Univerzita Karlova

Přírodovědecká fakulta

Biochemie

Bc. Marek Polák

Využití radikálového značení bílkovin pro strukturní biologii.
Utilization of protein radical footprinting for structural biology

Diplomová práce

Vedoucí práce: RNDr. Petr Novák, Ph.D.

Konzulant: Mgr. Zdeněk Kukačka, Ph.D.

Praha, 2020

Charles University

Faculty of Science

Biochemistry

Bc. Marek Polák

Utilization of protein radical footprinting for structural biology

Využití radikálového značení bílkovin pro strukturní biologii.

Diploma thesis

Supervisor: RNDr. Petr Novák, Ph.D.

Advisor: Mgr. Zdeněk Kukačka, Ph.D.

Prague, 2020

DECLARATION OF AUTHORSHIP

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 and will not be used for obtaining any other academic degree than Mgr. from Charles University in Prague.

In Prague,

Signature

Acknowledgments

Abstract (In Czech)

Mapování povrchu proteinů je jednou z metod strukturní biologie, které poskytují důležité informace o struktuře, dynamice, funkci a vazebných interakcích proteinů. V této práci jsme se k mapování povrchu proteinů rozhodli využít unikátní přístup, kterým je kombinace hmotnostní spektrometrie s rychlou fotochemickou oxidací proteinů (FPOP, z anglického Fast Photochemical Oxidation of Proteins), která k tomuto účelu využívá reaktivní radikály kyslíku.

Konkrétně byla v této práci metoda FPOP využita ke studiu interakce proteinu s DNA a to přesněji komplexu DNA-vázající domény proteinu FOXO4 s oligonukleotidem DAF16. V první fázi projektu byl produkován a purifikován protein, jehož vazba na DNA byla ověřena nativní elektroforézou a nativní hmotnostní spektrometrií. Dále byly optimalizovány podmínky oxidace proteinu, a to jak v přítomnosti DNA, tak samostatně a modifikace proteinu byly analyzovány hmotnostní spektrometrií pomocí přístupů top-down a bottom-up.

Přístupem bottom–up byla identifikována a kvantifikována modifikovaná residua, což odhalilo rozdíly v oxidaci jednotlivých aminokyselin v přítomnosti a nepřítomnosti DNA. Aby se předešlo možným artefaktům způsobených analýzou vícenásobně oxidovaného proteinu, byla kromě hmotnostně spektrometrické analýzy standardním přístupem bottom-up provedena rovněž analýza v tomto směru pokrokovým přístupem top-down, při kterém byl izolován a fragmentován kolizně indukovanou disociací a disociací záchytem elektronu pouze jedenkrát modifikovaný protein. Rovněž top-down přístup odhalil rozdíly v oxidaci proteinu, které nastaly v důsledku vazby proteinu na DNA, ale pouze na úrovni jednotlivých regionů, které ovšem naprosto korelovali s výsledky z přístupu bottom-up a ukázaly na regiony, které se nalézají na interakčním rozhraní protein-DNA. Naše data jsou rovněž ve shodě s již dříve publikovanou krystalovou strukturou a s daty z vodík-deuteriové výměny komplexu DNA-vázající domény proteinu FOXO4 s oligonukleotidem DAF16.

Naše výsledky ukazují, že oba přístupy bottom-up i top-down jsou užitečnými nástroji při zkoumání interakčního rozhraní mezi proteinem a nukleovou kyselinou. Zatímco bottom-up přístup umožňuje získat informaci o modifikaci s prostorovým rozlišením na úrovni jedné aminokyseliny, přístup top-down naproti tomu dovoluje analýzu pouze jedenkrát modifikovaného proteinu.

Klíčová slova: rychlá fotochemická oxidace proteinu, protein-DNA komplexy, FOXO4, transkripční faktory, kapilární průtokový reaktor, excimerový laser, bottom-up přístup, top-down přístup.

Abstract (In English)

The reaction of highly reactive oxygen radicals with protein solvent-accessible residues can be utilized to map protein landscape. Fast photochemical oxidation of proteins (FPOP) is an MS-based technique, which utilizes highly reactive radical species to oxidize proteins and map protein surface or its interactions with their interaction partners.

In this work, FPOP was employed to study protein-DNA interactions. First, a full-length of FOXO4-DBD was successfully expressed and purified. The ability of the protein to bind its DNA-response element was verified by electrophoretic and MS-based techniques, respectively. Optimal experimental conditions were achieved to oxidize the protein itself and in the presence of DNA, respectively. Oxidized samples were analyzed by bottom-up and top-down approach.

In the bottom-up experiment, modification of individual residues was precisely located and quantified. Different extend of modification was observed for protein alone and in complex with DNA. To avoid experimental artifacts analyzing multiply oxidized protein, standard bottom up approach was replaced by a progressive top-down technology. Only a singly oxidized protein ion was isolated, and further fragmented by collision-induced dissociation (CID) and electron-capture dissociation (ECD), respectively. Quantifying the extent of modification of neighboring sequence ions enabled identification of protein region shielded by DNA. Even the bottom-up approach reached better spatial resolution, both techniques pointed out the same protein regions responsible for DNA binding that are in agreement with previously published crystal structural model and hydrogen-exchange experiments of FOXO4-DBD•DAF16 complex. Our results indicate, that both bottom-up and top-down approaches are useful for probing protein-DNA interface. While bottom-up is able to reach single residue spatial resolution analyzing modified protein on peptide level, top-down allows gas phase purification of singly oxidized protein for further analysis.

Key words: Fast photochemical oxidation of proteins, protein-DNA complexes, FOXO4, transcription factor, quench flow system, protein footprinting, mass spectrometry, bottom-up, top-down.

Table of contents

Abbreviations

ACN	Acetonitrile
BRET	Bioluminescence resonance energy transfer
CASI	Continuous accumulation of selected ions
CCS	Collision cross-section
CID	Collision-induced dissociation
Cryo-EM	Cryo-electron microscopy
CXMS	Cross-linking/mass spectrometry
DBD	DNA-binding domain
DBE	DNA-binding element
DNA	Deoxyribonucleic acid
dsDNA	Double-strand DNA
DTIMS	Drift-tube ion-mobility spectrometry
ECD	Electron-capture dissociation
EDC	1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray ionization
ETD	Electron-transfer dissociation
FOX	Forkhead box proteins
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
GFP	Green fluorescent protein
HDX	Hydrogen/deuterium exchange
HFIP	1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP)
HFN	Hepatocyte nuclear factor
HTH	Helix-turn-helix
IMMS	Ion-mobility mass spectrometry
IPTG	Isopropyl β-D-1-thiogalactopyranoside

IRMPD	Infrared multiphoton dissociation
LB medium	Luria-Bertani medium
LC	Liquid chromatography
MALDI	Matrix-assisted laser desorption/ionization
MES	2-(*N*-morpholino)ethanesulfonic acid
MLL	Mixed-lineage leukemia
MS	Mass spectrometry
NES	Nuclear export sequence
nESI	Nano-ESI
NLS	Nuclear localization sequence
NMR	Nuclear magnetic resonance
O/N culture	Overnight culture
OD_{600}	Optical density at wavelength of 600 nm
PCA	Protein- fragment complementation assay
PMF	Peptide mass fingerprinting
PMSF	Phenylmethylsulfonyl fluoride
PPI's	Protein-protein interactions
PTMs	Post-translational modifications
RE	Response element
S/N	Signal to noise ratio
SAXS	Small angle X-ray scattering
SDS-PAGE	Polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate (SDS-PAGE)
SEC	Size-exclusion chromatography
SIRT	Silent information regulator-2, mammalian homolog of the yeast histone deacetylase Sir2
Skp2	S-phase kinase-associated protein 2
SNAP	Sophistical Numerical Annotation Procedure
SPR	Surface plasmon resonance
TBE buffer	Tris/Borate/EDTA buffer
TEA	Triethylamine
TEMED	1,2-bis(dimethylamino)-ethan

TFA	Trifluoracetic acid
TFs	Transcription factors
TOF	Time of flight
UVPD	Ultraviolet photodissociation
Y2H	Yeast two-hybrid system

1. Introduction

All biological processes in all organisms are intertwined with biomolecules, such as proteins, polysaccharides, lipids, or nucleic acids. These biomolecules may have unique roles as they differ on a structural level between each other. Lipids incorporated in lipid membranes play a protective role and intermediate communication between cells and tissues. DNA stores genetic information about the protein's structure and function and transfers this information to further generations. Proteins mostly form dynamic assemblies or networks that interact and modulate the function between other proteins and consequently regulate the activity of themselves or other biomolecules in the cell[1,2]. Therefore, it is crucial to understand the structure, function, and dynamical interactions of this biomolecular machinery.

The pursuit to understand these fundamentals started in 1951, when Linus Pauling and Robert Corey[3] discovered two basic units of protein composition using protein crystallography and called them α-helix and β-sheet. Two years later, James Watson and Francis Crick[4] explored DNA structure and identified it as a double-helical composition. This was a crucial milestone, which led to understanding how genetic information is stored in the DNA and is inherited in the succeeding generations.

Few years later, the first crystal structure of a protein - myoglobin[5] was developed. It revealed an extensive helical structure of myoglobin and approved the suggested prediction of Linus Pauling and Robert Corey. The protein structure of another protein, hemoglobin[6], was revealed and also confirmed the estimated globular-folded-protein, which also helped to understand how oxygen is transported through the bloodstream and clarified the physiological role of oxygen. This work outlined first insights into the pathology of cells based on a molecular level - -in this case sickle cell disease[7]. A crystal structure of the first enzyme - lysozyme - - was developed in the mid-1960s. The structure provides first insight into individual residues, which are responsible for the catalytic activity of the protein at atomic resolution. These milestones helped to understand how proteins work together, interact, and may be regulated among each other. In the case of enzymes, it helped us to understand how catalytic reactions are intermediated. A possible application can be to design new enzymes with modulated catalytic activity, as opposed to just wild type enzymes.

Such results had an enormous impact, which accelerated the emerging field of structural biology[8]. However, the milestones discussed above were reached by using a single biophysical technique. Nowadays, a comprehensive field of structural biology cannot be resolved by using only a single biophysical technique (Figure 1) due to the limitations of each technique concerning accuracy, resolution, or sample amount. Over the years, nuclear magnetic resonance (NMR) spectroscopy[9] enhances the protein crystallography[10] with additional techniques such as optical or fluorescence methods (Anisotropy, Förster resonance energy transfer - FRET)[11], small angle X-ray scattering (SAXS)[12], cryo-EM[13,14] and the highly emerging mass spectrometry (MS) methods[15], which will be discussed in detail in the following paragraph. The task for integrative biology is to integrate resulting data obtained by these biophysical techniques, combine and complete them with computational biology led by molecular dynamics and Monte Carlo simulations. Integrative biology proposes a huge number of models, and the task is to validate the result and estimate the most accurate model built based on our data[16].

Using integrative methods in structural biology, which refers to as unifying relevant biophysical techniques abridged by computational biology, can help us obtain the desired answers to biological questions[15–18].
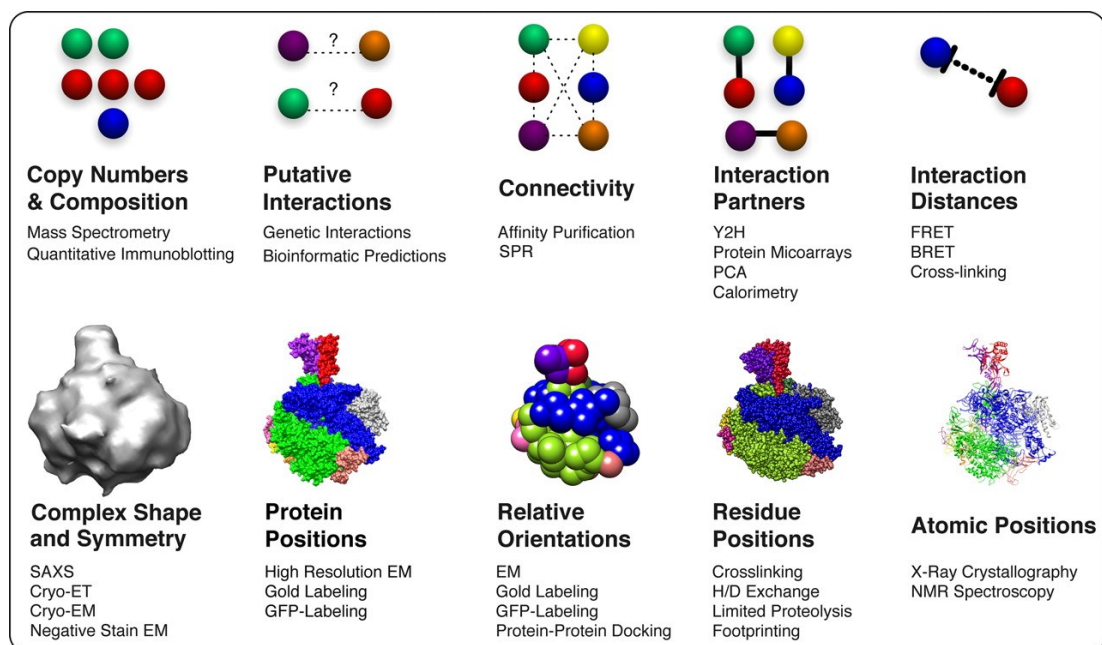
Figure 1 | Different biophysical approaches for obtaining structural information about protein structure, interactions and assemblies. From ref. [19].

1.1 Mass spectrometry approaches for studying protein structure and dynamics

Over the past two decades, mass spectrometry-based proteomics became a valuable approach to study biomolecules. Nowadays, MS-based methods form an individual field of structural biology. Thus, different MS-methods are rapidly emerging rather than other biophysical techniques, enabling study of higher organization of proteins or protein assemblies[1,15].

Increasing knowledge about protein structure, functions, and interactions is being mostly applied in drug design and the development of novel technologies for disease diagnostics[1]. Mass spectrometry approaches analyze the samples in two ways, which are discussed below. Different mass spectrometry methods are further discussed in the next chapters.

Complex proteomic samples can be analyzed in two ways, in a *bottom-up* or *top-down* approach.

Bottom-up approach. Bottom-up proteomics have become a standard in MS analysis. In the bottom-up approach, a protein is cut into smaller, yet overlapping pieces, which enables retrieving information about individual parts and consequently information about the overall protein structure. First, samples, such as proteins, protein mixtures, whole-cell lysates, cells, body fluid, or even organism[18,20–22] are proteolytically digested into peptides using proteases. When analyzing samples containing a massive amount of peptides and proteins, advanced separation techniques such as liquid chromatography, are used prior to mass spectrometry analysis[23]. This enhances a wide dynamic range and enables analysis of such complex samples. Peptides are trapped in a trapping column and desalted during the analysis. Peptides are then transferred to an analytical column for reverse-phase liquid chromatography. After the addition of an organic solution gradient, peptides are being transferred to the mass spectrometer, followed by ionization and monitoring. To obtain specific information about peptides (such as PTMs), tandem MS-experiment (MS/MS) should be performed, followed by data processing and statistical analysis[24].

Unlike the bottom-up approach, *top-down* proteomics can provide information on intact proteins. This requires a soft ionization into a gas phase, which is ensured

by electrospray-ionization (ESI)[25]. By applying high voltage, ESI forms solution -phase intact protein into to positively charged ions. The intact protein is sprayed into the MS in the presence of a volatile solvent (methanol, acetonitrile, isopropanol) with a small addition of acid. With the help of high voltage, temperature, and gas-flow, ions are further desolvated to positively and multiply charged protein ions[25,26]. ESI is being used in both approaches, *bottom-up* (LC-ESI-MS)[24] and *top-down* (intact analysis and fragmentations)[27].

To further investigate information on intact proteins in the gas phase, FT-based analyzers can be utilized. Gas-phase fragmentation provides deeper insights into the protein structure, such as protein sequence[28], post-translational modifications (PTMs)[29], non-covalent interactions[30], composition of assemblies, and the location of subunits (peripheral/core) in a protein assembly[31]. This can be achieved by fragmentation techniques, such as collision-induced dissociation (CID)[26], electron- -collision dissociation (ECD), electron-transfer dissociation (ETD)[26,32] or infrared multi-photon dissociation (IRMPD). A combination of FT-based analyzer and dissociation technique excels high resolution and accuracy and thus provides a powerful method to study protein interactions[30,33].

Several mass spectrometric software programs have been developed to process an extensive amount of mass spectrometric raw data. They are intended to separate, process, and present the MS data, which is currently a challenge for bioinformatics and statistics[34].

The next chapters describe individual MS-based experimental approaches which can be used for studying the 3D structure of proteins: ion mobility and native mass spectrometry, chemical cross-linking, and footprinting techniques.

1.1.1 Ion mobility and native mass spectrometry

In this chapter, ion mobility and native mass spectrometry are discussed individually, yet they are closely related.

Ion mobility. Ion mobility coupled to mass spectrometry (IMMS) provides structural insights of protein structure in the gas phase. IMMS studies the electrophoretic mobility and separation of ions in the gas phase in the presence of low electric field[35,36].

Drift tube ion mobility spectrometry (DTIMS) is the traditional instrumentation of IMS (Figure 2). Drift tube comprises of a series of ring electrodes with a uniform low electric field. Ionization is provided almost extensively by ESI, or nano-ESI. Molecules flow through the drift tube in a gas phase in the presence of ambient pressure; opposite drift gas (mostly nitrogen, also helium or argon) flows from the side of a detector. Due to differences in the ions' mobility, ions are being separated - smaller ions travel faster (higher mobility) and bigger ions travel slower (lower mobility). Ions are subsequently analyzed mostly in a time-of-flight (TOF) analyzer[36].

DTIMS can directly measure the collision cross-section (CCS) value, which is a typical layout of IMMS. CCS reflects the momentum transfer between the ion and gas particles averaged over all gas-ion relative thermal velocities[37].
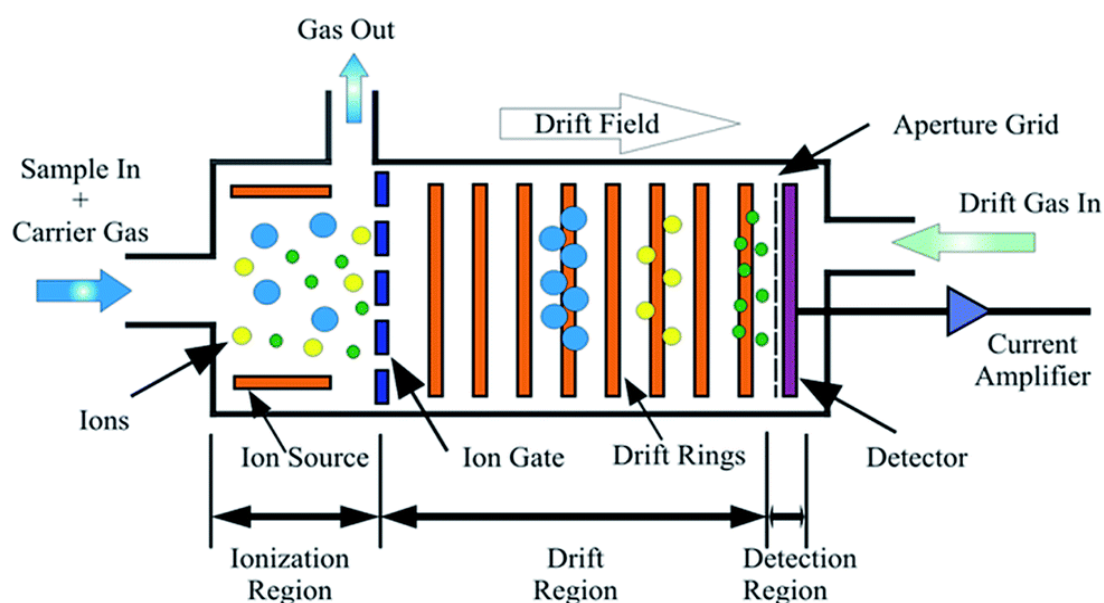
Figure 2 | Scheme of a drift-tube ion mobility spectrometer (DTIMS). After ionization of molecules, ions are separated according to their mobility in the drift tube. Predominantly used time-of-flight analyzer (not shown) provides detection of ions. Adapted from ref.[38].

The application of IMMS is mostly in conformer/isomer separation, such as lipids, carbohydrates, nucleic acids[35], but also in clinical diagnostics[39]. Protein

conformers can be studied as well as protein assemblies[40] or folding/unfolding of proteins[41].

Native mass spectrometry. Native mass spectrometry (native-MS) provides structural information about macromolecular assemblies. The term native, in this case, means gentle enough to preserve non-covalent interactions during the ionization process. It is a type of ESI, where analytes are sprayed in native, non-denaturing conditions. Hence we can determine the exact mass of intact assemblies, their precise stoichiometry, and interactions between individual subunits[15,42].

Nowadays, native-MS is almost exclusively connected with a nESI, because it preserves non-covalent interactions. Nano-ESI provides nanoflow soft ionization to a gas phase through a capillary with a smaller orifice diameter (1-5 μm) than conventional ESI. A reduced flow rate (~10 μl/hrs) results in a benefit of production of smaller droplets, which increase sensitivity, decrease salt adducts production, and allows higher buffer concentration[1,43–45].

Samples present in native conditions are being sprayed involving the use of volatile buffers, such as ammonium acetate (AmAc) or ammonium bicarbonate (AmBic) and physiological pH (~7)[46]. To preserve non-covalent interactions, the usage of both, acidic conditions and the presence of organic solvents are not allowed[42].

To study big macromolecular assemblies, mass spectrometry requires evolving mass spectrometers to increase their *m/z* range. The reason is that large protein complexes over 60 kDa typically form ions with an *m/z* ratio of over the 4000. A modern and modified instrument with detection *m/z* above 4000 can detect these assemblies. This instrumentation consists mostly of nESI-quadrupole-time-of-flight equipment (nESI-Q-TOF)[47]. Modified mass spectrometers with an increased *m/z* ratio enable detecting and investigating very large complexes, such as viruses[48] or ribosomes[49].

Native-MS is a powerful analytical approach because it can provide high-resolution data of huge macromolecular complexes. Native-MS data combined with data obtained by other biophysical techniques are an excellent powerful tool. Native-MS has been successfully used to solve protein assemblies composition (elF3-ref. [50], Hsp90 chaperon assemblies[51], Cascade protein[52]), macromolecular complexes (viruses[48], ribosomes[49]), protein-lipid complexes[53–55], DNA assemblies

like G-quadruplex[56], noncovalent nucleic acid-drug interactions[57], ligand interactions[58] and dynamics of protein dimerization (bacterial GlmS protein)[59].

1.1.2 Chemical Cross-linking

Chemical cross-linking/mass spectrometry (CXMS) provides information about the 3D structure of a protein or protein-protein interactions (PPIs). Even though NMR and X-ray crystallography provide information on the 3D structure, CXMS can also provide information about the PPIs network, e.g. in living cells. The main advantage of CXMS is that the protocol requires short reaction time and a minimal amount

of sample, compared to other techniques. Results can be further combined with the results from other techniques. Especially a combination of CXMS and cryo-EM provide insight into large biomolecular assemblies in cells, that other methods[60,61] cannot provide.

The main idea of a cross-linking reaction (Figure 3, p. 21) consists of a covalent bond of two amino acid residues, which are exposed to solvent and can react with the reactive group of a cross-linker. The result of the coupling reaction is a cross--linked protein, or protein molecules, that interact among each other[2]. The result of a cross-linking reaction is further analyzed mostly by one-dimensional gel electrophoresis (SDS-PAGE), or MALDI-TOF. The main purpose of this step is to visualize the result of the reaction, and eventually to optimize the yield of a reaction. It is essentialto optimize protein and cross-linker concentrations, because both can cause a production of artifacts[62]. Optimizing buffer composition and pH is also crucial, as they may induce protein denaturation[63].

The results of the reaction can be analyzed by both, *bottom-up* or *top-down* approach. After protein digestion in the *bottom-up* approach, the mixture contains unmodified peptides of the protein and cross-linked peptides of type 0 (modified solvent accessible amino acid by a cross-linker), type 1 (intra-peptide cross-linked amino acids in one peptide) and type 2 (inter-peptide cross-linked peptides)[60,63]. Such mixture undergoes LC-MS/MS analyses, followed by data processing in software such as StavroX[64], GPMAW[65], X-Link[66], or most recent Mango[67]. The software creates a list of peptides originated by protein *in-silico* digestion. Subsequently, the software compares obtained MS/MS spectra with molecular weights of peptides

from *in silico* digestion. If a match is found, cross-linked peptides are automatically matched and reported[63].

With the *top-down* approach, the sample is analyzed mostly by an ESI-FT-ICR instrument, where one-modified cross-linked product is isolated and submitted to fragmentation during the analysis[63].
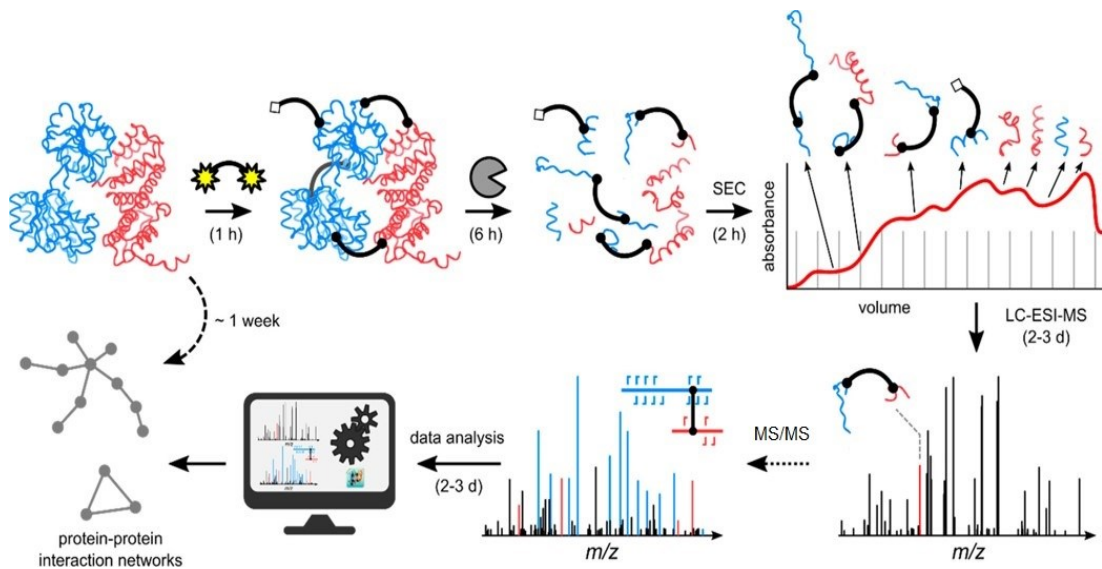
Figure 3 | The general cross-link workflow. Cross-linked reaction, followed by in-solution digestion, size-exclusion chromatography, and LC-ESI-MS/MS analysis. Raw data are then analyzed and interpreted. Adapted from ref. [68].

The result of the CXMS is a list of residues, which are in an appropriate spatial proximity and they are cross-linked. Based on the result, a 3D model of the protein and protein assemblies can be built. The cross-linking approach has been successfully used for mapping protein 3D structures, protein-protein interactions (PPIs)[60], proteome networks[69] in cell lysates[70], organelles[71], or tissues[72]. Notably, protein-DNA cross-linking among protein and specific functional groups of nucleic acids was also documented[73].

1.1.3 Footprinting techniques

1.1.3.1 Limited proteolysis

Limited proteolysis belongs to the lower-resolution methods. It was introduced over 60 years ago in order to study higher organizations of biomolecules. The concept is based on the fact that proteolytic digestion is not only based on a protein sequence

but also depends on surface-exposed amino acids. A protein in its native state produces less digested peptides or fragments compared to a denatured one. Surface-exposed and flexible regions are proteolytically digested the most. Proteolysis can also unravel individual protein features, such as domains, as they are proteolytically separated from the original complex. With the development and enhancement of mass spectrometers, limited proteolysis became possible to be coupled with MS[18,74].

1.1.3.2   Covalent labeling techniques

1.1.3.2.1   Hydrogen/deuterium exchange

During the last two decades, hydrogen/deuterium exchange (HDX) became a valuable tool to gain structural insights into the protein structure, function, and dynamics. The theoretical basis of hydrogen/deuterium exchange was established in the 1950s, when Kaj Ulrik Linderstrøm-Lang and his coworkers made their first HDX experiments by measuring the stability of the hydrogen bonds involved in the secondary structure of proteins. They found out that hydrogen exchanged for deuterium in the amide backbone of the protein reflects conformational changes caused by protein dynamics[75,76]. In the 80s, the H/D exchange was measured on NMR instruments. In the 80s, ionization techniques – MALDI[77] and ESI[25] were introduced and later commercialized. The combination of pH- and temperature-based sample quenching, followed by chromatographic separation, enables this technique in the field of structural mass spectrometry[78,79].

The secondary structure of a protein is stabilized by hydrogen bonds, which are formed between amide hydrogen and carbonyl oxygen (N–H···O=C stabilization). In a typical HDX experiment, the reaction starts by mixing protein and the solvent containing D_2O. The exposure of protein to deuterated solvent induces a spontaneous replacement of the backbone hydrogen (protium, ^1H) by deuterium (^2H). Mass spectrometry analysis can detect this exchange, as the deuterium is heavier than hydrogen.

However, not every hydrogen is easily exchangeable by deuterium. Hydrogen atoms bonds to carbon atoms too strong to be exchanged. Hydrogen atoms of labile groups (carboxy –COOH, hydroxy –OH, primary amine –NH, sulfhydryl –SH) exchange too fast, and their detection is complicated. The exceptions represent

C(2)–H hydrogen atoms in histidine, and also proline, of which hydrogen exchange is too slow.

Amides on a peptide backbone are great candidates to monitor HDX for at least four reasons: (i) amide backbone hydrogens can be found along the whole protein molecule, because they're bound to the polypeptide chain, (ii) the binding of any ligand to a protein molecule will cause even a small structural change in a polypeptide chain. Consequently, a small structural change will cause a change in the deuterium rates in the binding region, and the HDX will detects this; (iii) hydrogens on amide backbones can be exchanged, considering a neutral pH, on a wide time scale. Other hydrogens are exchanging "too slow" or "too fast"; (iv) During the HDX, amide backbone hydrogens are subjected to specific kinetic condition. This reaction can be quenched at a low pH and temperature, which enables monitoring protein dynamics at various time scales (Figure 4, p. 24)[2,78].

The reaction of the exchange of hydrogen for deuterium is catalyzed by a base, an acid, and water. The rate of exchange (k_{ch}) then follows the equation 1:

$$k_{ch} = k_{int,based}[OH^-] + k_{int,acid}[H_3O^+] + k_{int,water}[H_2O] \qquad (1)$$

where the terms $k_{int,based}$ and $k_{int,acid}$ mean rate coefficients for the base- and -acid-catalyzed reaction, $k_{int,water}$ means the intrinsic rate coefficient for the water--catalyzed reaction.

The HDX rate highly depends on the (i) pH and (ii) temperature[2,78,80].

(i) *Dependence on pH.* The acid-catalyzed H/D mechanism is highly preferable at pH 2.5-3. Above this pH range, base-catalyzed H/D mechanism is dominant. That means that H/D experiments provided at physiological pH (~7) are strongly base-catalyzed. Then, $k_{int,acid}$ and $k_{int,water}$ can be neglected. At pH 2.5-3 (pH$_{min}$) the exchange is the slowest (approximately $10^5\times$ slower than at pH 7) (Figure 4a, p. 24).

(i) *Dependence on temperature.* A drop of temperature from 25 °C to 0 °C leads to a 14× decrease of k_{ch}, according to a modified Arrhenius equation:

$$k_{ch} = k_{rc}(293)e^{\left(-\frac{E_A}{R}\left[\frac{1}{T}-\frac{1}{293}\right]\right)} \qquad (2)$$

where $k_{rc}(293)$ is the deuterium exchange rate constant, E_A is the activation energy of a catalyzed proton exchange dependent on pH, and R is the universal gas constant (8.314 J·K^{-1}·mol^{-1}). Overall, the HDX kinetics is approximately 14× slower at 0 °C than at 20 °C (Figure 4b)[78].

