Peer-reviewed version available at Biomolecules 2019, 9, 535; doi:10.3390/biom9100535

Article

MS-based approaches enable the structural characterization of transcription factor / DNA response element complex

Lukáš Slavata[1, 2], Josef Chmelík[1, 2], Daniel Kavan[1, 2], Růžena Filandrová[1, 2], Jan Fiala[1, 2], Michal Rosůlek[1, 2], Hynek Mrázek[1], Zdeněk Kukačka[1], Karel Vališ[1], Petr Man[1, 2], Michael Miller[3], William McIntyre[3], Daniele Fabris[3], Petr Novák[1, 2, *]

[1] Institute of Microbiology, The Czech Academy of Sciences, Prague, 14220, Czech Republic
[2] Faculty of Science, Charles University, Prague, 12843, Czech Republic
[3] RNA Institute, University at Albany, State University of New York, Albany, NY 12222, U.S.A.

* Correspondence: pnovak@biomed.cas.cz; Tel.: +42-0325-873-610

Abstract: The limited information available on the structure of complexes involving transcription factors and cognate DNA response elements represents a major obstacle in the quest to understand their mechanism of action at the molecular level. We implemented a concerted structural proteomics approach, which combined hydrogen-deuterium exchange (HDX), quantitative protein-protein and protein-nucleic acid cross-linking (XL), and homology analysis, to model the structure of the complex between the full-length DNA binding domain (DBD) of FOXO4 and its DNA binding element (DBE). The results confirmed that FOXO4-DBD assumes the characteristic forkhead topology shared by these types of transcription factors, but its binding mode differs significantly from those of other members of the family. The results showed that the binding interaction stabilized regions that were rather flexible and disordered in the unbound form. Surprisingly, the conformational effects were not limited only to the interface between bound components but extended also to distal regions that may be essential to recruiting additional factors to the transcription machinery. In addition to providing valuable new insights into the binding mechanism, this project provided an excellent evaluation of the merits of structural proteomics approaches in the investigation of systems that arematerial not directly amenable to traditional high-resolution techniques.

Keywords: transcription factor, protein, DNA, protein-nucleic acid cross-linking, cross-linking, transplatin, trans-dichlorodiamineplatinum(II), hydrogen-deuterium exchange, FOXO4, molecular modeling

1. Introduction

The central dogma of biology rationalizes the flow of information leading from the sequence encoded by a gene to the biosynthesis of the corresponding protein. This flow must be finely regulated to coordinate the proper development and functioning of any living organism. A key role in the regulatory machinery is covered by transcription factors (TFs), proteins that recognize target DNA sequences called response elements and establish specific interactions with additional factors to activate or inhibit the transcription process [1–4]. The species involved in the process have been unambiguously identified [5,6], but significant information is still lacking on the effects of structure/dynamics on specific recognition and mechanism of action. The Protein Data Bank contains the high-resolution structures of at least 483 TFs from different species [7], which include less than 10% of all predicted human TFs [6]. Of such structures, only one third include also the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 July 2019 doi:10.20944/preprints201907.0233.v1

Peer-reviewed version available at Biomolecules 2019, 9, 535; doi:10.3390/biom9100535

cognate DNA response element and only one fifth are available in both bound and unbound states (**Figure 1**). Due to the size and complexity of such systems, the vast majority of solved structures do not cover the entire TF sequence, but consist almost exclusively of the DNA binding domain (DBD). This fact reflects the modular organization of TFs, which includes discrete domains acting in rather independent manner [8,9]. While DBDs tend to be highly structured, other regions responsible for either modulating transcription activity, or supporting facultative ligand interactions, are rather flexible and assume well-defined conformations only upon binding to the intended factor [10]. The unstructured nature of these regions poses many challenges to conventional high-resolution approaches, which require adequate stability and homogeneity. In most cases, the natural interactions established *in vivo*, which are responsible for stabilizing well-defined functional conformations, cannot be properly replicated *in vitro*. These challenges explain the chronic lack of comprehensive information on full-fledged TF structures, which still hampers the elucidation of their mechanism of action at the molecular level.

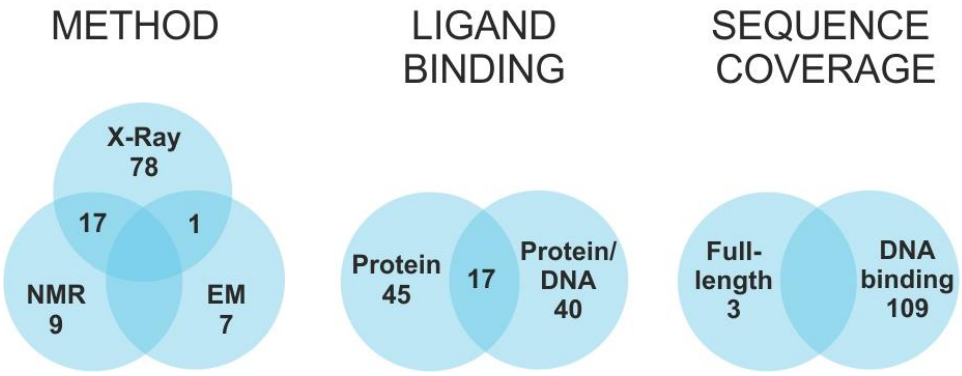

Figure 1: Transcription factors structures - current state: Statistics on high-resolution structures deposited in the Protein Data Bank (80), with consideration of selected methodology, presence of interaction partner, and sequence coverage. To date, the structure of only 112 human transcription factor (TF) have been solved out of a total predicted to be in the 1300 - 1900 range [6].

Powered by the development of new experimental strategies and mass spectrometric (MS) instrumentation, structural proteomics has rapidly become an essential approach for gathering valuable structural information for species that are not directly amenable to conventional high-resolution techniques [11–14]. In addition to hydrogen-deuterium exchange (HDX) [15], a combination of chemical and biochemical techniques broadly known as MS3D [16–18] have been effectively utilized to identify the regions of contact between bound biomolecules and reveal their mutual spatial organization. In HDX experiments, the exchange rate of backbone amide hydrogens, which is affected by solvent accessibility and possible involvement in hydrogen bonding, can be directly determined by MS analysis. This technique has been broadly employed to study conformational changes [19,20], protein folding [21,22] and protein-protein interactions [23,24], as well as nucleic acid-protein complexes [25–30]. Among the MS3D techniques, chemical and photo-activated cross-linking (XL) are employed to generate stable covalent bridges between contiguous functional groups, which can reveal their mutual placement in the targeted assembly [13,16]. A variety of bifunctional reagents with different spacing between reactive groups have been developed to determine the distance between susceptible residues. In this way, the sequence position of cross-linked residues and the length of the respective cross-linker provide valid constraints for building accurate molecular models through established computational methods [17,18,31,32]. The excellent versatility of these approaches has prompted the development of reagents capable of targeting functional groups present on protein as well as nucleic acid substrates [33–39]. Over time, capture tags and isotopic labels have been included in the cross-linker design to facilitate the isolation and analysis of cross-linked products[40,41]. Isotope-labelling, in particular, has provided a valuable tool for highlighting the presence of different conformational states and quantifying their partitioning on the basis of cross-linking probability [13].

In this study, we evaluated the concerted application of complementary structural proteomics techniques to overcome the challenges posed by full-fledged complexes between TFs and respective DNA response elements. We selected a model system consisting of the human transcription factor FOXO4 and its cognate Daf-16 family member-binding element (DBE) [42]. FOXO4 is part of the "O" subfamily of the forkhead box (FOX) class of transcription factors [42–44]. The DBD regions of this class are characterized by winged helix structures, which are formed by approximately 100 amino acids folded into helix-turn-helix motifs and β-sheet-bordered loops that make them resemble butterfly wings [45]. FOXO4 is involved in cell signaling pathways responsible for oxidative stress response [44,46], proliferation regulation [47,48], and angiogenesis [49]. Different mechanisms have been described to explain its activity and localization [50], which involve lysine acetylation [51–54], serine and threonine phosphorylation, as well as mono- [55] and poly-ubiquitination [56]. Enzymes responsible for FOXO4 phosphorylation are primary effectors in two important regulatory networks: the phosphatidylinositol-3-kinase [57–59] and Ras-like GTPase/c-Jun N-terminal kinase pathways [46]. The first high-resolution structure of FOXO4-DBD, which was obtained by NMR spectroscopy, confirmed the presence of a typical forkhead, winged helix fold [60]. At the same time, however, the report indicated that the N- and C- terminal regions of the DBD displayed chemical shifts consistent with highly flexible, disordered structures. The more recent identification of consensus sequences for the FOXO family [61] enabled the crystallization of a complex comprising a selected DBE duplex and a FOXO4-DBD construct that lacked the C-terminal region to facilitate crystal formation [62]. This high-resolution structure provided valuable details on the protein-DNA interaction, but revealed also numerous discrepancies with the binding modes exhibited by other members of the FOXO family [44], which were attributed to possible crystal-packing issues [62]. For these reasons, the FOXO4-DBD•DBE system offered an excellent opportunity for testing the ability of structural proteomics to probe the conformational effects of binding, which would help rectify or corroborate the observed discrepancies. On the other hand, it also afforded sufficient structural information to determine the validity of the new experimental constraints and evaluate the merits of the selected approaches.

The experimental strategies were selected for their ability to provide specific information on a typical protein-DNA complex. For instance, HDX was applied to recognize the regions of the protein affected by DNA binding, either through direct protection of the contact interface, or through allosteric conformational changes involving distal regions of the protein. Quantitative XL was used instead to identify possible variations between free and DNA-bound DBD structures, which would help elucidate the effects of the interaction on overall structure topology. In the case of the DNA component, the fast rate of back-exchange characteristic of nucleic acid hydrogens prevented the application of HDX to recognize the surface of the DBE duplex in direct contact with the DBD. As a possible alternative, we explored the application of transplatin (trans-dichlorodiamineplatinum(II), tPt) to generate protein-DNA cross-links that would help locate the mutual positions of interacting structural features [63]. The spatial constraints afforded by these determinations were combined to guide model-building operations and obtain a full-fledged structure for the complex. The results were compared to the available high-resolution structures to assess possible discrepancies and highlight the new information afforded by the selected techniques. It is necessary to point out here, that the FOXO4-DBD protein construct used in this study was identical to protein constructs used in previous high-resolution structural studies [60,62]. The outcome clearly demonstrated the benefits of structural proteomics to tackle the elucidation of structure and dynamics in systems that elude established high-resolution techniques.

2. Materials and Methods

2.1. Materials

Non-labelled and isotope-labelled cross-linkers di(N-succinimidyl) glutarate (DSGd0/DSGd4) and di(N-succinimidyl) suberate (DSSd0/DSSd4) were purchased form ProteoChem (USA). Modified protease trypsin (Gold, mass spectrometry grade) was purchased from Promega (USA).

Nuclease Bal-31 was obtained from New England BioLabs (USA). Liquid chromatography solvents of LC/MS grade were purchased from Thermo Fisher Scientific (USA). Other chemicals (highest available purity) were obtained from Sigma-Aldrich. All other chemicals, solvents and buffers for SDS-PAGE were obtained from Bio-Rad Laboratories, Inc. (USA). The pET-15b plasmid carrying His-tag, thrombin cleavage site, and FOXO4-DBD (Uniprot ID: P98177-1; residues 86 - 211) sequences was obtained from Prof. Obsil.

2.2. Sample preparation

Full-length DBD (residues 82 – 207 of the entire FOXO4 sequence) was expressed with an N-terminal His-tag from an appropriate pET-15b plasmid, and then affinity captured on a TALON Superflow Resin (Clontech Laboratories, USA) charged with Co^{2+}. The captured protein was submitted to thrombin digestion to eliminate the tag, followed by gel permeation chromatography. A more detailed description of all experimental procedures is included in the ***Appendix (Methods in detail)*** section. The identity, integrity, and purity of the final sample were verified by MS analysis (*vide infra*). A duplex DNA construct containing one of the DBE consensus sequences (i.e., TTG TTT AC) [42,64] was obtained by annealing complementary oligonucleotides (i.e., 5′-TTG GGT AAA CAA G-3′ forward and 5′-CTT GTT TAC CCA A-3′ reverse). Equimolar amounts were mixed and then heated to 95°C for 1 min. Finally, the sample was let cool to room temperature to form the 13 bp duplex DNA. The desired FOXO4-DBD•DBE complex was obtained by mixing equimolar amounts of protein and DNA samples in 10mM HEPES buffer with 50mM NaCl (pH adjusted to 7.4) to a final 33.7μM concentration, and incubating the mixture at 18°C for 1 hr.

2.3. Product characterization

Initial stocks of FOXO4-DBD, DBE construct, and FOXO4-DBD•DBE were analyzed to assess sample purity after expression/purification, annealing of the duplex structure, and proper complex formation. The products of XL reaction were also analyzed in the same fashion to assess the incidence of modification. Briefly, each stock was diluted to a final 10-μM concentration by adding a 7.5 mM solution of ammonium acetate (AA) with 50% MeOH (pH 6.85), and then analyzed on a Bruker Daltonics (Billerica, MA) 15T-Solarix XR Fourier transform ion cyclotron resonance (FTICR) mass spectrometer. FOXO4-DBD, as well as FOXO4-DBD•DBE, were analyzed in positive ion mode. Each sample was loaded onto a syringe and introduced into the electrospray ionization (ESI) source at a 2 μL/min flow rate. The FTICR analyzer was calibrated by using a solution of sodium trifluoroacetate (NaTFA), which afforded a typical 1 ppm accuracy. Mass spectra were acquired over a 250 – 4000 m/z range for 3 min.

