures, in the comparison with the AsLOV2 wt.

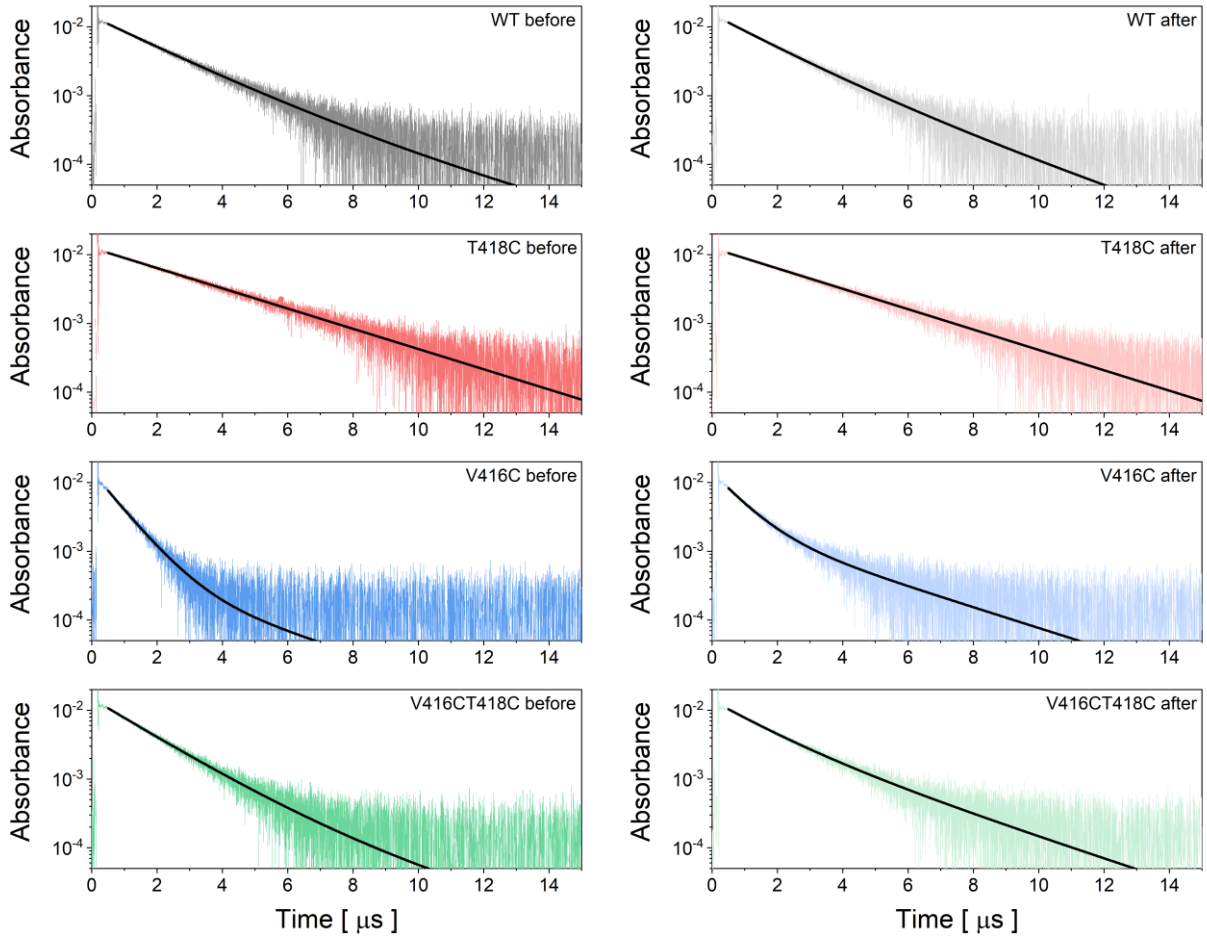


Figure S5. Transient absorption measurements after short irradiation time (left column) and after the whole irradiation experiment (right column). The solid lines represent fits obtained by the experimental data fitting by Equation 5.

