Published in final edited form as:

J Proteome Res. 2015 December 04; 14(12): 5048–5062. doi:10.1021/acs.jproteome.5b00903.

Comprehensive Cross-Linking Mass Spectrometry Reveals Parallel Orientation and Flexible Conformations of Plant HOP2-MND1

Evelyn Rampler^{#1,2}, Thomas Stranzl^{#1}, Zsuzsanna Orban-Nemeth^{1,3}, David Maria Hollenstein⁴, Otto Hudecz⁵, Peter Schlögelhofer^{3,*}, and Karl Mechtler^{1,5,*}

¹Institute of Molecular Pathology, Dr.-Bohr-Gasse 7, 1030 Vienna, Austria

²Department of Analytical Chemistry, Faculty of Chemistry, University of Vienna, Währingerstr. 42, 1090 Vienna, Austria

³Department of Chromosome Biology, Max F. Perutz Laboratories, University of Vienna, Dr.-Bohr-Gasse 9, 1030 Vienna, Austria

⁴Department of Biochemistry and Cell Biology, Max F. Perutz Laboratories, University of Vienna, Dr.-Bohr-Gasse 9, 1030 Vienna, Austria

⁵Institute of Molecular Biotechnology of the Austrian Academy of Sciences, Dr.-Bohr-Gasse 3, 1030 Vienna, Austria

These authors contributed equally to this work.

Abstract

The HOP2–MND1 heterodimer is essential for meiotic homologous recombination in plants and other eukaryotes and promotes the repair of DNA double-strand breaks. We investigated the conformational flexibility of HOP2–MND1, important for understanding the mechanistic details of the heterodimer, with chemical cross-linking in combination with mass spectrometry (XL–MS). The final XL–MS workflow encompassed the use of complementary cross-linkers, quenching, digestion, size exclusion enrichment, and HCD-based LC–MS/MS detection prior to data evaluation. We applied two different homobifunctional amine-reactive cross-linkers (DSS and BS²G) and one zero-length heterobifunctional cross-linker (EDC). Cross-linked peptides of four biological replicates were analyzed prior to 3D structure prediction by protein threading and protein–protein docking for cross-link-guided molecular modeling. Miniaturization of the size-exclusion enrichment step reduced the required starting material, led to a high amount of cross-linked peptides, and allowed the analysis of replicates. The major interaction site of HOP2–MND1 was identified in the central coiled-coil domains, and an open colinear parallel arrangement of HOP2 and MND1 within the complex was predicted. Moreover, flexibility of the C-terminal capping helices of both complex partners was observed, suggesting the coexistence of a closed complex conformation in solution.

*Corresponding Authors: Peter Schlögelhofer, peter.schloegelhofer@univie.ac.at, Phone: +43-1-4277-56240; Karl Mechtler, mechtler@imp.ac.at, Phone: 0043 1 79730, Fax: 0043 1 798 7153.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Keywords

HOP2-MND1; XL-MS; SEC; cross-linking; comparative modeling; DSS; EDC; BS²G

Introduction

Mass spectrometry (MS)-based proteomics research has previously been focused primarily on identification of proteins and their modifications yet advances in instrumentation and methodological approaches have more recently enabled researchers to use MS as a tool to gain structural information on proteins and protein complexes. In particular, chemical cross-linking combined with mass spectrometry (XL-MS) has been successfully utilized to stabilize protein-protein interaction, to identify protein interaction partners, to reveal interfaces of protein complexes, and to investigate interactions within entire protein networks^{1–5}. Additionally, XL-MS has the capacity to reflect conformational heterogeneity and capture transient interactions of proteins in solution^{6–10}. The technique is complementary to classical protein characterization methods and offers several advantages. While X-ray crystallography provides atomistic detail of protein structures, it is intolerant of dynamics and experiments rely on successful crystal formation. Nuclear magnetic resonance (NMR) offers high resolution data but suffers from the fact that structural analysis is limited by protein size¹¹. XL-MS contributes relatively low-resolution information, but it is advantageous in that it can provide information on multiple protein conformations and can be used in the characterization of protein complexes that are theoretically unlimited in size. Furthermore, it can be effectively coupled to other methods^{1,2,12,13} including affinity purification-MS (AP-MS) and hydrogen-deuterium exchange-MS (HDX-MS)^{11,14}.

In this study, we analyzed the heterologously expressed and purified HOP2-MND1 complex from *Arabidopsis thaliana*. The conserved HOP2-MND1 complex is crucial for DNA repair and genomic integrity. It is an essential factor for meiotic homologous recombination in plants^{15,16} and most other eukaryotes including humans^{17–21}. It assists DNA recombinases during homologous recombination by stabilizing ssDNA/DMC1 nucleoprotein filaments and stimulating strand invasion^{19,22–24}. Recently, it was shown that in mammalian cancer cells HOP2-MND1 acts together with Rad51 in the recombination-dependent telomere maintenance pathway called alternative lengthening of telomeres (ALT), thus revealing for the first time a meiosis independent function of the complex²⁵. Moreover, mutations in Hop2 have been identified in early onset familial breast and ovarian cancer patients^{16,26} and a single amino acid deletion in the C-terminal RAD51/DMC1 interaction domain of Hop2 has been associated with XX ovarian dysgenesis²⁷. These results highlight the importance of HOP2-Mnd1, and of its interaction with RAD51 and DMC1 in mammalian cells¹⁶. The biological function of the HOP2-MND1 heterodimer is well understood in most organisms^{18,21,26,28,29} and the biochemical function has been studied in depth using mouse and yeast proteins^{19,30–33}, yet the mechanistic details how it actually stimulates homologous recombination remain hypothetical.

The overall organization of HOP2 and MND1 is highly conserved¹⁶. The HOP2-MND1 complex is ca. 55 kDa in size and both proteins consist of three major domains of about 80

amino acid residues each^{23,33}. The N-terminal domains were found to be responsible for dsDNA binding, the coiled coil regions are required for complex formation, the C-terminus of HOP2 has been proposed to bind ssDNA and both C-termini together have been shown to be important for DMC1 recombinase interaction^{23,34}. In 2014, Zhao *et al* proposed an elongated V-like structure of mouse HOP2-Mnd1 utilizing small angle X-ray scattering (SAXS) and electron microscopy analysis³⁵. Based on the measured envelopes of differently tagged proteins the authors suggested an antiparallel orientation of Hop2 and Mnd1 within the complex³⁵. The interpretation of the data appeared counterintuitive since it placed the N-terminal dsDNA binding domains on opposite sides of the complex, while functionally they appeared to be well coordinated. On the other hand, a V-shaped configuration of the HOP2-Mnd1 complex, with both proteins in antiparallel configuration, would fold the N- and C-termini back and bring the terminal domains into close proximity. Recently, Kang *et al* proposed a parallel conformation for the *Giardia lamblia* HOP2-MND1 complex based on crystallography studies³⁶.

No structures have yet been solved for any plant HOP2-MND1 complexes. Hence, we sought to combine the latest X-ray crystallography data derived from *Giardia lamblia*³⁶ with our XL-MS derived distance restraints to (1) perform comparative modeling, (2) construct atomistic models to improve our understanding of HOP2-MND1 complex formation in *A. thaliana* and (3) address the question of antiparallel or parallel configuration. We chose a complementary cross-linking approach based on two homobifunctional amine-reactive cross-linkers disuccinimidyl suberate (DSS) and bis(sulfosuccinimidyl) glutarate (BS²G)^{2,37–39}, and a zero-length heterobifunctional cross-linker 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC)^{40–42}. DSS and BS²G are primary amine specific (reacting with lysine residues and N-termini of proteins) with spacer lengths of 11.4 and 7.7 Å, respectively. EDC cross-links amines to carboxylic acids (aspartic acid residues, glutamic acid residues and C-termini of proteins) and is a zero-length cross-linker. Different types of cross-linked products are formed during experiments and to date there are co-existing nomenclature schemas^{43–45}. We classify our cross-linking products into (1) cross-links designating links between two peptides further divided into intralinks (both peptides are located on one protein) and interlinks (between two different proteins), (2) loop-links connecting two amino acids within one single peptide and (3) mono-links describing hydrolyzed dead-end links on a single peptide⁴⁴. Amine-reactive linkers result in cross-links, loop-links and mono-links, whereas in EDC mono-links are chemically unstable, meaning that only cross-links and loop-links are formed.

We performed replicate analysis (n=4) for each cross-linking reagent and identified the respective products for DSS, BS²G, BS²G_{d0/d6} and EDC. A similar approach facilitating DSS and EDC was shown to be effective by Shi *et al* during integrative modeling of the nuclear pore complex⁴². As well as producing a comprehensive, high-quality cross-linking dataset using complementary cross-linking reagents, we have established an enrichment strategy suitable for limited sample availability. Our final cross-linking workflow consisted of the following steps: (1) purification of heterologously expressed HOP2-MND1 complex, optimization of cross-linking concentration and protein digestion, (2) enrichment of the cross-linked peptides using a miniaturized SEC approach, (3) XL-MS and data analysis and (4) application of the obtained distance restraints for structural modeling of the complex.

We demonstrate the power of these combined approaches by reporting on complex conformations in solution of the plant HOP2-MND1 complex. Our results underline the parallel orientation of the two proteins and reveal the coexistence of open and closed complex conformations in solution.

Material and Methods

Reagents

The cross-linking reagents DSS, BS²G_{d0}, BS²G_{d4}, EDC and N-hydroxysulfosuccinimide (sulfo-NHS) were purchased from Pierce. BS²G_{d0/d6} was obtained from Creative Molecules Inc. All buffers were of analytical grade. H₂O with 0.1% formic acid (FA) and acetonitrile (ACN) with 0.1% FA applied for nano-HPLC were Optima® LC-MS products purchased from Fisher Scientific. Trifluoroacetic acid (TFA), dimethyl sulfoxide (DMSO), ammonium bicarbonate, 2-(N-morpholino)-ethanesulfonic acid sodium salt (MES) and hydroxylamine were ordered from Sigma Aldrich. Milli-Q water was derived from MQ Millipore Advantage purification system. For buffer exchange Vivaspin 500 (Cut-off: 30 000, 11VS50019) filters were purchased from Satorius Stedim Biotech.

