

Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes

Alexander Leitner^{a,1}, Lukasz A. Joachimiak^b, Pia Unverdorben^c, Thomas Walzthoeni^a, Judith Frydman^b, Friedrich Förster^c, and Ruedi Aebersold^{a,d,1}

^aDepartment of Biology, Institute of Molecular Systems Biology, Eidgenössische Technische Hochschule Zürich, 8093 Zurich, Switzerland; ^bDepartment of Biology, Stanford University, Stanford, CA 94305; ^cDepartment of Molecular Structural Biology, Max-Planck-Institute of Biochemistry, 82152 Martinsried, Germany; and ^dFaculty of Science, University of Zurich, 8057 Zurich, Switzerland

Edited by Brian T. Chait, Rockefeller University, New York, NY, and accepted by the Editorial Board May 25, 2014 (received for review November 5, 2013)

The study of proteins and protein complexes using chemical cross-linking followed by the MS identification of the cross-linked peptides has found increasingly widespread use in recent years. Thus far, such analyses have used almost exclusively homobifunctional, amine-reactive cross-linking reagents. Here we report the development and application of an orthogonal cross-linking chemistry specific for carboxyl groups. Chemical cross-linking of acidic residues is achieved using homobifunctional dihydrazides as cross-linking reagents and a coupling chemistry at neutral pH that is compatible with the structural integrity of most protein complexes. In addition to cross-links formed through insertion of the dihydrazides with different spacer lengths, zero-length cross-link products are also obtained, thereby providing additional structural information. We demonstrate the application of the reaction and the MS identification of the resulting cross-linked peptides for the chaperonin TRiC/CCT and the 26S proteasome. The results indicate that the targeting of acidic residues for cross-linking provides distance restraints that are complementary and orthogonal to those obtained from lysine cross-linking, thereby expanding the yield of structural information that can be obtained from cross-linking studies and used in hybrid modeling approaches.

Proteins exert the majority of their functions in the form of protein complexes to control cellular signaling, protein synthesis, folding and degradation, and many more essential processes. Therefore, elucidating the composition and structure of such complexes has been a longstanding goal of biological research.

MS-based proteomics has emerged as one of the main techniques to identify and quantify proteins and their modifications in biological samples such as isolated complexes, proteome fractions, or whole proteomes. Various MS methods now provide structural information on protein assemblies (1–3). Among them, chemical cross-linking and identification of cross-linked peptides by MS (XL-MS) has been increasingly applied to determine the subunit arrangements of biologically relevant complexes (4–6). Such XL-MS experiments indicate the locations of cross-linking sites and thus the spatial proximity of reactive groups that are connected by a covalent bond. This information is then used to determine the positioning of subunits or locate interacting regions, alone or in combination with other techniques such as NMR spectroscopy, electron microscopy, and X-ray crystallography.

In the last few years, optimized protocols and new computational tools for the reliable analysis of XL-MS datasets resulted in significant advances of the XL-MS technology (4–6). These advances have contributed to the emergence of a robust, integrated XL-MS method that has been successfully applied for structure determination of a number of large protein complexes (7–11) and the detection of direct, physical interactions in whole cells (12–14). To date, the cross-linking chemistries applied in these studies have targeted primary amines. Predominantly, *N*-hydroxysuccinimide esters were used as reactive groups, although other chemistries, for example, based on amidates, have also been described (8, 15, 16). Cysteine-specific cross-linking is also well established but usually does not yield sufficient structural

information due to the low prevalence of cysteines in proteins and their involvement in the formation of disulfide bonds. Zero-length cross-linking by carbodiimide coupling (17–19) and photochemical cross-linking (20) are other strategies that have been described but have not yet found widespread application in the field.

The development of cross-linking chemistries that cross-link functional groups different from amino groups but maintain the efficiency achieved by amine-specific cross-linking are expected to be highly beneficial to increase the depth of structural information obtained from cross-linking experiments. Specifically, such a technique would generate distance restraints from protein regions that are refractive to amine-specific cross-linking under the conditions used and reduce the coverage bias in basic sites. Also, typically only a fraction of theoretically possible cross-links are experimentally observed. This effect is presumably caused by variations in the reactivity of individual lysine residues and/or the unsuitable properties of the resulting cross-linked peptides for MS analysis. An increase in the number of cross-links per substrate by the use of an orthogonal chemistry would therefore be beneficial for de novo identification of hetero-oligomer subunit architectures, as well as restraints in hybrid methods incorporating low-resolution structural information.

Residues with carboxyl-terminating side-chains (aspartic and glutamic acids) are attractive targets because of their high prevalence

Significance

Most proteins carry out their function by associating with other proteins in stable or transient complexes. Structural analysis of such complexes, for example by crystallography, is frequently performed to study their function or mode of action. However, many large complexes are refractory to traditional structural biology methods. Alternative methods have been developed that are often combined with computational methods in hybrid structural biology strategies. Among these, the combination of chemical cross-linking and MS (XL-MS) has shown to be particularly informative. Current XL-MS methods mainly rely on the coupling of lysine residues. Here we describe a chemistry to cross-link acidic residues that generates structural information complementary to that obtained by amine-specific cross-linking, thus significantly expanding the scope of XL-MS analyses.