2.4. Hydrogen-deuterium exchange

HDX reactions were performed on 20μM solutions of either FOXO4-DBD or FOXO4-DBD•DBE complex prepared in an H_2O-based buffer (pH 7.4) containing 10mM HEPES and 50mM NaCl. After pre-incubation for an hour at 20 °C, the exchange was initiated by diluting each sample 10-fold into a D_2O-based buffer (pD 7.4) containing 10mM HEPES and 50mM NaCl. The reaction was allowed to proceed at 20 °C, while small aliquots containing 100pmol of protein were taken at predetermined intervals (i.e., 0.33, 2, 5, 10, 30, 60, 180, and 300 min). Quenching was achieved by immediately mixing the aliquot with a 1M glycine/HCl buffer with pH 2.35, and then rapidly freezing the solution in liquid nitrogen. Samples were stored at -80°C. The analysis was performed by using columns with different immobilized proteases followed by liquid chromatography-mass spectrometry (LC-MS) according to ref. [65]. All experiments were performed as triplicate. A complete description of these procedures is included in the ***Appendix (Methods in detail)*** section.

2.5. Quantitative protein-protein cross-linking

Samples containing 20μM of either FOXO4-DBD or FOXO4-DBD•DBE complex prepared in 10mM HEPES buffer (pH 7.4) with 50mM NaCl were pre-incubated for an hour at 20°C before

Peer-reviewed version available at Biomolecules 2019, 9, 535; doi:10.3390/biom9100535

introducing the cross-linking reagent. Separate samples of FOXO4-DBD were treated with either DSGd0 or DSSd0 in their regular, non-labelled form, whereas FOXO4-DBD•DBE samples were reacted with the deuterium-labelled DSGd4 and DSSd4 versions. The reagents were dissolved in dimethyl-sulfoxide (DMSO) to 6.74 mM concentrations and then added to each substrate to achieve a 10:1 molar ratio. The cross-linking reaction was allowed to proceed undisturbed for 2 hrs, after which corresponding regular and deuterium-labelled samples (e.g., treated with DSSd0 and DSSd4) were mixed in a 1:1 ratio to enable quantification. In parallel, control samples were also examined, which were treated with pure DMSO lacking cross-linker, or matching cross-linker mixtures with a 1:1 ratio of either DSGd0/DSGd4 or DSSd0/DSSd4. All reactions solutions were analyzed by SDS-PAGE to check whether cross-linking had induced any unwanted formation of non-specific higher-order aggregates, or had prevented a sufficient degree of digestion necessary to enable subsequent analysis. Characterization of cross-linked conjugates was achieved according to a bottom up approach that employed trypsin digestion followed by LC-MS analysis [13]. All experiments were performed as triplicate. A complete description is included in the ***Appendix (Methods in detail)*** section.

2.6. Protein-DNA cross-linking

Samples containing 25μM of either FOXO4-DBD or FOXO4-DBD•DBE complex in 150 mM ammonium acetate (pH 6.85) were treated with a 1mM solution of trans-platinum(II)diammine dichloride (transplatin, tPt), which had been pre-incubated for an hour at 18°C. Reaction mixtures containing a final 200μM concentration of transplatin and 20μM of protein/complex were incubated at 18°C for 14 hr. Control samples devoid of transplatin were prepared at the same time in the same manner. Reaction and control samples were analyzed by native and denaturing DNA polyacrylamide gel electrophoresis, SDS polyacrylamide gel electrophoresis, and mass spectrometry. Characterization of cross-linked conjugates was achieved according to a bottom up approach that involved treatment with Bal-31 nuclease and trypsin to digest DNA and protein components, respectively, followed by LC-MS/MS analysis with data-independent acquisition with broad isolation window. A complete description is included in the ***Appendix (Methods in detail)*** section.

2.7. Data processing and interpretation

The SNAP 2.0 algorithm of the DataAnalysis 4.2 (Bruker Daltonics) software package was utilized to generate deconvoluted spectra and lists of monoisotopic masses from the acquired MS data. MASCOT 2.2 search engine was used to search MS/MS data and achieve the identification of on-line digest products from a theoretical library of digestion products. Deuteration rate was determined by using the home-built Deutex software (unpublished). The home-built LinX software (available online) and Stavrox (v. 3.6.0.1 by Michael Götze) software was then used to compare the experimental data with a library of theoretical cross-linking products to correctly identify the sought-after conjugates. The proportion of labelled versus un-labelled species was determined by applying mMass 5.4.1 [66] to the signals of such conjugates. In the case of peptide-DNA conjugates, deconvoluted spectra and monoisotopic masses were calculated by using a constant unit to mimic the presence of a certain oligonucleotide cross-linked by a transplatin equivalent and searched by LinX. A complete description of these procedures is included in the ***Appendix (Methods in detail)*** section.

2.8. Molecular modeling

The program Modeller [67] was employed to generate models of FOXO4-DBD and FOXO4-DBD•DBE according to the spatial constraints afforded by the HDX and cross-linking experiments. Initially, six different structures available in the Protein Data Bank were utilized as possible templates to guide homology modeling, which displayed different levels of sequence identity (s. i.) with our target. For instance, the 1E17 (100% s. i.), 2K86 (83% s. i.), and 2KIU (51% s. i.)

structures contain only the apo-protein with no bound DNA, whereas 3L2C (100% s. i.), 2UZK (83% s. i.), and 2A07 (48% s. i.) include also the latter. These structures, however, covered only the 101-176 residues of the DBD sequence, whereas our intended target covered the 82-207 section to include the additional flanking sequences that had eluded structural elucidation. The missing regions spanning G_{74}-Q_{100} at the N-terminus and N_{177}-A_{207} at the C-terminus of the DBD sequence were generated directly in Modeller by using distance restraints derived from our XL data. The online server *make-na* (http://structure.usc.edu/make-na/) was employed to generate the initial structure of DBE duplex in an ideal B-DNA conformation.

The types of cross-linking restraints informing these operations included C^{α} - C^{α} distances that were set to 20.5 ± 3.0 Å for DSG and 24.2 ± 3.0 Å for DSS, respectively. The maximal LYS C^{α}- LYS C^{α} cross-link distance was calculated as a sum of the spacer arm length [68] and the distances between C^{α} and N^{ζ} atoms (for both LYS involved in the cross-link) where all χ side-chain torsion angles are set to the trans conformation. The same approach was used for the calculation of the maximal C^{α}-C^{α} cross-link distances for other amino acids. For cross-links involving the N-terminal residue, the N_{term} - C^{α} distance was set to 14.1 ± 3.0 Å for DSG and 17.8 ± 3.0 Å for DSS, respectively. A distance of 5.5 ± 2.5 Å [69,70] was assigned between atoms bridged by the transplatin reagent. For all distance restraints, a 3-Å standard deviation was used to account for the intrinsic flexibility of linker spacers and side chains involved in the conjugate. The initial structures were submitted to limited molecular dynamics (MD) simulations in the Modeler package to eliminate any possible angle strains and steric clashes. Subsequently, this process a simulated annealing protocol in the torsion angle space was accomplished in CNS [71]. An ensemble of 50 structures was calculated for each starting model. During simulations, the coordinates afforded by the initial PDB templates and DBE structure, as well as the XL distances, were kept fixed. The resulting models were visualized by using Pymol [72].

The HADDOCK [73] program was utilized to perform docking experiments between FOXO4-DBD and DBE substrate. The process designated the D_{139}-L_{154} stretch as active residues on the basis of the results of HDX experiments. The THY_{16}-ADE_{21} region was designated as active because of its high degree of conservation in the observed consensus sequences. The passive residues were automatically defined around the active ones. The WeNMR/WestLife infrastructure [74] was used to carry out the computationally intensive docking calculations. The first structure of the cluster of the FOXO4-DBD•DBE complex, which displayed the best HADDOCK score for each run, was used for the subsequent modeling operations.

2.9. Data availability and software

- MSTools package – available at http://peterslab.org/MSTools
- LinX - available at http://peterslab.org/MSTools
- Stavrox (v. 3.6.0.1 by Michael Götze) – available at http://www.stavrox.com/
- DeutEx - In house developed program DeutEx is based on a Tcl macro. It requires protein sequence, list of identified peptides from search engines such as MASCOT or PEAKS. Basic overview of the workflow shown on unrelated example data can be found here - http://peterslab.org/downloads/SW/DeutEx.mp4
- Mass spectrometry data available at https://www.ebi.ac.uk/pride/archive/ (Project accession: PXD013969)

3. Results and Discussion

The crystal structure available for the FOXO4-DBD•DBE complex does not cover the entire sequence of the DNA binding domain [62], which spans the 82 - 207 section of FOXO4, but omits instead flanking regions hypothesized to promote the recruiting of additional components of the transcription machinery. At the same time, the NMR structure of full-length FOXO4-DBD provides limited information on the G_{138} - A_{144}, E_{166} - K_{170}, and the N- and C-terminal regions, which were described as rather flexible and disordered in solution [60]. Although DNA binding has been credited for stabilizing at least some of these regions, the structure of the bound form still displayed significant discrepancies with those of homologous members of the FOXO family [44], which were

possibly caused by crystal packing [62]. For this reason, we investigated such discrepancies by implementing biochemical approaches to probe the effects of ligand binding directly in solution. The study employed recombinant full-length DBD (i.e., residues G_{82} - A_{207}) and a duplex DNA construct containing the 5'-TAC CCA A-3' consensus sequence, which was obtained by annealing commercial oligo-deoxyribonucleotides. As shown in **Figure S-1A** and **Figure S-2** of ***Supporting Information***, mixing equimolar amounts of protein and duplex DNA provided the expected 1:1 species corresponding to the desired FOXO4-DBD•DBE complex. These samples were submitted to hydrogen-deuterium exchange, quantitative protein-protein cross-linking, protein-DNA cross-linking, and docking experiments to obtain complementary information on their mutual interactions and spatial arrangement. The results were compared to those obtained from the individual FOXO4-DBD protein to investigate the effects induced by specific DNA binding.

3.1. Combined online digestion by pepsin and nepenthesin I improves the HDX resolution

We initially pursued the identification of the regions of FOXO4-DBD, which were making direct contact with the DBE duplex, or were subjected to detectable microenvironment variations upon binding. Hydrogen-deuterium exchange (HDX) was performed on both free and bound forms of full-length FOXO4-DBD. The determinations followed a well-established protocol in which the exchange process was stopped at predetermined intervals to monitor the rate of exchange. Since, hydrogen-deuterium exchange has not yet been used to study the complex of protein and duplex DNA, we first tuned conditions for on-line digestion in order to obtain the best spatial resolution. It was achieved using the combination of nepenthesin I and pepsin where rather small and overlapping peptides were observed. The robustness of the setup was approved by multiple injections. The final analysis were carried out at low pH and temperature to minimize back-exchange and included protein digestion in consecutive on-line columns containing immobilized proteases (i.e., pepsin and nepenthesin-1), followed by LC-MS and LC-MS/MS analysis of digested peptides (see ***Materials and Methods*** and ***Appendix (Methods in detail)*** section) [75]. The ensuing peptide map demonstrated that the procedure afforded full coverage of the FOXO4-DBD sequence (**Figure 2**).

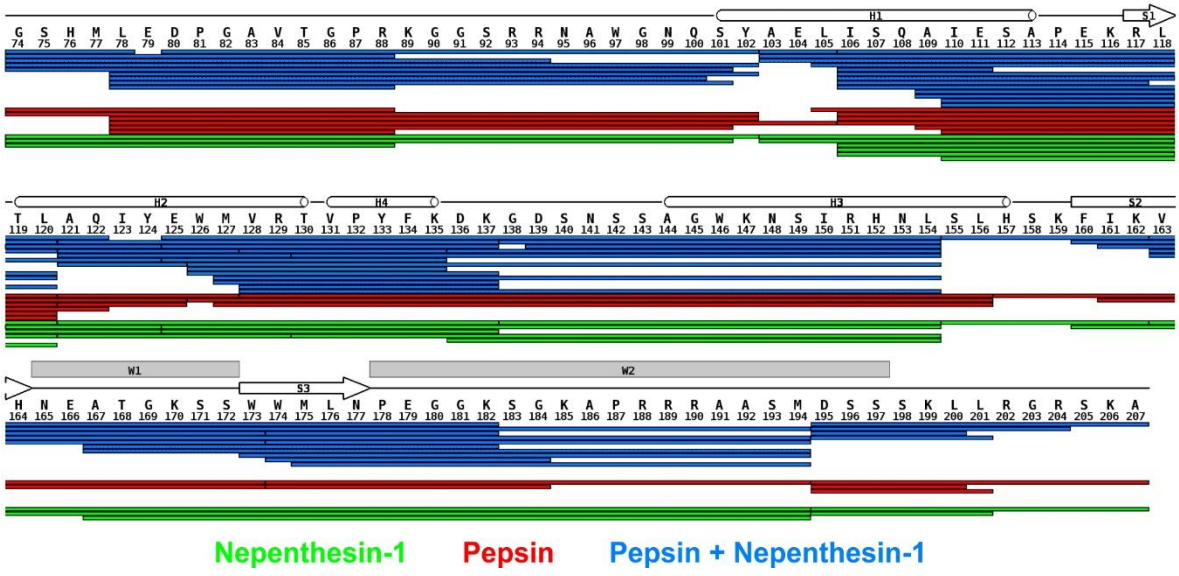

Figure 2: Combined online digestion by pepsin and nepenthesin I improve the HDX resolution. Comparison of hydrolytic products obtained by on-line digestion in columns derivatized with nepentesin-1, pepsin, or pepsin/nepenthesin-1 combination. The peptides are mapped along sequence and topological information. The peptides afforded 100% coverage of the protein sequence. Note that the G_{74} - P_{87} region of the construct was contributed by the recombinant-production vector and, thus, was not part of the wild-type FOXO4 sequence (see also **Figure S-33** of ***Supporting Information***).