HOP2-MND1 complex expression and purification

The generation of protein expression constructs of HOP2-MND1 complex from *Arabidopsis thaliana* was performed as described by Unanschou *et al.*²³ For expression the plasmids were transformed into chemically competent Rosetta (DE3) pLysS *E. coli* cells (Merck). Expression was started by autoinduction⁴⁶. On the next day, the cells were harvested by centrifugation (5000g, 30 min, 4°C). The cell paste was frozen at -20°C and thawed on ice. Subsequent re-suspension of the pellet in ice cold buffer A (1 g pellet per 5 mL of 50 mM phosphate buffer, pH 8.0, with 500 mM NaCl, 2 mM β-mercaptoethanol, 10% glycerol, 0.05% Nonidet P-40, and one tablet of Complete-Mini EDTA-free [Roche] per 50 mL of buffer) was followed by sonication on ice. Afterwards, an additional centrifugation step (20,000g, 30 min, 4°C) was conducted, and the resulting pre-cleared lysate was incubated with 5 mL of Ni-beads (Profinity IMAC Ni-charged resin; Bio-Rad) for 1 h at 4°C. The Ni-beads were washed four times with buffer A, and the recombinant 6xHis-tagged proteins were eluted with 800 mM imidazole, pH 8.0, in buffer A. After elution, the protein solution was concentrated with Vivaspin 20 10,000 MWCO polyethersulfone membrane centrifugal concentrators (Sartorius Stedim Biotech). Size exclusion (SEC) purification was performed using an ÄKTA fast protein liquid chromatography system, a Superdex 200 16/60 column (GE Healthcare Life Sciences), and buffer A. The fractions that contained the protein complex were pooled and the protein solution was concentrated. Aliquots were shock frozen in liquid nitrogen and stored at -80°C. The purity and concentration of the recombinant proteins was analyzed by SDS-PAGE using Coomassie Brilliant Blue staining and purified BSA as a standard and Bradford assay according to standard techniques. Additionally, to reduce the salt concentration and to remove detergent, the buffer was changed to 50 mM phosphate buffer, pH 8.0, with 25 mM NaCl, 2 mM β-mercaptoethanol using a second round of purification (SEC). The complex presence was confirmed with native PAGE using 1:1 dilution in sample buffer (2x Tris-glycine buffer, pH= 8.4) prior to application to a 10% gel (3M Tris-acetate, pH=7) running for 3 h at 140 V constant voltage in running buffer (1x

Tris-glycine buffer, pH=8.3) using a 200 Biorad machine. For cross-linker concentration optimization, SDS-PAGE was performed using a 10% containing SDS stacking gel (0.5 M Tris-HCl, pH=6.8) followed by a 15% running gel with 10% running buffer (1.5 M Tris-HCl, pH=8.8) for 1 h at 50 mA using a 200 Biorad system.

Cross-linking with amine-reactive homobifunctional cross-linkers

Purified HOP2-MND1 samples had a concentration of 20 μ M protein. For optimization of the cross-linking reaction the following conditions were tested: time points of 15 min, 30 min, 45 min, 1 h at 37°C; cross-linker concentrations of 0.1, 0.25, 0.5, 1.0, 2.0 mM. The optimized protocol comprised of the transfer of 10 μ L of the protein solution to a 0.6 mL low-binding tube and reaction started by adding 1 μ L of 10 mM of cross-linker (n=4), either DSS dissolved in DMSO or BS²G, BS²G_{d0d6} dissolved in H₂O. The samples were incubated for 30 min at 37°C and 750 rpm on a shaker. The heat experiment for quantitative cross-linking using differentially labeled BS²G was performed with heavy BS²G_{d4} reacted with HOP2-MND1 under the conditions described above, whereas light BS²G_{d0} was reacted with HOP2-MND1 at 95°C for 30 min and 750 rpm agitation. The samples were mixed, and the reaction was quenched by addition of 10 μ L fresh 100 mM ammonium bicarbonate buffer followed by incubation for 20 min at 37°C and 750 rpm on a shaker (pH= 8). Four biological replicates were performed.

Cross-linking with heterobifunctional zero-length cross-linker

The buffer was exchanged to 0.1 M MES, 0.5 NaCl (pH=6.0) for 200 μ L of purified HOP2-MND1 (20 μ M) using Vivaspin 500 filters and several centrifugation cycles at 3000g for 2 min (2x water washes, 3x MES buffer washes). For optimization of the cross-linking reaction the following conditions were tested: time points of 1, 2, 3, 4 h at room temperature; cross-linker concentrations of 1, 5, 10, 20, 50 mM and 5:1, 3:1, 1:1, 1:5 ratio of EDC to sulfo-NHS. For the optimized sample preparation, 15 μ L aliquots of the HOP2-MND1 complex were used and 1 μ L of a 100 mM EDC stock solution (500x molar excess) was added together with 1 μ L of 100 mM sulfo-NHS (optimized EDC:NHS ratio was: 1:1). After 2 h incubation at room temperature in the dark, 3 μ L of 100 mM hydroxylamine stock solution (~10 mM final) were added to quench the reaction (pH=7.5-8.5). Four biological replicates were performed.

Reduction, Alkylation & Digestion of cross-linked sample

The samples were reduced with 2 mM Tris-(2-carboxyethyl)-phosphine (TCEP) for 30 min at 37°C and shaken at 750 rpm in the dark. Alkylation was performed with 4.5 mM iodoacetamide at room temperature in the dark prior to trypsin digestion at 37 °C (2 μ L of 100 ng μ L⁻¹ stock were added to 20 μ L sample) over-night in the dark.

Size exclusion chromatography sample purification

Digested samples were acidified with 2 μ L of 10% TFA to stop trypsin digestion. Then, the samples were evaporated to dryness and diluted in 20 μ L 0.1% TFA. An aliquot of 10 μ L was transferred to a 0.2 mL PCR tube and fractionated via size exclusion. For size exclusion 1 μ L was injected on a Super SW2000 column (300 mm x 1 mm x 4 μ m) using an isocratic

gradient for 35 min with 30% ACN 70% water in 0.1% TFA applying a flow rate of 10 $\mu\text{L min}^{-1}$. An UltiMate 3000 HPLC μflow system was used for detection, monitoring wavelengths at 214 and 254 nm. The fractions from 10-17 min contained the cross-linked peptides and were evaporated to dryness and resuspended in 10-50 μL of 0.1% TFA prior to LC-MS/MS measurement. The miniaturized Super SW2000 column showed stable performance over several hundred injections, however for further downscaling column clogging and problems in enrichment reproducibility were observed.

Liquid chromatography tandem mass spectrometry- LC-MS/MS

A 10 μL aliquot of the sample was measured using a PepMap C18 column (500 mm x 75 μm x 3 μm) with guard column on an UltiMate 3000 HPLC RSLC nano system coupled to a Q-Exactive plus mass spectrometer (Thermo Fisher Scientific) with a Proxeon nanospray source employing gradient elution (buffer A: $\text{H}_2\text{O}/\text{FA} = 99.9/0.1$; buffer B: $\text{H}_2\text{O}/\text{ACN}/\text{FA} = 19.92/80/0.08$). Peptides were separated using a flow rate of 275 nL min^{-1} and the following gradient: 0-10 min: 2% B, 10-130 min: to 35% B, 130-135 min: to 90% B, 135-140 min: 90% B, 145-142 min: to 2% B, 142-165: 2% B. The mass spectrometer was operated in data-dependent mode (recording from 20-160 min of the LC method), selecting up to 12 precursor ions from MS1 scan (resolution = 70000) in the range of m/z 350-2000 for higher-energy collisional dissociation (HCD). Fragment ions were analyzed with resolution set to 17 500 and an ion threshold of 1e4 excluding singly and doubly charged ions. HCD was performed with 28% normalized collision energy, an isolation window of 1.6 m/z and a fixed first mass of 150 m/z . The underfill ratio was set to 10%, dynamic exclusion to 30 s and peptide match was switched off. Ion target values were 1e6 (maximum injection: 60 ms) for full scans and 5e4 (maximum injection: 500 ms) for MS/MS scans.

Data Analysis

To verify the presence of HOP2 and MND1 and possible other proteins or contaminants present in the analyzed samples RAW-files were loaded into Proteome Discoverer (version 1.4.0.288, Thermo Scientific). All MS/MS spectra were searched using MS Amanda47 against the E.coli swissprot protein sequence database (4309 sequences) concatenated with HOP2 and MND1 sequence and common contaminants (contaminants file as provided by MaxQuant48). The following search parameters were used: Carbamidomethylation on cysteine was set as a fixed modification, oxidation on methionine was set as variable modification. Monoisotopic masses were searched within unrestricted protein masses for tryptic peptides. The peptide mass tolerance was set to ± 5 ppm and the fragment mass tolerance to ± 0.03 Da. The maximal number of missed cleavages was set to 2. The result was filtered to 1% FDR using Percolator algorithm integrated in Thermo Proteome Discoverer. For the determination of cross-linking enrichment, the same Proteome Discoverer and MS Amanda47 search parameters were applied, except of changing the database to TAIR (*Arabidopsis thaliana*). Cross-link analysis was performed after Thermo Xcalibur .raw files were converted using Thermo Proteome Discoverer to generate mgf input files for pLink49 (version 1.16) analysis. In the pLink file xlink.ini the cross-links (and mono-links) for the cross-linkers of interest were configured. For lysine reactive cross-linkers mono-links and cross-links were identified (DSS: cross-link=138.0681 u, mono-link=156.0786 u; BS²G_{d0}: cross-link= 96.0206 u, mono-link=114.0312 u; BS²G_{d4}: cross-