Author contributions: A.L. and R.A. designed research; A.L., L.A.J., P.U., and T.W. performed research; A.L., L.A.J., P.U., T.W., J.F., and F.F. contributed new reagents/analytic tools; A.L., L.A.J., P.U., T.W., J.F., and F.F. analyzed data; and A.L. and R.A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. B.T.C. is a guest editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

¹To whom correspondence may be addressed. E-mail: leitner@imsb.biol.ethz.ch or aebersold@imsb.biol.ethz.ch.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1320298111/-DCSupplemental.

in most proteins. In the most recent release of the SwissProt database (21), version 2014_04, 5.5% and 6.7% of all residues are Asp and Glu, respectively, compared with 5.8% for Lys. However, the low intrinsic chemical reactivity of carboxylic acids poses practical challenges for cross-linking reactions, requiring the use of a coupling reagent. Novak and Kruppa used different dihydrazides as cross-linking reagents in combination with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) for activation (22). However, to obtain sufficiently high reaction yields, cross-linking was carried out at relatively low pH (5.5 in ref. 22), which is incompatible with pH-sensitive assemblies. Furthermore, the method was only applied to a single protein, ubiquitin, resulting in the identification of two cross-links.

Here we introduce a cross-linking chemistry that connects proximal carboxyl groups [acidic cross-linking (AXL)], whereby side-chains of Asp and Glu residues are cross-linked with dihydrazides using the coupling reagent 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMTMM) (23) (Fig. 1A). In contrast to EDC, DMTMM is able to couple carboxylic acids with hydrazide-based cross-linkers at neutral pH (7–7.5), ensuring good reaction yields and biocompatibility. Experiments with model proteins yielded numbers of cross-linked peptides that were in the same range to those generated by the well-established Lys-specific cross-linking using the reagent disuccinimidyl suberate (DSS). On top of that, a second set of cross-linking restraints is observed in the form of zero-length cross-links between Lys and Asp or Glu residues, respectively.

To show the practical relevance of the method, we applied it to multisubunit complexes in the megadalton range that have been recently probed with lysine-specific cross-linking and for which structural information is available. The results indicate that for both the chaperonin TRiC/CCT from *Bos taurus* and the *Schizosaccharomyces pombe* 26S proteasome, cross-links were identified that are in agreement with the available structures of these complexes. Acidic and zero-length cross-links provided orthogonal sets of structural restraints that are complementary to a cross-linking chemistry targeting lysines. We therefore expect that chemical cross-linking of acidic residues will become an important method for a more comprehensive structural analysis of protein complexes by XL-MS.

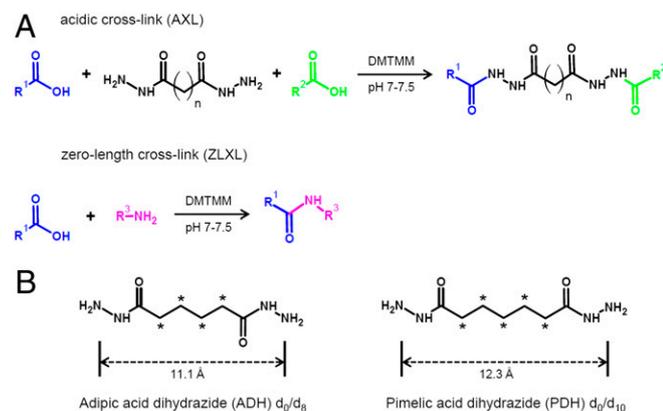


Fig. 1. (A) Cross-linking reactions involving coupling of carboxyl groups using dihydrazides (acidic cross-linking, AXL; *Top*) and zero-length cross-linking (ZLXL) with DMTMM as coupling reagent. R^1 and R^2 denote acidic residues (Asp, Glu) in a single or two different proteins, R^3 denotes a primary amine (usually Lys). (B) Structure of the two dihydrazide reagents used in this study. Asterisks denote positions where hydrogen atoms are exchanged for deuterium atoms in the heavy form of the reagent. The spacer length of the reagents (calculated between the terminal nitrogen atoms) is also given.

Results and Discussion

Establishment of Cross-Linking Chemistry and Optimization of Reaction Conditions. To cross-link proximal carboxyl groups under native conditions, cross-linking reagents with different functional groups can be conceived. (Di)hydrazides are interesting candidates, as they are intrinsically less basic than amines and introduce a bio-orthogonal functional group that shares some chemical similarities with primary amines. However, the previously described cross-linking protocol [using EDC as coupling reagent and 50 mM pyridine HCl buffer, pH 5.5 (22)] is not compatible with preserving many native protein structures. A different activation chemistry is therefore required for the cross-linking reaction at neutral, near-physiological pH.

We tested the more reactive coupling reagent, DMTMM (23), for coupling hydrazides to carboxylic acids in proteins and protein complexes (Fig. 1A). Recently, DMTMM has been applied to many challenging synthetic problems, for example, in peptide synthesis and bioconjugation chemistry (24–26). Hayama et al. have shown its use in the coupling of the fluorescent derivatization agent, 4-(1-pyrene)butanoic acid hydrazide, to carboxyl groups of microcystins (27). However, to our knowledge, DMTMM has not been used for conjugation of hydrazides to proteins. We tested the suitability of DMTMM on the lysine-free peptide angiotensin II (DRVYIHPF), using adipic acid dihydrazide (ADH) as a homobifunctional cross-linking reagent. Comparing the results of this coupling chemistry with those obtained with EDC coupling, the data indicate that, although EDC only showed significant coupling of ADH to the peptide under acidic conditions up to a pH of 5.8, DMTMM also allowed the reaction at near-neutral pH of 7.4 in PBS buffer (Fig. S1).

