Probing the Conformation of the ISWI ATPase Domain With Genetically Encoded Photoreactive Crosslinkers and Mass Spectrometry*⑤

Ignasi Forné‡, Johanna Ludwigsen§, Axel Imhof‡§, Peter B. Becker§, and Felix Mueller-Planitz§¶

We present a strategy for rapidly gaining structural information about a protein from crosslinks formed by genetically encoded unnatural amino acids. We applied it to ISWI, a chromatin remodeling enzyme involved in chromatin assembly, DNA replication and transcription. ISWI is part of the vast Snf2 family of helicase-related proteins, many of which constitute the catalytic cores of chromatin remodeling complexes. Structural information about this family is scarce, hampering our mechanistic understanding of chromatin remodeling. Making use of cells that harbor a special tRNA/aminoacyl-tRNA synthetase pair, several residues within the ATPase domain of ISWI were individually substituted with the UV-reactive unnatural amino acid *p*-benzoyl-*p*-phenylalanine. Intramolecular crosslinks could be mapped with amino acid precision by high resolution tandem mass spectrometry and the novel bioinformatic tool "Crossfinder." Most crosslinks were fully consistent with published crystal structures of ISWI-related ATPases. A subset of crosslinks, however, disagreed with the conformations previously captured in crystal structures. We built a structural model using the distance information obtained from the crosslinks and the structure of the closest crystallized relative, Chd1. The model shows the ATPase lobes strongly rotated against each other, a movement postulated earlier to be necessary to achieve a catalytically competent state. The minimal requirements for solubility and protein amounts make our approach ideal for studying structures and conformations of proteins that are not amenable to conventional structural techniques.   *Molecular & Cellular Proteomics 11: 10.1074/mcp.M111.012088, 1–11, 2012.*

Crosslinking methods have been powerful tools for decades to obtain information about the structural organization of proteins (1). Under the assumption that crosslinks only form between neighboring subunits, rough topological models of protein complexes could be delineated (2). With advances in

MS instrumentation and computational analysis of the MS data, it became possible to precisely determine the residues involved in the crosslink (3–5). Crosslinks provide constraints on the through-space distance of the attachment sites, and this information can aid structure prediction (6, 7), can distinguish protein conformations (8), and can identify the interaction surface between proteins (9, 10) and between proteins and their ligands (11). Crosslinking-MS-based methods are widely applicable as they only require microgram quantities of protein, are not limited by protein size or solubility and are relatively tolerant against sample heterogeneity. Moreover, crosslinking approaches can also be applied *in vivo* (12–16).

Two strategies are employed to crosslink proteins. Most simply, a bifunctional chemical compound is added to the protein sample. Alternatively, an amino acid with a photoreactive side chain moiety is site-specifically incorporated into the polypeptide during synthesis. The latter method has several advantages (see also Discussion). Most importantly, with only a single reactive group present per polypeptide, at most one crosslink will be formed per molecule, eliminating the risk that multiple consecutive crosslinks within one molecule distort its native structure. In addition, the crosslinking amino acid can be placed at any position throughout the entire polypeptide chain during synthesis, whereas crosslinkers added in *trans* will preferentially react with surface-exposed residues.

Photoreactive amino acids have traditionally been incorporated into peptides by chemical synthesis. A number of years ago however, Schultz and Chin introduced a technique allowing the unnatural amino acid *p*-benzoyl-*p*-phenylalanine (Bpa)[1] to be genetically encoded (17). Bpa is incorporated into the polypeptide chain by the ribosome during translation (Fig.
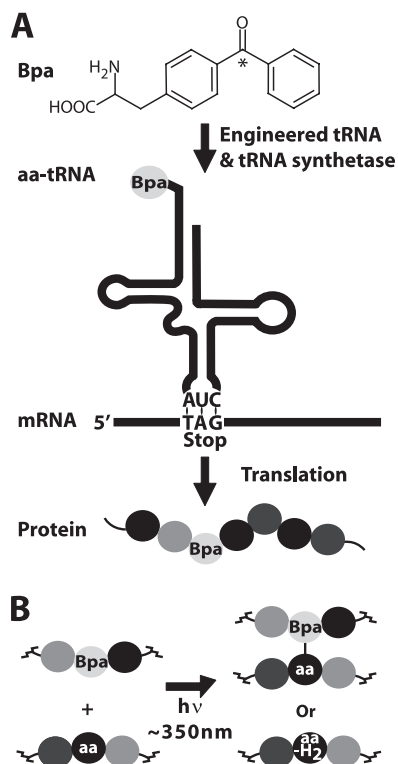
FIG. 1. **Photo-crosslinking with a genetically encoded unnatural amino acid.** *A*, The photo-crosslinkable amino acid Bpa (*p*-benzoyl-*p*-phenylalanine) is covalently joined to a mutant tRNA by a Bpa-specific aminoacyl-tRNA synthetase (17, 23). The anticodon of the tRNA is complementary to the TAG stop codon, allowing incorporation of Bpa during translation. The asterisk marks the carbonyl group that forms a triplet state diradical upon UV excitation. *B*, Photoactivation of Bpa by UV light can result in radical-mediated crosslinking to a nearby amino acid (aa) or in hydrogen elimination (-H$_2$) without formation of a crosslink.

1*A*). No laborious chemical synthesis is needed and the length of the polypeptide is not restricted. When irradiated with long-wavelength UV light, Bpa can crosslink to aliphatic side chains of other amino acids. Alternatively, the activated chromophore can sequentially abstract two hydrogen atoms from nearby amino acids (Fig. 1*B*) (18). However, since its conception the potential of this technique to provide structural information has not been fully exploited as no simple and automatable pipeline for mapping the acceptor sites of the crosslinks existed.

ISWI belongs to the Snf2 family, a family that shares a conserved helicase-related ATPase domain (Fig. 2*A*) (19). Snf2 enzymes constitute the catalytic core of ATP-dependent chromatin remodeling complexes. Remodeling complexes influence the structure and dynamics of chromatin and thus are involved in a multitude of genetic processes. Despite considerable effort by many laboratories, structures of the ATPase region of only three members of the Snf2 family have been solved (Fig. 2*A*). The structures show that the ATPase region contains two conserved lobes. However, the lobes were crys-

tallized in drastically different orientations to each other, the significance of which is currently unclear (20–22). No structure of the ISWI ATPase region is available.

To learn about the structure of the ATPase region of ISWI in solution, we introduced the genetically encoded unnatural amino acid Bpa at strategic positions. UV-induced crosslinks were mapped by high resolution tandem mass spectrometry and a novel software tool termed "Crossfinder." Crossfinder provides a fully automated pipeline to identify crosslinks. Although most crosslinks were consistent with existing structures, some were not. The distance information obtained from the crosslinks was used to build a structural model for the ATPase region of ISWI that satisfied the experimental constraints.

Besides revealing a novel conformation of an ATPase domain of the Snf2 family, our study demonstrates the general feasibility of gaining structural insights from crosslinking of genetically encoded unnatural amino acids. The clarity of the data, automation of the mapping procedure and requirements of modest amounts of starting material (as little as 1 μg), make the described method straightforward to use and amenable to a wide range of applications.

EXPERIMENTAL PROCEDURES

Mutagenesis—pPROEX-HTb-based expression plasmids with genes encoding full-length *Drosophila* ISWI or amino acids 26 to 648 of ISWI were kindly provided by Christoph Mueller (EMBL Heidelberg). The encoded proteins are referred to as ISWI FL and ISWI$^{26-648}$, respectively. Both genes are fused N-terminally to a 6xHis-TEV tag. These constructs served as the template for mutagenesis. TAG stop codons were inserted at the appropriate positions by Quickchange mutagenesis (Fig. 2*B*). Mutants were fully sequenced and no secondary mutations were found.

Enzyme Expression and Purification—pSUP-Bpa (kind gift of Peter Schultz, The Scripps Research Institute) encoding the suppressor tRNA and an engineered aminoacyl-tRNA synthetase that charges the suppressor tRNA with Bpa (23) was transformed into BL21-Gold(DE3) competent cells (Stratagene, Heidelberg, Germany). Chemically competent cells were prepared from a single transformant colony according to standard protocols. The mutagenized ISWI plasmids were individually transformed into this strain. Transformants were selected for on chloramphenicol and ampicillin for 24 h at 37 °C. Two colonies per construct were used for small scale expression tests. Large scale expression was performed from glycerol stocks in 1 L LB medium supplemented with 34 mg/L chloramphenicol, 200 mg/L ampicillin, and 1 mM Bpa (Bachem). Cells were induced with 0.2 mM isopropyl β-D-thiogalactoside at 20 °C over night.

ISWI$^{26-648}$ proteins were purified as follows. A cell lysate was prepared using a French Press (Thermo Spectronic) and ultrasonication (Branson, Danbury, CT). The lysate was clarified by centrifugation (30 min, SS34 rotor). Nickel affinity purification was performed by FPLC using a 1 ml HisTrap HP column (GE Healthcare) in 15 mM Tris-chloride pH 8, 130 mM Potassium acetate, 10% glycerol, 0.05% Tween 20, and 20 to 400 mM imidazole. Pooled fractions were loaded on a 1 ml HiTrap Q FF column (GE Healthcare). The flow-through was applied onto a Superdex 200 10/300 GL gel filtration column (GE Healthcare) equilibrated in 20 mM Hepes-KOH pH 7.6, 200 mM potassium chloride, 0.2 mM EDTA, 5 mM dithiothreitol. The UV light of the FPLC remained switched off during the entire procedure to protect the Bpa residue. Fractions were pooled and concentrated in Amicon

Ultra-4, 30000 MWCO (Millipore). Aliquots were flash frozen in liquid nitrogen and stored at $-80\ °C$. Two Bpa constructs did not express (W192B, L439B) presumably because of inefficient suppression, protein instability, or misfolding. The yields per liter culture medium varied for the other mutants between 20 μg (W199B) and 1.2 mg (Q190B), yields that fell one to two orders of magnitude below that of the wild-type ISWI construct.

Full-length ISWI was purified as above with the following modifications. Nickel affinity purification was performed at higher ionic strength (300 mM Sodium chloride). The pooled eluate of the nickel column was TEV protease digested overnight at 4 °C. The digest was reapplied to the Nickel matrix. The flow-through was purified over a Mono S 5/50 GL ion exchange column (GE Healthcare) before being injected onto the Superdex 200 column.

UV Crosslinking, SDS-PAGE and Trypsin Digestion—Crosslinking was performed in uncoated 384-well plates (Greiner) on ice with a Blak-Ray C50 365 nm UV light (Ultra-Violet Products, Inc.) in 25 mM Hepes-KOH pH 7.6, 100 mM potassium acetate, 1.5 mM magnesium acetate, 0.1 mM EDTA, 10% glycerol, 10 mM β-mercaptoethanol for 30 to 180 min. The 180 min time point is displayed in Fig. 3*B* and was used for all quantifications and the automated crosslink identification.

Approximately 1 μg of ISWI$^{26-648}$ and 20 μg of the ISWI FL constructs were separated by SDS-PAGE and stained with Coomassie blue. Protein bands were excised and trypsin digested in-gel as described elsewhere with minor modifications (24). The digested samples were resuspended in 10 μl 0.1% formic acid. All experiments were performed at least in duplicates.

Additional gel bands became visible upon UV treatment for the F361B ISWI$^{26-648}$ and the M578B ISWI FL mutants (Fig. 3*B*). For M578B ISWI FL, the two bands indicated by an asterisk were separately analyzed by LC/MS-MS. The analysis showed that the upper band was enriched in the crosslinks XL9, XL13 and XL15 that bridge a long stretch in the primary structure of ISWI explaining the altered gel mobility (supplemental Fig. S8A). No attempts were made to separately analyze the two ISWI-derived bands of the F361B mutant.

MS Analysis and Peptide Quantification—For MS/MS analysis 3 μl of the digest were injected in an Ultimate 3000 HPLC system (LC Packings Dionex). Samples were desalted on-line in a C18 micro column (300 μm i.d. x 5 mm, packed with C18 PepMap™, 5 μm, 100 Å by LC Packings), and peptides were separated with a gradient from 5 to 60% acetonitrile in 0.1% formic acid over 40 min at 300 nl/min on a C18 analytical column (75 μm i.d. $\times$ 15 cm, packed with C18 PepMap™, 3 μm, 100 Å by LC Packings). The effluent from the HPLC was directly electrosprayed into the LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific). The MS instrument was operated in the data-dependent mode to automatically switch between full scan MS and MS/MS acquisition. Survey full scan MS spectra (from *m/z* 300–2000) were acquired in the Orbitrap with resolution $r = 60,000$ at *m/z* 400. The six most intense peptide ions with charge states between 2 and 5 were sequentially isolated to a target value of 10,000 and fragmented in the linear ion trap by collision induced dissociation (CID).

Product ion spectra were recorded in the Orbitrap part of the instrument. For all measurements with the Orbitrap detector, 3 lock-mass ions from ambient air (*m/z* = 371.10123, 445.12002, 519.13882) were used for internal calibration as described (25). Typical mass spectrometric conditions were: spray voltage, 1.4 kV; no sheath and auxiliary gas flow; heated capillary temperature, 200 °C; normalized collision energy, 35% for CID in LTQ. The ion selection threshold was 10,000 counts for MS2. An activation $q = 0.25$ and activation time of 30 ms were used.