316

317 *3.2. HDX identified the interaction interface and long-distance stabilization of protein structure*

318 For each digestion product, a relative deuteration rate was calculated by considering the
319 number of hydrogens exchanged with deuterium atoms against the total number of exchangeable
320 amide hydrogens in the peptide. The relative deuteration rates versus exchange time were
321 calculated at both the peptide and amino acid levels to recognize possible variations between the
322 free and bound FOXO4-DBD (see **Figure S-3** to **S-6** of ***Supporting Information***, respectively). This
323 task was facilitated by calculating actual differences for each amino acid in the sequence, which were
324 visualized onto a 3D model of FOXO4-DBD•DBE by using an appropriate color palette (**Figure S-7**
325 of ***Supporting Information***).

326 The results were summarized in HDX difference plot that provided a comprehensive view of
327 the variations of solvent accessibility induced by the specific interactions between FOXO4-DBD and
328 its cognate DBE duplex (**Figure 3**). Starting from the N-terminus, the G_{74} - Y_{102} region displayed
329 relatively high levels of deuteration regardless of reaction time, with no significant differences
330 between free and bound form. These observations indicated that this region was rather exposed and
331 capable of exchanging freely with the solvent in both forms. In contrast, the next section spanning
332 the A_{103} - T_{130} residues displayed the most extensive differences in deuteration rates, which increased
333 significantly as a function of time. This sequence folds helix H1 and H2, strand S1, and intervening
334 loops (see topology annotation in **Figure 2**). According to the crystal structure, none of these
335 distinctive features is supposed to make direct contact with the duplex DNA [62], which would help
336 explain the drop of deuteration by invoking a simple protection effect. In the absence of direct
337 contact, the observed loss of solvent accessibility must be attributed to indirect conformational
338 effects induced by binding. The fact that the difference in deuteration levels increased gradually
339 with time and stabilized after 60 min suggests that, in the free form, this set of secondary structures
340 may undergo slow mutual dynamics that delay the exchange of susceptible hydrogens. In the bound
341 form, such dynamics may be stabilized by interactions with contiguous structures that, in turn, make
342 direct contact with the DNA ligand. Like falling dominoes, a series of relatively minor
343 conformational variations linked together may ultimately induce observable inhibition of the
344 exchange reaction. This long-distance effect is clearly evident, for example, in the relative
345 deuteration plot of peptide A_{103} - L_{118} (2-3), which shows increasing uptake in the free FOXO4-DBD
346 as a function of time, but constant low-level deuteration in the bound form (**Figure S-3** of ***Supporting***
347 ***Information***).

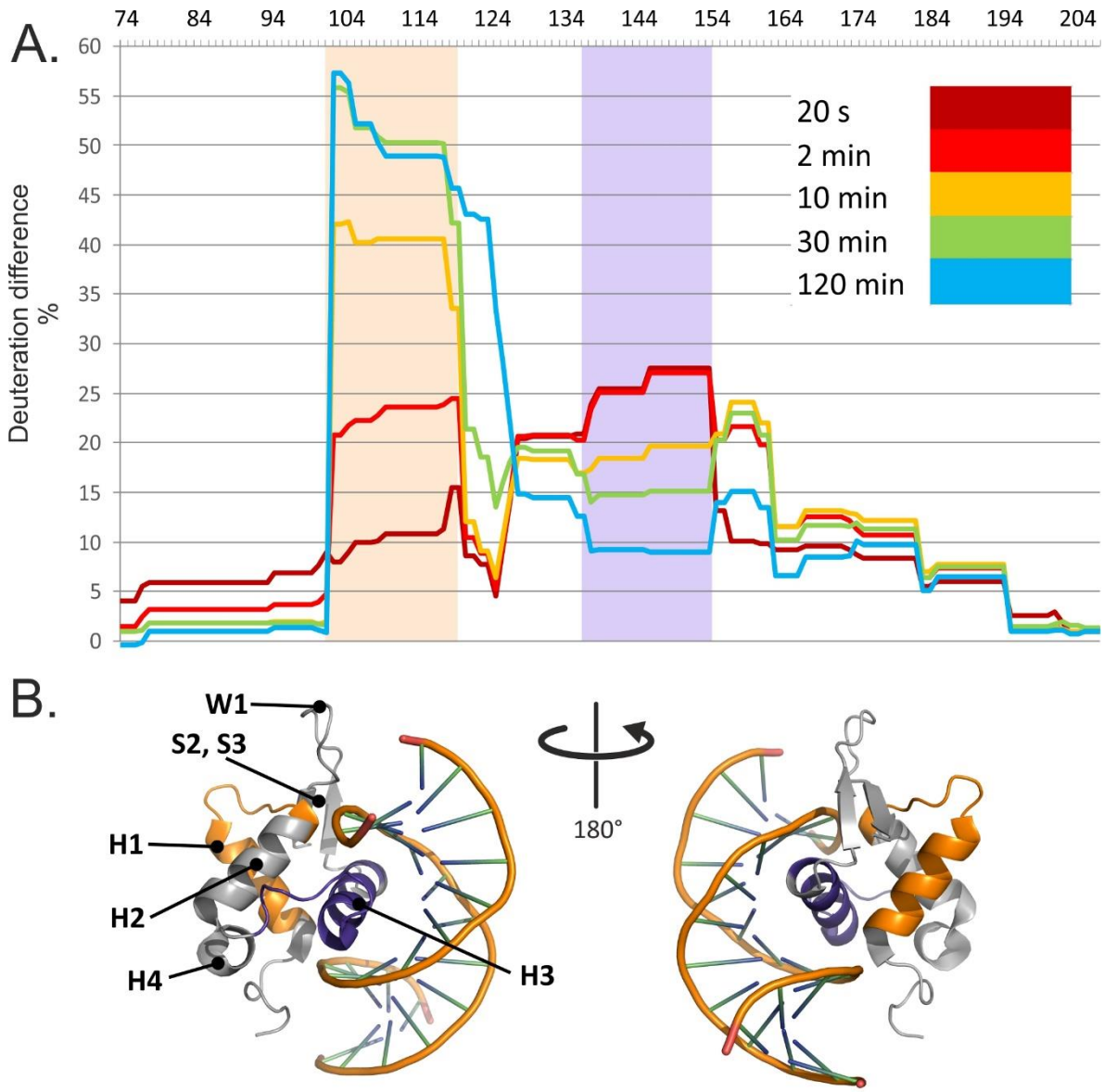

Figure 3: HDX identified the interaction interface and long-distance stabilization of protein structure. **A.** Relative deuteration differences [DR(FOXO4-DBD) – DR(FOXO4-DBD•DBE)] plotted along the FOXO4-DBD sequence and their evolution in time. **B.** FOXO4-DBD•DBE structure with highlighted regions showing significant differences in deuteration levels. Note that the G_{74} - P_{87} region of the construct was contributed by the recombinant-production vector and, thus, was not part of the wild-type FOXO4 sequence (see also **Figure S-3** to **Figure S-7** of ***Supporting Information***).

The next region spanning the V_{131} - K_{159} residues also manifested significant differences between free and bound forms, but their time-dependence displayed a rapid increase at shorter intervals, followed by a decline near initial levels at longer reaction times (**Figure 3A**). In particular, the residues forming helix H3, which the crystal structure places directly in the major groove of the DBE duplex [62], experienced the largest differences. For this reason, direct steric protection induced by bound DNA could explain the uptake inhibition observed for such residues. In contrast, the outcome observed for the contiguous helix H4 and intervening loop could indirectly result from the stabilization of the H3 conformation, which could constrain the placement of such residues and restrict their solvent accessibility. The crystal structure identified also a handful of contacts that were mediated by water molecules trapped in the binding interface. While these interactions have been

suggested to further stabilize the dynamics of helix H3 and flanking regions, it is not clear how trapping D_2O versus H_2O present in the solvent may affect the observed deuteration rates. It should be also pointed out that FOXO4 differs from other FOXO homologues by the insertion of five amino acids (K_{137} - N_{141}) in the H4-H3 loop (**Figure 2**). It is not clear whether this insertion may be responsible for the unusual conformation assumed by helix H3 in the bound form, which differs from those assumed in other members of the family [62]. However, the observed inhibition pattern supports a role in stabilizing the fold of FOXO4-DBD and reducing its overall flexibility upon complex formation.

The F_{160} - D_{195} section corresponds to sequences located on the C-terminal side of helix H3, which manifested smaller but still perceptible differences of deuteration rates. This region contains strands S2 and S3, as well as the W1 and W2 wings, which the crystal structure placed far removed from the DNA binding interface. Also, in this case, an overall decrease of structural flexibility upon binding could explain the reduced deuterium uptake. It should be noted that, in addition to conferring FOXO4-DBD its winged look and fine tuning of the interaction with DNA [61], W1 and W2 could contribute to constitute possible regions of contact for auxiliary components of the transcription complex. In this context, the DNA binding could represent a possible mechanism for modulating the recruitment of such components. The final section covered by HDX determinations consisted of the S_{196} - A_{207} sequence and displayed marginal deuteration differences. This observation indicated the absence of any protection or conformational effects induced by DNA binding.

3.3. Protein-DNA cross-linking revealed the mutual placement of protein and DNA components

The HDX experiments identified the regions affected directly and indirectly by DNA binding, which experienced clearly detectable variations of solvent accessibility. However, these types of determinations could not identify the structures responsible for limiting the access of solvent to a specific region. In other words, these experiments could not reveal the mutual spatial relationships between such structures, nor assess the effects of binding on such relationships. For this reason, we employed different types of cross-linking strategies to probe the organization of the various structures and recognize their mutual placement in the overall fold. The first approach employed transplatin to generate putative protein-DNA conjugates that may be capable of constraining the position of the DBE ligand onto the FOXO4-DBD substrate. The reactivity of platinum compounds towards specific functional groups of nucleic acids is well documented [76] and involves the preferential attack of the N7 position of guanine base [77]. Although the characteristics of their reactivity towards protein residues are still unclear [63], amino acids with electron-rich S, N and O atoms, such as Cys, Met, His, and Thr, have been described as preferred targets [78]. We treated samples of FOXO4-DBD•DBE complex, as well as free FOXO4-DBD and DBD, with a 10:1 transplatin to substrate molar ratio in 150 mM ammonium acetate (pH 6.85) and incubated at 18 °C for 14 hr (see ***Materials and Methods*** and ***Appendix (Methods in detail)*** section). The mixtures were analyzed by both ESI-MS and SDS-PAGE to assess the distribution of transplatin adducts and estimate the proportion of sought-after intermolecular cross-links. The representative data in **Figure 4A** provides a view of the typical product distributions obtained from these probing reactions, which included monofunctional "dangling" adducts displaying a still unreacted chloride function (i.e., marked as tPt-Cl adducts), as well as bifunctional conjugates in which both functions had effectively reacted (i.e., marked as tPt adducts).