link=100.0453 u, mono-link=118.0558 u; BS²G_{dc} CL=102.0576 u, ML=120.0682 u). For heterobifunctional EDC linkage lysine residues and acidic residues (D, E) under water removal only cross-links were identified, as the mono-link species are not stable (EDC: cross-link=-18.0106 u). In instrument.ini the instrument parameters were chosen, for amine-reactive cross-linkers “HCD” standard parameter were used, for EDC the “HCD-common” feature was used. In the pLink.ini file the cross-linkers for analysis was selected, standard settings for peptide tolerance were used, the modifications were chosen (Fixed: carbamidomethylation; variable: methionine oxidation) and the FDR filter was set to 1%. Cross-linking derived LC-MS/MS raw data and pLink results files are available at ProteomeXchange via PRIDE50 partner repository with identifier PXD001538. The data was further filtered for precursor mass deviations smaller than 5 ppm and pLink e-value scores lower than 10⁻² and PSM > 2 present in at least 2 samples (n=4) for EDC was applied. Data analysis was performed using in-house python scripts utilizing the library numpy51. Figures were generated with the python libraries matplotlib(1.4.0), matplotlib_venn (0.11)52 and seaborn (0.6.0)53. XL-MS derived cross-links were visualized with Circos (0.67-7)54. Validation of the heat experiment was performed after analyzing peptide elution features generated by the openMS (1.11.1)55 node “featureFinderCentroided”56. MS2 Spectra identified by pLink were used to annotate peptide features by matching m/z, elution time and charge state. Annotated features were combined into charge state clusters and pairs of heavy and light peptide clusters were grouped together, missing label partners were added if possible. Peptides not identified in certain runs were added by using the m/z value, charge and normalized retention time information from runs with an annotated feature. Ratios for each charge state were defined as heavy feature area divided by light feature area. To obtain the final peptide ratio the log2 ratios of all charge states were averaged. If multiple technical replicates were available, the final peptide ratio was calculated as the average log2 ratio of all charge states in all technical replicates. We estimated the 95% confidence interval for the median by bootstrapping (n=20,000). A detailed description of the data evaluation workflow for quantitative XL-MS can be found in the Supporting Information SI1.

Comparative modeling of HOP2 and MND1

The 3D structural models of the *Arabidopsis thaliana* HOP2 and MND1 proteins were individually generated using I-TASSER (version 4.3) predictions based on a threading algorithm57. The models with the highest I-TASSER C-score were selected for step-wise structural refinement based on XL-MS derived cross-links within one protein domain (either HOP2 or MND1). These intralinks were classified and filtered by Euclidean distance (ED) and solvent accessible surface distance (SASD) calculation (ED: -c1 AB -c2 AB -aa1 LYS -aa2 LYS (or ASP#GLU#LYS) -a1 CB -a2 CB -max 60; SASD: -c1 AB -c2 AB -aa1 LYS -aa2 LYS (or ASP#GLU#LYS) -a1 CB -a2 CB -max 500 -euc) applying Xwalk (command line version 0.6)58. The following cut-off values for the Xwalk derived distances were chosen: (1) DSS with ED < 35 Å and SASD < 40 Å combined with BS²G using ED < 30 Å, SASD < 35 Å, (2) EDC with ED < 25 Å, SASD < 30 Å. Intralinks not fitting the models were used as input to investigate conformational change of HOP2 and MND1 with I-TASSER. Intralinks satisfying the threshold levels were used for two additional rounds of I-TASSER predictions for step-wise structural refinement. Firstly, we applied combined DSS

and BS²G distance restraints followed by a round of EDC derived distance restraints for each HOP2 and MND1. The best scored models were used for protein-protein docking using the expert interface with HADDOCK (version 2.2)⁵⁹. Protein-protein docking was performed applying all protein-protein interlink derived restraints (derived from EDC, BS²G, DSS). Active and passive residues were determined using the prediction tool CPORT based on the I-TASSER derived models⁶⁰. Afterwards, another Xwalk calculation was performed to gain a refined cross-linker list applying previously mentioned threshold levels⁵⁸. These cross-links were used for structural refinement leading to the final model of HOP2-MND1. A detailed description of the workflow eventuating in the HOP2-MND1 result model is found below (Figure 5). The quality of the model was validated by calculating the number of matching cross-links and superimposition to the *Giardia lamblia* crystal structure (4y66, chain D, chain C)³⁶. Structure visualization and superimposition were performed by UCSF Chimera MatchMaker function using the Needleman-Wunsch algorithm, BLOSUM-62 Matrix and 30% secondary structure score inclusion⁶¹.

Results

The aim of the research was to develop a sensitive and robust analytical strategy for the structural analysis of the *A. thaliana* HOP2-MND1 complex. The overall strategy consists of the following steps: (1) purification of HOP2-MND1 heterologously expressed in *E. coli*, optimization of cross-linker concentration and protein digestion, (2) enrichment of the cross-linked peptides using a miniaturized SEC approach, (3) XL-MS and data analysis and (4) application of the resulting distance restraints for comparative structural modeling of the complex.

Purification of HOP2-MND1 and cross-linking optimization

Since the expression and purification of plant MND1 alone in *E. coli* (or other expression systems) was not successful due to its instability and propensity to aggregate, HOP2 and MND1 were co-expressed. The instability of singly-expressed MND1 is consistent with the situation *in vivo*, since MND1 in *A. thaliana* was not detected in the absence of HOP2 (Western blot analysis; meiotic chromosome spreads), demonstrating the crucial role of HOP2 in stabilizing and regulating MND1 in plants^{15,23}. After HOP2-MND1 complex purification and desalting, a final protein concentration of 1 mg mL⁻¹ (20 µM) was determined. To verify the presence of the complex prior to cross-linking, we performed a native PAGE and observed a broad major band ~ 65-75 kDa, a minor band ~ 60-65 kDa and a small fraction of complexes with higher molecular weight (Figure 1). The difference to the theoretically calculated size of HOP2-MND1 of 55kDa can be explained with different complex conformations and/or difference of HOP2-MND1 in charge and size compared to the marker protein. After complex purification, optimum cross-linking times (DSS and BS²G: 15 min, 30 min, 45 min, 1 h at 37°C, EDC: 1, 2, 3, 4 h at room temperature) and cross-linker concentrations (DSS and BS²G: 0.1, 0.25, 0.5, 1.0, 2.0 mM, EDC: 1, 5, 10, 20, 50 mM and 5:1, 3:1, 1:1, 1:5 ratio of EDC to sulfo-NHS) were investigated. The optimum conditions were based on the lowest cross-linking time and concentration needed for complete cross-linking of HOP2-MND1 which were determined to be 1 mM at 37 °C for 30 min for the amine-reactive cross-linkers (DSS and BS²G) and 10 mM at room temperature

for 2 h (with 1:1 sulfo-NHS) for the zero-length cross-linker EDC (Figure S1). Afterwards, the protein complex was digested in-solution, producing a mixture of cross-linked and non-cross-linked peptides. To estimate protein abundance, we used peptide spectra matches (PSMs), reflecting the number of MS/MS spectra measured for peptides of a specific protein. We unambiguously identified HOP2 and MND1 as the major proteins within our samples using a concatenated database comprised of *E. coli* proteins (4309 sequences), HOP2 and MND1 sequences and common contaminants (i.e. keratins). The sequence coverage for both proteins in all samples was between 94-98% based on 2500 PSMs for HOP2 and 2150 peptides for MND1 averaged over all cross-linked samples (DSS, BS²G, BS²G_{d0/d6}, EDC) and replicates (n=4 per cross-linker). Quantitative analysis of the top 3 peptide areas of HOP2 and MND1 in the top 5 identified proteins revealed a relative purity of over 99.7% within all samples and replicates (Table S1).

Enrichment of cross-linked products using a miniaturized SEC approach

Enrichment of cross-linked peptides is crucial due to the fact that the yield of cross-linked peptides in a cross-linking experiment is typically low⁶². There are several enrichment strategies available, namely (1) strong cation-exchange chromatography^{63,64}, (2) cross-linking reagents with affinity tag^{65–67}, (3) size exclusion (SEC)^{2,62,68,69}. Here, we used SEC to purify and enrich cross-linked peptides since this technique takes advantage of the higher molecular mass of the cross-linked peptides and their more bulky nature. Both factors cause them to elute earlier than non-cross-linked peptides. The chromatograms of the peptides of the HOP2-MND1 reaction mixtures and a non-cross-linked control were superimposed for detection of differences in their elution profiles (Figure S2). LC-MS/MS analysis of SEC fractions (1 min) revealed that the peptides eluted within 10-19 min, whereas the unreacted cross-linking reagents eluted later, after 20 min of chromatography. The removal of excess cross-linker, salts or other contaminants is another advantage of performing an SEC guided enrichment step prior to XL-MS. Significantly higher absorption values (~ factor 1.5-5) were measured at 214 nm within a retention of 10 -14 min for the cross-linked sample compared to the control, and EDC led to a higher cross-linked peptide fraction compared to BS²G and DSS (Figure S2). These findings are consistent with reports from Leitner *et al* showing that SEC enabled substantial depletion of peptides that are not relevant for cross-linking studies with removal of ~ 80% of the low molecular weight fraction⁴⁴. This increase of larger or bulkier peptides in the SEC experiment together with the results obtained via SDS-PAGE (Figure S1) confirm that intra- and inter-protein cross-links between residues present in the HOP2-MND1 complex are being formed during the cross-linking step of the experiment. Furthermore, upon MS analysis of the enriched peptide fractions eluting between 10-17 min we found that the yield of the spectral counts from cross-linked peptide spectra matches (PSMs of mono-links, loop-links and cross-links from pLink, FDR 1%) within the total spectral counts (PSMs of cross-linked products plus peptides PSMs derived from MS Amanda, FDR 1%) was $45 \pm 8\%$ (for all four replicates and all cross-linking reagents) meaning that almost 50% of the PSMs in the cross-linked samples correspond to peptides reacted with a cross-linker (as mono, loop or cross-link). The high proportion of cross-linked PSMs compared to peptide PSMs as well as the increased absorption values in the cross-linked samples compared to the non-cross-linked control suggests a successful enrichment of the cross-linked products via SEC.