Peptides were quantified using the peak area from the corresponding extracted ion chromatograms ($\pm$10 ppm). To avoid differences originating from the amount of material, digestion efficiency and spray fluctuation during the LC-MS/MS analysis, peptides were normalized to the peak area of the eight most intense uncrosslinked ISWI peptides.

Automated Mapping of Crosslinks—MS/MS raw files were preprocessed with Decon2 SL 1.0 to extract individual MS/MS spectra and their precursor ion masses and charge states (26). The newly developed software "Crossfinder," coded in Matlab 7.10, was used to map crosslinks with the following settings: MS1 and MS2 mass accuracy, 10 ppm; enzyme, trypsin; allowed number of missed cleavage sites, four; fixed modifications, carbamidomethylation on cysteine; variable modifications, oxidation on methionine and tryptophan; number of top MS2 peaks considered per 50 Da mass window, twelve; minimum number of assignable MS2 product ions for each of the two peptides in a crosslink, two; minimum crosslinking score, 500; search for H_2-eliminated peptides, disabled. Scores were calculated for b- and y-product ions as previously described (27). The source code of Crossfinder is available to noncommercial users upon request.

Multiple Sequence Alignments—Proteins related to the ATPase region of *Drosophila* ISWI (amino acids 100–640) were identified by PSI-BLAST (www.ncbi.nlm.nih.gov/BLAST) using the Swissprot database and a PSI Blast threshold of 10^{-40}. Duplicate sequences and sequences with only partial sequence homology to the ATPase region of ISWI were removed from the list manually. Homologous regions as identified by the PSI-BLAST algorithm were excised. These stretches were submitted to T-Coffee (http://tcoffee.vital-it.ch) with default parameters for multiple sequence alignment. The C terminus of the ATPase region (amino acids 600–655) was separately aligned as above. The multiple sequence alignments were used to relate the positions crosslinked in ISWI to the corresponding positions in the three crystallized ISWI-related proteins (Chd1, Sso1653 and Rad54; Table II) and for structural modeling. Variation of the start and end amino acids of the entire procedure (amino acids 120 to 600 or 80 to 690) did not affect the results.

Structural Modeling—Modeler 9.10 was used for homology modeling and subsequent structural refinement (28). Models were built by satisfaction of spatial restraints, which Modeler extracted from the Chd1 crystal structure. Results from crosslinking experiments were incorporated into the model as a set of additional spatial restraints between the Cα atoms of the two crosslinked amino acids. The restraints were implemented in the form of upper bounds with a single-sided standard deviation of the potential function of 0.01 Å. The "automodel" class in the default mode was used for homology modeling. Default settings were also used for the structural refinement. A more thorough structural refinement (library_schedule = autosched.slow, max_var_iterations = 300, md_level = refine.slow, repeat_optimization = 2, max_molpdf = 1e6) did not affect results and conclusions.

To derive conservative limits for the spatial restraints provided by the crosslinks, the worst case scenario was always assumed. In this scenario, all amino acid side chains are fully extended, and crosslinks take place at the atom most distal to the Cα atom. The maximal distance spanned by individual amino acids was estimated from crystal structures. Individual maximal distances between Cα atoms and the most distal carbon- or heteroatom atoms were: Arg, 7.4 Å; Glu, 5.0 Å; Gly, 2.4 Å; His, 4.7 Å; Ile, 4.0 Å; Leu, 4.0 Å; Lys, 6.5 Å; Pro, 2.5 Å. The maximal distance spanned by Bpa (9.6 Å) was added to these values. It is given by the distance between the Cα atom of Bpa and the most distal atom from which a new covalent bond can initiate, the *para* C atom of the benzoyl moiety (18). The value was calculated from the sum of the maximal distances that a phenylalanine and a benzoyl moiety can bridge.

For XL3 (Table I) the precise acceptor amino acid of the crosslink was uncertain. To derive the most conservative distance restraint in this case, distance restraints to both possible acceptor amino acids (A631 or K632) were calculated as above, 3.9 Å (the maximal distance

between neighboring Cα atoms) was added per ambiguous residue, and the largest restraint was selected. All distance constraints are summarized in Table II.

The target peptides of XL8 in the ISWI26–648 construct and XL17 in full-length ISWI were identical but only[26–648] for XL8 could we delineate the precise acceptor amino acid (R589). Distance constraints for XL8 and XL17 were derived as above, and structural modeling was repeated using the restraint from either XL8 or XL17. The degree of compaction and rotation of the ATPase lobes and all other conclusions were not affected when the tight distance restraint derived from XL8 (17 Å between amino acids 578 and 589) was substituted with a looser distance restraint derived from XL17 (24.8 Å; data not shown). Pymol and Matlab were used to visualize structures and to measure distances.

Steady State ATP Hydrolysis Assays—ATP hydrolysis was measured in crosslinking buffer (see above) by a coupled ATPase assay using Mg^{2+}-ATP (1 mM), pyruvate kinase (15.5 U/ml), PEP (6 mM), lactate dehydrogenase (15.5 U/ml), NADH (1.2 mM) and 100 to 200 nM full-length ISWI in 30 µl reactions (29). The absorption of NADH over time was read by a Biotek PowerWave HT 384 well plate reader in flat bottom, uncoated 384-well plates (Greiner). Linearized plasmid DNA (pUC18) or chromatin assembled by salt gradient dialysis on linearized pUC18 (30) were used as cofactors. The mass ratio of histones to DNA in the chromatin was 0.23. The concentration of chromatinized DNA was expressed in terms of its DNA content by measuring the UV absorbance at 260 nm.

The purity of the M578B ISWI26–648 mutant proved to be insufficient for ATP turnover measurements. The preparation of the ISWI26–648 M578B appeared to hydrolyze ATP several-fold faster than the wild-type construct in the presence of saturating concentrations of DNA (data not shown). However, an unrelated Bpa control mutant with comparable amounts of impurities present in the preparation appeared to be similarly hyperactive. In addition, the DNA concentrations necessary to elicit half-maximal activities were markedly lower in both mutants. We considered it unlikely that Bpa mutagenesis would result in a stimulated ATPase activity and an increased DNA affinity in both mutants. Rather, the data strongly pointed to the presence of contaminating ATPases in the protein preparations that overwhelmed the signal. Indeed, a number of additional protein bands were visible by SDS-PAGE. We therefore characterized the mutation in the context of full-length ISWI, which we succeeded in purifying to a significantly higher degree (Fig. 3B and supplemental Fig. S9).

RESULTS

Site-specific Incorporation of Bpa Into the ATPase Domain of ISWI and Bpa Photo-crosslinking—We inserted the photo-activatable amino acid Bpa into the ATPase domain of ISWI to obtain structural information from the crosslinks it forms. Guided by available crystal structures of Snf2 family members, six positions along the interface of the two ATPase lobes were chosen (Fig. 2B). We selected six bulky amino acids that are in proximity of the neighboring ATPase lobe. Their codons were individually replaced with TAG stop codons to generate single point mutants of ISWI. Mutagenesis was carried out for all six positions in a construct that spans the entire ATPase region of ISWI from *Drosophila melanogaster*. We refer to it as ISWI[26–648] (Fig. 2B). One mutation (M578B) was also introduced in the full-length ISWI protein. This mutant served as an important control (see below).

The mutagenized ISWI constructs were individually transformed into bacteria carrying a Bpa-specific suppressor tRNA and aminoacyl-tRNA synthetase, which allow incorporation of
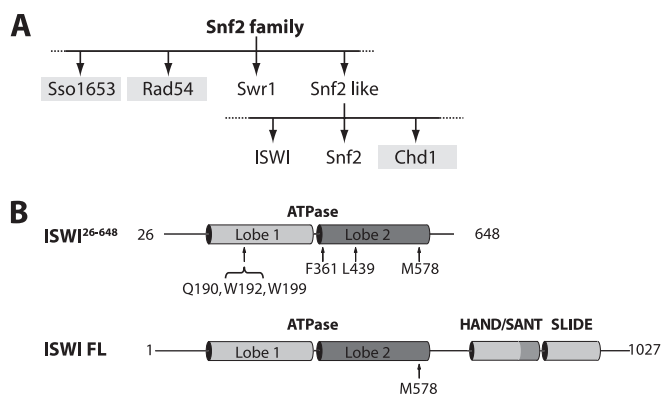
FIG. 2. **Construct design.** *A*, The ATPase domain of ISWI belongs to a diverse family of Snf2-type ATPases(19). Crystal structures of the ATPase regions are available for the three shaded proteins. *B*, The indicated sites within lobes 1 and 2 of the ATPase domain of ISWI were chosen for Bpa mutagenesis. Mutagenesis was carried out in two ISWI construct. One is referred to as ISWI[26–648]. It spans the entire ATPase region but lacks part of the N terminus and the C-terminal HAND/SANT and SLIDE domains. Bpa mutagenesis was also applied to position 578 in the context of the full-length protein. All ISWI[26–648] and ISWI FL constructs contain an N-terminal hexahistidine purification tag followed by a TEV protease site.

Bpa in place of the TAG stop codons (23). Cells were grown in the presence or absence of Bpa in the medium. Protein with similar SDS-PAGE mobility as the wild-type protein was only purified in the presence of Bpa, indicating successful suppression of the stop codon (Fig. 3A, arrow). Subsequent MS analyses confirmed the identity of the protein and incorporation of Bpa at the expected position (see below).

Five out of seven Bpa constructs could be expressed and purified. Crosslinking was induced by irradiation with UV light. The samples along with unirradiated control samples were separated by SDS-PAGE (Fig. 3B). Bands corresponding to ISWI were excised from the gel, trypsin digested and submitted to LC-MS/MS analysis. We employed an Orbitrap XL mass spectrometer to achieve a high mass accuracy for peptide precursor ions and their fragment ions (see Methods), crucial for unambiguous assignment of the masses and mapping of crosslinks with high confidence.

An Automated Workflow for Mapping Crosslinks—To quickly and effortlessly map crosslinks we devised an automated workflow. It is implemented in the newly developed analysis software "Crossfinder" (Fig. 4 and supplemental Figs. S1 to S4). Based on user-specified protein sequence(s) and the protease used, a database was calculated that contained all possible combinations of peptides that could be covalently linked. The masses of these hypothetical crosslinks were calculated, taking fixed or variable modifications of amino acids into consideration. Crosslink candidates were selected by matching experimental precursor masses against the database within a specified mass tolerance (Fig. 4A), then by matching theoretical and experimental peptide fragmentation spectra (Fig. 4B).
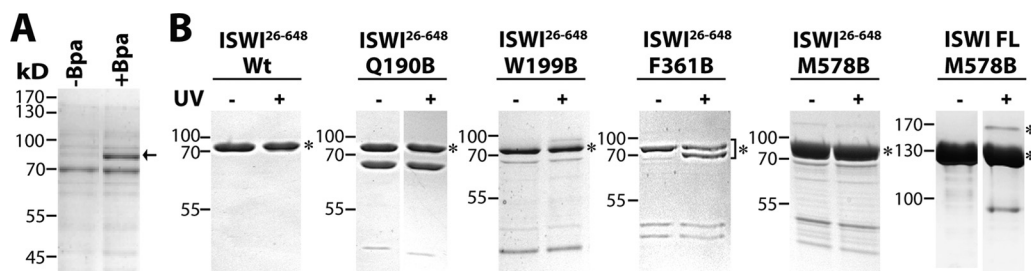
FIG. 3. **Protein purification and UV-induced crosslinking.** *A*, Proof-of-principle experiment to demonstrate successful co-translational incorporation of Bpa. Bacteria carrying the ISWI$^{26-648}$ construct with a TAG stop codon at position 190 were grown in the absence or presence of Bpa in the medium. Cell extracts were incubated with Ni-NTA beads. The beads were washed and boiled in SDS sample buffer. Proteins were separated by SDS-PAGE and stained with Coomassie blue. Protein with the expected mobility could only be enriched from cells grown in the presence of Bpa (arrow). *B*, SDS-PAGE gels of wild-type protein and Bpa mutants before and after UV treatment. ISWI-derived bands were excised from the gel for MS analysis (asterisks). The mobility of the isolated bands was consistent with monomeric but not dimeric ISWI. The crosslinks that we mapped therefore formed intra- and not intermolecularly (Table I). Additional gel bands were observed upon UV irradiation for the ISWI FL M578B and ISWI$^{26-648}$ F361B mutants indicating the accumulation of crosslinked species that have altered mobility during SDS-PAGE.

We applied next a scoring algorithm that calculated the statistical significance of each annotation of a product ion in a fragmentation spectrum as previously described (27). The algorithm takes into account that the probability of a false positive annotation increases with the number of experimental and theoretical product ions to be matched in a given m/z window. Scoring significantly improved the signal-to-noise ratio (compare Figs. 4*B* and 4*C*). In addition, it allowed us in most cases to pick the amino acid that Bpa crosslinked to. Crosslink candidates, in which the Bpa was incorrectly linked to the target peptide, produced fewer matching product ions and thus received lower scores. Lastly, the data were filtered. We required that fragments of both peptides engaged in the crosslink were present in the MS/MS spectrum and that a minimum score was achieved (Fig. 4*D*).