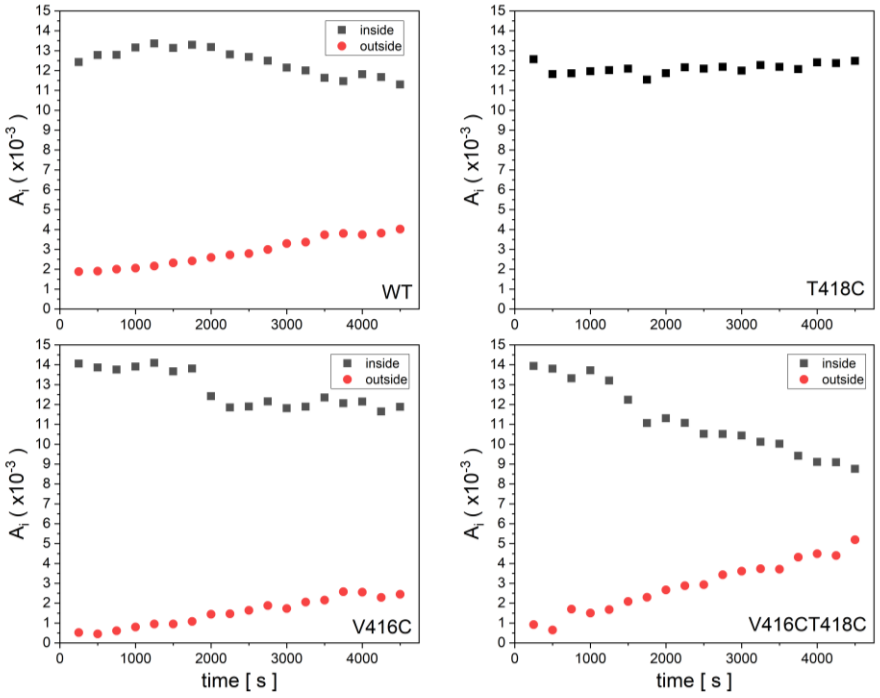


Figure S6. Time evolution of FMN triplet states amplitudes obtained from curve fitting of the experimental data for each AsLOV2 variant by Equation 5. Black squares represent the amplitudes of FMN population inside the protein matrix and red points represent the amplitudes of free FMN in solvent.

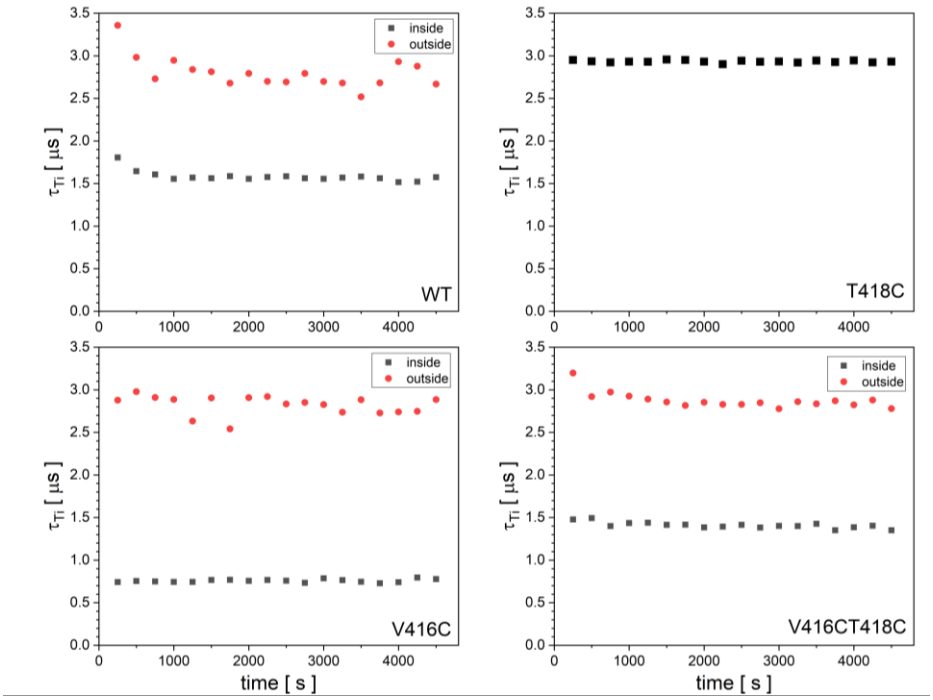


Figure S7. Time evolution of FMN triplet states lifetimes obtained from curve fitting of the experimental data for each AsLOV2 variant by Equation 5. Black squares represent the population of FMN inside the protein matrix and red points represent free FMN in solvent.