This kinetic information is further utilized in HDX-quenching conditions: pH drop to 2.5-3 and temperature drop to 0 °C.
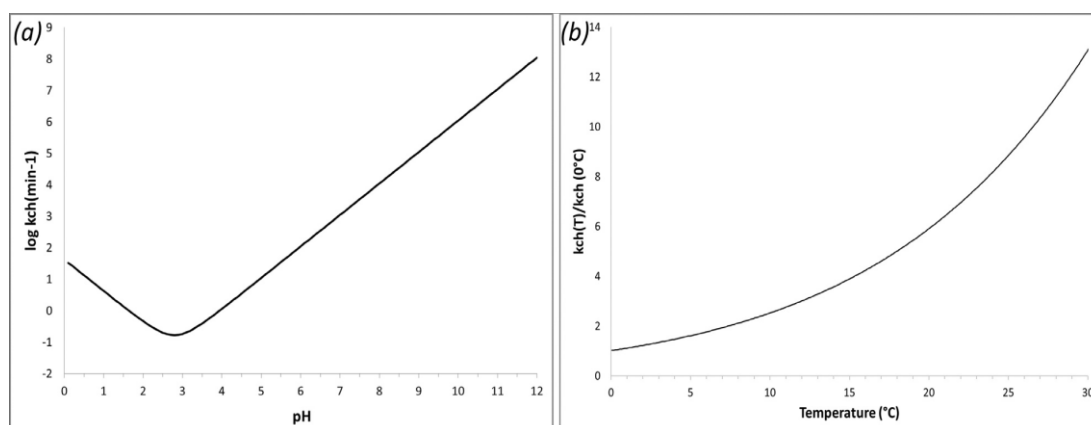
Figure 4 | The dependence of HDX kinetic rates on (a) pH and (b) temperature. The slowest HD exchange is between pH 2.5-3 and at a temperature of 0 °C. Adapted from ref. [78].

HDX kinetics. The first HDX kinetics model was developed by Linderstrøm-Lang and his colleagues in the 1960s. They discovered that proteins' amide backbones exist or fluctuate between two states, closed-one (N–H$_{closed}$) and open-one (N–H$_{open}$), where proton can be exchanged for a deuterium[2]. In the open state, deuteration is possible with the rate constant k_{ch} according to the following equation:

$$\text{N}-\text{H}_{\text{closed}} \underset{k_{cl}}{\overset{k_{op}}{\rightleftarrows}} \text{N}-\text{H}_{\text{open}} \underset{\text{D2O}}{\overset{k_{ch}}{\rightleftarrows}} \text{N}-\text{D}_{\text{open}} \underset{k_{op}}{\overset{k_{cl}}{\rightleftarrows}} \text{N}-\text{D}_{\text{closed}}$$

(3)
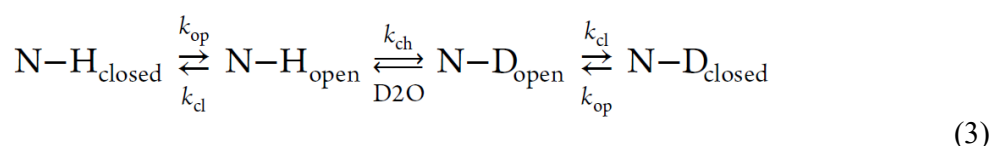
where N-H and N-D represent backbone amides, which are hydrogenated or deuterated, respectively. Kinetic constants k_{cl} and k_{op} reflect the structural changes upon folding and unfolding. Rate constant k_{ch} represents the speed of the HDX reaction itself[2,78].

24

HDX can be analyzed by both *bottom-up*[81] and *top-down*[82] approaches. The *bottom-up* approach is predominant. The labeled and quenched samples undergo online digestion and analysis by LC-MS/MS[2,78]. Figure 5 describes the representative HDX workflow.

The top-down approach is quite rare and even nowadays only a few research groups are developing this method. Major problems arise mostly during the gas-phase fragmentation, where hydrogen/deuterium migrates along the entire polypeptide chain ("scrambling")[78].
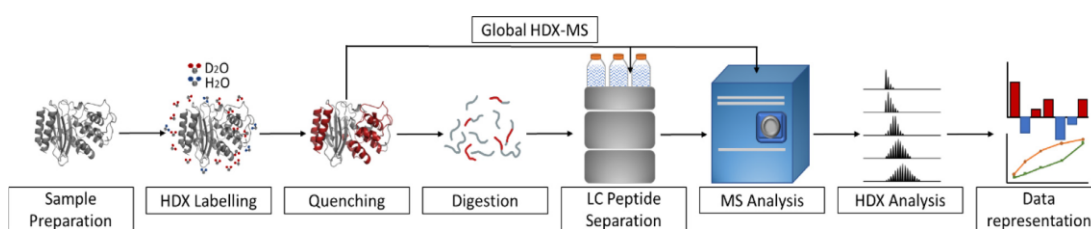
Figure 5 | HDX workflow depicting the usual *bottom-up* analysis. Adapted from ref. [78].

HDX has been successfully applied to monitor protein dynamics[83] and conformational changes[84], protein-protein interactions[85], protein-DNA interactions[86], membrane proteins[87], protein-small molecule interactions[88], or drug discovery[79,89].

1.1.3.2.2 Stable covalent labeling

Covalent labeling techniques coupled to mass spectrometry are footprinting techniques that utilize radicals to label protein surface. Footprinting techniques like other S-based techniques can provide information about protein structure and dynamics. HDX primarily monitors structure and dynamics on a secondary structural level, while footprinting techniques monitor solvent accessibility of proteins[90]. Data obtained by these two techniques can be complementary to each other and provide structural insight about protein structure or dynamics[91]. Covalent labeling can be further divided based on a using probe as a chemical or radical labeling, respectively.

1.1.3.2.2.1   Chemical labeling

Chemical modification of amino acid residues was described to probe surface-topology. Firstly, amino-acetylation with acetic anhydride[92] or succinylation[93] reaction of ε-amine of lysine was described. Not only lysine residues can be modified, but also arginine can be modified by 1,2-cyclonhexanedione to form N7, N8-(dihydroxy-1,2-cyclohexylidene)arginine[92].

Secondly, aromatic functional groups can react with tetranitromethane and creates stable nitrated aromatic groups. Aromatic residues in proteins can be nitrated, because the reaction can be carried out under gentle conditions, which is suitable for biomolecules[94].

Lastly, solvent accessible carboxyl groups of aspartate and glutamate can be irreversibly modified. A protein is modified using glycine ethyl ester (GEE) upon the previous activation mediated by EDC (1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide). The mass shift is +85 Da (+GEE), or +57 Da (+glycine), respectively[95].

1.1.3.2.2.2   Radical labeling

Generated reactive radicals can also serve as a probes for protein structure determination. Carbene footprinting is a method with the lack of specificity, which modifies all residues. The source of methylene radicals can be diazirine[96] or photoleucine[97]. Because of their high reactivity, methylene radicals are created from diazirine/photoleucine *in situ* by laser photolysis at wavelength of 355 nm, followed by the protein footprinting reaction at nanosecond time scale, and terminated by a self-quenching reaction with solvent H_2O. The footprinting method modifies residues with variable mass shift (+14 Da for methylene)[96].

Photochemical iodination of biomolecules was also documented. Iodination was performed directly in quench flow capillary system, followed by radical generation by photolysis of iodobenzoic acid at the wavelength of 248 nm, and quenched by histidine. Iodine radicals are more site-specific than hydroxyl radicals, and modify only histidine and tyrosine residues[98].

The common footprinting technique is hydroxyl radical footprinting. The protein undergoes oxidation intermediated by reactive •OH radicals, which can be generated

by various methods. A technique utilizing hydroxyl radicals generated by UV flash photolysis, provided in a continuous flow-capillary system, is called Fast Photochemical Oxidation of Proteins (FPOP) and is described in the following chapter. The next chapter describes the possibilities of hydroxyl radicals' generations for performing FPOP.

Generation of hydroxyl radicals

Nowadays, hydroxyl radical footprinting is the most used irreversible covalent labeling technique. There are several ways to generate hydroxyl radicals *in situ*.

(i) *Fenton chemistry* was the first method to generate hydroxyl radicals in aqueous solution[99,100].

$$Fe(II) + H_2O_2 \rightarrow Fe(III) + OH^- + \bullet OH \tag{eq. 4}$$

$$Fe(III) + H_2O_2 \rightarrow Fe(II) + HO_2^{\bullet} + H^+ \tag{eq. 5}$$

$$H_2O_2 + \bullet OH \rightarrow HO_2^{\bullet} + H_2O \tag{eq. 6}$$

$$HO_2^{\bullet} \rightarrow O_2^{\bullet -} + H^+ \tag{eq. 7}$$

$$Fe(III) + HO_2^{\bullet} \rightarrow Fe(II) + O_2 + H^+ \tag{eq. 8}$$

$$Fe(III) + O_2^{\bullet -} \rightarrow Fe(II) + O_2 \tag{eq. 9}$$

$$Fe(II) + \bullet OH + H^+ \rightarrow Fe(III) + H_2O \tag{eq. 10}$$
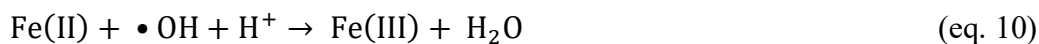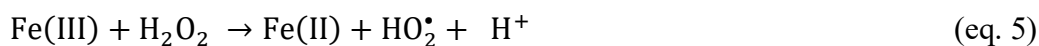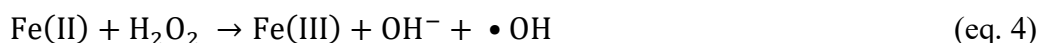
In an initial reaction called Fenton-Haber-Weis reaction, Iron(II) is oxidized to iron(III) by hydrogen peroxide, forming hydroxide ion and hydroxyl radical (equation 4). In the reaction, the metal ion is present in a complex with EDTA. The reaction in equation 4 is a free radical-mediated initiation reaction, reactions in equations 5-9 are propagation reactions and may be included in the reaction chain. Reaction in equation 10 presents a radical terminal reaction[100].

The goal is to implement the recycle of Fe(III) to Fe(II), and the addition of ascorbate can achieves this. The reaction in equation 4 is then self-sustained with reaction time from several minutes to dozens of minutes, in the dark and the optimal pH range of 3-4. Hydroxyl radicals, which are created, subsequently oxidize solvent-accessible residues of proteins[100–102]. Hydroxyl radicals generated by Fenton reaction have been successfully utilized to study protein surface mapping[102].

(ii) *Radiolysis of water by a synchrotron.* Synchrotron generates ionizing energy with an energy range of 3-30 keV. The reaction is fast, and oxidation is completed within a few milliseconds. Comparing the concentration of protein to the concentration of water in the sample, direct contact of X-ray and the protein is minimal and provides typical conditions for the FPOP experiment with no/minimal sample damage[103].

$$H_2O + hv \longrightarrow H_2O^{\bullet+} + e^-_{dry} \qquad \text{(eq. 11)}$$

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

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

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

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

$$2\,OH^\bullet \longrightarrow H_2O_2 \qquad \text{(eq. 16)}$$
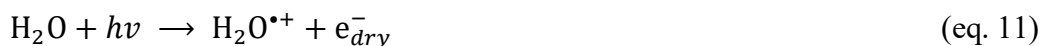
Radiolysis of water generates a dry electron and a water radical-cation (equation 11) which reacts with another water molecule and generates a hydroxyl radical (equation 12). The reaction of dry electron results in a hydrated electron e^-_{aq} and excited water molecules (equation 13), which subsequently produce hydroxyl radicals and hydrogen atom radicals (equation 14). Generated hydroxyl radicals oxidize solvent-accessible residues. In water, hydroxyl radicals can also be quenched according to equation 15 and 16, respectively.

Synchrotron equipment as a hydroxyl radical generator has been successfully used to map protein surface[104], protein interactions[105,106], conformational changes[107–109], and dynamics[109,110].

(iii) *Electron pulse radiolysis.* This technique utilizes pulses of electrons, which are accelerated by a linear accelerator, approximately at ~3 MeV. In a diluted aqueous solution, electrons interact with water molecules and creates hydroxyl radicals. The reaction mechanism is similar to the radiolysis of water by X-rays on synchrotron (equations 11-14)[100]. Electron pulse radiolysis of water has been successfully utilized to generate hydroxyl radicals and perform footprinting of proteins[111].

(iv) *Photolysis of hydrogen peroxide.* The most used approach of generating hydroxyl radicals is flash photolysis of hydrogen peroxide. The generation of hydrogen radicals,

called Haber-Weis chain reaction, also consists of a self-quenching step, which is illustrates equations 17-20 (ref.[100,112]). In the reaction, hydroxyl radicals are generated by flash photolysis by an excimer laser at a wavelength of 248 nm.

$$H_2O_2 + h\nu \rightarrow 2\,HO^\bullet \qquad\qquad (eq.17)$$

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

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

$$2\,HO^\bullet \rightarrow H_2O_2 \qquad\qquad (eq.\ 20)$$

Fast photochemical oxidation of proteins (FPOP)

Fast Photochemical Oxidation of Proteins (FPOP) is a hydroxyl radical based footprinting technique, that utilizes an excimer laser to dissociate hydrogen peroxide to hydroxyl radicals, which modify solvent accessible residues on the protein surface.

Early in the 1980s, the purpose of hydroxyl radical "footprinting technique" was to study ligand-biomolecule protection and to map protein-DNA interactions[113,114]. Hydroxyl radicals generated by Fenton chemistry were later accepted as a reagent to monitor protein structure, interactions, and dynamics. Footprinting was then used to study protein-DNA interactions[115] and radical-induced nucleic acids damages[116]. The first report of protein hydroxyl radical footprinting was published in 1988 (ref.[117]).

FPOP was first developed by Aye, Low, and Sze[118], who first connected oxidative labeling with the continuous flow setup. Consequently, FPOP is mostly performed in the quench-flow capillary system. The setup comprises of syringes filled with the sample of interest, hydrogen peroxide (~10 mM) and a quencher; syringe pumps, capillary flow path, and excimer laser. In a flow system setup, a protein of interest is mixed with hydrogen peroxide. A flow path leads the mixture through a transparent window, where the excimer laser triggers photolysis of hydrogen peroxide. The generated hydrogen radicals subsequently oxidize amino acids on a protein surface. To avoid further post-footprinting oxidation, the leftover hydroxyl radicals are quenched with a scavenger right after the transparent window. The quencher is typically a free amino acid – methionine or glutamine at a ~50mM

concentration. To prevent any post-FPOP oxidations, a catalase can be added to the collecting tube as an enzyme which dissociates hydrogen peroxide to water and oxygen[112,119].

Using a capillary flow system and hydrogen peroxide as a source of hydroxyl radicals has an advantage compared to Fenton reaction because it has shorter exposure time. The capillary system provides continuous flow and exposure of the protein to hydroxyl radicals for a short time because the scavenger immediately quenches the reaction. Water radiolysis by X-ray or electron beam provides fast oxidation with minimal secondary reactions, and the overall pattern of oxidation is similar to laser photolysis[90].

An excimer laser is significantly cheaper compared to synchrotron sources, which are also not so easily available. KrF excimer laser produces a UV pulse at a wavelength of 248 nm. These properties of the UV beam can be directly controlled by the frequency or energy of the laser, or even pulse width. Connection of these parameters with the full control of width pulse gives this setup a huge advantage[112,119]. The lifetime and amount of created hydroxyl radicals in the Fenton reaction cannot be controlled in this way[90,112,119].

The amino acid reactivity can differ as they oxidize differently. To investigate the reactivity of the amino acid residues, Xu and Chance[120] utilized X-ray or Cesium-137 γ-ray source to oxidize individual amino acids and FT-ICR-MS/MS to further investigate oxidative modifications. It was found out, that 14 out of 20 amino acids are useful in a typical FPOP experiment.

For aliphatic amino acids (Figure 6, p. 31), the carbon of aliphatic residues is oxidized by oxygen to form hydroxyl or ketone/aldehyde group and creates a mass shift of +16 Da or +14 Da, respectively[120].

Aliphatic side chains

+16 Da   +14 Da
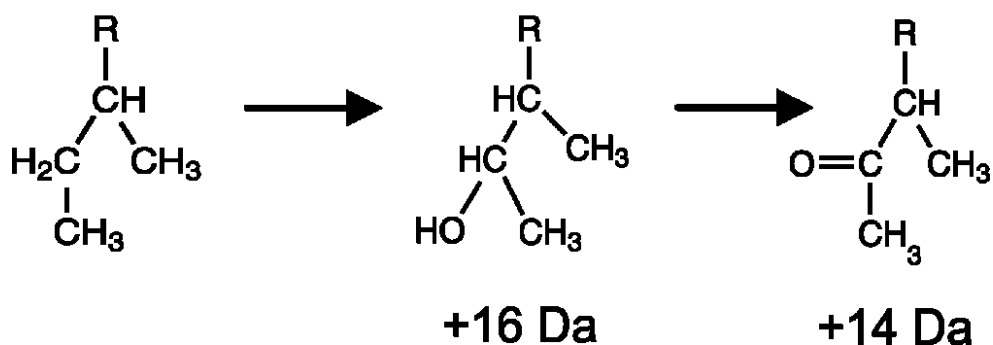
Figure 6 | Oxidation of aliphatic amino acids. Oxidation on carbon atom leads to the formation of a hydroxyl or ketone group. Created in ChemDraw® Professional 2016 according to ref. [120].

Cysteine and methionine as sulfur-containing residues are the most oxidizable of all amino acids. After exposure to hydroxyl radicals, they form a typical mass shift of +16/+32 Da for methionine (Figure 7). Cysteine can be oxidized three times with a +64 Da mass shift to create sulfonic acid or create serine as a minor product of the reaction, respectively[120,121].

Methionine (Cysteine)

+16 Da   +32 Da
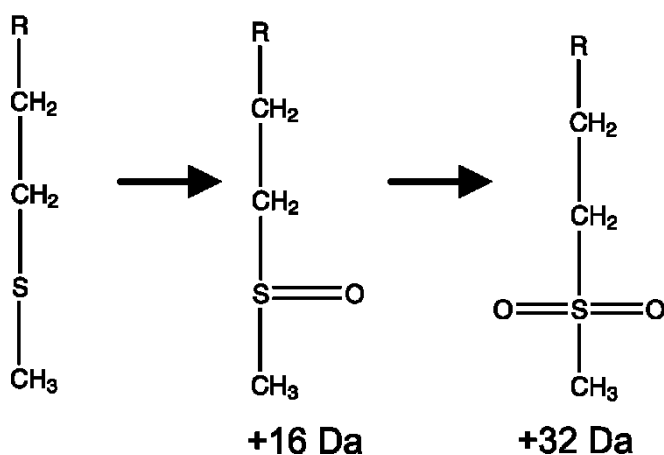
Figure 7 | Oxidation of sulfur-containing amino acids. The sulfur atom can be oxidized once or twice, for cysteine also three times. Created in ChemDraw® Professional 2016 according to ref. [120,121].

The reaction of hydroxyl radicals with aromatic residues (Phe, Tyr, Trp) results in the addition of +16 Da to an aromatic ring (Figure 8, p. 32). A single oxidation at any position is possible, as well as multiple oxidations (+32, +48 Da)[120].

Phenylalanine
(aromatic residues)

+16 Da
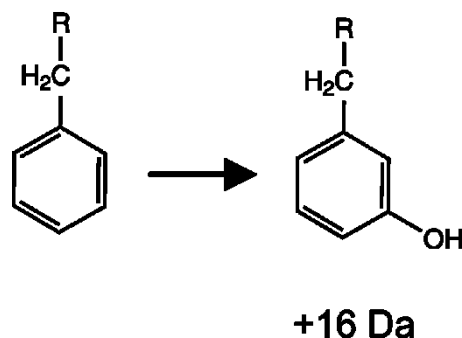
Figure 8 | Oxidation of aromatic amino acids results in an addition of oxygen on aromatic rings. Multiple oxidations are possible as well. Created in ChemDraw® Professional 2016 according to ref[120].

Oxidation of acidic residues (Asp, Glu) results in an oxidative decarboxylation (loss of CO_2) (Figure 9). The mass shift is then -30 Da. Methylene groups of Asp and Glu can undergo oxidation forming a hydroxyl group (+16 Da) or a ketone group (+14 Da)[122].

Glutamic (Aspartic) acid
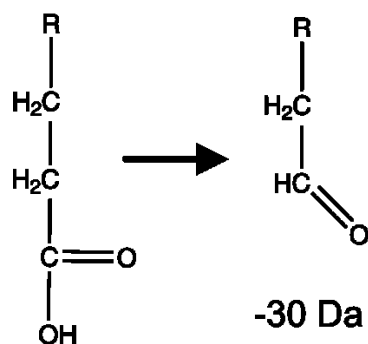
-30 Da

Figure 9 | Oxidation of acidic amino acids (Glutamic or Aspartic acid) results mostly in an oxidative decarboxylation with a loss of -30 Da. Created in ChemDraw® Professional 2016 according to ref. [122].

Reactions of hydroxyl radicals with basic amino acid residues can differ. Histidine (Figure 10, A, p. 33) reacts with hydroxyl radicals, involving imidazole ring rupture resulting in possible multiple products at different masses. Arginine (Fig. 10, B, p. 33) can be oxidized in two different ways. The first involving +16Da/+14Da oxidation on an aliphatic residue. The second way involving an attack of a hydroxyl

radical on δ-carbon following in guanidine-group elimination with 43 Da mass shift loss[123].
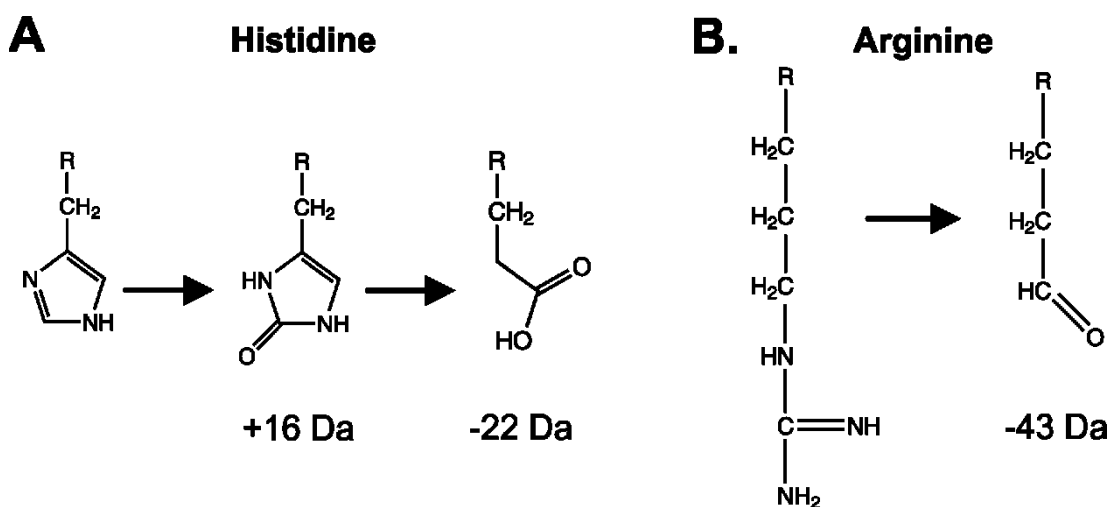
Figure 10 | Oxidation of basic amino acids (Histidine and Arginine). Oxidation of Histidine (**A**) is complicated, results in multiple possible products. Arginine (**B**) mostly results in a loss of guanidine group with a -43 Da loss. Created in ChemDraw® Professional 2016 according to ref. [123].

According to a publication by Xu and Chance[120], the reactivity of individual amino acids varies. Biomolecular rate constants of the reaction of all 20 amino acids with hydroxyl radical and hydrated electron at neutral pH have been examined. Xu and Chance calculated rate constants based on the decay of individual amino acids as a function of time (typical dose-response study). Rate constants of oxidized reaction were obtained after nonlinear fitting of an unmodified compound as a function of exposure time regarding to the first-order kinetics. The amount of unmodified compound and product of oxidation were calculated based on the intensity. The fraction of the unmodified compound was calculated as a ratio of signal intensity of unmodified compound to the sum of intensities of modified and unmodified compound[120]. Six amino acids (Asp, Asn, Ala, Gly, Thr, Ser) are not good probes for FPOP, due to their lower reactivity. However, the remaining 14 amino acids report higher reactivity, which includes ~65 % of the typical protein sequence, and thus can provide a reliable sequence resolution for examining the protein surface[90,120]. Figure 11, p. 34 visualizes the overall relative reactivity of all 20 amino acids based on their calculated rate constants.

Figure 11 | The relative reactivity of the amino acid side chains. The most reactive are cysteine, methionine, and tryptophan, almost non-reactive are alanine and glycine. Adapted from ref.[120].

Analysis and data interpretation.

After an FPOP experiment, the collecting tube contains an oxidized sample, the rest of hydrogen peroxide, and methionine as a quencher (can also contain catalase). Samples can be frozen, or directly submitted to further mass spectrometric analysis. Typical for FPOP experiments is to divide the sample into two part: one half for an intact analysis and the second half for bottom-up analysis.

For intact analysis, the sample is desalted using solid-phase extraction and sprayed via electrospray ionization to MS. Intact analysis visualizes yield of oxidative modification.

To locate and quantify the site of oxidative modification, a bottom-up approach coupled with MS is used for further analysis. Oxidized samples are further digested for bottom-up approach. To increase the intensity of peptides, specific proteases, such as trypsin, AspN, LysC, GluC, or their combinations are being used. Typical LC-MS/MS is performed to further localize and quantify the site of modification per each peptide[112,124,125].

Quantification of the extent of modification can be calculated for a certain peptide, from their intensities, according to equation 21, p. 35 (ref.[112]):

$$\text{Extent of modification} = \frac{\sum I_{\text{OX}}}{\sum (I_{\text{OX}} + I)} \tag{eq. 21}$$

where $\sum I_{\text{OX}}$ is the sum of intensities of modified peptide/fragment and $\sum (I_{\text{OX}} + I\ )$ are sum of intensities of both unmodified and modified peptide.

A recently published article by Khaje et al.[126] from the year 2018 points out the main problem of the irreproducibility in FPOP analysis, which can be caused by a small sample amount suited for LC-MS analysis. Published data show an LC-MS analysis of two samples after an FPOP experiment. The mixture of tryptic digested peptides was analyzed by LC-MS in two different volumes: smaller amount against 2.5 times higher sample volume. LC-MS analysis was performed, and data was processed with an appropriate statistical data evaluation. The data shows, that increasing an overall amount of the injected sample significantly decreased the coefficient of variability of each peptide. The recommendation of the study is to prefer using higher sample amount of FPOP samples during the LC-MS analysis. Smaller oxidation yield in FPOP experiments can cause smaller intensities of individual peptides and therefore variable coefficient of variability.

Preserving native conditions during FPOP

In 2018, Chea and Jones published an article[127], in which they dealt with preserving native conformation of proteins after the FPOP experiment. In their FPOP experiment, they used two model protein enzymes, lysozyme and invertase.

Immediately after the FPOP experiment, an enzymatic assay was used to monitor the catalytic activity of both enzymes. The enzymatic activity was normalized with respect to protein concentration and folded protein control (protein not submitted to FPOP experiment). Lysozyme retained 47%±0.2 of the activity of the native protein after oxidation. Over-oxidized lysozyme retained only 6%±0.02 of the original activity.

Immediately after FPOP, invertase was subjected to an assay to measure its catalytic activity. Compared to the normalized sample, the oxidized sample retained 34%±2.8 of its original activity, while the overoxidized sample showed significant loss dropping to 9%±1.3 of its original activity[127]. According to Chea and Jones,

it can be assumed that the FPOP experiment does preserve the native conformation of proteins.

To this day, FPOP has been used for probing protein interface[128], conformational changes[109], protein aggregation[129], membrane proteins[130,131], mapping the antibody epitope[132,133], *in-cell*[134], or *in-vivo*[135] FPOP. Comparing HDX and FPOP has been also described in the literature[91]. Protein unfolding[136,137] has been documented as well, as the utilization of FPOP high-resolution data for further modeling and molecular dynamics simulations[138–140].

1.2 Transcription factors

Transcription factors (TFs) are cellular proteins, which recognize and bind specific DNA sequences to regulate transcription of the gene in a positive, or negative manner. TFs form a complex system, that controls the expression of the genome. TFs report a different function in various cell lines in the body. TFs may also regulate different genes in some way due to the fact that they can bind to specific DNA sequences. That truly verifies an idea, that this family of proteins has a dynamic purpose in cells[141,142].

The sequence, which TF typically recognizes *in vivo*, is called the response element (RE)[142,143]. TFs recognize small DNA sequences (typically 6-12 bp)[144] in different genes, which may have roles as promoters, enhancers, silencers, or regulatory regions[145].

The region in which TF binds to DNA is called a DNA binding domain (DBD). In 2013, Wingender et al.[145] provided a bioinformatic analysis of a huge number of TFs and classified them into 9 superclasses, 40 classes, and 111 families. As was published in above mentioned article[145], DBDs have been identified and classified into 9 superclasses (Figure 12, p. 37): (i) basic domain, (ii) Zinc--coordinating domain, (iii) Helix-turn-helix domain, (iv) other all-α-helical DBDs, (v) α-Helices exposed by β-structures, (vi) immunoglobulin fold domain, (vii) β-Hairpin exposed by an α/β--scaffold, (viii) β-Sheet binding to DNA domain, (ix) β-Barrel DBD. The most explored DBDs are the first three – basic domains, zinc-coordinating domains, and helix-turn-helix domains, which are closer described below[145].
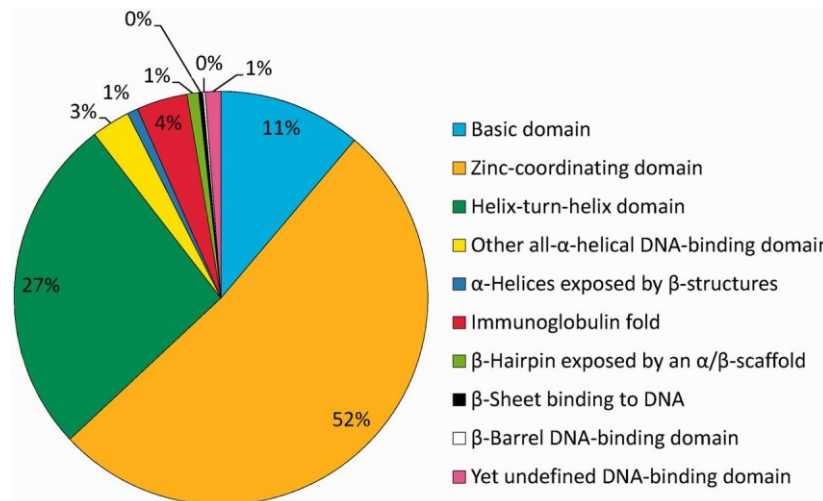
Figure 12 | The percentual representation of individual superclasses of DBDs, according to the bioinformatics study from Wingender et al.[145].

Basic domains contain a region consisting of basic amino acid residues, which binds to DNA. Without the presence of DNA, this region exists as an unstructured region. Upon DNA binding, the basic region forms an α-helical structure, which is stabilized by DNA. Moreover, the dimerization of the molecule is necessary for DNA binding, which ensures the motif structure similar to a leucine--zipper[146].

Zinc-coordinating domains, alternatively called zinc finger domains are also well explored. These domains consist of an antiparallel β-sheet, which contains a loop stabilized by cysteines bridge, and an α-helix containing a His-His loop. These two structures form a motif due to: (i) stabilization by hydrophobic core formed by Tyr, Phe, and Leu between sheet and helix; (ii) zinc coordination, where zinc is coordinated by two histidines and two cysteines in the molecule center.

These 7 amino acids (Tyr, Phe, Leu, 2× cysteine, 2× histidine) are highly conserved and stabilize the tertiary structure of the zinc finger. Other variable amino acids determine the specificity of each zinc finger domain, which binds to a specific DNA motif. The crystal structure of a zinc finger-DNA complex observed primary contact of protein with DNA that is intermediated by an α-helix binding into the major groove in DNA. Direct interactions are intermediated by 3 hydrogen bonds between α-helix and DNA[147].

Helix-turn-helix (HTH) motif was found as another domain found in both prokaryotes and eukaryotes. The base of the domain is formed by a right-handed

helical bundle which rises from helix-turn-helix conformation. A sharp turn between 2nd and 3rd helices is characteristic for the HTH motif. Helix H3 typically recognizes and creates a protein-DNA interface by binding into the major groove. The DNA recognition and binding specificity may vary by based on the residue diversity forming helix. Additional contacts of the other parts of a protein are also important[148].

1.2.1 Mechanism of action of transcription factors

As mentioned above, transcription factors contain a DNA binding domain, which recognizes a specific DNA sequence and binds to it. Transcription factors can also contain specific regions necessary for the activation of transcription. Regulation domain can stimulate transcription in a positive, or negative manner[141].

Positively, these activation domains simply activate transcription. Several activation domains were defined, which differ by occurrence of proline, glutamine or acidic residues, respectively. Activation domains can also directly interact with small molecules or interact with the basal transcriptional complex. Nevertheless, these results in an activation of transcription[142].

Binding of TF to DNA can also inhibit the transcription activity. This can be achieved by binding negatively acting binders to prevent the binding of positively acting TFs to DNA. To prevent the activation of transcription, some proteins can form protein-protein interaction of both factors (the positive and negative), to achieve this. Neutralization of activation by repressor on DNA, called quenching is also possible[142]. Transcription repressors were also discovered, as the proteins which repress transcription activity in cells[142,149].