For one Bpa derivative, W199B, the presence of the tryptic peptide ^{199}BCPSLR proved successful incorporation of Bpa into the protein, but no crosslinks were observed for this construct (data not shown). Presumably no amino acid target is in proximity, or nearby amino acids do not fulfill the chemical requirements for crosslinking.

The entire workflow illustrated in Fig. 4 provided an excellent signal-to-noise ratio. No candidate crosslinks were found in the unirradiated sample. UV-treated and untreated wild-type proteins were used to independently assess the noise in the experiment. We searched for the presence of peptide masses in these datasets that would be identified as crosslinks in any of the successfully purified Bpa constructs. Not a single UV-specific peptide was identified by this procedure (supplemental Fig. S4).

Validation of Mapped Crosslinks—In total, 17 crosslinked peptides were detected in three ISWI$^{26-648}$ derivatives and one ISWI FL mutant (XL1 to XL17, Table I). Many linkage sites were identified more than once, significantly increasing the confidence in the correct assignments of the crosslinked peptides. For instance, XL1 and XL2 in the Q190B dataset both show a covalent link between amino acids 190 and 209. They

could be separately detected because trypsin missed a cleavage site in one of the peptides.

Moreover, we found evidence for Bpa-induced elimination of H$_2$ at the target residue of XL1 and XL2, Ile209 (supplemental Fig. S5A and S5C). Whereas the detection of H$_2$ elimination represented independent support for the correct assignments of XL1 and XL2, one should not take linear peptides lacking two hydrogen masses as the only evidence for spatial proximity to a Bpa residue. We sporadically detected peptides (*W*CPSLR and AVC*LIGDQDTR) that were short of 2.016 Da within error at the underlined positions also in the absence of UV and in the wild-type controls (data not shown). We speculate that oxidation (+15.995 Da) followed by water loss (–18.011 Da) may lead to the observed mass difference of these residues.

Reassuringly, all crosslinked peptides detected in the M578B ISWI$^{26-648}$ protein were also present in the M578B full-length protein (Table I; note that XL7 has a different mass than XL12 as the target peptide of XL7 is the most C-terminal peptide of the ISWI$^{26-648}$ protein and the corresponding tryptic peptide from the full-length protein is longer). Thanks to significantly more material used for the analysis, we also discovered several new crosslinked peptides in the full-length dataset (Table I, supplemental Figs. S3 and S8).

All crosslink candidates were validated in two additional ways. First, we made sure that the candidate crosslinks were only present in the UV-irradiated samples by quantifying the amounts of the peptides. Extracted ion chromatograms of the characteristic *m/z* values of the peptides were used for quantification. Indeed, we could detect the crosslink candidates only in the UV-treated sample. At the same time, UV treatment reduced the signals for the uncrosslinked Bpa peptides by >60%, suggesting an efficient crosslinking reaction (Fig. 5*A* and supplemental Figs. S5A, S6A, and S8A). Second, all ions of the MS/MS spectra were manually matched with the product ions calculated by Crossfinder to confirm the identity of the suggested crosslinked peptide and the precise attachment site
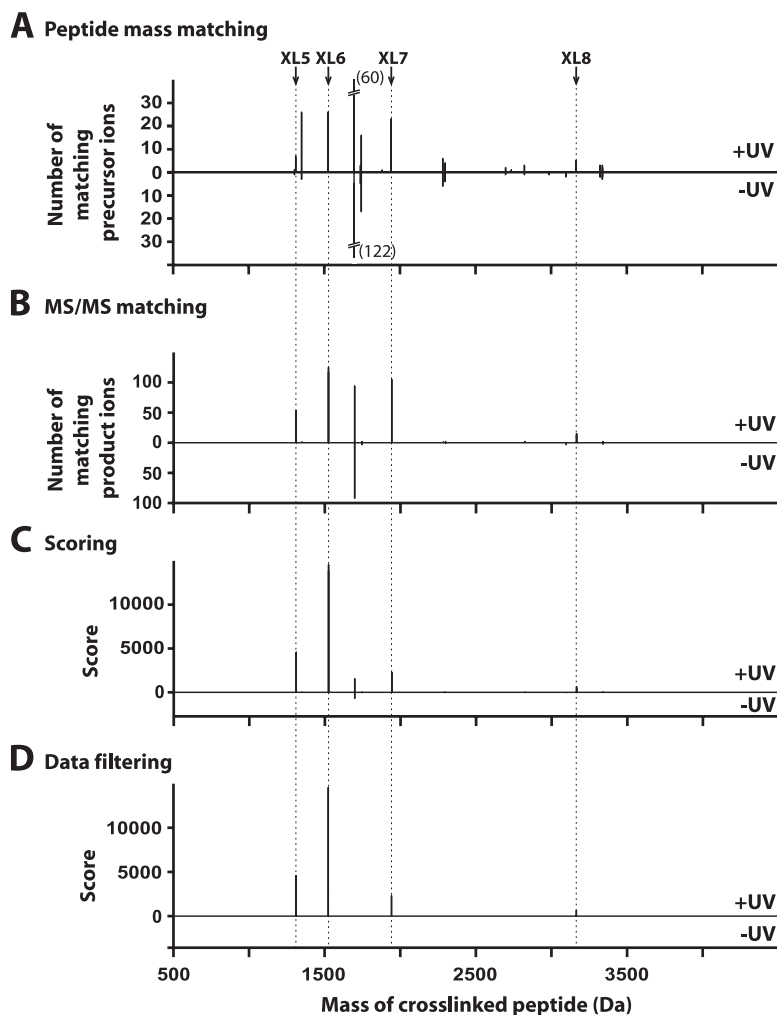
FIG. 4. **Mapping of crosslinks by LC-MS/MS and Crossfinder.** Shown are results obtained from the M578B ISWI[26−648] data set. *A*, Proteins contained in the gel bands from the previous figure were digested with trypsin and submitted to LC-MS/MS analysis. Based on the protein sequence and the protease used, a database of all theoretically possible crosslinks was calculated. The masses of observed precursor ions were matched against this database of crosslink candidates within a given mass tolerance (10 ppm). For each crosslink candidate, the number of matching precursor ions was plotted as a bar above the candidate's mass. Unirradiated protein (-UV) served as a control and is displayed below the X axes in all panels. UV-specific peaks could be observed at masses that correspond to crosslinks XL5–8 (dotted lines; Table I) but numerous other crosslink candidates also passed this step of the algorithm. *B*, Product ions that would be expected to be generated during CID were calculated for each crosslink candidate. They were matched against the observed product ions of the precursors from the previous panel within a tolerance of 10 ppm. The total number of matched product ions was plotted for each crosslink candidate. The number of false positive candidates present in the first panel could be significantly reduced by this step. *C*, A score was calculated for each matched product ion according to Olsen and Mann (27; see also Results section for a brief explanation). Displayed is the sum of these scores for each crosslink candidate. Scoring strongly increased the signal-to-noise ratio. *D*, Results after data filtering. The filter required that the crosslink candidate achieved a minimum score and that at least two product ions were found for each peptide engaged in the crosslink (see also Methods section). Four crosslinks, XL5–8, were found (see Table I for details). No false positive candidates were detected.

(Fig. 5*C*, supplemental Figs. S5 to S8, and data not shown). The manual analysis also validated two different acceptor amino acids predicted by Crossfinder in XL7 and XL12, G644 and G645. The two linkages were named XL7a and b and XL12a and b, respectively (supplemental Fig. S7B and data not shown).

A Novel Conformation of a Snf2-type ATPase Domain—We next inspected the crystal structures of Chd1, Rad54, and Sso1653 for their consistency with the crosslinks we mapped in ISWI. The maximal distance between two Cα atoms in a

crosslink is given by the distance that Bpa and the target amino acids can bridge while being fully extended (see Methods). Crosslinks thus provide upper limits for the distances between two Cα atoms. The target residues for all crosslinks except for XL3 could be unambiguously identified through MS/MS assignments (see above; note that XL8 and XL17 are redundant). The distance constraint derived from XL3 is therefore larger (see Methods) and the spatial resolution it offers correspondingly lower.

TABLE I
Crosslinks within the ATPase domain of ISWI

Construct		Crosslinked peptides					
		Mass (Da)	Error (ppm)	Bpa peptide		Target peptide	
				Sequence[a,b]	Site	Sequence[b]	Site
XL1	ISWI[26–648]	2734.3062	+0.7	STL<u>B</u>NWVNEFK	190	AVCL<u>I</u>GDQDTR	209
XL2	Q190B	2862.4012	+0.3	STL<u>B</u>NWVNEFKK	190	AVCL<u>I</u>GDQDTR	209
XL3	ISWI[26–648]	1826.9779	+0.4	I<u>B</u>VGLSK	361	GE<u>A</u>KTAEQK[c]	631–632[c]
XL4	F361B	2014.0875	+0.7	I<u>B</u>VGLSK	361	<u>L</u>ITESTVEEK	555
XL5		1309.6932	+0.1	<u>B</u>VIQGGR	578	<u>H</u>FK	172
XL6	ISWI[26–648]	1522.8621	+0.1	<u>B</u>VIQGGR	578	<u>L</u>RLDK	573
XL7a	M578B	1940.9845	+3.7	<u>B</u>VIQGGR	578	AALDSL<u>G</u>ESSL	644
XL7b		1940.9845	+3.7	<u>B</u>VIQGGR	578	AALDSLG<u>E</u>SSL	645
XL8		3163.6562	+2.7	<u>B</u>VIQGGR	578	LVDN<u>R</u>SNQLNKDEMLNIIR	589
XL9		1309.6932	+7.1	<u>B</u>VIQGGR	578	<u>H</u>FK	172
XL10		1494.7944	+5.1	<u>B</u>VIQGGR	578	<u>L</u>VDNR	585
XL11		1522.8621	+5.6	<u>B</u>VIQGGR	578	<u>L</u>RLDK	573
XL12a		2097.0856	+6.3	<u>B</u>VIQGGR	578	AALDSL<u>G</u>ESSLR	644
XL12b	ISWI FL[d]	2097.0856	+6.3	<u>B</u>VIQGGR	578	AALDSLG<u>E</u>SSLR	645
XL13	M578B	2246.1737	+8.7	LDK<u>B</u>VIQGGR	578	<u>F</u>DASPAYIK	120
XL14		2453.2916	+9.0	LDK<u>B</u>VIQGGR	578	AALDSL<u>G</u>ESSLR	644
XL15		2640.4793	+9.0	<u>B</u>VIQGGR	578	TLQTISLLGYLK<u>H</u>FK	172
XL16		2722.4768	+5.0	LRLDK<u>B</u>VIQGGR	578	AALDSL<u>G</u>ESSLR	644
XL17		3163.6562	+7.7	<u>B</u>VIQGGR	578	LVDN<u>RS</u>NQLNKDEMLNIIR[c]	589–591[c]

[a] B symbolizes Bpa.
[b] Crosslinked amino acids are underlined.
[c] Precise attachment sites not distinguishable from data.
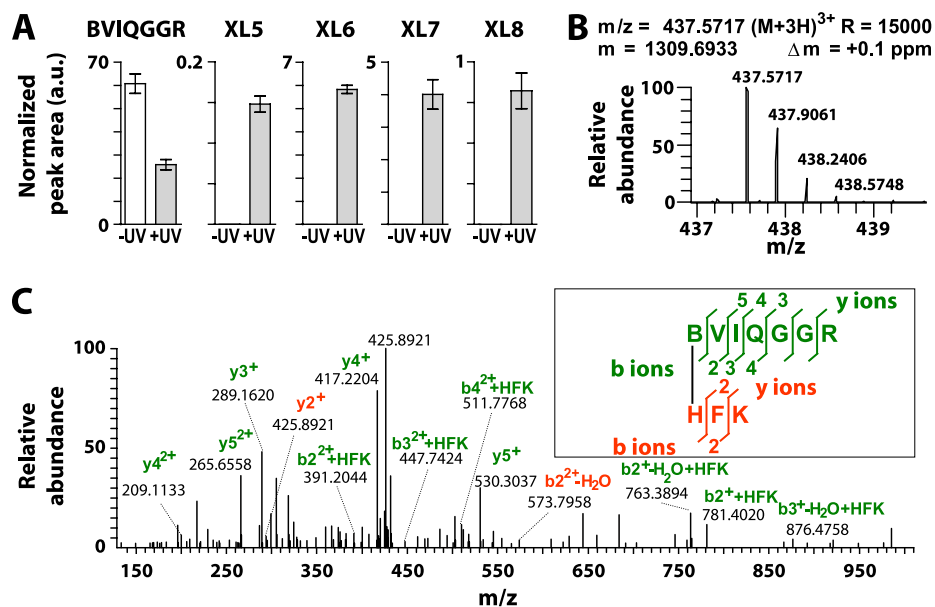[d] Real time calibration by lock masses not available leading to an increased error.