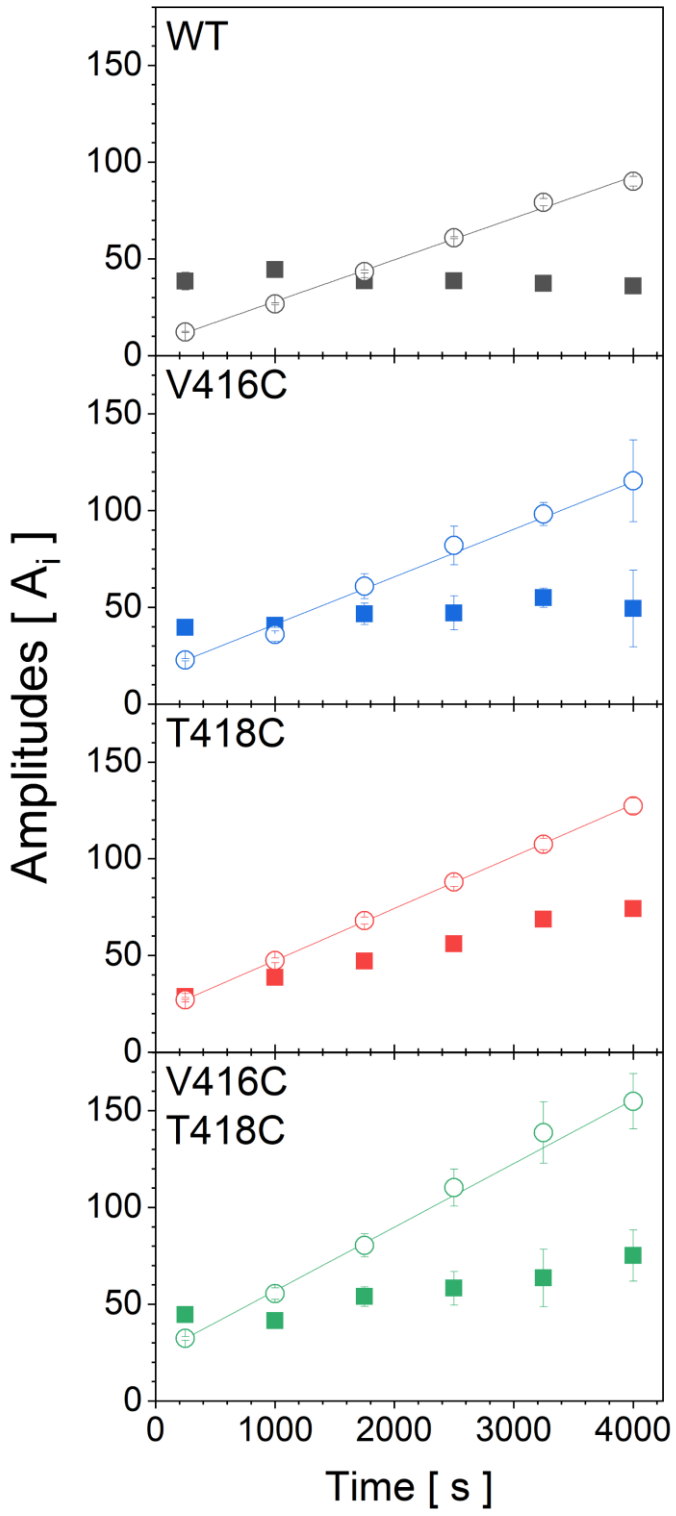


Figure S8. Time evolution of FMN amplitudes obtained from curve fitting of the experimental data for each AsLOV2 variant by Equation 6. Solid squares represent the amplitudes of FMN population inside the protein matrix and empty points represent the amplitudes of free FMN in solvent.