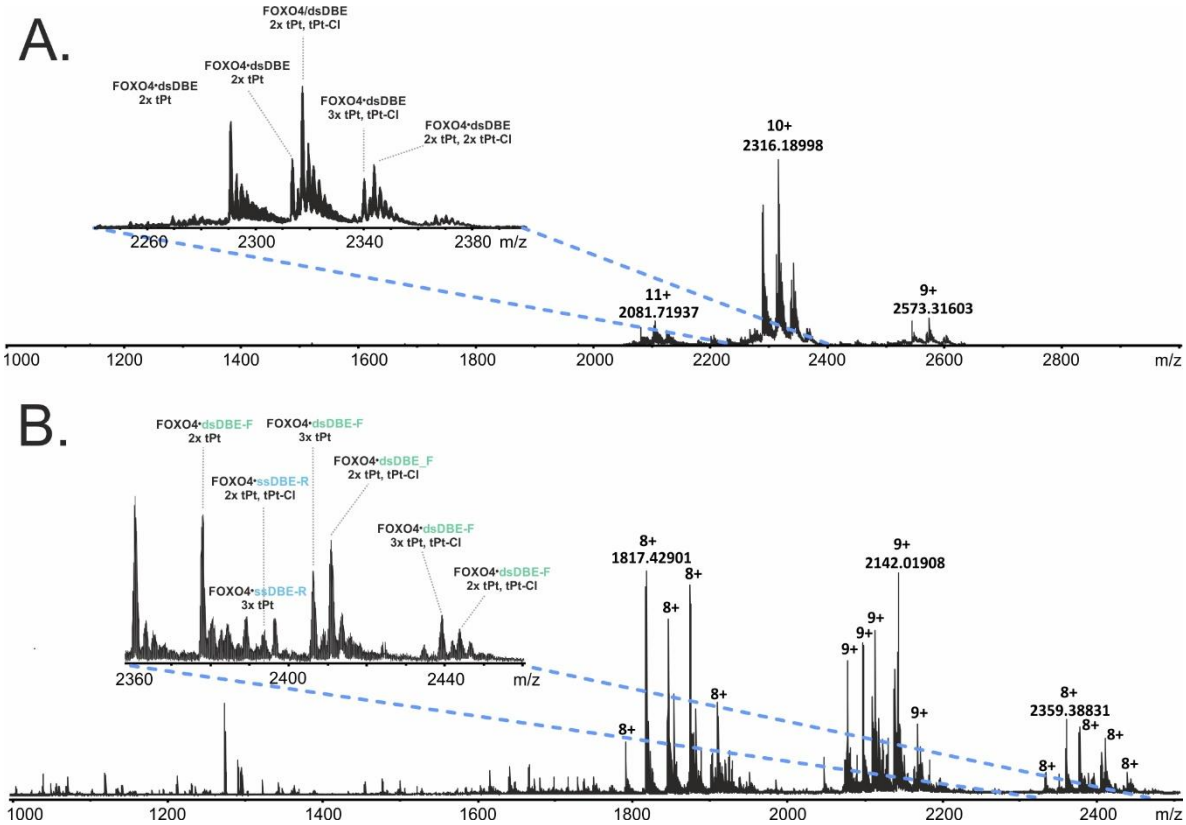

Figure 4: Transplatination generates protein-DNA cross-links. **A.** Mass spectrum of FOXO4-DBD•DBE treated with transplatin and sprayed in 1% acetic acid in 7.5 mM ammonium acetate solution. **B.** Mass spectrum of FOXO4-DBD•DBE treated with transplatin and sprayed in 1% acetic acid and 50% methanol in 7.5 mM ammonium acetate solution. Abbreviations: tPt – bifunctionally bound transplatin cross-linker ($Pt(NH_2)_2$); tPt-Cl – monofunctionally bound transplatin cross-linker ($Pt(NH_2)_2Cl$); dsDBD – duplex DNA; ssDBD-F – forward oligonucleotide strand; ssDBD-R – reverse oligonucleotide strand. Species denoted here as FOXO4 correspond to the form of FOXO4-DBD shortened by three C-terminal amino acids (see also **Figure S-1**, **Figure S-2**, **Figure S-8**, **Figure S-9** of ***Supporting Information***.

Based on mass alone, it is not typically possible to distinguish desired intermolecular conjugates from intramolecular crosslinks, which share the same elemental composition. For this reason, ESI-MS analysis was repeated in the presence of 50% methanol to achieve mild denaturing conditions. In this way, products stabilized by bridging bifunctional cross-links were still detected intact by virtue of their covalent nature, such as the conjugates containing FOXO4-DBD and individual DBE-F or DBE-R strands produced by dissociation of the DBE duplex. In contrast, no signal was observed for complexes devoid of any intermolecular conjugation, whereas adducts of their free unbound components were individually detected, such as those of FOXO4-DBD protein, DBE-F, and DBE-R strand (**Figure 4B**). In analogous fashion, the reaction mixtures were also analyzed by PAGE under both native and denaturing conditions (**Figure S-1** of ***Supporting Information***). Direct data comparison enabled the identification of bands that eluded dissociation by elevated concentrations of SDS or urea and, thus, could be attributed to the presence of bridging bifunctional crosslinks. These data enabled us to estimate that the desired intermolecular cross-links amounted to less than 5% of typical transplatin reactions mixtures.

A classic bottom-up strategy was carried out to complete the characterization of cross-linked products, which included digesting the material with protein- and nucleic acid-specific enzymes to obtain samples amenable to LC-MS and LC-MS/MS analysis. In particular, reaction mixtures were treated with trypsin to map the position of peptides conjugated to DNA strands (see ***Materials and***

Peer-reviewed version available at Biomolecules 2019, 9, 535; doi:10.3390/biom9100535

Methods and ***Appendix (Methods in detail)*** section). In subsequent experiments, the size of the oligonucleotide moieties was reduced by treatment with Bal-31 nuclease to facilitate analysis. The representative data in (**Figure S-8** of ***Supporting Information***) illustrates the challenges faced by the MS/MS analysis of these types of hetero-conjugates. Upon gas-phase activation, a precursor ion consisting of G_{74} - R_{88} cross-linked to the DBE_R strand underwent dissociation around the bridging Pt atom, rather than along the backbones of the bridged moieties. The absence of sequence information afforded by this type of fragmentation prevented the identification of the actual residues involved in the cross-linking reaction. Nevertheless, the identity of the conjugated components still represented valuable information on the mutual relationships between contiguous regions (summarized in **Figure S-9** of ***Supporting Information***).

The detected peptide-oligonucleotide conjugates were examined in the context of the results afforded by the HDX determinations and other structural information available for the system. For instance, peptide N_{148} - K_{159} spanning helix H3 was found conjugated to the forward strand of DBE, consistent with the placement of H3 directly into the major groove of the DNA duplex in the crystal structure [62]. This finding agreed also with the prominent protection effects observed in this region during HDX experiments (**Figure 3A**). Residues S_{149} and H_{152}, in particular, should represent excellent conjugation sites by virtue of their susceptibility to transplatin reaction and favorable orientation facing the duplex's major groove. Peptide F_{160} - K_{170} covering the end of S2 and beginning of W1 was also cross-linked to the forward strand of DBE, in spite of the absence of any direct contact in the crystal structure. In this peptide, reactivity and orientation considerations would point towards H164 and T168 as possible conjugation sites, if their distances from susceptible DNA structures were sufficiently favorable. In this direction, the HDX data indicated that this region experienced a detectable decline in deuterium uptake consistent with the adoption of a rather constrained conformation upon DNA binding (**Figure 3A**). The new conformation could place susceptible groups within mutual striking distance, thus promoting the formation of the observed cross-linked product. A similar explanation is applicable also to the S_{171} - K_{182} peptide spanning the end of W1, beginning of W2, and intervening S3 strand, which formed cross-links with both forward and reverse strands of DBE. Also, this region experienced a significant decrease of deuterium exchange upon binding, which was not explainable by direct steric protection, but rather by indirect conformational effects transmitted through contiguous structures. The crystal structure orients S_{171} and S_{172} to face the minor groove of the duplex construct, which would represent prime positions for promoting conjugation with either strand. Also, in this case, the respective functional groups could be placed within striking distance by the more constrained conformation revealed by HDX experiments. The remaining products consisted of the G_{74} - R_{88} peptide cross-linked to either the forward or reverse strand. These products are a testament to the flexibility of the N-terminal loop, which was supported by the lack of any significant variation of deuteration patterns reported by the HDX experiments.

3.4. DNA binding induced significant effects on protein conformation

The observed protein-DNA cross-links provided valuable information not only on the reciprocal positions of protein and DNA components, but also on the significant changes induced by binding on the initial protein conformation. We employed protein-specific reagents to evaluate the extent of such variations and enable a better appreciation of indirect conformational effects. Our quantitative crosslinking approach involved the concerted application of the homobifunctional reagents DSGd0/d4 and DSSd0/d4 [di(N-succinimidyl) suberate and di(N-succinimidyl) glutarate] to bridge susceptible amino or hydroxy groups that may be placed within 20.5 ± 3.0 or 24.2 ± 3.0 Å, respectively, from one another (see ***Appendix (Methods in detail)*** section). The utilization of reagents with different bridging spans offered the ability to determine an average distance between residues. At the same time, the isotopic labels facilitated the identification of cross-linked products in complex digestion mixtures from their characteristic 4-Da spacing and enabled the acquisition of unbiased quantitative data on the incidence of cross-linking in the free or bound FOXO4-DBD. Initially, separate samples were treated with 1:1 mixtures of matching unlabelled/labelled reagents

of same length to complete a survey of the regions susceptible to cross-linking (see ***Materials and Methods*** and ***Supporting Information***). A total of 39 conjugates were identified, which bridged lysine and serine residues, as well as the N-terminal amino group (**Table S-1** and **Figure 10** to **Figure 26** of ***Supporting Information***). The majority of them were detected in matching pairs generated by reagents of either length, and were observed in both free and bound samples. However, a small portion was unique for just one form and/or cross-linker length. Next, individual aliquots of free FOXO4-DBD were treated with either DSGd0 or DSSd0, whereas those of FOXO4-DBD•DBE complex were separately treated with either DSGd4 or DSSd4 (see ***Materials and Methods*** and ***Appendix (Methods in detail)*** section). Corresponding samples treated with the same unlabelled/labelled reagent were mixed in a 1:1 molar ratio prior to protease digestion and analysis to compare the incidence of each conjugate in either free or bound samples. The proportion of each of the 39 conjugates identified earlier was determined for at least one of the cross-linker lengths, as summarized in (**Table S-2** of ***Supporting Information***).

A close examination of the results revealed distinctive cross-linking patterns associated with the presence of bound DNA. In particular, the incidence of some conjugates decreased significantly upon binding, while others increased. Among the former, the DGS conjugates bridging K147 to either K170 or N-term dropped from 95.9% and 92.8% to 4.1% and 7.2%, respectively (**Table S-2** of ***Supporting Information***). The cross-linking inhibition manifested by K_{147} cannot be merely ascribed to its location on helix H3, in direct contact with the duplex construct, because this residue was still capable of supporting conjugation with both K_{135} and K_{137}. A more plausible explanation could be that DNA binding forced K_{147} out of the reach of either K_{170} on helix H2, or K_{162} on the adjacent S2 region. The limited nature of such conformational changes was revealed by the fact that K_{147} was pushed out of N-term's reach for the shorter DSG reagent, but was still sufficiently close for the longer DSS, with incidence of cross-linking dropping from 92.8 to 7.2% and 72.7 to 27.3%, respectively (see **Table S-2** of ***Supporting Information***). The limited extent of these changes was also evident in the subtler cross-linking variations between K_{147} and either K_{135} or K_{137} located on the H3-H4 intervening loop.

In other cases, DNA binding increased the incidence of specific conjugates by placing residues within mutual striking distance in the complex, which were marginally susceptible or inert in the free protein. For example, the conjugates bridging residue K_{182} with N-term, K_{89}, K_{116}, K_{159}, K_{162}, or K_{182} were greatly enhanced by the presence of DNA duplex (**Table S-2** of ***Supporting Information***). For the majority of these positions, the levels of cross-linking observed with the longer DSS reagent displayed more significant variations than those with the shorter DSG. Considering that K_{182} is located on the W2 wing region, these observations offered further evidence of the long-range conformational effects of DNA binding suggested by the results of HDX and protein-DNA cross-linking experiments. Another example was provided by the numerous conjugates involving the N-term, which suggested that DNA binding had prominent stabilizing effects on a region that was rather flexible in free FOXO4-DBD. Consistent with the HDX data, the variations of cross-linking patterns confirmed that DNA binding induced significant effects on protein conformation not only within the contact interface, but also in rather distal positions.

3.5. Structural proteomics could effectively guide model-building operations to produce very high-quality 3D models

The HDX and XL experiments provided a wealth of information that was used to guide the molecular modelling of full-fledged FOXO4-DBD and FOXO4-DBD•DBE complex. Our approach took advantage of available high-resolution structures that, although incomplete in their coverage of the protein sequence, still represented excellent templates for homology modelling operations (see ***Materials and Methods*** and ***Appendix (Methods in detail)*** section). In particular, the templates were used to obtain the coordinates of what could be defined as the structured core of the complex, a region spanning approximately from R_{93} to N_{177}, which displayed limited discrepancies across the available structures. In contrast, the regions, which were either absent from the templates, or displayed significant variations, or had been predicted to possess a high degree of flexibility by

Peer-reviewed version available at Biomolecules 2019, 9, 535; doi:10.3390/biom9100535

PSIPRED [79], were modeled according to the HDX and cross-linking information (see ***Materials and Methods*** and ***Appendix (Methods in detail)*** section). These regions corresponded to the G_{74} - Q_{100} and N_{177} - A_{207} sections located respectively at the N- and C-terminal ends of the DBD sequence. These operations were performed in the Modeler suite [67], which was also used to eliminate possible strains and steric clashes introduced during model building. The program applied the DOPE scoring algorithm to identify the best possible structures that were subsequently employed in docking and simulated annealing procedures. The former was carried out to place the DBE structure, which was created separately by using the make-na server (http://structure.usc.edu/make-na/), onto the putative binding site of the protein. This operation was accomplished in HADDOCK [73] by designating as active those residues that had experienced reduced rates of exchange (**Figure 3A**) upon DNA binding (see **Figure S-27** of ***Supporting Information***). The mutual positioning between the DBE and FOXO4-DBD components was further refined according to the results of the protein-DNA cross-linking experiments, which were introduced by using Modeller software. Finally, the structures of both FOXO4-DBD and FOXO4-DBD•DBE complex were submitted to simulated annealing and energy minimization in CNS [71] to generate the sought-after model ensembles (**Figure 5**).