Subsequent cross-linking experiments using a set of three model proteins confirmed the initial observations made with a synthetic peptide, demonstrating for the first time the chemical cross-linking of acidic residues at neutral pH in folded proteins. The influence of experimental parameters such as reagent excess, protein concentration, reaction time, and temperature were systematically varied. Similar to lysine-lysine cross-linking using succinimide-based reagents, highest reaction yields were obtained at temperatures of 25 °C or 37 °C and at protein concentrations at or above 1 mg/mL (Fig. S2). Although succinimide esters are subject to hydrolysis in aqueous solutions, the hydrazide reaction can be carried out for longer periods of time without depletion of the reagents, if desired. This stability provides additional flexibility in the experimental protocol. Instead of a quenching step, we opted for the removal of excess reagents by gel filtration to efficiently terminate the cross-linking reaction (*SI Materials and Methods*). Overall, these experiments established a robust and optimized protocol that was suitable for the application to more complex samples.

Results for Model Proteins and Comparison with Lysine-Lysine Cross-Linking.

We then evaluated the performance of the acidic cross-linking method using a set of eight model proteins that was previously found to be highly suitable for this purpose (28). Three dihydrazide reagents with different spacer lengths were tested. Initial experiments showed inferior performance for the longest-chain reagent, suberic acid dihydrazide, due to its very low solubility in aqueous medium. This reagent was therefore not included in subsequent experiments. To take advantage of isotope coding that improves bioinformatic data analysis using the previously described software xQuest (29, 30), we obtained deuterated analogs of ADH and pimelic acid dihydrazide (PDH) (Fig. 1B) by custom synthesis.

An overview of the experimental workflow is given in Fig. 2. Mixtures of the eight model proteins were cross-linked using one of the two hydrazides under the conditions outlined in *SI Materials and Methods*. Cross-linked proteins were digested with Lys-C and

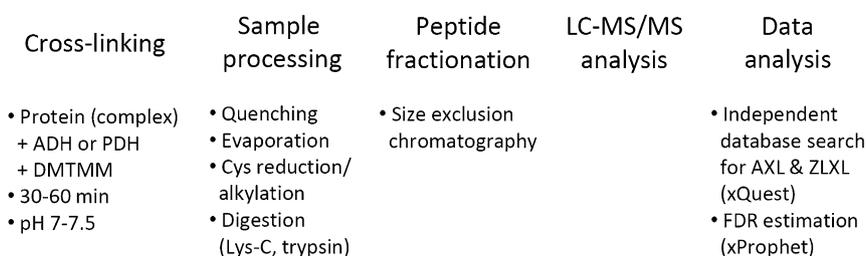


Fig. 2. Experimental workflow for the chemical cross-linking of acidic side-chains and the identification of cross-linked peptides by MS.

trypsin, and the resulting peptides were fractionated by size exclusion chromatography (SEC). Three SEC fractions that were most strongly enriched in cross-linked peptides were analyzed by LC-MS/MS, essentially as described previously for lysine cross-linked substrates (28).

Data analysis was carried out using xQuest. In addition to searches for the reaction products that incorporated the hydrazide reagents, we also carried out searches for zero-length cross-links (ZLXL). DMTMM has not been used before in ZLXL studies, where typically carbodiimides such as EDC are used to introduce cross-links between spatially proximal lysine and aspartic or glutamic acid side-chains. Because ZLXL does not allow for the direct introduction of an isotope signature, we opted for a two-step data analysis strategy: MS/MS spectra from light/heavy scan pairs (determined by xQuest) were searched to obtain hydrazide cross-links, whereas all MS/MS spectra were searched separately for zero-length cross-links. Both searches were independently evaluated using xProphet to determine appropriate score thresholds for a chosen false discovery rate (for details, see *SI Materials and Methods*). This data analysis strategy avoids more complex experimental schemes and does not require specialized software as used, for example, in a recent ZLXL study (31). AXL and ZLXL identifications obtained for the two hydrazide reagents are summarized in Table 1 and compared with previously published data generated with the amine-reactive reagent, DSS, on the same instrumentation (28). A detailed list of all identified cross-links from the eight model proteins treated with ADH and PDH is provided in *Dataset S1 A–D*, and results are further compared in *Table S1*.

For PDH, cross-links were obtained for all eight proteins. In contrast, for ADH, the coverage was less extensive (six of eight proteins), whereas zero-length cross-links were identified for all proteins except catalase from both samples. The number of identified cross-linked peptides was lower for PDH, 94, and ADH, 60, respectively, than previously obtained for DSS, 147, under similar conditions, although the differences were clearly protein specific. For example, although BSA and chicken transferrin gave similarly high numbers for both cross-linking chemistries, the coverage for

catalase was much more extensive for the AXL approach (25 vs. 3). In addition to identifying cross-linked peptides by xQuest, we also performed additional data analysis using the unrestricted modification search engine, MOD^a (32), which confirmed that the cross-linking chemistry was exclusively specific to Asp and Glu residues (*Fig. S3* and *Dataset S2*). The hydrazide cross-links were complemented by additional zero-length cross-links formed in the same reactions; 70 were identified in the ADH sample and 52 in the PDH sample. It appears that the number of ZLXL products inversely correlates with the formation of AXL products. The absence of zero-length cross-links on catalase is in line with the low number of DSS cross-links observed on this protein. Therefore, there seems to be a lack of suitably positioned or relative lysine residues in catalase.