FIG. 5. **Validation of crosslinks.** *A*, UV-dependent decrease of the signal for the Bpa-containing peptide BVIQGGR and concomitant increase of the signals for the crosslinked peptides XL5-XL8 in the M578B ISWI[26–648] data set. Extracted ion chromatograms of the ions corresponding to the peptides of interest were used for the quantification. The signals were normalized against the amount of analyzed protein (see Methods). *n* = 2 and S.D. as error bars. *B*, Mass, charge and measurement error determination of the crosslinked peptide XL5. Displayed is the isotopic distribution of the crosslink candidate, from which the mass-to-charge ratio (*m/z*), the charge (3+) and the monoisotopic mass value (*m*) were derived. Δ*m*: difference between the expected and the measured masses; *R*: resolution of the MS measurement. *C*, High resolution, high accuracy MS2 fragmentation spectrum of the precursor ion shown in *B*). A series of b and y product ions generated by fragmentation were detectable for both peptides involved in the crosslink providing a high confidence in its correct identification. The inset shows the observed product ions mapped onto the sequence of the crosslinked peptide. B symbolizes Bpa.

The distance constraints were compared with the observed distances between homologous positions in the crystal structures. As might be expected, most crosslinks were fully consistent with the crystal structures (Table II). However, a number of crosslinks that formed in the M578B constructs were inconsistent with an ISWI conformation

TABLE II
Cα distances of crosslinked sites in ISWI and its relatives

ISWI				Chd1			Sso1653			Rad54		
Sites		Distance (Å)		Sites		Dist. (Å)	Sites		Dist. (Å)	Sites		Dist. (Å)
1	2	Exp.[a]	Model[b]	1	2	3MWY[c]	1	2	1Z6A[c]	1	2	1Z3I[c]
190	209	≤13.6	10.6	438	457	10.4	505	524	13.1	218	238	9.0
361	632	≤20.0	11.5	615	901	11.9	669	n/a[g]	n/a	407	n/a[g]	n/a
361	555	≤13.6	6.0	615	819	6.0	669	855	5.8	407	615	6.3
578	120	≤12.1	12.1	842	368	≥27.9[d,e]	878	437	**35.8[e]**	638	140	**37.4[e]**
578	172	≤14.3	13.2	842	420	≥29.0[d,e]	878	488	**42.7[e]**	638	196	**37.9[e]**
578	573	≤13.6	13.6	842	837	≥5.5[d]	878	873	8.0	638	633	6.8
578	585	≤13.6	13.3	842	849	n/a[f]	878	885	**14.8[e]**	638	645	**16.2[e]**
578	589	≤17.0	16.8	842	858	≥13.6[d]	878	888	**18.2[e]**	638	648	16.0
578	644	≤12.0	12.0	842	914	≥21.5[d,e]	878	n/a[g]	n/a	638	n/a[g]	n/a
578	645	≤14.6	14.3	842	915	≥24.6[d,e]	878	n/a[g]	n/a	638	n/a[g]	n/a

[a] Upper limits derived from crosslinking experiments. See Methods.
[b] Distances from model in Fig. 6B.
[c] PDB codes.
[d] A842 is missing from the structure precluding exact measurements. Lower limits were calculated from the distances to Y841 minus the maximal distance between neighboring Cα atoms (3.9Å).
[e] Bold face values exceed the distance limits measured for ISWI.
[f] Not available. Target amino acid is not present in the crystal structure.
[g] Not available. No sequence homology exists between ISWI and Sso1653/Rad54 in that region.

that resembled the published crystal structures of its relatives (bold face values).

In principle, the observed deviation from the crystal structure could be explained by a structural perturbation present in our constructs. First, truncating the protein may allow the ATPase domain to assume an artificial conformation. ATP hydrolysis parameters of the wild-type ISWI[26−648] construct, however, were within a factor of two unperturbed relative to full-length ISWI (unpublished results), providing strong evidence against this possibility. Moreover, we observed all crosslinks from the truncated M578B mutant also in full-length ISWI, invalidating this concern.

Alternatively, substituting M578 with Bpa may induce a structural change. To test this possibility, we performed ATP hydrolysis assays comparing the wild-type and M578B mutant proteins. As the ATP hydrolysis signal by the ISWI[26−648] M578B preparation was overwhelmed by contaminating foreign ATPases (see Methods), we characterized the effect of the mutation using the significantly purer full-length mutant (Fig. 3B). The supplemental Fig. S9 shows comparable levels of ATP hydrolysis of the wild-type and mutant proteins in the absence and presence of DNA. As an additional, more stringent test we probed ATP hydrolysis in the presence of chromatin, the natural substrate of ISWI. Both enzymes were strongly stimulated by chromatin above the levels of DNA, and ATP turnover rates in the mutant differed by less than a factor of two. The mutation therefore has no major impact on the stimulatory effects of DNA and chromatin and on ATP turnover, suggesting that the structural integrity of the M578B enzyme is not significantly affected.
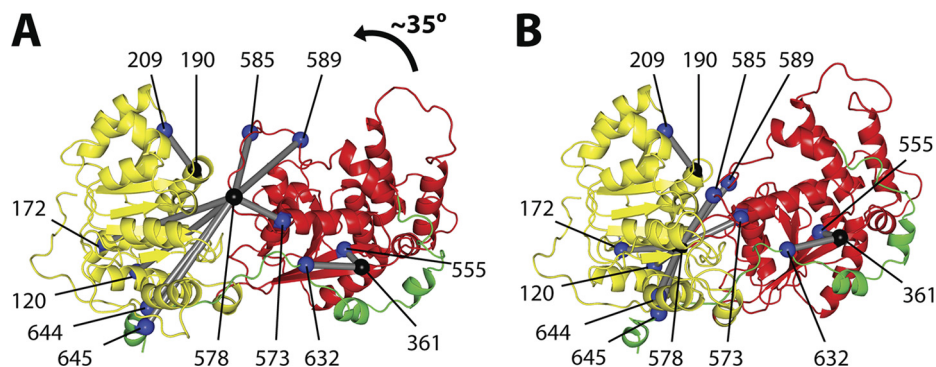
A Structural Model of the ATPase Region of ISWI in Solution—To visualize the ISWI conformation in solution we turned to experimentally constrained structural modeling. Chd1 is the closest ISWI relative whose ATPase domain has been crystallized. Its sequence similarity also extends to the region C-terminal to the second ATPase lobe, which is not the case for Sso1653 and Rad54 (21). Chd1 was crystallized in a relatively open conformation with a large cleft between the two ATPase lobes. Conserved motifs located in both lobes are believed to directly contact the ATP substrate and catalyze its hydrolysis. The large cleft between the lobes, however, would not allow such a configuration. The authors therefore hypothesized that the ATPase lobes must close the cleft to attain a catalytically active conformation. They suggest that the lobes do so by rotating 52° with respect to each other (21). By visualizing the corresponding crosslinked sites in the Chd1 crystal structure, we noted that the crosslinks XL9, 12a, 12b, and 13 could be better satisfied by precisely such a rotation of the lobes (data not shown). Structural modeling of the ISWI ATPase region was used to test that idea.

Homology models of the ISWI ATPase domain were constructed using "Modeler" (see Methods). During modeling, spatial restraints derived from the crosslinks were either not considered (Fig. 6A) or were enforced (Fig. 6B). As expected, all distance constraints were entirely satisfied in the restrained model (Table II). A comparison of the two models demonstrated that the restraints induced a significant compaction of the structure. The lobes were rotated by ∼35° against each other. This rotation closed the cleft between the lobes as postulated from the analysis of the Chd1 crystal structure, albeit to a lower degree. However, the distance restraints were very conservatively derived, so the true magnitude of the rotation was likely underestimated. Our crosslinking data would be fully consistent with an even closer packing of both lobes.

During modeling, poorly restrained parts of the protein would adopt different conformations if modeling were to be

FIG. 6. **Structural model.** The ATPase region of ISWI was modeled onto the homologous structure of Chd1 neglecting distance constraints from crosslinking experiments (*A*) or enforcing them (*B*). Both structures were aligned on lobe 1 (yellow). Enforcing the distance constraints during modeling led to a closer packing of the two ATPase lobes via a rotation of lobe 2 (red) by ~35°. The C-terminal extension of lobe 2 that according to the Chd1 structure bridges both lobes is shown in green. Bpa positions are indicated by black spheres, target amino acids of the crosslinks by blue spheres, and crosslinks by rods. The distance restraints used for modeling are given in Table II.

repeated. To demonstrate that the data robustly restrained the overall orientation of the lobes, modeling was repeated several times. All modeled conformations featured the compaction of the structure and the rotation of the two lobes against each other suggesting that the crosslinking and crystallization data sufficiently restrained the orientation of the lobes (supplemental Fig. S10).

Besides the closure of the cleft between the lobes, the structural model predicts that a loop harboring M578 invades a cavity in lobe 1 to fulfill crosslinks XL9, 12a, 12b, and 13. A comparison of the Chd1 crystal structure and the model in Fig. 6B showed that no significant structural rearrangements in lobe 1 were necessary for this conformation (supplemental Fig. S11). Additional research will be needed to validate this proposed configuration.

DISCUSSION

We developed a novel procedure that uses crosslinks formed by genetically encoded unnatural amino acids to extract structural information for the purpose of testing and refining structural models. We applied the method to probe the solution structure of the ATPase domain of ISWI. ISWI is a Snf2-type chromatin remodeling enzyme and among the best-studied remodelers (31, 32). Its ATPase domain powers the remodeling reaction. Yet, despite much effort in the field no molecular structure of the ISWI ATPase domain is available, significantly hampering our mechanistic insight into the remodeling reaction.

Even within the extensive Snf2 family of proteins (19), the structure of ATPase regions of only three enzymes have been solved (20–22). The ATPase region contains two apparently flexibly joined domains, which were crystallized in radically different conformations. It remains unclear if these conformations are also preferred in solution and if ISWI assumes similar conformations.

To probe the solution structure of the ATPase domain of ISWI, the unnatural photoreactive amino acid Bpa was incor-

porated at strategic positions (17, 23). Crosslinking was induced by irradiation with long-wavelength UV light (365 nm). Nucleotides, DNA and proteins do not absorb in the near UV region and thus are compatible with this technique.

Crosslinks were mapped by high resolution and accuracy MS. Initially, we pursued a manual strategy to map the crosslinks in the MS data sets. This approach proved to be slow and cumbersome and thus impractical to be done on a larger scale (analysis not shown). Instead we designed a dedicated software tool that mines the MS data sets in an unbiased and comprehensive fashion. With a minimum of user-specified parameters, our software tool "Crossfinder" automatically identifies crosslinks and thus streamlines the limiting step in the process. By this bioinformatic approach we confirmed the crosslinked species we found manually and identified several additional ones.

The entire procedure provides an exceptional signal-to-noise ratio. Moreover, the possibility of false discoveries can be rigorously examined by three types of controls for each sample: unirradiated Bpa mutants, and irradiated and unirradiated wild-type protein. The latter two control for possible UV-induced artifacts and, as additional and independent negative controls, they safeguard against experimental noise and the stochastic nature of MS detection.

Comparing the distance information obtained from the crosslinks with atomic distances in related crystal structures showed that none of the available structures satisfied all distance constraints. ISWI evidently can assume a previously undescribed conformation in solution. Our model for the solution structure of ISWI is based on the experimental distance constraints and the crystal structure of Chd1, the closest ISWI relative with known structure. The two ATPase lobes were found to be strongly rotated with respect to each other when distance restraints were applied. Remarkably, such a swivelling motion was previously predicted to be necessary for Chd1 to become catalytically active (21). The modeled ISWI conformation may therefore lie closer to the catalytically com-

petent state. If this state predominates in solution or is a short lived catalytic intermediate remains an open question.

Many crosslinking applications, for instance mapping of subunit interfaces or binding pockets for ligands, rely on the resolution the crosslinks provide and thus tremendously benefit from pinpointing the precise linkage sites (33). This becomes especially important when the crosslinking information is to be used for docking studies (34) or structural modeling (see above). In the past, limited proteolysis (35), mutagenesis (36) or elaborate, but unautomated MS analysis (33, 37) have been employed to determine the linkage sites. Our approach in contrast provided the precise attachment points of most crosslinks effortlessly, which allowed us to derive distance constraints that were sufficiently small to be useful for structural modeling. Furthermore, we did not rely on the often used assumption that benzophenones would selectively react with methionine side chains (38).

Though generating Bpa-substituted proteins requires time at the outset of the project, crosslinking with site-specifically incorporated crosslinkers offers several key advantages over crosslinking with bivalent chemical compounds. For example, the outstanding signal-to-noise ratio and absence of false positive assignments are direct consequences of the smaller number of crosslinking possibilities that have to be considered. With a single crosslinkable residue incorporated in the protein chain, the number of theoretically possible crosslinks grows only linearly, not quadratically with protein size. Side reactions that occur during chemical crosslinking (*e.g.* dead-end crosslinks, attachment of multiple crosslinking molecules per peptide, attachment to unconsidered residues because of limited selectivity of the compound (39)) further add to the number of false positive identifications. Additionally, mutagenesis can position the unnatural amino acid throughout the protein and can specifically target an area of interest, while chemical compounds may not have access to certain areas, *e.g.* the protein's interior, tight interaction surfaces, or regions that lack the reactive groups attacked by the crosslinker. Moreover, most chemical crosslinkers are relatively long, lowering the resolution of that technique.