Overall, the size exclusion procedure enables cross-linked peptide concentration on the basis of their mass and size, removal of the excess cross-linking agent which could otherwise interfere with the ionization process, and general desalting of the sample. Size exclusion enrichment of cross-linked peptides employing a 3.2 mm ID column have previously been reported, that require a minimum amount of 10 to 200 μg (42,62). To reduce the required protein amount to less than 1 μg , we employed a miniaturized 1 mm SEC column. This miniaturized approach reduced the required, injected peptide amount by a factor of 20-200 (calculated from injected peptide amounts corresponding to 1 μg of protein and compared to a 3.2 mm ID column). This represents a significant technical advancement, since purified protein complexes, especially when purified directly from the organism of interest, are often cumbersome and difficult to obtain.

XL-MS of HOP2-MND1

In-solution digestion followed by SEC and subsequent nano-HPLC/nano-ESI-Orbitrap MS/MS detection allowed the analysis of the generated cross-linked peptides. We applied complementary cross-linking reagents (DSS, BS²G, BS²G_{d0/d6}, EDC) in biological replicates (n=4) to capture potential protein-protein interaction sites and to gain structural information utilizing the varying distance restraints of the cross-linkers. In XL-MS, it is of utmost importance to evaluate the quality of the scoring of the cross-linking software. False positive cross-link identifications need to be reduced as far as possible. We chose a stringent filtering approach using a pLink false discovery rate of 1%, precursor mass deviations smaller than 5 ppm and pLink e-value scores lower than 10^{-2} to decrease the chance of false positive hits within our XL-MS derived restraints. The results of the cross-linker analysis after filtering are summarized in Table S2. The summary of all identified cross-links, loop-links and mono-links in terms of spectral counts, unique links and individual amino acid sites is shown in Table S2 and Figure 2 B. Most cross-linking were identified using the zero-length cross-linker EDC with 606 unique cross-links (DSS: 67; BS²G_{d0}: 57; BS²G_{d0/d6}: 34, Table S2) corresponding to 58 individual amino acid sites for HOP2 and 49 sites for MND1. One explanation is the larger repertoire of potential cross-linking sites of EDC compared to amine-reactive linkers (K-E, K-D compared to K-K for amine-reactive linker). Another possible reason is the difference in the linking reaction (pH, buffer, incubation time). When different cross-linkers are used, it is important to prove comparability of distance restraints for structural determination. This is not a trivial task as different conditions are demanded for specific cross-linking reagents potentially leading to conformational change or destabilization of the protein of interest. In the case HOP2-MND1, we optimized the cross-linking conditions and found stability of the complex for all cross-linkers. Moreover, comparison of HOP2-MND1 lysine reactivity by EDC and the amine-reactive cross-linkers reveals a good match of sites for mono-, loop- and cross-links (Figure 2 A). From 32 possible lysine sites for HOP2 and 23 for MND1, 27 out of 29 for HOP2 and 16 out of 19 for MND1 are shared by all cross-linkers.

The amine-reactive cross-linkers DSS and BS²G_{d0} performed relatively similar in terms of cross-linked products (DSS: 67 unique cross-links and 23 loop-links; BS²G_{d0}: 57 unique cross-links and 21 loop-links) and 60% of all cross-links identified with BS²G were also identified by DSS. Mono-links accounted for the majority of cross-linking products with

~58% of all spectral counts for all amine-reactive linkers (DSS, BS²G, BS²G_{d0d6}). Interestingly, we found comparable numbers of mono-links for DSS, BS²G and BS²G_{d0d6} (Table S2 and Figure 2 B.) The matching numbers and mono-link positions indicate that these hydrolyzed dead-end links do not depend on the length of the spacer arms but rather on the surface accessibility of the protein of interest.

Comparing BS²G_{d0} with its d0/d6 labeled version, the labeled version revealed more identified cross-linked spectra but less unique cross-linking sites (BS²G: 608 cross-link PSMs corresponding to 20 HOP2 and 12 MND1 sites; BS²G_{d0/d6}: 727 cross-link PSMs with 16 HOP2 and 11 MND1 individual sites) shown in Table S2 and Figure 2 B. The cross-links, loop-links and mono-links formed by BS²G_{d0/d6} match those formed by unlabeled BS²G (except one intraprotein link and one mono-link) and by DSS (Figure S3). The match of cross-links of the differentially labeled BS²G version further validates our data and at the same time reveals some characteristics of labeled cross-linkers. Although, BS²G_{d0d6} overall led to more PSMs than BS²G_{d0} (Table S2), less unique cross-linking sites were observed. Out of 92 BS²G_{d0d6} and 122 BS²G_{d0} cross-linking products, BS²G_{d0d6} shared all products with the d0 version except 2. Moreover, the remaining d0 hits were predominantly corresponding to lower numbers of PSMs i.e. more than 80% of the individual d0 cross-links were observed for spectral counts lower than 20 in four biological replicates. Hence, we assume that the low abundant cross-linked peptides are the ones that were not identified with the isotopic mixture of BS²G. This means that upon gain in identification confidence, loss in sensitivity is caused due to signal doublets in MS. To further investigate the potential of using differentially labeled cross-linking reagents during XL-MS (BS²G_{d0}, BS²G_{d4}), we performed heat denaturation of HOP2-MND1. More precisely, we compared control samples (mixing of BS²G_{d0}, BS²G_{d4} prior to cross-linking at standard conditions to examine the mixing effect, n=4), cross-linked samples at standard conditions (separate cross-linking at standard conditions for BS²G_{d0} and BS²G_{d4} containing samples prior to mixing, n=4) and heated samples (BS²G_{d0} containing sample was heated, BS²G_{d4} sample was treated at standard conditions, subsequent mixing for quantitative comparison, n=4) for HOP2-MND1. The median of the control and the sample treated at standard conditions were matching, showing that the mixing effect (control: XL-Mix, sample: Mix-XL) is negligible (Figure S4). The comparison of the sample containing heated HOP2-MND1 (BS²G_{d0}) versus HOP2-MND1 treated at standard conditions (BS²G_{d4}) led to 20 quantifiable cross-links (Figure S4), most of them within the same protein. Only two interprotein cross-links MND1(12)-HOP2(235) and MND1(154)-HOP2(235) remained after heating the complex, both being down-regulated by a factor of 1.6 and 3.4. Besides the decrease of the total numbers of identified cross-links by a factor of 10, we observed increasing numbers of mono-links and loop-links in the heated sample by a factor of 2. The anticipated loss of coiled coil interactions, suggested by (near complete) absence of cross-links in these domains, is evidence for successful complex heat denaturation. Furthermore, this specific decrease of identified cross-link PSMs upon heat denaturation of the complex underlines the relevance of the identified cross-links in the non-heated samples.

Moreover, we investigated the potential of replicates and XL-MS. Out of 764 hits (derived from all cross-linkers and four biological replicates) 53% unique cross-linking products (mono, loop and cross-links at specified amino acid positions) were shared by all four

replicates, 74% by three replicates and 93% by at least two of the replicates as can be deduced from Table S2. Interestingly, though overall a good replicate overlay was observed, we found significant higher numbers of unique cross-links by replicate analysis (Figure 3). Conducting one replicate increases unique cross-link identifications by ~15-20% for all amine-reactive cross-linkers and ~20% unique identifications for EDC. The analysis of a second replicate leads to a comparable gain in cross-linking with ~15% for amine-reactive cross-linker and ~10% identifications for EDC. The third replicate reveals stagnating individual EDC and ~10% more cross-linking products for BS²G_{d0} and DSS. For loop-links and mono-links a very low gain in unique hits of maximum 5% per replicate in all cross-linking reagents was observed, explained by lower complexity of the reactions (for mono-links, only one peptide-cross-linker reaction; for loop-links, two reactions with residues in close proximity) and hence earlier saturation.

A Circos Plot⁵⁴ allows comprehensive visualization of XL-MS derived cross-links (Figure 4 A and B). Depicted cross-links for the HOP2-MND1 heterodimer indicate that both linker chemistries lead to similar cross-linked regions, supporting the applicability of different linkers for structural modeling of the complex. Major sites of interaction (25% of all protein-protein links) were identified between the coiled coil regions of HOP2 and MND1. Moreover, the three most abundant interprotein links in terms of summed spectral counts over four replicates were HOP2(118)-MND1(127) with 81 PSMs (derived from EDC), HOP2(138)-MND1(127) with 74 PSMs (derived from BS²G) and HOP2(122)-MND1(127) with 66 PSMs (derived from EDC), all located in the coiled coil regions. These findings demonstrate that the most prominent regions of protein-protein interaction of the HOP2-MND1 heterodimer are the coiled coil domains, confirming previous data that identified the coiled-coil regions as indispensable for HOP2-MND1 complex formation^{23,34}. Considering the links between the coiled coil regions of HOP2 and MND1, it is important to note that the two complex partners are interlinked in a co-linear, parallel manner with the two N-terminal parts of the coiled coils, the two middle parts and the two C-terminal parts being connected to each other (Figure 4 A and B).

Additionally, major interaction between the N-terminal part of MND1 and the C-terminal domain of HOP2 were observed by all cross-linking reagents (HOP2(235)-MND1(61)/MND1(57)). Also several high abundant links from the C-terminal domain of HOP2 (position 235) to the coiled coil region of HOP2 (HOP2(235)-HOP2(155)/HOP2(152)/HOP2(138)) and the coiled coil region of MND1 (HOP2(235)-MND1(139)/MND1(137)/MND1(134)) were identified. Collectively the data suggests high flexibility of the C-terminal regions of HOP2 and MND1 and also of the N-terminal region of MND1. EDC-derived cross-links further identified cross-links for the the C-terminal region of MND1, which is otherwise mostly depleted of lysine residues. The C-terminal position 229 (MND1) has been found cross-linked abundantly to the coiled coil and N-terminal regions of both proteins. The lack of further cross-links in the MND1 C-terminal domain can be explained by low lysine abundance (only 3 lysines residues in the entire C-terminus) (Figure 4 A and B). Importantly, based on the results described above, we suggest a parallel orientation of the partner proteins HOP2 and MND1 within their complex and high flexibility for both of the C-terminal domains of HOP2-MND1 in *A. thaliana*.