1.2.2 Regulation of transcription factors

There are two possible ways to regulate transcription factor activity. The first is to regulate the synthesis of transcription factors. This regulation is secured by the synthesis of a protein in one specific tissue or cell type, but not in another. This mechanism is a widespread way of the regulating transcription factors, which control a specific cell type or gene expression[142].

The second way is to control the transcription factor activity. Pre-existing transcription factors in the cell are activated by specific signals, which allow *de novo*

transcription. That can be mediated by following mechanisms: (i) protein-protein interaction[150]; (ii) ligand binding[151]; (iii) posttranslational modifications (PTMs)[152]. Furthermore regulation of the transcription factor activity via PTMs, phosphorylation, O-/N-glycosylation, ubiquitinoylation, acetylation, methylation, or sumoylation is also possible[152].

1.2.3 Forkhead box/winged-helix transcription factors

Forkhead box (FOX, also calls winged helix) transcription factors are a class of proteins consisting of an approximately 110 amino acids long winged helix DNA binding domain. The DBD forms a well-known helix-turn-helix motif[153–155].

The first member of this family of which crystal structure with its target DNA has been solved in 1993 (ref.[156]) was a hepatocyte nuclear factor 3γ (HFN3γ, today known as a FOXA3). This structure forms a typical helix-turn-helix motif and two characteristic large loops, called "wings" (winged helix proteins). This protein family shares a highly conserved DNA binding region with a forkhead protein, whose gene was first found and identified in *Drosophila*. These forkhead proteins found in *Drosophila* are involved in proper formation of the anterior and the posterior gut in *Drosophila* fly embryos[154,157,158].

Nevertheless, the official name of winged helix proteins remains Forkhead box (FOX) proteins, both contain "winged-helix" structural motif and are found in both, prokaryotes and eukaryotes. All FOX proteins share a characteristic, 110 amino acids long, highly-conserved winged helix DNA binding domain. Considering the phylogenetic analysis, over 100 members of FOX proteins were unified, and today they are classified into 18 subgroups (FOXA-FOXR)[155,159,160]. Individual subgroups members are numbered (FOXE1)[155].

1.3 FOXO subgroup of FOX family

FOXO subfamily transcription factors play an important role in metabolism control, cell-differentiation, stress response, cell survival, DNA damage repair response, and apoptosis. The subgroup consists of four members, FOXO1, FOXO3, FOXO4, and FOXO6[161].

The FOXO "O" subfamily are vertebrate orthologs of DAF-16 transcription factors isolated from *Caenorhabditis elegans*, which share similar, highly conserved DBD[162]. FOXO subfamily recognizes and binds to a responsive element (RE), so-called DAF-16: 5`-GTAAACAA-3`. This sequence includes a core sequence 5`-(A/C)AA(C/T)A-3`, which recognizes all FOX proteins. FOXO proteins can also bind to an insulin-response element (IRE), which is present in the IGFBP-1 promoter region (5`-(C/A)(A/C)AAA(C/T)AA-3`). FOXO proteins bind to IRE with lower affinity than to DAF16 DBE[153,161].

FOXO1, FOXO3, and FOXO4 are expressed widely. FOXO1 is present mostly in adipose tissue, FOXO3 in the brain, kidney, heart, and spleen, FOXO4 in skeletal muscle[163]. FOXO6 is specific to tissue development in adult brains and neural tissues, playing an important role in the nervous system[163–165].

1.3.1 FOXO4

FOXO4 has been identified according to translocating studies on 11q23 in acute mixed-lineage leukemia (MLL)[165,166]. At least, 29 chromosome translocations have been found in MLL, which result in a fusion with several different partner genes. The fusion of AFX (FOXO4) gene to MLL in acute leukemia was found, translocation t(X;11)(q13;q23)[166].

FOXO4 is a member of the "O" subfamily of FOX transcription factors[155]. Its role is mostly in cell differentiation, cell cycle progression, and negative regulation of cell proliferation[160]. FOXO4 is a·~500 amino-acid residues long protein. FOXO4 contains a typical Forkhead domain consisting of three α-helices (H1-H3), three β-strands (S1-S3) and two wing loops (W1, W2). FOXO4 (and some other FOXO proteins) have a typical additional 3_{10}-type helix (H4) located between α-helices H2 and H3. The arrangement of FOXO4-DBD follows the H1-S1-H2-H4-H3-S2-W1--S3-W2 topology[153,154,167].

FOXO4 overall consists of four domains: (i) highly conserved Forkhead DBD; (ii) nuclear localization signal (NLS); (iii) nuclear export sequence (NES); (iv) C-terminal transactivation domain. Figure 13, p. 41 shows the arrangement of the individual domains. Multiple sequence alignment analysis shows that DBD, N-terminal region, NLS, and part of C-terminal transactivation domains are highly conserved[153].
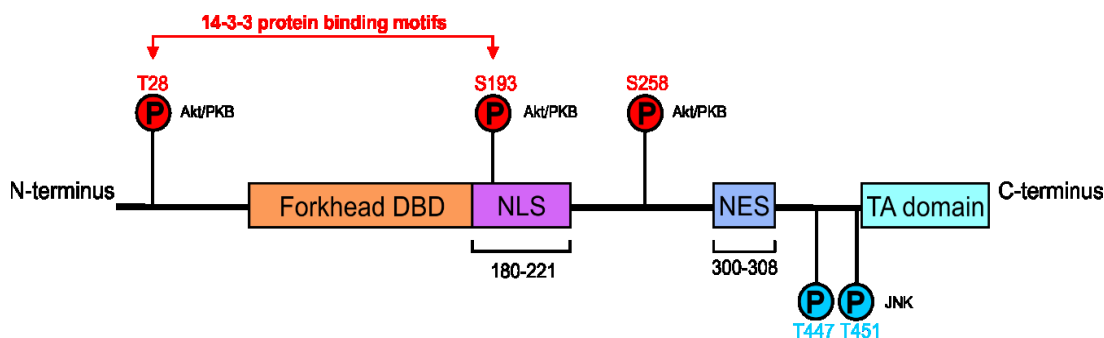
Figure 13 | The arrangement of individual domains in FOXO4. The most extensive domain forms the DBD, followed by the nuclear localization sequence (NLS), the third domain is the nuclear export sequence (NES) and the last is the C-terminal transactivation domain. Adapted from ref.[153].

1.3.2 Modulation of FOXO4-DBD•DNA interactions through PTMs

The importance of the transcriptional activity of FOXO4 remains in control of cell cycle regulation, integration of different signaling pathways, cancer, and the control of life span[153,168].

The activity of FOXO4 is regulated at a level of posttranslational modification (PTMs) through phosphorylation, acetylation/deacetylation, and ubiquitination.

1.3.2.1 Regulation by phosphorylation

To control the transcriptional activity, FOXO4 can be phosphorylated by several kinases: (i) protein kinase B (PKB, also referred to as Akt), (ii) c-Jun N-terminal kinase in case of Ral-JNK pathway.

Protein kinase B (PKB) is a member of the phosphatidylinositol 3-kinase pathway. Akt/PKB induced phosphorylation of FOXO4 occurs predominantly at three sites. The first site is located at the N-terminus (T28), second site is on the wing W2 (S193), and the third one at S258. Both T28 and S193 phosphorylations are highly important. Phosphorylation at only T28 residue does not interfere with DNA binding affinity. However, phosphorylation at S193 (near to region containing basic amino acids region) only reduce the DNA binding affinity[153,168].

Phosphorylation status at both N-terminus (T28) and C-terminus (S193) creates two binding motifs for 14-3-3ζ (zeta isoform) proteins. Phosphorylation of both T28 and S193 is recognized, followed by a stable association of two 14-3-3ζ molecules,

which affect the DNA binding affinity; and the function of FOXO4 is being regulated in this way. Upon binding of 14-3-3ζ molecules, FOXO4 completely inhibits DNA binding affinity by interaction with DBD. Resulting FOXO4-DBD-(14-3-3ζ)$_2$ complex masks the NLS (one of the binding sites of 14-3-3ζ, S193, is located at NLS) and cause rapid decrease of transcription activity by transporting the complex into the cytoplasm with interfering and masking NLS[153,161,168–171].

Phosphorylation at only one residue (T28 or S193) and complex formation FOXO4-14-3-3ζ doesn't significantly affect FOXO4-DNA binding affinity[172].

In the case of the Ral/JNK pathway, FOXO4 is phosphorylated by c-Jun N-terminal kinase on two residues, T447 and T451, respectively. Phosphorylation at these residues is essential for the transcriptional activity of FOXO4. Phosphorylation results in translocation from cytoplasm to nucleus and enhances the transcriptional activity[161,173,174]. JNK can also phosphorylate 14-3-3ζ. This can lead to decrease of binding affinity with its interaction partners[175].

1.3.2.2 Regulation by acetylation/deacetylation

Acetylation/deacetylation is another process which can the control transcription activity through PTMs. To decrease transcription activity, acetylation of FOXO4 is being utilized. Several acetyltransferases exist and can acetylate the FOXO4 protein (histone acetyltransferase, p300, the cyclic-AMP responsive element binding-binding protein, cyclic-AMP responsive element binding-binding protein-associated factor)[153].

Therefore, a deacetylase, such as SIRT1 (mammalian homolog of the yeast histone deacetylase Sir2, silent information regulator-2) regulates and balances this action. 14-3-3 protein is also required for FOXO4-SIRT1 interaction.

Also, there have been suggestions that acetylation of FOXO4 can also enhance transcription activity. However, counter opinions are still matter of debates[153,161,168].

1.3.2.3 Regulation by ubiquitination

Both monoubiquitination and polyubiquitination are possible in the case of FOXO4. Cells exposed to increased oxidative stress reported monoubiquitination of FOXO4 protein, relocate FOXO4 from the cytoplasm to the nucleus, and enhance its transcription activity.

Monoubiquitination occurs at lysine residues K199 and K211 at C-terminus of the DBD. The same lysine residues are acetylated to decrease transcription activity, so we can say that these regulations are in some manner opposite-acting[153].

Polyubiquitination results in proteasomal degradation of FOXO4, which is ensured by a specific E3 ligase, called Skp2. This happens in the case of the Akt/PKB signaling pathway, in which FOXO is relocalized into the cytoplasm, where it can be subsequently degraded. In both cases, mono- and polyubiquitination, a USP7 specific enzyme can binds to FOXO4 and ensure deubiquitination[153,161].

1.4   Structure of FOXO4-DBD•DAF16 complex

An NMR solution structure of FOXO4-DBD[167], as well as crystal structure of FOXO4-DBD·DAF16, has been solved at 1.9Å resolution[176].

In this case, the construct at length of 82-183 has been used for the successful crystallization of FOXO4, lacking a part of the C-terminal region of wing W2. This terminal region was impossible to crystallize due to high flexibility, and it does not directly interact with DNA[169,176].

The crystal structure of FOXO4-DBD (82-183) (Figure 14, p. 44) provides a typical winged-helix topology of three α-helices (H1-H3), additional 3_{10}-type helix (H4), three β-strands (S1-S3) and wing W1. FOXO4-DBD binds to a 13bp of DAF16 (5`-TTG GGT AAA CAA G-3`, 5`-CTT GTT TAC CCA A-3`). As it has been observed in other solved FOX-DBD-DNA structures, α-helix H3 creates the main contact with DNA, docking the major groove of a duplex. Other parts of the DBD just participate in DNA binding. Due to FOXO4-DBD binding to DNA, approximately 13° bend of DNA is observed in a crystal structure[176].

The main recognition of FOXO4-DBD bound to DNA is by α-helix H3. H3 helix contains a conserved sequence (N^{168}-X-X-R-H-X-X-S/T^{175}) found also in other FOXO-DBDs. Amino acids form direct and water-mediated hydrogen bonds and van der Waals contacts with DNA bases[176]. Wing W1 and N-terminal loop also participate and interact with DNA, N-terminal amino acids create direct and water-mediated hydrogen bonds with phosphate groups of DNA[169,176]. FOXO4 also contains an H4 helix, which is located between H2 and H3 helices. Its function

is unclear because H4 helix can be found only in FOXO proteins. In the case of FOXO3 and FOXO1, H4 helix does not create any protein-DNA contact. In FOXO4, S142 at H4 helix interacts with the phosphate group of the DNA. Water molecules occupied protein-DNA interface and create a highly ordered network of hydrogen bonds between protein residues and bases[176].
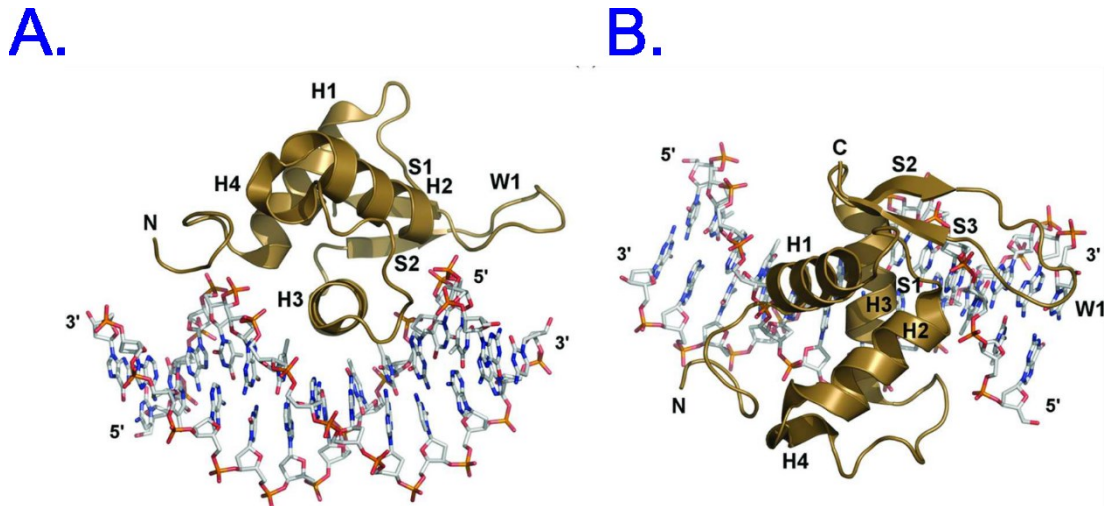
Figure 14 | A. The overall view of FOXO4-DBD bound to DNA. B. 90° rotation of complex in A., toward its horizontal axis. Adapted and modified from ref.[176].

2. Aims of the thesis

The overall goal of this thesis was adopting a Top-down approach for footprinting utilizing fast photo-oxidation of proteins on protein-DNA complexes.

The specific goals were:

1. Expression of DNA-binding domain of FOXO4 protein.

2. Characterization of the ability of protein binding to DNA.

3. Laser condition optimization to achieve an ideal extent of modification.

4. The use of bottom-up analysis to localize and quantify modified amino acids side chains.

5. Utilization of top-down approach to localize and quantify oxidized protein regions.

6. Finding amino acid side chains or protein regions of FOXO4 DBD differently modified in absence/presence of DNA and comparison with previous reports [86,167,176].

3. Material

3.1 Instruments and consumables

Analytical weighing scale ML 104/01	Mettler Toledo, Czech Republic
15T-solariX XR FT-ICR mass spectrometer	Bruker Daltonics, Billerica, USA
Automatic pipettes	Mettler Toledo, Czech Republic
Buchner funnel kit, 250 ml	Sigma-Aldrich, USA
CCD camera ChemiDoc MP Imaging System	Bio-Rad Laboratories, USA
Centrifuge (ultra) Avanti J-26 XP	Beckman Coulter, USA
Centrifuge 5920R	Eppendorf, USA
Centrifuge Mini spin	Eppendorf, USA
Concentrators Amicon Ultra	Merck, USA
Cooling/Heating Dry Block Biosan CHH-100	Biosan, Latvia
Excimer laser Compex 50 KrF	Coherent, Inc., USA
Falcon® Conical Centrifuge Tubes	VWR, Czech Republic
FPLC Enrich SEC 70 10 × 300 column	Bio-Rad Laboratories, USA
FPLC system NGC Quest 10	Bio-Rad Laboratories, USA
Fused silica capillaries	Polymicro Technologies, USA
HPLC system 1200 series pump	Agilent Technologies, USA
Laser puller P-2000	Sutter Instrument, USA
Luna® Omega, particle size 3 µm Polar C18, pore size100 Å (0.3 × 150 mm)	Phenomenex, USA
Luna® Omega, particle size 5 µm Polar C18, pore size100 Å (0.3 × 30 mm)	Phenomenex, USA
Native-PAGE apparatus	Bio-Rad Laboratories, USA
nESI glass capillaries	Sutter Instrument, USA
nESI source	AffiPro s.r.o., Czech Republic
Peristaltic pump Econo gradient pump	Bio-Rad Laboratories, USA
pH electrode InLab Expert	Mettler Toledo, Czech Republic
pH meter Orion 2 Star	Thermo Fischer Scientific, USA

PowerPac 1000 Electrophoresis Power Supply	Bio-Rad Laboratories, USA
SDS-PAGE apparatus	Bio-Rad Laboratories, USA
Shaker Multitron Pro	Biotrace, Czech Republic
Spectrophotometer DeNovix DS-11 FX+ Thermomixer	DeNovix, USA
Sterile syringe disk filters, 0.22 μm	Corning® Incorporated, USA
Syringe pumps	New Era Pump Systems, Inc., USA
Syringes	Hamilton, USA
Trap column (protein)	Optimize technologies, USA
Ultrasonic homogenizer UP200S	Heilscher, Germany
Vortex MS 3 digital	IKA, Germany

3.2 Chemicals and enzymes

1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP)	Sigma-Aldrich, USA
1,2-bis(dimethylamino)-ethan (TEMED)	Sigma-Aldrich, USA
1.4-dithiothreitol (DTT)	Sigma-Aldrich, USA
2-(*N*-morpholino)ethanesulfonic acid (MES)	Sigma-Aldrich, USA
Acetic acid (MS grade)	VWR, Czech Republic
Acetonitrile (ACN, MS grade)	EMD Millipore Corp, USA
Acrylamide	Serva, Germany
Ammonium acetate	Sigma-Aldrich, USA
Ammonium persulfate	VWR, Czech Republic
Ampicillin	Biotika, Czech Republic
Bacto - Yeast extract	VWR, Czech Republic
Bromophenol blue	Serva, Germany
Cobalt(II) chloride	VWR, Czech Republic
Coomassie Brilliant Blue R-250	Fluka, Switzerland
Disodium phosphate	VWR, Czech Republic
DNAse I	Thermo Fischer Scientific, USA
Ethanol	VWR, Czech Republic

Ethylenediaminetetraacetic acid (EDTA)	VWR, Czech Republic
Formic acid (FA)	Sigma-Aldrich, USA
GelRed® dye	Biotinum, USA
Hydrogen peroxide	Sigma-Aldrich, USA
Imidazole	Fluka, Switzerland
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	VWR, Czech Republic
Leupeptin	VWR, Czech Republic
Loading dye (6×) for native PAGE	New England BioLabs, Inc., USA
LysC	Promega, USA
Lysozyme	Thermo Fischer Scientific, USA
Magnesium(II) chloride	VWR, Czech Republic
Methanol	Lach-Ner, Czech Republic
Methanol (MS-grade)	EMD Millipore Corp, USA
Methionine	Sigma-Aldrich, USA
N,N-methylene-bis-acrylamide	Sigma-Aldrich, USA
NaTFA tuning mix (MS grade)	Agilent Technologies, USA
Orto boric acid	VWR, Czech Republic
Phenylmethylsulfonyl fluoride (PMSF)	VWR, Czech Republic
Precision Plus Protein™ Unstained Standards	Bio-Rad Laboratories, USA
RNAse A	Thermo Fischer Scientific, USA
Sodium chloride	VWR, Czech Republic
Sodium dodecyl sulfate	VWR, Czech Republic
Talon® Superflow resin	Clontech Laboratories, USA
TEMED	Bio-Rad Laboratories, USA
Thrombin	EMD Millipore Corp, USA
Triethylamine	Sigma-Aldrich, USA
Trifluoroacetic acid	Sigma-Aldrich, USA
Tris(hydroxymethyl)aminomethane	VWR, Czech Republic
Tris-Glycine 10× running buffer (TGS)	Bio-Rad Laboratories, USA
Trypsin	Promega, USA
Trypton	Oxoid, UK

| Water (MS grade) | EMD Millipore Corp, USA |

All chemicals were purchased in p.a. purity, all solvents for LC-MS analysis were MS grade

3.3 Bacterial cell lines

| *Epicurian Coli* BL21-Gold (DE3) | Stratagene, USA |

Genotype: F– ompT hsdS(rB – mB –) dcm+ Tetr gal λ(DE3) endA Hte

3.4 Vectors and DNA oligonucleotides.

pET-15b_FOXO4-DBD	Donated from Prof. Obšil group
KV_F (5′-TTG GGT AAA CAA G-3′)	IDT, USA
KV_R (5′-CTT GTT TAC CCA A-3′)	IDT, USA

3.5 Solutions, buffers and medium composition

- 200 mM H_2O_2 (working stock solution): dissolved in H_2O (MS grade). Stored in Aluminium foil. For other dilution, diluted in ammonium acetate buffer 150mM, pH 6.8.

- 75mM Methionine: dissolved in ammonium acetate, 150mM, pH 6.8, filtered via sterile syringe disk filters, degassed.

- Acrylamide solution (30%) for SDS-PAGE: 70% (v/v) H_2O, 29% (w/v) acrylamide, 1 % (w/v) *N,N*-methylene-bis-acrylamide.

- Acrylamide solution (40%) for native gel electrophoresis: 60% (v/v) H_2O, 38% (w/v) acrylamide, 2% (w/v) *N,N*-methylene-bis-acrylamide.

- Ammonium acetate buffer 150mM, pH 6.8: ammonium acetate dissolved in H_2O (MS grade). Adjusted pH to 6.8 with acetic acid. Filtered via sterile syringe disk filters (0.22μm), degassed, saturated by Helium.

- Destaining solution for SDS-PAGE: 55% (v/v) H_2O, 35% (v/v) ethanol, 10% (v/v) acetic acid.

- DNA trap-column stripping solution: 75% MeOH, 1% HFIP 0.025% TEA, adjusted pH to 7.5 with acetic acid.

- LB medium: 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, pH 7.4, sterilized.

- Sample buffer (5×): 50mM Tris-HCl, 12% (v/v) glycerol, 4% (w/v) SDS, 100mM DTT, 0.1% (w/v) bromophenol blue, pH 6.8.

- Solution A for LC-MS analysis: 2% ACN, 0.1% FA.

- Solution B for LC-MS analysis: 98% ACN, 0.1% FA.

- Staining solution for native gel electrophoresis: dissolved 5 μl of GelRed® dye (10000× concentrated) in 50 ml distilled water.

- Staining solution for SDS-PAGE: 45% (v/v) methanol, 10% (v/v) acetic acid, 0.25% (w/v) Coomassie brilliant blue R-250.

- TBE buffer for native gel electrophoresis (10× concentrated): 0.89M tris-HCl, 0.89M Orto boric acid, 0.02M EDTA, pH 8.8, filtered.

- Wash buffer (for pellets solubilization): 20mM phosphate buffer, 500mM NaCl, pH 7.4, lysozyme (10 μg/ml), $MgCl_2$ (20mM), DNAse (300 U), RNAse (300 U), leupeptin (10nM), PMFS (1mM).

- Wash buffer + 30mM/200mM imidazole: 20mM phosphate buffer, 500mM NaCl, 30mM/200mM imidazole, adjusted pH to 7.4, filtered.

- Wash buffer: 20mM phosphate buffer, 500mM NaCl, pH 7.4, filtered.

3.6 Software

ChemDraw® Professional 16.0	Perkin Elmer, USA
Data analysis 5.0	Bruker Daltonics, Billerica, USA
DrawMap	Ref. [178]
GPMAW 8.0	Developed by Suraj Peri[65]
ImageLab software	Bio-Rad Laboratories, USA
Mascot 2.0	Matrix Science, UK
Ms2Links	Developed by Dr. Malin Young, Sandia National Laboratories, California[177]
OriginPro 2015	Origin Lab Corporation, USA
OxIntComp, in-house-built software	Lab 163, IMIC, Czech Republic
Peaks® X+ software	Bioinformatics Solution Inc., Canada
ProteinScape 4	Bruker Daltonics, Billerica, USA

| Pymol 2.1.0 | Schrödinger LLC, USA |
| Syringe pump controller | New Era Pump Systems, Inc., USA |

4. Methods

4.1 Protein expression and purification

The DNA binding domain of protein FOXO4 at the length of 82-207 was chosen as a model protein for the footprinting experiment. The protein is encoded in pET-15b plasmid carried by N-terminal his-tag for further and easier purification. The his-tag sequence also contains a specific cleavage site recognized by thrombin protease. Plasmid was already transformed in *Epicurian Coli* BL21-Gold (DE3) competent cells (Stratagene, USA), which was donated by Prof. Obšil's group. Figure 15 visualizes the full sequence of the produced FOXO4-DBD protein. Please note, that after his-tag removal, the N-terminal part still contains first eight residues as the rest of cleavage site. These residues are numbered from -8 to -1 and the sequence continues with real numbering of FOXO4 starting at G^{82} position. Theoretical monoisotopic $[M+H]^+$ of FOXO4-DBD construct is 16709.80818 Da, or [M]-FOXO4-DBD is 16568.39438 Da, respectively.

MGSSHHHHHHSSGLVPRGSHMLEDPG^{82}AVTGPRKGGSRRNAWGN QSYAELISQAIESAPEKRLTLAQIYEWMVRTVPYFKDKGDSNSSAGWK NSIRHNLSLHSKFIKVHNEATGKSSWWMLNPEGGKSGKAPRRRAASM DSSSKLLRGRSKA^{207}

Figure 15 | The sequence of the expressed FOXO4-DBD protein construct. Green highlighted sequence represents the FOXO4-DBD protein sequence. Black highlighted sequence represents encoded N-terminal his-tag with specific additional sequence (in red), which is recognized and cleaved by thrombin protease. Red arrow represents the cleavage site for thrombin protease.

4.1.1 Overnight culture preparation

Ten µl of transformed cells was directly pipetted from a cryo-can into 5 ml of sterile LB medium containing antibiotic (Ampicillin, final concentration 100 µg·ml⁻¹). The tube was incubated overnight in a shaker at 220 rpm, 37 °C for 16 hours.

4.1.2 Protein expression

One ml of overnight culture (O/N) was transferred under sterile conditions to Erlenmeyer flasks containing 500 ml of sterile LB medium with Ampicillin (concentration 100 µg·ml⁻¹). Four inoculated flasks were incubated at 37 °C at 220 rpm. Optical density was measured in each flask at a wavelength of 600 nm (OD_{600}) to reach 0.6-0.8 value.

Induction of protein expression was reached by addition of IPTG (Isopropyl β-D-1-thiogalactopyranoside) inductor to its final concentration of 0.5mM. The temperature was decreased to 30 °C. The protein expression was carried out in shakers at 30 °C, 220 rpm for 12 hours. After 12 hours shakers were automatically stopped and cooled down to 4 °C.

4.1.3 Cell harvesting and protein isolation

Bacteria were harvested from cell cultures by centrifugation. Bacteria cultures were transferred into 1 l centrifugation bottles and centrifuged at 8 000 × g for 15 minutes. The supernatant was carefully removed and cell pellets from 1 l were resuspended in 20 ml of wash buffer (20 mM phosphate buffer, 500 mM NaCl, pH 7.4) containing lysozyme (10 µg·ml⁻¹), $MgCl_2$ (20mM), DNAse (300 U), RNAse (300 U), leupeptin (10nM) and PMFS (1mM). Resuspended pellets were transferred into small centrifuge tubes and each tube was sonicated on ice for 10 min during the cycle: 1 s sonication, 1 s pause.

In order to separate insoluble membranes and organelles from soluble solution, the tubes were subsequently centrifuged at 70 000 × g, 4 °C, for 70 min. After centrifugation, the protein was present in a cytosolic fraction – intracellular fluid.

The solution was transferred from centrifuged tubes into 50 ml Falcon tubes. To clear the solution from any contamination the lysate was filtered three times via 3 different sterile syringe disk filters with an 0.22µm mesh membrane.

After this step, cell lysate was finally prepared for further affinity purification.

4.1.4 Affinity chromatography

FOXO4-DBD construct contains the N-terminal his-tag. Affinity chromatography was chosen as an optimal step for further protein purification, where Talon® Superflow resin (Clontech Laboratories, USA) was used. Five ml of Talon affinity resin was freshly charged by Co^{2+} according to a manufacturer manual. Table 1 summarizes the procedure of individual steps. The solvent flow was 2 ml·min^{-1}. Absorbance at a wavelength of 280 nm was measured automatically during the whole affinity chromatography.

Fractions of various volumes were collected into individual Falcon tubes. Ten µl of each fraction was further analyzed by SDS-PAGE.

Table 1 | Affinity chromatography purification protocol.

Step	Buffer	Purpose
A.	Wash buffer	Equilibration
B.	Cell lysate	Loading
C.	Wash buffer	Cleaning step
D.	Wash buffer + 30mM imidazole	Elution of proteins bound by non-specific interactions
E.	Wash buffer	Cleaning step
F.	Wash buffer + 200mM imidazole	Elution of FOXO4-DBD
G.	50mM MES	Removing the imidazole
H.	Wash buffer	Equilibration
WASH BUFFER – 20 mM phosphate buffer, 500 mM NaCl, pH 7.4		

A fraction containing the protein of interest was further concentrated on Amicon Ultra concentrators with 3 kDa membrane cut-off (Merck, USA) at 4 °C. The final volume was approximately 3 ml of the initial volume (approximately 10 ml).

4.1.5 His tag removal of the FOXO4-DBD fusion protein

FOXO4-DBD's N-terminal his-tag contains a cleavage site for thrombin protease.

Six hundred µl of protein sample was treated with thrombin (2U per milligram of recombinant protein) for14 hours at 4 °C. After the cleaving reaction, the Eppendorf tube was centrifuged at 13 400 × g for 5 min, 4 °C to separate any precipitated proteins or contaminants and the supernatant was transferred into a new Eppendorf tube.

4.1.6 Size exclusion chromatography (SEC)

SEC was performed on ENrich SEC 70 10 × 300 column (Bio-Rad Laboratories, USA). The column had been previously equilibrated by 100 ml of mobile phase (ammonium acetate, 150 mM, pH 6.8).

Six hundred µl were injected into the loading loop. The constant flow was set at 800 µl·min^{-1}. Eluate was monitored at 280 nm. Column was connected to a fraction collector, where 800 µl per fraction were collected.

After the SEC, 3 fractions corresponding to the highest absorbance of protein concentration according to a spectrophotometer were stored on ice, 10 µl from each of the three fractions were analyzed by SDS-PAGE.

The three collected fractions were pooled into one tube and further concentrated using 0.5ml Amicon Ultra concentrators with 3 kDa membrane cut-off (Merck, USA), at 4 °C to the final volume of approximately 300 µl.

Finally, protein concentration was measured on a spectrophotometer Denovix DS-11 (molar absorption coefficient of FOXO4-DBD, ε_{280} = 32 290 M^{-1}·cm^{-1}). The quality of the prepared protein was further analyzed by native-MS and intact protein analysis.

4.2 Formation and characterization of FOXO4-DBD•DBE complex

FOXO4-DBD, in a native state, is capable of binding 13bp DNA duplex called DAF16.

4.2.1 Double-strand DNA preparation

The solution of 500µM dsDNA was prepared by mixing both single DNA strands, KV_f, and KV_r in molar ratio 1:1 in water, heating at 90 °C for 3 min, and let cool to room temperature to form double-strand DNA.

4.2.2 Assembly of FOXO4-DBD•DAF16 complex

Complex FOXO4-DBD•DAF16 was mixed to the final 30μM concentration of both components. Protein at a concentration of 170 μM and the dsDNA at a concentration of 500 μM were mixed in 1:1 molar ratio and diluted into 30μM concentration with ammonium acetate, 150mM, pH 6.8.

The complex remained at laboratory temperature for 10 min and was then stored on ice. In order to confirm complex formation, native electrophoresis and native-MS was used. Four μl of a sample was used for native gel electrophoresis. For native-MS, 2 μl of protein·DNA sample was diluted into 8 μl of ammonium acetate, 150mM, pH 6.8 buffer, and gently mixed.

4.3 Gel electrophoresis

4.3.1 Polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate (SDS-PAGE)

An SDS-electrophoresis gel consists of two parts – a stacking gel and a separating gel. Separation gel can have various concentrations and stacking gel commonly has 5% concentration of acrylamide. For our purpose, 15% separation gel and 5% stacking gel were prepared. Table 2, p. 56 describes the composition of individual gels used for SDS-PAGE. The apparatus was filled by 1× concentrated running buffer (TGS buffer, Bio-Rad Laboratories, USA). The voltage for this technique was set up to 200 V and the electrophoresis had been running for approximately 90 min.