A distinctive advantage of having only one crosslinkable residue per protein is that the protein structure can be probed under single-hit conditions. At most one crosslink is formed per protein molecule. In contrast, many crosslinks are formed per molecule using chemical crosslinkers, bearing the risk of distorting the native structure. Also, a number of functional and structural assays are available to characterize the effects of a point mutation on the protein structure. It would be exceptionally difficult on the other hand to isolate the structural or functional effects of a particular crosslink among all other crosslinks formed by a chemical compound. Lastly, chemical crosslinkers may bind and thus disrupt the hydrophobic core of proteins because of their hydrophobicity (1).

The approach we employed can be applied to proteins expressed in bacteria, yeast and mammalian cells (23). Other photoactivatable amino acids carrying aryl azides and diazirin derivatives can be genetically encoded, and their use may provide even tighter distance constraints (14, 23). As photo-crosslinking with genetically encoded unnatural amino acids can be performed also *in vivo* (12–14), the protein structure can be probed in its natural environment. Conformational changes may be discovered by comparing crosslinking yields in different functional states of the protein. The clarity of the data, the automated mapping process, and the modest protein amounts it demands make the method suitable for high-throughput applications.

¶ To whom correspondence should be addressed: Adolf-Butenandt Institute, Ludwig-Maximilians University, Schillerstr. 44, 80336 Munich, Germany. Tel.: ++49 89 218075431; Fax: ++49 89 218075425; E-mail: Felix.Mueller-Planitz@med.uni-muenchen.de.
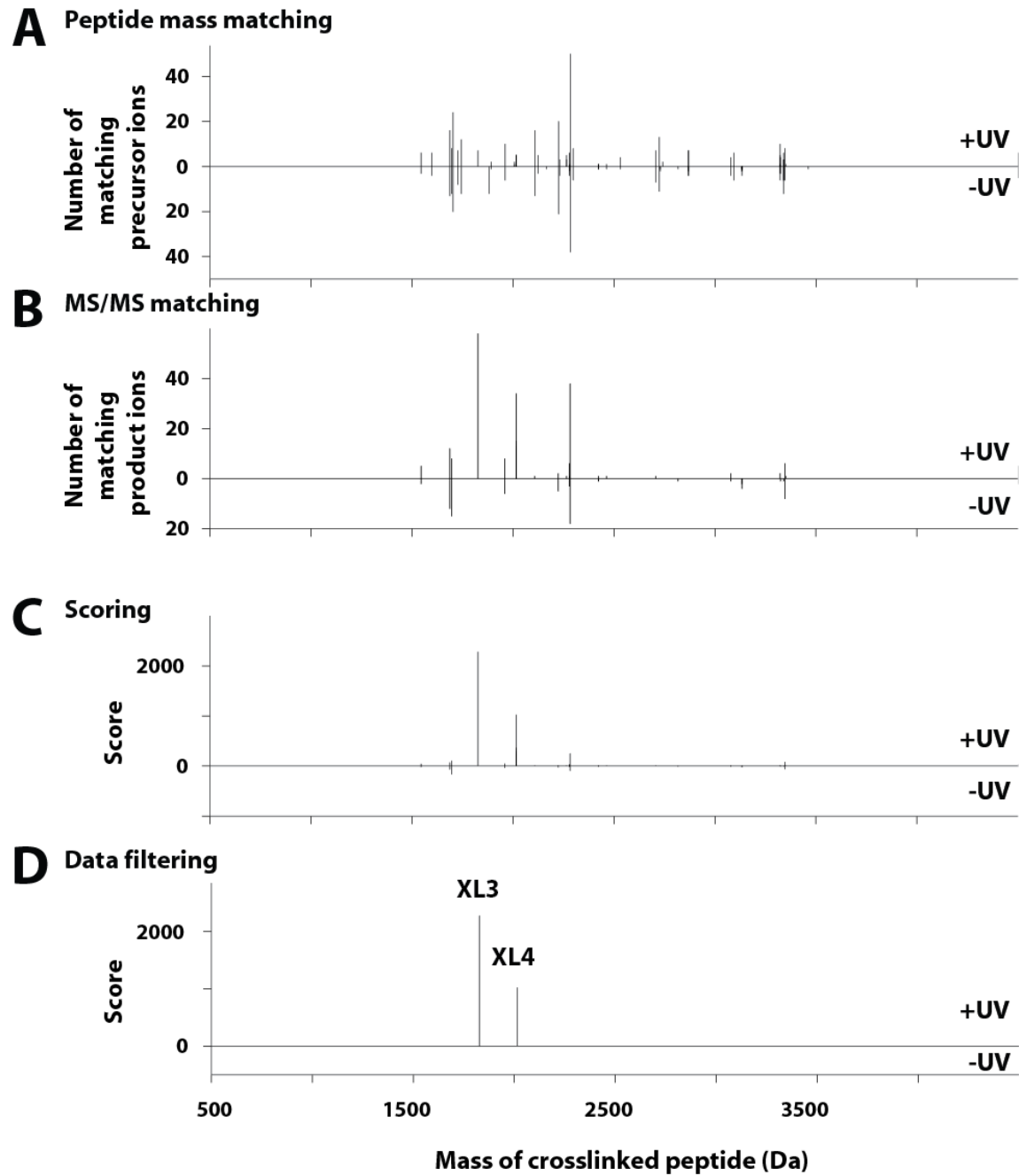
REFERENCES

1. Fasold, H., Klappenb. J, Meyer, C., and Remold, H. (1971) Bifunctional Reagents for Crosslinking of Proteins. *Angew Chem. Int. Edit.* **10,** 795–801
2. Lutter, L. C., and Kurland, C. G. (1975) Chemical Determination of Protein Neighborhoods in a Cellular Organelle. *Mol. Cell. Biochem.* **7,** 105–116
3. Leitner, A., Walzthoeni, T., Kahraman, A., Herzog, F., Rinner, O., Beck, M., and Aebersold, R. (2010) Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell. Proteomics* **9,** 1634–1649
4. Rappsilber, J. (2011) The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* **173,** 530–540
5. Singh, P., Panchaud, A., and Goodlett, D. R. (2010) Chemical cross-linking and mass spectrometry as a low-resolution protein structure determination technique. *Anal. Chem.* **82,** 2636–2642
6. Young, M. M., Tang, N., Hempel, J. C., Oshiro, C. M., Taylor, E. W., Kuntz, I. D., Gibson, B. W., and Dollinger, G. (2000) High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **97,** 5802–5806
7. Kalkhof, S., Haehn, S., Paulsson, M., Smyth, N., Meiler, J., and Sinz, A. (2010) Computational modeling of laminin N-terminal domains using sparse distance constraints from disulfide bonds and chemical cross-linking. *Proteins* **78,** 3409–3427
8. Müller, M. Q., de Koning, L. J., Schmidt, A., Ihling, C., Syha, Y., Rau, O., Mechtler, K., Schubert-Zsilavecz, M., and Sinz, A. (2009) An Innovative Method To Study Target Protein-Drug Interactions by Mass Spectrometry. *J. Med. Chem.* **52,** 2875–2879
9. Chen, Z. A., Jawhari, A., Fischer, L., Buchen, C., Tahir, S., Kamenski, T., Rasmussen, M., Lariviere, L., Bukowski-Wills, J. C., Nilges, M., Cramer, P., and Rappsilber, J. (2010) Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* **29,** 717–726
10. Chu, F., Shan, S. O., Moustakas, D. T., Alber, F., Egea, P. F., Stroud, R. M., Walter, P., and Burlingame, A. L. (2004) Unraveling the interface of signal recognition particle and its receptor by using chemical cross-linking and tandem mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 16454–16459
11. Robinette, D., Neamati, N., Tomer, K. B., and Borchers, C. H. (2006) Photoaffinity labeling combined with mass spectrometric approaches as a tool for structural proteomics. *Expert Rev. Proteomics* **3,** 399–408

12. Chin, J. W., and Schultz, P. G. (2002) In vivo photocrosslinking with unnatural amino acid mutagenesis. *Chembiochem.* **3,** 1135–1137

13. Hino, N., Okazaki, Y., Kobayashi, T., Hayashi, A., Sakamoto, K., and Yokoyama, S. (2005) Protein photo-cross-linking in mammalian cells by site-specific incorporation of a photoreactive amino acid. *Nat. Methods* **2,** 201–206

14. Sinz, A. (2010) Investigation of protein-protein interactions in living cells by chemical crosslinking and mass spectrometry. *Anal. Bioanal. Chem.* **397,** 3433–3440

15. Tanaka, Y., and Kohler, J. J. (2008) Photoactivatable crosslinking sugars for capturing glycoprotein interactions. *J. Am. Chem. Soc.* **130,** 3278–3279

16. Nittis, T., Guittat, L., LeDuc, R. D., Dao, B., Duxin, J. P., Rohrs, H., Townsend, R. R., and Stewart, S. A. (2010) Revealing novel telomere proteins using in vivo cross-linking, tandem affinity purification, and label-free quantitative LC-FTICR-MS. *Mol. Cell. Proteomics* **9,** 1144–1156

17. Chin, J. W., Martin, A. B., King, D. S., Wang, L., and Schultz, P. G. (2002) Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 11020–11024

18. Dormán, G., and Prestwich, G. D. (1994) Benzophenone photophores in biochemistry. *Biochemistry* **33,** 5661–5673

19. Flaus, A., Martin, D. M., Barton, G. J., and Owen-Hughes, T. (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res.* **34,** 2887–2905

20. Dürr, H., Körner, C., Müller, M., Hickmann, V., and Hopfner, K. P. (2005) X-ray structures of the Sulfolobus solfataricus SWI2/SNF2 ATPase core and its complex with DNA. *Cell* **121,** 363–373

21. Hauk, G., McKnight, J. N., Nodelman, I. M., and Bowman, G. D. (2010) The chromodomains of the Chd1 chromatin remodeler regulate DNA access to the ATPase motor. *Mol. Cell* **39,** 711–723

22. Thomä, N. H., Czyzewski, B. K., Alexeev, A. A., Mazin, A. V., Kowalczykowski, S. C., and Pavletich, N. P. (2005) Structure of the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54. *Nat. Struct. Mol. Biol.* **12,** 350–356

23. Liu, C. C., and Schultz, P. G. (2010) Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* **79,** 413–444

24. Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996) Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* **379,** 466–469

25. Olsen, J. V., de Godoy, L. M., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005) Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **4,** 2010–2021

26. Zimmer, J. S., Monroe, M. E., Qian, W. J., and Smith, R. D. (2006) Advances in proteomics data analysis and display using an accurate mass and time tag approach. *Mass Spectrom. Rev.* **25,** 450–482

27. Olsen, J. V., and Mann, M. (2004) Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 13417–13422

28. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234,** 779–815

29. Mueller-Planitz, F., and Herschlag, D. (2006) Interdomain communication in DNA topoisomerase II. DNA binding and enzyme activation. *J. Biol. Chem.* **281,** 23395–23404

30. Luger, K., Rechsteiner, T. J., and Richmond, T. J. (1999) Expression and purification of recombinant histones and nucleosome reconstitution. *Methods Mol. Biol.* **119,** 1–16

31. Clapier, C. R., and Cairns, B. R. (2009) The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* **78,** 273–304

32. Yadon, A. N., and Tsukiyama, T. (2011) SnapShot: Chromatin remodeling: ISWI. *Cell* **144,** 453–453 e451

33. Davidson, W., McGibbon, G. A., White, P. W., Yoakim, C., Hopkins, J. L., Guse, I., Hambly, D. M., Frego, L., Ogilvie, W. W., Lavallée, P., and Archambault, J. (2004) Characterization of the binding site for inhibitors of the HPV11 E1-E2 protein interaction on the E2 Transactivation domain by photoaffinity labeling and mass spectrometry. *Anal. Chem.* **76,** 2095–2102

34. Al-Mawsawi, L. Q., Fikkert, V., Dayam, R., Witvrouw, M., Burke, T. R., Borchers, C. H., and Neamati, N. (2006) Discovery of a small-molecule HIV-1 integrase inhibitor-binding site. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 10080–10085

35. Clérico, E. M., Szymanska, A., and Gierasch, L. M. (2009) Exploring the interactions between signal sequences and E. coli SRP by two distinct and complementary crosslinking methods. *Biopolymers* **92,** 201–211

36. Bisello, A., Adams, A. E., Mierke, D. F., Pellegrini, M., Rosenblatt, M., Suva, L. J., and Chorev, M. (1998) Parathyroid hormone-receptor interactions identified directly by photocross-linking and molecular modeling studies. *J. Biol. Chem.* **273,** 22498–22505