Application of cross-linking restraints for structural modeling

There are different approaches to integrate distance restraints obtained in XL-MS experiments into protein structure predictions such as (1) comparative modeling, (2) de-novo modeling with partial structural information and (3) protein-protein docking⁷⁰. The structure of a target protein can be predicted using comparative modeling, if data of a structurally related protein is available from X-ray crystallography or NMR. Recently, Kang et al solved the crystal structure of the HOP2-MND1 complex from *Giardia lamblia*³⁶. This allows combining the Glhop2/Glmnd1 X-ray crystallography data with XL-MS derived distance restraints presented in the given study to construct enhanced atomistic models and improve our understanding of HOP2-MND1 interactions in general and in the model plant *A. thaliana*. We predicted HOP2 and MND2 independently by protein threading using the I-TASSER server⁵⁷ recently ranked top in the 11th Critical Assessment of Techniques for Protein Structure Prediction (CASP 11, <http://www.predictioncenter.org/casp11/>) and conducted step-wise structural refinement by different cross-linker lengths prior to protein-protein docking using the program HADDOCK⁵⁹ (Figure 5).

The first individual protein models were obtained by selecting the best-scoring models for HOP2 and for MND1 without restraints. Afterwards, classification of our XL-MS derived distance restraints was performed based on Euclidean distance (ED) and solvent accessible surface distance (SASD) thresholds. The recommended Euclidean distance (ED) cut-off for DSS is 30 Å, however longer distances were reported^{2,42,45,71}. The 30 Å distant restraint is calculated from the length of two extended lysine side chains (two times 5.5 Å for Ca-Cα) plus the spacer length (11.4 Å) adding some flexibility². Kahraman et al performed a comprehensive study of XL-MS based modeling workflows and recommended structural feature probing for DSS using a distance ruler of 34 Å based on solvent accessible surface calculations (SASD)⁷⁰. Not surprisingly, we observed higher SASD distance values compared to ED (SASD considers amino acids surface “barriers” within the protein structure, ED only calculates shortest distances) in the range of 5 Å by Xwalk⁵⁸ calculation of our intraprotein cross-link (cross-links within the same protein) list. Accordingly, we applied different ED and SASD cut-offs for all cross-linkers to classify protein intralinks (Figure 5) consistent with previously reported distance restraints for DSS and EDC⁴². In the first step, combined amine-reactive cross-linker constraints derived from DSS und BS²G were used. In the second step, the best models for HOP 2 and MND1 were further processed with EDC derived constraints. The best final protein models were selected for HOP2 (C-score=0.42; Estimated TM-score = 0.77±0.10) and MND1 (C-score=-0.06; Estimated TM-score = 0.71±0.12) and revealed good confidence scores (C-score) as well as high structural similarity (given by TM values > 0.5) between the plant HOP2-MND1 model and the *Giardia lamblia* structure (predicted as the protein structure template by I-TASSER⁵⁷). Subsequently, HADDOCK⁵⁹ protein-protein docking of these models was performed using the complete list of protein-protein (interprotein) cross-links. Xwalk⁵⁸ was applied to calculate amino-acid distances for the best scored model and selected a list of 49 cross-links (38 EDC, 11 amine-reactive derived links) fitting the defined (Figure 5) ED and SAS cut-offs. After performing iterative modeling steps, the final HOP2-MND1 model was chosen based on another Xwalk run utilizing all resulting clusters selecting with most fitting cross-links and lowest cross-linker distances. The resulting model proposes an elongated, open

HOP2-MND1 complex with two N-terminal winged-helix domains (WHDs) (Figure 6 A). HOP2 and MND1 interaction is predominantly formed *via* the central region consisting of two split coiled coils (coiled coil 1 and 2) in parallel orientation. The C-terminus of the heterodimer is composed of the third coiled coil together with capping helices. Interestingly, for all docking result clusters (using all cross-links or the refined cross-link list) parallel orientation of the complex was predicted. The quality of the calculated model was tested by superimposition of the docking-derived *A. thaliana* HOP2-MND1 complex structure on the *Giardia lamblia* HOP2-Mnd1 complex crystal structure³⁶. To quantify the difference between the *Giardia* (4Y66) and the plant complex we used RMSD calculation by best aligning pairs of chains and found 53 atom pairs with 0.951 Å (Cα-Cα) for *Giardia* MND1 and *Arabidopsis* MND1. Comparing the HOP2 proteins of the two species the average RMSD between 104 atom pairs was 1.191 Å. Hence, we conclude that our model is in good agreement with the *Giardia* crystal structure (Figure 6 B). However, compared to the *Giardia* crystal structure no twisting of the coiled coil regions was observed. The orientation of the N-terminal winged-helix domains, the conserved hydrophobic interface between the WHDs and the basic patches on their surface are similar to that showed for the *G. lamblia* crystal structure³⁶ and are a prerequisite of dsDNA-binding activity. Although parts of the hydrophobic regions could be reconstructed on the resulting 3D-model of HOP2-MND1 in *A. thaliana*, differences in spacing and angles were observed. As a control, we performed protein-protein docking of *Giardia* Mnd1 and Hop2 chains with our modeling workflow, additionally simulating fitting Lys-Lys constraints (SAS < 40, ED < 35). Interestingly, these docking results led to comparable chain arrangements for Hop2 and Mnd1 in *Giardia* as the protein domains were oriented in a parallel manner with similar spacing compared to the crystal structure. There are several possible explanations for these differences between the *Giardia* crystal structure and our XL-MS derived HOP2-MND1 model, namely (1) XL-MS is by definition a low-resolution technique due to longer distance constraints (given by the amino acid flexibility and the cross-linker spacer length) compared to high-resolution techniques such as X-ray crystallography¹¹ leading to limited protein-protein docking capabilities, (2) the simultaneous prediction of the protein structures in complex is hampered as our workflow relies on the determination of independent structures for HOP2 and MND1, (3) the *Giardia* structure was obtained by crystallization only reflecting one specific protein structure thus not monitoring protein flexibility. Although structural variance of HOP2-MND1 structures derived from crystallography and XL-MS was observed, both methods revealed a parallel domain orientation and an open conformation. Hence, our comparative modeling approach based on the *Giardia lamblia* crystal structure³⁶ and XL-MS derived distance restraints strongly supports parallel domain orientation and presence of open HOP2-MND1 in *A. thaliana*.

Discussion

While extensive studies have been conducted on HOP2-MND1, structural information on the complex is limited, hampering mechanistic understanding of the complex. The recently published *Giardia lamblia* crystal structure³⁶ allowed fascinating insights and opened new roads to mechanistic understanding of the complex, yet possible dynamic changes of the native complex in solution cannot be uncovered by crystallography. XL-MS is especially

suited to investigate protein structures and conformational changes in solution. In case a crystal structure is available, it can add crucial information on possible protein (complex) flexibility and dynamic changes. In case, no prior detailed structure information (X-ray or NMR) is available, the data obtained by XL-MS will provide important insights in low-resolution protein (complex) structure, orientation of interaction domains and surface accessibility. The workflow presented in the given study is suitable even for low amounts of protein input, further extending the broad usability of XL-MS. More specifically, we optimized the XL-MS workflow and demonstrated its capacity by structural analysis of the *Arabidopsis thaliana* HOP2-MND1 complex. No prior structure information has been available for the plant HOP2-MND1 complex, but the available crystal structure of the related *Giardia*36 complex made it a very suitable object to study the capacity of XL-MS, apart from being a genuinely interesting and relevant meiotic factor. To obtain non-ambiguous results, the plant HOP2-MND1 complex has been expressed in *E.coli* and thoroughly purified. Subsequently, cross-linking conditions were systematically tested and optimized. Our XL-MS approach facilitating different cross-linkers (EDC, BS²G and DSS) together with replicate analysis led to a comprehensive set of distance restraints for structural modeling of HOP2-MND1, once more highlighting the potential of complementary cross-linkers⁴². In this work, we could successfully miniaturize the enrichment step for cross-linked peptides, thereby enabling replicate analysis of HOP2-MND1, subsequently increasing the structural information and at the same time validating the obtained XL-MS results (Figure 3, Table S2). Miniaturization of specific sample preparation steps is a possible solution for the larger quantity of sample material required by replicate XL-MS analysis. Replicate analysis in general is a promising tool for extracting maximum structural information from a XL-MS experiment. Moreover, the growing interest in quantitative XL-MS to compare different protein states^{8,72–74} makes replicate analysis indispensable. Data validations such as filtering for spectral counts per replicate offer confirmation of biological findings and will be crucial if XL-MS methods develop into automated workflows.