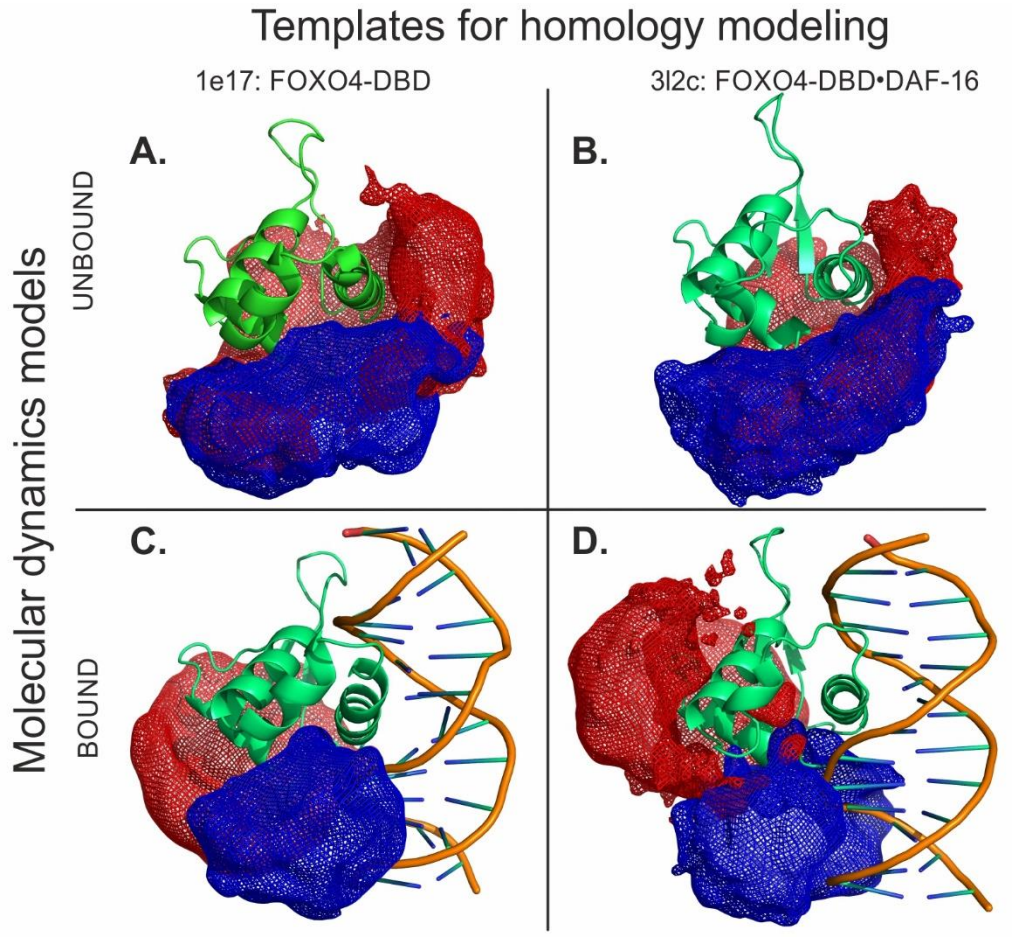

Figure 3: Structural proteomics could effectively guide model-building operations to produce very high-quality 3D models. Models of FOXO4-DBD and FOXO4-DBD•DBE were obtained by combining homology modelling with experimental constraints and molecular dynamics simulations. These models incorporated extensive information from protein-DNA cross-links, quantitative protein-protein cross-links, and hydrogen-deuterium exchange. The green structures show representative models for unbound (**A.** and **B.**) and bound (**C.** and **D.**) forms based on corresponding 1E17 (**A.** and **C.**) or 3L2C (**B.** and **D.**) high-resolution templates. The mesh areas in blue and red colors represent the spaces occupied by all the models in the ensembles, which provided a measure of the flexibility of the N- and C- terminal regions (see also **Figure S-10** to **Fig S-32** and **Table S-1** and **S-2** of ***Supporting Information***).

Peer-reviewed version available at Biomolecules 2019, 9, 535; doi:10.3390/biom9100535

The structures obtained for FOXO4-DBD and FOXO4-DBD•DBE complex were compared to the corresponding high-resolution structures to assess the robustness of our structural proteomics approach. In the case of individual FOXO4-DBD, the overall topology of the ensemble reflected the typical forkhead structure of the FOXO family and matched very closely that of the NMR structure used as homology template (see **Figure 28A** and **B** of ***Supporting information***), thus supporting the validity of the HDX and cross-linking constraints. A more detailed comparison was obtained by calculating the root mean square deviation (RMSD) between the coordinates of corresponding heavy atoms located in the backbone of each ensemble model and the various templates. The representative plot in **Figure S-29** of ***Supporting Information***, for example, shows that the model obtained from the 1e17 structure deviated very little from the initial template. The fact that the experimental constraints introduced during modelling did not force any significant variation of the initial coordinates indicates that the probing operations did not cause any perturbation of the substrate's 3D structure and corroborated the excellent stability of the structured core of FOXO4-DBD. In contrast, larger RMSD values were obtained when models based on other templates were compared with the initial FOXO4 models used in the study (3L2C, 1E17), as expected from the discrepancies between the various NMR and crystal structures available (see **Figure S-30** in ***Supporting Information***). Although this type of analysis was not possible for the regions that were absent from the templates, the excellent match manifested by the regions present in both model and template warranted a high level of confidence in the entire structures produced by our approach.

The ensemble of the FOXO4-DBD•DBE complex was examined in similar fashion. Also, in this case, the overall topology matched very closely that of the corresponding high-resolution template (i.e., 3l2c), with the DBE component oriented in the proper direction and placed in the correct position onto the FOXO4-DBD's binding site (see **Figure 28C** and **D** of ***Supporting information***). Also, in this case, RMSD comparisons between the model and crystal structure revealed excellent match for the regions present in both, thus ruling out the possibility of inadvertent perturbations introduced by the probing procedures. Additionally, we determined the distances between residues that had been conjugated by the cross-linking reagents, and then compared them with the corresponding distances measured on the crystal structure. The resulting RMSD values revealed excellent agreement across the board, with the sole exception of the distances between the DBE molecule (DBE_F) and specific residues of the H_{164} - M_{175} loop (H_{164}, T_{168}, S_{171}, S_{172} or M_{175}), which were somewhat longer in our model (see **Table S-1** in ***Supporting Information***). These discrepancies, however, were consistent with the high degree of flexibility possessed by the loop, which were manifest also in the higher B-factors displayed by this region in the crystal structure. In agreement with the crystal structure, the models showed that helix H3 represents the main interaction interface, as indicated by both HDX and cross-linking data (see **Table S-1** in ***Supporting Information***). The DBE structure employed here replicated the consensus binding sequences for all related FOX factors, which contain a general TGTTT motif surrounded by more variable sequences (see **Figure S-31** in ***Supporting Information***). Whereas FOXO3 and FOXO6 recognize two nucleotides located after this consensus sequence, FOXP1 only recognize the second nucleotide but not the first one. In contrast, our model indicates that FOXO4 may recognize one nucleotide before and one after the consensus motif, thus affording additional evidence of the uniqueness of the interactions established by this member of the FOXO family. Our results are supported with binding models estimated probability of individual bases at each position in sequence according Position Count Matrix (PCM) values (see **Figure S-31** in **Supporting Information**).

A close comparison of the structures of FOXO4-DBD and FOXO4-DBD•DBE complex obtained by our approach allowed us to further explore the effects of binding on protein conformation. The examination confirmed that the N- and C-terminal sequences remained largely unstructured even after DBE binding, as shown by the mesh representation obtained from our models (**Figure 5**) The main interface region represented by H3 showed limited variations between unbound and bound forms. Similar outcomes were also observed for the contiguous H1 - H2 loop and H2 helix. In contrast, loop H2 – H4 – H3and the S2 and S3 strands of wing W1 showed rather large variations upon binding. Additionally, also the N- and C- terminal regions manifested extensive variations.

Peer-reviewed version available at Biomolecules 2019, 9, 535; doi:10.3390/biom9100535

These observations were consistent with the HDX data that revealed clearly peculiar time dependencies. For instance, the initially increasing rates in the H1 helix, H1 – H2 loop, and H2 helix (A_{103} – T_{130}, in particular) suggested variations of dynamics upon DBE binding, whereas their subsequent decreasing rates were consistent with the actual structural stabilization resulting from the presence of bound DNA. The RMSD values calculated for corresponding heavy backbone atoms provided an excellent measure of these conformational effects (see **Figure S-32** of ***Supporting Information***). The values obtained from flanking regions near the interface, indicating major changes in loop H4 - H3 and W1 wing and minor changes in loops H1 - H2 and H3 – S1, highlighted the indirect effects of binding, which were consistent with the results of HDX and XL experiments. The conformational changes revealed by this type of treatment were consistent with a classic adaptive binding mechanism by which rather sizeable conformational changes may be necessary to establish specific substrate-ligand interactions. The fact that the observed conformational changes were not limited only to the sequences in direct contact with the ligand DBE, but involved also contiguous regions, supports mechanisms by which binding events may trigger associated activities through allosteric effects, or place bordering regions in positions necessary to mediate the recruiting of additional factors, such as the acetyl transferases that are known to interact with other members of the FOXO family [80,81].

5. Conclusions

The structural investigation of FOXO4-DBD and FOXO4-DBD•DBE provided a thorough assessment of the ability of structural proteomics techniques to obtain valid information on systems that are not directly amenable to classic high-resolution approaches. The outcome showed that the concerted application of HDX and XL could effectively guide model-building operations to produce very high-quality 3D models. Our multi-step strategy involved the utilization of high-resolution templates to carry out initial homology modelling. The sections that were not present in the templates were generated from experimental constraints and integrated with the initial structures to cover the entire DNA-binding domain. Although the high-resolution templates did not cover the entire structure folded by our construct, they still provided sufficient overlap to enable an unbiased assessment of the validity of the results afforded by our experimental/computational workflow. In fact, the excellent match between the templates and our structures ruled out the possibility that cross-linking procedures might have introduced unwanted artefacts or structure distortion. The excellent agreement between the templates and the corresponding portions of our models ruled out such possibilities and confirmed the ability of the selected computational strategies to translate these types of experimental constraints into actual 3D structures. Validating the approach on the "known" portions of the structures was essential to support the validity of the "unknown" sections that were conspicuously absent from the templates. The fact that the results of HDX and cross-linking experiments were in consistent mutual agreement provided further proof of the robustness of our concerted approach. For these reasons, our all-atom models represent excellent comprehensive structures of full-fledged FOXO4-DBD and FOXO4-DBD•DBE, which have thus far eluded crystallography and NMR analysis.

The pictures painted by the high-resolution templates (identical protein construct was used), which were obtained by NMR and crystallography, are not only incomplete, but also static. Our extensive data provided a wealth of new information on the conformational dynamics of the protein in both unbound and bound forms. Our experiments clearly differentiated regions that are conformationally stable from those that undergo significant conformational changes upon DNA binding. The most important finding is that binding affects not only the interface region, but also the conformation of regions that are located away from the interface. This information will be essential to understand the allosteric properties of the complex and their role in recruiting additional transcriptional factors.

For example, our models confirmed that full-length FOXO4-DBD adopts the classic forkhead topology characteristic of this family of transcription factors, but corroborated also its unusual DNA binding mode that is unique among those manifested by the highly homologous FOXO proteins

Peer-reviewed version available at Biomolecules 2019, 9, 535; doi:10.3390/biom9100535

[44]. The close match between our model and the crystal structure of FOXO4-DBD•DBE ruled out the possibility that crystal packing might be the cause of the significant differences noted between the types of interactions established by FOXO4 versus the other members of the family [62]. Our models confirmed the prominent role of helix H3 in such interactions and highlighted the involvement of neighbouring regions, which may be responsible for fine-tuning the sequence-specific recognition of the correct DNA counterpart.

Beyond the structural insights obtained by this study, the nature of the selected techniques enabled the acquisition of valuable information of the regions involved in the binding interactions, which will be essential to understand the mechanism of action at the molecular level. The greater conformational stability exhibited by the bound complex than the unbound protein may substantiate the selectivity of FOXO4-DBD toward DBE over alternative DNA sequences. In contrast, the increased dynamics detected in regions away from the binding site may support the execution of additional functions associated with the binding event, the significance of which remains to be investigated.