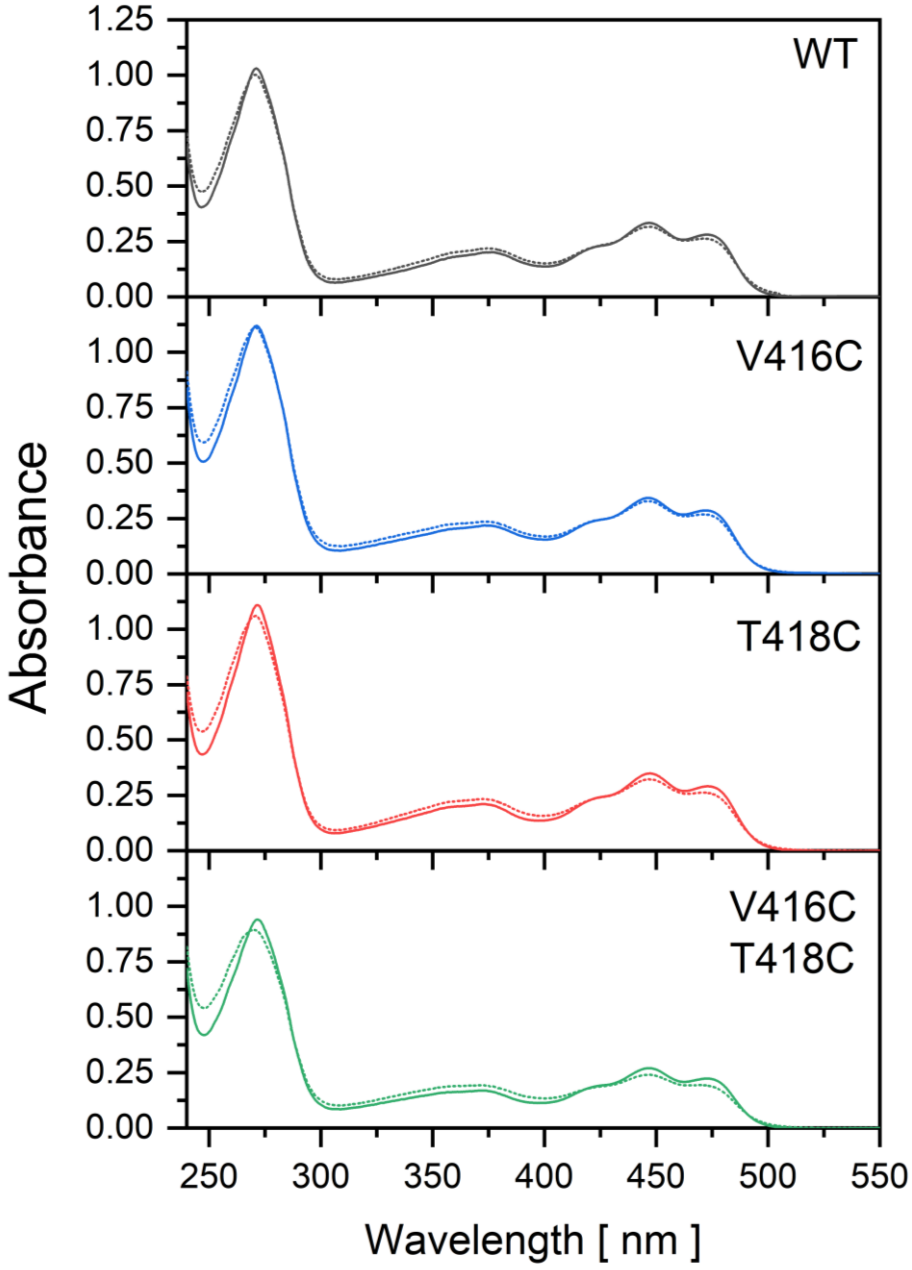


Figure S9. Absorption spectra of AsLOV2 wt and its variants before (solid lines) and after (dasehd lines) irradiation.