To verify that the cross-linking protocol provides representative structural information, we validated the cross-linking identifications on the known structures of the test proteins. Using BSA as an example, we observed that nearly all cross-links were compatible with the model derived from X-ray crystallography. For ADH, all 22 restraints (unique combination of cross-linking sites) fell within a calculated C_α-C_α distance of less than 21 Å (range, 9.1–20.6 Å). For PDH, 30 restraints were observed that covered a range of 8.7–20.5 Å, plus an additional one with a theoretical C_α-C_α distance of 35.5 Å, the single likely false-positive assignment. Fifteen zero-length cross-links on BSA bridged distances of ≤16 Å exclusively (range, 6.5–16.0 Å).

We then expanded the structural validation to the remaining standard proteins. *Fig. S4* visualizes the cross-link identifications on the seven proteins with available high-resolution 3D structures and shows that the cross-linking chemistries with different specificities are highly complementary as different regions of the protein are covered.

In total, 65 nonredundant distance restraints (defined as unique contacts between acidic side chains and measured as C_α-C_α distances) could be validated on available 3D structures for the PDH-AXL data set (48 for the ADH data). The majority of the data unequivocally fit the models with distances of less than 30 Å: 63 of 65 PDH restraints were below this threshold, whereas two were

Table 1. Nonredundant cross-linked peptides identified from eight model proteins

Protein	ADH AXL	ADH ZLXL	PDH AXL	PDH ZLXL	DSS*
Catalase, bovine	14	0	25	0	3
Creatine kinase, rabbit	0	5	1	6	14
Fructose-bisphosphate aldolase A, rabbit	0	8	1	3	12
Lactoferrin, bovine	5	7	5	7	21
Ovotransferrin, chicken	10	18	13	15	19
Pyruvate kinase, rabbit	3	8	1	2	7
Serotransferrin, bovine	5	9	14	7	33
Serum albumin, bovine	23	15	34	12	38
Total	60	70	94	52	147

*Data for comparable SEC fractions taken from ref. 28. ADH, adipic acid dihydrazide; AXL, acidic cross-link; DSS, disuccinimidyl suberate; PDH, pimelic acid dihydrazide; ZLXL, zero-length cross-link.

in the range of 33–36 Å. For ADH, 46 of 48 validated restraints corresponded to distances below 21 Å; two cross-links in bovine lactoferrin significantly exceeded this distance. In comparison with PDH, but also with the lysine-reactive DSS, the restraints of ADH were found to be significantly shorter (two-sample *t* test PDH vs. ADH distances, $P = 0.02$; DSS vs. ADH, $P = 0.0001$). This difference could be explained by the shorter distance between the α -carbons (16 atoms distance for Glu-ADH-Glu vs. 18 atoms for Lys-DSS-Lys). Conformational flexibility within the side-chains may also play a role, because the spacer length of all three reagents differs only by roughly 1 Å (ADH: 11.1 Å, DSS: 11.4 Å, and PDH 12.3 Å). This information is valuable for the use of the restraints in various modeling methods (33, 34). In our hands, using shorter chain amine-reactive succinimide reagents did not result in a comparable reduction in accuracy (4). However, increased structural flexibility of large complexes may partially offset this advantage. In contrast to ADH, PDH appears to provide restraints of similar maximum length than DSS [~ 30 Å as upper bound, $P = 0.09$ (35)].

Structural validation of the zero-length cross-links on the proteins tested did not provide a clear distance cutoff value. Although more than two thirds of the distances were below 15 Å, the remaining third were above this value, so that a more appropriate cutoff may lie closer to 25 Å. Zero-length restraints were comparable in length in both ADH and PDH data sets ($P = 0.25$), which is not unexpected because they are independent of the hydrazide. Interestingly, they were not significantly shorter than ADH cross-links ($P = 0.34$), but shorter than PDH cross-links ($P = 0.003$). Especially for larger protein assemblies, structural flexibility of the complex may be a more important factor for the accuracy of the restraint than spacer length, as is also confirmed by the proteasome data discussed below.

Fig. 3 compares the observed distance distributions for the complete model protein data set. We also carried out additional comparisons of the cross-linked peptide data sets obtained using different chemistries, for example, regarding the efficiency of the SEC fractionation or the mass and charge state distributions (*SI Text*, Table S1, and Figs. S5 and S6). The higher frequency of Asp and Glu residues as cross-linkable sites also poses additional challenges to the exact localization of cross-linking sites, not unlike the assignment of posttranslational modifications (36). This issue is discussed in detail in *SI Text*.