Replicate cross-linking analysis of HOP2-MND1 revealed that most cross-linking hits were derived from EDC compared to DSS and BS²G (Table S2). Although differences in cross-linking reaction exist, we assume comparability of the structural results supported by our cross-linking optimization (Figure S1) and matching lysine reactivity of EDC and the amine-reactive cross-linkers for mono-, loop- and cross-links (Figure 2). The loss of cross-links and coiled coil interaction of HOP2 and MND1 upon heating the complex using differentially labeled BS²G underlines the relevance of the identified cross-links in the non-heated samples and once more highlights the potential of labeled cross-linkers for quantitative XL-MS^{8,72–74} to compare different protein state as experimental conditions can be changed much easier compared to a spectroscopy experiment, whereas with crystallography conditional change is impossible. This puts XL-MS at the forefront for the analysis of certain structural changes in proteins and complexes by small chemical co-factors (e.g. divalent ions). Visualization of all interaction sites derived by our complementary set of cross-linkers (amine-reactive cross-linkers, zero-length) answered the question of parallel³⁶ or antiparallel³⁵ orientation for HOP2-MND1 in *A. thaliana*. The most prominent regions of protein interaction between HOP2 and MND1 was found to reside in their coiled coil

domains, corroborating earlier data^{23,34}. In the coiled coil regions predominant interlinks between the two N-terminal domains, the two middle and the two C-terminal parts of the two partners were identified (Figure 4 A and B), best explained with a co-linear, parallel arrangement of HOP2 and MND1 within the complex. Furthermore, all results obtained employing molecular docking (HADDOCK⁵⁹) either using all cross-links or the refined cross-link list, predicted parallel domain orientation of the two partner proteins HOP2 and MND1 in the complex. Hence, both visual inspection of the distance restraints (Figure 4 A and B) and iterative modeling (Figure 6 A and B) led to a parallel complex association. Given this strong evidence we propose a parallel orientation of HOP2-MND1 in *A. thaliana*.

It is important to note that the presented open HOP2-MND1 complex model matches 50% of all amine-reactive and 20% of all EDC derived links. Interestingly, utilizing the spectral counts of all linked peptides, the open HOP2-MND1 complex configuration is in agreement with 70% of the amine reactive and 30% of the EDC cross-linking hits. This means, that the open model can be explained by a high number of identified cross-links (Figure 6 A), but a considerable number of cross-links violate the distance restraints dictated by the open complex configuration. Additionally, we identified pronounced cross-links between the C-terminal domain of HOP2 (confirmed with all cross-linkers) and the HOP2 coiled coil domains, the MND1 coiled coil domains and the MND1 N-terminus. We also observed strong cross-links between the MND1 C-terminus (identified with EDC) and the HOP2 coiled coil domain and the HOP2 N-terminus. These data suggest versatile flexibility of C- and N-terminal domains of both HOP2 and MND1. These observations are difficult to reconcile with an open and elongated model of HOP2-MND1 due to the long distances between the N- and C-termini in case of an anticipated rigid structure. However, postulating flexibility in the regions between the coiled coil domains, and the co-existence of a folded closed HOP2-MND1 complex in solution, would explain a large fraction of identified high-abundant peptide cross-links. In this sense, we also mapped these remaining protein intralinks on the protein models of HOP2 and MND1 obtained during the first protein-threading step (without applying distance restraints). Interestingly, also protein models with closed conformations were predicted by I-TASSER (Figure S5 A and B, AtHOP2 and AtMND1 with no distance restraints). These models explain 80% amine-reactive and 50% EDC intraprotein cross-links for MND1 and 50% of amine-reactive and 40% for EDC intraprotein cross-links for HOP2. The modified kink in the hinge region enables the interaction between the N-terminal winged-helix domains (WHDs) and the C-terminal coiled coil 3 with capping helices. Parallel orientation of the major interaction sites (central split coiled coil 1 and coiled coil 2) of HOP2 and MND1 observed in the elongated open model is preserved by the anticipated flexibility of the N- and C-terminal domains. Hence, we propose a tighter arrangement of HOP2-MND1 (closed complex version) enabling N- and C-terminal flexibility in solution. The remaining cross-links not fitting the two proposed models can be explained by several reasons, namely (1) other conformations or intermediate complex states are present in solution, (2) differences in reaction conditions (as the linking reactions were performed at different pH potentially influencing complex conformation), (3) protein aggregates or false positive identification not identified by our workflow. Interestingly, native PAGE analysis (Figure 1) of the purified HOP2-MND1 revealed a smaller distinct band ~60-65 kDa and a major broad band ~65-75 kDa, although overall

HOP2 and MND1 were identified in over 99% relative purity (Table S1). This means that although two distinct proteins bands were observed, HOP2-MND1 was the major product in solution further supporting the hypothesis that several conformation and/or intermediate states of HOP2-MND1 are present in solution. Moreover, replicate analysis, stringent filtering and matching lysine reactivity of the different cross-linkers reduce false positive interaction identifications. Therefore, we argue for high flexibility of the HOP2-MND1 complex in solution, with two dominant coexisting configurations: open and elongated and closed and folded. Currently it is unclear if these configurations are a mere representation of the complex' versatile flexibility, or if they are regulated and correlated with different molecular functions. For instance, it is anticipated that the HOP2-MND1 complex configuration will change upon interaction with a DMC1 coated ssDNA nuclear-protein filament. The situation may be further complicated in vivo because post-translational modifications (PTMs) may promote or prevent certain configurations or interactions. In two different large-scale screens, two different PTMs have been identified on MND1: for human MND1, an acetylation has been identified on the conserved serine on position 275 for Arabidopsis MND1, a phosphorylation has been identified on the conserved serine on position 8.76 The functional implications of these PTMs is currently unclear and demand future investigation. Moreover, future efforts will be needed to determine structural changes of HOP2-MND1 and their functional significance in conjunction with DMC1 and DNA. Following the methodology outlined in this study, XL-MS could potentially be used to follow conformational change induced upon DNA-binding. This aspect is of great importance to build comprehensive models of the mechanism of somatic and meiotic DNA repair reactions depending on HOP2 and MND1. Very recently, the HOP2-MND1 complex gained attention not only as a meiotic repair factor, but also for its role in promoting alternative telomere lengthening (ALT) in cancer cell lines²⁵. In this sense, we believe that our study will be very relevant to different areas of research, including cancer biology. In summary, XL-MS is a powerful technique to determine the structure of protein and protein complexes in solution, especially in combination with high-resolution techniques.

Conclusions

Here we present an optimized XL-MS workflow for structural analysis of the *A. thaliana* HOP2-MND1 complex. We successfully applied different cross-linking reagents (EDC, BS²G, DSS), a miniaturized enrichment strategy and performed four replicates gaining a comprehensive data set for structural modeling. Importantly, we present a novel XL-MS workflow for comparative modeling based on protein threading with I-TASSER⁵⁷ prior to protein-protein docking with HADDOCK⁵⁹. The use of complementary cross-linking reagents enabled us to perform a two-step refinement for iterative modeling based on a *Giardia lamblia* crystal structure³⁶. Replicate analysis enhanced structural information tremendously due to increased identification of unique cross-links. Our results strongly support the findings of a parallel orientation of the HOP2 and MND1 moieties within the plant complex, confirming previous findings concerning the related *Giardia lamblia* complex^{23,34}. Additionally, we propose versatile HOP2-MND1 conformations based on the anticipated flexibility of the C- and N-termini of both HOP2 and MND1 in solution. Future

studies will probe if this observed flexibility is relevant for the mode of DNA binding and activation of the DNA recombinases RAD51 and DMC1.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Research in the Mechtler laboratory was funded by the Austrian Science Fund (SFB F3402; P2465-B24; TRP 308-N15) and the European Commission via the Seventh Framework Programme (FP7/2007-2013), projects MEIOsys (222883-2) and PRIME-XS (262067). We want to thank IMP and IMBA for general funding as well as all technicians of the protein chemistry and mass spectrometry facility for continuous laboratory support. Research in the Schloegelhofer laboratory was supported by Rijk Zwaan, the Austrian Science Fund (I1468-B16; SFB F3402) and the European Commission via the Seventh Framework Programme (FP7/2007-2013; MEIOsys Project #222883-2). Especially, we acknowledge J. Fuchs for sample preparation support, T. Köcher and J. Stadlmann for useful and inspiring scientific discussion as well as A.S., R.B., B.L., for manuscript revision.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository⁵⁰ with the dataset identifier PXD001538.

Funding Sources

Austrian Science Fund FWF (SFB F3402, P2465-B24, TRP 308-N15, I1468-B16) the European Commission via the Seventh Framework Programme (FP7/2007-2013), projects MEIOsys (222883-2), PRIME-XS (262067)

Abbreviations

SEC	Size exclusion
XL-MS	Cross-linking mass spectrometry
DSS	disuccinimidyl suberate
BS²G	bis(sulfosuccinimidyl) glutarate
EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide

References

- (1). Petrotchenko EV, Borchers CH. Cross-linking Combined With Mass Spectrometry For Structural Proteomics. 2010:862–876.
- (2). Herzog F, Kahraman A, Boehringer D, Mak R, Bracher A, Walzthoeni T, Leitner A, Beck M, Hartl F-U, Ban N, et al. Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science*. 2012; 337:1348–1352. [PubMed: 22984071]
- (3). Schulz DM. Protein - Protein Interaction Studies by Chemical Cross-Linking and Mass Spectrometry. 2007
- (4). Ihling C, Schmidt A, Kalkhof S, Schulz DM, Stingl C, Mechtler K, Haack M, Beck-Sickinger AG, Cooper DMF, Sinz A. Isotope-labeled cross-linkers and Fourier transform ion cyclotron resonance mass spectrometry for structural analysis of a protein/peptide complex. *J Am Soc Mass Spectrom*. 2006; 17:1100–1113. [PubMed: 16750914]
- (5). Sinz A. The advancement of chemical cross-linking and mass spectrometry for structural proteomics: from single proteins to protein interaction networks. *Expert Rev Proteomics*. 2014; 11:1–11. [PubMed: 24351106]