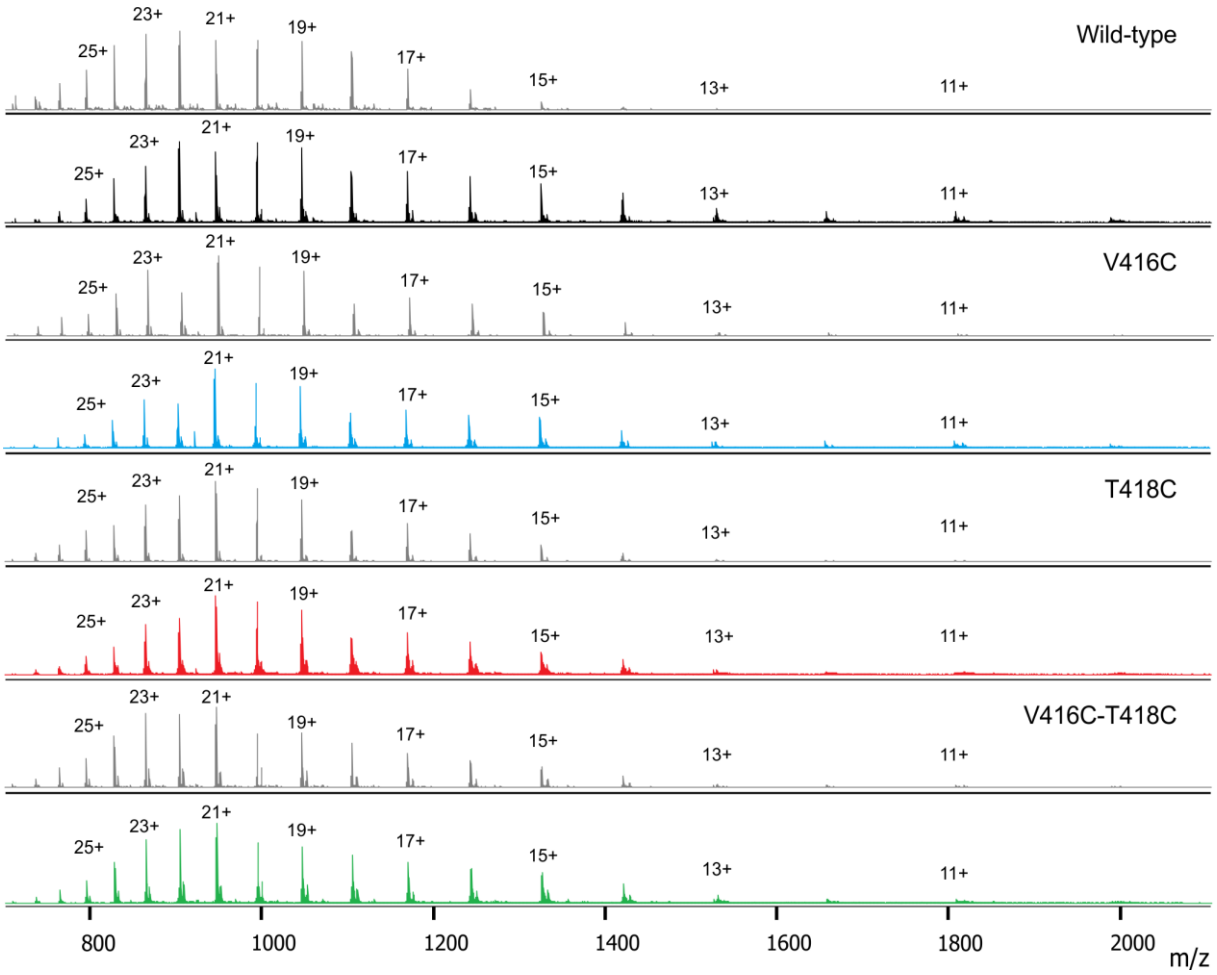


Figure S10. Broadband ESI-MS spectra of AsLOV2 protein and mutants spanning 700-2100 m/z mass region. The grey spectra represent proteins profile before the irradiation. The black, blue, red, and green spectra represent spectra after the irradiation of WT, V416C, T418C, V416/T418C proteins, respectively.