In summary, we demonstrate for the first time, to our knowledge, a cross-linking chemistry for coupling acidic residues in proteins

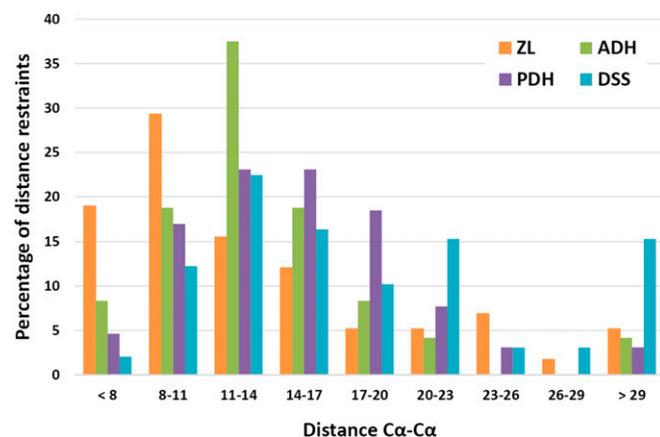


Fig. 3. Histogram depicting the distribution of C_{α} - C_{α} distances as obtained by validation of structures of the model proteins. Data are shown for the acidic cross-linking reagents ADH (green) and PDH (magenta), zero-length cross-links (ZLXL, from ADH experiment, in orange), and previous data from a lysine-lysine cross-linking study using DSS (blue; taken from ref. 28).

maintained at near neutral pH that provides numbers of cross-links and therefore distance restraints that are on par with those generated with established lysine-specific methods.

Applications to TRiC Chaperonin and Proteasome Complexes. To demonstrate the applicability of the AXL protocol to multi-subunit protein complexes of high biological relevance, we selected two complexes that were previously studied in-house.

The chaperonin TRiC/CCT is a complex consisting of 16 subunits arranged in two back-to-back stacked rings. Each ring consists of one copy of eight homologous, but not identical, subunits, averaging about 60 kDa in mass. The subunit composition of the two rings is identical so that the intact complex has two copies of each subunit. Lysine-lysine cross-linking data on TRiC/CCT has recently been obtained in three independent studies (37–39). These studies provided convincing evidence that subunit arrangements proposed earlier were in fact incorrect, and the new subunit arrangement was supported by a reexamination of published X-ray crystallography data (38, 40).

In the present work we performed AXL on the open state configuration (i.e., in the absence of ATP) of the bovine TRiC complex as an example. PDH was used as the cross-linking reagent. Analysis of three SEC fractions enriched for cross-linked peptides resulted in the identification of 25 intrasubunit and 22 intersubunit cross-links containing the PDH spacer at an estimated false discovery rate of less than 5%. For the analyses, 50 μ g of total protein was used as starting material (see *Dataset S1E* for a complete list of identifications). Forty-four of these 47 cross-links could be validated on a homology model of the complex, of which 40 were found to have C_{α} - C_{α} distances of not more than 28.2 Å. (Five cross-links have an ambiguous topology with a maximum inaccuracy of 4.3 Å.) Two additional cross-links corresponded to distances of ~ 33 Å, but are still plausible because they were located in the apical region where significantly higher structural flexibility is expected (41) and the available structural model might be less accurate (Fig. S7). Only two cross-links (4.3%) were found to be incompatible with the structure, in line with the expected false discovery rate of 5% for the whole data set.

Searches for ZLXL products revealed exceptionally large numbers: 111 intrasubunit and 63 intersubunit cross-links were identified at 5% false discovery rate (*Dataset S1F*). We assume that this correlates with the accessibility and reactivity of the lysine residues, because previous DSS experiments also resulted in high numbers of cross-links (38). One hundred four of 111 intrasubunit cross-links could be validated on the TRiC model, of which 14 exceeded a 25-Å distance. In contrast, only 18 of the 53 intersubunit cross-links that could be validated were within this distance. However, only three cross-links clearly violated the subunit arrangement of the complex as determined by us and others (37–39). In fact, all but four of the large-distance cross-links (>30 Å) lie within the apical region, defined here as >30 Å from the interranging interface region (see annotations in *Dataset S1F* and Fig. S7). This part of the complex is known to be highly flexible and conformationally dynamic in the open state (41). We speculate that the distance violations are caused by the substantial conformational flexibility in this region in the open state of the complex that is not accounted for in the present model.

Interestingly, the complementarity of the DSS and PDH data sets initially observed with the reference protein set described above was also high for this sample. The highest number of AXL intersubunit cross-links was found between subunits CCT5 and CCT7 (Fig. 4A), which was the only interranging subunit contact not covered by DSS cross-links in our earlier work (38), further highlighting the power of using an orthogonal XL-MS cross-linking chemistry to improve coverage and thus confidence in identifying de novo assemblies. Theoretical calculations show that for carboxyl- and amine-specific chemistries, many cross-linkable

residues are close enough in space, so that the absence of observed Lys-Lys cross-links may result from differences in reactivity of specific targeted residues or mass spectrometric properties of the generated cross-linked peptides.

To test the method on an even more complex supramolecular assembly, we studied the 2.5-MDa 26S proteasome. This complex consists of more than 60 individual subunits (two copies per subunit are present in the complex) that are assembled in two major subcomplexes: the 19S regulatory particle and the 20S core particle. Only recently, the positioning of most subunits in the 26S holocomplex could be established in two studies (42, 43). We and others have also performed cross-linking experiments on proteasomes from various organisms using lysine-specific reagents (28, 30, 43–48).

Here, we cross-linked a 26S proteasome preparation from *S. pombe* using the shorter spacer-length hydrazide ADH as cross-linking reagent. From a combination of four separate experiments (total starting amount, ~200 μ g), 83 nonredundant ADH-cross-linked peptides were identified (10 of which with ambiguous cross-linking site localization; see [Dataset S1G](#) for details). Of the 83 ADH-cross-linked peptides, 60 contacts were within a subunit, whereas 23 were between different subunits. In addition, we observed 71 intrasubunit and 34 intersubunit ZL cross-linked peptides (5 with ambiguous topology; see [Dataset S1H](#)).