All samples were prepared in the same way. Ten μl of sample was mixed with 2 μl of 5× concentrated sample buffer. To fully denature proteins, samples were mixed, sonicated for 3 min, then heated at 95 °C for 5 min. In order to identify the proteins' relative molecular weight, 10 μl of protein marker (Bio-Rad Laboratories, USA) was loaded on the gel as well.

For protein visualization, gels were stained in staining solution containing Coomassie Brilliant Blue R-250 dye for 10 min. Right after staining, gels were washed with distilled water. Gels were subsequently destained by overnight incubation in a destaining solution. All gels were visualized using a ChemiDoc imaging CCD

camera (Bio-Rad Laboratories, USA) against a white background. For long-term storage, gels were stored at laboratory temperature in 1% acetic acid.

Table 2 | List of individual components used for preparation of separating (15% acrylamide) and stacking (5% acrylamide) gels.

Component	The volume of individual components (ml)	
	Separating gel (15%)	Stacking gel (5%)
H_2O	1.1	0.68
30% Acrylamide	2.5	0.17
1.5M Tris (pH 8.8)	1.3	-
1.0M Tris (pH 6.8)	-	0.13
10% Sodium dodecyl sulfate	0.05	0.01
10% Ammonium persulfate	0.05	0.01
TEMED	0.002	0.001

4.3.2 Native gel electrophoresis

Native gel electrophoresis was used to confirm the presence of the FOXO4-DBD•DAF16 complex. To run native gel electrophoresis, a fresh 15% acrylamide gel was prepared. Table 3 describes all components used for gel preparation.

Table 3 | List of individual components used for a native gel electrophoresis

Component	The volume of individual components (ml)
	Separating gel (15%)
H_2O	5
40% Acrylamide	4
10×TBE (pH 8.8)	1
10% Ammonium persulfate	0.08
TEMED	0.008

The volume corresponding to 2 μg of DNA was pipetted for ssDNA, dsDNA, and complex samples. Ammonium acetate was added to the final volume of 10 μl. One and half μl of loading dye (6× concentrated) was added to each sample. Freshly prepared solutions were immediately applied into the wells, apparatus was closed, and native electrophoresis had been running in the presence of 1× TBE buffer for 40 min at 4 °C. The voltage was set to 150 V per gel.

After 40 minutes, the gel was pulled out from the apparatus, and DNA was visualized by staining in GelRed® dye (Biotinum, USA) for 15 min in the presence of 50 ml of 10 000× diluted solution. The gel was visualized using a ChemiDoc imaging CCD camera (Bio-Rad Laboratories, USA) against a black background.

4.4 Fast photochemical oxidation of proteins

4.4.1 FPOP quench flow setup

Flow reactor comprised syringe pumps (New Era Pump Systems, Inc., USA), syringes (Hamilton, USA) and fused silica capillaries (Polymicro Technologies, USA), (Figure 16, p. 58). Syringe one at volume of 250 μl contained protein/protein·DNA sample. A silica capillary 17 cm long with an inner diameter of 150 μm led the sample into a mixer. The second syringe at volume of 500 μl contained hydrogen peroxide and was connected with a mixer by capillary 15 cm long and with an inner diameter of 100 μm. Sample and hydrogen peroxide were mixed and they were further led by a capillary of an overall length 13.7 cm and an inner diameter of 75 μm to be mixed with a quencher. After a 7.0 cm of flow pathway a transparent window was created by removing the silica protective layer, and the solution was subsequently irradiated at a width of a laser beam of 0.75 cm and led through the rest of 5.95 cm of silica capillary to be quenched. The third syringe contained 1ml of methionine as a quencher, which was led through a capillary 17 cm long and with an inner diameter of 75 μm and quenched the irradiated solution. The quenched solution was subsequently collected into an Eppendorf tube. The flow rate of sample and hydrogen peroxide syringe was set to 10 μl·min^{-1}, the third syringe was set to 20 μl·min^{-1}.
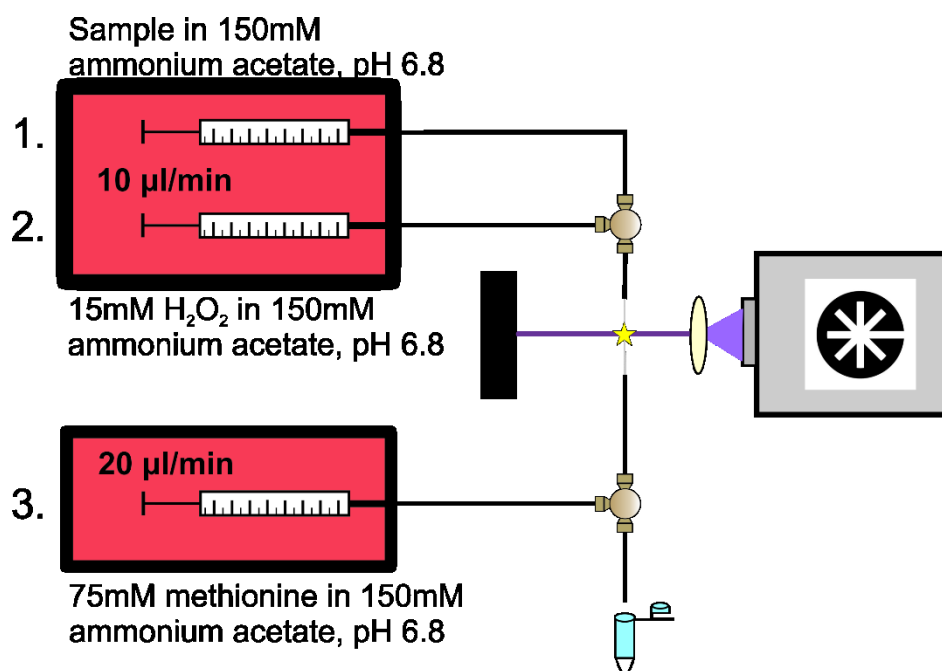
Figure 16 | Scheme of a general FPOP quench flow system.

In the beginning, the excimer laser was turned on and allowed 5 min to heat up, according to the manufacturer's manual. During the irradiation experiments, the laser was cooled with a continuous water-cooling circuit. Excimer laser at a wavelength of 248 nm was set to different frequencies and energies. Before and after each experiment, syringes were filled with water (LC grade) and the flow reactor was washed within the continual flow for 20 minutes to avoid carry-over.

4.4.2 FPOP

Three syringes filled with appropriate solutions were used in the FPOP experiment. Syringe one was filled with a protein/protein·DNA sample at a concentration of 0.4 mg·ml^{-1} and volume of 250 µl. The second syringe was filled with 500 µl of 15mM hydrogen peroxide. The third syringe was filled with methionine at a concentration of 75 mM and volume of 1 ml. All solutions were diluted in 150mM ammonium acetate, pH 6.8, which was freshly degassed.

For FPOP optimization, an excimer laser was set to frequencies 15 Hz and 20 Hz, and energies 15 mJ, 20 mJ, 30 mJ, and 40 mJ were set for each frequency.

After 1 min of solution mixing, an Eppendorf tube was placed to collect a reaction mixture with no laser irradiation for 1 min. The laser was turned on in an appropriate frequency and energy, following 30 sec of mixing the solution and 1 min for sample

collecting. Overall, 5 μg of sample was collected during one minute of collection. After one minute, the laser was turned off, the energy (or frequency) was changed followed by laser being turned on, 30 sec of mixing, and 1 min of sample collecting.

After optimization, the FPOP experiment was carried out for specific frequency 20 Hz and energy 30 mJ. Ten μg of the sample was collected without laser irradiation, which was subsequently subjected to sample desalting and intact protein analysis. Overall, 40 μg of the irradiated sample was collected, of which 10 μg were further desalted and analyzed by top-down fragmentation. 30 μg of the sample was proteolytically digested for further bottom-up analysis.

Between both experiments, syringes were cleaned by water and the reactor was let to wash under continuous flow for 20 min to avoid carry-over.

4.5 Sample preparation for mass spectrometry

4.5.1 Digestion of samples

LysC (Promega, USA) or a combination of LysC/Trypsin (Promega, USA) was added to an irradiated sample containing 15 μg of protein to the final ratio protease:protein 1:20. After incubation at 37 ºC for 2 hours, another load of the same protease was added in a protease:protein 1:20 ratio to the final concentration ratio of 1:10. The mixture was incubated for 4 hours at 37 ºC.

The reaction was quenched by the addition of TFA to a final concentration of 0.1%. All samples were diluted in solvent A (2% ACN, 0.1% TFA) to a final concentration of 0.5 μg·μl^{-1} and subsequently transferred to vials for further LC-MS analysis.

4.5.2 Sample desalting

Protein samples were desalted for further mass spectrometric analysis. Desalting was performed on a protein reverse-phase microtrap column (Optimize technologies, USA) with a capacity of 20 μg of total protein sample. Table 4, p. 60 describes the desalting procedure in detail.

Table 4 | Detailed description of desalting procedure for intact protein analysis and fragmentation.

Step	Buffer	Purpose
1.	Sample mixed with 100 µl of 0.1% FA	Sample denaturation
2.	1× 250 µl of 80% ACN, 0.1% FA	Cleaning the column from the previous usage.
3.	3× 250 µl 0.1% FA	Equilibration
4.	Loading the sample	Protein should interact with the column resin
5.	3× 250 µl 0.1% FA	Desalting the sample
6.	50 µl of 80% ACN, 0.1% FA	Elution of protein
7.	1× 250 µl of 80% ACN, 0.1% FA	Column cleaning
8.	2× 250 µl 75% MeOH, 1% HFIP 0.025% TEAA, pH 7.5	DNA stripping
9.	3× 250 µl 0.1% FA	Equilibration

The sample has been eluted into a new Eppendorf tube and stored at -20 °C for further usage.

4.6 Mass spectrometry measurements

All measurements were performed using a 15T-SolariX XR FT-ICR mass spectrometer (Bruker Daltonics, Billerica, USA), which was calibrated by using a solution of sodium trifluoroacetate (NaTFA) to achieve at least 1 ppm accuracy.

4.6.1 Native mass spectrometry

For a native-MS, an in-house-built quartz capillary was prepared using a laser puller. The orifice diameter of the capillary was less than 1 µm. The quartz capillary was filled with 10 µl of sample and put into a holder. The capillary was shortly centrifuged in the holder (15 sec, 2500 × g) to avoid any air bubbles and placed into an in-house built nESI holder. Steel wire was placed into the solution inside the capillary. The tip

of the capillary was gently set closely in the proximity of the entry of the ESI source. Voltage source was turned on and high voltage was set between 600-700 V. The capillary tip and position in front of ESI source is shown in Figure 17.

Mass spectra were acquired in a positive mode using nESI source, over the 200-3500 m/z range. To achieve a higher m/z range of ions a higher accumulation time of ions (1 s) and time of flight (2.0 ms) was applied. The temperature of the glass capillary was 25 °C. Collision voltage was -2.5V and data were stored with 1M data points transient accumulating 128 scans.
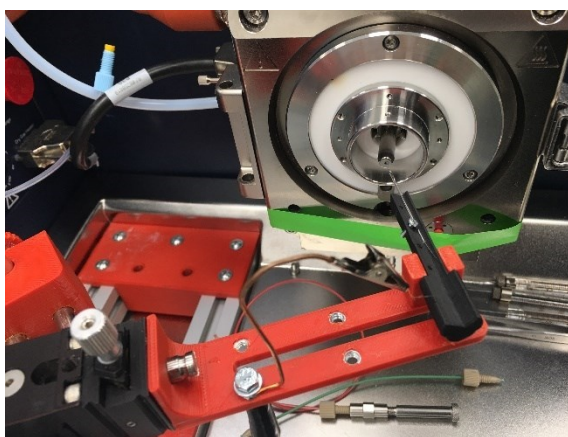
Figure 17 | A view at an in-house built nESI holder with a capillary tip. The tip is placed in front of the ESI source.

4.6.2 Intact protein analysis

Desalted samples were diluted 2× or 5× using 50% MeOH, 0.1% FA solution and sprayed using nESI source. Mass spectra were acquired in a positive mode over 200-2500 m/z range. Collision voltage was set to -2.5V using 0.1 s accumulation of ions in collision cell with a 1M data points transient. The temperature of the glass capillary was 200 °C. Spectra were acquired within 128 scans collection.

4.6.3 Top-Down mass spectrometry

All samples were sprayed using nESI source for the top-down approach. Samples were diluted (5× or 10×) using 50% MeOH, 0.1% FA solution. After data acquisition of broad m/z range, isolation of individual charge states was achieved.

The charge state of interest was isolated using a broader isolation window. The center of isolation was set up with respect to the distribution of all modifications. Isolation of +16 charge state was achieved by setting up the center of isolation in the center of isotopic distribution. Isolation window was set as ±0.7 Da for unmodified ion and ±0.9 Da for singly oxidized ion, respectively. Using the accumulation of ions in the collision cell for 1.0 s, time of flight 1.1 s, collision voltage 2.5 V, 1M data point transient, 128 scans were recorded. In order to fragment an unmodified or singly oxidized ion, electron-capture dissociation (ECD) or collision-induced dissociation (CID) was used.

Collision induced dissociation (CID). After the stable isolation of +16 unmodified/singly oxidized ion, collision voltage 15 V was applied, and CID was performed. Data were acquired using 1M data point transient, accumulation of ions for 5 s, and 128 scans and spectra recorded as a technical triplicate.

Electron capture dissociation (ECD). Unmodified ion/+16 Da modification was isolated using appropriate *m/z* and isolation windows. ECD cathode was turned on to heat up. After the required time, ECD parameters were slowly increased (ECD pulse 0.075 V, bias: 0.7 V and lens: 13 V) to achieved protein fragmentation with respect to preserving both, a precursor ion in the ICR cell and S/N ratio. Spectra were recorded as a technical triplicate by 128 scans data acquisition, time of flight 1.1 s, collision voltage -5 V, 1M data points transient and accumulation of ions for 15 s.

4.6.4 Bottom-up mass spectrometry

HPLC separation was performed using an Agilent 1200 series HPLC system (Agilent Technologies, USA). The sample was loaded twice in a total volume of 1 μl. The sample was injected onto a reverse-phase trap column (Luna® Omega Polar C18, 0.3 × 30 mm, Phenomenex, USA) followed by reverse-phase analytical column (Luna® Omega Polar C18, 0.3 × 30 mm, Phenomenex, USA), both heated to 60 °C. Two solvents were used: A (2% ACN, 0.1% TFA), and B (98% ACN, 0.1% FA). The initial flow rate was set to 10 μl/min with the following solution composition: 95% A, 5% B. The LC run consisted of a 35-minute separation gradient of 5-40 % of solvent B (98% ACN, 0.1% FA), a 3-minute spike of 40-95% of solvent B, 3-minute

washing flow at 95% solvent B, a 1-minute drop of 95-2% of solvent B and equilibration of columns in 2% B for 10 min.

HPLC system was directly connected to the FT-ICR mass spectrometer. Analyses were performed in the data-dependent analysis (DDA) and mass spectra were acquired in *m/z* range 250-2500, with 1M data points transient, ion accumulation 0.2 s. Four scans were accumulated per spectrum. Three MS/MS spectra were obtained after one mS scan. To quantify the ratio between unoxidized and oxidized modification, LC-MS analysis was performed as a technical triplicate in the same aforementioned conditions.

4.7 Data analysis/processing

4.7.1 Bottom-up data analysis

LCMS data were processed using DataAnalysis 5.0 (Bruker Daltonics, Billerica, USA). First, an *in-silico* protease digest of FOXO4-DBD sequence was obtained using GPMAW 8.0 (ref.[65]).

Using LC-MS/MS analysis, data were processed using Peaks® X+ software. The software precisely searched for predefined modifications[100,120] and these matches were manually confirmed in spectra for each peptide. Masses of regular and modified peptides were searched in LC-MS chromatograms.

To quantify the extent of modification for each peptide, an extracted ion chromatogram was performed. Extracted ion chromatogram with *m/z* ± 0.005 Da precision was used for each peptide, both unoxidized and oxidized. The same procedure was used for all samples, including a technical triplicate of each sample and control sample.

$$\text{Extent of modification} = \frac{\sum I_{OX}^p}{\sum (I_{OX}^p + I^p)} \qquad \text{(eq. 22)}$$

An individual intensity of monoisotopic peptides was taken for each unoxidized and oxidized peptide. The extent of modification for the two most intense charge states was calculated and then an arithmetic average was created. To calculate the extent of modification, equation 22 was used, where $\sum I_{OX}^p$ is the sum of intensities of two

charge states of modified peptide and $\sum(I_{OX}^p + I^p)$ is the sum of intensities for two charge states of both unmodified and modified peptide. The data were plotted in OriginPro 2015 and presented in the form of *mean ± SD*. In order to determine a significant difference for each modification between both forms, data were statistically analyzed by unpaired t-test using GraphPad Prism 8 software. Residues which were found to be oxidized are highlighted in the crystal structure (Figure 33, p.85 and 34, p. 86, respectively) using Pymol 2.1.0 software (Schrödinger LLC, USA).

4.7.2 Intact and top-down data analysis

DataAnalysis 5.0 (Bruker Daltonics, Billerica, USA) was used to process top-down and intact spectra. MS spectra were deconvoluted using Sophistical Numerical Annotation Procedure method (SNAP, Bruker Daltonics) with quality factor threshold set to 0.3 and S/N threshold set to 0. The maximum charge was set to 30 for intact protein analysis, and 15 for native-MS.

Processing of fragmented spectra. Top-down raw spectra were deconvoluted using SNAP 2.0 algorithm (DataAnalysis 5.0) and re-calibrated using high-intensive fragments with well-known monoisotopic *m/z.* Theoretical monoisotopic masses were obtained by *in-silico* fragmentation of FOXO4-DBD sequence in GPMAW 8.0[65] software.

After re-calibration, a *mascot generic file* was exported for each spectrum and further used to generate a list of unmodified fragment ions using MS2Links software[177].

To get an input file for OxIntComp, oxidized protein fragment spectra were exported as x,y coordinates into a text file containing individual masses and their intensities using DataAnalysis 5.0. In this case, spectra were recalibrated using an S/N threshold set to 2, relative intensity threshold as 0.01 %.

$$\text{Extent of modification} = \frac{\sum I_{OX}^f}{\sum(I_{OX}^f + I^f)} \qquad \text{(eq. 23)}$$

Last but not least, both MS2Links file and .txt file generated by FTMS algorithm were used and uploaded into an in-house-built OxIntComp software for searching, visualizing and exporting individual fragments and its oxidation. The software searched for 1 modification (+16 Da) in spectra of oxidized ions against spectra of unmodified ions within 3ppm accuracy.

Data were quantified using equation 23, p.64, where $\sum I_{OX}^{f}$ is the sum of intensity of modified fragment and $\sum (I_{OX}^{f} + I^{f})$ is the sum of intensities of both unmodified and modified fragments. The data were processed as a technical triplicate presented in form of *mean ± SD* and plotted in OriginPro 2015 software. In order to determine a significant change between both forms, the extent of modification of holo form was subtracted from the extent of modification of apo form for each fragment. The regions of significant change between both forms were further highlighted in the crystal structure (Figure 39, p. 95) using Pymol 2.1.0 software (Schrödinger LLC, USA).

5. Results

5.1 Protein expression and characterization

5.1.1 Expression, isolation and purification of FOXO4-DBD

Solution of competent cells was transferred into sterile LB medium of a small volume to create an O/N culture. Next day, LB medium was inoculated by O/N culture, and cells were let to grow, until the protein expression was induced by IPTG to produce FOXO4-DBD.

After a required time, cells were harvested from LB medium, resuspended in Wash buffer, disrupted by sonication, and centrifuged to separate insoluble components from cell lysate. Subsequently, cell lysate containing FOXO4-DBD was loaded onto a Talon® Superflow resin, where protein was washed and eluted. Figure 18 visualize procedure of affinity chromatography, where fraction F should contain eluted protein. Fractions B-F were collected into individual Falcon tubes, stored on ice and samples analyzed by SDS-PAGE (Figure 19, A, p. 67).
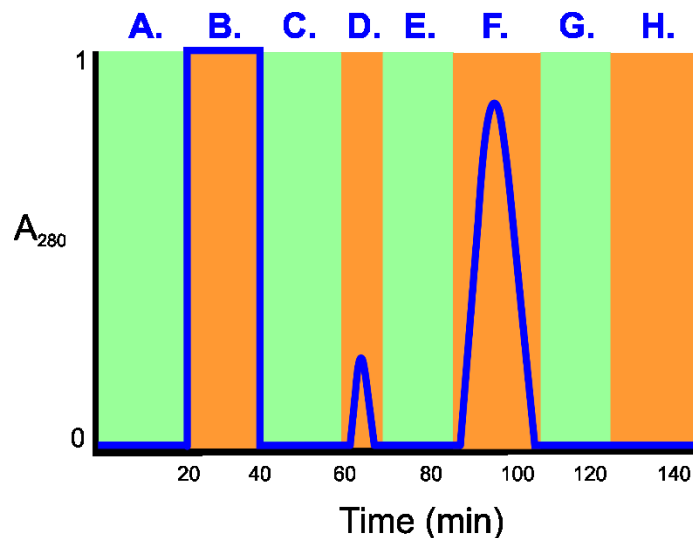

Figure 18 | Record of FOXO4-DBD affinity purification. A – equilibration of Talon® resin by Wash buffer, B – loading of cell lysate, C – washing by Wash buffer, D – 30mM imidazole, E – washing by Wash buffer, F – elution step using 200 mM imidazole, G – 50 mM MES, H – equilibration by wash buffer. Blue line represents signal detected by spectrophotometer at a wavelength of 280 nm.
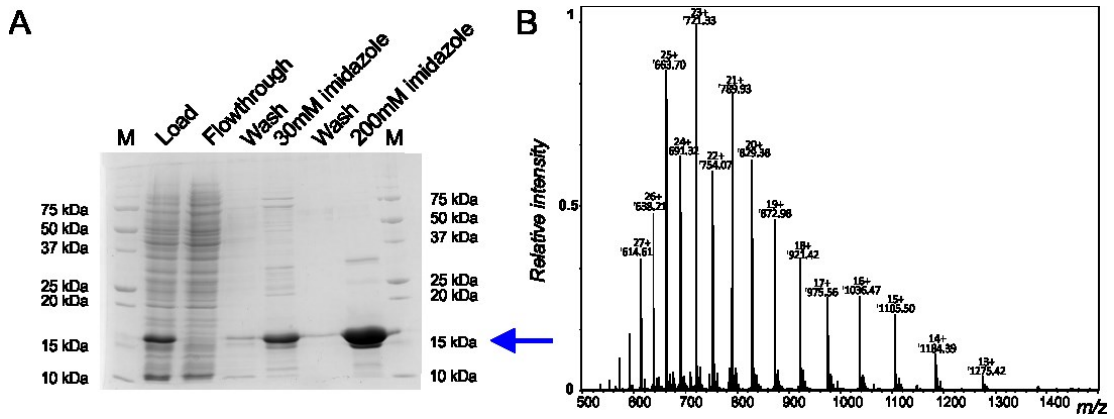
Figure 19 | **A**. SDS-PAGE of collected fractions from affinity chromatography. M - Precision Plus Protein™ standard marker, Load - load of cell lysate. Flowthrough - cell lysate lacking a recombinant protein. Wash - wash by wash buffer (20 mM phosphate buffer, 500 mM NaCl, pH 7.4). Elution of protein was proceeded by the addition of 200mM imidazole into the washing buffer. **B**. Deconvoluted MS spectrum of expressed FOXO4-DBD in range of 500-1500 *m/z* using an FT-ICR instrument.

According to Figure 18, p. 66 the eluted fraction F should contain a protein of interest. Performed SDS-PAGE (Figure 19, A) of fractions B-F revealed that the eluted fraction F contains protein at molecular weight, approximately 17 kDa. This molecular weight and mobility on gel of eluted protein corresponds to a molecular weight of FOXO4-DBD, almost 17 kDa. To confirm the presence of FOXO4-DBD in fraction F, sample was desalted, 2× diluted in spraying solution and submitted to MS analysis. Theoretical $[M+H]^+$ molecular wight of mass [M]-FOXO4-DBD is 16568.39438 Da. Deconvoluted mass spectrum determined $[M+H]^+$ as 16568.39576 with 0.09 ppm error and this can be considered, that the FOXO4-DBD to be successfully expressed and isolated.

Eluted fraction was further concentrated to 2.5 mg/ml and volume of approximately 3 ml. Final FOXO4-DBD yield from 1l of LB medium was 3.5 mg of recombinant protein.

5.1.2 His-tag removal and gel filtration of FOXO4-DBD

After an overnight cleaving reaction with thrombin, sample was subsequently centrifuged and gel filtration was performed. Figure 20 shows gel filtration record.
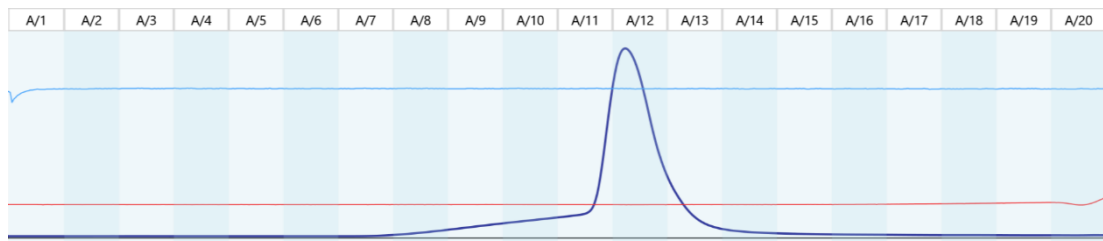
Figure 20 | Record of the gel filtration. Blue line represents a FOXO4-DBD elution absorbance measured by spectrophotometer at wavelength of 280 nm (A_{280}). Red line represents a conductivity of salts/buffer and light blue line represents the pressure on a separation column.
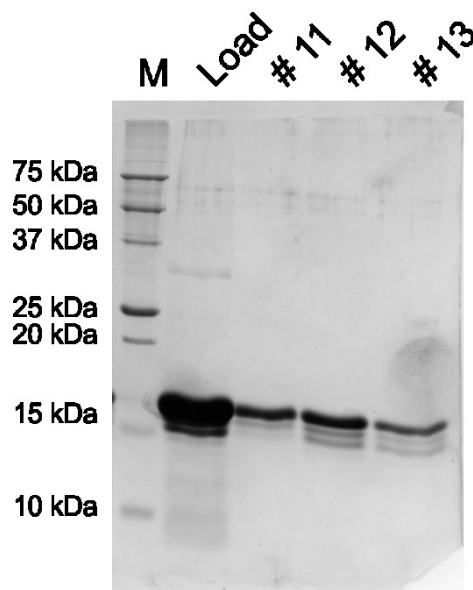
Figure 21 | SDS-PAGE of three fractions collected after SDS-PAGE. As a protein marker (lane M) has been used Precision Plus Protein™ standard marker.

Procedure of gel filtration was recorded (Figure 20), and visualizes one major eluted protein, which was collected in fractions #11-#13. These fractions should correspond to yielded FOXO4-DBD protein. SDS-PAGE of these fractions in Figure 21 revealed, that the gel filtration of sample was successful and fully removed any other protein contaminants from FOXO4-DBD sample. FOXO4-DBD molecular weight should be ~15 kDa. However, electrophoretic mobility of FOXO4-DBD

observed the same ~17 kDa molecular weight with any change in electrophoretic mobility. Therefore, we decided to examine the accurate molecular weight of FOXO4-DBD by MS analysis.

All three fractions (#11, #12, #13) were collected into one solutions and further concentrated at 4 °C into a protein concentration 2 mg·ml^{-1} and volume of 300 µl. Approximately 1.5 mg of recombinant protein was mixed with thrombin, and 600 µg of pure recombinant protein was obtained. More than 50 % of original protein were lost during gel filtration and further concentration, yet still enough to be further characterized and use for FPOP.

5.1.3 FOXO4-DBD characterization

After gel permeation chromatography and final protein concentration, protein sample was desalted, 5× diluted in spraying solution and submitted to MS analysis (Figure 22). Theoretical [M+H]$^+$ mass is 14818.5635 Da. Deconvoluted mass spectrum determined [M+H]$^+$ 14818.5568 with 0.45ppm error and that confirms the FOXO4-DBD to be successfully expressed and prepared.
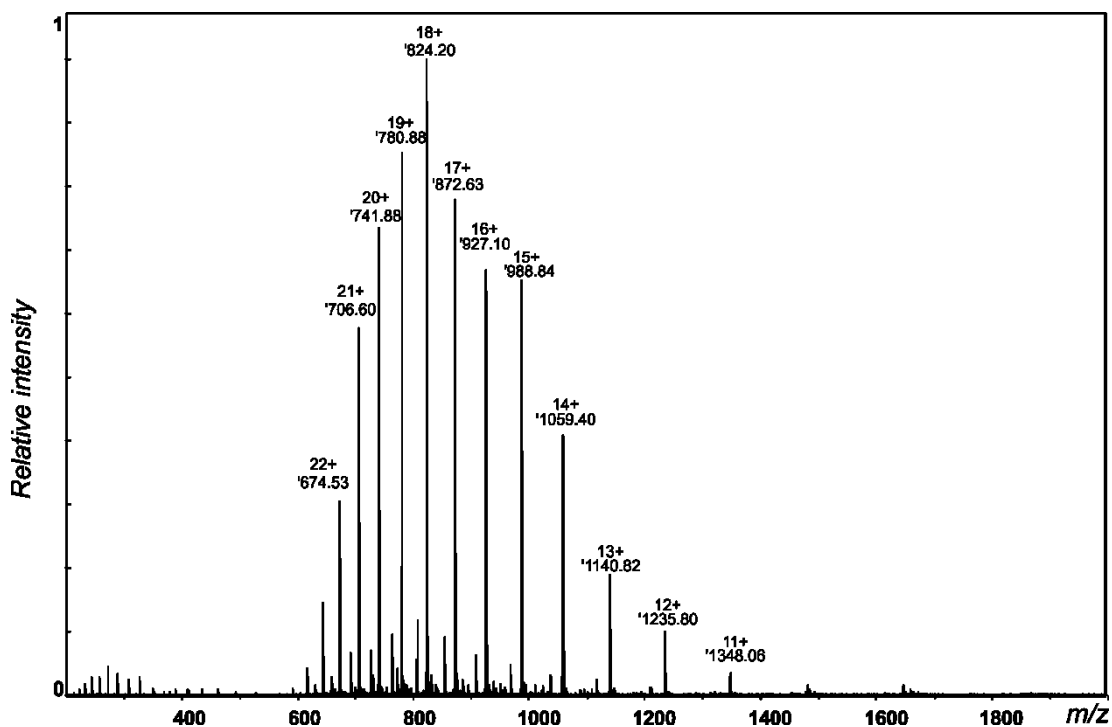
Figure 22 | Deconvoluted mass spectrum of FOXO4-DBD protein obtained using FT-ICR MS instrument.

5.1.4 Complex FOXO4-DBD•DAF16 characterization

FOXO4-DBD and DAF16 duplex was mixed in equimolar ratio 1:1 and further characterized. Overall, 5× diluted samples were sprayed in the presence of ammonium acetate using an in-house built nESI instrument. Sample was further analyzed by native gel electrophoresis to further confirm the presence of complex.
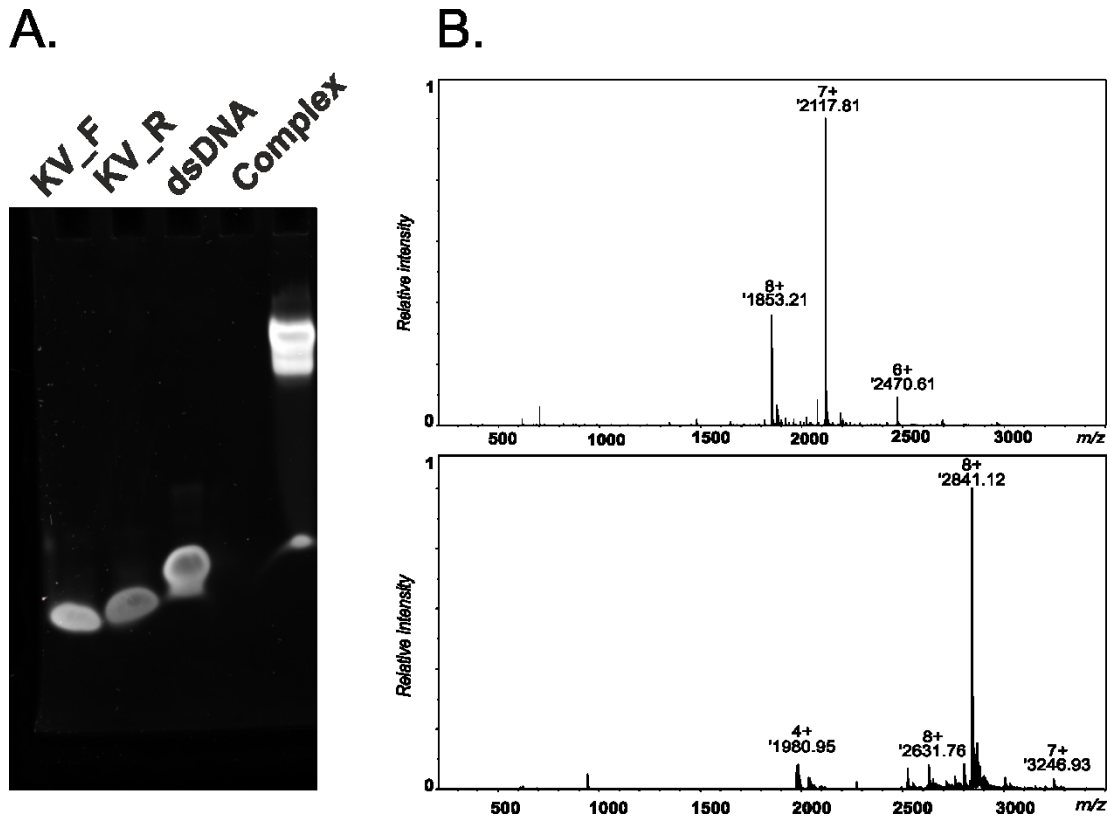
Figure 23 | **A.** Native gel electrophoresis confirming the FOXO4-DBD•DAF16 complex formation. **B.** Deconvoluted mass spectra of native free form (FOXO4-DBD, up) and native bound form (FOXO4-DBD•DAF16, down) obtained by using FT-ICR MS instrument.