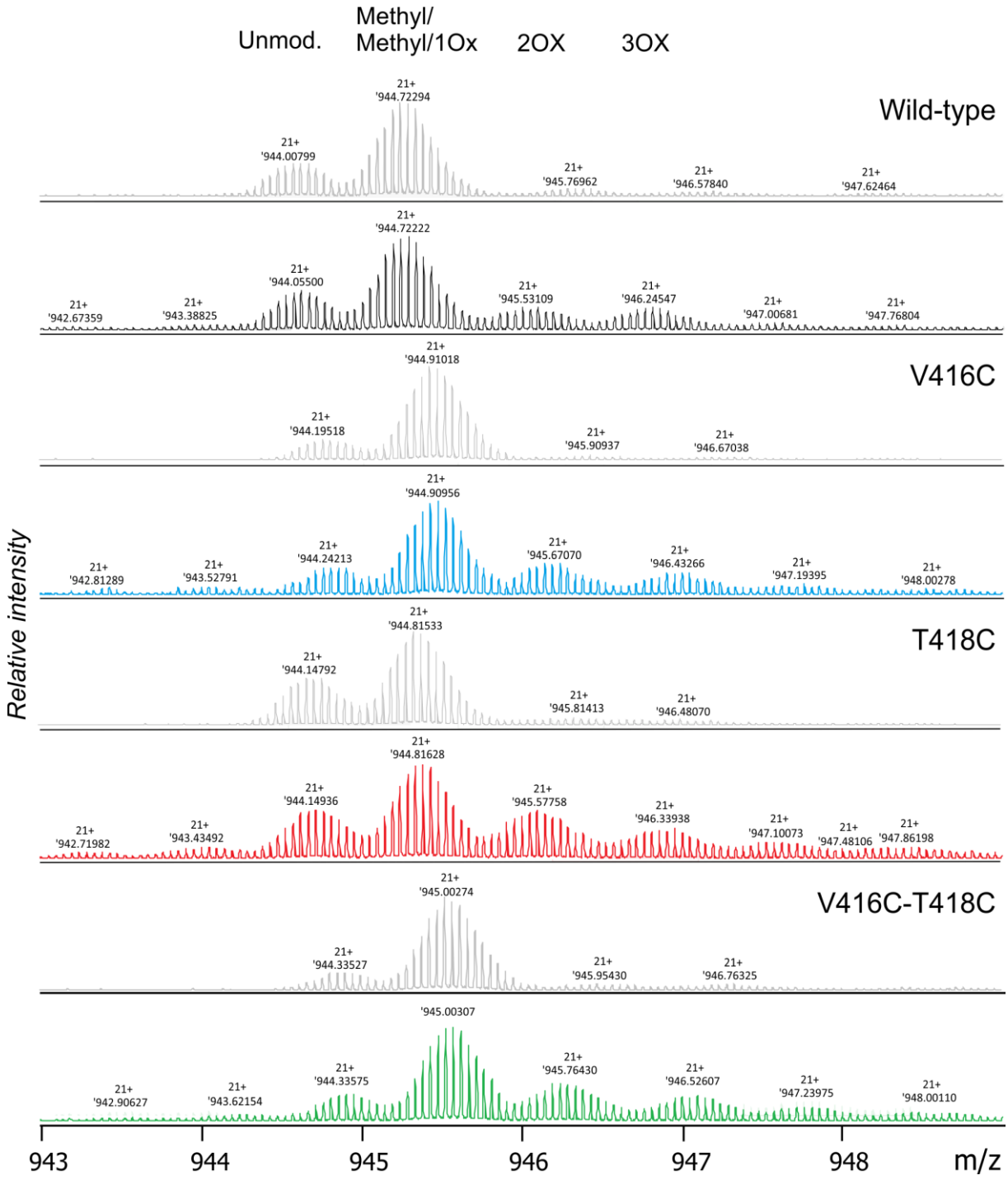


Figure S11. A zoom into the ESI-MS spectra of 21+ charge state, mass region at m/z 943 – 947, before the irradiation (grey spectra) and after the irradiation of either AsLOV2 wt or its variants (black, blue, red and green spectra).