- (6). Löster K, Hofmann W, Calvete JJ, Reutter W. Chemical cross-linking detects different conformational arrangements of platelet integrin alpha IIb beta III (gpIIb/IIIa). *Biochem Biophys Res Commun.* 1996; 229:454–459. [PubMed: 8954919]
- (7). Chavez JD, Liu NL, Bruce JE. Quantification of protein-protein interactions with chemical cross-linking and mass spectrometry. *J Proteome Res.* 2011; 10:1528–1537. [PubMed: 21222489]
- (8). Schmidt C, Zhou M, Marriott H, Morgner N, Politis A, Robinson CV. Comparative cross-linking and mass spectrometry of an intact F-type ATPase suggest a role for phosphorylation. *Nat Commun.* 2013; 4:1985. [PubMed: 23756419]
- (9). Schmidt C, Robinson CV. Dynamic protein ligand interactions--insights from MS. *FEBS J.* 2014; 281:1950–1964. [PubMed: 24393119]
- (10). Doberenz C, Zorn M, Falke D, Nannemann D, Hunger D, Beyer L, Ihling CH, Meiler J, Sinz A, Sawers RG. Pyruvate Formate-Lyase Interacts Directly with the Formate Channel FocA to Regulate Formate Translocation. *J Mol Biol.* 2014; 426:2827–2839. [PubMed: 24887098]
- (11). Sali A, Glaeser R, Earnest T, Baumeister W. From words to literature in structural proteomics. *Nature.* 2003; 422:216–225. [PubMed: 12634795]
- (12). Konermann L, Vahidi S, Sowole Ma. Mass spectrometry methods for studying structure and dynamics of biological macromolecules. *Anal Chem.* 2014; 86:213–232. [PubMed: 24304427]
- (13). Sinz A. Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. *J Mass Spectrom.* 2003; 38:1225–1237. [PubMed: 14696200]
- (14). Aebersold R, Mann M. Mass spectrometry-based proteomics. 2003; 422
- (15). Kerzendorfer C, Vignard J, Pedrosa-Harand A, Siwiec T, Akimcheva S, Jolivet S, Sablowski R, Armstrong S, Schweizer D, Mercier R, et al. The Arabidopsis thaliana MND1 homologue plays a key role in meiotic homologous pairing, synapsis and recombination. *J Cell Sci.* 2006; 119:2486–2496. [PubMed: 16763194]
- (16). Vignard J, Siwiec T, Chelysheva L, Vrielynck N, Gonord F, Armstrong SJ, Schlögelhofer P, Mercier R. The interplay of RecA-related proteins and the MND1-HOP2 complex during meiosis in Arabidopsis thaliana. *PLoS Genet.* 2007; 3:1894–1906. [PubMed: 17937504]
- (17). Pan S, Chen R, Aebersold R, Brentnall TA. Mass spectrometry based glycoproteomics--from a proteomics perspective. *Mol Cell Proteomics.* 2011; 10 R110.003251.
- (18). Leu JY, Chua PR, Roeder GS. The meiosis-specific Hop2 protein of *S. cerevisiae* ensures synapsis between homologous chromosomes. *Cell.* 1998; 94:375–386. [PubMed: 9708739]
- (19). Petukhova GV, Pezza RJ, Vanevski F, Ploquin M, Masson J-Y, Camerini-Otero RD. The Hop2 and Mnd1 proteins act in concert with Rad51 and Dmc1 in meiotic recombination. *Nat Struct Mol Biol.* 2005; 12:449–453. [PubMed: 15834424]
- (20). Petukhova GV, Romanienko PJ, Camerini-Otero RD. The Hop2 protein has a direct role in promoting interhomolog interactions during mouse meiosis. *Dev Cell.* 2003; 5:927–936. [PubMed: 14667414]
- (21). Tsubouchi H, Roeder GS. The Mnd1 Protein Forms a Complex with Hop2 To Promote Homologous Chromosome Pairing and Meiotic Double-Strand Break Repair. *Mol Cell Biol.* 2002; 22:3078–3088. [PubMed: 11940665]
- (22). Ploquin M, Petukhova GV, Morneau D, Déry U, Bransi A, Stasiak A, Camerini-Otero RD, Masson JY. Stimulation of fission yeast and mouse Hop2-Mnd1 of the Dmc1 and Rad51 recombinases. *Nucleic Acids Res.* 2007; 35:2719–2733. [PubMed: 17426123]
- (23). Uanschou C, Ronceret A, Von Harder M, De Muyt A, Vezon D, Pereira L, Chelysheva L, Kobayashi W, Kurumizaka H, Schlögelhofer P, et al. Sufficient amounts of functional HOP2-MND1 complex promote interhomolog DNA repair but are dispensable for intersister DNA repair during meiosis in Arabidopsis. *Plant Cell.* 2013; 25:4924–4940. [PubMed: 24363313]
- (24). Chan YL, Brown MS, Qin D, Handa N, Bishop DK. The third exon of the budding yeast meiotic recombination gene HOP2 is required for calcium-dependent and recombinase Dmc1-specific stimulation of homologous strand assimilation. *J Biol Chem.* 2014; 289:18076–18086. [PubMed: 24798326]
- (25). Cho NW, Dilley RL, Lampson Ma, Greenberg Ra. Interchromosomal homology searches drive directional ALT telomere movement and synapsis. *Cell.* 2014; 159:108–121. [PubMed: 25259924]

- (26). Chen Y-K, Leng C-H, Olivares H, Lee M-H, Chang Y-C, Kung W-M, Ti S-C, Lo Y-H, Wang AH-J, Chang C-S, et al. Heterodimeric complexes of Hop2 and Mnd1 function with Dmc1 to promote meiotic homolog juxtaposition and strand assimilation. *Proc Natl Acad Sci U S A*. 2004; 101:10572–10577. [PubMed: 15249670]
- (27). Zhao W, Sung P. Significance of ligand interactions involving Hop2-Mnd1 and the RAD51 and DMC1 recombinases in homologous DNA repair and XX ovarian dysgenesis. *Nucleic Acids Res*. 2015; 43:gvk259.
- (28). Gerton JL, DeRisi JL. Mnd1p: an evolutionarily conserved protein required for meiotic recombination. *Proc Natl Acad Sci U S A*. 2002; 99:6895–6900. [PubMed: 12011448]
- (29). Zierhut C, Berlinger M, Rupp C, Shinohara A, Klein F. Mnd1 is required for meiotic interhomolog repair. *Curr Biol*. 2004; 14:752–762. [PubMed: 15120066]
- (30). Enomoto R, Kinebuchi T, Sato M, Yagi H, Kurumizaka H, Yokoyama S. Stimulation of DNA strand exchange by the human TBPIP/Hop2-Mnd1 complex. *J Biol Chem*. 2006; 281:5575–5581. [PubMed: 16407260]
- (31). Pezza RJ, Petukhova GV, Ghirlando R, Camerini-Otero RD. Molecular activities of meiosis-specific proteins Hop2, Mnd1, and the Hop2-Mnd1 complex. *J Biol Chem*. 2006; 281:18426–18434. [PubMed: 16675459]
- (32). Bugreev DV, Huang F, Mazina OM, Pezza RJ, Voloshin ON, Camerini-Otero RD, Mazin AV. HOP2-MND1 modulates RAD51 binding to nucleotides and DNA. *Nat Commun*. 2014; 5:4198.
- (33). Pezza RJ, Voloshin ON, Volodin AA, Boateng KA, Bellani MA, Mazin AV, Camerini-Otero RD. The dual role of HOP2 in mammalian meiotic homologous recombination. *Nucleic Acids Res*. 2014; 42:2346–2357. [PubMed: 24304900]
- (34). Moktan H, Guiraldelli MF, Eyster Ca, Zhao W, Lee C-Y, Mather T, Camerini-Otero RD, Sung P, Zhou DH, Pezza RJ. Solution Structure and DNA-binding Properties of the Winged Helix Domain of the Meiotic Recombination HOP2 Protein. *J Biol Chem*. 2014; 289:14682–14691. [PubMed: 24711446]
- (35). Zhao W, Saro D, Hammel M, Kwon Y, Xu Y, Rambo RP, Williams GJ, Chi P, Lu L, Pezza RJ, et al. Mechanistic insights into the role of Hop2-Mnd1 in meiotic homologous DNA pairing. *Nucleic Acids Res*. 2014; 42:906–917. [PubMed: 24150939]
- (36). Kang H-aShin H-C, Kalantzi a-S, Toseland CP, Kim H-M, Gruber S, Dal Peraro M, Oh B-H. Crystal structure of Hop2-Mnd1 and mechanistic insights into its role in meiotic recombination. *Nucleic Acids Res*. 2015:1–16.
- (37). Mädler S, Bich C, Touboul D, Zenobi R. Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. *J Mass Spectrom*. 2009; 44:694–706. [PubMed: 19132714]
- (38). Rudashevskaya EL, Sacco R, Kratochwill K, Huber ML, Gstaiger M, Superti-Furga G, Bennett KL. A method to resolve the composition of heterogeneous affinity-purified protein complexes assembled around a common protein by chemical cross-linking, gel electrophoresis and mass spectrometry. *Nat Protoc*. 2013; 8:75–97. [PubMed: 23237831]
- (39). Peng L, Rasmussen MI, Chailan A, Houen G, Højrup P. Probing the structure of human protein disulfide isomerase by chemical cross-linking combined with mass spectrometry. *J Proteomics*. 2014; 108:1–16. [PubMed: 24792702]
- (40). Grabarek Z, Gergely J. Zero-length cross-linking procedure with the use of active esters. *Anal Biochem*. 1990; 185:131–135. [PubMed: 2344038]
- (41). Marekov LN. Determination of protein contacts by chemical cross-linking with EDC and mass spectrometry. *Curr Protoc Protein Sci*. 2007; Chapter 19:Unit 19.16.
- (42). Shi Y, Fernandez-Martinez J, Tjioe E, Pellarin R, Kim SJ, Williams R, Schneidman-Duhovny D, Sali A, Rout MP, Chait BT. Structural Characterization by Cross-linking Reveals the Detailed Architecture of a Coatmer-related Heptameric Module from the Nuclear Pore Complex. *Mol Cell Proteomics*. 2014; 13:2927–2943. [PubMed: 25161197]
- (43). Schilling B, Row RH, Gibson BW, Guo X, Young MM. MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically cross-linked peptides. *J Am Soc Mass Spectrom*. 2003; 14:834–850. [PubMed: 12892908]