In conclusion, this study demonstrated the merits of structural proteomics approaches for the elucidation of protein-nucleic acid complexes. The utilization of specific hydrolytic procedures to complete the characterization of HDX and cross-linking products virtually eliminates any limitation pertaining the size of the species of interest, which may instead hamper NMR analysis. The effects of sample flexibility and heterogeneity, which may hamper crystallization, are not an issue for techniques designed to detect structure dynamics. Another favorable feature of these techniques consists of the modest sample demands as compared to those posed by NMR and crystallography. Propelled by continued advances in the computational approaches employed to translate the experimental results into all-atoms models, structural proteomics has rapidly emerged as a valid complement, and often alternative, to the classic high-resolution techniques. By clearly demonstrating the applicability to transcription factor-response element complexes, we hope that this study will lead to a broader utilization of structural proteomics to mitigate the chronic dearth of information on these essential regulatory systems.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Figure S-1: Native and denaturing polyacrylamide gel electrophoresis (PAGE) analysis of transplatinated FOXO4-DBD•DBE complex, Figure S-2: Native ESI-MS analysis of FOXO4-DBD•DBE complex, Figure S-3: Relative deuteration rates of peptides (1/3), Figure S-4: Relative deuteration rates of peptides (2/3), Figure S-5: Relative deuteration rates of peptides (3/3), Figure S-6: HDX butterfly plot, Figure S-7: Variations of relative deuteration rates as a function of time mapped onto the protein structure, Figure S-8: Collision-induced dissociation (CID) of peptide-oligonucleotide cross-link by transplatin, Figure S- 9: Representative data obtained from protein-DNA cross-linking by transplatin, Figure S-10: Mass spectra of cross-links dissociation products (1/17), Figure S-11: Mass spectra of cross-links dissociation products (2/17), Figure S-12: Mass spectra of cross-links dissociation products (3/17), Figure S-13: Mass spectra of cross-links dissociation products (4/17), Figure S-14: Mass spectra of cross-links dissociation products (5/17), Figure S-15: Mass spectra of cross-links dissociation products (6/17), Figure S-16: Mass spectra of cross-links dissociation products (7/17), Figure S-17: Mass spectra of cross-links dissociation products (8/17), Figure S-18: Mass spectra of cross-links dissociation products (9/17), Figure S-19: Mass spectra of cross-links dissociation products (10/17), Figure S-20: Mass spectra of cross-links dissociation products (11/17), Figure S-21: Mass spectra of cross-links dissociation products (12/17), Figure S-22: Mass spectra of cross-links dissociation products (13/17), Figure S-23: Mass spectra of cross-links dissociation products (14/17), Figure S-24: Mass spectra of cross-links dissociation products (15/17), Figure S-25: Mass spectra of cross-links dissociation products (16/17), Figure S-26: Mass spectra of cross-links dissociation products (17/17), Figure S-27: Protein-DNA docking, Figure S-28: Comparison between models of full-fledged FOXO4-DBD in either free or DBE-bound form with the corresponding high-resolution structures, Figure S-29: Average RMSD between the coordinates of backbone heavy atoms in our unbound FOXO4-DBD model and the corresponding 1E17 high-resolution template, Figure S-30: Average RMSD between the coordinates of backbone heavy atoms in six created unbound FOXO4-DBD models based on different FOX structures related to FOXO4 and the 3L2C or 1E17 high-resolution template, Figure S-31: Putative FOXO4 consensus sequences, Figure S-32: Average RMSD between the coordinates of backbone heavy atoms in our unbound FOXO4-DBD and bound FOXO4-DBD•DBE complex, Figure S-33: Protein sequence encoded by the express vector pET-15b, Table S-1: Deviations between theoretical and back-calculated distances of cross-linked residues in unbound

and bound forms obtained from our structural models, Table S-2: Quantitative protein-protein cross-linking of unbound FOXO4-DBD and bound FOXO4-DBD•DBE (complex).

Author Contributions: Conceptualization, D. F. and P. N.; Methodology D. F., H. M., P. N. and P. M.; Software D. K. and M. R.; Validation M. M. and W. M.; Formal Analysis J. C. and K. V.; Investigation J. F., J. C., L. S., R. F. and Z. K.; Resources M. M. and W. M.; Data Curation J. F., L. S., M. R. and Z. K.; Writing – Original Draft L. S.; Writing – Review & Editing D. F.; Visualization J. C. and L. S.; Supervision P. M. and P. N.; Project Administration P. N.; Funding Acquisition P. N.

Funding: This research was mainly funded by the Czech Science Foundation (grant number 16-24309S), the Ministry of Education of the Czech Republic (project LH15010, programs "NPU II" project LQ1604 and "NPU I" project LO1509), European Commission H2020 (European Network of Fourier-Transform Ion-Cyclotron-Resonance Mass Spectrometry Centers - project agreement No.731077 and European Proteomics Infrastructure Consortium providing Access - project agreement No.823839), the Charles University Grant Agency (932316, 1618218) and, in part, by the Czech Academy of Sciences (RVO61388971). This work was also funded by the National Institutes of Health R01 GM121844–01.

Acknowledgments: We acknowledge the Centre of molecular structure Core Facility at BIOCEV, a facility funded by European Regional Development Funds (CZ.1.05/1.1.00/02.0109 BIOCEV) and supported by the Czech Infrastructure for Integrative Structural Biology (LM2015043 CIISB for CMS BIOCEV funded by MEYS CR).The FP7 WeNMR (project# 261572) and H2020 West-Life (project# 675858) European e-Infrastructure projects are acknowledged for the use of their web portals, which make use of the EGI infrastructure and DIRAC4EGI service with the dedicated support of CESNET-MetaCloud.

Conflicts of Interest: The authors declare no conflict of interest.

6. Appendix (Methods in detail)

6.1. *Protein recombinant expression and purification*

Full length DBD of FOXO4 (residues 82 – 207) was recombinantly produced as a N-terminal His-tag fusion protein in BL21 competent *E. coli* (DE3). The pET-15b plasmid carrying His-tag, thrombin cleavage site and FOXO4-DBD sequences was used as a vector for transformation. After initial growth at 37°C, protein expression in transformed cells was induced by addition of IPTG to the cell culture cooled to 30°C and performed for 12 hrs. Subsequent affinity purification was performed on a column filled with a TALON Superflow Resin (Clontech Laboratories, USA) charged with Co^{2+} according to the manufacturer protocol. Cleavage of His-tag was accomplished overnight at 4°C by treatment with human thrombin (2U per mg of recombinant protein). Removal of thrombin, cleaved His-tag and buffer exchange were carried out by gel permeation chromatography on a ENrich SEC 70 10 × 300 column (Bio-Rad Laboratories, USA).

6.2. *Design of oligonucleotides and duplex DNA preparation*

Both, forward and reverse, complementary oligonucleotide strands (5'-TTG GGT AAA CAA G-3' and 5'-CTT GTT TAC CCA A-3', respectively) were purchased from Integrated DNA Technologies (USA) in standard desalted purity. Reverse oligonucleotide sequence containing DBE (TTG TTT AC) originating from human NOXA promoter localized on chromosome 18 (+ strand) in the position 59,899,552 to 59,899,564 [80]. Both strands were tested for secondary structure presence to avoid interfering structures. Forward and reverse strands were dissolved in water and mixed in an equimolar ratio and then heated up to 95 °C for 1 min. Next, the mixture was left to cool to the room temperature to form the 13 bp duplex DNA.

6.3. *FOXO4 consensus binding sequence determination, validation and comparison*

The consensus binding sequence of FOXO4 was obtained by searching the HOmo sapiens COmprehensive MOdel COllection (HOCOMOCO) v11 [82], which contains the binding models for 680 human transcription factor (TFs). First, we scanned for TF binding models with a Position Count Matrix (PCM) similar to that of FOXO4 by using HOCOMOCO and Matrix CompaRisOn.

Peer-reviewed version available at Biomolecules 2019, 9, 535; doi:10.3390/biom9100535

Approximate P-value Estimation (MACRO-APE) software [83]. We confirmed that the TGTTT consensus sequence was presented in binding models of other FOX factors. Second, we searched for TF binding models with PCM similar to FOXO4 consensus sequence using HOCOMOCO and MACRO-APE to predict TFs with binding mode highly similar to FOXO4.

6.4. Electrospray ionization (ESI) Fourier transform ion cyclotron (FTICR) MS analysis

ESI-FTICR MS analysis was carried out to verify the identity of recombinant protein samples and the integrity of the FOXO4-DBD•DBE complex. This type analysis was repeated before and after cross-linking reactions to monitor the outcome and verify the composition of intact samples before protease digestion. Samples were typically concentrated by ultrafiltration on Amicon Ultra 0.5 mL centrifugal filters (Merck, USA) and then diluted to a final 10μM protein concentration in either a 7.5mM AA, 2% AcOH solution that preserved complex association, or a 7.5mM AA, 1% AcOH, 50% MeOH solution to induce complex denaturation. Each sample was introduced into the ESI source by using a syringe pump operating at a 2μL/min flow rate. The analysis was carried out on a Bruker Daltonics (Billerica, USA) 15T-solariX XR FTICR mass spectrometer, which was calibrated by using a solution of sodium trifluoroacetate (NaTFA) to achieve a typical 1 ppm mass accuracy. Mass spectra were acquired in positive mode over the 250 - 4000 m/z range for 3 min. DataAnalysis 4.2 was employed to accomplish data interpretation and processing.

6.5. Hydrogen-deuterium exchange

6.5. 1. HDX reaction

FOXO4-DBD (20μM) or FOXO4-DBD/DBE complex (20μM, 1:1 molar ratio) were pre-incubated for 1 hr at 20°C in an H_2O-based buffer at pH 7.4 consisting of 10mM HEPES and 50mM NaCl. Deuterium exchange was initiated by 10-fold dilution of such samples into an equivalent D_2O-based buffer (10mM HEPES, 50mM NaCl, pD 7.4). The reaction was allowed to proceed at 20°C and quenched at pre-determined intervals (i.e., 0.33, 2, 5, 10, 30, 60, 180, 300 min) by removing aliquots containing 100 pmol of protein and mixing them immediately with a 1M glycine/HCl buffer at pH 2.35, followed by rapid freezing in liquid nitrogen. All HDX experiments were performed as triplicate.

6.5.2. UHPLC/ESI-FTICR MS analysis.

Samples were stored at -80°C and thawed immediately before LC-MS analysis. Each sample was injected into a home-built LC system based on a LC-20AD pump (Shimadzu corp., JPN), which was maintained at 0°C to minimize hydrogen back-exchange. Protein digestion was carried out in column(s) with immobilized protease(s) at a flow rate of 100 μL/min of 0.4% formic acid (FA) in water. The ensuing peptide products were trapped and desalted on a C8 reversed-phase microtrap (Michrom BioResources, USA) for 3 min under the same flow rate and solvent composition. Trapped peptides were eluted at a constant flow rate of 15 μL/min by using a 1200 HPLC pump (Agilent Technologies, USA), which provided a linear 5 min gradient from 5-35% of solution B (0.4% FA, 95% MeCN in water), followed by a quick step to 99% B. The released peptides were finally separated on a reversed-phase analytical column (Jupiter C18, 0.5 × 5 mm, 5 μm, 300 Å, Phenomenex, USA). Solution A consists of 0.4% FA, 2% MeCN in water. The same protocol was used to accomplish the peptide mapping of non-deuterated samples, except that H_2O-based buffer was used instead of the D_2O-based equivalent and that 300 pmol of protein was injected into the LC system. Separated peptides were introduced into the ESI source of a Bruker Daltonics (Billerica, USA) 9.4T Apex Ultra Qe FTICR mass spectrometer. Peptide mapping was carried out in positive ion mode by performing data-dependent broadband analysis. In this experiment, each MS scan was followed by six MS/MS scans, in which the six most abundant peptide ions detected in the MS scan were submitted to collision-induced dissociation.

6.5.3. Data processing.

The MS/MS data acquired in the 300 - 1800 m/z range were processed by DataAnalysis 4.2 (Bruker Daltonics, USA) and then searched by MASCOT 2.2 search engine against a library of possible digestion products obtained from the FOXO4-DBD sequence. The hits were further processed by using the MSTools package [84]. Deuteration rate was determined in the course of positive broadband MS mode analysis (300 - 1800 m/z range) and the ensuing data were processed by using the home-built DeutEx software (unpublished). For each product, the relative deuteration rate was determined as a percentage of maximum possible number of exchangeable amide hydrogens in each peptide, which was then recalculated for each amino acid residue as already described in reference [85].

6.6. Quantitative protein-protein cross-linking

6.6.1. Cross-linking reaction

FOXO4-DBD (20μM) or FOXO4-DBD/DBE complex (1 : 1 molar ratio) in a buffer solution at pH 7.4 (10mM HEPES, 50mM NaCl) were pre-incubated for 1 hr at 20°C and then mixed with either DSG or DSS cross-linking reagents. Unbound FOXO4-DBD samples were mixed with the non-labelled cross-linkers (DSGd0 or DSSd0), whereas bound FOXO4-DBD•DBE complex was mixed with the deuterium-labelled cross-linkers (DSGd4 or DSSd4). All cross-linkers were dissolved in dimethyl sulfoxide (DMSO) to a 6.74mM concentration. The amount of cross-linkers added to each sample provided a 10-fold molar excess over the protein and the addition resulted in a final 5% content of DMSO and 20μM protein concentration in the reaction mixture. After 2 hr of cross-linking reaction, corresponding unlabelled and deuterium-labelled samples (e.g., DSSd0- and DSSd4-traeted) were mixed in 1 : 1 ratio to enable direct quantification. Control samples obtained by adding pure DMSO in the absence of cross-linker, as well the samples treated with 1 : 1 mixtures of cross-linkers (e.g., DSGd0/DSGd4 or DSSd0/DSSd4) were prepared at the same time and handled in the same manner. All cross-linking reactions were performed as triplicate. Reaction samples were analyzed by SDS-PAGE to monitor the formation of higher oligomeric structures as a result of cross-linking. This analysis was repeated after trypsin digestion to verify the formation of peptide products to be submitted to LC-MS.

6.6.2. Proteolytic digestion

The pH of samples was adjusted to 8.3 by addition of 100mM ethylmorpholine with 20% of MeCN. Next, trypsin dissolved in water was added to reach a final 1 : 20 (trypsin : protein) weight ratio. Digestion was performed overnight at 37°C and quenched by adjusting the solution to an acidic pH with 1% FA.