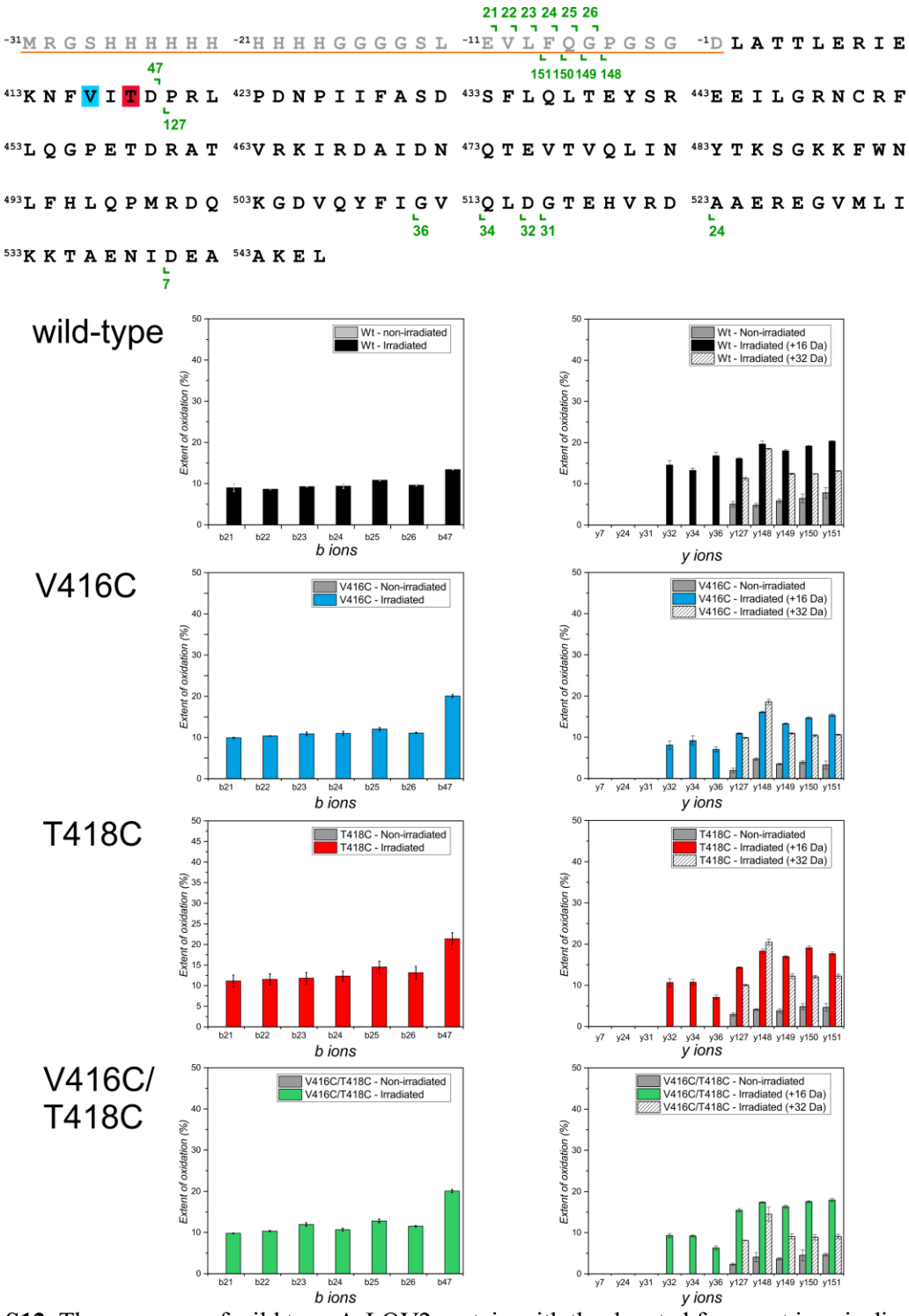


Figure S12. The sequence of wild type AsLOV2 protein with the denoted fragment ions is displayed at the top of the figure. The His-tag, which does not represent the AsLOV2 sequence, is colored in grey and underlined by red line. The position of mutation, V41C, T418, are colored by blue and red squares, respectively. The histograms represent the extent of oxidation for both b and y ion series observed upon multiCASI CID fragmentation. The grey histograms represent the extent of oxidation of b and y fragment ions that originates from non-irradiated samples. The black, blue, red, and green spectra represent extent of oxidation of b and y ions of WT, V416C, T418C, V416/T418C proteins, respectively, after the irradiation. The data were measured in triplicates and are plotted as an average ± SD.