As can be seen in the summary given in Fig. 4B, intersubunit cross-links covered all regions of the large complex. Subunit contacts were identified that connected subunits in the 20S core particle, the ATPase ring of the 19S regulatory particle, as well as among the non-ATPase subunits of the 19S subcomplex. In addition, cross-links between ATPase and non-ATPase subunits and connecting the 19S and 20S regions were found.

We validated these data using a model that was obtained by homology modeling from the previously published *Saccharomyces cerevisiae* assembly (49) ([SI Materials and Methods](#)). Sixty-seven ADH-cross-links could be mapped on the structure. Overall, three cross-links (3.6% of the total) were found to be clearly incompatible with the model, when score thresholds corresponding to an estimated 5% false discovery rate for the data set were chosen. All structurally validated cross-links except one were found to bridge distances of less than 24 Å, in line with the relatively short ADH-derived distance restraints observed in model proteins. The single outlier with a distance of 35.9 Å is located at the interface of Prs4 (Rpt2) and Rpn2 that is known to undergo substantial conformational transition on binding and hydrolysis of ATP (50). In addition, 89 ZL-cross-links could be validated, of which 9 exceeded a 25-Å threshold. Six of these nine violations occurred in the region of the subunit Prs6A (Rpt5), either as intra- or intersubunit contacts. A recent study examined the conformational flexibility of the 26S holocomplex based on deep classification of a large cryo-electron microscopy data set (51). This work revealed that the proximity of Prs6A is one of the regions that undergoes substantial

conformational rearrangement in different states. Therefore, it is highly likely that at least some of the cross-links in this region are indeed plausible taking this flexibility into account. Generally, the accuracy of the distance restraints for ADH-AXL and ZLXL seem to be very similar, in line with the observations from the model protein data set.

We again observe a high degree of complementarity of the AXL/ZLXL data to the DSS data. As an example, among the 10 different interprotein contacts represented by 20 ADH intersubunit cross-links that are compatible with the model, 7 were not found in a previous study (30) from our laboratory that used the DSS reagent. One restraint solidifies the positioning of the Rpn8/Rpn11 heterodimer. These two MPN domain-containing subunits form a heterodimer that could be positioned in the initial 26S model (43), although the exact location of the two subunits remained ambiguous at this point. In this earlier study, four DSS cross-links were observed between Rpn11 and Rpt3, one of the AAA-ATPases. Here, the ADH cross-link observed between Rpn8 and Rpn9 supports the relative localization of these two subunits, in line with recent electron microscopy data from the *S. cerevisiae* proteasome (49).

The difference in the number of hydrazide cross-links for the proteasome is partially explained by the shorter spacer length, in line with data from the model proteins, but comes with the benefit of (generally) higher-resolution distance restraints. In addition, the lower reactivity of acidic residues and bioinformatic limitations play a role. xQuest score thresholds need to be raised when cross-linking acidic residues due to their higher frequency, resulting in a larger search space. Additional, likely correct identifications are found below the score threshold set for intersubunit cross-links (also listed in [Dataset S1](#)). Similar restrictions apply to the identification of zero-length cross-links, where the search space is even larger (Lys, Asp, and Glu are potential cross-linking sites). In contrast to the situation where isotope-coded linkers are inserted into the structure, all MS/MS spectra need to be searched, thus further increasing the chance for random matches. The use of high mass accuracy on the MS/MS level could compensate for this effect. However, with the instrumentation used in this work, acquisition of high mass accuracy fragment ion spectra comes with a severe penalty in scan speed and sensitivity, making it impractical to use. Further improvements in sensitivity and mass accuracy of mass spectrometers will certainly improve the outcome of AXL studies on larger complexes.

Conclusions. We have successfully implemented a cross-linking strategy targeting acidic residues. The chemistry combines dihydrazides as cross-linking reagents with activation by the coupling reagent DMTMM. Application of the method to model proteins confirmed that the reaction conditions preserved the structural integrity of the proteins. The AXL chemistry offers flexibility regarding spacer lengths and reaction times because the hydrazides and the coupling reagent are stable under the

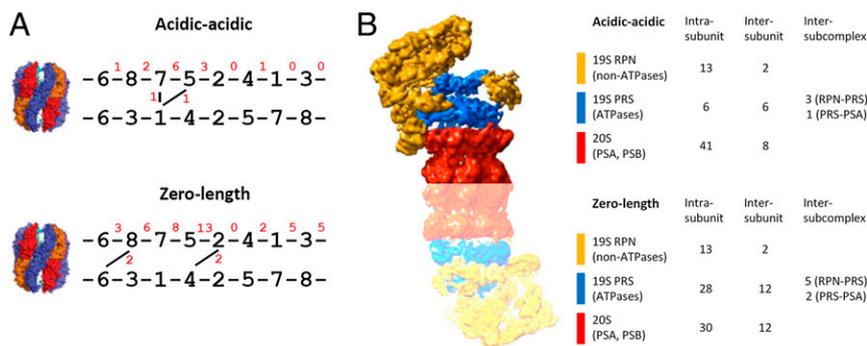


Fig. 4. Cross-linking of acidic residues and zero-length cross-linking applied to the TRiC/CCT and 26S proteasome complexes. (A) Connectivity map for TRiC/CCT depicting the intersubunit contacts observed. (B) Overview of intersubunit cross-links observed in different regions of the 26S proteasome. Structures of the complexes are used for illustrative purposes only.

reaction conditions. Longer-distance cross-links could potentially be realized if more hydrophilic spacers are used to increase solubility. In addition, zero-length cross-links, which provide a second layer of restraints, are obtained from the same reactions.