Figure 23, A shows native gel electrophoresis. Both forward and reverse single strands report the same mobility, but the duplex DNA reports slower mobility compared to single strands, of which the presence of duplex DAF16 is confirmed. By binding of protein to DNA, a huge mobility shift is made between duplex DNA and protein·DNA sample, which confirmed the complex formation, where protein slowing down the mobility of duplex DNA. Last but not least, Figure 23, B shows native mass spectra of both FOXO4-DBD (up) and FOXO4-DBD•DAF16 complex (down), respectively. In the presence of ammonium acetate, FOXO4-DBD report

lower charge states in mass spectra, where the most abundant is +7charge state, compared to ~+20 charge state of denatured protein. This can be assumed, that FOXO4-DBD exists in solution in native state. Compared to molecular weight of almost 15 kDa FOXO4-DBD, DAF16 element has molecular weight only 8 kDa. In summary, 23 kDa mass of FOXO4-DBD•DAF16, was observed in native spectra predominantly +8 charged. Both molecular weight of protein·DNA complex and lower charge state of complex confirming the presence of native FOXO4-DBD•DAF16 complex.

5.2 Fast photochemical oxidation of protein (FPOP)

Before the FPOP experiment itself, it was necessary to optimize the laser energy and frequency to cause an optimal extent of oxidations. Experiment was provided on both FOXO4-DDB (apo form) and its complex with DAF16 (holo form). The irradiation of free and bound forms was performed at different frequencies (15 Hz and 20Hz) and energies (15, 20, 30 and 40 mJ), samples were collected, desalted and sprayed in ESI positive mode using a nESI instrument. The mass spectra of sample, sample non-irradiated by laser and irradiated samples at frequency of 20 Hz shows Figure 24, p. 72, and at frequency of 15 Hz shows Fig. S1, which can be found in supplementary material.

Irradiation at frequency 15 Hz did not cause rapid increase of modifications with the increasing energy. Frequency of 20 Hz was found to be more efficient. Frequency of 20 Hz and energy 30 mJ was found to cause an optimal extents (Gaussian distribution) of oxidation for both form. For frequency 20 Hz, this energy is visualized in Figure 24, p. 72 in purple for apo form, and in orange for holo form. At the defined frequency of 20 Hz and energy of 30 mJ, FPOP of protein and protein·DNA complex was performed, where an increased amount of sample was collected. Oxidative modifications of both apo and holo forms at defined conditions were further examined by mass spectrometric analysis.
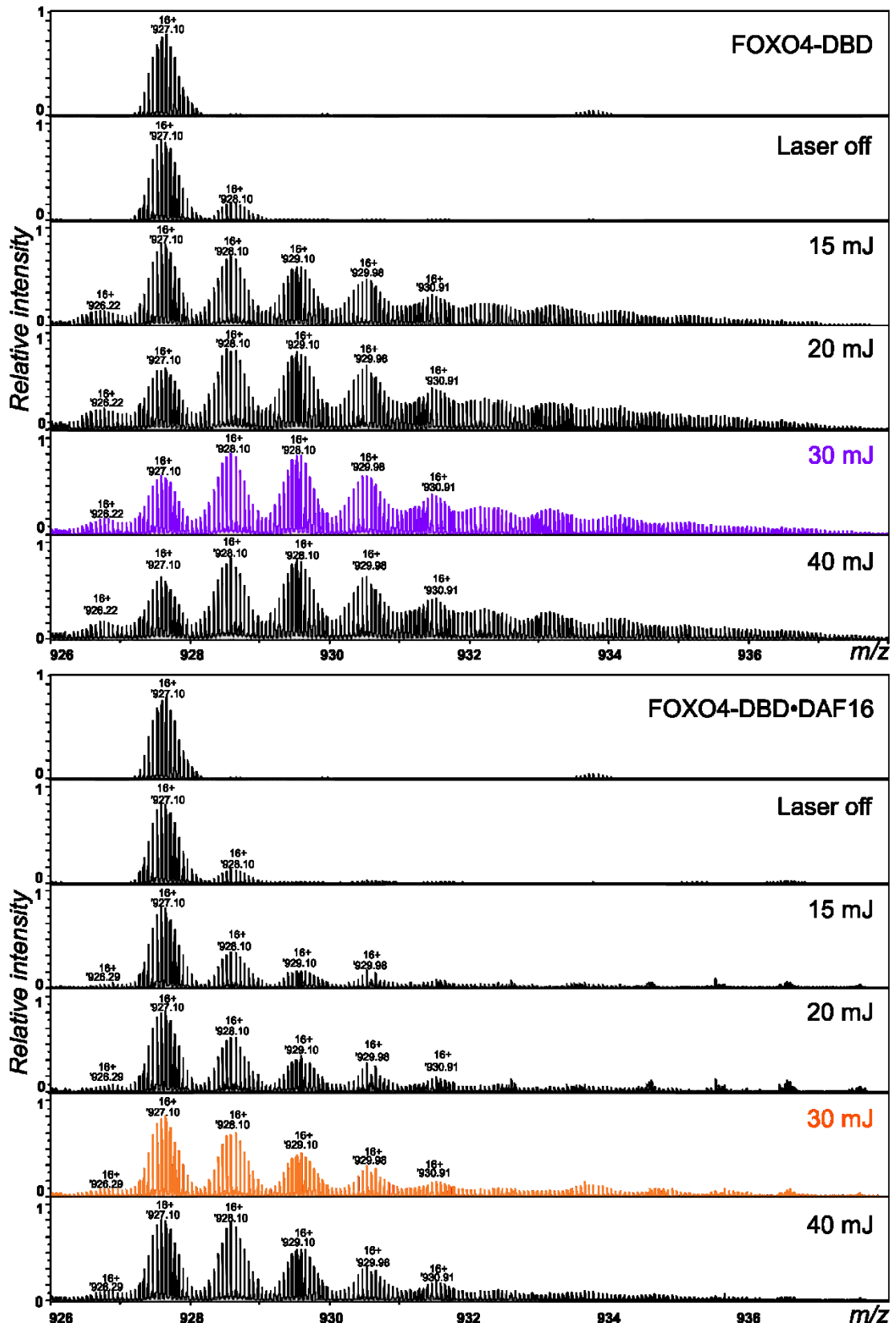
Figure 24 | Mass spectra of both, FOXO4-DBD and FOXO4-DBD•DAF16 submitted FPOP experiment at the frequency of 20 Hz. Spectra of the energy 30 mJ are highlighted in purple (apo form) or in orange (holo form), respectively.

5.3 Mass spectrometric analysis

Oxidized samples obtained by FPOP have been submitted to bottom-up and top-down analysis.

5.3.1 Bottom-up analysis

Irradiated samples undergo proteolysis by LysC or combination of LysC/Trypsin. Samples undergo LC-MS/MS analysis. After data interpretation, seven peptides were detected in both LysC and LysC/Trypsin digest covering an 82 % and 66 % of FOXO4-DBD sequence, respectively. Figure 25 shows the peptide map obtained by proteases digestion completed by the additional FOXO4-DBD secondary structure topology. Moreover, 24 and 27 modifications were found for FOXO4-DBD in LysC and LysC/Trypsin digest, respectively. LC-MS/MS analysis enabled the precise localization of oxidative modifications on each peptide. Collision spectra of each localized modification can be found in supplementary material in Figures S2-S5.
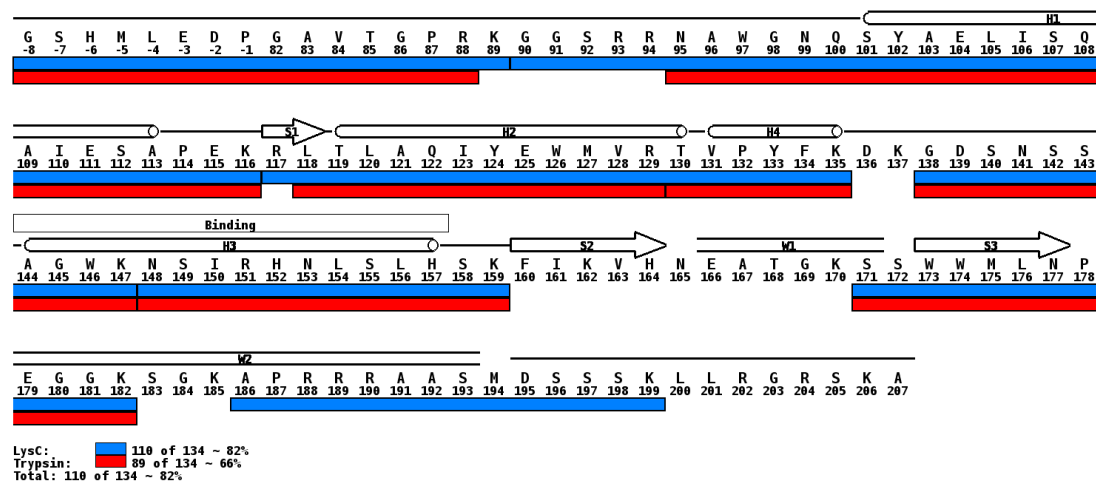
Figure 25 | Peptide coverage map obtained by proteolytic digestion of FOXO4-DBD. Blue lines represent peptides obtained by LysC digestion, red lines digestion by LysC/Trypsin. Created in Draw Map[178], web-based software of MSTools package.

Additional to LC-MS/MS analysis, LC-MS was performed in technical triplicate to quantify the ratio between modified and unmodified peptides (Figure 26, p. 74). The extents of conversion of individual modified residues were quantified using the equation 22, p. 63 for apo and holo form, respectively. Thus, each result was verified by unpaired t-test, which assigned an asterisk to each modified residue a significant difference between apo and holo form, respectively.

First, singly and doubly oxidized residues were observed: (H-6), (M-5), (M-5)+2O, W97, Y102, Y102+2O, Y124, W126, W126+2O, M127, Y133, H152, H157, W173, W173+2O, W174, M175, M194. However, hydroxyl radicals introduced also other modifications than +16 Da/+32 Da modification, respectively. Namely, it was observed histidine conversion to aspartate (-22 Da) at (H-6), H152 and H157 residues, (L-4) residue oxidation to ketone form (+14 Da), and deguanidylation (-43 Da) of arginine R94 residue. LC-MS separation could effectively separate also oxidized isomers. The residues W97, Y102 and W173, and their isomers are highlighted by "#" to be distinguished in the Figure 26.

From overall 16 modified residues and 30 observed modifications, 3 residues undergo no significant difference either in apo or holo forms: (H-6) and (H-6) oxidation to aspartate, (L-4) oxidation to carbonyl, W97/W97# oxidation, as well as doubly oxidized (M-5) and double oxidation of (H-6), (M-5) residues. Some residues were observed to be oxidized at different positions (W97/W97#, Y102/Y102#, W173/W173#). In holo form, some residues were affected by the presence of DNA and their exposure to solvent was decreased. These residues are: Y102(Y102#), Y124, W126, M127, H152, H157, W173#,W173, W174, M175and doubly oxidized W126 and W173 residues. Various oxidized modifications of residues W173, W174 and M175 were observed (W173 and M175, W174 and M175, W173+2O and M175). DNA also effected some residues by increasing their exposure into solvent, which were found to be more oxidized in holo form: (M-5), R94 (-Gnd), Y133, W173, M194.
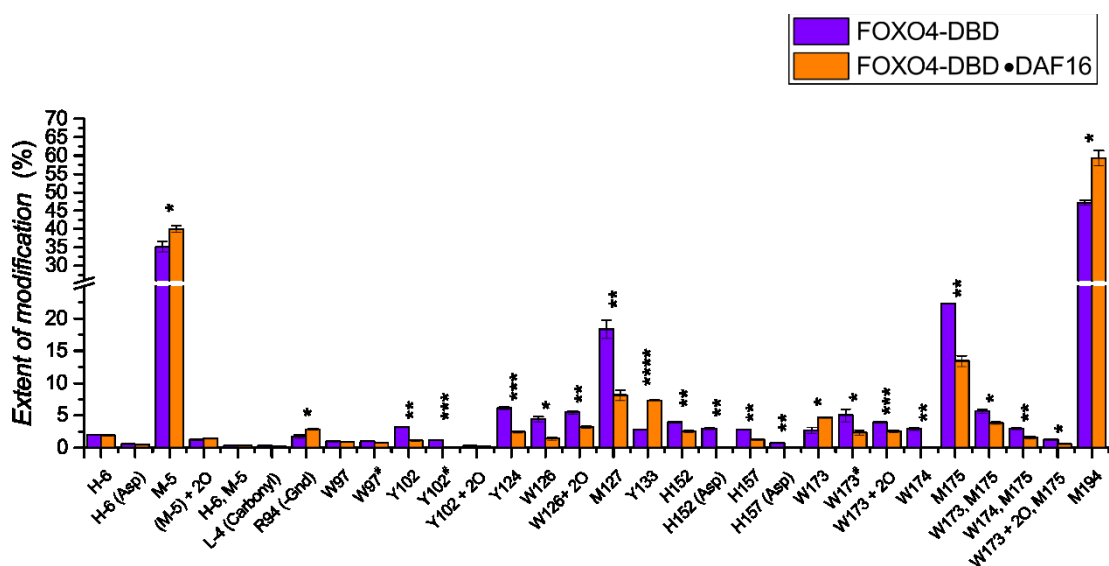
Figure 26 | Histograms represents quantified extents of modifications obtained by bottom-up analysis. Purple bars represent FOXO4-DBD, orange bars FOXO4-DBD

in the presence of DNA. Modifications different than +16 Da are closer described in brackets. Modifications marked by [#] are oxidations quantified on different positions of that residue All bars are presented in the form of *mean ± SD*, obtained by a technical triplicate measurements. Statistical analysis of unpaired t-test was provided for each residue. Significant difference between free and bound form is indicated above the histograms by number of asterisks. T-test legend: * (P ≤ 0.05), ** (P ≤ 0.01), *** (P ≤ 0.001), **** (P ≤ 0.0001).

5.3.2 Top-down analysis

5.3.2.1 Intact protein fragmentation

Intact protein charge state +16 at *m/z* 927.6 Da was isolated and further fragmented by CID and ECD to obtain individual fragments across the protein molecule. Raw spectra were processed and deconvoluted to $[M+H]^+$ ions and the lists of deconvoluted ions obtained by both fragmentation methods can be find in supplementary material as a Table S1. A 27 %, and 60 % of fragment coverage has been reached for CID (Figure 27, A, p. 76) and for ECD (Figure 27, B, p. 76), respectively.

From overall 266 possible fragments, 143 fragments (27 b ions, 55 y ions, 61 internal ions) were observed in CID spectra and 168 fragments (96 c ions, 67 ions and 5 y ions) were observed in ECD spectra, respectively. Raw spectra were deconvoluted using SNAP 2.0 algorithm, further assigned by the Ms2Links software[177], and these masses were further used as a list of unique fragments for subsequent quantification of oxidation by the OxIntComp software.

A.
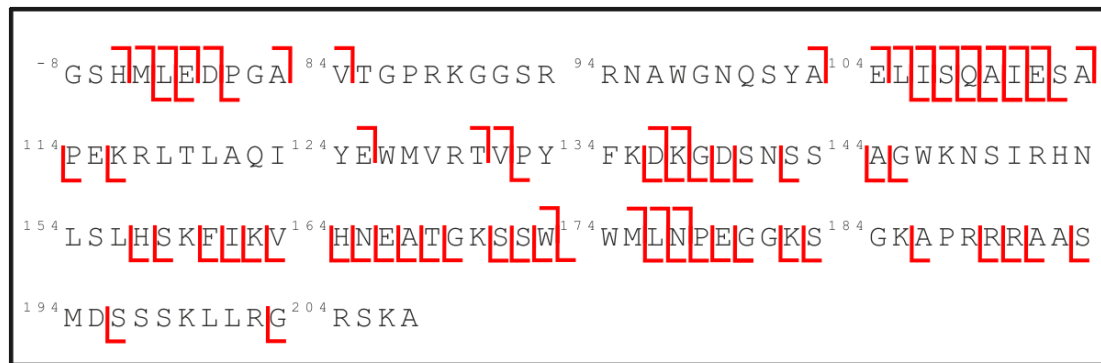
B.
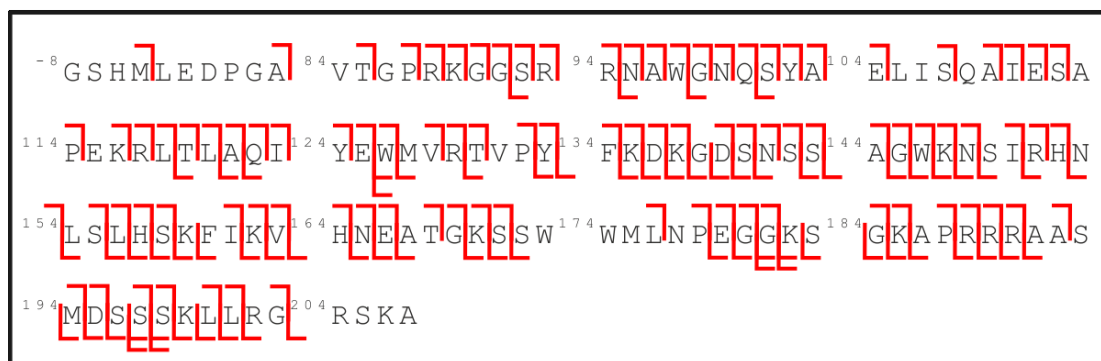
Figure 27 | Fragment coverage maps obtained by top-down fragmentations techniques of CID (**A**), and ECD (**B**), respectively. Double reverse fragments in ECD fragments spectra refers to the y ions, which can be also created during ECD.

5.3.2.2 Single-oxidized protein fragmentation

In order to localize the sites and extent of oxidation, CID has been performed. After collecting of broad *m/z* MS spectra for both, apo (Figure 28, A, p.77) and holo (Figure 28, B, p. 77) forms, singly oxidized 16+ charge state at *m/z* 928.90 Da mass selection has been isolated for both, apo (Figure 28, C, p. 77) and holo (Figure 28, D, p. 77) form, respectively. Next, CID has been provided. Fragmented raw spectra can be found in supplementary material as Figure S6 for both forms.

Figure 28 | Mass spectra of +16 charge state submitted to CID fragmentation. MS spectrum of +16 charge state apo form (**A**, purple) with following isolation of first oxidation (**C**, purple). Same, MS spectrum of +16 charge state of the holo form (**C**, orange) following stable isolation of first oxidation (**D**, orange). Solid line reflects the center of the isolation window with both sides following the dash lines, which reflect the width of an isolated window.

After a fragmentation, data were processed using an OxIntComp software. The list of assigned fragment and the list of all ions of oxidized protein served as an input files. Software searched for modified fragments in oxidized spectra using unmodified fragments from control. In the positive match, calculated ratio between oxidized and unoxidized fragments.

Overall, 50 oxidized ions (26 b ions, 15 y ions and 9 internal ions) were observed in CID spectra of singly oxidize ion, of which only some had a sufficient intensity to be further consider for quantification. These fragments (11 b ions, 9 y ions and 4 internal ions, Figure 29, A, p. 78 were quantified and plotted for both apo (in purple) and holo form (in orange) against the unmodified ions. Figure 29, B, p. 78 visualizes plotted fragments for both forms. The extents of modifications were compared for each obtained fragment between apo form, holo form and internal fragments (Figure 29, B, bottom).
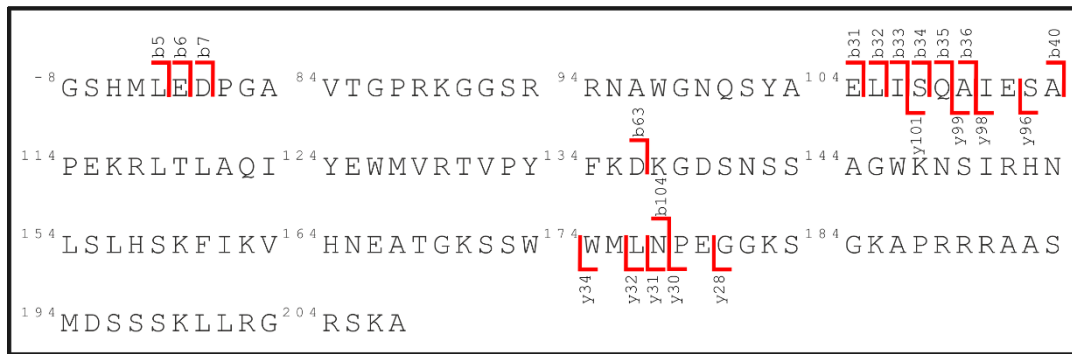
Figure 29 | **A.** Fragment coverage maps achieved by CID fragmentation of singly oxidized protein. Note, that first 8 residues are the rest of his-tag and numbering starts by -8 residue and continue with real numbering of FOXO4 with highlighted fragments. **B.** Histograms visualizing extents of oxidations obtained by CID fragmentation. Extents of modification of apo form for b, y and internal ions are highlighted in purple. Extents of modification of holo form for b, y and internal ions are highlighted in orange. The comparison of oxidized b, y fragments and internal ions are summarized at the bottom of the figure.

To achieve more efficient fragmentation, ECD was utilized as additional fragmentation technique. First, MS spectra were recorded for both, apo (Figure 30, A) and holo (Figure 30, B) form. Then, singly oxidized 16+ charge state was isolated for both, apo (Figure 30, C), and holo (Figure 30, D) form at *m/z* 928.90 Da. Both forms were subsequently fragmented as a technical triplicate. MS/MS spectra for both, apo and holo form can be find in supplementary material as a Figure S7.
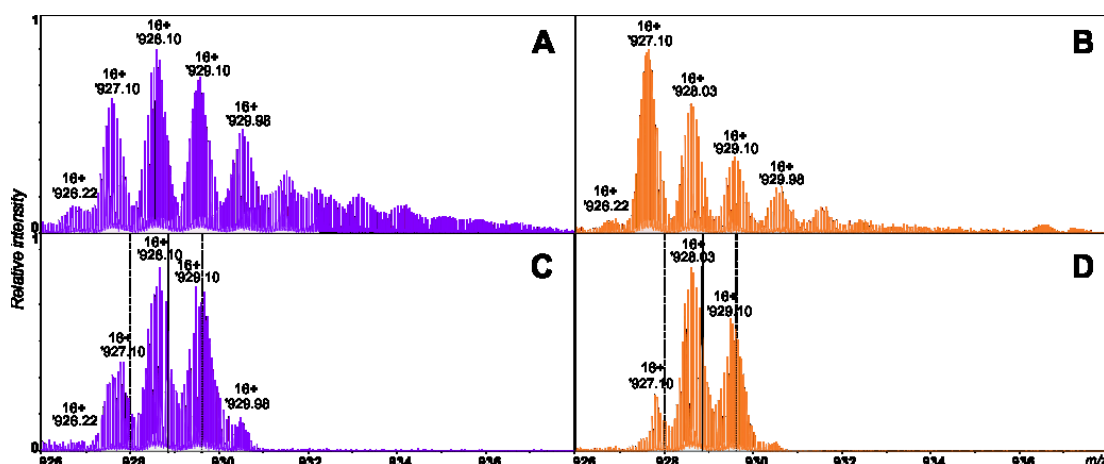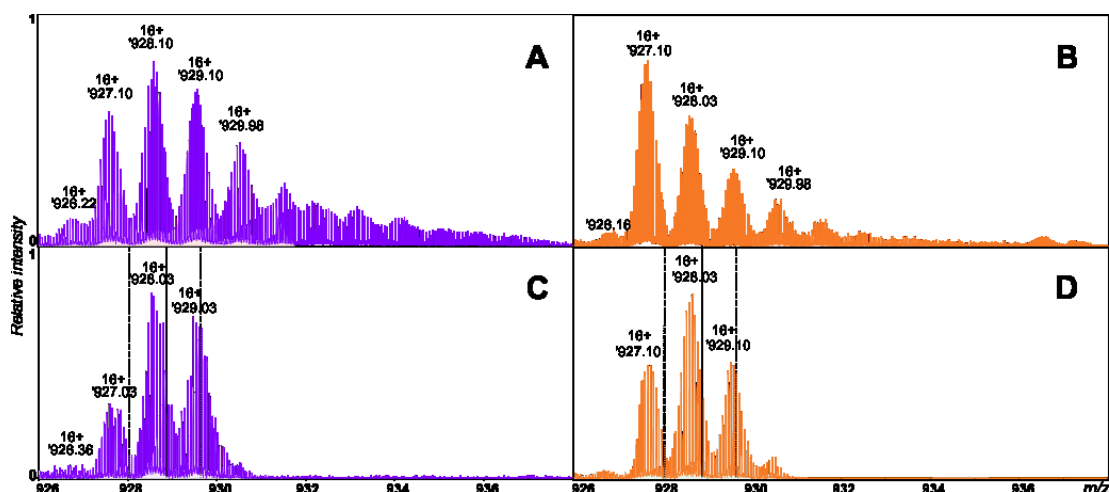

Figure 30 | Mass spectra of +16 charge state further submitted to ECD fragmentation. MS spectrum of +16 charge state apo form (**A**, purple) with following isolation of the first oxidation (**C**, purple). Same, MS spectrum of +16 charge state of the holo form (**B**, orange) following stable isolation of first oxidation (**D**, orange). Solid line reflects the center of the isolation window with both sides following the dash line, which reflects the width of an isolated window. Isolated modifications were further fragmented.

After a fragmentation, ECD data were processed using an OxIntComp software. The list of assigned fragment and the list of all ions of oxidized protein served as an input files. Software searched for modified fragments in oxidized spectra using unmodified fragments from control. In the positive match, calculated ratio between oxidized and unoxidized fragments.

In ECD spectra of singly oxidize ion, 112 ions (71 c ions, 41 z ions) were observed of which only some had a sufficient intensity to be further consider for quantification. These fragments (29 c ions, 13 z ions, Figure 31, A, p. 80) were quantified and plotted for both apo (in purple) and holo form (in orange) against

the unmodified ions. Figure 31, B visualized plotted fragments for both forms. The extents of modifications were compared for each obtained fragment between the apo and holo form, respectively (Figure 31, B, bottom).

A.
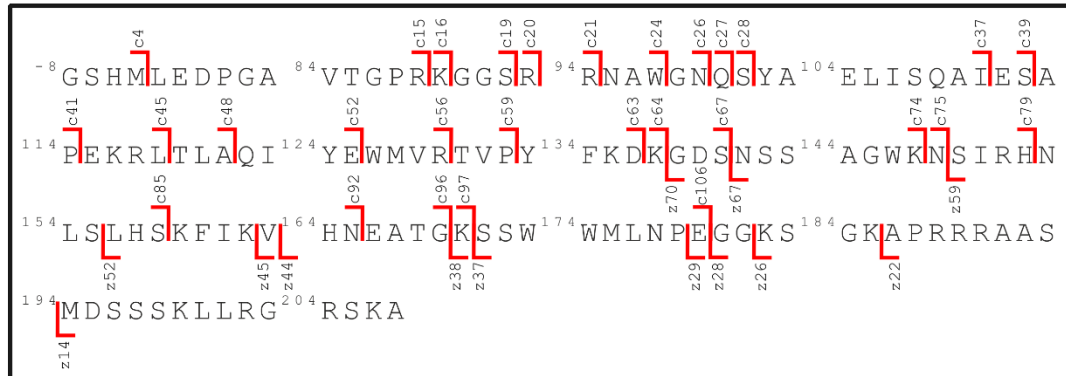
B.

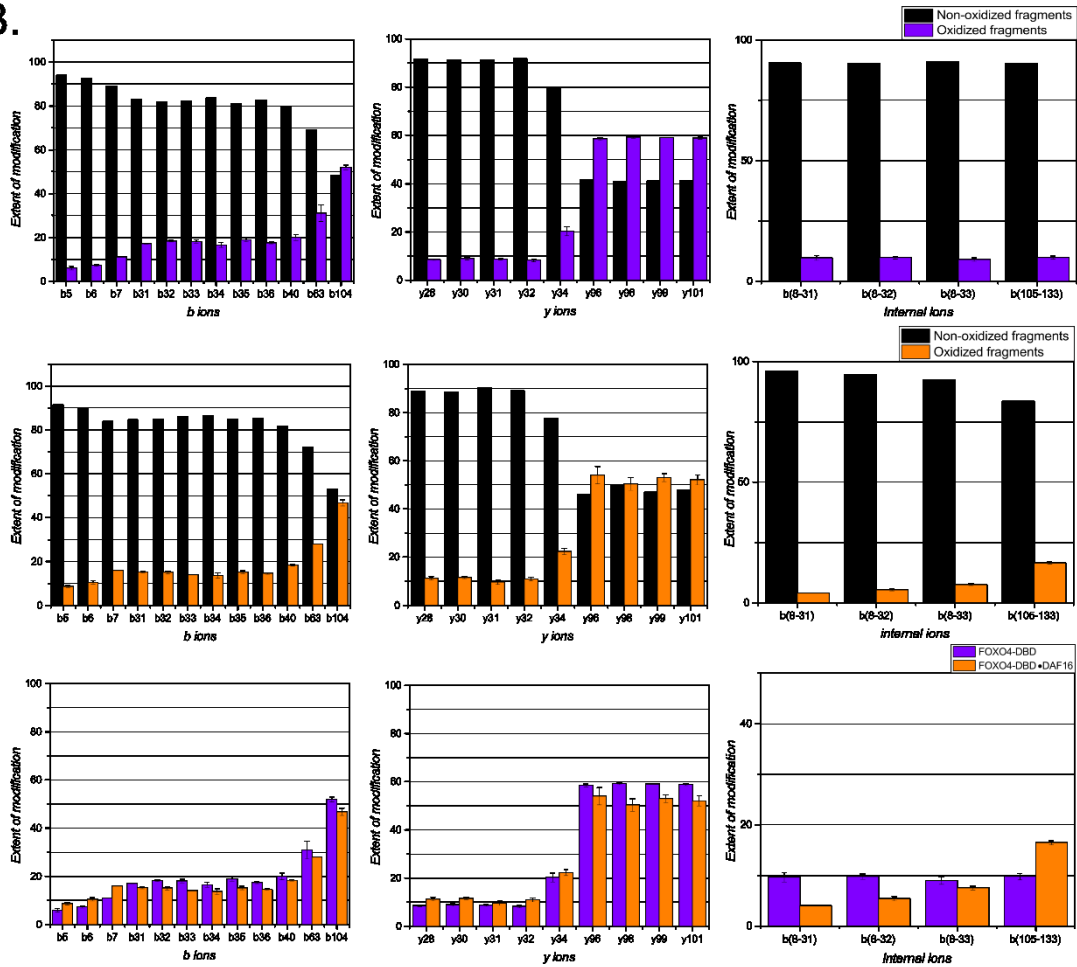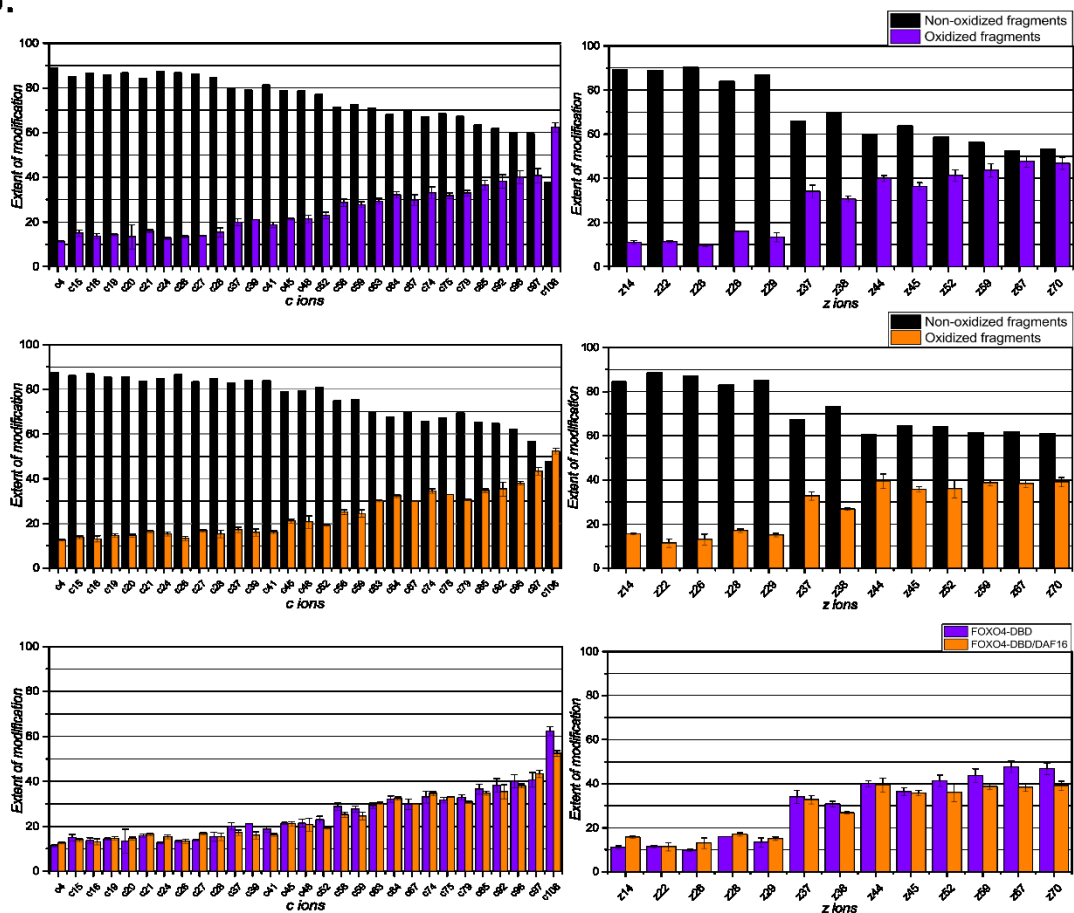Figure 31 | **A**. Fragment coverage maps achieved by ECD fragmentation of singly oxidized protein. Note, that first 8 residues are the rest of hi-tag and numbering starts by -8 residue and continues with real numbering of FOXO4 with highlighted fragments. **B**. Histograms visualizing extents of oxidations obtained by ECD fragmentation. Extents of oxidation of apo form for c and z ions are highlighted in purple. Extents of modification of holo form for c and z ions are highlighted

in orange. The comparison of oxidized c and z fragments of each form can be found and at the bottom of figure.