SUPPORTING TABLES

Table S1. The FMN triplet state lifetimes (μs) inside and outside the protein matrix together with singlet oxygen lifetimes at different stages of the irradiation experiment.

	WT			T418C			V416C			V416C/T418C			
	τ_{FMNin}	τ_{FMNout}	$\tau_{\Delta in}$	τ_{FMNin}	τ_{FMNout}	$\tau_{\Delta in}$	τ_{FMNin}	τ_{FMNout}	$\tau_{\Delta in}$	τ_{FMNin}	τ_{FMNout}	$\tau_{\Delta in}$	$\tau_{\Delta out}$
1	1.69	3.02	0.57	2.94	2.94	1.14	0.75	2.92	1.62	1.46	3.03	1.75	3.5
2	1.56	2.87	0.93	2.94	2.94	1.16	0.75	2.81	2.76	1.43	2.89	2.49	3.5
3	1.57	2.72	1.31	2.93	2.93	1.40	0.76	2.79	2.89	1.40	2.83	2.74	3.5
4	1.57	2.73	1.64	2.93	2.93	1.53	0.76	2.84	2.98	1.40	2.82	2.87	3.5
5	1.57	2.63	1.94	2.93	2.93	1.60	0.75	2.78	2.83	1.39	2.86	2.97	3.5
6	1.54	2.83	2.21	2.93	2.93	1.82	0.77	2.79	3.04	1.38	2.83	2.99	3.5

Table S2. The extent of single and double oxidation for b- and y-fragment ions generated by multiCASI-CID. Data are shown only for irradiated samples and are presented in form AVG±SD.

	wt-Irradiated				V416C-Irradiated				T418C-Irradiated				V416C/T418C-Irradiated			
	1OX		2OX		1OX		2OX		1OX		2OX		1OX		2OX	
	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD
b21	9.00	0.93	0	0	9.93	0.13	0	0.00	11.15	1.44	0	0	9.80	0.08	0.00	0.00
b22	8.65	0.13	0	0	10.38	0.05	0	0.00	11.55	1.33	0	0	10.33	0.13	0.00	0.00
b23	9.28	0.10	0	0	10.90	0.42	0	0.00	11.83	1.36	0	0	11.98	0.44	0.00	0.00
b24	9.40	0.57	0	0	10.98	0.50	0	0.00	12.33	1.21	0	0	10.70	0.34	0.00	0.00
b25	10.83	0.17	0	0	12.05	0.37	0	0.00	14.53	1.42	0	0	12.80	0.48	0.00	0.00
b26	9.63	0.17	0	0	11.10	0.16	0	0.00	13.15	1.50	0	0	11.53	0.15	0.00	0.00
b47	13.40	0.10	0	0	20.07	0.35	0	0.00	21.37	1.50	0	0	20.08	0.38	0.00	0.00
y7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0	0	0
y24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0	0	0
y31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0	0	0
y32	4.40	0.87	0.00	0.00	8.08	0.99	0.00	0.00	10.67	0.98	0	0	9.3	0.5	0	0
y34	3.45	1.10	0.00	0.00	9.20	1.18	0.00	0.00	10.78	0.68	0	0	9.2	0.21	0	0
y36	3.83	0.23	0.00	0.00	7.03	0.64	0.00	0.00	7.08	0.55	0	0	6.3	0.48	0	0
y127	12.00	0.50	11.30	0.32	10.93	0.17	9.87	0.12	14.30	0.17	10.05	0.19	15.43	0.42	8.13	0.06
y148	13.73	0.38	18.50	0.10	16.13	0.25	18.57	0.67	18.28	0.56	20.47	0.72	17.38	0.12	14.5	1.73
y149	13.20	0.08	12.45	0.13	13.30	0.14	10.93	0.17	16.97	0.23	12.23	0.58	16.33	0.35	9.1	0.66
y150	15.20	0.12	12.43	0.06	14.70	0.28	10.43	0.22	19.1	0.47	12.05	0.35	17.53	0.29	8.9	0.70
y151	15.58	0.15	13.13	0.10	15.38	0.37	10.63	0.12	17.68	0.43	12.25	0.48	17.92	0.40	9.08	0.53