The suitability of the protocol for analyzing large multisubunit protein complexes was demonstrated using the TRiC/CCT chaperonin and the 26S proteasome as examples. The results suggest that this method is a very promising alternative to established lysine-specific cross-linking chemistries and will be a valuable addition to the expanding range of chemical cross-linking methods that are suitable for applications in structural biology. Clearly, additional cross-link restraints will improve the accuracy of newly established topologies and will increase the chances of observing interactions in specific regions of a complex, such as ligand binding sites, where high cross-link coverage is crucial. The protocol will

also enable the design of new functionalized cross-linking reagents incorporating gas-phase cleavable or affinity-tagged linkers based on hydrazide chemistry, as has been demonstrated for lysine-specific reagents.

Materials and Methods

Detailed experimental procedures are included in the *SI Materials and Methods*. This section covers the cross-linking reaction conditions, mass spectrometric analysis, and identification of cross-linked peptides by database search. Identified cross-linked peptides are listed in [Dataset S1](#).

ACKNOWLEDGMENTS. We thank Franz Herzog for valuable discussions, Ramya Kumar for help in the bovine TRiC purification, and Eunok Paek for providing the MOD³ software. Financial support from the European Union (PROSPECTS: Proteomics Specification in Space and Time, Grant HEALTH-F4-2008-201648; and European Research Council Advanced Grant Proteomics v3.0, Grant 233226) (to R.A.) is acknowledged. F.F. was supported by the Human Frontiers Science Program and the Deutsche Forschungsgemeinschaft (Graduiertenkolleg 1721).