6. Discussion

In this thesis, DNA binding domain of transcription factor FOXO4 (FOXO4-DBD) has been chosen as a model protein for FPOP experiment. This protein has been chosen for a reason of well-known defined NMR solution structure of free protein[167], the crystal structure of its complex with DNA[176], and the multitude of publications[86,169,179] investigating interactions between FOXO4-DBD and its DNA binding element, DAF16.

Protein preparation and characterization. FOXO4-DBD at length 82-207 has been encoded and prepared using protein expression and purification protocol. Protein was carried by an N-terminal his-tag for further purification using affinity chromatography. SDS-PAGE of the affinity chromatography showed, that not only FOXO4-DBD protein was eluted (Figure 19 A, p. 67). This is caused by the presence of the proteins with naturally occurred histidines in spatial proximity, which bind to affinity resin non-specifically. Nevertheless, the gel filtration successfully purified FOXO4-DBD from other protein contaminants (Figure 21, p. 68) and mass spectrometry of intact protein fully confirmed the presence of FOXO4-DBD (Figure 22, p. 69). Notably, SDS-PAGE of fractions #11-#13 (Figure 21, p. 68) also revealed smaller, minor bands under the major, FOXO4-DBD band. These bands correspond to a shorter versions of FOXO4-DBD, created by spontaneous cleaving of amino acids from the C-terminus. This part is not directly interacting with DNA and as our data shows, the truncated products do not interfere with the full-length DBD hence the acquired data can be fully interpreted.

The desired FOXO4-DBD•DAF16 complex was prepared by annealing protein and duplex DNA in equimolar 1:1 ratio. Native forms of both free and bound forms were further examined by native nESI (Figure 23, B, p. 70) and reported mass shift to higher *m/z*; and in the case of free form lower charge compared to denatured one. Native gel electrophoresis (Figure 23, A, p. 70) revealed significant mobility shift of protein·DNA complex compared to DNA duplex (dsDNA). However, the lane containing the desired complex contains a small amount of free dsDNA. This

may happen due to spontaneous dissociation of dsDNA from complex due to a high voltage during the electrophoresis, which subsequently heats the gel during the separation process.

Laser condition optimization. The prepared FOXO4-DBD and FOXO4-DBD•DAF16 complex have been further utilized for the following FPOP experiment. We initially tested the laser conditions at a different frequency and energy. Within the constant 15 Hz frequency, no impact of oxidative modifications was observed even with increasing energy in range of 15-40 mJ (Figure S1). For this reason, we did not perform further experiments at the defined 15 Hz frequency.

However, at a constant frequency of 20 Hz (Figure 24, p. 72), higher impact on distribution of protein oxidative modifications was observed. The energy of 15 mJ caused desired extent of oxidation for apo form, but not for holo form. The energy of 20 mJ also caused sufficient extent of oxidation for apo form, where the protein was present predominantly in a singly and doubly oxidized form. Nevertheless, DNA protection provided lower yield of oxidation in holo form. This could be a problem within the isolation of singly oxidized modification during top-down analysis. The desired oxidative impact for both forms was observed at a frequency of 20 Hz and energy of 30 mJ. These conditions were further applied to oxidize protein and protein·DNA complex. Figure 24, p. 72 shows that apo (in purple) and holo form (in orange) differ individually in the extent of oxidation. While the apo form is predominantly oxidized, DNA physically protects protein in the holo form, hence the protein is less oxidized. Energy of 40 mJ revealed higher oxidative damage accompanied by mass losses.

Experimental strategies. Apo and holo forms were further analyzed by both bottom-up and top-down approaches. The advantage of bottom-up approach lies in the spatial resolution of individual peptides. By digestion of protein into smaller peptides, we can thoroughly focus on each peptide. By performing an MS/MS analysis we can obtain specific information about each modification located on individual peptides. The disadvantage of bottom-up approach is that any modifications caused by multiple oxidations are analyzed as a mixture, and thus their origin cannot be tracked. Thus, multiple modifications can be burdened by changes in protein states.

If so, our assigned modification could be created by an influence of structural changes on protein surface, not by actual presence of that residue in the protein surface.

In pursuit to prevent that, top-down approach has been chosen as a novel unique approach to localize the sites of oxidation. The advantage of this approach is that we can theoretically obtain fragments through the entire molecule, even where information from bottom-up approach can be missing due to missing peptides. Multiple oxidation events can alter protein conformation; hydroxyl radicals can oxidize amino acid residues, which were not originally solvent accessible[102]. The undesirable artifacts of further oxidative processes can be minimized using the top-down technology where only singly oxidized ion is isolated.

In order to investigate the effect of DNA bound to protein, the results of bound form (holo-form) were compared to unbound form (apo-form). Bottom-up results were compared to those obtained by top-down approach. Finally, combined results obtained by FPOP were compared to a previous HDX study (Figure 32)[86].
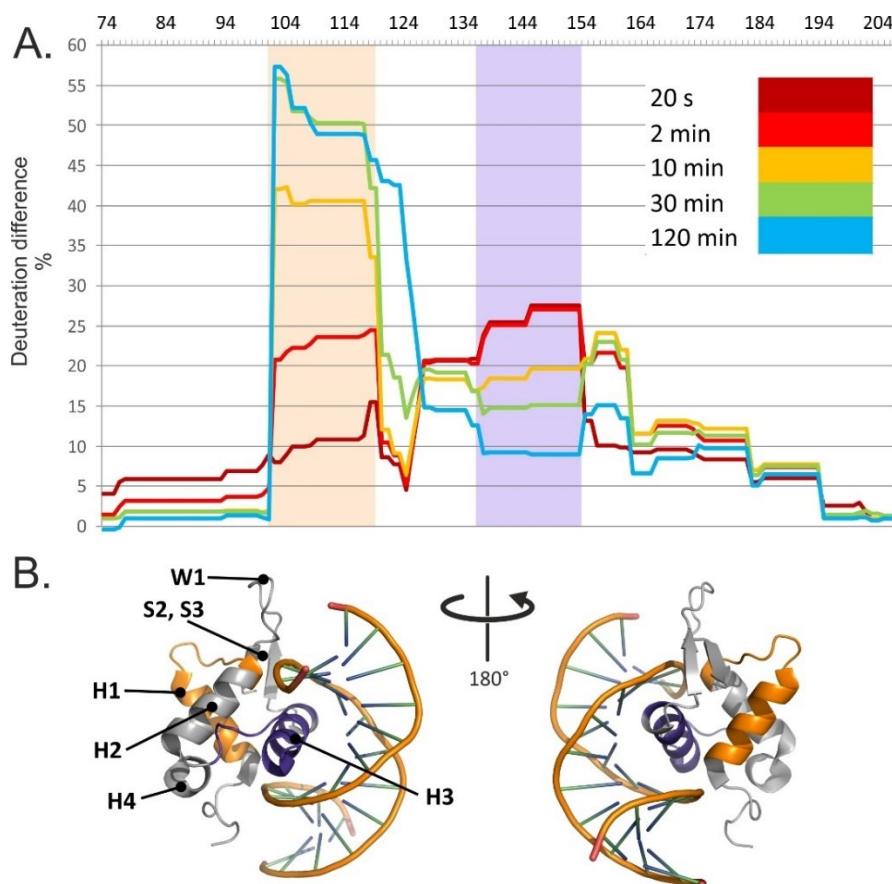
Figure 32 | **A.** Deuteration difference (DR) between FOXO4-DBD and FOXO4-DBD•DAF16 complex. DR is visualized along protein sequence

and their evolution in time. Both orange and purple regions indicate two regions with the most significant deuterium difference. **B.** Crystal structure of FOXO4-DBD•DAF16 complex with highlighted topology and significant deuterium differences. Adapted with permission from Slavata et al.[86].

Bottom-up approach. Classical bottom-up approach has been utilized as a current gold standard for analyzing FPOP samples. Irradiated samples underwent proteolytic digestion by LysC, or LysC/Trypsin mixture reaching 82% sequence coverage (Figure 25, p. 73). C-terminal peptides were missing in the peptide map due to an abundance of basic residues (R, K) in this part of the sequence, and obviously such small peptides were not retained in the trap column. In this case, even LysC did not increase spatial resolution. Interestingly, peptide 160-170 was missing as well, possibly due to high polarity of this peptide. To increase spatial resolution, another specific protease may be tested.

LC-MS/MS analysis precisely located sites of oxidation on certain residues (Figure 26, p. 74), which were subsequently highlighted in the crystal structure (Figure 33, p. 85 and Figure 34, p. 86, respectively). Beside the +16 Da oxidation, other modifications were also detected: keto-oxidation of (L-4) residue (+14 Da), deguanidylation of R94 residue (-43 Da) and conversion of H152 and H157 to aspartate distinguished with 22 Da mass lost.

Certain residues retained their solvent accessibility, and they did not report rapid changes either in apo or holo form ((H-6) oxidation, (H-6) conversion to aspartate, (L-4) oxidation to carbonyl, W97/W97[#] oxidation, (M-5)+2O and also double oxidation of (H-6) and (M-5) residues). Upon DNA binding to protein, the effect of DNA caused that some residues showed changes in their solvent accessibility and they are individually discussed below.

The crystal structure[176] indicates, that residue R94 is directly involved in DNA binding via its guanidyl group and it was found to be more oxidized in the presence of DNA. While guanidyl group interacts with DNA, δ-carbon of side chain (which binds guanidyl group) is reactive and it seems to be more rotated into the solvent. Therefore, δ-carbon is the first target for radical attack[123], while guanidyl group still participates in DNA binding (Figure 37, B, p. 91)[176].
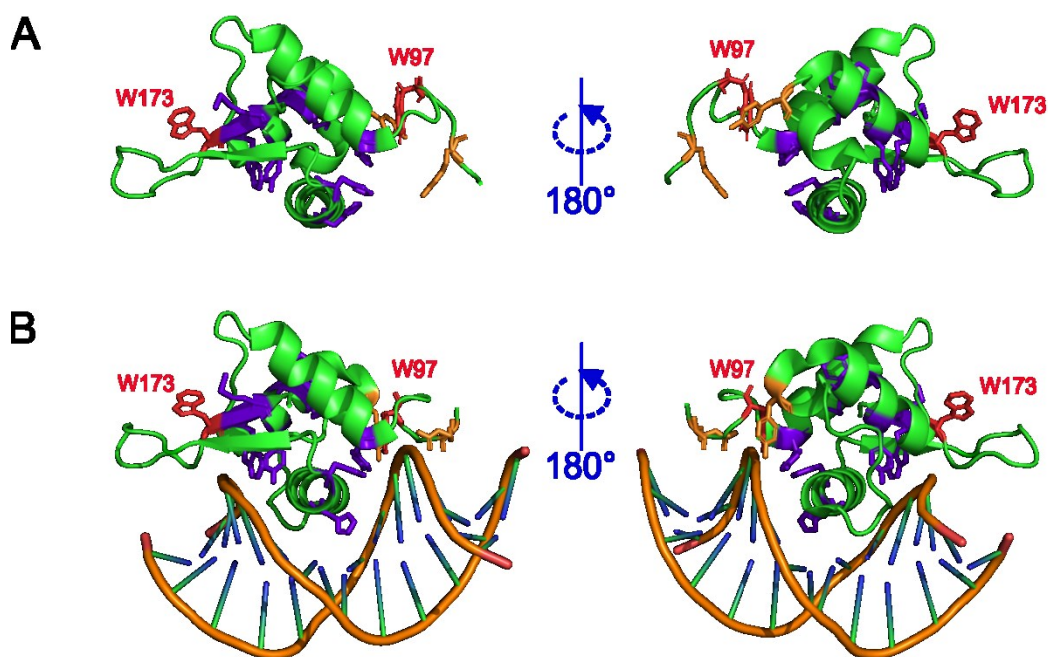
Figure 33 | LC-MS/MS analysis revealed sites of oxidation. **A** NMR structure of FOXO4-DBD[167] and **B** crystal structure of FOXO4-DBD•DAF16 complex[176], respectively. Residues, which were found to be more oxidized in apo form are highlighted in purple, or more oxidized in holo form are highlighted in orange, respectively. No significant difference in both forms was observed for residue W97 (highlighted in red). On the contrary, W173 residue was found to be oxidized in two isoforms – one more and one less oxidized.

LC-MS analysis could effectively separate and detect peptide modified on Y102 residue at two isoforms, which were both found in singly (Figure 26, p. 74 displayed as Y102 and Y102[#]) and doubly oxidized states. In the presence of DNA, Y102 residue was observed to be resistant to oxidation and hence was not exposed to solvent (see Figure 37, B, p. 91). This is in agreement with the crystal structure, where Y102 is crucially involved in DNA binding. HDX data (see also Figure 32, p. 83 and supplementary material in ref.[86]) revealed lower deuterium exchange at short reaction time in region 94-102, but provided no further specific information. Nevertheless, FPOP offers a closer look on individual residues, and shows that Y102

residue is in the presence of DNA responsible for lower solvent accessibility of this region.
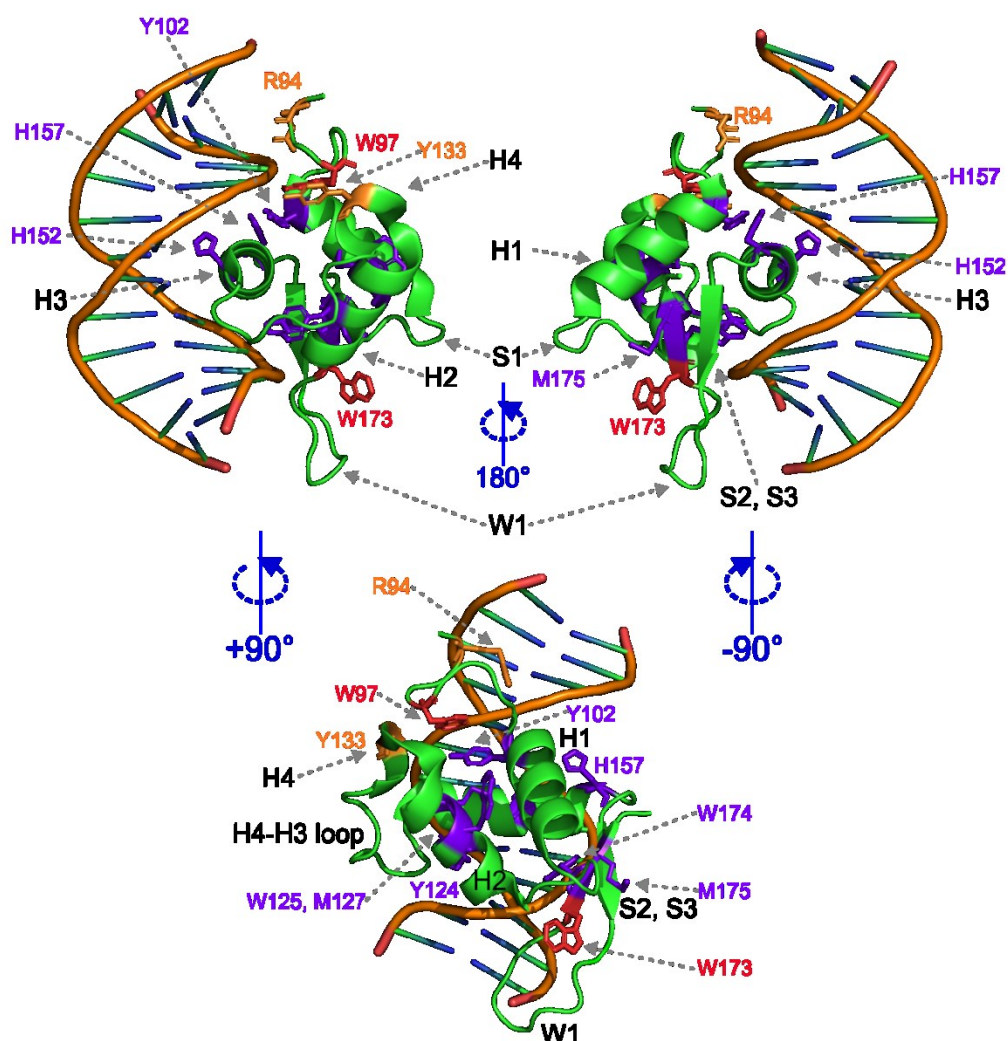
Figure 34 | The crystal structure of FOXO4-DBD•DAF16 complex[176] with highlighted topology (in black) and residues, which undergo FPOP oxidation. Residues more oxidized in apo form are highlighted in purple, residues more oxidized in holo form are highlighted in orange. Red labelled residue W97 showed no significant difference in oxidation either in apo or holo form. Residue W173 was observed in two isoforms, of which one was more and one less oxidized in the presence of DNA.

The next described section spanning region of A103-T130 follows the topology of helix H1, intervening loop, strand S1 and helix H2. Residues Y124, W126 and M127 are located in helix H2 (Figure 34). All three residues are very reactive[120] and have been found to be oxidized (W126 predominantly doubly-oxidized). Further

information on H1 helix, strand S1 or intervening loop is missing in bottom-up approach. As found in the HDX study (Figure 32, p. 83), helix H1, intervening loop, strand S1 and helix H2 spanning region A103-T130 showed the highest differences in deuteration during the HDX experimental conditions[86]. HDX data points out this is an indirect conformational effect of DNA binding throughout the allosteric effect. It seems that upon DNA binding, these regions loses their flexibility and become less solvent accessible. FPOP data clearly suggests the same; neither Y124, W126 or M127 is participating in DNA binding[176], they lose their solvent accessibility upon DNA binding, and thus they are less oxidized (Figure 26, p. 74, Figure 34, p. 86).

Residues 131-135 form an additional 3_{10}-type short helix, commonly called helix H4, which places this region between helices H2 and H3 and can be found in FOXO4 and in some other FOX proteins. By placing the helix H3 in the duplex DNA, structural change occurs in helix H4, and this region is reported to be rather unstructured in the holo form, while in the apo form is reported to be rather helical[179]. Consequently, Y133 residue was found to be more oxidized in the holo form. HDX data indicates this region with lower deuteration in shorter times (20 s – 30 min), which consequently shows this regions to be rather more exposed to solvent in the presence of DNA.

In addition, FOXO4-DBD forms a cluster of hydrophobic residues containing W97, Y102, Y133, F134 and W146. The cluster partly plays a role in protein-DNA interaction and these individual residues interact with the DNA backbone. Based on our results, no differences have been observed in the case of W97 residue (N-terminus) in the apo or in the holo form. The solvent accessibility decreased for residue Y102 (helix H1), on the other hand Y133 (helix H4) residue increased its solvent accessibility. Both F134 residue located at H4 helix and W146 residue (helix H3) showed to be completely inert to oxidation, as they are not exposed to solvent either in apo or holo forms.

FOXO4-DBD places its helix H3 into the major groove of duplex DNA, which creates most of the protein-DNA interface. Residues H152 and H157 located on H3 helix can be easily oxidized in FPOP and were found to be mass shifted by +16 Da (addition of oxygen) and -22 Da (conversion to aspartate). This indicates full solvent accessibility of helix H3 in apo form. In holo form, both H152 and H157 decrease their solvent accessibility and they are less oxidized, due to direct protection by DNA. Moreover, no conversion to aspartate was observed in bound form

(Figure 35, p. 88). HDX data support our observations, whereas this region showed lower deuterium exchange, which indicates that helix H3 is directly protected by DNA (Figure 32, p. 83).
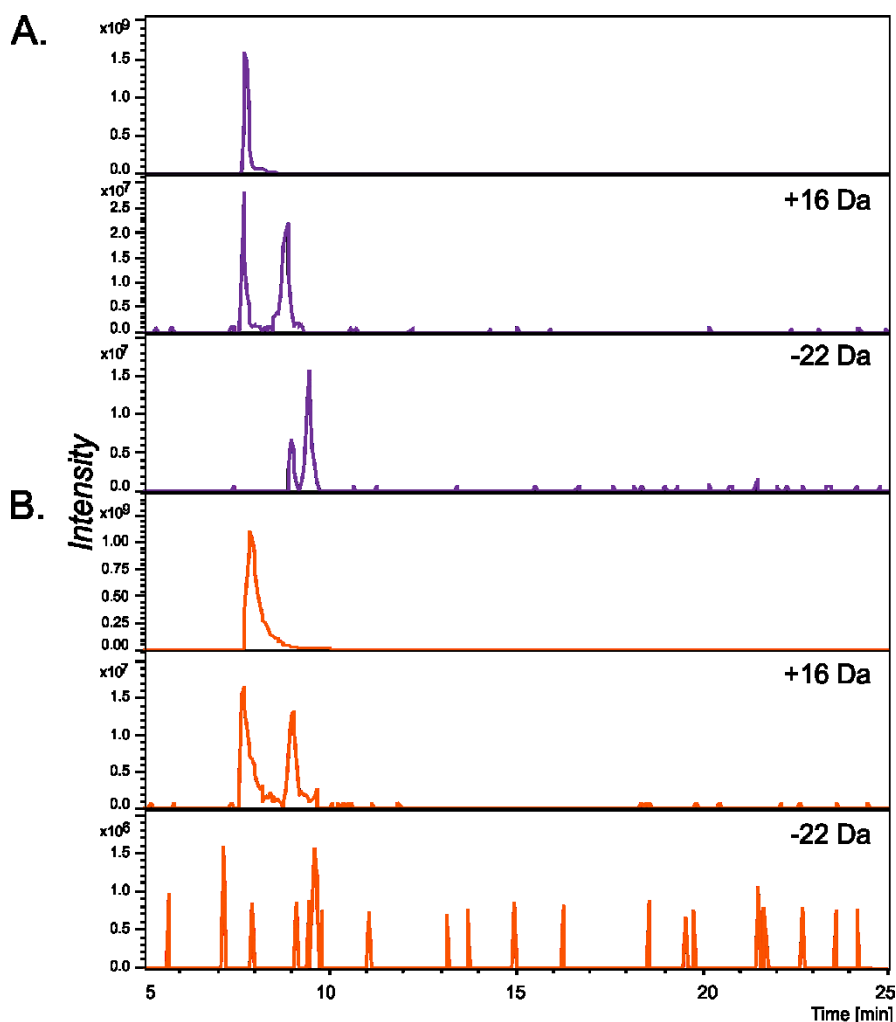
Figure 35 | Extracted ion chromatograms of apo (A., in purple) and holo (B, in orange.) forms ranging retention time of 5-25 min of peptide spanning region 152-159. Using extracting ion chromatograms, +16 Da (addition of oxygen) and -22 Da (conversion to aspartate) modifications were found. In the presence of DNA, decreased extents of +16 Da modification was found, accompanied by decreasing overall peptide intensity. Nevertheless, no conversion of H152/157 to aspartate was observed due to DNA protection.

In strand S3 and part of wing W2 covered by peptide 171-182, three residues were found to be oxidized. Modification +16 Da and various oxidative combinations were detected on residues W173, W174 and M175 (Figure 26, p. 74). Residue W173

appeared to be oxidized in two isoforms; one isoform (W173) was more oxidized in the holo form, while the second one was less oxidized (W173[#]) as well as doubly oxidized residue. In pursuit of clarifying this, one possible explanation establishes the direct DNA stabilization of this region. Where W173 residue is turned more into the solvent and therefore gets more oxidized, while the second side is less exposed to solvent and therefore gets less oxidized. Residue W174 is completely covered by DNA and it was found no to be oxidized. Compared to the crystal structure, the residue W174 is oriented into protein-DNA interface and directly interacts with the phosphate group of DNA[176]. Similarly as W173 residue, M175 residue was found to be less oxidized in the holo form, yet oxidized enough to be solvent accessible (Figure 37, A, p. 91). Our findings are confirmed by detecting oxidized variations of these three residues in extracted ion chromatograms (W173/M175, W174/M175, W173+2O/M175), where their exposure to solvent is reduced in the presence of DNA. Such differences of extents for W173, W174 and M175 residues are visible in Figure 36 presenting extracted ion chromatograms, of which time ranging of peptide elution is visualized.
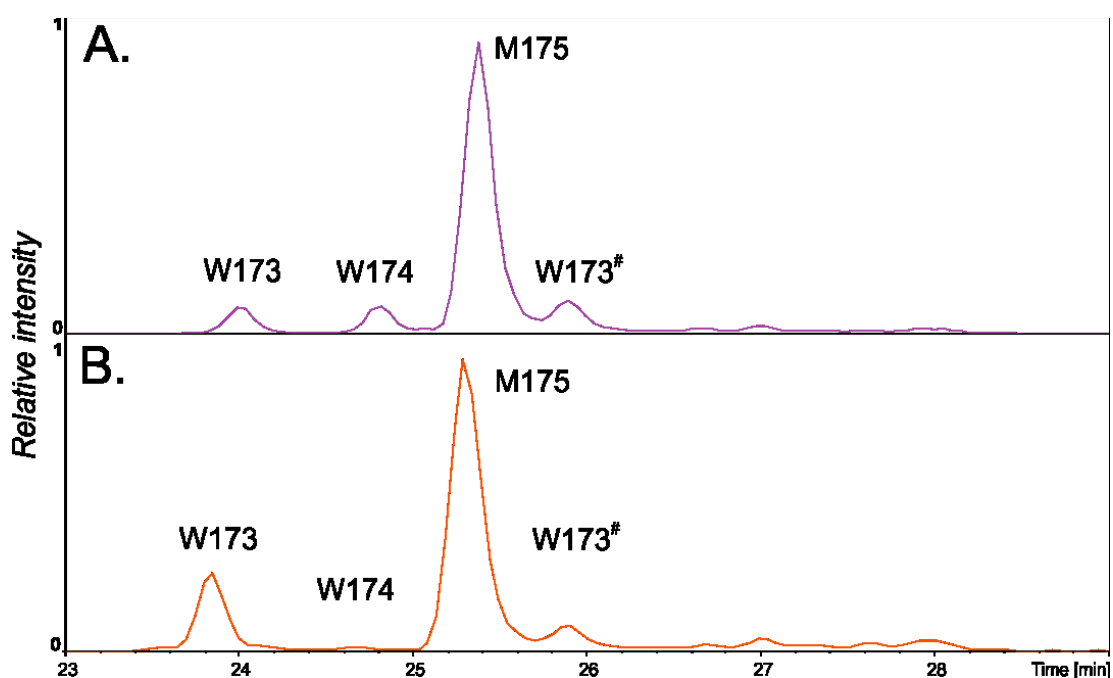
Figure 36 | Extracted ion chromatograms of the apo (A., in purple) and the holo (B., in orange.) forms in retention time ranging 23-29 min, where the oxidized peptide 173-182 had been eluted. It is visible even from the chromatogram, that residue W173 is more oxidized in the holo form and its isoform W173[#] is slightly less oxidized

in the holo form. The exposure to solvent of residue W174 in the holo form is completely disabled and the residue is almost unoxidized. For M175 residue no change is visible, hence the extent was calculated based on the peptide intensity.

Last oxidized residue revealed by bottom-up approach is M194 (Figure 26, p. 74). M194 is located near the region of basic residues R188-R190, which create ionic unspecific contacts with DNA phosphate groups. Notably, basic residues and phosphorylated S193 are motifs for kinase and subsequent the 14-3-3 protein recognition[169,170,180]. This basic region helps to stabilize protein-DNA interaction, and M194 seems to be rather solvent accessible in the presence of DNA. The crystal structure lacks the region 178-207 due to high flexibility[176]. Nevertheless, we cannot further explain the increased solvent accessibility of this region. The 14-3-3 protein binding motif may be responsible for this behavior, however with no further evidence. This could be a motivating goal for further investigation of not only the FOXO4-DBD structure, but also other transcription factors binding to its cognate DNA binding partner. Notably, according to protein-DNA cross-linking, quantitative cross-linking and HDX data[86], a homology model containing C-terminal region was built, where this region shows to be still fully exposed to solvent.

Overall, the C-terminal region containing strands S2 and S3 forming antiparallel β-sheet, and wings W1 and W2 seem to be still exposed to solvent. Higher deuterium rates pointed out the stabilized region of wing W1 and strand S2, which correspond to decreased flexibility of this region when placing a helix H3 into the major groove of DNA. In the region downstream from the wing W2, no further differences in deuterium exchange were revealed, according to HDX data[86].
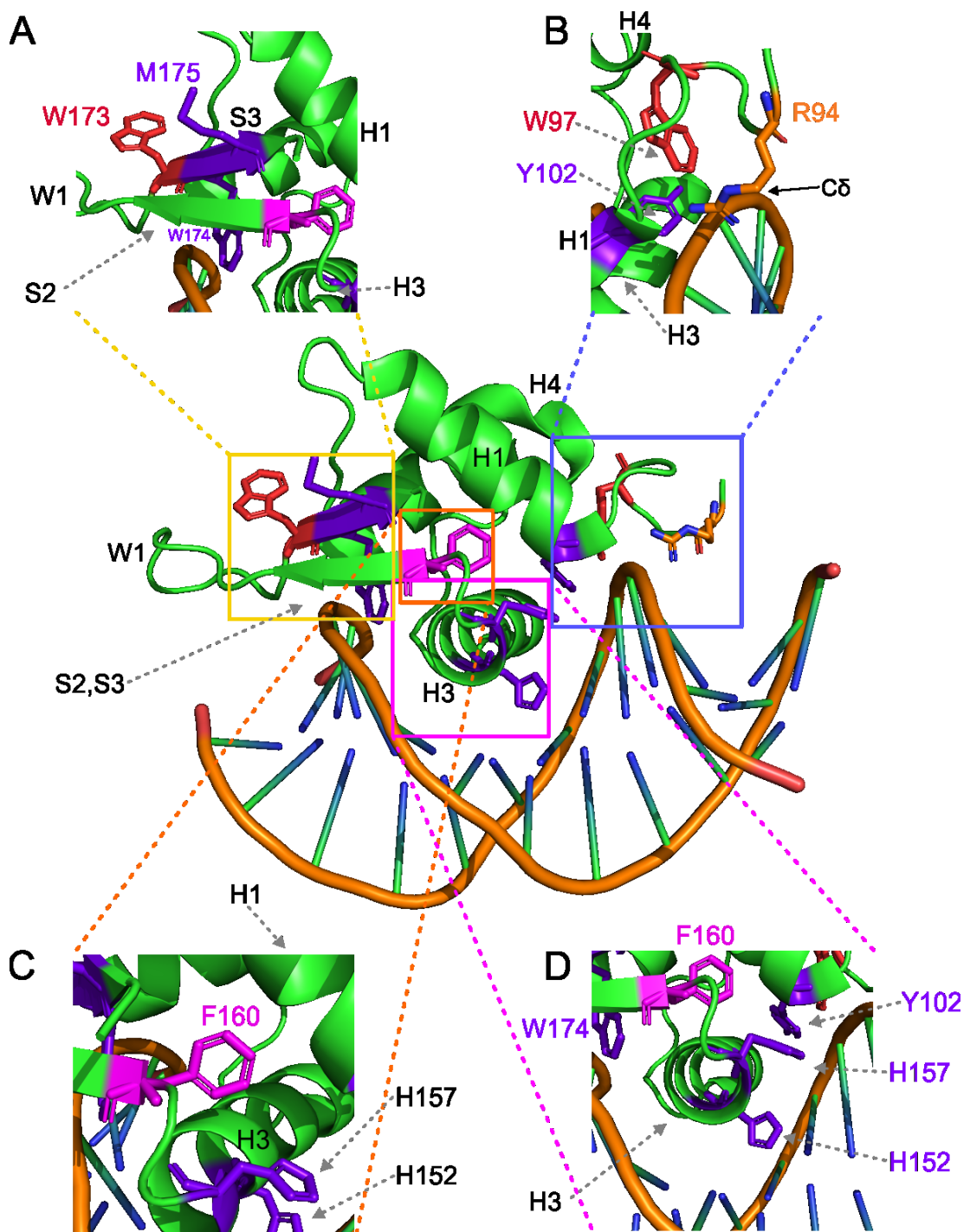
Figure 37 | Closer view on the crystal structure of FOXO4-DBD•DAF16 complex[176].
A Closer view on S2, W1 and S3 region forming antiparallel β-sheet. Residues W173, W174 and M175 are highlighted. W173 and M175 are exposed to solvent, while W174 is directly involved in DNA binding. **B** Residue R94 directly interacts with DNA. In the holo form, Cδ carbon is rotated more into the solvent, and in the holo form, R94 residue was deguanidylated more. Residue W97 (in red) shows the same exposure to solvent both in the apo or the holo form. Residue Y102 (in purple) is crucially involved in protein-DNA interaction. **C** Top-down approach revealed residue F160

(in magenta) to be solvent accessible and further oxidized. In bottom-up approach, information is missing due to missing peptide. **D** Zoom to helix H3, which is placed into the duplex DNA. Residues H152 and H157 (in purple) are protected by DNA, therefore they are less oxidized.