- (44). Leitner A, Walzthoeni T, Kahraman A, Herzog F, Rinner O, Beck M, Aebersold R. Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol Cell Proteomics*. 2010; 9:1634–1649. [PubMed: 20360032]
- (45). Rappsilber J. The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J Struct Biol*. 2011; 173:530–540. [PubMed: 21029779]
- (46). Studier FW. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif*. 2005; 41:207–234. [PubMed: 15915565]
- (47). Dorfer V, Pichler P, Stranzl T, Stadlmann J, Taus T, Winkler S, Mechtler K. MS Amanda, a Universal Identification Algorithm Optimized for High Accuracy Tandem Mass Spectra. 2014
- (48). Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008; 26:1367–1372. [PubMed: 19029910]
- (49). Yang B, Wu Y-J, Zhu M, Fan S-B, Lin J, Zhang K, Li S, Chi H, Li Y-X, Chen H-F, et al. Identification of cross-linked peptides from complex samples. *Nat Methods*. 2012; 9:904–906. [PubMed: 22772728]
- (50). Vizcaino JA, Côté RG, Csordas A, Dianes JA, Fabregat A, Foster JM, Griss J, Alpi E, Birim M, Contell J, et al. The Proteomics Identifications (PRIDE) database and associated tools: Status in 2013. *Nucleic Acids Res*. 2013; 41
- (51). Van Der Walt S, Colbert SC, Varoquaux G. The NumPy array: A structure for efficient numerical computation. *Comput Sci Eng*. 2011; 13:22–30.
- (52). Hunter JD. Matplotlib: A 2D Graphics Environment. *Comput Sci Eng*. 2007; 9:90–95.
- (53). Waskom M, Evans C, Warmenhoven J, Yarkoni T, kjemmett, Meyer K, Rocher L, Hobson P, Halchenko Y, Koskinen M, et al. seaborn: v0.6.0 (June 2015). 2015
- (54). Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. Circos: An information aesthetic for comparative genomics. *Genome Res*. 2009; 19:1639–1645. [PubMed: 19541911]
- (55). Sturm M, Bertsch A, Gröpl C, Hildebrandt A, Hussong R, Lange E, Pfeifer N, Schulz-Trieglaff O, Zerck A, Reinert K, et al. OpenMS - an open-source software framework for mass spectrometry. *BMC Bioinformatics*. 2008; 9:163. [PubMed: 18366760]
- (56). Weisser H, Nahnsen S, Grossmann J, Nilse L, Quandt A, Brauer H, Sturm M, Kenar E, Kohlbacher O, Aebersold R, et al. An automated pipeline for high-throughput label-free quantitative proteomics. *J Proteome Res*. 2013; 12:1628–1644. [PubMed: 23391308]
- (57). Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*. 2008; 9:40. [PubMed: 18215316]
- (58). Kahraman A, Malmström L, Aebersold R. Xwalk: Computing and visualizing distances in cross-linking experiments. *Bioinformatics*. 2011; 27:2163–2164. [PubMed: 21666267]
- (59). Dominguez C, Boelens R, Bonvin AMJJ. HADDOCK: a protein-protein docking approach based on biochemical or biophysical data. *J Am Chem Soc*. 2003; 125:1731–1737. [PubMed: 12580598]
- (60). De Vries SJ, Bonvin AMJJ. Cport: A consensus interface predictor and its performance in prediction-driven docking with HADDOCK. *PLoS One*. 2011; 6
- (61). Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera - A visualization system for exploratory research and analysis. *J Comput Chem*. 2004; 25:1605–1612. [PubMed: 15264254]
- (62). Leitner A, Reischl R, Walzthoeni T, Herzog F, Bohn S, Forster F, Aebersold R. Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. *Mol Cell Proteomics*. 2012; 11 M111 014126.
- (63). Chen ZA, Jawhari A, Fischer L, Buchen C, Tahir S, Kamenski T, Rasmussen M, Lariviere L, Bukowski-Wills J-C, Nilges M, et al. Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J*. 2010; 29:717–726. [PubMed: 20094031]

- (64). Fritzsche R, Ihling CH, Götze M, Sinz A. Optimizing the enrichment of cross-linked products for mass spectrometric protein analysis. *Rapid Commun Mass Spectrom.* 2012; 26:653–658. [PubMed: 22328219]
- (65). Hurst GB, Lankford TK, Kennel SJ. Mass spectrometric detection of affinity purified cross-linked peptides. *J Am Soc Mass Spectrom.* 2004; 15:832–839. [PubMed: 15144972]
- (66). Tang X, Munske GR, Siems WF, Bruce JE. Mass spectrometry identifiable cross-linking strategy for studying protein-protein interactions. *Anal Chem.* 2005; 77:311–318. [PubMed: 15623310]
- (67). Tang X, Bruce JE. A new cross-linking strategy: protein interaction reporter (PIR) technology for protein-protein interaction studies. *Mol Biosyst.* 2010; 6:939–947. [PubMed: 20485738]
- (68). Leitner A, Walzthoen T, Aebersold R. Lysine-specific chemical cross-linking of protein complexes and identification of cross-linking sites using LC-MS/MS and the xQuest/xProphet software pipeline. *Nat Protoc.* 2014; 9:120–137. [PubMed: 24356771]
- (69). Zorn M, Ihling CH, Golbik R, Sawers RG, Sinz A. Mapping Cell Envelope and Periplasm Protein Interactions of *Escherichia coli* Respiratory Formate Dehydrogenases by Chemical Cross-Linking and Mass Spectrometry. 2014
- (70). Kahraman A, Herzog F, Leitner A, Rosenberger G, Aebersold R, Malmström L. Cross-Link Guided Molecular Modeling with ROSETTA. *PLoS One.* 2013; 8
- (71). Hofmann T, Fischer AW, Meiler J, Kalkhof S. Protein structure prediction guided by cross-linking restraints – A systematic evaluation of the impact of the cross-linking spacer length. *Methods.* 2015
- (72). Petrotchenko EV, Serpa JJ, Borchers CH. Use of a combination of isotopically coded cross-linkers and isotopically coded N-terminal modification reagents for selective identification of inter-peptide cross-links. *Anal Chem.* 2010; 82:817–823. [PubMed: 20050626]
- (73). Fischer L, Chen ZA, Rappsilber J. Quantitative cross-linking/mass spectrometry using isotope-labelled cross-linkers. *J Proteomics.* 2013; 88:120–128. [PubMed: 23541715]
- (74). Petrotchenko EV, Serpa JJ, Makepeace KaT, Brodie NI, Borchers CH. (14)N(15)N DXMSMS Match program for the automated analysis of LC/ESI-MS/MS cross-linking data from experiments using (15)N metabolically labeled proteins. *J Proteomics.* 2014:1–7.
- (75). Van Damme P, Lasa M, Polevoda B, Gazquez C, EloseguiArtola a, Kim DS, De Juan-Pardo E, Demeyer K, Hole K, Larrea E, et al. N-terminal acetylome analyses and functional insights of the N-terminal acetyltransferase NatB. *Proc Natl Acad Sci U S A.* 2012; 109:12449–12454. [PubMed: 22814378]
- (76). Engelsberger WR, Schulze WX. Nitrate and ammonium lead to distinct global dynamic phosphorylation patterns when resupplied to nitrogen-starved *Arabidopsis* seedlings. *Plant J.* 2012; 69:978–995. [PubMed: 22060019]
- (77). Vizcaino J, Deutsch E, Wang R. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat.* 2014; 32:223–226.

Figure 1.

Native PAGE analysis of heterologously expressed HOP2– MND1. The complex was expressed in Rosetta (DE3) pLysS *E. coli* cells and purified via two rounds of SEC (Superdex 200 16/60 column). 10% resolving gel; S1; 4 μ L of 20 μ M HOP2–MND1; S2; 2 μ L of 20 μ M HOP2–MND1; M: molecular weight marker.

Figure 2. Venn diagrams of reactive lysine sites and amine-reactive cross-linking products. (A) Overlap of reactive lysine sites in HOP2– MND1 for EDC, DSS, and BS²G (combined BS²G_{d0} and BS²G_{d0d6}). (B) Identified unique amino acid sites observed for mono-, loop-, and cross-links in HOP2 and MND1.

Figure 3.

Gain of unique protein positions per replicates. The plot shows the gain in unique cross-linking sites for each cross-linker and cross-linking product. The given number of unique cross-linking sites is the average over the replicates. For every replicate shown, the average is based on all relevant combinations of the four replicates.

Figure 4.

Circos plot⁵⁴ of the amine-reactive- (A) and EDC- (B) derived cross-links within the three domains of HOP2 (red) and MND1 (brown). Line thickness corresponds to the number of identified cross-link spectra. Lysine positions are marked in yellow. (A) Amine-reactive links: Gray lines correspond to cross-links derived from DSS, whereas green lines indicate BS²G cross-linking products. (B) EDC links: Black lines correspond to cross-links derived from EDC. MND1_N: N-terminus of MND1 protein. MND1_CC: coiled-coil region of MND1. MND1_C: C-terminal domain of MND1. HOP2_N: N-terminal domain of HOP2. HOP2_CC: coiled-coil region of HOP2. HOP2_C: C-terminus of HOP2.

Figure 5.

Comparative modeling workflow to analyze the open conformation of the HOP2–MND1 heterodimer. (A) Modeling of the individual proteins HOP2 and MND1. (B) Protein–protein docking for the HOP2–MND1 model in complex.

Figure 6.

Structural models for HOP2-MND1 in *A. thaliana*. (A) *A. thaliana* HOP2-MND1 model predicted by HADDOCK.59 Matching distance restraints are shown as black lines. AtHOP2 is shown in red and AtMND1 in blue. (B) *A. thaliana* HOP2-MND1 superimposed on the *G. lamblia* using GIHOP2 as the reference chain. The *A. thaliana* HOP2 chain is labeled in red, MND1 in dark blue, the *Giardia* Hop2 shown in orange, and the Mnd1 in light blue. HOP2 and MND1 are organized in an elongated, open complex with two N-terminal winged-helix domains (WHDs). The central region consists of two split coiled coils (coiled coil 1 and 2) in parallel orientation and represent the major interaction site. Coiled coil 3, together with the capping helices, form the C-terminus of the heterodimer.