6.6.3. HPLC/ESI-FTICR MS analysis

Peptide samples diluted in 0.1% FA to a final 0.05 μg/μL concentration were injected (0.1 μg per injection) onto a trap column (ZORBAX 300SB-C18, 5 μm, 0.30 x 5 mm cartridge, Agilent, USA), in which the products were desalted by a 10 μL/min flow of 0.1% FA in water for 3 min (LC-20AD HPLC pump, Shimadzu corp., JPN). Desalted peptides were eluted from the trap column at a constant flow rate of 0.5 μL/min (HPLC pump 1200, Agilent Technologies, USA) by a two-step linear gradient. The program consisted of a slow ramp from 2 to 45% of solution B (0.1% FA, 98% MeCN in water) in 43 min, followed by a fast step from 45 to 95% in 3 min. Solution A consisted of 0.1% FA and 2% MeCN in water. The peptides were then separated on-line on an analytical column (ZORBAX 300SB-C18, 0.3 x 150 mm, 3.5 μm, Agilent, USA). Finally, the peptide products were introduced into the ESI source of a Bruker Daltonics (Billerica, USA) 15T-solarix XR FTICR mass spectrometer. Mass spectral data was acquired in positive mode across a 250 - 2000 m/z range.

6.6.4. Data processing and quantification

MS data were processed by DataAnalysis 4.2 (Bruker Daltonics, USA). In particular, the SNAP 2.0 algorithm was used to generate deconvoluted spectra and corresponding lists of singly-charged monoisotopic masses. Data obtained from control and cross-linked samples were compared with theoretical libraries by using the home-built LinX software (available online) to identify the sough-after cross-linking products. All spectra containing the signal of identified cross-links were summed and further processed by mMass 5.4.1 [66]. For each cross-linking product, the ratio between light (d0) and heavy (d4) cross-linker incorporation was calculated by using the envelope fit tool, which operated by generating theoretical profiles and then fitting them to the experimental data according to linear combination and least-square fitting. Control samples were also checked for possible 1:1 ratios of light and heavy forms.

6.7. Protein-DNA chemical cross-linking

6.7.1. Transplatin reaction

Substrate solutions containing 25μM of either FOXO4-DBD or FOXO4-DBD•DBE complex in 150mM ammonium acetate (AA) at pH 6.85 and reagent solutions containing 1mM trans-platinum(II)diammine dichloride (transplatin) prepared by dilution of 20mM stock in dimethylformamide (DMF) were pre-incubated separately for 1 hr at 18°C. Aliquots were then mixed to obtain final concentrations of 20μM substrate and 200μM reagent in 150mM AA and 1% DMF. The reaction was carried out for 14 hr at 18°C. Control samples containing the same amount of DMF in water (no transplatin) were prepared at the same time and handled in the same manner. Reaction and control samples were analyzed by native and denaturing DNA polyacrylamide gel electrophoresis, SDS polyacrylamide gel electrophoresis, ESI-FTICR MS and HPLC/ESI-FTICR MS.

6.7.2. Complex digestion

The nucleic acid component of the FOXO4-DBD•DBE complex was digested with specific nucleases to facilitate the identification of nucleotides that were directly conjugated with the protein component by the cross-linking reaction. The reaction mixture was added with a solution of Bal-31 nuclease (0.15 U per 1 μg of duplex DNA) in a 40mM TRIS/HCl buffer at pH 8.0, which contained 24mM $MgCl_2$, 24mM $CaCl_2$, and 2mM EDTA. Nuclease digestion was carried out for 1 hr at 30°C. The samples were subsequently subjected to trypsin digestion as described above.

6.7.3. HPLC/ESI-FTICR MS analysis

The digestion mixtures were analyzed in both positive and negative ion mode to better capitalize on the charging of products that may have a predominant peptide (positive) or oligonucleotide (negative) character. The former was accomplished as described above for trypsin digests, with only slightly different conditions: 0.375 μg of material per injection; 10 μL/min two-step linear gradient from 5 to 30% in 25 min and from 30 to 95% in 5 min. Mass spectral data were acquired in positive, data-independent mode over the 250 - 2500 m/z range. In this experiment, each MS scan (0.2 s accumulation) was followed by an MS/MS scan (0.8 s accumulation) with a broad isolation window (250 - 2500 m/z) and a fixed 15eV collision-induced dissociation energy. The analysis in negative ion mode was accomplished as described above, but with the following variations: 0.375 μg of material per injection; 10 mM ammonium bicarbonate at pH 7.5 for trapping and desalting operations; separation accomplished on a ZORBAX 300Extend-C18, 0.3 x 150 mm, 3.5 μm (Agilent, USA) analytical column by using a 10 μL/min two step linear gradient from 10 to 55% in 30 min and from 55 to 99% in 5 min; A solution (4% hexafluoroisopropyl alcohol (HFIP) and 0.1% triethylammonium acetate (TEAA) in water at pH 7.5), B solution (1% hexafluoroisopropyl alcohol (HFIP), 0.025% triethylammonium acetate (TEAA), 75% and MeOH in water at pH 7.5). Mass spectral data were acquired in negative, data-independent mode over the 200 - 2500 m/z range. Also, in this experiment, each MS scan (0.2 s accumulation) was followed by an MS/MS scan (0.8 s accumulation) with a broad isolation window (250 - 2500 m/z) and a fixed 15eV collision-induced dissociation energy.

6.7.4. Data interpretation

MS and MS/MS data were processed by using DataAnalysis 4.2 (Bruker Daltonics, USA). In particular, the SNAP 2.0 algorithm was employed to generate deconvoluted spectra of detected species to be compared to a theoretical library of putative peptide-oligonucleotide conjugates. The process required combining the average content of biogenic isotopes in average peptides with a fictitious additional constant corresponding to $C_{129}H_{160}N_{54}O_{74}P_{12}PtN_2H_4$, which represented a single strand of the DBE component conjugated by the transplatin cross-linker. This expedient ensured the proper determination of monoisotopic masses of putative protein-DNA cross-links products. The lists of monoisotopic masses provided by the MS determinations were searched for possible protein-DNA cross-links by using the home-built LinX software (available online). Further validation of initial hits was achieved by manually comparing the experimental isotopic patterns with the model isotopic patterns suggested by LinX, which were expected to manifest unique features conferred by the phosphate and platinum contributions from the oligonucleotide and cross-linker moieties. The MS/MS data were visually inspected in similar fashion to identify fragmentation products corresponding to the sought-after protein-DNA cross-links.

References

1. Lambert, S.A.; Jolma, A.; Campitelli, L.F.; Das, P.K.; Yin, Y.; Albu, M.; Chen, X.; Taipale, J.; Hughes, T.R.; Weirauch, M.T. The Human Transcription Factors. *Cell* **2018**, *172*, 650–665.
2. Latchman, D.S. Transcription factors: an overview. *Int. J. Biochem. Cell Biol.* **1997**, *29*, 1305–12.
3. Latchman, D.S. Transcription factors: bound to activate or repress. *Trends Biochem. Sci.* **2001**, *26*, 211–3.
4. Babu, M.M.; Luscombe, N.M.; Aravind, L.; Gerstein, M.; Teichmann, S.A. Structure and evolution of transcriptional regulatory networks. *Curr. Opin. Struct. Biol.* **2004**, *14*, 283–291.
5. Brivanlou, A.H.; Darnell, J.E. Signal transduction and the control of gene expression. *Science* **2002**, *295*, 813–8.
6. Vaquerizas, J.M.; Kummerfeld, S.K.; Teichmann, S. a; Luscombe, N.M. A census of human transcription factors: function, expression and evolution. *Nat. Rev. Genet.* **2009**, *10*, 252–63.
7. Heck, A.J.R. Native mass spectrometry: a bridge between interactomics and structural biology. *Nat. Methods* **2008**, *5*, 927–33.
8. Brent, R.; Ptashne, M. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* **1985**, *43*, 729–36.
9. Hollenberg, S.M.; Evans, R.M. Multiple and cooperative trans-activation domains of the human glucocorticoid receptor. *Cell* **1988**, *55*, 899–906.
10. Ma, J.; Ptashne, M. A new class of yeast transcriptional activators. *Cell* **1987**, *51*, 113–9.
11. Rozbeský, D.; Adámek, D.; Pospíšilová, E.; Novák, P.; Chmelík, J. Solution structure of the lymphocyte receptor Nkrp1a reveals a distinct conformation of the long loop region as compared to in the crystal structure. *Proteins* **2016**, *84*, 1304–11.
12. Rozbesky, D.; Man, P.; Kavan, D.; Chmelik, J.; Cerny, J.; Bezouska, K.; Novak, P. Chemical cross-linking and H/D exchange for fast refinement of protein crystal structure. *Anal. Chem.* **2012**, *84*, 867–70.
13. Kukacka, Z.; Rosulek, M.; Strohalm, M.; Kavan, D.; Novak, P. Mapping protein structural changes by quantitative cross-linking. *Methods* **2015**, *89*, 112–20.
14. Vandermarliere, E.; Stes, E.; Gevaert, K.; Martens, L. Resolution of protein structure by mass spectrometry. *Mass Spectrom. Rev.* **2016**, *35*, 653–665.
15. Konermann, L.; Pan, J.; Liu, Y.-H. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem. Soc. Rev.* **2011**, *40*, 1224–34.
16. Zhang, Q.; Yu, E.T.; Kellersberger, K.A.; Crosland, E.; Fabris, D. Toward building a database of

bifunctional probes for the MS3D investigation of nucleic acids structures. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1570–1581.

17. Fabris, D.; Yu, E.T. Elucidating the higher-order structure of biopolymers by structural probing and mass spectrometry: MS3D. *J. Mass Spectrom.* **2010**, *45*, 841–60.
18. Sinz, A. Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom. Rev.* **2006**, *25*, 663–82.
19. Giladi, M.; van Dijk, L.; Refaeli, B.; Almagor, L.; Hiller, R.; Man, P.; Forest, E.; Khananshvili, D. Dynamic distinctions in the Na+/Ca2+ exchanger adopting the inward- and outward-facing conformational states. *J. Biol. Chem.* **2017**, *292*, 12311–12323.
20. Kadek, A.; Kavan, D.; Felice, A.K.G.; Ludwig, R.; Halada, P.; Man, P. Structural insight into the calcium ion modulated interdomain electron transfer in cellobiose dehydrogenase. *FEBS Lett.* **2015**, *589*, 1194–1199.
21. Engen, J.R.; Wales, T.E.; Chen, S.; Marzluff, E.M.; Hassell, K.M.; Weis, D.D.; Smithgall, T.E. Partial cooperative unfolding in proteins as observed by hydrogen exchange mass spectrometry. *Int. Rev. Phys. Chem.* **2013**, *32*, 96–127.
22. Engen, J.R.; Wales, T.E.; Chen, S.; Marzluff, E.M.; Hassell, K.M.; Weis, D.D.; Smithgall, T.E. Partial cooperative unfolding in proteins as observed by hydrogen exchange mass spectrometry. *Int. Rev. Phys. Chem.* **2013**, *32*, 96–127.
23. Lennartz, F.; Bengtsson, A.; Olsen, R.W.; Joergensen, L.; Brown, A.; Remy, L.; Man, P.; Forest, E.; Barfod, L.K.; Adams, Y.; et al. Mapping the Binding Site of a Cross-Reactive Plasmodium falciparum PfEMP1 Monoclonal Antibody Inhibitory of ICAM-1 Binding. *J. Immunol.* **2015**, *195*, 3273–83.
24. Kacirova, M.; Kosek, D.; Kadek, A.; Man, P.; Vecer, J.; Herman, P.; Obsilova, V.; Obsil, T. Structural Characterization of Phosducin and Its Complex with the 14-3-3 Protein. *J. Biol. Chem.* **2015**, *290*, 16246–60.
25. Zhang, J.; Chalmers, M.J.; Stayrook, K.R.; Burris, L.L.; Wang, Y.; Busby, S.A.; Pascal, B.D.; Garcia-Ordonez, R.D.; Bruning, J.B.; Istrate, M.A.; et al. DNA binding alters coactivator interaction surfaces of the intact VDR-RXR complex. *Nat. Struct. Mol. Biol.* **2011**, *18*, 556–63.
26. Graham, B.W.; Tao, Y.; Dodge, K.L.; Thaxton, C.T.; Olaso, D.; Young, N.L.; Marshall, A.G.; Trakselis, M.A. DNA Interactions Probed by Hydrogen-Deuterium Exchange (HDX) Fourier Transform Ion Cyclotron Resonance Mass Spectrometry Confirm External Binding Sites on the Minichromosomal Maintenance (MCM) Helicase. *J. Biol. Chem.* **2016**, *291*, 12467–80.
27. Zhu, M.M.; Chitta, R.; Gross, M.L. PLIMSTEX: a novel mass spectrometric method for the quantification of protein–ligand interactions in solution. *Int. J. Mass Spectrom.* **2005**, *240*, 213–220.
28. Sperry, J.B.; Shi, X.; Rempel, D.L.; Nishimura, Y.; Akashi, S.; Gross, M.L. A mass spectrometric approach to the study of DNA-binding proteins: interaction of human TRF2 with telomeric DNA. *Biochemistry* **2008**, *47*, 1797–807.
29. Zheng, J.; Yong, H.Y.; Panutdaporn, N.; Liu, C.; Tang, K.; Luo, D. High-resolution HDX-MS reveals distinct mechanisms of RNA recognition and activation by RIG-I and MDA5. *Nucleic Acids Res.* **2015**, *43*, 1216–30.
30. Morton, V.L.; Burkitt, W.; O'Connor, G.; Stonehouse, N.J.; Stockley, P.G.; Ashcroft, A.E. RNA-induced conformational changes in a viral coat protein studied by hydrogen/deuterium exchange mass spectrometry. *Phys. Chem. Chem. Phys.* **2010**, *12*, 13468–75.
31. Novak, P.; Kruppa, G.H. Intra-molecular cross-linking of acidic residues for protein structure studies.