37. Krauth, F., Ihling, C. H., Rüttinger, H. H., and Sinz, A. (2009) Heterobifunctional isotope-labeled amine-reactive photo-cross-linker for structural investigation of proteins by matrix-assisted laser desorption/ionization tandem time-of-flight and electrospray ionization LTQ-Orbitrap mass spectrometry. *Rapid Commun. Mass Spectrom.* **23,** 2811–2818

38. Wittelsberger, A., Thomas, B. E., Mierke, D. F., and Rosenblatt, M. (2006) Methionine acts as a "magnet" in photoaffinity crosslinking experiments. *FEBS Lett.* **580,** 1872–1876

39. Kalkhof, S., and Sinz, A. (2008) Chances and pitfalls of chemical crosslinking with amine-reactive N-hydroxysuccinimide esters. *Anal. Bioanal. Chem.* **392,** 305–312
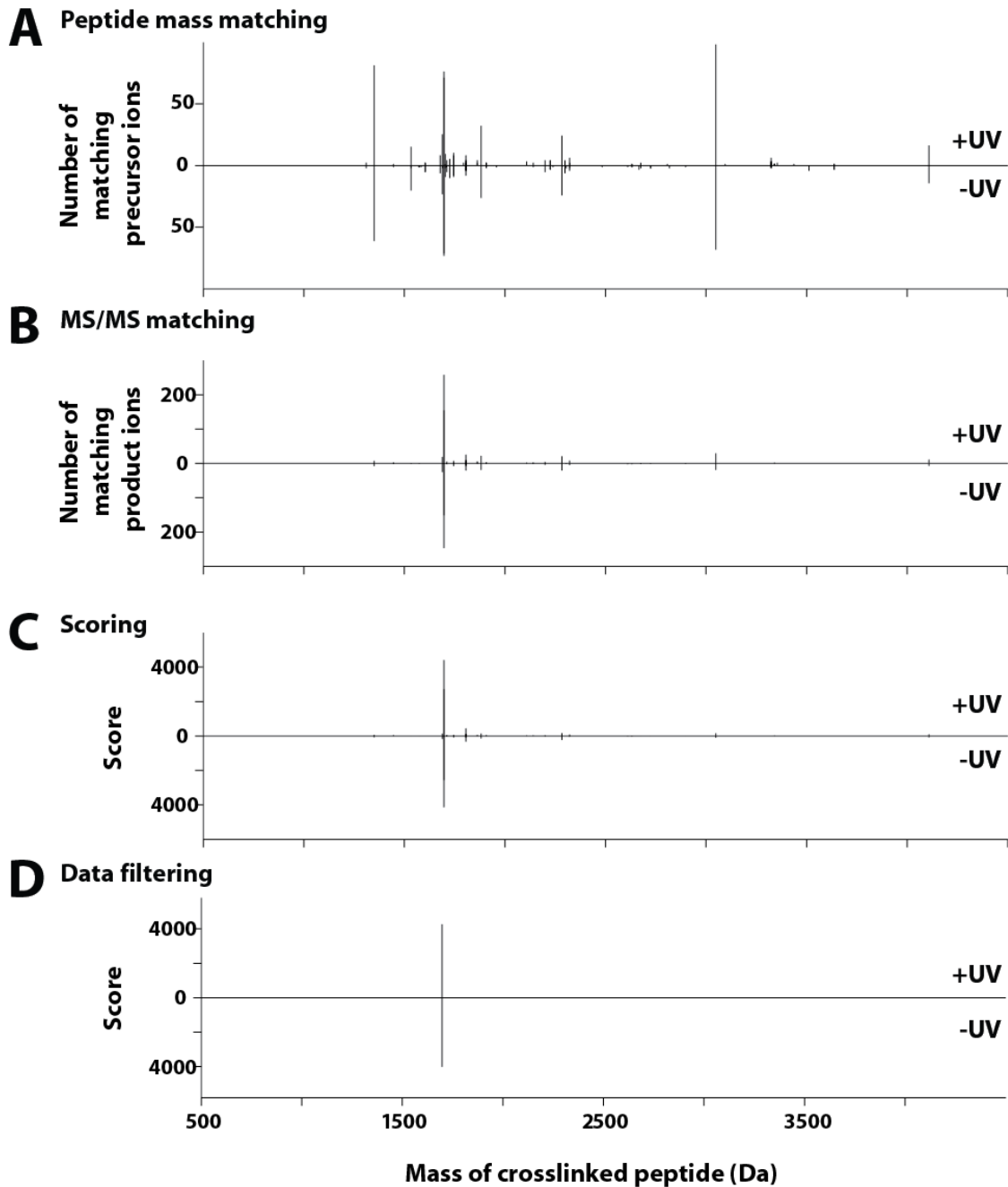
A Peptide mass matching

B MS/MS matching

C Scoring

D Data filtering
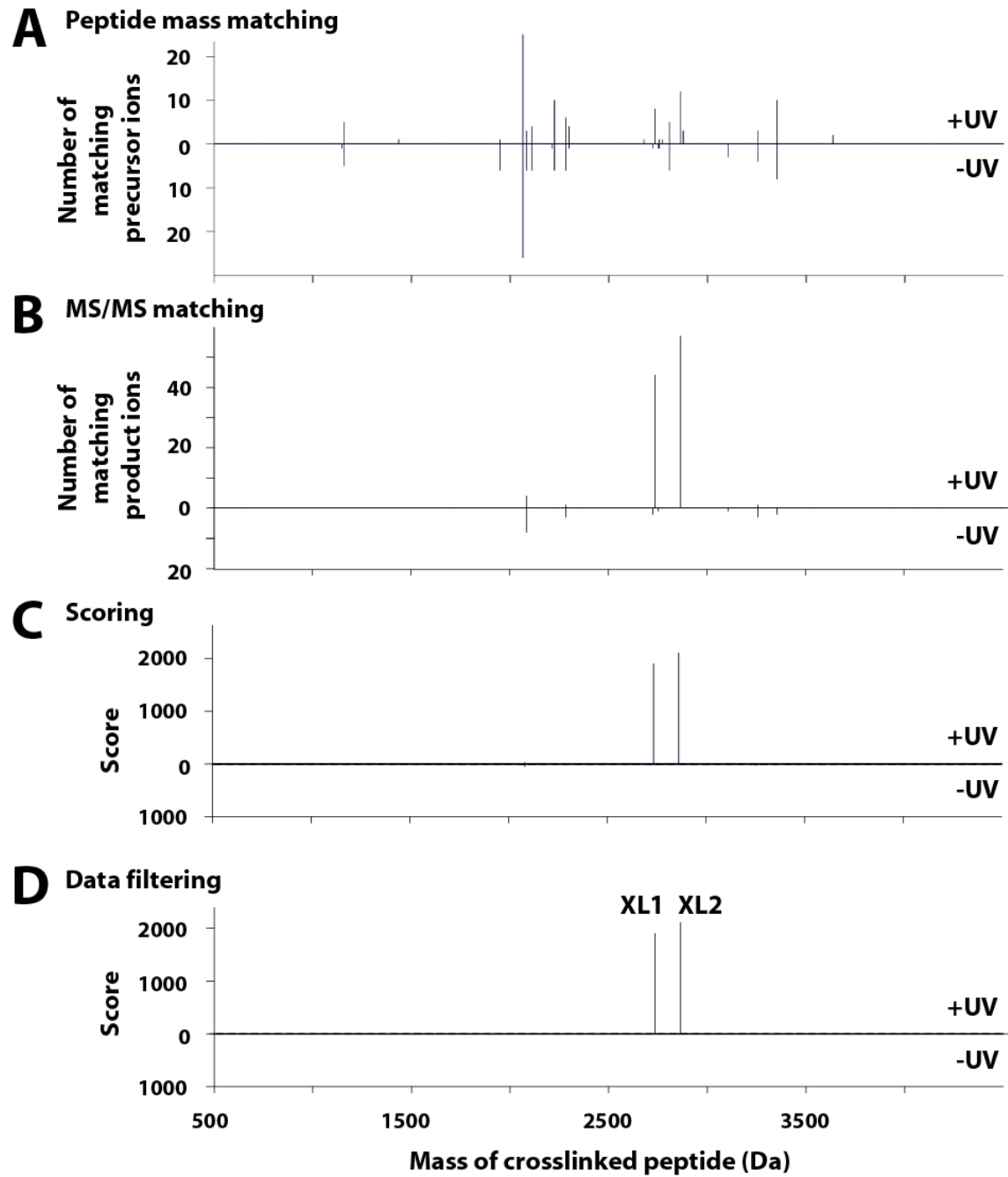
Supplementary Figure 1: Automated mapping of UV-induced crosslinks in the Q190B ISWI[26-648] dataset. For description of the panels, see Fig. 4. Two candidate crosslinks, XL1 and XL2, were identified. See also Table 1.

A Peptide mass matching

B MS/MS matching

C Scoring

D Data filtering

Mass of crosslinked peptide (Da)

Supplementary Fig. 2: Automated mapping of UV-induced crosslinks in the F361B ISWI[26-648] dataset. For description of the panels, see Fig. 4. Two candidate crosslinks, XL3 and XL4, were identified. See also Table 1.
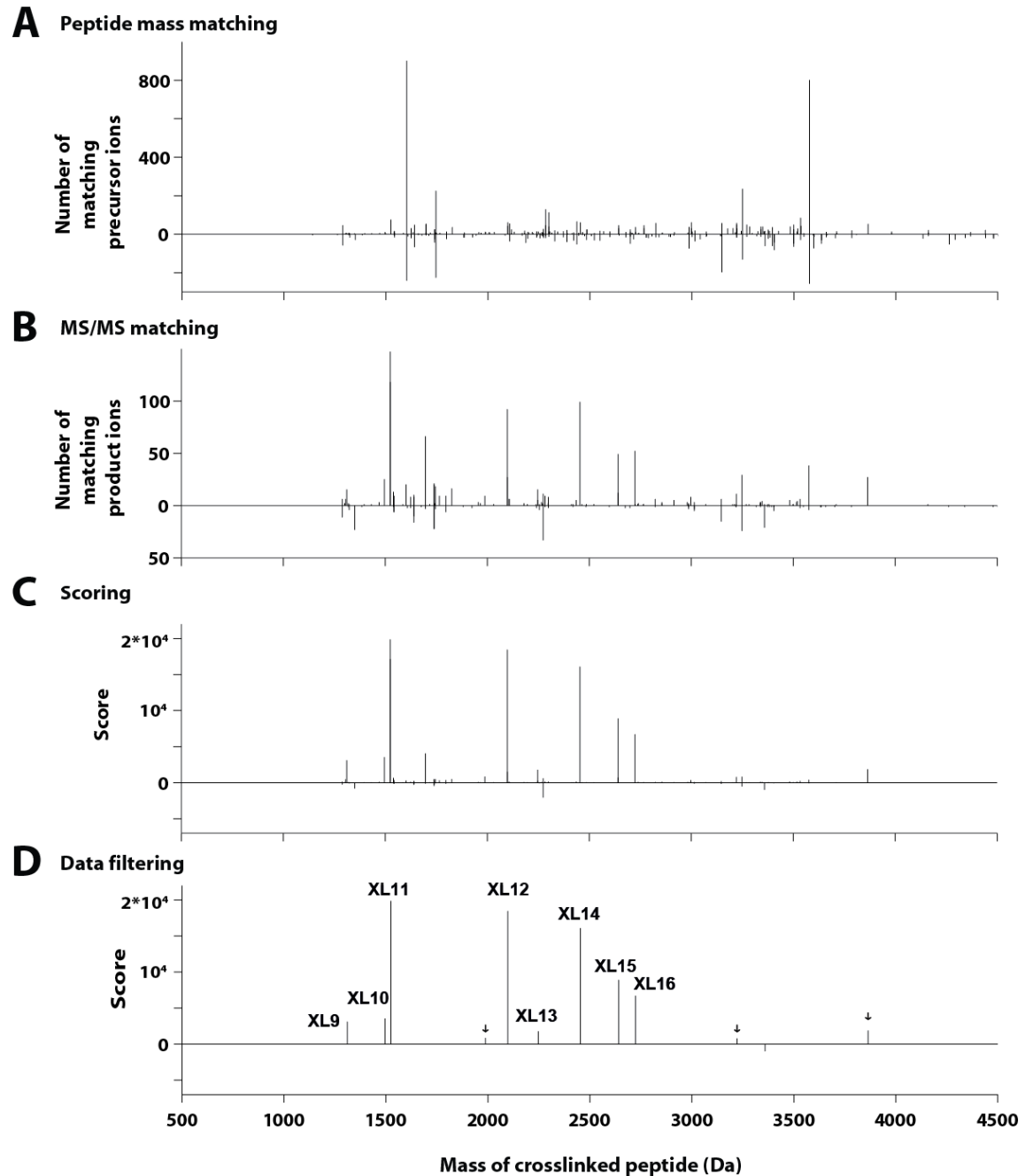
Supplementary Figure 3: Automated mapping of UV-induced crosslinks in the M578B ISWI FL dataset. For description of the panels, see Fig. 4. The crosslinks XL9 to XL16 were identified (Table 1). Three additional crosslinks were mapped (arrows). They map to regions outside of the ATPase domain and were therefore not further considered for structural modeling purposes. The precursor ion with the expected m/z and charge for XL17 was present in the dataset (Suppl. Fig. 8A), but it was not selected for fragmentation by the MS instrument. XL17 was therefore not present in the Crossfinder analysis. A dedicated MS method to fragment this ion, however, confirmed the presence of XL17 in the dataset (data not shown).