Top-down approach. Top-down approach offers the benefit of isolating the first oxidation event which is not affected by possible protein structure perturbation potentially caused by previous and/or multiple modifications[102]. After a stable isolation, free and bound forms were fragmented by CID and ECD, respectively. Charge state +16 was isolated due to dominantly uniform distribution of fragments obtained by ECD along the entire protein molecule. Other charge states (+17, +18, +19) did not provide such rich fragment distribution along the entire protein molecule, especially in the region of helix H3 which creates the main protein-DNA interface. CID did not provide such fragments along the whole molecule, but on a both protein termini. To increase spatial fragment resolution over the whole protein, different fragmentation techniques may be used, such as UVPD[181,182].

Isolation of singly oxidized +16 charge state ion at *m/z* 928.90 Da and isolation window of ±0.9 Da was performed. However, this modification was not sufficiently isolated. To achieve precise and efficient isolation of such a small modification, decreasing the width of isolation window to ±0.5 Da is necessary. Unfortunately, in the case of FOXO4-DBD with molecular weight of ~15 kDa combined with naturally abundant isotopic distribution, it is not technologically possible to isolate such a charge state at certain *m/z* and such small isolation window in a quadrupole analyzer. Based on our observation during ECD fragmentation, the isolation window was adjusted to strive both: (i) isolation of singly modified event, and partially the second event due to no-information yield of unmodified ion; (ii) preserving the S/N ratio. To perform even more precise isolation in quadrupole analyzers, it is possible to decrease the overall number of isotopes by using protein isotope depletion of ^{13}C and^{15}N[183,184].

A high number of fragments were observed in raw spectra, but only some of the fragments had sufficient intensity to be further analyzed, quantified and plotted. Due to the high complexity of ECD fragment spectra and low fragmentation yield,

several fragments did not have an ideal isotopic distribution and intensity to be further used for quantification. These fragments should be eliminated, because they do not follow the overall trend of increasing oxidation, which are for certain fragments discussed further and critically evaluated. Some individual fragments of the apo and holo forms revealed differences in their extents compared to each other. These fragments are plotted and visualized in Figure 29, B, p. 78 for CID and Figure 31, B. p. 80 for ECD, respectively. Both CID and ECD fragments were compared and further analyzed, and regions with significant oxidative differences are indicated in Figure 38. The most suitable candidate for oxidation is chosen and discussed for each region, according to relative reactivity of amino acid residues: Cys~Met~Trp>Tyr>Phe>His>Leu~Ile>Arg~Lys~Val>Ser~Thr~Pro>Gln~Glu>Asp~Asn>Ala~Gly[120]. Resulting data were subsequently compared to data from bottom-up approach and HDX experiments[86], as discussed above.
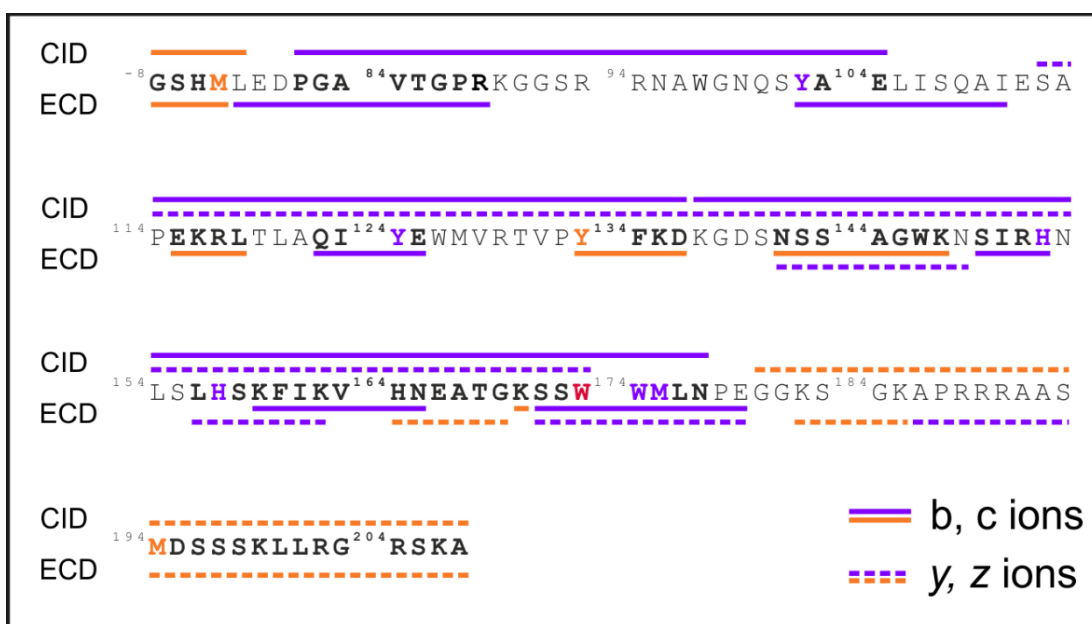
Figure 38 | Sequence map of FOXO4-DBD, with highlighted regions. Individual fragments from both ECD and CID visualize overall regions, which are differently oxidized in the apo or holo form. Sequence, where specific change has been detected for apo or holo form is bolded. Solid lines indicate b/c ions, dash lines indicate y/z ions. Orange/purple residues represent detected residues found in bottom-up approach in these regions. Residue W173 (in red) represents change detected in bottom-up approach in two different oxidized isoforms (W173 and W173[#]). One of them was found to be more oxidized and one less oxidized in apo form.

First region with a significant difference between the apo and holo forms detected in top-down data is covered by fragments c4/b5 (Figure 38, p. 91). Both CID and ECD fragments report to be more oxidized in holo form, therefore they are highlighted in orange. If taking varying amino acid reactivity into account, methionine (M-5) residue would be oxidized. Bottom-up data also decipher (H-6) and (L-4) residues to be oxidized (Figure 26, p. 74), but neither of them showed relevant changes between both forms. To localize any other oxidation located in this region, additional top-down data are essential. Next fragment c15 obtained by ECD spectra spans c5-c15 region, where the apo form is more oxidized. In CID, fragment b31 was detected and covers b8-b31 region, where the apo form is more oxidized. Intersection of fragments c15 and b31 are residues (P-2) - R88, where no modified residue was detected in bottom-up approach. When considering fragment b31, a jump for the apo form from b7 is significantly visible and this extent keeps until b40 fragment, where no significant change is visible in both forms. In comparison to previous fragment c28, fragment c37 shows to be more oxidized in benefit to the apo form. Combining this change with the change between b7-b31 fragments, whose extent also prevails in the apo form, a change occurred in the intersecting c29-c31 region. Region c29-c31 spans residues Y102, A103 and E104 (highlighted in bold in Figure 38, p. 93 and labeled in purple in Figure 39, p. 95). Amino acid reactivity[120] indicates Y102 residue would be oxidized. Bottom-up data also clearly shows that the Y102 residue is oxidized more in the apo form, while in the holo form it is almost not oxidized at all. Both bottom-up and top-down data corresponds with the crystal structure[176] and HDX data[86], where this residue is directly involved in DNA binding.

CID fragment b63, which spans region of b41-b63; b104 spanning region of b64-b104 as well as y96 spanning region of y35-y96 were observed in spectra. These fragments cover long distance regions and do not provide any further information despite the fact, that this large regions undergo oxidation, which is visibly more prevalent in the apo form.

Next region, where diverse significant differences were observed, spans a section of helix H1, intervening loop, strand S1 and helix H2. Region covered by fragment b63 (b41-b63) provides just overall information about increased solvent accessibility in this region. Region of intervening loop and strand S1 located between helices H1 and H2 spans E115, K116, R117 and L118 residues, and undergoes more significant

oxidation in the presence of DNA. This was observed in ECD plots, where the extent increased between c41-c45 fragments in the benefit of holo form. Nevertheless, further information regarding biological role of this region, or further location of oxidation is missing. Next region which undergoes significant difference is located between c48-c52 fragments. This region covers first half of H2 helix and includes residues Q122, I123, Y124 and E125 (Figure 39). According to amino acid reactivity[120] of residues located in this region, Y124 residues can be easily oxidized. In this case, bottom-up data are consistent with our findings from ECD measurement, because Y124 residue was found to be less oxidized in the holo form. Bottom-up data also revealed that residues W126 and M127 are oxidized. In top-down data, region between fragments c52-c56, where higher extent of oxidation has occurred, but with no significant difference between the apo and holo forms.
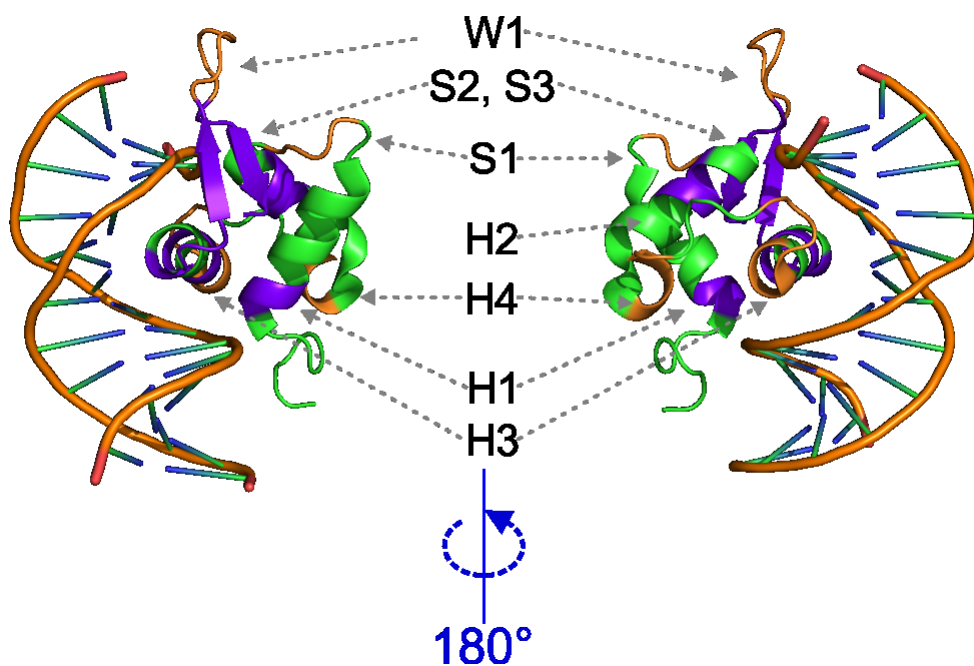
Figure 39 | The crystal structure of FOXO4-DBD•DAF16[86] complex with illustrated topology of individual regions. Highlighted regions represent regions, which undergo FPOP oxidation and significantly differ in the apo or holo form, respectively. Regions, which undergo higher oxidation in the apo form are highlighted in purple, while regions, which were more oxidized in the holo form are highlighted in orange.

To summarize the region spanning A103-T130 residues of helix H1, intervening loop, strand S2 and helix H2, it undergoes the biggest differences in deuteration during the HDX (Figure 32, p. 83). In this region, HDX focused on dynamics and detected loss of solvent accessibility in bound form due to conformational change. Nevertheless, HDX did not provide any closer insight on individual residues. Bottom-up data revealed certain residues to be less oxidized, which is related to loss of solvent accessibility in this region. Additionally, top-down data provided even closer view on this region. Fragments c37, c39 and c41 undergo no significant change between both forms in helix H1. Intervening loop and strand S2 (c41-c45 fragments) increased its solvent accessibility in the benefit of the holo form, and finally helix H2 decrease the extents and determining that Y124 residue shows to be less exposed to solvent, and consequently less oxidized.

Fragment c63, compared to previous fragment c59 shows a significant difference between both forms. The difference points out that the region between fragments c59-c63 is rather exposed to solvent in the presence of DNA. Bottom-up data also revealed this region to be more solvent accessible and consequently Y133 was found to be more oxidized in the presence of DNA (Figure 39, p. 95). This region (c60-c63) forms a well-known H4 helix, which is rather unstructured upon DNA binding, as is determined by kinetic studies[179]. HDX data (Figure 32, p. 83)[86] indicate the region forming helix H4 to increase its solvent accessibility following increased deuteration in the holo form.

Increase of extent was observed in region of fragments c67-c74 (Figure 38, p. 93). This change indicates the holo form to be more oxidized. Bottom-up data did not reveal this region to be oxidized at all. One candidate for oxidation in this region could be residue W146 located at the beginning of H3 helix, but according to the crystal structure[176], W146 is part of a cluster of hydrophobic residues (contains also W97, Y102, Y133, F134), which partly interacts with DNA. This extent change was observed between fragments c67-c74 due to decreased extent of fragment c67. This fragment also decreased its extent compared to previous fragment c64, which should not have happened at all. Each fragment should inherit information about its extent from the previous fragment, followed by increase of extent if oxidized at a certain residue or remains in the same value, if no change was occurred. If fragment c67 was eliminated from the ion list, due to not-following the trend

as mentioned above, no change would then be observed between fragments c64 and c74. This region is also covered CID fragment b64-b104, where visible higher oxidation in the apo form was observed. Moreover, this region is covered by fragment z67 (z59-z67 region), where the apo and holo form significantly differ in favor of the apo form. Notably, fragment z70 did not reveal any increase in the apo form as opposed to the previous fragment z67. That must explain that no significant change happened in the region N141-K147. Compared to c67 fragment decreased intensity, z ions are more credible in this statements, as they follow the trend of increasing/retaining extent.

Helix H3, which places into the major groove of duplex DNA, is sequenced by fragments c79 and c85, where significant difference between c75-c79 fragments is observed in benefit of the holo form (Figure 38, p. 93). Next fragment c85 continues sequencing of helix H3 following the same trend. Residues H152 and H157 located in this region were also found in bottom-up approach, where +16 Da (addition of oxygen) and -22 Da (His→Asp) modifications were assigned (Figure 26, p. 74, Figure 37, p. 91). Top-down data confirmed that both H152 and H157 are oxidized in the apo form, while in the holo form, the extents in this region are delayed (Figure 31, p. 80). Our findings are consistent with bottom-up data, and both approaches confirm that helix H3 creates main protein-DNA interface. Helix H3 is also sequenced by fragments z45 and z52, where H157 residue is located, and the increase of the extent is visible for the apo form, while in the holo form the extent does not increase (Figure 31, p. 80, Figure 38, p. 93). Further fragments z67 and z70 retain the same trend and that confirms this region is not oxidized. Notably, HDX data (Figure 32, p.83) displayed the same trend of significant deuterium difference in helix H3.

One of the advantages of bottom-up approach is that the extracted ion chromatograms can be used for any mass addition/loss (in case of -22 Da loss), while top-down data focuses only on +16 Da modifications. Any residue losses affected by oxidation are outside of the isolation window and were not isolated.

Fragment c92 covering region of strand S2 indicates it follows the trend of previous fragments, where the apo form is more oxidized. Residue F160 located on strand S2 (Figure 37, C, p. 91) was considered to be oxidized. Bottom-up approach did not reveal any information about this region due to missing peptides,

as well as about wing W1 region. Nevertheless, the top-down approach covered the region of wing W1 fragments z38 and z44. Notably, the intersection of c92 and z44 fragments is H164 and N165. The data are inconsistent in this case, because c92 fragment indicates that H164 is more oxidized in the apo form, while z44 fragment indicates H164 is more oxidized in the holo form. To explain that, H164 is part of a region covered by c92 fragment, where other reactive species are located. On the other hand, z44 fragment covering z38-z44 region, whereas H164-G169 residues are located and, in this case, H164 is the most reactive residue. Residue F160 is then oxidized in c92 fragment in the apo form, while H164 is just a part of this fragment ion. Fragment z44 points shows wing W1 is more oxidized in the apo form. All above assumed, strand S2 (with located residue F160) is more oxidized in the apo form, while wing W1 is rather solvent accessible in the holo form (Figure 38, p. 93 and Figure 39, p. 95). In this case, top-down data reveals the F160 to be the oxidized residue, while bottom-up did not provide any information about this region.

Last fragment obtained by ECD is c106 ion, which shows to be much more oxidized in the apo form than in the holo form. This region is also covered by z37 ion, where the extent jump occurred between z29-z37 fragments, but extensively higher for apo form. This region covered strand S3 and as shown protein topology in Figure 26, p. 74 and contains residues W173, W174 and M175. Bottom-up data revealed two conformers of W173, one of which is more and one less oxidized in the holo form. Residue W174 directly interacts with DNA and in the holo form is not oxidized and M175 residue is less oxidized in the holo form but still solvent accessible. Top-down data revealed only overall higher oxidation impact in this region in the apo form, with no further information about the particular residues due to low sequence resolution. In this case, M175 would be highlighted as a candidate for oxidation. CID y-ion series indicates an extensive jump between y32 and y34 fragment, which is caused by M175 oxidation. HDX data (Figure 32, p.83) contributes to the statement that strand S3 is less solvent accessible due to decreased overall flexibility upon DNA binding. Any further information is not provided by HDX of the C-terminus containing strand S2, wing W1 and strand S3 despite the fact, that this region indicates decreased structural flexibility in the presence of DNA[86].

The final C-terminal region contains residues of wing W2 and unstructured region. The truncated crystal structure available in PDB database[176] was missing the information about this region, but it was possible to model this region according to MS data[86]. This region is overall more solvent accessible in the holo form, where M194 residue was found to be more oxidized. The C-terminal region did not provide such extensive fragment ions as the N-terminus (see Figure 29, A, p. 78 for CID and Figure 33, A, p. 80, respectively). The first obtained CID fragment is y28 and provides information about overall higher solvent accessibility of this region in the holo form. However, further CID fragments are missing. The first fragment obtained by ECD is z14 containing M194 residue, y28 fragment in CID and bottom-up data agree that M194 is oxidized in the presence of DNA. The next z22 fragment decreases its extent in the holo form towards the apo form. This change was observed in the region of a cluster of basic residues (R188-190), which unspecifically interact with DNA. This may point out that this region is less solvent accessible due to DNA interaction. In this case, bottom-up data did not reveal any other residue to be oxidized except M194. Next fragment is z26, and again, it shows higher extent of oxidation and therefore increased solvent accessibility of this region. In summary, the bottom-up determined oxidized residue, the top-down sequencing data provided overall solvent accessibility of C-terminal region including W2 wing, where interacting residues are still less solvents accessible due to protein-DNA interaction. In this case, HDX data did not reveal any further information except lower solvent accessibility in the region containing M194 residue and almost no difference in deuteration in the region containing R188-R190 residues.

In conclusion, these results indicate that FPOP is a valuable tool to study protein-DNA interactions. Classical bottom-up approach was utilized to focus on each region and find, eventually verify, residues, which undergo FPOP oxidation. Some of these residues show lower or higher extent of oxidation compared to unbound form. Moreover, a novel top-down approach was used to isolate singly oxidized ion following fragmentation and further localize regions, where the oxidation differences occurred due to the presence of DNA. To conclude this, top-down fragmentation as a novel approach to obtain information about protein surfaces was successfully utilized. Both approaches were then compared with already published HDX data[86].

This concludes, that FPOP is a valuable tool to study protein-DNA interaction and is complementary to HDX. While HDX probes mostly protein backbone, FPOP focuses on protein side chains responsible for DNA recognition.

7. Conclusion

The aim of this diploma thesis was to investigate the potential of top-down approach for FPOP protein-ligand analysis.

The goals are achieved:

- DNA binding domain of FOXO4 transcription factor was successfully expressed.
- The ability of FOXO4-DBD to bind DNA was confirmed.
- Laser conditions was optimized to sufficiently oxidize free and bound form of protein with no further protein, or DNA damage.
- FPOP experiment was successfully performed on protein and protein-DNA complex.
- Bottom-up approach was used to localize individual residues of oxidation.
- Top-down approach was utilized as a novel approach to localize residues or at least regions, which are differently oxidized in the presence of DNA.
- Data were combined to map the protein-DNA surface.
- Finally, data were compared with HDX data and well known structural models of the apo (PDB entry: 1e17, ref. [167]) and holo (PDB entry: 3l2c, ref. [176]) forms.

In conclusion, fast photochemical oxidation of protein was used to monitor protein-DNA surface. Our findings indicate that data from this method can be combined with other MS-based techniques such as H/D exchange, due to mutual complementarity of both techniques.

8. Bibliography

(1)   Leney, A. C.; Heck, A. J. R. Native Mass Spectrometry: What Is in the Name? *J. Am. Soc. Mass Spectrom.* **2017**, *28* (1), 5–13. https://doi.org/10.1007/s13361-016-1545-3.

(2)   Konermann, L.; Vahidi, S.; Sowole, M. A. Mass Spectrometry Methods for Studying Structure and Dynamics of Biological Macromolecules. *Anal. Chem.* **2014**, *86* (1), 213–232. https://doi.org/10.1021/ac4039306.

(3)   Pauling, L.; Corey, R. B. The Pleated Sheet, a New Layer Configuration of Polypeptide Chains. *Proc. Natl. Acad. Sci. U. S. A.* **1951**, *37* (5), 251–256. https://doi.org/10.1073/pnas.37.5.251.

(4)   Watson, J. D.; Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171* (4356), 737–738. https://doi.org/10.1038/171737a0.

(5)   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.

(6)   Perutz, M. F.; Rossmann, M. G.; Cullis, A. F.; Muirhead, H.; Will, G.; North, A. C. T. Structure of Hemoglobin: A Three-Dimensional Fourier Synthesis at 5.5-Å. Resolution, Obtained by X-Ray Analysis. *Nature* **1960**, *185* (4711), 416–422. https://doi.org/10.1038/185416a0.

(7)   Finch, J. T.; Perutz, M. F.; Bertles, J. F.; Dobler, J. Structure of Sickled Erythrocytes and of Sickle-Cell Hemoglobin Fibers. *Proc. Natl. Acad. Sci.* **1973**, *70* (3), 718–722. https://doi.org/10.1073/pnas.70.3.718.

(8)   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.

(9)   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.

(10) Parker, M. W. Protein Structure from X-Ray Diffraction. *J. Biol. Phys.* **2003**, *29* (4), 341–362. https://doi.org/10.1023/A:1027310719146.

(11) 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.

(12) 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.

(13) Milne, J. L. S.; Borgnia, M. J.; Bartesaghi, A.; Tran, E. E. H.; Earl, L. A.; Schauder, D. M.; Lengyel, J.; Pierson, J.; Patwardhan, A.; Subramaniam, S. Cryo-Electron Microscopy - a Primer for the Non-Microscopist. *FEBS J.* **2013**, *280* (1), 28–45. https://doi.org/10.1111/febs.12078.

(14) Elmlund, D.; Elmlund, H. Cryogenic Electron Microscopy and Single-Particle Analysis. *Annu. Rev. Biochem.* **2015**, *84* (1), 499–517. https://doi.org/10.1146/annurev-biochem-060614-034226.

(15) Benesch, J. L. P.; Ruotolo, B. T. Mass Spectrometry: Come of Age for Structural and Dynamical Biology. *Curr. Opin. Struct. Biol.* **2011**, *21* (5), 641–649. https://doi.org/10.1016/j.sbi.2011.08.002.

(16) Rout, M. P.; Sali, A. Principles for Integrative Structural Biology Studies. *Cell* **2019**, *177* (6), 1384–1403. https://doi.org/10.1016/j.cell.2019.05.016.

(17) Jiang, J.; Chan, H.; Cash, D. D.; Miracco, E. J.; Ogorzalek Loo, R. R.; Upton, H. E.; Cascio, D.; O'Brien Johnson, R.; Collins, K.; Loo, J. A.; et al. Structure of Tetrahymena Telomerase Reveals Previously Unknown Subunits, Functions, and Interactions. *Science (80-.).* **2015**, *350* (6260), 529–539. https://doi.org/10.1126/science.aab4070.

(18) 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.

(19) 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.

(20) Hebert, A. S.; Richards, A. L.; Bailey, D. J.; Ulbrich, A.; Coughlin, E. E.; Westphall, M. S.; Coon, J. J. The One Hour Yeast Proteome. *Mol. Cell. Proteomics* **2014**, *13* (1), 339–347. https://doi.org/10.1074/mcp.M113.034769.

(21) Washburn, M. P.; Wolters, D.; Yates, J. R. Large-Scale Analysis of the Yeast Proteome by Multidimensional Protein Identification Technology. *Nat. Biotechnol.* **2001**, *19* (3), 242–247. https://doi.org/10.1038/85686.

(22) Cox, J.; Mann, M. MaxQuant Enables High Peptide Identification Rates, Individualized p.p.b.-Range Mass Accuracies and Proteome-Wide Protein Quantification. *Nat. Biotechnol.* **2008**, *26* (12), 1367–1372. https://doi.org/10.1038/nbt.1511.

(23) McCormack, A. L.; Schieltz, D. M.; Goode, B.; Yang, S.; Barnes, G.; Drubin, D.; Yates, J. R. Direct Analysis and Identification of Proteins in Mixtures by LC/MS/MS and Database Searching at the Low-Femtomole Level. *Anal. Chem.* **1997**, *69* (4), 767–776. https://doi.org/10.1021/ac960799q.

(24) 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.

(25) 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.

(26) Toby, T. K.; Fornelli, L.; Kelleher, N. L. Progress in Top-Down Proteomics and the Analysis of Proteoforms. *Annu. Rev. Anal. Chem.* **2016**, *9* (1), 499–519. https://doi.org/10.1146/annurev-anchem-071015-041550.

(27) 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.; et al. 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.

(28) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry. *J. Am. Chem. Soc.* **1999**, *121* (4), 806–812. https://doi.org/10.1021/ja973655h.

(29) 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.

(30) Snijder, J.; Heck, A. J. R. Analytical Approaches for Size and Mass Analysis of Large Protein Assemblies. *Annu. Rev. Anal. Chem.* **2014**, *7* (1), 43–64. https://doi.org/10.1146/annurev-anchem-071213-020015.

(31) Lorenzen, K.; Duijn, E. van. Native Mass Spectrometry as a Tool in Structural Biology. *Curr. Protoc. Protein Sci.* **2010**, *62* (1), 1–17. https://doi.org/10.1002/0471140864.ps1712s62.

(32) 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), 5014–5030. https://doi.org/10.1039/c3cs35477f.

(33) 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.

(34) Bruce, C.; Stone, K.; Gulcicek, E.; Williams, K. Proteomics and the Analysis of Proteomic Data: 2013 Overview of Current Protein-Profiling Technologies. *Curr. Protoc. Bioinforma.* **2013**, *41* (1), 1–17. https://doi.org/10.1002/0471250953.bi1321s41.

(35) Dodds, J. N.; Baker, E. S. Ion Mobility Spectrometry: Fundamental Concepts, Instrumentation, Applications, and the Road Ahead. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (11), 2185–2195. https://doi.org/10.1007/s13361-019-02288-2.

(36) Kanu, A. B.; Dwivedi, P.; Tam, M.; Matz, L.; Hill, H. H. Ion Mobility-Mass Spectrometry. *J. Mass Spectrom.* **2008**, *43* (1), 1–22. https://doi.org/10.1002/jms.1383.

(37) Gabelica, V.; Marklund, E. Fundamentals of Ion Mobility Spectrometry. *Curr. Opin. Chem. Biol.* **2018**, *42*, 51–59. https://doi.org/10.1016/j.cbpa.2017.10.022.

(38) Cumeras, R.; Figueras, E.; Davis, C. E.; Baumbach, J. I.; Gràcia, I. Review on Ion Mobility Spectrometry. Part 1: Current Instrumentation. *Analyst* **2015**, *140* (5), 1376–1390. https://doi.org/10.1039/C4AN01100G.

(39) Chouinard, C. D.; Wei, M. S.; Beekman, C. R.; Kemperman, R. H. J.; Yost, R. A. Ion Mobility in Clinical Analysis: Current Progress and Future Perspectives. *Clin. Chem.* **2016**, *62* (1), 124–133. https://doi.org/10.1373/clinchem.2015.238840.

(40) Bleiholder, C.; Dupuis, N. F.; Wyttenbach, T.; Bowers, M. T. Ion Mobility-Mass Spectrometry Reveals a Conformational Conversion from Random Assembly to β-Sheet in Amyloid Fibril Formation. *Nat. Chem.* **2011**, *3* (2), 172–177. https://doi.org/10.1038/nchem.945.

(41) Dixit, S. M.; Polasky, D. A.; Ruotolo, B. T. Collision Induced Unfolding of Isolated Proteins in the Gas Phase: Past, Present, and Future. *Curr. Opin. Chem. Biol.* **2018**, *42*, 93–100. https://doi.org/10.1016/j.cbpa.2017.11.010.

(42) Heck, A. J. R. Native Mass Spectrometry: A Bridge between Interactomics and Structural Biology. *Nat. Methods* **2008**, *5* (11), 927–933. https://doi.org/10.1038/nmeth.1265.

(43) Konijnenberg, A.; Butterer, A.; Sobott, F. Native Ion Mobility-Mass Spectrometry and Related Methods in Structural Biology. *Biochim. Biophys. Acta - Proteins Proteomics* **2013**, *1834* (6), 1239–1256. https://doi.org/10.1016/j.bbapap.2012.11.013.

(44) Hilton, G. R.; Benesch, J. L. P. Two Decades of Studying Non-Covalent Biomolecular Assemblies by Means of Electrospray Ionization Mass Spectrometry. *J. R. Soc. Interface* **2012**, *9* (70), 801–816. https://doi.org/10.1098/rsif.2011.0823.

(45) Boeri Erba, E.; Petosa, C. The Emerging Role of Native Mass Spectrometry in Characterizing the Structure and Dynamics of Macromolecular Complexes. *Protein Sci.* **2015**, *24* (8), 1176–1192. https://doi.org/10.1002/pro.2661.

(46) Konermann, L. Addressing a Common Misconception: Ammonium Acetate as Neutral PH "Buffer" for Native Electrospray Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (9), 1827–1835. https://doi.org/10.1007/s13361-017-1739-3.

(47) van den Heuvel, R. H. H.; van Duijn, E.; Mazon, H.; Synowsky, S. A.; Lorenzen, K.; Versluis, C.; Brouns, S. J. J.; Langridge, D.; van der Oost, J.; Hoyes, J.; et al. Improving the Performance of a Quadrupole Time-of-Flight Instrument for Macromolecular Mass Spectrometry. *Anal. Chem.* **2006**, *78* (21), 7473–7483. https://doi.org/10.1021/ac061039a.

(48) Snijder, J.; Rose, R. J.; Veesler, D.; Johnson, J. E.; Heck, A. J. R. Studying 18 MDa Virus Assemblies with Native Mass Spectrometry. *Angew. Chemie Int. Ed.* **2013**, *52* (14), 4020–4023. https://doi.org/10.1002/anie.201210197.