Eur. J. Mass Spectrom. (Chichester, Eng). **2008**, *14*, 355–65.

32. Young, M.M.; Tang, N.; Hempel, J.C.; Oshiro, C.M.; Taylor, E.W.; Kuntz, I.D.; Gibson, B.W.; Dollinger, G. High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5802–6.
33. Yu, E.T.; Zhang, Q.; Fabris, D. Untying the FIV frameshifting pseudoknot structure by MS3D. *J. Mol. Biol.* **2005**, *345*, 69–80.
34. Steen, H.; Petersen, J.; Mann, M.; Jensen, O.N. Mass spectrometric analysis of a UV-cross-linked protein-DNA complex: tryptophans 54 and 88 of E. coli SSB cross-link to DNA. *Protein Sci.* **2001**, *10*, 1989–2001.
35. Lenz, C.; Kühn-Hölsken, E.; Urlaub, H. Detection of protein-RNA crosslinks by NanoLC-ESI-MS/MS using precursor ion scanning and multiple reaction monitoring (MRM) experiments. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 869–81.
36. Sharma, K.; Hrle, A.; Kramer, K.; Sachsenberg, T.; Staals, R.H.J.; Randau, L.; Marchfelder, A.; van der Oost, J.; Kohlbacher, O.; Conti, E.; et al. Analysis of protein-RNA interactions in CRISPR proteins and effector complexes by UV-induced cross-linking and mass spectrometry. *Methods* **2015**, *89*, 138–48.
37. Loeber, R.; Michaelson, E.; Fang, Q.; Campbell, C.; Pegg, A.E.; Tretyakova, N. Cross-Linking of the DNA Repair Protein O 6 -Alkylguanine DNA Alkyltransferase to DNA in the Presence of Antitumor Nitrogen Mustards. *Chem. Res. Toxicol.* **2008**, *21*, 787–795.
38. Loeber, R.; Rajesh, M.; Fang, Q.; Pegg, A.E.; Tretyakova, N. Cross-Linking of the Human DNA Repair Protein O 6 -Alkylguanine DNA Alkyltransferase to DNA in the Presence of 1,2,3,4-Diepoxybutane. *Chem. Res. Toxicol.* **2006**, *19*, 645–654.
39. Michaelson-Richie, E.D.; Ming, X.; Codreanu, S.G.; Loeber, R.L.; Liebler, D.C.; Campbell, C.; Tretyakova, N.Y. Mechlorethamine-induced DNA-protein cross-linking in human fibrosarcoma (HT1080) cells. *J. Proteome Res.* **2011**, *10*, 2785–96.
40. Müller, D.R.; Schindler, P.; Towbin, H.; Wirth, U.; Voshol, H.; Hoving, S.; Steinmetz, M.O. Isotope-tagged cross-linking reagents. A new tool in mass spectrometric protein interaction analysis. *Anal. Chem.* **2001**, *73*, 1927–1934.
41. Kang, S.; Mou, L.; Lanman, J.; Velu, S.; Brouillette, W.J.; Prevelige, P.E. Synthesis of biotin-tagged chemical cross-linkers and their applications for mass spectrometry. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 1719–26.
42. 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*, 629–34.
43. 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*, 195–202.
44. Zhang, X.; Tang, N.; Hadden, T.J.; Rishi, A.K. Akt, FoxO and regulation of apoptosis. *Biochim. Biophys. Acta* **2011**, *1813*, 1978–86.
45. Kaestner, K.H.; Knochel, W.; Martinez, D.E. Unified nomenclature for the winged helix/forkhead transcription factors. *Genes Dev.* **2000**, *14*, 142–6.
46. 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*, 4802–12.

47. Dijkers, P.F.; Medema, R.H.; Pals, C.; Banerji, L.; Thomas, N.S.; Lam, E.W.; Burgering, B.M.; Raaijmakers, J.A.; Lammers, J.W.; Koenderman, L.; et al. Forkhead transcription factor FKHR-L1 modulates cytokine-dependent transcriptional regulation of p27(KIP1). *Mol. Cell. Biol.* **2000**, *20*, 9138–48.

48. Schmidt, M.; Fernandez de Mattos, S.; van der Horst, A.; Klompmaker, R.; Kops, G.J.P.L.; Lam, E.W.-F.; Burgering, B.M.T.; Medema, R.H. Cell cycle inhibition by FoxO forkhead transcription factors involves downregulation of cyclin D. *Mol. Cell. Biol.* **2002**, *22*, 7842–52.

49. Tang, T.T.-L.; Lasky, L. a The forkhead transcription factor FOXO4 induces the down-regulation of hypoxia-inducible factor 1 alpha by a von Hippel-Lindau protein-independent mechanism. *J. Biol. Chem.* **2003**, *278*, 30125–35.

50. 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*, 11608–17.

51. Daitoku, H.; Sakamaki, J.; Fukamizu, A. Regulation of FoxO transcription factors by acetylation and protein-protein interactions. *Biochim. Biophys. Acta* **2011**, *1813*, 1954–60.

52. van der Horst, A.; Tertoolen, L.G.J.; de Vries-Smits, L.M.M.; Frye, R. a; Medema, R.H.; Burgering, B.M.T. FOXO4 is acetylated upon peroxide stress and deacetylated by the longevity protein hSir2(SIRT1). *J. Biol. Chem.* **2004**, *279*, 28873–9.

53. Nasrin, N.; Ogg, S.; Cahill, C.M.; Biggs, W.; Nui, S.; Dore, J.; Calvo, D.; Shi, Y.; Ruvkun, G.; Alexander-Bridges, M.C. DAF-16 recruits the CREB-binding protein coactivator complex to the insulin-like growth factor binding protein 1 promoter in HepG2 cells. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 10412–7.

54. Wang, F.; Marshall, C.B.; Yamamoto, K.; Li, G.-Y.; Gasmi-Seabrook, G.M.C.; Okada, H.; Mak, T.W.; Ikura, M. Structures of KIX domain of CBP in complex with two FOXO3a transactivation domains reveal promiscuity and plasticity in coactivator recruitment. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 6078–83.

55. van der Horst, A.; de Vries-Smits, A.M.M.; Brenkman, A.B.; van Triest, M.H.; van den Broek, N.; Colland, F.; Maurice, M.M.; Burgering, B.M.T. FOXO4 transcriptional activity is regulated by monoubiquitination and USP7/HAUSP. *Nat. Cell Biol.* **2006**, *8*, 1064–73.

56. Huang, H.; Tindall, D.J. Regulation of FOXO protein stability via ubiquitination and proteasome degradation. *Biochim. Biophys. Acta* **2011**, *1813*, 1961–4.

57. Matsuzaki, H.; Ichino, A.; Hayashi, T.; Yamamoto, T.; Kikkawa, U. Regulation of intracellular localization and transcriptional activity of FOXO4 by protein kinase B through phosphorylation at the motif sites conserved among the FOXO family. *J. Biochem.* **2005**, *138*, 485–91.

58. Kops, G.J.; de Ruiter, N.D.; De Vries-Smits, A.M.; Powell, D.R.; Bos, J.L.; Burgering, B.M. Direct control of the Forkhead transcription factor AFX by protein kinase B. *Nature* **1999**, *398*, 630–4.

59. Takaishi, H.; Konishi, H.; Matsuzaki, H.; Ono, Y.; Shirai, Y.; Saito, N.; Kitamura, T.; Ogawa, W.; Kasuga, M.; Kikkawa, U.; et al. Regulation of nuclear translocation of forkhead transcription factor AFX by protein kinase B. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 11836–41.

60. Weigelt, J.; Climent, I.; Dahlman-Wright, K.; Wikström, M. 1H, 13C and 15N resonance assignments of the DNA binding domain of the human forkhead transcription factor AFX. *J. Biomol. NMR* **2000**, *17*, 181–2.

61. 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*, 8265–75.

62. 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. D. Biol. Crystallogr.* **2010**, *66*, 1351–7.
63. Scalabrin, M.; Dixit, S.M.; Makshood, M.M.; Krzemien, C.E.; Fabris, D. Bifunctional cross-linking approaches for mass spectrometry-based investigation of nucleic acids and protein-nucleic acid assemblies. *Methods* **2018**, *144*, 64–78.
64. Obsil, T.; Obsilova, V. Structural basis for DNA recognition by FOXO proteins. *Biochim. Biophys. Acta* **2011**, *1813*, 1946–53.
65. Kadek, A.; Kavan, D.; Marcoux, J.; Stojko, J.; Felice, A.K.G.; Cianférani, S.; Ludwig, R.; Halada, P.; Man, P. Interdomain electron transfer in cellobiose dehydrogenase is governed by surface electrostatics. *Biochim. Biophys. acta. Gen. Subj.* **2017**, *1861*, 157–167.
66. Strohalm, M.; Kavan, D.; Novák, P.; Volný, M.; Havlícek, V. mMass 3: a cross-platform software environment for precise analysis of mass spectrometric data. *Anal. Chem.* **2010**, *82*, 4648–51.
67. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. In *Current Protocols in Bioinformatics*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2016; pp. 5.6.1–5.6.37.
68. Pierce Chemicals *Double agents cross-linking reagents selection guide.*; Pierce Chemicals: Rockford, IL, 1999;.
69. Dans, P.D.; Crespo, A.; Estrin, D.A.; Coitiño, E.L. Structural and Energetic Study of Cisplatin and Derivatives: Comparison of the Performance of Density Funtional Theory Implementations. *J. Chem. Theory Comput.* **2008**, *4*, 740–50.
70. Rosen, M.S.; Spokoyny, A.M.; Machan, C.W.; Stern, C.; Sarjeant, A.; Mirkin, C.A. Chelating Effect as a Driving Force for the Selective Formation of Heteroligated Pt(II) Complexes with Bidentate Phosphino-Chalcoether Ligands. *Inorg. Chem.* **2011**, *50*, 1411–1419.
71. Brünger, A.T.; Adams, P.D.; Clore, G.M.; DeLano, W.L.; Gros, P.; Grosse-Kunstleve, R.W.; Jiang, J.S.; Kuszewski, J.; Nilges, M.; Pannu, N.S.; et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D. Biol. Crystallogr.* **1998**, *54*, 905–21.
72. W.L. DeLano The PyMOL Molecular Graphics System 2002.
73. Dominguez, C.; Boelens, R.; Bonvin, A.M.J.J. HADDOCK: A Protein−Protein Docking Approach Based on Biochemical or Biophysical Information. *J. Am. Chem. Soc.* **2003**, *125*, 1731–1737.
74. Wassenaar, T.A.; van Dijk, M.; Loureiro-Ferreira, N.; van der Schot, G.; de Vries, S.J.; Schmitz, C.; van der Zwan, J.; Boelens, R.; Giachetti, A.; Ferella, L.; et al. WeNMR: Structural Biology on the Grid. *J. Grid Comput.* **2012**, *10*, 743–767.
75. Kadek, A.; Mrazek, H.; Halada, P.; Rey, M.; Schriemer, D.C.; Man, P. Aspartic Protease Nepenthesin-1 as a Tool for Digestion in Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal. Chem.* **2014**, *86*, 4287–4294.
76. 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*, 2224–48.
77. Deubel, D. V. Factors governing the kinetic competition of nitrogen and sulfur ligands in cisplatin binding to biological targets. *J. Am. Chem. Soc.* **2004**, *126*, 5999–6004.
78. Nafisi, S.; Norouzi, Z. A comparative study on the interaction of cis- and trans-platin with DNA and RNA. *DNA Cell Biol.* **2009**, *28*, 469–77.

79. Buchan, D.W.A.; Minneci, F.; Nugent, T.C.O.; Bryson, K.; Jones, D.T. Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res.* **2013**, *41*, W349–W357.

80. Valis, K.; Prochazka, L.; Boura, E.; Chladova, J.; Obsil, T.; Rohlena, J.; Truksa, J.; Dong, L.F.; Ralph, S.J.; Neuzil, J. Hippo/Mst1 stimulates transcription of the proapoptotic mediator NOXA in a FoxO1-dependent manner. *Cancer Res.* **2011**, *71*, 946–954.

81. Eijkelenboom, A.; Burgering, B.M.T. FOXOs: signalling integrators for homeostasis maintenance. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 83–97.

82. Kulakovskiy, I. V.; Vorontsov, I.E.; Yevshin, I.S.; Sharipov, R.N.; Fedorova, A.D.; Rumynskiy, E.I.; Medvedeva, Y.A.; Magana-Mora, A.; Bajic, V.B.; Papatsenko, D.A.; et al. HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res.* **2018**, *46*, D252–D259.

83. Vorontsov, I.E.; Kulakovskiy, I. V.; Makeev, V.J. Jaccard index based similarity measure to compare transcription factor binding site models. *Algorithms Mol. Biol.* **2013**, *8*, 23.

84. Kavan, D.; Man, P. MSTools—Web based application for visualization and presentation of HXMS data. *Int. J. Mass Spectrom.* **2011**, *302*, 53–58.

85. Trcka, F.; Durech, M.; Man, P.; Hernychova, L.; Muller, P.; Vojtesek, B. The assembly and intermolecular properties of the Hsp70-Tomm34-Hsp90 molecular chaperone complex. *J. Biol. Chem.* **2014**, *289*, 9887–901.