A Peptide mass matching

B MS/MS matching

C Scoring

D Data filtering
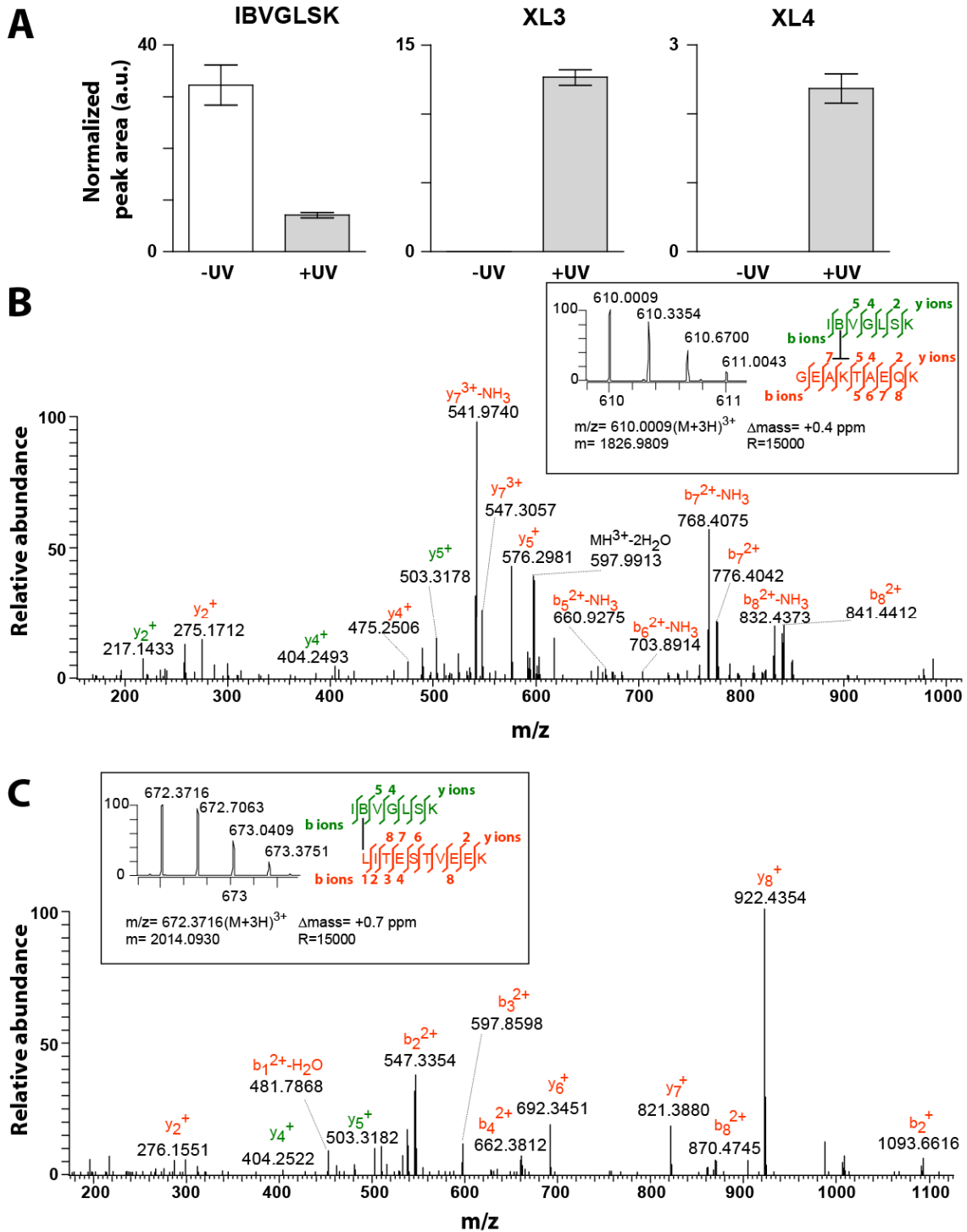
Mass of crosslinked peptide (Da)

Supplementary Figure 4: No false positive crosslinks were identified in the wild-type control. Crosslinks were attempted to be identified in UV-treated and untreated wild-type ISWI[26-648] protein. The datasets were searched against all theoretically possible crosslinks in all mutants that were successfully purified (Q190B, W199B, F361B, and M578B; see Fig. 3B). Only one candidate peptide was found by the mapping procedure (D). It was however identified with equal score in the presence and absence of UV, and the corresponding crosslink candidate can be discarded on these grounds. For description of the panels, see Fig. 4.
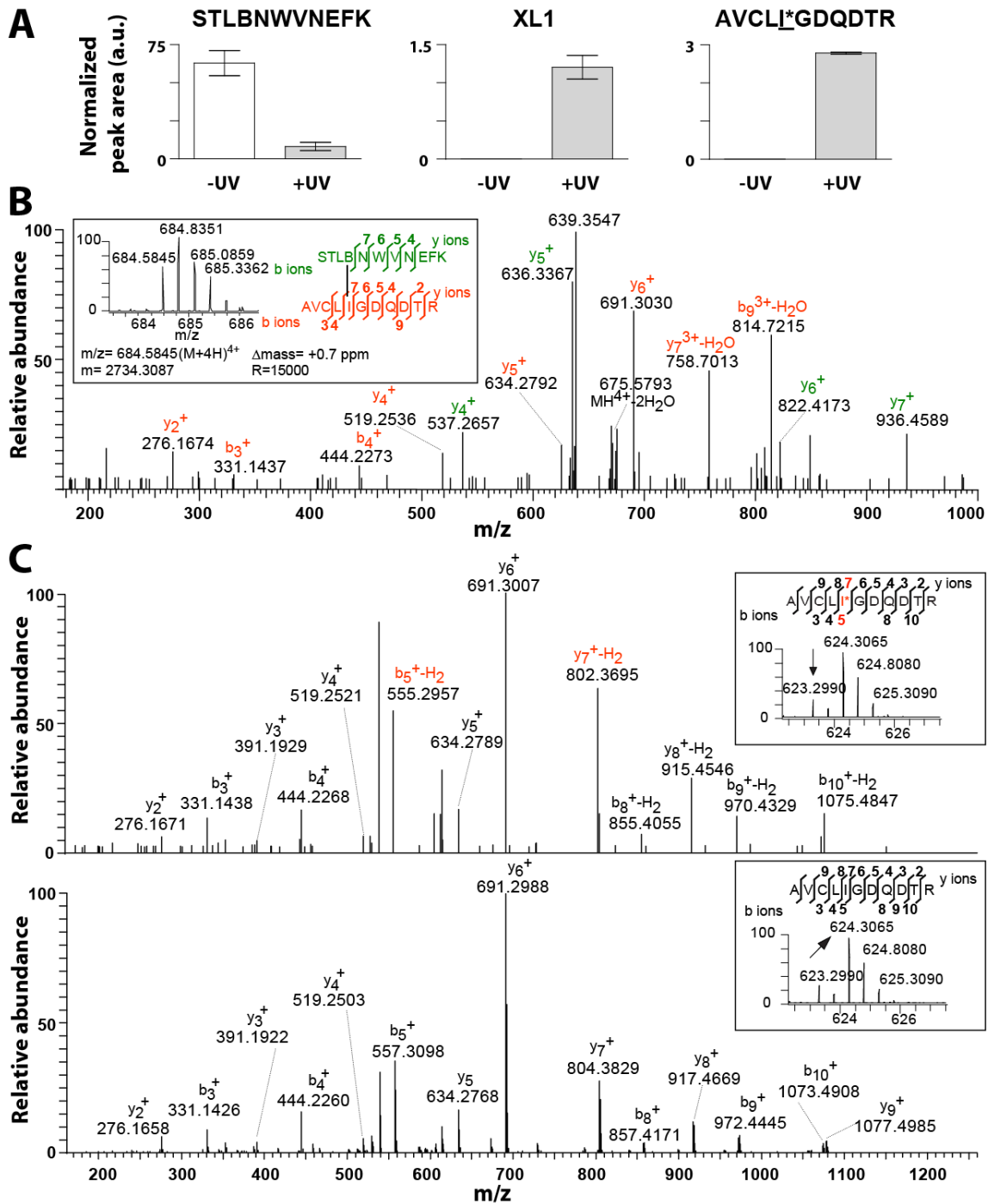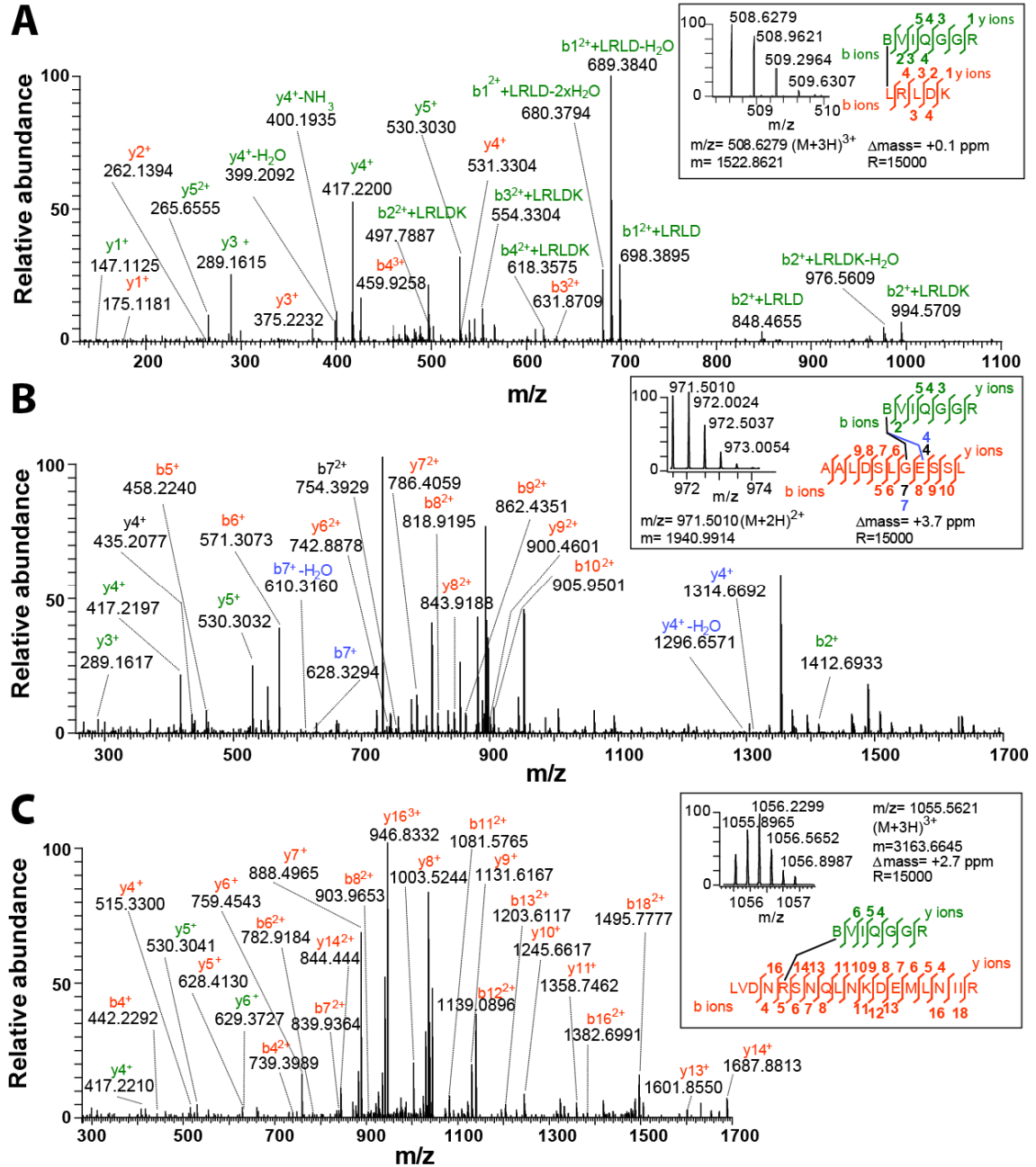
Supplementary Figure 5: Validation of UV reaction products present in the Q190B ISWI[26-648] dataset. A) UV-dependent decrease of the Bpa-containing peptide and concomitant increase of the crosslinked peptide XL1 and the peptide AVCLI*GDQDTR, in which Bpa abstracted two hydrogen atoms at the underlined position. Quantification of the peptides was performed as in Fig. 5A. N=2 and SD as error bars. B) High resolution, high accuracy MS2 fragmentation spectrum of the XL1 crosslinked peptide. Inset: Mass, charge and measurement error determination of the unfragmented crosslinked peptide (see Fig. 5B for details) and mapping of the observed fragment ions onto the sequence of the crosslinked peptide. B symbolizes Bpa. C) Comparison of the MS2 fragmentation spectra of the H_2-eliminated peptide AVCLI*GDQDTR (top) and the corresponding unmodified peptide (bottom). Their charges, masses and the measurement errors were derived as in panel B (insets). The site of hydrogen elimination (I209) could be unambiguously assigned by the $b5^+$ and $y7^+$ fragment ions.
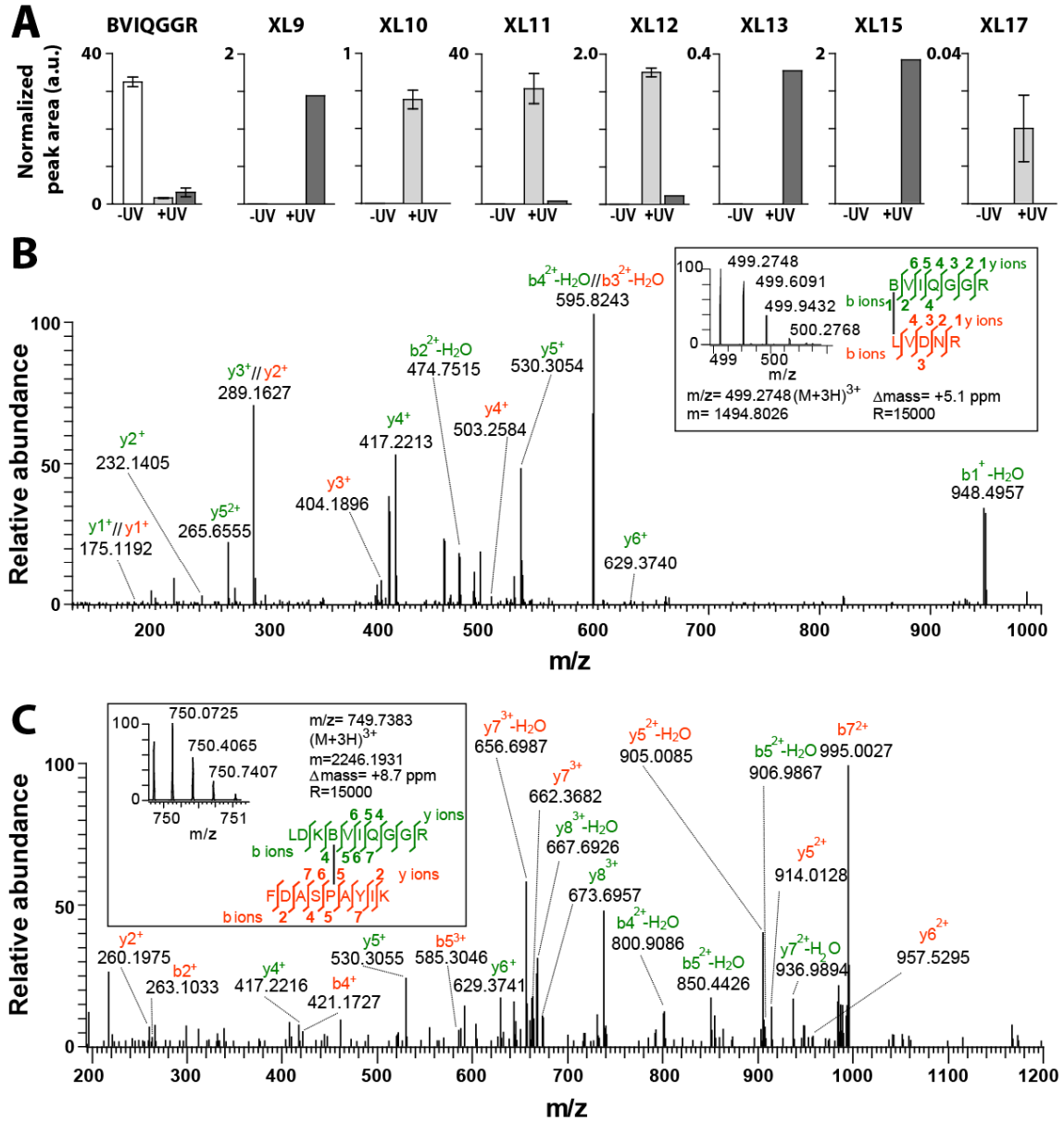
Supplementary Figure 6: Validation of crosslinks detected in the F361B ISWI[26-648] dataset. A) UV-dependent decrease of the Bpa-containing peptide and concomitant increase of the crosslinked peptides XL3 and XL4. Quantification of the peptides was performed as in Fig. 5A. N=3 and SD as error bars. B and C) High resolution, high accuracy MS2 fragmentation spectra of the XL3 and XL4 crosslinked peptides, respectively. The residues involved in the crosslink were assigned based on the observed b and y fragment ions. The acceptor amino acid is ambiguous for XL3 (A or K) due to the absence of distinguishing fragment ions. The insets show the mass, charge and measurement error determination of the unfragmented crosslinked peptides (see Fig. 5B for details) and mapping of the observed fragment ions onto the sequences of the crosslinked peptides. B: Bpa.
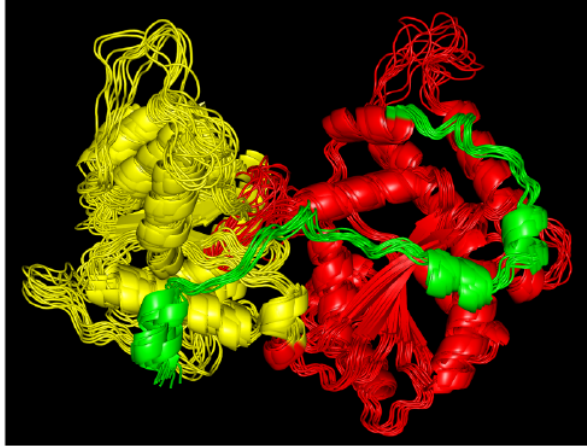
Supplementary Figure 7: Validation of the crosslinks XL6 (A), XL7a and XL7b (B) and XL8 (C) in the M578B ISWI[26-648] dataset through assignments of product ions in high resolution, high accuracy MS2 fragmentation spectra. The residues involved in the crosslink were assigned based on the observed b and y fragment ions. Two sets of mutually exclusive b and y ions were found for XL7 in panel B (colored black and blue). They indicate two alternative target amino acids of the crosslink, G644 and E645. The two crosslinks were named XL7a and XL7b, respectively, in Table 1. The insets show the mass, charge and measurement error determination of the unfragmented crosslinked peptides (for details see Fig. 5B) and mapping of the observed fragment ions onto the sequences of the crosslinked peptides. B symbolizes Bpa.