(49) Rostom, A. A.; Fucini, P.; Benjamin, D. R.; Juenemann, R.; Nierhaus, K. H.; Hartl, F. U.; Dobson, C. M.; Robinson, C. V. Detection and Selective Dissociation of Intact Ribosomes in a Mass Spectrometer. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (10), 5185–5190. https://doi.org/10.1073/pnas.97.10.5185.

(50) Zhou, M.; Sandercock, A. M.; Fraser, C. S.; Ridlova, G.; Stephens, E.; Schenauer, M. R.; Yokoi-Fong, T.; Barsky, D.; Leary, J. A.; Hershey, J. W.; et al. Mass Spectrometry Reveals Modularity and a Complete Subunit Interaction Map of the Eukaryotic Translation Factor EIF3. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (47), 18139–18144. https://doi.org/10.1073/pnas.0801313105.

(51) Ebong, I. O.; Morgner, N.; Zhou, M.; Saraiva, M. A.; Daturpalli, S.; Jackson, S. E.; Robinson, C. V. Heterogeneity and Dynamics in the Assembly of the Heat Shock Protein 90 Chaperone Complexes. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (44), 17939–17944. https://doi.org/10.1073/pnas.1106261108.

(52) Jore, M. M.; Lundgren, M.; Van Duijn, E.; Bultema, J. B.; Westra, E. R.; Waghmare, S. P.; Wiedenheft, B.; Pul, Ü.; Wurm, R.; Wagner, R.; et al. Structural Basis for CRISPR RNA-Guided DNA Recognition by Cascade. *Nat. Struct. Mol. Biol.* **2011**, *18* (5), 529–536. https://doi.org/10.1038/nsmb.2019.

(53) Demmers, J. A. A.; van Dalen, A.; de Kruijff, B.; Heck, A. J. R.; Killian, J. A. Interaction of the K+ Channel KcsA with Membrane Phospholipids as Studied

(46) Konermann, L. Addressing a Common Misconception: Ammonium Acetate as Neutral PH "Buffer" for Native Electrospray Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (9), 1827–1835. https://doi.org/10.1007/s13361-017-1739-3.

(47) van den Heuvel, R. H. H.; van Duijn, E.; Mazon, H.; Synowsky, S. A.; Lorenzen, K.; Versluis, C.; Brouns, S. J. J.; Langridge, D.; van der Oost, J.; Hoyes, J.; et al. Improving the Performance of a Quadrupole Time-of-Flight Instrument for Macromolecular Mass Spectrometry. *Anal. Chem.* **2006**, *78* (21), 7473–7483. https://doi.org/10.1021/ac061039a.

(48) Snijder, J.; Rose, R. J.; Veesler, D.; Johnson, J. E.; Heck, A. J. R. Studying 18 MDa Virus Assemblies with Native Mass Spectrometry. *Angew. Chemie Int. Ed.* **2013**, *52* (14), 4020–4023. https://doi.org/10.1002/anie.201210197.

(49) Rostom, A. A.; Fucini, P.; Benjamin, D. R.; Juenemann, R.; Nierhaus, K. H.; Hartl, F. U.; Dobson, C. M.; Robinson, C. V. Detection and Selective Dissociation of Intact Ribosomes in a Mass Spectrometer. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (10), 5185–5190. https://doi.org/10.1073/pnas.97.10.5185.

(50) Zhou, M.; Sandercock, A. M.; Fraser, C. S.; Ridlova, G.; Stephens, E.; Schenauer, M. R.; Yokoi-Fong, T.; Barsky, D.; Leary, J. A.; Hershey, J. W.; et al. Mass Spectrometry Reveals Modularity and a Complete Subunit Interaction Map of the Eukaryotic Translation Factor EIF3. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (47), 18139–18144. https://doi.org/10.1073/pnas.0801313105.

(51) Ebong, I. O.; Morgner, N.; Zhou, M.; Saraiva, M. A.; Daturpalli, S.; Jackson, S. E.; Robinson, C. V. Heterogeneity and Dynamics in the Assembly of the Heat Shock Protein 90 Chaperone Complexes. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (44), 17939–17944. https://doi.org/10.1073/pnas.1106261108.

(52) Jore, M. M.; Lundgren, M.; Van Duijn, E.; Bultema, J. B.; Westra, E. R.; Waghmare, S. P.; Wiedenheft, B.; Pul, Ü.; Wurm, R.; Wagner, R.; et al. Structural Basis for CRISPR RNA-Guided DNA Recognition by Cascade. *Nat. Struct. Mol. Biol.* **2011**, *18* (5), 529–536. https://doi.org/10.1038/nsmb.2019.

(53) Demmers, J. A. A.; van Dalen, A.; de Kruijff, B.; Heck, A. J. R.; Killian, J. A. Interaction of the K+ Channel KcsA with Membrane Phospholipids as Studied

by ESI Mass Spectrometry. *FEBS Lett.* **2003**, *541* (1–3), 28–32. https://doi.org/10.1016/s0014-5793(03)00282-5.

(54) Laganowsky, A.; Reading, E.; Allison, T. M.; Ulmschneider, M. B.; Degiacomi, M. T.; Baldwin, A. J.; Robinson, C. V. Membrane Proteins Bind Lipids Selectively to Modulate Their Structure and Function. *Nature* **2014**, *510* (7503), 172–175. https://doi.org/10.1038/nature13419.

(55) Cousido-Siah, A.; Ayoub, D.; Berberián, G.; Bollo, M.; Van Dorsselaer, A.; Debaene, F.; DiPolo, R.; Petrova, T.; Schulze-Briese, C.; Olieric, V.; et al. Structural and Functional Studies of ReP1-NCXSQ, a Protein Regulating the Squid Nerve Na$^+$/Ca^{2+} Exchanger. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68* (9), 1098–1107. https://doi.org/10.1107/S090744491202094X.

(56) Marchand, A.; Gabelica, V. Native Electrospray Mass Spectrometry of DNA G-Quadruplexes in Potassium Solution. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (7), 1146–1154. https://doi.org/10.1007/s13361-014-0890-3.

(57) Ilag, L. L.; Ubarretxena-Belandia, I.; Tate, C. G.; Robinson, C. V. Drug Binding Revealed by Tandem Mass Spectrometry of a Protein-Micelle Complex. *J. Am. Chem. Soc.* **2004**, *126* (44), 14362–14363. https://doi.org/10.1021/ja0450307.

(58) Cheng, X.; Chen, R.; Bruce, J. E.; Schwartz, B. L.; Anderson, G. A.; Hofstadler, S. A.; Gale, D. C.; Smith, R. D.; Gao, J.; Sigal, G. B.; et al. Using Electrospray Ionization FTICR Mass Spectrometry To Study Competitive Binding of Inhibitors to Carbonic Anhydrase. *J. Am. Chem. Soc.* **1995**, *117* (34), 8859–8860. https://doi.org/10.1021/ja00139a023.

(59) Chevreux, G.; Atmanene, C.; Lopez, P.; Ouazzani, J.; Van Dorsselaer, A.; Badet, B.; Badet-Denisot, M. A.; Sanglier-Cianférani, S. Monitoring the Dynamics of Monomer Exchange Using Electrospray Mass Spectrometry: The Case of the Dimeric Glucosamine-6-Phosphate Synthase. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (3), 431–439. https://doi.org/10.1007/s13361-010-0054-z.

(60) Piotrowski, C.; Sinz, A. Structural Investigation of Proteins and Protein Complexes by Chemical Cross-Linking/Mass Spectrometry. In *Advances in Experimental Medicine and Biology*; Springer New York LLC, 2018; Vol. 1105, pp 101–121. https://doi.org/10.1007/978-981-13-2200-6_8.

(61) Greber, B. J.; Boehringer, D.; Leitner, A.; Bieri, P.; Voigts-Hoffmann, F.; Erzberger, J. P.; Leibundgut, M.; Aebersold, R.; Ban, N. Architecture of the Large Subunit of the Mammalian Mitochondrial Ribosome. *Nature* **2014**, *505* (7484), 515–519. https://doi.org/10.1038/nature12890.

(62) Rozbeský, D.; Rosůlek, M.; Kukačka, Z.; Chmelík, J.; Man, P.; Novák, P. Impact of Chemical Cross-Linking on Protein Structure and Function. *Anal. Chem.* **2018**, *90* (2), 1104–1113. https://doi.org/10.1021/acs.analchem.7b02863.

(63) Sinz, A. Chemical Cross-Linking and Mass Spectrometry to Map Three-Dimensional Protein Structures and Protein–Protein Interactions. *Mass Spectrom. Rev.* **2006**, *25* (4), 663–682. https://doi.org/10.1002/mas.20082.

(64) Götze, M.; Pettelkau, J.; Schaks, S.; Bosse, K.; Ihling, C. H.; Krauth, F.; Fritzsche, R.; Kühn, U.; Sinz, A. StavroX-A Software for Analyzing Crosslinked Products in Protein Interaction Studies. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (1), 76–87. https://doi.org/10.1007/s13361-011-0261-2.

(65) Peri, S.; Steen, H.; Pandey, A. GPMAW – a Software Tool for Analyzing Proteins and Peptides. *Trends Biochem. Sci.* **2001**, *26* (11), 687–689. https://doi.org/10.1016/S0968-0004(01)01954-5.

(66) Taverner, T.; Hall, N. E.; O'Hair, R. A. J.; Simpson, R. J. Characterization of an Antagonist Interleukin-6 Dimer by Stable Isotope Labeling, Cross-Linking, and Mass Spectrometry. *J. Biol. Chem.* **2002**, *277* (48), 46487–46492. https://doi.org/10.1074/jbc.M207370200.

(67) Mohr, J. P.; Perumalla, P.; Chavez, J. D.; Eng, J. K.; Bruce, J. E. Mango: A General Tool for Collision Induced Dissociation-Cleavable Cross-Linked Peptide Identification. *Anal. Chem.* **2018**, *90* (10), 6028–6034. https://doi.org/10.1021/acs.analchem.7b04991.

(68) Götze, M.; Iacobucci, C.; Ihling, C. H.; Sinz, A. A Simple Cross-Linking/Mass Spectrometry Workflow for Studying System-Wide Protein Interactions. *Anal. Chem.* **2019**, *91* (15), 10236–10244. https://doi.org/10.1021/acs.analchem.9b02372.

(69) Iacobucci, C.; Götze, M.; Sinz, A. Cross-Linking/Mass Spectrometry to Get a

Closer View on Protein Interaction Networks. *Current Opinion in Biotechnology*. Elsevier Ltd June 1, 2020, pp 48–53. https://doi.org/10.1016/j.copbio.2019.12.009.

(70) Liu, F.; Rijkers, D. T. S.; Post, H.; Heck, A. J. R. Proteome-Wide Profiling of Protein Assemblies by Cross-Linking Mass Spectrometry. *Nat. Methods* **2015**, *12* (12), 1179–1184. https://doi.org/10.1038/nmeth.3603.

(71) Schweppe, D. K.; Chavez, J. D.; Lee, C. F.; Caudal, A.; Kruse, S. E.; Stuppard, R.; Marcinek, D. J.; Shadel, G. S.; Tian, R.; Bruce, J. E. Mitochondrial Protein Interactome Elucidated by Chemical Cross-Linking Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (7), 1732–1737. https://doi.org/10.1073/pnas.1617220114.

(72) Chavez, J. D.; Lee, C. F.; Caudal, A.; Keller, A.; Tian, R.; Bruce, J. E. Chemical Crosslinking Mass Spectrometry Analysis of Protein Conformations and Supercomplexes in Heart Tissue. *Cell Syst.* **2018**, *6* (1), 136–141. https://doi.org/10.1016/j.cels.2017.10.017.

(73) Timerbaev, A. R.; Hartinger, C. G.; Aleksenko, S. S.; Keppler, B. K. Interactions of Antitumor Metallodrugs with Serum Proteins: Advances in Characterization Using Modern Analytical Methodology. *Chem. Rev.* **2006**, *106* (6), 2224–2248. https://doi.org/10.1021/cr040704h.

(74) Fontana, A.; De Laureto, P. P.; Spolaore, B.; Frare, E.; Picotti, P.; Zambonin, M. Probing Protein Structure by Limited Proteolysis. In *Acta Biochimica Polonica*; 2004; Vol. 51, pp 299–321. https://doi.org/035001299.

(75) Hvidt, A.; Nielsen, S. O. Hydrogen Exchange in Proteins. In *Advances in Protein Chemistry*; 1966; Vol. 21, pp 287–386. https://doi.org/10.1016/S0065-3233(08)60129-1.

(76) 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.

(77) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. Protein and Polymer Analyses up to m/z 100 000 by Laser Ionization Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1988**, *2* (8), 151–

153. https://doi.org/10.1002/rcm.1290020802.

(78)   Oganesyan, I.; Lento, C.; Wilson, D. J. Contemporary Hydrogen Deuterium Exchange Mass Spectrometry. *Methods* **2018**, *144*, 27–42. https://doi.org/10.1016/j.ymeth.2018.04.023.

(79)   Masson, G. R.; Jenkins, M. L.; Burke, J. E. An Overview of Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) in Drug Discovery. *Expert Opin. Drug Discov.* **2017**, *12* (10), 981–994. https://doi.org/10.1080/17460441.2017.1363734.

(80)   Percy, A. J.; Rey, M.; Burns, K. M.; Schriemer, D. C. Probing Protein Interactions with Hydrogen/Deuterium Exchange and Mass Spectrometry—A Review. *Anal. Chim. Acta* **2012**, *721*, 7–21. https://doi.org/10.1016/j.aca.2012.01.037.

(81)   Brown, K. A.; Wilson, D. J. Bottom-up Hydrogen Deuterium Exchange Mass Spectrometry: Data Analysis and Interpretation. *Analyst* **2017**, *142* (16), 2874–2886. https://doi.org/10.1039/C7AN00662D.

(82)   Pan, J.; Han, J.; Borchers, C. H.; Konermann, L. Hydrogen/Deuterium Exchange Mass Spectrometry with Top-Down Electron Capture Dissociation for Characterizing Structural Transitions of a 17 KDa Protein. *J. Am. Chem. Soc.* **2009**, *131* (35), 12801–12808. https://doi.org/10.1021/ja904379w.

(83)   Vankova; Salido; Timson; Man; Pey. A Dynamic Core in Human NQO1 Controls the Functional and Stability Effects of Ligand Binding and Their Communication across the Enzyme Dimer. *Biomolecules* **2019**, *9* (11), 728. https://doi.org/10.3390/biom9110728.

(84)   Man, P.; Montagner, C.; Vitrac, H.; Kavan, D.; Pichard, S.; Gillet, D.; Forest, E.; Forge, V. Accessibility Changes within Diphtheria Toxin T Domain upon Membrane Penetration Probed by Hydrogen Exchange and Mass Spectrometry. *J. Mol. Biol.* **2011**, *414* (1), 123–134. https://doi.org/10.1016/j.jmb.2011.09.042.

(85)   Trcka, F.; Durech, M.; Vankova, P.; Chmelik, J.; Martinkova, V.; Hausner, J.; Kadek, A.; Marcoux, J.; Klumpler, T.; Vojtesek, B.; et al. Human Stress-Inducible Hsp70 Has a High Propensity to Form ATP-Dependent Antiparallel

Dimers That Are Differentially Regulated by Cochaperone Binding. *Mol. Cell. Proteomics* **2019**, *18* (2), 320–337. https://doi.org/10.1074/mcp.RA118.001044.

(86) Slavata; Chmelík; Kavan; Filandrová; Fiala; Rosůlek; Mrázek; Kukačka; Vališ; Man; et al. 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.

(87) Laganowsky, A.; Reading, E.; Hopper, J. T. S.; Robinson, C. V. Mass Spectrometry of Intact Membrane Protein Complexes. *Nat. Protoc.* **2013**, *8* (4), 639–651. https://doi.org/10.1038/nprot.2013.024.

(88) Masson, G. R.; Maslen, S. L.; Williams, R. L. Analysis of Phosphoinositide 3-Kinase Inhibitors by Bottom-up Electron-Transfer Dissociation Hydrogen/Deuterium Exchange Mass Spectrometry. *Biochem. J.* **2017**, *474* (11), 1867–1877. https://doi.org/10.1042/BCJ20170127.

(89) Hoofnagle, A. N.; Resing, K. A.; Ahn, N. G. Protein Analysis by Hydrogen Exchange Mass Spectrometry. *Annual Review of Biophysics and Biomolecular Structure*. 2003, pp 1–25. https://doi.org/10.1146/annurev.biophys.32.110601.142417.

(90) Wang, L.; Chance, M. R. Structural Mass Spectrometry of Proteins Using Hydroxyl Radical Based Protein Footprinting. *Anal. Chem.* **2011**, *83* (19), 7234–7241. https://doi.org/10.1021/ac200567u.

(91) 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.

(92) 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.

(93) Glocker, M. O.; Borchers, C.; Fiedler, W.; Suckau, D.; Przybylski, M.

Molecular Characterization of Surface Topology in Protein Tertiary Structures by Amino-Acylation and Mass Spectrometric Peptide Mapping. *Bioconjug. Chem.* **1994**, *5* (6), 583–590. https://doi.org/10.1021/bc00030a014.

(94)   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.

(95)   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.

(96)   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.

(97)   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.

(98)   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.

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

(100) Xu, G.; Chance, M. R. Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chem. Rev.* **2007**, *107* (8), 3514–3543. https://doi.org/10.1021/cr0682047.

(101) 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.

(102) 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.

(103) 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. *Anal. Biochem.* **2001**, *289* (2), 103–115. https://doi.org/10.1006/ABIO.2000.4910.

(104) 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.

(105) 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.

(106) 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.

(107) 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.

(108) 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.

(109) 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.

(110) 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.

(111) 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.

(112) 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.

(113) 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.

(114) Humayun, Z.; Kleid, D.; Ptashne, M. Sites of Contact between λ Operators and λ Repressor. *Nucleic Acids Res.* **1977**, *4* (5), 1595–1608. https://doi.org/10.1093/nar/4.5.1595.

(115) 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.

(116) Aruoma, O. I.; Halliwell, B.; Gajewski, E.; Dizdaroglu, M. Damage to the Bases in DNA Induced by Hydrogen Peroxide and Ferric Ion Chelates. *J. Biol. Chem.* **1989**, *264* (34), 20509–20512.

(117) Sheshberadaran, H.; Payne, L. G. Protein Antigen-Monoclonal Antibody Contact Sites Investigated by Limited Proteolysis of Monoclonal Antibody-Bound Antigen: Protein "Footprinting". *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85* (1), 1–5. https://doi.org/10.1073/pnas.85.1.1.

(118) 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.

(119) 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**, *144*, 94–103. https://doi.org/10.1016/j.ymeth.2018.05.016.

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

(121) 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.

(122) 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.

(123) 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.

(124) 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.

(125) Kiselar, J.; Chance, M. R. High-Resolution Hydroxyl Radical Protein Footprinting: Biophysics Tool for Drug Discovery. *Annu. Rev. Biophys.* **2018**, *47* (1), 315–333. https://doi.org/10.1146/annurev-biophys-070317-033123.

(126) Abolhasani Khaje, N.; Mobley, C. K.; Misra, S. K.; Miller, L.; Li, Z.; Nudler, E.; Sharp, J. S. Variation in FPOP Measurements Is Primarily Caused by Poor Peptide Signal Intensity. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (9), 1901–1907. https://doi.org/10.1007/s13361-018-1994-y.

(127) Chea, E. E.; Jones, L. M. Modifications Generated by Fast Photochemical Oxidation of Proteins Reflect the Native Conformations of Proteins. *Protein Sci.* **2018**, *27* (6), 1047–1056. https://doi.org/10.1002/pro.3408.

(128) 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.

(129) Li, K. S.; Rempel, D. L.; Gross, M. L. Conformational-Sensitive Fast Photochemical Oxidation of Proteins and Mass Spectrometry Characterize Amyloid Beta 1–42 Aggregation. *J. Am. Chem. Soc.* **2016**, *138* (37), 12090–12098. https://doi.org/10.1021/jacs.6b07543.

(130) 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.

(131) 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.

(132) Li, J.; Wei, H.; Krystek, S. R.; Bond, D.; Brender, T. M.; Cohen, D.; Feiner, J.; Hamacher, N.; Harshman, J.; Huang, R. Y. C.; et al. 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.

(133) 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.

(134) Espino, J. A.; Mali, V. S.; Jones, L. M. In Cell Footprinting Coupled with Mass Spectrometry for the Structural Analysis of Proteins in Live Cells. *Anal. Chem.* **2015**, *87* (15), 7971–7978. https://doi.org/10.1021/acs.analchem.5b01888.

(135) Espino, J. A.; Jones, L. M. Illuminating Biological Interactions with in Vivo Protein Footprinting. *Anal. Chem.* **2019**.

https://doi.org/10.1021/acs.analchem.9b00244.

(136) 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.

(137) Gau, B. C.; Sharp, J. S.; Rempel, D. L.; Gross, M. L. Fast Photochemical Oxidation of Protein Footprints Faster than Protein Unfolding. *Anal. Chem.* **2009**, *81* (16), 6563–6571. https://doi.org/10.1021/ac901054w.

(138) Heinkel, F.; Gsponer, J. Determination of Protein Folding Intermediate Structures Consistent with Data from Oxidative Footprinting Mass Spectrometry. *J. Mol. Biol.* **2016**, *428* (2), 365–371. https://doi.org/10.1016/j.jmb.2015.10.022.

(139) Aprahamian, M. L.; Chea, E. E.; Jones, L. M.; Lindert, S. Rosetta Protein Structure Prediction from Hydroxyl Radical Protein Footprinting Mass Spectrometry Data. *Anal. Chem.* **2018**, *90* (12), 7721–7729. https://doi.org/10.1021/acs.analchem.8b01624.

(140) Misra, S. K.; Sood, A.; Soares, P. A.; Pomin, V. H.; Woods, R. J.; Sharp, J. S. Mapping of the Fondaparinux Binding Site of JR-FL Gp120 by High Resolution Hydroxyl Radical Protein Footprinting and Computational Docking. *bioRxiv* **2017**, 207910. https://doi.org/10.1101/207910.

(141) 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.

(142) Latchman, D. S. Transcription Factors: An Overview. *Int. J. Biochem. Cell Biol.* **1997**, *29* (12), 1305–1312. https://doi.org/10.1016/S1357-2725(97)00085-X.

(143) Todeschini, A.-L.; Georges, A.; Veitia, R. A. Transcription Factors: Specific DNA Binding and Specific Gene Regulation. *Trends Genet.* **2014**, *30* (6), 211–219. https://doi.org/10.1016/j.tig.2014.04.002.

(144) Spitz, F.; Furlong, E. E. M. Transcription Factors: From Enhancer Binding to Developmental Control. *Nat. Rev. Genet.* **2012**, *13* (9), 613–626.

https://doi.org/10.1038/nrg3207.

(145) Wingender, E.; Schoeps, T.; Dönitz, J. TFClass: An Expandable Hierarchical Classification of Human Transcription Factors. *Nucleic Acids Res.* **2013**, *41* (D1), 165–170. https://doi.org/10.1093/nar/gks1123.

(146) Weiss, M. A.; Ellenberger, T.; Wobbe, C. R.; Lee, J. P.; Harrison, S. C.; Struhl, K. Folding Transition in the DMA-Binding Domain of GCN4 on Specific Binding to DNA. *Nature* **1990**, *347* (6293), 575–578. https://doi.org/10.1038/347575a0.

(147) Klug, A. The Discovery of Zinc Fingers and Their Applications in Gene Regulation and Genome Manipulation. *Annu. Rev. Biochem.* **2010**, *79* (1), 213–231. https://doi.org/10.1146/annurev-biochem-010909-095056.

(148) Aravind, L.; Anantharaman, V.; Balaji, S.; Babu, M.; Iyer, L. The Many Faces of the Helix-Turn-Helix Domain: Transcription Regulation and Beyond. *FEMS Microbiol. Rev.* **2005**, *29* (2), 231–262. https://doi.org/10.1016/j.femsre.2004.12.008.

(149) Gaston, K.; Jayaraman, P.-S. Transcriptional Repression in Eukaryotes: Repressors and Repression Mechanisms. *Cell. Mol. Life Sci.* **2003**, *60* (4), 721–741. https://doi.org/10.1007/s00018-003-2260-3.

(150) Yamamoto, Y.; Gaynor, R. B. IκB Kinases: Key Regulators of the NF-KB Pathway. *Trends Biochem. Sci.* **2004**, *29* (2), 72–79. https://doi.org/10.1016/j.tibs.2003.12.003.

(151) Grad, I.; Picard, D. The Glucocorticoid Responses Are Shaped by Molecular Chaperones. *Mol. Cell. Endocrinol.* **2007**, *275* (1–2), 2–12. https://doi.org/10.1016/j.mce.2007.05.018.

(152) Filtz, T. M.; Vogel, W. K.; Leid, M. Regulation of Transcription Factor Activity by Interconnected Post-Translational Modifications. *Trends Pharmacol. Sci.* **2014**, *35* (2), 76–85. https://doi.org/10.1016/j.tips.2013.11.005.

(153) Obsil, T.; Obsilova, V. Structure/Function Relationships Underlying Regulation of FOXO Transcription Factors. *Oncogene* **2008**, *27* (16), 2263–2275. https://doi.org/10.1038/onc.2008.20.

(154) Gajiwala, K. S.; Burley, S. K. Winged Helix Proteins. *Curr. Opin. Struct. Biol.* **2000**, *10* (1), 110–116. https://doi.org/10.1016/S0959-440X(99)00057-3.

(155) Kaestner, K. H.; Knochel, W.; Martinez, D. E. Unified Nomenclature for the Winged Helix/Forkhead Transcription Factors. *Genes Dev.* **2000**, *14* (2), 142–146. https://doi.org/10.1101/GAD.14.2.142.

(156) Clark, K. L.; Halay, E. D.; Lai, E.; Burley, S. K. Co-Crystal Structure of the HNF-3/Fork Head DNA-Recognition Motif Resembles Histone H5. *Nature* **1993**, *364* (6436), 412–420. https://doi.org/10.1038/364412a0.

(157) Weigel, D.; Jäckle, H. The Fork Head Domain: A Novel DNA Binding Motif of Eukaryotic Transcription Factors? *Cell* **1990**, *63* (3), 455–456. https://doi.org/10.1016/0092-8674(90)90439-L.

(158) Weigel, D.; Jürgens, G.; Küttner, F.; Seifert, E.; Jäckle, H. The Homeotic Gene Fork Head Encodes a Nuclear Protein and Is Expressed in the Terminal Regions of the Drosophila Embryo. *Cell* **1989**, *57* (4), 645–658. https://doi.org/10.1016/0092-8674(89)90133-5.

(159) Tuteja, G.; Kaestner, K. H. SnapShot:Forkhead Transcription Factors I. *Cell* **2007**, *130* (6), 1160–1160. https://doi.org/10.1016/j.cell.2007.09.005.

(160) Tuteja, G.; Kaestner, K. H. SnapShot: Forkhead Transcription Factors II. *Cell* **2007**, *131* (1), 192–192. https://doi.org/10.1016/j.cell.2007.09.016.

(161) Calnan, D. R.; Brunet, A. The FoxO Code. *Oncogene* **2008**, *27* (16), 2276–2288. https://doi.org/10.1038/onc.2008.21.

(162) Furuyama, T.; Nakazawa, T.; Nakano, I.; Mori, N. Identification of the Differential Distribution Patterns of MRNAs and Consensus Binding Sequences for Mouse DAF-16 Homologues. *Biochem. J.* **2000**, *349* (2), 629. https://doi.org/10.1042/0264-6021:3490629.

(163) Fu, Z.; Tindall, D. J. FOXOs, Cancer and Regulation of Apoptosis. *Oncogene* **2008**, *27* (16), 2312–2319. https://doi.org/10.1038/onc.2008.24.

(164) van der Heide, L. P.; Jacobs, F. M. J.; Burbach, J. P. H.; Hoekman, M. F. M.; Smidt, M. P. FoxO6 Transcriptional Activity Is Regulated by Thr26 and Ser184, Independent of Nucleo-Cytoplasmic Shuttling. *Biochem. J.* **2005**, *391* (3), 623–

629. https://doi.org/10.1042/BJ20050525.

(165) Zhang, X.; Tang, N.; Hadden, T. J.; Rishi, A. K. Akt, FoxO and Regulation of Apoptosis. *Biochim. Biophys. Acta - Mol. Cell Res.* **2011**, *1813* (11), 1978–1986. https://doi.org/10.1016/j.bbamcr.2011.03.010.

(166) Borkhardt, A.; Repp, R.; Haas, O. A.; Leis, T.; Harbott, J.; Kreuder, J.; Hammermann, J.; Henn, T.; Lampert, F. Cloning and Characterization of AFX, the Gene That Fuses to MLL in Acute Leukemias with a t(X;11)(Q13;Q23). *Oncogene* **1997**, *14* (2), 195–202. https://doi.org/10.1038/sj.onc.1200814.

(167) Weigelt, J.; Climent, I.; Dahlman-Wright, K.; Wikström, M. Solution Structure of the DNA Binding Domain of the Human Forkhead Transcription Factor AFX (FOXO4). *Biochemistry* **2001**, *40* (20), 5861–5869. https://doi.org/10.1021/bi001663w.

(168) Vogt, P. K.; Jiang, H.; Aoki, M. Triple Layer Control: Phosphorylation, Acetylation and Ubiquitination of FOXO Proteins. *Cell Cycle* **2005**, *4* (7), 908–913. https://doi.org/10.4161/cc.4.7.1796.

(169) 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.

(170) 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.

(171) Kops, G. J. P. L.; De Ruiter, N. D.; De Vries-Smits, A. M. M.; Powell, D. R.; Bos, J. L.; Burgering, B. M. T. Direct Control of the Forkhead Transcription Faotor AFX by Protein Kinase B. *Nature* **1999**, *398* (6728), 630–634. https://doi.org/10.1038/19328.

(172) Obsil, T.; Ghirlando, R.; Anderson, D. E.; Hickman, A. B.; Dyda, F. Two 14-3-3 Binding Motifs Are Required for Stable Association of Forkhead Transcription Factor FOXO4 with 14-3-3 Proteins and Inhibition of DNA Binding. *Biochemistry* **2003**, *42* (51), 15264–15272.

https://doi.org/10.1021/bi0352724.

(173) Essers, M. A. G.; Weijzen, S.; de Vries-Smits, A. M. M.; Saarloos, I.; de Ruiter, N. D.; Bos, J. L.; Burgering, B. M. T. FOXO Transcription Factor Activation by Oxidative Stress Mediated by the Small GTPase Ral and JNK. *EMBO J.* **2004**, *23* (24), 4802–4812. https://doi.org/10.1038/sj.emboj.7600476.

(174) De Ruiter, N. D.; Burgering, B. M. T.; Bos, J. L. Regulation of the Forkhead Transcription Factor AFX by Ral-Dependent Phosphorylation of Threonines 447 and 451. *Mol. Cell. Biol.* **2001**, *21* (23), 8225–8235. https://doi.org/10.1128/mcb.21.23.8225-8235.2001.

(175) Tsuruta, F.; Sunayama, J.; Mori, Y.; Hattori, S.; Shimizu, S.; Tsujimoto, Y.; Yoshioka, K.; Masuyama, N.; Gotoh, Y. JNK Promotes Bax Translocation to Mitochondria through Phosphorylation of 14-3-3 Proteins. *EMBO J.* **2004**, *23* (8), 1889–1899. https://doi.org/10.1038/sj.emboj.7600194.

(176) 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.

(177) Kellersberger, K. A.; Yu, E.; Kruppa, G. H.; Young, M. M.; Fabris, D. Top-Down Characterization of Nucleic Acids Modified by Structural Probes Using High-Resolution Tandem Mass Spectrometry and Automated Data Interpretation. *Anal. Chem.* **2004**, *76* (9), 2438–2445. https://doi.org/10.1021/ac0355045.

(178) Kavan, D.; Man, P. MSTools—Web Based Application for Visualization and Presentation of HXMS Data. *Int. J. Mass Spectrom.* **2011**, *302* (1–3), 53–58. https://doi.org/10.1016/j.ijms.2010.07.030.

(179) 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.

(180) Tsai, K.-L.; Sun, Y.-J.; Huang, C.-Y.; Yang, J.-Y.; Hung, M.-C.; Hsiao, C.-D. Crystal Structure of the Human FOXO3a-DBD/DNA Complex Suggests the

Effects of Post-Translational Modification. *Nucleic Acids Res.* **2007**, *35* (20), 6984–6994. https://doi.org/10.1093/nar/gkm703.

(181) 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.

(182) 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.

(183) 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.

(184) Akashi, S.; Takio, K.; Matsui, H.; Tate, S.; Kainosho, M. Collision-Induced Dissociation Spectra Obtained by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry Using a 13 C, 15 N-Doubly Depleted Protein. *Anal. Chem.* **1998**, *70* (15), 3333–3336. https://doi.org/10.1021/ac980215f.