- Serpa JJ, et al. (2012) Mass spectrometry-based structural proteomics. *Eur J Mass Spectrom* (Chichester, Eng) 18(2):251–267.
- Stengel F, Aebersold R, Robinson CV (2012) Joining forces: Integrating proteomics and cross-linking with the mass spectrometry of intact complexes. *Mol Cell Proteomics* 11(3):R111.014027.
- Walzthoeni T, Leitner A, Stengel F, Aebersold R (2013) Mass spectrometry supported determination of protein complex structure. *Curr Opin Struct Biol* 23(2):252–260.
- Leitner A, et al. (2010) Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol Cell Proteomics* 9(8):1634–1649.
- 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(3):530–540.
- Calabrese AN, Pukala TL (2013) Chemical cross-linking and mass spectrometry for the structural analysis of protein assemblies. *Aust J Chem* 66(7):749–759.
- Chen ZA, et al. (2010) Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J* 29(4):717–726.
- Laubler MA, Reilly JP (2011) Structural analysis of a prokaryotic ribosome using a novel amidinating cross-linker and mass spectrometry. *J Proteome Res* 10(8):3604–3616.
- Jennebach S, Herzog F, Aebersold R, Cramer P (2012) Crosslinking-MS analysis reveals RNA polymerase I domain architecture and basis of rRNA cleavage. *Nucleic Acids Res* 40(12):5591–5601.
- Ciferri C, et al. (2012) Molecular architecture of human polycomb repressive complex 2. *eLife* 1:e00005.
- Schmidt C, et al. (2013) Comparative cross-linking and mass spectrometry of an intact F-type ATPase suggest a role for phosphorylation. *Nat Commun* 4:1985.
- Bruce JE (2012) In vivo protein complex topologies: Sights through a cross-linking lens. *Proteomics* 12(10):1565–1575.
- Weisbrod CR, et al. (2013) In vivo protein interaction network identified with a novel real-time cross-linked peptide identification strategy. *J Proteome Res* 12(4):1569–1579.
- Chavez JD, Weisbrod CR, Zheng C, Eng JK, Bruce JE (2013) Protein interactions, post-translational modifications and topologies in human cells. *Mol Cell Proteomics* 12(5):1451–1467.
- Laubler MA, Reilly JP (2010) Novel amidinating cross-linker for facilitating analyses of protein structures and interactions. *Anal Chem* 82(18):7736–7743.
- Laubler MA, Rappsilber J, Reilly JP (2012) Dynamics of ribosomal protein S1 on a bacterial ribosome with cross-linking and mass spectrometry. *Mol Cell Proteomics* 11(12):1965–1976.
- Haladová K, et al. (2012) The combination of hydrogen/deuterium exchange or chemical cross-linking techniques with mass spectrometry: Mapping of human 14-3-3 ζ homodimer interface. *J Struct Biol* 179(1):10–17.
- Zhao C, et al. (2012) Cross-linking mass spectrometry and mutagenesis confirm the functional importance of surface interactions between CYP3A4 and holo/apo cytochrome b₅. *Biochemistry* 51(47):9488–9500.
- Rozbesky D, et al. (2013) Structural model of lymphocyte receptor NKR-P1C revealed by mass spectrometry and molecular modeling. *Anal Chem* 85(3):1597–1604.
- Pham ND, Parker RB, Kohler JJ (2013) Photocrosslinking approaches to interactome mapping. *Curr Opin Chem Biol* 17(1):90–101.
- Apweiler R, et al.; UniProt Consortium (2013) Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Res* 41(Database issue):D43–D47.
- Novak P, Kruppa GH (2008) Intra-molecular cross-linking of acidic residues for protein structure studies. *Eur J Mass Spectrom* (Chichester, Eng) 14(6):355–365.
- Kunishima M, et al. (1999) 4-(4,6-Dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride: An efficient condensing agent leading to the formation of amides and esters. *Tetrahedron* 55(46):13159–13170.
- Sturm M, Leitner A, Lindner W (2011) Development of an indole-based chemically cleavable linker concept for immobilizing bait compounds for protein pull-down experiments. *Bioconjug Chem* 22(2):211–217.
- Zivanovic A (2012) 4-(4,6-Dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride. *Synlett* 23(16):2426–2427.
- Perciani CT, et al. (2013) Conjugation of polysaccharide 6B from *Streptococcus pneumoniae* with pneumococcal surface protein A: PspA conformation and its effect on the immune response. *Clin Vaccine Immunol* 20(6):858–866.
- Hayama T, et al. (2012) Liquid chromatographic determination of microcystins in water samples following pre-column excimer fluorescence derivatization with 4-(1-pyrene)butanoic acid hydrazide. *Anal Chim Acta* 755:93–99.
- Leitner A, et al. (2012) Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. *Mol Cell Proteomics* 11(3):M111.014126.
- Rinner O, et al. (2008) Identification of cross-linked peptides from large sequence databases. *Nat Methods* 5(4):315–318.
- Walzthoeni T, et al. (2012) False discovery rate estimation for cross-linked peptides identified by mass spectrometry. *Nat Methods* 9(9):901–903.
- Sriswasdi S, Harper SL, Tang H-Y, Speicher DW (2014) Enhanced identification of zero-length chemical cross-links using label-free quantitation and high-resolution fragment ion spectra. *J Proteome Res* 13(2):898–914.
- Na S, Bandeira N, Paek E (2012) Fast multi-blind modification search through tandem mass spectrometry. *Mol Cell Proteomics* 11(4):M111.010199.
- Schneidman-Duhovny D, et al. (2012) A method for integrative structure determination of protein-protein complexes. *Bioinformatics* 28(24):3282–3289.
- Kahraman A, et al. (2013) Cross-link guided molecular modeling with ROSETTA. *PLoS ONE* 8(9):e73411.
- Merkley ED, et al. (2014) Distance restraints from crosslinking mass spectrometry: Mining a molecular dynamics simulation database to evaluate lysine-lysine distances. *Protein Sci* 23(6):747–759.
- Chalkley RJ, Clauser KR (2012) Modification site localization scoring: Strategies and performance. *Mol Cell Proteomics* 11(5):3–14.
- Kalishman N, Adams CM, Levitt M (2012) Subunit order of eukaryotic TRiC/CCT chaperonin by cross-linking, mass spectrometry, and combinatorial homology modeling. *Proc Natl Acad Sci USA* 109(8):2884–2889.
- Leitner A, et al. (2012) The molecular architecture of the eukaryotic chaperonin TRiC/CCT. *Structure* 20(5):814–825.
- Herzog F, et al. (2012) Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science* 337(6100):1348–1352.
- Kalishman N, Schröder GF, Levitt M (2013) The crystal structures of the eukaryotic chaperonin CCT reveal its functional partitioning. *Structure* 21(4):540–549.
- Meyer AS, et al. (2003) Closing the folding chamber of the eukaryotic chaperonin requires the transition state of ATP hydrolysis. *Cell* 113(3):369–381.
- Lander GC, et al. (2012) Complete subunit architecture of the proteasome regulatory particle. *Nature* 482(7384):186–191.
- Lasker K, et al. (2012) Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proc Natl Acad Sci USA* 109(5):1380–1387.
- Sharon M, Taverner T, Ambroggio XI, Deshaies RJ, Robinson CV (2006) Structural organization of the 19S proteasome lid: Insights from MS of intact complexes. *PLoS Biol* 4(8):e267.
- Bohn S, et al. (2010) Structure of the 26S proteasome from *Schizosaccharomyces pombe* at subnanometer resolution. *Proc Natl Acad Sci USA* 107(49):20992–20997.
- Kao A, et al. (2011) Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. *Mol Cell Proteomics* 10(1):M110.002212.
- Karadzic I, et al. (2012) Chemical cross-linking, mass spectrometry, and in silico modeling of proteasomal 20S core particles of the haloarchaeon *Haloferax volcanii*. *Proteomics* 12(11):1806–1814.
- Kao A, et al. (2012) Mapping the structural topology of the yeast 19S proteasomal regulatory particle using chemical cross-linking and probabilistic modeling. *Mol Cell Proteomics* 11(12):1566–1577.
- Beck F, et al. (2012) Near-atomic resolution structural model of the yeast 26S proteasome. *Proc Natl Acad Sci USA* 109(37):14870–14875.
- Śledź P, et al. (2013) Structure of the 26S proteasome with ATP- γ S bound provides insights into the mechanism of nucleotide-dependent substrate translocation. *Proc Natl Acad Sci USA* 110(18):7264–7269.
- Unverdorben P, et al. (2014) Deep classification of a large cryo-EM dataset defines the conformational landscape of the 26S proteasome. *Proc Natl Acad Sci USA* 111(15):5544–5549.