Supplementary Figure 8: Validation of the crosslinks in the M578B ISWI FL dataset. A) UV-dependent decrease of the Bpa-containing peptide and concomitant increase of selected crosslinked peptides. Quantifications were performed on the dataset derived from the main (white and light grey boxes) and up-shifted bands in Fig. 3B (dark grey). XL9, 13 and 15 were enriched in the shifted band. They span large distances in the primary structure explaining their altered gel mobility (Table 1). B and C) Assignments of product ions in high resolution, high accuracy MS2 fragmentation spectra of XL10 and XL13, respectively. The residues involved in the crosslink were assigned based on the observed b and y fragment ions. The insets show the mass, charge and measurement error determination of the unfragmented crosslinked peptides (see Fig. 5B for details) and mapping of the observed fragment ions onto the sequences of the crosslinked peptides. B: Bpa.
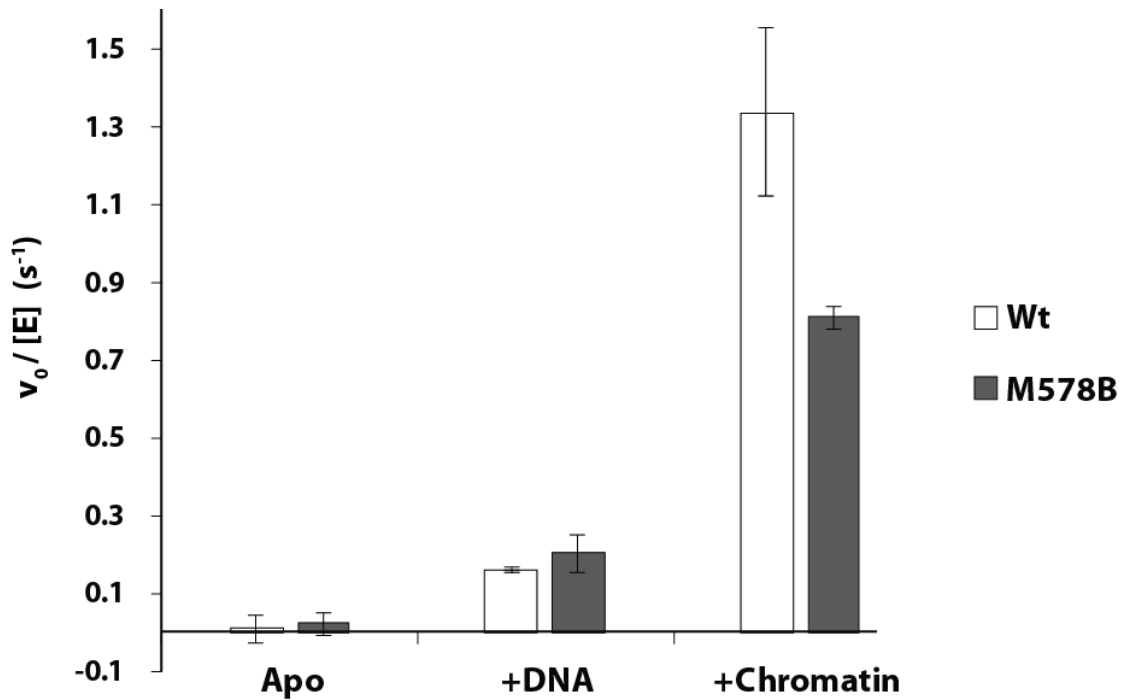
Supplementary Figure 9: The M578B mutation in full-length ISWI does not significantly affect ATP turnover at steady state. Displayed are the initial velocities for ATP hydrolysis normalized by the enzyme concentration in the absence of nucleic acids (Apo), in the presence of linearized plasmid DNA (0.2 mg/mL), or chromatinized plasmid DNA (0.1 mg/mL). Control experiments using four-fold lower DNA or chromatin concentrations confirmed the conclusions and showed that the chosen DNA concentration was fully saturating while the chromatin concentration was near saturation. Error bars: standard deviation of at least three measurements.

Supplementary Figure 10: Sampling of the conformational space available to the ATPase lobes of ISWI in the restrained conformation. Ten structural models were generated based on the structural information from crosslinks and the Chd1 crystal structure as in Fig. 6B. The models were structurally aligned to lobe 1 (yellow). All modeled conformations showed a very similar conformation of lobe 2 (red) and the C-terminal extension (green) with respect to lobe 1. Differences among the models arose primarily in loops that were missing in the crystal structure of Chd1 and were therefore relatively unconstrained during modeling.
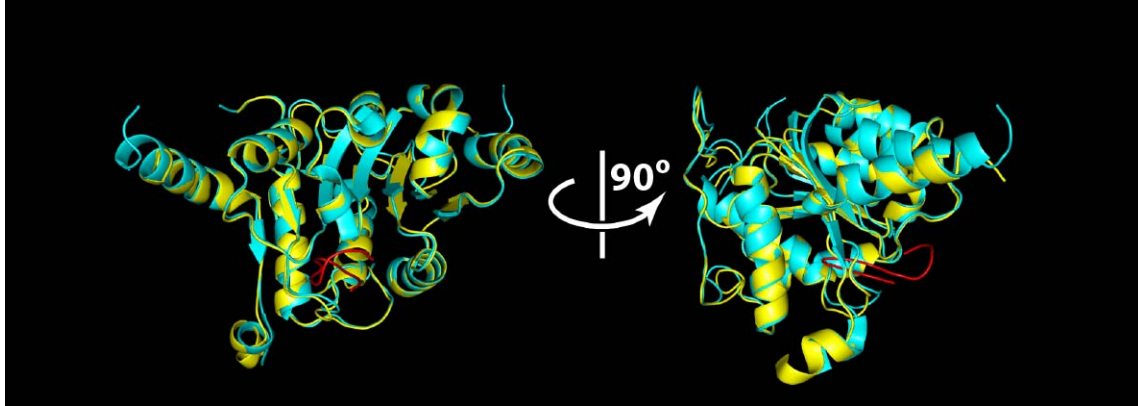
Supplementary Figure 11: A loop harboring amino acid M578 in lobe 2 of ISWI was found to insert itself deeply into a cavity present in lobe 1 to satisfy crosslinking constraints in the restrained model. Minimal structural rearrangements were necessary in lobe 1 for this configuration. Depicted is a superposition of the structures of Chd1 (cyan; PDB code 3MWY) and of the restrained structural model of ISWI (yellow; Fig. 6B). Amino acids 575-585 that are part of the loop in the second ATPase lobe of ISWI are colored red. For clarity only the loop and lobes 1 of ISWI and Chd1 are shown. Left: front view, right: side view of the interface between lobes one and two.